Transactions of the ASME®

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).
 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 69 • NUMBER 4 • JULY 2002

Special Issue: Nanomechanics of Surfaces and Interfaces

TECHNICAL PAPERS

(**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.

Journal of Applied Mechanics Volume 69, Number 4 JULY 2002

BRIEF NOTES

ANNOUNCEMENTS AND SPECIAL NOTES

Journal of Applied Mechanics

Foreword

NANOMECHANICS OF SURFACES AND INTERFACES

An engineering revolution is currently underway in that devices at the nanometer scale are fabricated. Successful design and development of nanoscale devices and systems will have a significant impact in our lives. The microelectronics industry is currently pursuing 0.15-micron process technology and this will be extended to the 100-nm process technology in the next few years. The information storage industry is investigating control and processing of nanometer grain magnetic thin films for ultra high-density data storage. Biotechnology of DNA and protein microarrays involving oligonucleotide molecules attached to glass or plastic substrates promises to monitor the whole human genome on a single chip so that researchers can have a better picture of the interactions among thousands of genes simultaneously. We, as students of Mechanics, have an important role to play in these new developments. This is not unexpected; even Aristotle gave Mechanics a sense of wonder: *Nature works against man's needs, because it always takes its own course. Thus, when it is necessary to do something that goes beyond Nature, the difficulties can be overcome with the assistance of Engineering. Mechanics is the name of the Engineering discipline that helps us over those difficulties; as the poet Antiphon put it "Engineering brings the victory that Nature impedes."*

The emerging field of nanomechanics is concerned with extending concepts and methods in traditional mechanics to describe the mechanical response of nanoscale devices (e.g., the deformation modes of nanotubes). At nanometer length scales, conventional computational methods (e.g., finite element method) for design and analysis are no longer appropriate. There is a compelling need for novel modeling and simulation techniques suitable for nanoscale science and engineering as well as innovative experiments to validate these methods.

Modeling and simulation of materials at the nanoscale will require not only atomic and molecular modeling, but also modeling at both the mesoscopic and continuum scales. This is especially true of mechanical properties that depend on phenomena at all possible length scales. For example, it is known that atomic and molecular arrangements (and energies) at the cores of dislocations and at the tips of cracks play a dominant role on the mechanical properties of materials. These effects manifest themselves through the properties and behavior of imperfections and through the collective behavior of these imperfections at the continuum scale. The grand challenge of nanoscale modeling and simulation is to develop computational methods capable of spanning different length scales. The validation of such computational methods will also require comparison with experiments, especially ones that directly reveal material behavior at multiple length scales. One example is the nanoindentation experiment that, besides being a widely used experimental method to measure nanoscale material properties, also appears to be ideal for validating modeling and simulation methods. We need more experimental methods that provide direct information about mechanical properties of materials at all length scales of interest, from the atomic and molecular scale, through the mesoscopic to the continuum scale.

The group of eight papers that follows is intended as a preliminary overview of some of the current topics and opportunities of interest to the researchers from a traditional mechanics background. The authors have utilized their varying expertise in applied physics, materials, and mechanics to provide an exciting view in the world of nanomechanics. We hope that the readers of the Journal will find this work useful and stimulating.

We would like to thank Professor Wheeler for his encouragement and continuous support, the Journal of Applied Mechanics for providing the venue for this interaction, and particularly the authors for their contribution and patience.

Demitris Kouris, Associate Editor
Department of Mechanical Engineering
University of Wyoming
Laramie, WY 82071-3295
Tel: (307) 766-2122
e-mail: kouris@uwyo.edu

Huajian Gao, Professor
Max Planck Institute for Metals Research
Seestr 92, 70174 Stuttgart
Germany
Tel: +49-711-2095-519
e-mail: hjgao@mf.mpg.de

S. P. Baker
Department of Materials Science and Engineering,
Cornell University,
Bard Hall,
Ithaca, NY 14853
e-mail: shefford.baker@cornell.edu
Mem. ASME

X. Wang
Alventive, Inc.,
Galeria Parkway, Suite 400,
Atlanta, GA 30339
e-mail: xwang@alventive.com

C.-Y. Hui
Department of Theoretical and Applied Mechanics,
Cornell University,
Kimball Hall,
Ithaca, NY 14853
e-mail: ch45@cornell.edu

Effect of Nonlinear Elastic Behavior on Bilayer Decohesion of Thin Metal Films From Nonmetal Substrates

Nonlinear unloading behavior has been observed in thin metal films on substrates. In the present work, the effects of this nonlinear unloading behavior on the strain energy release rate in bilayer decohesion experiments, in which a highly stressed overlayer ("driver") is used to decohere a layer ("target") from a substrate, is modeled. Cases where either the driver or the target layer are nonlinear are considered. For particular combinations of stiffnesses and thicknesses, the difference between linear and nonlinear unloading behavior can be quite large (several hundred percent) at experimentally observed stress levels. For practical cases of CR/CU and CU/glass driver/target layer combinations, the maximum difference is about 25%. [DOI: 10.1115/1.1468998]

1 Introduction

The adhesion between thin metal films and nonmetal substrates is a major technological concern and a number of analytical and experimental investigations on this topic have been reported ([1–7]). Decohesion of a thin metal film from a nonmetal substrate is a very complicated process ([4]) which depends on many factors including: the cohesive energy at the interface, elastic, and plastic properties of the film and substrate, thickness of the film, stress level in the film, the geometry of the crack, and environmental effects. Furthermore, two decohesion modes have been observed. Depending on the adhesion between the thin film and the substrate and the stresses, cracks may initiate and propagate along the film/substrate interface ([5,6]) or start at the interface, extend into the substrate and evolve into a trajectory parallel to the interface ([1–3]).

An example of much current interest is provided by the emergence of copper as the next generation metallization material in microelectronics. Because of its excellent electrical conductivity and resistance to electromigration, Cu is used in a wide array of applications including Si-based integrated circuits, flat panel displays and other large area electronics, multichip modules, circuit boards, and individual high-performance thin film devices (e.g., transistors, resistors, capacitors). In these applications, Cu layers are deposited on, and are in turn covered with, other layers of many different types of material including oxide and nitride barriers, polymer dielectrics, and metallic seed and adhesion layers. Due to differential thermal expansion between Cu and the typical nonmetal substrates (e.g., Si, glass, polymer) that are used in such applications, very high stresses may arise in the metallization which, in turn, may lead to decohesion failure. A number of recent studies have focused on adhesion between Cu and various other layers ([6,7]).

In recent studies of the thermomechanical behavior of Cu thin films, a very strong Bauschinger-like effect has been observed in which the yield stress of the film on reverse loading is reduced to the point that it becomes negative: i.e., during unloading, the film begins to yield plastically in compression while the applied stresses are still tensile ([8,9]). An example is shown in Fig. 1. Here, the stress in a 1-μm thick Cu film encapsulated by thin Si_3N_4 barrier (underlayer) and passivation (overlayer) films on a Si substrate is shown as a function of temperature during the second thermal cycle between room temperature and 600°C ([10]). Also shown is the ''thermoelastic line'' that the stress would follow if the deformation during heating were purely elastic. It is evident that the data deviate from the thermoelastic line beginning at a temperature of about 100°C, where the stress on the film is still tensile. The fact that the data lie above this line (relaxation towards more tensile stresses) immediately after yielding indicates that the plastic strains are compressive.

As in classical descriptions of the Bauschinger effect, this ''negative yielding'' behavior indicates deformation that is driven by energy stored in the film during the previous loading (cooling) cycle. A mechanism for this type of behavior based on the storage and recovery of energy in the form of dislocation line length at interfaces has been proposed elsewhere ([8]). In the present work, we consider the effect of this nonlinear unloading behavior on the driving force for decohesion crack propagation in two ''driver film'' bilayer decohesion problems: In one, a highly stressed linear elastic overlayer (the ''driver'') is used to drive a decohesion crack between an underlying metal film (the ''target'') having unloading behavior like that shown in Fig. 1 and the substrate. In the other, the metal film is the driver and is used to cause an underlying linear elastic layer to decohere from the substrate. In both cases, the effect of the nonlinear unloading in the metal film shown in Fig. 1 on the strain energy release rate during decohesion is considered. General solutions for these bilayer problems are presented, along with predictions for particular cases involving Cu films.

2 Bilayer Modeling

The model can be described using the thought experiment illustrated in Fig. 2. Figure 2(*a*) shows a bilayer film on a substrate with a decohesion crack running from the right to the left along the bilayer/substrate interface. Due to internal stresses, the decohered bilayer curls away from the substrate. Our general strategy will be to compare the strain energy per unit film area, $U_{+\infty}$, far

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, March 15, 2001; final revision, December 15, 2001. Associate Editor: D. A. Kouris. Discussion on the paper should be addressed to the Editor, Professor Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

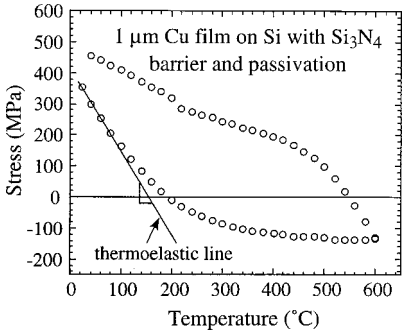


Fig. 1 Thermomechanical behavior of a 1-μm thick Cu film encapsulated between Si_3N_4 layers on a Si substrate ([10]). The solid line indicates the expected thermoelastic behavior of the film. The data deviate upward from the thermoelastic line beginning at about 150°C during initial heating indicating the onset of compressive plastic yielding although the film is still in tension. This nonlinear unloading behavior is included in the subsequent decohesion modeling.

ahead of the decohesion crack (shaded area on the left in Fig. 2(*a*)) with the strain energy per unit area far behind the decohesion crack, $U_{-\infty}$ (shaded area on the right in Fig. 2(*a*)). The strain energy per unit area available to drive the crack is then just $G = U_{+\infty} - U_{-\infty}$.

To find $U_{+\infty}$ and $U_{-\infty}$, we need to know the stresses in the layers. We assume that, before decohesion, layers 1 and 2 support uniform stresses σ_1^o and σ_2^o due to the constraint imposed by the substrate (Fig. 2(*b*)). To find the stresses in the decohered bilayer, we imagine releasing both layers from the substrate as shown in Fig. 2(*c*). We then apply equal and opposite moments, *M*, and uniaxial forces, *S*, per unit width to the two layers as shown in Fig. 2(*d*) in order to match their lengths and curvatures along the layer 1/layer 2 interface. When these conditions are met, the stresses in the layers are the same as those in the decohered bilayer.

For a bilayer film, the stress state before decohesion is, in general, biaxial in the plane of the film. When the bilayer is released from the substrate and *S* and *M* are applied, it curves along the direction perpendicular to the crack front. The stresses are relaxed in this direction, but parallel to the crack front the decohered bilayer has no curvature and hence continues to support stresses in this direction. In other words, the stress state is approximately that corresponding to plane strain in the plane with normal in the direction parallel to the crack front. However, to make this problem more tractable, we neglect the stresses parallel to the crack front. This uniaxial treatment is similar to treating the sample as if it were a decohering narrow strip ([5]).

2.1 Model for the Linear Driver Film Experiment. In the driver film experiment, a highly stressed ''driver'' film (layer 1 in Fig. 2) is deposited onto a ''target'' film of interest (layer 2) in order to measure the adhesion between the target film and the substrate ([11]). In general, the driver film is selected such that $\sigma_1^o > \sigma_2^o$ so that the bilayer curls up when it decoheres from the substrate. A typical driver film is a refractory metal (e.g., Cr) evaporated at low temperature so as to support very high stresses after deposition. Such a film does not show the nonlinear unloading behavior shown in Fig. 1 ([11]). Thus, we model the driver as linear elastic and compare the cases where the target film exhibits linear and nonlinear unloading behavior.

2.1.1 Linear Unloading in Target Film. We begin with the case where layers 1 and 2 are both linear elastic and have uniform initial stresses σ_1^o and σ_2^o, Young's moduli E_1 and E_2, and thicknesses h_1 and h_2, respectively. The strain energy per unit area far ahead of the crack (the shaded area on the left side of Fig. 2(*a*)) is

$$U_{+\infty} = \frac{(\sigma_1^o)^2 h_1}{2E_1} + \frac{(\sigma_2^o)^2 h_2}{2E_2}. \tag{1}$$

The stresses in the two layers after the layers have been released from the substrate (Fig. 2(*c*)) and *S* and *M* are applied (Fig. 2(*d*)) can be written as

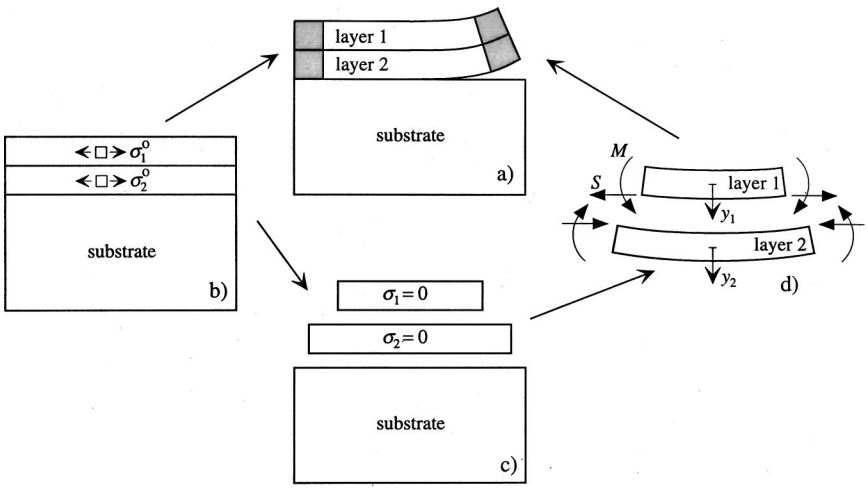


Fig. 2 Schematic of the bilayer adhesion problem. (*a*) The strain energy release rate can be obtained by comparing the strain energy in the film far ahead of and far behind the decohesion crack front. (*b*) Far ahead of the crack, the strain energy is just that arising from the initial stresses in the two layers. We then imagine (*c*) releasing the films from the substrate and (*d*) applying edge forces *S* and moments *M* to ensure that the two layers have the same length and curvature along the interface. When these two conditions are met, the layers can be joined, and have the configuration of the bilayer behind the crack. The strain energy can be calculated from *S* and *M* and the properties of the layers.

$$\sigma_1 = \frac{S}{h_1} + \left(\frac{Sh_1}{2} - M\right)\frac{y_1}{I_1}, \tag{2a}$$

$$\sigma_2 = -\frac{S}{h_2} + \left(\frac{Sh_2}{2} + M\right)\frac{y_2}{I_2}, \tag{2b}$$

where $I_1 = h_1^3/12$ and $I_2 = h_2^3/12$ are the moments of inertia per unit width, and y_1 and y_2 are coordinates normal to the plane of each layer (measured down from the center of each layer as shown in Fig. 2(*d*)), for layer 1 and layer 2, respectively. The requirement that the change in strain along the interface plane be the same in both layers (that is, at $h_1 = y_1/2$ and $h_2 = -y_2/2$) when the layers are removed from the substrate and rejoined as shown in Fig's 2(*b*)–2(*d*) can be expressed as

$$\frac{1}{E_1}\left[\frac{S}{h_1} + \left(\frac{Sh_1}{2} - M\right)\frac{h_1}{2I_1}\right] - \varepsilon_1^T = \frac{1}{E_2}\left[-\frac{S}{h_2} - \left(\frac{Sh_2}{2} + M\right)\frac{h_2}{2I_2}\right] - \varepsilon_2^T, \tag{3}$$

where

$$\varepsilon_1^T = \frac{\sigma_1^o}{E_1} \tag{4a}$$

and

$$\varepsilon_2^T = \frac{\sigma_2^o}{E_2}. \tag{4b}$$

The requirement that the curvatures be the same along the layer 1/layer 2 interface in Fig. 2(*d*) can be written as

$$\frac{1}{E_1 I}\left(\frac{Sh_1}{2} - M\right) = \frac{1}{E_2 I_2}\left(\frac{Sh_2}{2} + M\right). \tag{5}$$

To find $U_{-\infty}$, we must solve Eq's. (3) and (5) for S and M, and use these values in the expressions for the stresses in the layers in Eq. (2). The strain energy per unit width far behind the crack (the shaded area on the right side of Fig. 2(*a*)) can then be found from

$$U_{-\infty} = \frac{1}{2}\int_{-h_1/2}^{h_1/2}\frac{\sigma_1^2}{E_1}dy_1 + \frac{1}{2}\int_{-h_2/2}^{h_2/2}\frac{\sigma_2^2}{E_2}dy_2. \tag{6}$$

The solution can be put into a considerably more useful form by making the following substitutions:

$$\delta = \frac{h_1}{h_2}, \tag{7a}$$

$$\omega = \frac{E_1}{E_2}, \tag{7b}$$

$$\frac{1}{E} = \frac{1}{E_1} + \frac{\delta}{E_2} = \frac{1}{E_1}(1 + \delta\omega), \tag{7c}$$

$$\chi = \frac{6(1+\delta)}{\delta(1+\omega\delta)}, \tag{7d}$$

$$\gamma = \frac{\omega\delta^2(1+\delta)}{2(1+\omega\delta^3)}, \tag{7e}$$

and

$$\Delta\varepsilon^T = \varepsilon_1^T - \varepsilon_2^T = \frac{\sigma_1^o}{E_1} - \frac{\sigma_2^o}{E_2} = \frac{1}{E_1}(\sigma_1^o - \omega\sigma_2^o). \tag{7f}$$

If we then let

$$s = \frac{S}{Eh_1\Delta\varepsilon^T} \tag{8a}$$

and

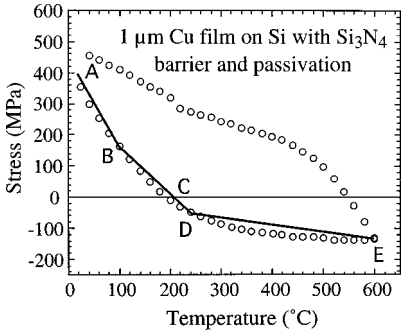


Fig. 3 The linearization of the Cu film unloading problem is accomplished by approximating substrate curvature $\sigma - T$ data with straight lines during heating

$$m = \frac{M}{Eh_1^2\Delta\varepsilon^T}, \tag{8b}$$

then by combining Eqs. (4) and (6) and using the substitutions in Eqs. (7) and (8), after substantial manipulation we find solutions for the stresses in the layers after decohesion (Eqs. (2*a*) and (2*b*)) which can be written as

$$\sigma_1 = \frac{\sigma_1^o - \omega\sigma_2^o}{(1+\omega\delta)(1+\chi\gamma)}\left(1 + 12\gamma\frac{y_1}{h_1}\right) \tag{9a}$$

and

$$\sigma_2 = \frac{\sigma_1^o - \omega\sigma_2^o}{(1+\omega\delta)(1+\chi\gamma)}\left(-\delta + \frac{12\gamma}{\omega\delta}\frac{y_2}{h_2}\right). \tag{9b}$$

The solution to Eq. (6) is then

$$U_{-\infty} = \frac{h_1}{2E_1}\left[\frac{\sigma_1^o - \omega\sigma_2^o}{(1+\omega\delta)(1+\chi\gamma)}\right]^2\left[\frac{1 + \omega\delta + 12\gamma^2\left(1 + \frac{1}{\omega\delta^3}\right)}{(1+\omega\gamma)^2}\right]. \tag{10}$$

Finally, using Eqs. (1), (7*a*), (7*b*), and (10) we find the strain energy release rate, $G = U_{+\infty} - U_{-\infty}$, for the linear elastic case to be

$$G = \frac{h_1}{2E_1}\left[(\sigma_1^o)^2 + \frac{\omega}{\delta}(\sigma_1^o)^2 - \left[\frac{\sigma_1^o - \omega\sigma_2^o}{(1+\omega\delta)(1+\chi\gamma)}\right]^2 \times\left[\frac{1 + \omega\delta + 12\gamma^2\left(1 + \frac{1}{\omega\delta^3}\right)}{(1+\omega\gamma)^2}\right]\right]. \tag{11}$$

2.1.2 Nonlinear Unloading Behavior in Target Film. We assume that a film undergoing isothermal unloading (as during decohesion) would follow the behavior shown in Fig. 1, at least at strains corresponding to low temperatures. We approximate this behavior by a series of straight lines as shown in Fig. 3. Using the known thermal expansion coefficients of the Cu film and the Si substrate, we can convert the stress-temperature behavior along the path ABCDE, into stress-strain behavior with the zero elastic strain axis at point C, as shown in Fig. 4 ([11]). Since the decohered film will never be reloaded, we can treat the nonlinear unloading behavior shown in Fig. 1 as if it were elastic. That is, although the nonlinear behavior may arise from dislocation motion during unloading ([8,10]), the stresses will decrease along a path similar to that shown in Fig. 4 and this may be treated as a nonlinear elastic unloading process.

The initial slope in Fig. 1 (segment AB in Fig 4) is seen to correspond to that which would be predicted from the well-known

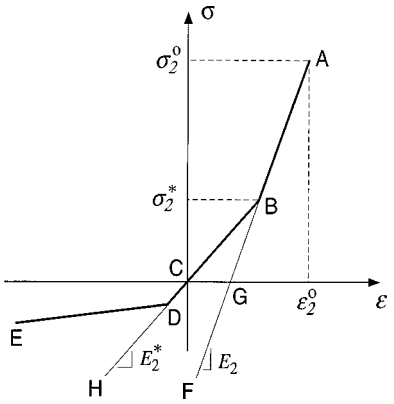


Fig. 4 The $\sigma-\varepsilon$ behavior corresponding to the approximation of the $\sigma-T$ data in Fig. 3 along the path ABCDE obtained using the temperature-dependent thermal expansion coefficients of Cu and Si [11]

elastic constants and thermal expansion coefficients of Cu and Si. Thus, the effect of the nonlinear unloading on the strain energy release rate can be visualized by comparing the case of linear unloading (along line AF) with that of nonlinear unloading which initially follows the line AB, but then deviates to a second branch, BD, and then to the third branch, DE. Since $\sigma_1^o > \sigma_2^o$, the stresses in layer 2 after decohesion will be more compressive than those in layer 1. Thus, compared with the linear case, an element in the nonlinear layer provides additional strain energy (the area in the triangle BCG) as it is unloaded to zero stress, and costs significantly less strain energy as the stress becomes compressive due to the higher compliance of the second and third branches.

In the present model, we consider only the first and second branch, allowing the stress to continue along the second branch as far as necessary (in the direction BDH). The stresses in layer 2 will then, in general, lie along the path ABH after unloading. However, the difference between the linear and non linear cases will be largest if all of the stresses in this layer lie along the second branch. Thus, we only consider solutions where $\sigma_2^{\max} < \sigma_2^*$, where $\sigma_2^{\max}$ is the (algebraic) maximum stress and σ_2^* the stress at the top of the second branch (point B in Fig. 4).

The strain energy per unit area far ahead of the crack front for the bilayer now becomes

$$U_{+\infty}^* = \frac{(\sigma_1^o)^2 h_1}{2E_1} + \frac{h_2}{2}\left[\frac{(\sigma_2^o)^2}{E_2} + (\sigma_2^o)^2\left(\frac{1}{E_2^*} - \frac{1}{E_2}\right)\right], \qquad (12)$$

where E_2^* is the compliance of layer 2 on the second branch. Since we limit solutions to cases where $\sigma_2^{\max} < \sigma_2^*$, Eqs. (2)–(6) above can be modified for the nonlinear unloading case simply by substituting E_2^* for E_2, S^* for S, and M^* for M. The change in strain along the interface in layer 2 during the unloading part of the thought experiment (Fig. 2(b–c)) becomes

$$(\varepsilon_2^T)^* = \frac{\sigma_2^*}{E_2^*} + \frac{\sigma_2^o - \sigma_2^*}{E_2} \qquad (13)$$

(compare with Eq. 4(*b*)). Otherwise, Eqs. (2)–(6) remain the same. The substitutions given by Eqs. (7) now become

$$\delta = \frac{h_1}{h_2}, \qquad \text{(unchanged from (7}a\text{))} \qquad (14a)$$

$$\omega^* = \frac{E_1}{E_2^*}, \qquad (14b)$$

$$\frac{1}{E^*} = \frac{1}{E_1} + \frac{\delta}{E_2^*} = \frac{1}{E_1}(1 + \delta\omega^*), \qquad (14c)$$

$$\chi^* = \frac{6(1+\delta)}{\delta(1+\omega^*\delta)}, \qquad (14d)$$

$$\gamma^* = \frac{\omega^*\delta^2(1+\delta)}{2(1+\omega^*\delta^3)}, \qquad (14e)$$

and

$$(\Delta\varepsilon^T)^* = \varepsilon_1^T - (\varepsilon_2^T)^* = \frac{\sigma_1^o}{E_1} - \frac{\sigma_2^o}{E_2} + \sigma_2^*\left(\frac{1}{E_2} - \frac{1}{E_2^*}\right)$$

$$= \frac{1}{E_1}(\sigma_1^o - \omega\sigma_2^o - \sigma_2^*(\omega^* - \omega)). \qquad (14f)$$

If as before we now let

$$s^* = \frac{S^*}{E^* h_1 (\Delta\varepsilon^T)^*} \qquad (15a)$$

and

$$m^* = \frac{M^*}{E^* h_1^2 (\Delta\varepsilon^T)^*}, \qquad (15b)$$

and follow a similar procedure to that described above, we find solutions for the stresses in the layers after decohesion (compare with Eqs. (9*a*) and (9*b*)) which can be written as

$$\sigma_1 = \frac{\sigma_1^o - \omega\sigma_2^o - \sigma_2^*(\omega^* - \omega)}{(1-\delta\omega^*)(1+\chi^*\gamma^*)}\left(1 + 12\gamma^*\frac{y_1}{h_1}\right) \qquad (16a)$$

and

$$\sigma_2 = \frac{\sigma_1^o - \omega\sigma_2^o - \sigma_2^*(\omega^* - \omega)}{(1-\delta\omega^*)(1+\chi^*\gamma^*)}\left(-\delta + \frac{12\gamma^*}{\omega^*\delta}\frac{y_2}{h_2}\right). \qquad (16b)$$

The strain energy per unit area in the decohered bilayer then becomes

$$U_{-\infty}^* = \frac{h_1}{2E_1}\left[\frac{\sigma_1^o - \omega\sigma_2^o - \sigma_2^*(\omega^* - \omega)}{(1-\delta\omega^*)(1+\chi^*\gamma^*)}\right]^2 \times\left[\frac{1 + \omega^*\delta + 12\left(1 + \frac{1}{\omega^*\delta^3}\right)(\gamma^*)^2}{(1+\omega^*\gamma^*)^2}\right]. \qquad (17)$$

Finally, using Eqs. (12), (14*a*), (14*b*), and (17), we find the strain energy release rate $G^* = U_{+\infty}^* - U_{-\infty}^*$ for the case where layer 2 exhibits nonlinear unloading as shown in Fig. 4 to be

$$G^* = \frac{h_1}{2E_1}\left[\begin{array}{c}(\sigma_1^o)^2 + \frac{\omega}{\delta}(\sigma_2^o)^2 + (\sigma_2^*)^2\frac{(\omega^* - \omega)}{\delta} \\ -\left[\frac{\sigma_1^o - \omega\sigma_2^o - \sigma_2^*(\omega^* - \omega)}{(1-\delta\omega^*)(1+\chi^*\gamma^*)}\right]^2\left[\frac{1 + \omega^*\delta + 12(\gamma^*)^2\left(1 + \frac{1}{\omega^*\delta^3}\right)}{(1+\omega^*\gamma^*)^2}\right]\end{array}\right]. \qquad (18)$$

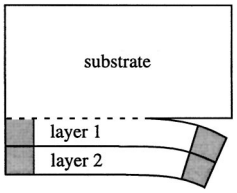


Fig. 5 Schematic of the bilayer decohesion problem when layer 2 is the driver and layer 1 is the target. For this configuration, Eqs. (1)–(18) still hold and the solution can still be used. Only the range of validity of the solutions (Sections 2.1.3 and 2.2) changes.

2.1.3 Limits on Solutions. We implement the requirement that $\sigma_2^{\max}<\sigma_2^*$ by observing that the maximum stress in layer 2 occurs at $y_2=h_2/2$. Thus, from Eq. (16*b*), we only consider solutions for which

$$\sigma_2^{\max}=\frac{\sigma_1^o-\omega\sigma_2^o-\sigma_2^*(\omega^*-\omega)}{(1-\delta\omega^*)(1+\chi^*\gamma^*)}\left(-\delta+\frac{6\gamma^*}{\omega^*\delta}\right)<\sigma_2^* \quad (19)$$

Note that, since we are only considering films that curl up, we also limit our solutions to cases where $\Delta\varepsilon^T$ (Eq. (7*f*)) and $(\Delta\varepsilon^T)^*$ (Eq. (14*f*)) are both positive. We also limit solutions to cases where $E_2^*\leqslant E_2$ so that $\omega^*\geqslant\omega$.

2.2 Model for a Nonlinear Driver Film. We now consider the case where the driver film is nonlinear and the target layer is linear. This would be the case where a metal film with behavior like that shown in Fig. 1 is used as a driver to cause a linear target layer to decohere. The lower layer may be distinct, but it could also be part of the substrate; *i.e.*, decohesion occurs by extension of a crack along a plane parallel to the film/substrate interface but located a distance below it in the substrate. This mechanism has been observed, for example, in decohesion of thin metal films from silicate glass substrates ([1–4]).

We could derive new solutions for this problem by using the thought experiment of Fig. 2 and assuming that layer 1 is nonlinear and layer 2 is linear. However, we can use the derivations above to describe this problem if we simply invert the geometry of the sample as shown in Fig. 5. Layer 1 is (still) linear, and we consider the cases where layer 2 is (still) linear or nonlinear as before (paths AF and ABH, respectively, in Fig. 4). The solutions to this problem are the same as those shown in Section 2.1 and the solutions for the strain energy release rates G and G^* given in Eqs. (11) and (18), respectively, still apply. However, the range of variables over which the solution is valid changes. The requirement that the bilayer curves *down* upon decohesion from the substrate means that $\Delta\varepsilon^T$ (Eq. (7*f*)) and $(\Delta\varepsilon^T)^*$ (Eq. (14*f*)) must both be negative. The maximum (algebraic) stress in layer 2 occurs at the interface ($y_2=-h_2/2$), so the requirement that $\sigma_2^{\max}<\sigma_2^*$ becomes

$$\sigma_2^{\max}=\frac{E^*(\Delta\varepsilon^T)^*}{1+\chi^*\gamma^*}\left(-\delta-\frac{6\gamma^*}{\omega^*\delta}\right)<\sigma_2^*. \quad (20)$$

As before, we require that $\omega^*\geqslant\omega$.

3 Results

We wish to determine the conditions under which the difference between the linear and nonlinear solutions becomes significant. In our derivation, the stresses, σ_1^o, σ_2^o, and σ_2^* are explicit, so we specify these stresses and calculate strain energy release rates for various values of δ, ω, and ω^*.

3.1 Linear Driver Film. We begin by considering the normalized values of the strain energy release rates obtained by multiplying both sides of Eqs. (11) and (18) by $2E_1/h_1$. As a frame of reference, we use values from experimental measurements that are reported elsewhere ([11]). In these experiments, Cr driver films were used to cause Cu films to decohere from glass substrates. The Cr films were found to have a stress as-deposited of 1280 GPa, independent of thickness. For Cu we use the values obtained from Fig. 4. The normalized strain energy release rates,

$$\mathcal{G}=\frac{2E_1}{h_1}G \quad \text{and} \quad \mathcal{G}^*=\frac{2E_1}{h_1}G^*, \quad (21)$$

for the case of a bilayer curling up with $\sigma_1^o=1280$, $\sigma_2^o=400$, and $\sigma_2^*=150$ (Fig. 4) are plotted in Fig. 6 as a function of ω and ω^* for a thickness ratio of $\delta=1$. It is evident that $\mathcal{G}$, which has no dependence on ω^*, goes to zero as ω decreases. This is because $G\rightarrow$constant as $E_1\rightarrow 0$ (layer 1 becomes infinitely compliant) and $G\rightarrow 0$ as $E_2\rightarrow\infty$ (layer 2 becomes infinitely stiff). On the other hand, $\mathcal{G}^*$ increases with ω^* at all ω values and decreases to a constant as ω decreases. The latter arises because E_2^* must scale with E_1 in order for ω to decrease at constant ω^*. Thus, $G^*\rightarrow\infty$ as $E_1\rightarrow 0$ (such that the product is constant) and $G^*\rightarrow$constant as $E_2\rightarrow\infty$. The region in which the solution is valid is limited at lower values of ω and ω^* by the requirement that $\sigma_2^{\max}<\sigma_2^*$, at higher values of ω and lower values of ω^* by the requirement that $\omega^*\geqslant\omega$, and at higher values of ω and ω^* by the requirement that $\Delta\varepsilon^T$ and $(\Delta\varepsilon^T)^*$ both be positive.

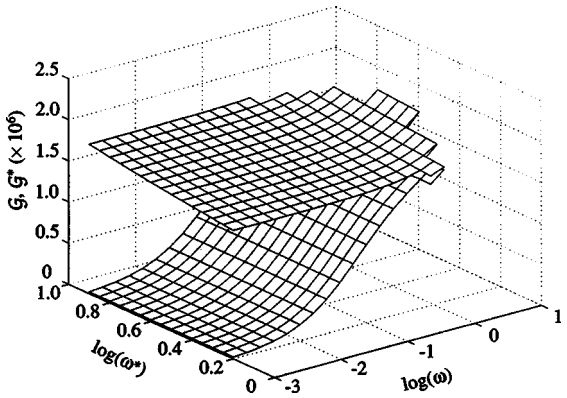


Fig. 6 Normalized strain energy release rates $\mathcal{G}$ (lower surface) and $\mathcal{G}^*$ (upper surface) as defined in Eq. (21) for linear and nonlinear bilayers, respectively, with $\sigma_1^o=1280$, $\sigma_2^o=400$, and $\sigma_2^*=150$ as a function of the stiffness ratios ω (Eq. 7(*b*)) and ω^* (Eq. 14(*b*)) at a thickness ratio of $\delta=1$ (Eq 7(*a*)). The bilayer curves up as shown in Fig. 2. As $\omega\rightarrow 0$, $\mathcal{G}\rightarrow 0$ and $\mathcal{G}^*\rightarrow$constant. $\mathcal{G}^*$ increases with increasing ω^*.

It is evident that the difference between the linear and nonlinear cases increases as ω decreases and ω^* increases. Figure 7 shows the fractional difference in the strain energy release rate, given by

$$\mathcal{F}=\frac{\mathcal{G}^*}{\mathcal{G}}-1, \quad (22)$$

over a narrower range of ω and ω^* for the same stresses and thickness ratio as in Fig. 6. As expected, $\mathcal{F}=0$ when $\omega=\omega^*$, and increases as ω decreases and ω^* increases. For the Cr/Cu system, we have $E_1=248$ GPa, $E_2=230$ GPa, $E_2^*=106$ GPa, so $\log(\omega)=0.018$ and $\log(\omega^*)=0.369$. The difference in strain energy release rates between the linear and nonlinear case for this system is about 5%. For systems with smaller ω or larger ω^*, this difference rises quickly.

Figure 8 shows $\mathcal{G}$ and $\mathcal{G}^*$ as a function of δ and ω^* for $\omega=1$. These values both increase as $\delta\rightarrow 0$, and decrease as $\delta\rightarrow\infty$ as a result of the normalization by $1/h_1$. In Fig. 9, $\mathcal{F}$ is shown as a

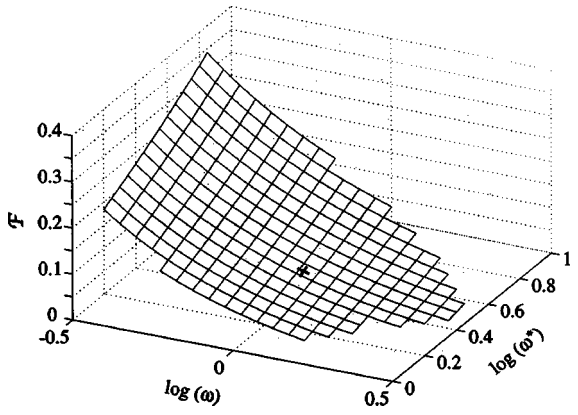


Fig. 7 Fractional difference in strain energy release rate, $\mathcal{F}$ (Eq. (22)), as a function of the elastic stiffness ratios ω and ω^* corresponding to the data shown in Fig. 6 ($\delta=1$). The difference is zero along the line $\omega=\omega^*$ and increases rapidly as ω decreases and ω^* increases. The point marked with a cross indicates the stiffness ratios for the case of a Cr/Cu bilayer. The strain energy release rate is about 5% higher for nonlinear than for linear unloading (branches BH and BF, respectively, in Fig 4) at this point. Note that the scale and perspective are different from Fig. 6.

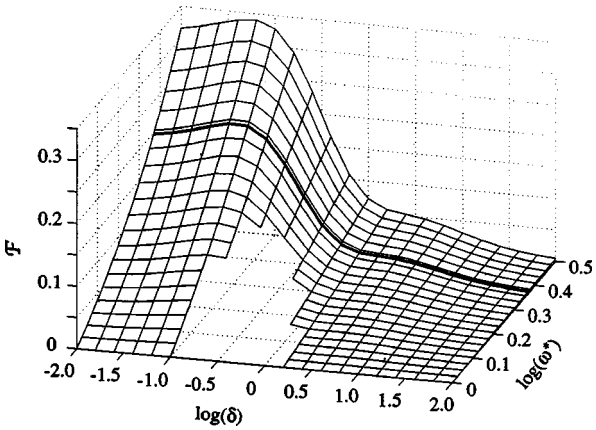


Fig. 9 Fractional difference in strain energy release rate, $\mathcal{F}$ (Eq. (22)), as a function of the elastic stiffness ratio ω^* and the thickness ratio δ corresponding to the data shown in Fig. 8 ($\omega=1$). The difference is zero along the line $\log(\omega^*)=0$ and increases as ω^* increases. $\mathcal{F}$ increases as δ increases and the upper linear layer dominates. As δ decreases, $\mathcal{F}$ approaches a constant that depends on ω^*. The dark line indicates the stiffness ratios for the case of a Cr/Cu bilayer. The difference in strain energy release rate reaches a maximum of about 24% at about $\log(\delta)=-1.2$. Note that the scale and perspective are different from Fig. 8.

function of δ and ω^* for $\omega=1$. As expected, along the line $\omega=\omega^*=1$, $\mathcal{F}=0$. A small maximum occurs at all values of ω^* greater than 0 at about $\log(\delta)=-1.2$. The heavy line at $\log(\omega^*)=0.369$ indicates the behavior of the Cr/Cu system. The maximum along this line (at $\log(\delta)\approx-1.2$) is about 24%. The gap in the solution arises from the requirement that $\sigma_2^{\max}<\sigma_2^*$.

3.2 Nonlinear Driver Film. We now consider the case shown in Fig. 5 where the target film (now layer 1) is linear and the driver film (now layer 2) is either linear or nonlinear. These solutions are similar in form to those described in Section 3.1. Figure 10 (compare with Fig. 6) shows $\mathcal{G}$ and $\mathcal{G}^*$ (Eq. (21)) as functions of ω and ω^* for $\delta=1$, $\sigma_1^o=0$, $\sigma_2^o=400$, and $\sigma_2^*=150$. These stresses correspond to a nonlinear metal film like that shown in Fig 1 decohering an unstressed layer. The solution for $\mathcal{G}$ does not depend on ω^* and goes to zero as ω decreases, and that for $\mathcal{G}^*$ becomes constant with ω as ω decreases and constant with ω^* as ω^* increases. As can be seen in Fig. 11 (compare with Fig. 7), the fractional difference $\mathcal{F}$ is a smooth function that is zero for $\omega=\omega^*$, and goes to infinity as ω decreases and ω^* increases. For the case of a nonlinear Cu driver film decohering a glass target layer ($E_1=70$ GPa, $E_2=230$ GPa, $E_2^*=106$ GPa) the strain energy release rate is about 15% higher than if the Cu layer were to unload linearly.

In Figs. 12 and 13 (compare with Fig's 8 and 9), the same calculations of $\mathcal{G}$ and $\mathcal{G}^*$, and $\mathcal{F}$, respectively, are shown as functions of δ and ω^* for $\omega=0.3$ ($\approx E_1/E_2=70/230$). In Fig. 13, the heavy line indicating $\omega^*=0.66$ ($\approx E_1/E_2=70/106$) represents the Cu/glass bilayer system. Along this line, $\mathcal{F}$ increases rapidly as δ decreases for $\delta>1$, exceeding 15% at $\log(\delta)\approx-0.2$, decreases

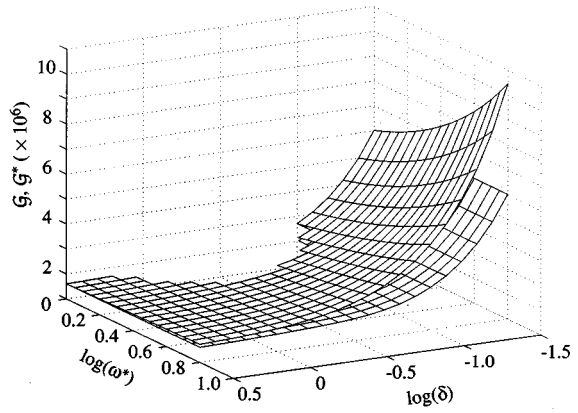


Fig. 8 Normalized strain energy release rates $\mathcal{G}$ (lower surface) and $\mathcal{G}^*$ (upper surface) as defined in Eq. (21) for linear and nonlinear bilayers, respectively, with $\sigma_1^o=1280$, $\sigma_2^o=400$, and $\sigma_2^*=150$ as a function of the stiffness ratio ω^* (Eq. (14*b*)) and the thickness ratio δ (Eq. (7*a*)) for $\omega=1$ (Eq. (7*b*)). The bilayer curves up as shown in Fig. 2. Both $\mathcal{G}$ and $\mathcal{G}^*$ go to zero as δ goes to zero and increase as δ increases.

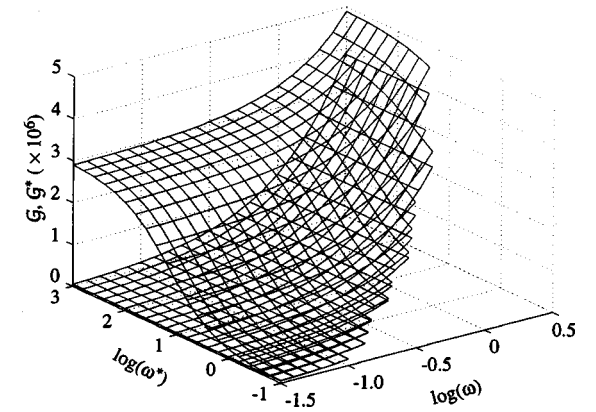


Fig. 10 Normalized strain energy release rates $\mathcal{G}$ (lower surface) and $\mathcal{G}^*$ (upper surface) as defined in Eq. (21) for linear and nonlinear bilayers, respectively, with $\sigma_1^o=0$, $\sigma_2^o=400$, and $\sigma_2^*=150$ as a function of the stiffness ratios ω (Eq. (7*b*)) and ω^* (Eq. (14*b*)) at a thickness ratio of $\delta=1$ (Eq. (7*a*)). The bilayer curves down as shown in Fig. 5. As $\omega\to0$, $\mathcal{G}\to0$ and $\mathcal{G}^*\to$constant. $\mathcal{G}^*$ increases with increasing ω^*.

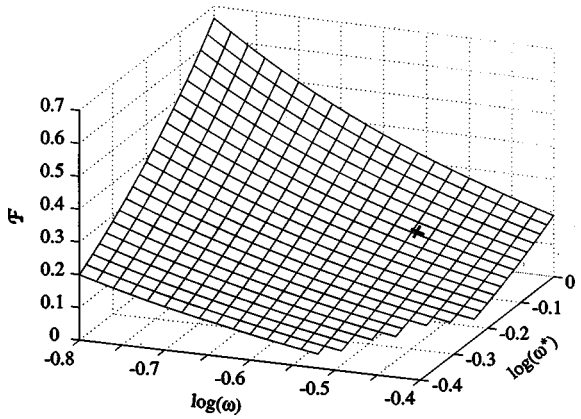


Fig. 11 Fractional difference in strain energy release rate, $\mathcal{F}$ (Eq. (22)), as a function of the elastic stiffness ratios ω and ω^* corresponding to the data shown in Fig. 10 ($\delta=1$). The difference increases rapidly as ω decreases and ω^* increases. The point marked with a cross indicates the stiffness ratios for the case of a Cu/glass bilayer. The strain energy release rate is about 15% higher for nonlinear than for linear unloading (branches BH and BF, respectively, in Fig 4) at this point. Note that the scale and perspective are different from Fig. 10.

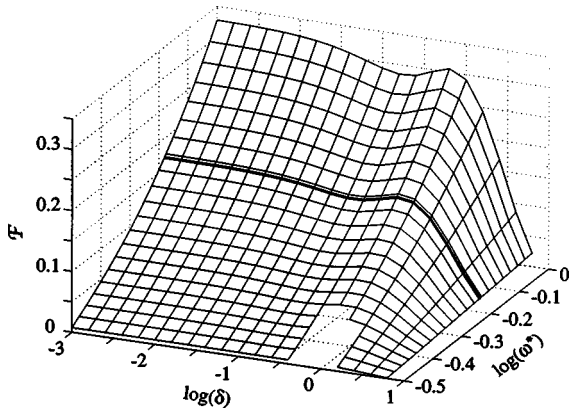


Fig. 13 Fractional difference in strain energy release rate, $\mathcal{F}$ (Eq. (22)), as a function of the elastic stiffness ratio ω^* and the thickness ratio δ corresponding to the data shown in Fig. 12 ($\omega=0.3$). The difference is zero along the line $\log(\omega^*)=0$ and increases as ω^* increases. $\mathcal{F}$ increases as δ increases and the upper linear layer dominates. As δ decreases, $\mathcal{F}$ approaches a constant that depends on ω^*. The dark line indicates the stiffness ratios for the case of a Cu/glass bilayer. The difference in strain energy release rate reaches a maximum of about 17% as $\delta\rightarrow 0$. Note that the scale and perspective are different from Fig. 12.

slightly and then increases again slightly, becoming constant at about 17% (the difference between linear and nonlinear Cu layers alone) as $\delta\rightarrow 0$.

4 Discussion

The calculations above show that non-linear unloading as shown in Figs. 1, 3, and 4 can have a significant effect on the strain energy release rate of a bilayer depending on the stresses, elastic constants and thicknesses of the layers.

In each case, the general form of the solution is very similar regardless of whether it is the driver or the target film which is non-linear (compare Figs. 6–9 with Figs. 10–13). This is to be expected since the same solutions are used; only the stresses and the range of parameters where the solutions are valid differ. Less obviously, the general shapes of the solutions shown in Figs. 6–13 do not change significantly over a wide range of reasonable tensile stress values, although, again, the magnitudes of $\mathcal{G}$, $\mathcal{G}^*$, and $\mathcal{F}$, as well as the range of parameters where the solutions are valid change. Thus, the solutions shown are representative of the behavior that might be expected for bilayers containing films with unloading behavior like that shown in Fig. 1.

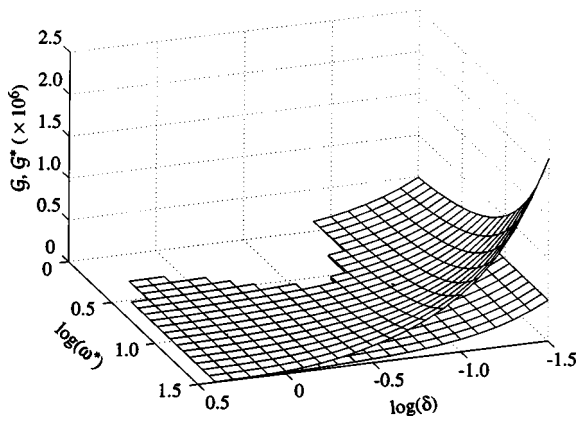


Fig. 12 Normalized strain energy release rates $\mathcal{G}$ (lower surface) and $\mathcal{G}^*$ (upper surface) as defined in Eq. (21) for linear and nonlinear bilayers, respectively, with $\sigma_1^o=0$, $\sigma_2^o=400$, and $\sigma_2^*=150$ as a function of the stiffness ratio ω^* (Eq. (14*b*)) and the thickness ratio δ (Eq. (7*a*)) for $\omega=0.3$ (Eq. (7*b*)). The bilayer curves down as shown in Fig. 5. Both $\mathcal{G}$ and $\mathcal{G}^*$ go to zero as δ goes to zero and increase as δ increases.

Interpretation of the effect of the thickness ratio δ on the fractional difference in the energy release rate $\mathcal{F}$ (Figs. 9 and 13) is straightforward. As the lower metal layer becomes thicker relative to the upper layer ($\delta\rightarrow 0$), the solutions for $\mathcal{G}$ and $\mathcal{G}^*$ approach the behavior of the lower layer alone, becoming constant in δ. At a constant low δ, the difference between the linear and nonlinear solutions, and thus $\mathcal{F}$, increases as ω^* increases. As δ increases, the strain energy becomes concentrated in the upper linear layer and the behavior of a nonlinear lower layer becomes irrelevant. The strain energy difference becomes small at all ω^* as the contribution of the lower layer becomes small. The location of the maximum in $\mathcal{F}$ depends on whether the nonlinear layer is the driver or the target. If the nonlinear layer is the target (Fig. 2), the maximum occurs at an intermediate value of δ (Fig. 9). This is due to the upper layer compressing the lower layer after unloading, driving it further along the line BH in Fig. 4, thereby increasing the strain energy difference in comparison to the linear film, which unloads along BF. If the nonlinear layer is the driver (Fig. 5), the maximum difference between the linear and nonlinear cases (Fig. 13) occurs as $\delta\rightarrow 0$ since the stress is more tensile in layer 2 and reducing the influence of layer 1 allows the strain energy in layer 2 to be further reduced.

For the specific cases of Cr/Cu and Cu/glass driver/target combinations, this model predicts that the nonlinear unloading in Fig. 1 increases the strain energy release rate by 15–25% compared with linear unloading at particular values of δ (Figs. 9 and 13). Of course, the total strain energy release rate available depends on the total thicknesses of the layers, so for a given target layer thickness, effects of this magnitude will only be seen in experiments when the fracture toughness of the interface lies in an appropriate range. Since low thickness ratios are needed to maximize the differences and the stresses in the Cu layers are relatively modest, the effects of nonlinear behavior will be most evident for systems with low fracture toughness.

Larger effects than those shown for the Cr/Cu and Cu/glass bilayers are, in principle, possible. In general, the solutions reveal that the differences between the linear and nonlinear unloading

cases increase as ω^* becomes high at constant ω, or ω becomes low at constant ω^*, in the range of valid solutions. When ω^* is high, the difference between the slopes of the first (AF) and second (BH) unloading branches in Fig. 4, and thus the difference in the overall elastic behavior, is maximized. When ω is low, the always-linear upper layer (layer 1) is very compliant compared to the linear or nonlinear lower metal layer (layer 2). If the metal film is the target, a compliant driver pushes the metal film farther into compression, thus maximizing the difference between the linear and nonlinear unloading cases. If the non-linear metal film is the driver, a compliant target allows the metal to unload farther towards the compressive regime with the same outcome.

In practice, larger values of $\mathcal{F}$ may be difficult to achieve. One can reduce ω by decreasing E_1. However, if E_1 is reduced, both ω and ω^* are reduced. Interestingly, this does not change the value of $\mathcal{F}$ significantly. Calculations with $E_1 = 10$ and 1 GPa, and layer 2 values for Cu as above ($E_2 = 230$ GPa, $E_2^* = 106$ GPa) both return values of $\mathcal{F} \approx 25\%$ and $\mathcal{F} \approx 17\%$ for cases where the metal film is the target or the driver, respectively. For the case of the metal target, the solutions were only valid when the stress in the driver layer (σ_1^o) was reduced to about 50 and 25 Mpa for $E_1 = 10$ and 1 GPa, respectively.

One could imagine increasing ω^*, by decreasing E_2^*. However, this would require changing the material behavior and may be difficult to do (although maybe not impossible if it is a microstructural effect ([8,9]). But we note that our crude approximation includes only the second branch (line BH) in Fig. 4). If the compliance of the third branch were used (line DE in Fig. 4) ([11]), the difference between the linear and nonlinear cases would be larger.

It is important to note that the mechanism which has been proposed to account for the nonlinear unloading behavior in a metal film depends on the energy stored in the form of dislocations at the film/substrate and film/overlayer interfaces. If the decohesion crack tip runs along an interface with the metal, the dislocations could escape through the free surface created at the crack tip and the stored energy needed to provide the nonlinear behavior would not be available. Thus, the model provided here applies strictly only to cases where the decohesion crack does not run along an interface with the non-linear metal film. This is inherent when the nonlinear layer is the driver (Fig. 5), but it could also apply when the nonlinear layer is the target as long as a third layer constraining dislocations to lie in the film is present. For example, in their study of adhesion of Cu interconnect systems with thin layers of SiO_2 and TaN separating the Si substrate from the Cu film, Lane and Dauskardt [7] saw that decohesion occurred along the SiO_2/TaN interface. To the extent that the adherent TaN layer constrains dislocations to lie in the Cu film, nonlinear unloading like that shown in Fig. 1 could affect the strain energy release rate.

The accuracy of the solutions presented here is limited by our assumption of a uniaxial stress state and by our rough linear approximation of the shape of the unloading curve in Fig. 3. Because of the approximation of the stress state, we used measured biaxial moduli in our calculations ([11]) rather than Young's moduli. Another possible limit to the ability of these model calculations to accurately predict experimental data is the fact that we ignore the plastic work that is done when deforming the film. However, this approach could be valid in some practical cases: Plastic deformation at the crack tip may be insignificant when the nonlinear layer is the driver (Fig. 5), and may be at least reduced when a thin adherent layer (as TaN in the example above) remains adhered to the decohering nonlinear target.

Despite these limitations, the comparisons (Figs. 7, 9, 11, 13) are expected to quantify the relative differences between the nonlinear and linear cases quite well. To the extent that nonlinear behavior such as that shown in Fig. 1 occurs in bilayer decohesion experiments, the present model should provide a good estimate of the difference relative to a simple linear analysis.

5 Conclusions

A model that can be used to determine the strain energy release rate for elastic decohesion of a bilayer in which one of the layers may show nonlinear elastic unloading has been presented. The model is based on the nonlinear unloading behavior observed during thermomechanical testing of Cu thin films. In this model, the nonlinear behavior is treated as two different linear branches. For a Cr overlayer driving decohesion of a nonlinear Cu film, the model predicts that the strain energy release rate could be increased by about 25% relative to a linear Cu film. For a Cu film driving decohesion of a glass layer, the difference between linear and nonlinear Cu is about 15%. For bilayers having a nonlinear layer with more pronounced nonlinearity the strain energy release rate can be several times that which would be predicted assuming both layers to be linear. For a given nonlinearity, the solution is not very sensitive to the compliance of the linear layer. Nonlinear unloading as observed in thin metal films may have a significant effect on strain energy release rate in bilayer decohesion experiments.

Acknowledgments

This work was supported by the National Science Foundation by a CAREER award (DMR-9875119) and through the Cornell Center for Materials Research (CCMR), a Materials Research Science and Engineering Center (DMR-0079992).

References

[1] Hu, M. S., Thouless, M. D., and Evans, A. G., 1988, ''The Decohesion of Thin Films From Brittle Substrates,'' Acta Metall., **36**, pp. 1301–1307.
[2] Drory, M. D., Thouless, M. D., and Evans, A. G., 1988, ''On the Decohesion of Residually Stressed Thin Films,'' Acta Metall., **36**, pp. 2019–2028.
[3] Suo, Z., and Hutchinson, J. W., 1989, ''Steady-State Cracking in Brittle Substrates Beneath Adherent Films,'' Int. J. Solids Struct., **25**, pp. 1337–1353.
[4] Evans, A. G., and Dalgleish, B. J., 1992, ''The Fracture Resistance of Metal-Ceramic Interfaces,'' Acta Metall. Mater., **40**, pp. S295–S305.
[5] Bagchi, A., Lucas, G. E., Suo, Z., and Evans, A. G., 1994, ''A New Procedure for Measuring the Decohesion Energy for Thin Ductile Films on Substrates,'' J. Mater. Res., **9**, pp. 1734–1741.
[6] Kriese, M. D., Gerberich, W. W., and Moody, N. R., 1999, ''Quantitative Adhesion Measures of Multilayer Films: Part II. Indentation of W/Cu, W/W, Cr/W,'' J. Mater. Res., **14**, pp. 3019–3026.
[7] Lane, M., and Dauskardt, R. H., 2000, ''Adhesion and Reliability of Copper Interconnects With Ta and TaN Barrier Layers,'' J. Mater. Res., **15**, pp. 203–211.
[8] Baker, S. P., Keller, R.-M., and Arzt, E., 1998, ''Energy Storage and Recovery in Thin Metal Films on Substrates,'' *Proceedings, Thin Films: Stresses and Mechanical Properties VII, R. C. Cammarata* et al. eds., Materials Research Society, Warrendale PA, pp. 605–610.
[9] Shu, J., Clyburn, S., Mates, T., and Baker, S. P., 1999, ''Effects of Thickness and Oxygen Content on Thermomechanical Behavior of Thin Cu Gilms Passivated With AlN,'' *Proceedings, Materials Reliability in Microelectronics IX, D. D. Brown*, et al., eds., Materials Research Society, Pittsburgh, PA, pp. 707–712.
[10] Keller, R.-M., Baker, S. P., and Arzt, E., 1998, ''Quantitative Analysis of Strengthening Mechanisms in Thin Cu films: Effects of Film Thickness, Grain Size and Passivation,'' J. Mater. Res., **13**, pp. 1307–1317.
[11] Wang, X., 2000, ''Bilayer Modeling and Measurements of Adhesion of Copper Thin Films to Glass,'' M.S. thesis, Cornell University, Ithaca, NY.

R. C. Cammarata
Department of Materials Science and Engineering and Department of Mechanical Engineering,
Johns Hopkins University,
Baltimore, MD 21218-2689
e-mail: rcc@jhu.edu

K. Sieradzki
Department of Mechanical and Aerospace Engineering and Center for Solid State Science,
Arizona State University,
Tempe, AZ 85287-6106
e-mail: karl@icarus.eas.asu.edu

Thermodynamics of Thin Film Epitaxy

The mechanics of thin film epitaxy is developed using an equilibrium thermodynamics formalism and linear elasticity. A virtual variation approach is employed that leads to a direct identification of the important volume and surface thermodynamic parameters characterizing mechanical equilibrium. In particular, the equilibrium volume stress state of an epitaxial film as a function of the film thickness, surface free energies, and surface stresses is obtained. It is shown how this formalism can be used to determine the critical thickness for epitaxy. [DOI: 10.1115/1.1468997]

Introduction

The mechanics of thin film epitaxy has been an area of experimental and theoretical interest for many years. One of the central issues involves determining the equilibrium coherency stress and strain states of the film as a function of the film thickness. A standard approach to analyzing this problem has been to minimize, with respect to the coherency strain, the sum of the volume strain energy of the film and the energy of the film-substrate interface for a given film thickness ([1]). However, it has been pointed out that this approach ignores the strain energy associated with the free surface of the film ([2,3]). This energy contribution can have a significant effect on the equilibrium state of the system, especially at small film thicknesses.

Recently, it has been shown that the thermodynamics of thin film epitaxy ([2,3]), as well as the kinetics of coherency strain relaxation ([3]), can be conveniently analyzed using the concepts of surface free energy and surface stress ([2–12]). The surface free energy is a scalar property defined to be the reversible work per unit area to create new surface. For a solid-fluid interface, the surface free energy is the surface work performed during processes such as cleavage and creep. The surface stress is a tensor that is associated with the reversible work to introduce an in-plane elastic strain at the surface. Physically, for the solid-fluid interface, the surface stress is related to the work to change the surface atom density while the surface free energy can be taken as the work to change the number of atoms at the surface. Although the surface free energy for a solid-fluid interface must be positive (otherwise the solid would spontaneously cleave), the components of the surface stress tensor can be positive or negative.

In a manner analogous to that used for a solid-fluid surface, it is possible to define a surface free energy and surface stresses for a solid-solid interface. Since there are two solid phases that can be independently strained, there are in general two distinct surface stresses for the solid-solid interface ([2,3,6,9]). The type of interface stress of interest for the epitaxy problem can be associated with the reversible surface work when one of the phases is stretched along the interface keeping the other phase fixed. A simple model for this interface stress has been given based on the reversible work associated with changing the density of interfacial misfit dislocations as a film is strained parallel to the interface, keeping the substrate fixed ([3]). (The other type of interface stress can be associated with the reversible surface work to stretch both phases along the interface equally; this type of interface stress can affect the intrinsic stress state of nonepitaxial films ([2,13]).)

In this paper, the mechanics of thin film epitaxy is developed using a formal equilibrium thermodynamics approach incorporating linear elasticity. The problem is formulated in a general way, and then the constraints and approximations needed to reduce unnecessary complexity while retaining the essential physics are identified. By carefully considering each of the deformations needed to introduce an elastic strain state in a film while keeping the in-plane area fixed, all of the important volume and surface thermodynamic parameters are obtained in a natural way. This analysis extends and generalizes previous discussions of the problem that treated the special case of an elastically isotropic system ([2,3]).

Coherency Stress and Strain States

Consider a Cartesian coordinate system with the 1-axis and 2-axis in the plane of the film and the 3-axis perpendicular to the plane of the plane of the film. Suppose an epitaxial film is partially or completely lattice matched to the substrate, resulting in uniform coherency stress and strain states in the film. Let the components of the coherency stress and strain tensors be denoted as σ_{ij} and ε_{ij}, respectively. The degree of lattice matching at the interface will determine the values of the in-plane strain components ε_{11}, ε_{12}, and ε_{22}. Since lattice matching results in a state of plane stress, the components σ_{13}, σ_{23}, and σ_{33} will all be zero. The remaining strain components ε_{13}, ε_{23}, and ε_{33} and the remaining stress components σ_{11}, σ_{12}, and σ_{22} can be determined by solving the simultaneous set of equations obtained from Hooke's law (for plane stress):

$$\sigma_{ij}=\sum_{k,l=1}^{3} C_{ijkl}\varepsilon_{kl}; \quad \varepsilon_{ij}=\sum_{k,l=1}^{2} S_{ijkl}\sigma_{kl}; \tag{1}$$

where C_{ijkl} and S_{ijkl} are the components of the elastic stiffnesses and compliances tensors, respectively, for the film. Although it is often convenient to use a Lagrangian measure of strain in surface thermodynamics problems involving solids ([8,9]), a Eulerian measure will used here since the surface areas in the problem to be treated will then be invariant.

For most cases of interest, it will be energetically favorable for an epitaxial film with a small film thickness to be completely lattice-matched with the substrate if the misfit between film and substrate is not too large. Since the volume strain energy of the film will scale with thickness, there is a critical thickness above which it is energetically favorable to reduce the degree of lattice matching by introducing misfit dislocations at the film-substrate interface ([1]). This is because the resulting reduction in volume strain energy of a film will more than compensate the increase 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, March 15, 2001; final revision, December 10, 2001. Associate Editor: D. A. Kouris. Discussion on the paper should be addressed to the Editor, Professor Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

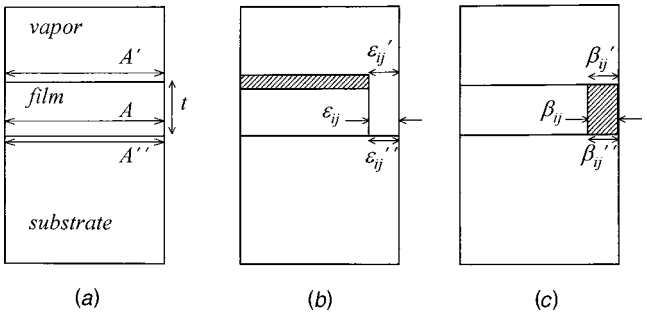


Fig. 1 (*a*) Thermodynamic system for thin film epitaxy, (*b*) application of in-plane elastic strains to the film and surfaces, (*c*) application of in-plane creep strains to the film and surfaces that return the in-plane film area and surface areas to their initial values.

the energy resulting from the creation of the misfit dislocations. However, it is also important to recognize that there will be a contribution to the change in the energy of the system resulting from the concomitant elastic deformation at the free surface of the film that can be positive or negative, and thus is an effect that can either impede or promote lattice matching.

Thermodynamic System

Figure 1(*a*) schematically illustrates the system under consideration. It is composed of a thin solid film of thickness t and in-plane area A that is bounded by two parallel, planar surfaces. One surface of area A' is the boundary between the film and a vapor phase (henceforth referred to as the free surface of the film) and the other surface of area A'' is the solid-solid interface between the film and a much thicker substrate. The system is constrained so that during any change in state, all of the surfaces remain parallel and planar, and the areas A, A', and A'' are equal to each other and remain fixed. This means that a change in the volume of one of the phases is a result of a change in dimension along the 3-axis. In the case of the thin film phase, the volume of can be expressed as At, and a change in volume of the film is associated with a change in film thickness t. It will also be assumed that the in-plane dimension of the film is much larger than the thickness so that thin film edge effects can be neglected.

Suppose the film experiences a uniform, in-plane elastic strain as shown schematically in Fig. 1(*b*). Let the in-plane components of this strain be denoted as ε_{ij}, where $i,j=1,2$. In order to maintain the constraint that the in-plane area A remains unchanged, it is necessary for some atomic rearrangement, analogous to a creep process, to result in an in-plane ''creep'' strain with components $\beta_{11}=-\varepsilon_{11}$ and $\beta_{22}=-\varepsilon_{22}$ (see Fig. 1(*c*)). It is noted that this atomic rearrangement leaves the volume of the film (after it has been elastically strained) unchanged. The in-plane area as a function of strain can be expressed as

$$A=A_o\sum_{i,j=1}^{2}[1+(\varepsilon_{ij}+\beta_{ij})\delta_{ij}], \tag{2}$$

where A_o is the area at zero strain and δ_{ij} is the Kronecker delta. Since the area of the free surface of the film, A', and the film-substrate interface area A'' are constrained to remain constant and equal to A, they can also be expressed as functions of surface elastic strains ε_{ij}', ε_{ij}'' and surface creep strains β_{ij}', β_{ij}'':

$$A'=A_o'\sum_{i,j=1}^{2}[1+(\varepsilon_{ij}'+\beta_{ij}')\delta_{ij}], \tag{3}$$

$$A''=A_o'\sum_{i,j=1}^{2}[1+(\varepsilon_{ij}''+\beta_{ij}'')\delta_{ij}]. \tag{4}$$

Thermodynamic Formalism

An important principle of thermodynamics is that for a system that is isolated from its surroundings and is constrained to have a fixed total entropy, the equilibrium state is the one for which the internal energy is a minimum. From this principle, it is possible to obtain the conditions for thermodynamic equilibrium by using the virtual variation method developed by Gibbs [14–17]. The physical interpretation of a virtual variation has been discussed in detail by Reiss [16]. Mathematically, it involves setting an infinitesimal variation in the internal energy U of a system composed of n components equal to zero, holding the entropy S, volume V, and number of moles of the components N_1, N_2, ... N_n constant:

$$\delta U_{S,V,N_1,N_2,\ldots N_n}=0. \tag{5}$$

In addition to the keeping S, V, N_1, N_2, ... N_n fixed, it is also necessary to include the effects of all other special constraints imposed on the system during the variation.

The internal energy of a thermodynamic system is a function of the extensive variables ([15,17]). For the system of Fig. 1, these variables are S, V, N_1, N_2, ... A, A', and A''. Since the areas are functions of the strains ε_{ij}, ε_{ij}', ε_{ij}'', β_{ij}, β_{ij}', β_{ij}'', the internal energy can be expressed as

$$\begin{aligned}U=U(&S,V,N_1,N_2,\ldots N_n,A_o\varepsilon_{11},A_o\varepsilon_{12},A_o\varepsilon_{22},A_o\beta_{11},\\&A_o\beta_{12},A_o\beta_{22},A_o'\varepsilon_{11}',A_o'\varepsilon_{12}',A_o'\varepsilon_{22}',A_o'\beta_{11}',A_o'\beta_{12}',\\&A_o'\beta_{22}',A_o''\varepsilon_{11}'',A_o''\varepsilon_{12}'',A_o''\varepsilon_{22}'',A_o''\beta_{11}'',A_o''\beta_{12}'',A_o''\beta_{22}'').\end{aligned} \tag{6}$$

It has been noted previously that lattice matching between the film and substrate results in an in-plane coherency stress, so that the film is in a nonhydrostatic stress state. As a result, the conditions for thermodynamic equilibrium are more complex than those obtained for the case of a solid that is hydrostatically stressed ([18]). For example, for a system involving a hydrostatically stressed crystalline solid in equilibrium with a fluid, it is possible to define the chemical potential of the solid for component k as $\mu_k^s=(\partial U^s/\partial N_k^s)$, where the superscript s refers to the value for the solid, as is done in conventional bulk thermodynamics. However, for a nonhydrostatically stressed crystalline solid, μ_k^s for a component that predominantly occupies lattice sites is in general not well defined. Instead, it is necessary to define a different parameter, called the diffusion potential M_{kh}, for components h and k that occupy lattice sites in the solid. M_{kh} involves the change in internal energy of the system when an atom of component h in the crystal is exchanged with an atom of component k in the fluid at constant entropy, volume, and number of moles of the other components. Further complicating the analysis is that there is in general a coupling among variations involving the number of moles of the different components and variations associated with the different strains.

In order to reduce unnecessary complexity, it is useful to introduce certain simplifying constraints and assumptions that do not significantly alter the basic physics of the problem. Each of the phases will be constrained not to exchange matter with the other phases, so that there can be no variation in the number of moles of any of the components within the phases. In addition, it will be assumed that the bulk of the phases are compositionally uniform, and that both of the surfaces can be treated as a Gibbs ''dividing surface'' ([5,7,19]). Using this construct, the thermodynamic quantities associated with bulk phases are calculated assuming the phases are uniform, even though the actual values of these quantities may vary in the proximity of the surfaces. Any ''excess'' values of the thermodynamic quantities of the system, equal to the difference between the actual values for the system and those obtained assuming the bulk phases are uniform, are then considered surface thermodynamic quantities. Gibbs showed that for fluid systems, this allows a proper accounting of all of the thermodynamic quantities and will rigorously produce the conditions for describing thermodynamic equilibrium. Although it may be

that there is not the same degree of rigor when applying this approach to systems involving nonhydrostatically stressed solids, it should be an adequate approximation for the present problem [20]. Finally, it will be assumed that the bulk of the substrate remains undeformed during any variation in state of the system, and that the vapor pressure is much smaller than the magnitude of the coherency stresses in an epitaxial film.

Conditions for Equilibrium

The virtual variation of the internal energy U as given in Eq. (6), incorporating the constraint that there can be no exchange of matter among the phases, can be expressed as

$$\delta U = T\,\delta S - P\,\delta V + \sum_{i,j=1}^{2}\left[\left(\frac{\partial U}{\partial \varepsilon_{ij}}\right)\delta\varepsilon_{ij} + \left(\frac{\partial U}{\partial \beta_{ij}}\right)\delta\beta_{ij} + \left(\frac{\partial U}{\partial \varepsilon'_{ij}}\right)\delta\varepsilon' + \left(\frac{\partial U}{\partial \beta'_{ij}}\right)\delta\beta'_{ij} + \left(\frac{\partial U}{\partial \varepsilon''_{ij}}\right)\delta\varepsilon''_{ij} + \left(\frac{\partial U}{\partial \beta''_{ij}}\right)\delta\beta''_{ij}\right], \tag{7}$$

where the definitions $T \equiv (\partial U/\partial S)$ and $P \equiv -(\partial U/\partial V)$ have been used. According to Eq. (6), the conditions for equilibrium can be obtained by setting the variation in Eq. (7) equal to zero, keeping S and V fixed. Since the entropy can be varied independently of the other variations on the right-hand side of Eq. (7), one condition for equilibrium is that

$$T\,\delta S = T^{v}\,\delta S^{v} + T^{f}\,\delta S^{f} + T^{s}\,\delta S^{s} = 0, \tag{8}$$

when the total entropy of the system is held fixed:

$$\delta S = \delta S^{v} + \delta S^{f} + \delta S^{s} = 0. \tag{9}$$

In Eqs. (8) and (9), the superscripts v, f, and s denote vapor, film, and substrate, respectively. The only way in which Eqs. (8) and (9) can be simultaneously satisfied is for $T^{v} = T^{f} = T^{s}$. Identifying T as the temperature means that at equilibrium, the temperature is uniform in the system (condition for thermal equilibrium).

Substituting Eq. (8) into Eq. (7) and setting the variation δU equal to zero gives

$$-P\,\delta V + \sum_{i,j=1}^{2}\left[\left(\frac{\partial U}{\partial \varepsilon_{ij}}\right)\delta\varepsilon_{ij} + \left(\frac{\partial U}{\partial \beta_{ij}}\right)\delta\beta_{ij} + \left(\frac{\partial U}{\partial \varepsilon'_{ij}}\right)\delta\varepsilon' + \left(\frac{\partial U}{\partial \beta'_{ij}}\right)\delta\beta'_{ij} + \left(\frac{\partial U}{\partial \varepsilon''_{ij}}\right)\delta\varepsilon''_{ij} + \left(\frac{\partial U}{\partial \beta''_{ij}}\right)\delta\beta''_{ij}\right] = 0. \tag{10}$$

The terms in Eq. (10) with the variations $\delta\varepsilon_{ij}$ are the work to elastically strain the volume of the film, and the terms with the variations $\delta\varepsilon'_{ij}$ and $\delta\varepsilon''_{ij}$ are the concomitant work terms to stretch the free surface and film-substrate interface, respectively. Since the process of atomic rearrangement, as illustrated in Fig. 1(*b*), does not change the volume of the film, there is no work associated with the terms involving the creep strain $\delta\beta_{ij}$, and thus $(\partial U/\partial\beta_{ij}) = 0$. The terms with the variations $\delta\beta'_{ij}$ and $\delta\beta''_{ij}$ represent the work to create new surface area by adding atoms at the free surface and film-substrate interface, respectively. Since the shear strain variations $\delta\beta'_{12}$ and $\delta\beta''_{12}$ do not create new area, $(\partial U/\partial\beta'_{12}) = (\partial U/\partial\beta''_{12}) = 0$. It is possible to associate all of the nonzero partial derivatives with certain film or surface parameters, so that Eq. (10) can be rewritten as

$$-P\,\delta V + \sum_{i,j=1}^{2}\left[V^{f}\sigma_{ij}\delta\varepsilon_{ij} + A'f_{ij}\delta\varepsilon'_{ij} + A'\gamma\delta\beta'_{ij} + A''g_{ij}\delta\varepsilon''_{ij} + A''\psi\delta_{ij}\delta\beta''_{ij}\right] = 0. \tag{11}$$

In Eq. (11), σ_{ij} is the volume stress state of the film, f_{ij} and γ are the surface stress and surface free energy, respectively, for the free surface of the film, and g_{ij} and ψ are the surface stress and surface free energy, respectively, for the film-substrate interface. Consistent with the discussion given previously regarding the physical meaning of the surface thermodynamic properties, it is noted that the partial derivatives in the work terms associated with creating new surface area by introducing the creep strains $\delta\beta'_{ij}$ and $\delta\beta''_{ij}$ are related to the scalar quantities γ and ψ, while the partial derivatives in the work terms involving elastic deformations are related to the stress tensors σ_{ij}, f_{ij}, and g_{ij}. It is also recalled that the surface stress g_{ij} is associated with the surface work when one phase (the film) is elastically stretched keeping the phase (the substrate) fixed.

Consideration is now given to the effects of constraining the areas A, A', and A'' to remain fixed at the constant value A_o during the virtual variation. It is seen from Eqs. (2), (3), and (4) that these area constraints lead to the relations $\delta\beta_{11} = -\delta\varepsilon_{11}$, $\delta\beta_{22} = -\delta\varepsilon_{22}$, $\delta\beta'_{11} = -\delta\varepsilon''_{11}$, $\delta\beta'_{22} = -\delta\varepsilon'_{22}$, $\delta\beta''_{11} = -\delta\varepsilon''_{11}$, and $\delta\beta''_{22} = -\delta\varepsilon''_{22}$. Using these relations, expressing $V^{f} = At = A_o t$, and employing the assumption that $-P\,\delta V$ is negligible compared to the other work terms, Eq. (11) can be rewritten as

$$A_o \sum_{i,j=1}^{2}\left[t\sigma_{ij} + f_{ij} + g_{ij} - (\gamma+\psi)\delta_{ij}\right]\delta\varepsilon_{ij} = 0, \tag{12}$$

which leads to the key result:

$$t\sigma_{ij} = (\gamma+\psi)\delta_{ij} - f_{ij} - g_{ij}. \tag{13}$$

Equation (13) characterizes the system at mechanical equilibrium, and can be considered as representing a force balance between the film stress σ_{ij} and a "surface pressure" $[f_{ij} + g_{ij} - (\gamma+\psi)\delta_{ij}]/t$.

Critical Film Thickness

It is seen from Eq. (13) that in the limit of an infinitely large film thickness t, the equilibrium coherency stress, and therefore the equilibrium coherency strain, of the film approaches zero. As the film thickness is reduced, the magnitude of the components of the equilibrium coherency stress and strain tensors increase until a critical thickness is reached, corresponding to the largest thickness for total lattice matching between film and substrate to be thermodynamically favorable.

Equation (13) can be used to determine the critical thickness when the stress tensors and surface free energies are evaluated at the strain state corresponding to a completely lattice matched film. Let the in-plane coherency strain components for such a film be denoted as ε^{c}_{ij} for $i, j = 1,2$. The other strain components ε^{c}_{13}, ε^{c}_{23}, and ε^{c}_{33}, as well as the components of the coherency stress tensor σ^{c}_{ij} can be obtained using Hooke's law for plane stress as given in Eqs. (1). A surface force tensor B_{ij} can be defined as

$$B_{ij} = (\gamma+\psi)\delta_{ij} - f_{ij} - g_{ij}. \tag{14}$$

Let B^{c}_{ij} denote the value of the surface force tensor evaluated at the strain state for a completely lattice matched film. In general, it is expected that f_{ij} and γ will be only weakly dependent on the coherency strain. Recently, a simple model for the surface thermodynamic parameters of a solid-solid interface suggested that although ψ is expected to have a strong dependence on the strain, $g_{ij} - \psi\delta_{ij}$ can be taken as approximately independent of strain ([3]). The critical thickness for epitaxy corresponds to the value of t when the representation ellipse (quadric) for B^{c}_{ij}/t makes tangential contact with the representation ellipse for σ^{c}_{ij}. This is illustrated in Fig. 2. Figure 2(*a*) refers to the general case when σ^{c}_{ij} and B^{c}_{ij} do not share the same set of principal axes. If σ^{c}_{ij} and B^{c}_{ij} do share the same set of principal axes, tangential contact is made along one of those axes, as illustrated in Fig. 2(*b*). In this case, the critical thickness can be determined from Eq. (13) using the principal values for the tensor components. If the system displays threefold or higher rotational symmetry about the 3-axis, the surface stress tensors can be taken as scalars f and g, and Hooke's

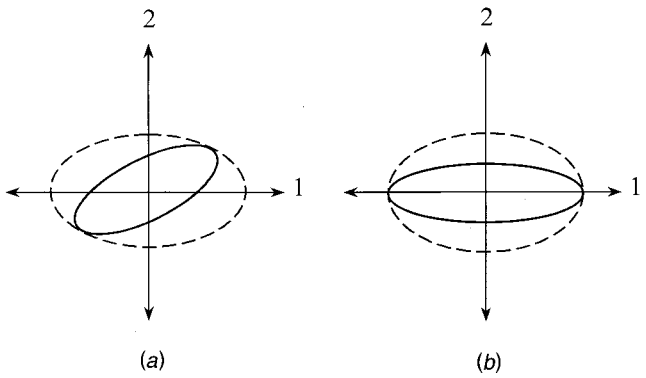


Fig. 2 Representation ellipses for σ_{ij}^c (dashed) and B_{ij}^c/t (solid) when t is equal to the critical thickness; (*a*) general system where σ_{ij}^c and B_{ij}^c/t do not share the same set of principal axes, (*b*) system where σ_{ij}^c and B_{ij}^c/t share the same set of principal axes

law for the coherency stress can be written as $\sigma^c = Ym$, where Y is the biaxial (plane stress) elastic modulus and m is the total film-substrate misfit. The critical thickness t^c will then be

$$t^c = (\gamma + \psi - f - g)/Ym. \tag{15}$$

It has been assumed that the surface force, equal to $(\gamma + \Psi - f - g)$ for an isotropic system, has the same sign as that for the components of the strain tensor when the film is partially or completely coherent, and this is expected to be true for the vast majority of systems. As mentioned earlier, a model has been presented ([2,3]) in which an expression for g was derived by considering the surface work to change the density of misfit dislocations at the film-substrate interface. For most cases of interest, the model predicts that g will be larger in magnitude than $(\gamma + \Psi - f)$ and have a sign opposite of m. However, the model expression for g has dependence on the film thickness t such that the magnitude of g can be significantly reduced for small values of t. In certain systems this results in the surface force $(\gamma + \Psi - f - g)$ and misfit m having opposite signs, so that Eq. (15) gives the unphysical result $t^c < 0$. What actually occurs is the development of an in-plane elastic strain in the film that will have a sign opposite that for the misfit m. This anomalous coherency strain behavior generated in such systems has been discussed elsewhere ([3]).

Acknowledgments

The authors thank Profs. J. W. Cahn, D. Kouris, K. T. Ramesh, F. Spaepen, J. B. Spicer, and K. M. Unruh for useful discussions. R. C. Cammarata gratefully acknowledges support from the National Science Foundation administered through the Materials Research Science and Engineering Center at Johns Hopkins.

References

[1] Matthews, J. M., 1975, "Coherent Interfaces and Misfit Dislocations," *Epitaxial Growth, Part B*, J. M. Matthews, eds., Academic Press, San Diego, CA, pp. 559–609.
[2] Cammarata, R. C., 1994, "Surface and Interface Stress Effects in Thin Films," Prog. Surf. Sci., **46**, pp. 1–38.
[3] Cammarata, R. C., Sieradzki, K., and Speapen, F., 2000, "Simple Model for Interface Stresses With Application to Misfit Dislocation Generation in Epitaxial Thin Films," J. Appl. Phys., **87**, pp. 1227–1234.
[4] Shuttleworth, R., 1950, "The Surface Tension of Solids," Proc. R. Soc. London, Ser. A, **63**, pp. 444–457.
[5] Herring, C., 1952, "The Use of Classical Macroscopic Concepts in Surface-Energy Problems," *Structure and Properties of Solid Surfaces*, edited by R. Gomer and C. S. Smith, eds., University of Chicago Press, Chicago, pp. 5–81.
[6] Brooks, H., 1963, "Theory of Boundaries," *Metals Interfaces*, American Society of Metals, Metals Park, Ohio, pp. 20–64.
[7] Mullins, W. W., 1963, "Solid Surface Morphologies Governed by Capillarity," *Metal Surfaces: Structure, Energetics, and Kinetics*, ASM, Cleveland, OH, pp. 17–64.
[8] Cahn, J. W., 1980, "Surface Stress and the Chemical Equilibrium of Small Crystals: I. The Case of the Isotropic Surface," Acta Metall., **28**, pp. 1333–1338.
[9] Cahn, J. W., and Larché, F., 1982, "Surface Stress and Chemical Equilibrium of Small Crystals. II. Solid Particles Embedded in a Solid Matrix," Acta Metall., **30**, pp. 51–56.
[10] Suo, Z., and Lu, W., 2000, "Composition Modulation and Nanophase Separation in a Binary Epilayer," J. Mech. Phys. Solids, **48**, pp. 211–232.
[11] Kouris, D., Peralta, P., and Sieradzki, K., 2000, "Surface Islands and Their Elastic Interaction With Adatoms," Surf. Sci., **445**, pp. 420–429.
[12] Spaepen, F., 2000, "Interfaces and Stresses in Thin Films," Acta Mater., **48**, pp. 31–42.
[13] Cammarata, R. C., Trimble, T. M., and Srolovitz, D. J., 2000, "Surface Stress Model for Intrinsic Stresses in Thin Films," J. Mater. Sci., **15**, pp. 2468–2474.
[14] Gibbs, J. W., 1961, "On the Equilibrium of Heterogeneous Substances," *The Scientific Papers of J. Willard Gibbs, Volume 1: Thermodynamics*, Dover, New York, pp. 55–353.
[15] Callen, H. B., 1985, *Thermodynamics and an Introduction to Thermostatics*, John Wiley and Sons, New York.
[16] Reiss, H., 1996, *Methods of Thermodynamics*, Dover, New York.
[17] Tschoegl, N. W., 2000, *Fundamentals of Equilibrium and Steady State Thermodynamics*, Elsevier, Amsterdam.
[18] Larché, F. C., and Cahn, J. W., 1985, "The Interactions of Composition and Stress in Crystalline Solids," Acta Metall., **33**, pp. 331–357.
[19] Zangwill, A., 1988, *Physics at Surface*, Cambridge University Press, Cambridge, UK, pp. 7–12.
[20] Johnson, W. C., 2000, "Superficial Stress and Strain at Coherent Interfaces," Acta. Mater., **48**, pp. 434–444.

Domain Dynamics in a Ferroelastic Epilayer on a Paraelastic Substrate

Y. F. Gao

Z. Suo

Mechanical and Aerospace Engineering Department and Princeton Materials Institute, Princeton University, Princeton, NJ 08544

This paper models the domain dynamics in a ferroelastic epilayer within the time-dependent Ginzburg-Landau (TDGL) framework. Constrained on a paraelastic substrate of square symmetry, the epilayer has rectangular symmetry, and forms domains of two variants. The domain wall energy drives the domains to coarsen. The spontaneous strains induce an elastic field, which drives the domains to refine. The competition between coarsening and refining selects an equilibrium domain size. We model the epilayer-substrate as a nonequilibrium thermodynamic system, evolving by the changes in the elastic displacements and the order parameters. The free energy consists of two parts: the bulk elastic energy, and the excess surface energy. The elastic energy density is taken to be quadratic in the strains. The surface energy density is expanded into a polynomial of the order parameters, the gradients of the order parameters, and the strains. In this expansion, the surface stress is taken to be quadratic in the order parameters. The evolution equations are derived from the free energy variation with respect to the order parameters. The elastic field is determined by superposing the Cerruti solution. Examples of computer simulation are presented. [DOI: 10.1115/1.1469000]

1 Introduction

The thermodynamics and kinetics of bulk ferroelectric materials have been extensively studied ([1]). However, there is little systematic study on the ferroelectric thin films ([2,3]). This paper considers some important issues on the domain dynamics in a ferroelastic epilayer on a paraelastic substrate. Consider a ferroelastic epilayer on a paraelastic substrate. The epilayer is much thinner than the substrate. They are coherent, accommodating the lattice mismatch by elastic deformation. At high temperatures, the epilayer and the substrate surface both have square symmetry, but have different lattice constants. To compensate for the lattice mismatch, the epilayer is under a uniform, biaxial stress state, and the substrate is unstressed. At low temperatures, the epilayer has rectangular symmetry, but the substrate surface still has square symmetry. Because of the broken symmetry, the epilayer now has two variants, equivalent upon a 90-deg rotation (Fig. 1). Within the Ginzburg-Landau framework ([1]), we characterize the ferroelastic state by two order parameters, (p_1, p_2), taken to be the components of a vector lying in the plane of the layer. If the epilayer were unconstrained by the substrate, the two variants would have different spontaneous strain states, equivalent after the 90-deg rotation. This paper considers the epilayer constrained on the substrate. It is sometimes assumed in the literature that the energy ground state is a single-variant epilayer on the substrate. This state is illustrated in Fig. 2(*a*). The stress state in the epilayer is uniform and biaxial, with unequal components in the two directions. The substrate is unstressed. Assuming that the epilayer is of a single variant, Pertsev et al. [2] showed that the constraint of the substrate could alter the Curie temperature and even induce new phases.

However, the epilayer of a single variant on the substrate is not the energy ground state ([3–7]). This is readily understood as follows. In Fig. 2(*b*), the two variants coexist and form domains. To aid the argument, imagine that suitable forces are applied on the domain walls, so that the stress state in the old variant is unchanged, the stress state in the new variant is equivalent to the old one after the 90-deg rotation, and the substrate remains unstressed. Consequently, the state in Fig. 2(*b*) has the same elastic energy as the state in Fig. 2(*a*). Now gradually reduce the forces on the domain walls, and allow the epilayer and the substrate to deform. After the forces are completely removed (Fig. 2(*c*)), the displacements are generally in directions opposite from the applied forces in Fig. 2(*b*). Consequently, after elastic relaxation, the two-variant state in Fig. 2(*c*) has lower elastic energy than the single-variant state in Fig. 2(*a*). That is, when the symmetry is broken, due to the substrate constraint, the single-variant epilayer is no longer the energy ground state. This suggests that Pertsev et al.'s assumption of a single variant is excessively simplistic and questions the validity of their results.

Figure 3 illustrates a domain pattern of a periodic array of alternating variants. The smaller the domain size, the more elastic energy is relaxed. Consequently, elasticity drives domains to refine. On the other hand, variants coexist at the cost of adding the domain wall energy. The smaller the domain size, the longer the collective domain walls. That is, the domain wall energy drives the domain to coarsen. The competition between the elastic energy and the domain wall energy selects an equilibrium domain size, which minimizes the combined energy ([3–14]). It is also expected that the minimization of the combined energy can select an equilibrium domain pattern.

This paper introduces a model within the time-dependent Ginzburg-Landau (TDGL) framework to simulate domain dynamics in the epilayer. The epilayer may have different properties from the bulk, particularly if the epilayer is only a few monolayers thick. It may also be possible that a marginally paraelastic bulk crystal is ferroelastic within a few surface layers. Motivated by these considerations, we will model the epilayer-substrate as a single system. The model parallels that for an epilayer with composition modulation ([12–14]). In the present model, the free energy of the system consists of two parts: the bulk elastic energy and the excess surface energy. The elastic energy density is taken to be quadratic in the strains. The excess surface energy density is expanded into a polynomial of the order parameters, the gradients

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, March 15, 2001; final revision, December 22, 2001. Associate Editor: D. A. Kouris. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

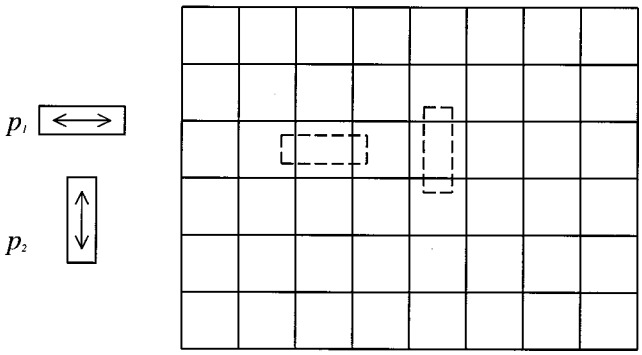


Fig. 1 Broken symmetry and lattice mismatch. The substrate has square symmetry. The epilayer has rectangular symmetry, with two variants.

of the order parameters, and the strains. In this expansion, the surface stress is taken to be quadratic in the order parameters. We derive the evolution equations from the free-energy variation associated with variation in the order parameters, and determine the elastic field by superposing the Cerruti solution. To illustrate the model, we present preliminary results from computer simulations.

The TDGL framework has been applied to bulk ferroelastic crystals ([15–17]). In a bulk ferroelastic single crystal, however, elasticity does not refine domains. In fact, the single-domain crystal is the energy ground state. The bulk elasticity problem lacks an intrinsic length scale. For an infinite, polydomain crystal, the total elastic energy in the crystal is independent of the domain size so long as the domain pattern is self-similar as domains coarsen. In this case, domains coarsen to reduce the domain wall energy. The situation is different if the system is not an infinite single crystal. For a thin ferroelastic film on a substrate ([4–9]), the film thickness provides a length scale. In a polycrystalline crystal, the grain size provides a length scale ([18,19]). In both cases, elasticity does limit the size of the domains. In the model introduced in this paper, the length scale is provided by the introduction of the surface stress, as will be identified in a later section.

2 Free Energy and Kinetic Law

Even for a single crystal, atoms in a few surface layers have different energy from atoms in the bulk. The deposition of the epilayer further changes the energy state. We model the substrate-epilayer system as a bulk solid coupled with a superficial object. The total free energy consists of two parts: the bulk elastic energy and the excess surface energy, written as

$$G=\int WdV+\int \Gamma dA, \tag{1}$$

where W is the elastic energy density, and Γ is the excess surface energy density. The first integral is over the volume of the entire system, and the second integral is over the surface area covered by the epilayer. Following [14], we interpret the surface energy as the excess free energy relative to bulk elastic energy. Thus, Γ includes the effects of the mismatch between the two materials, as well as the presence of the empty space above.

In the rectangular coordinate system, the surface coincides with the (x_1,x_2) plane, and the material occupies the half-space $x_3<0$. Reference the displacement vector in the system u_i from an infinite, unstressed crystal of the same composition as the substrate. The strain tensor ε_{ij} relates to the displacement gradient, namely,

$$\varepsilon_{ij}=\tfrac{1}{2}(u_{i,j}+u_{j,i}). \tag{2}$$

The Latin subscripts run from 1 to 3.

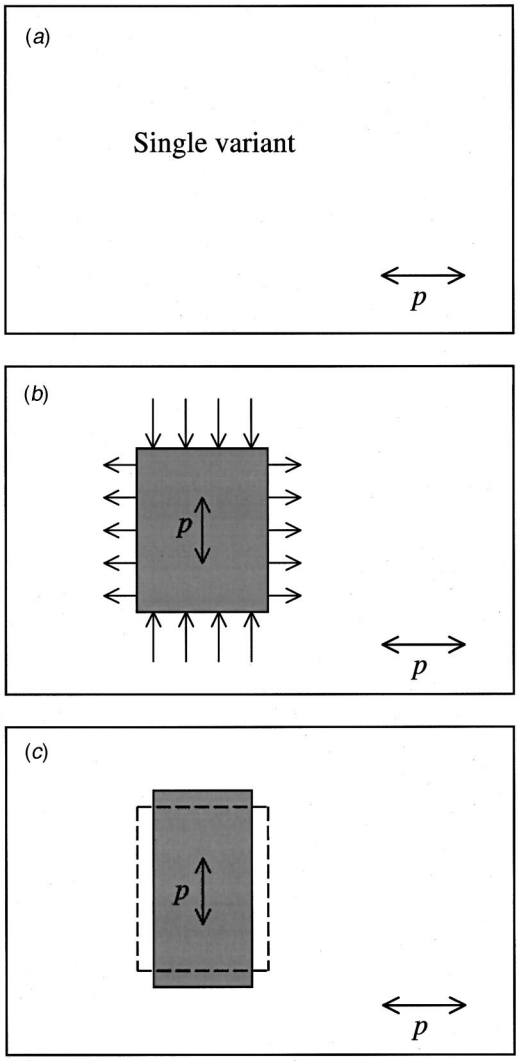


Fig. 2 Schematic illustration of elastic refining; (*a*) a single-variant epilayer, (*b*) a two-variant epilayer with appropriate forces applied around the new domain, (*c*) the external forces removed

We assume that the substrate is elastically isotropic, μ being the shear modulus, and ν Poisson's ratio. The elastic energy density W is quadratic in the strain tensor:

$$W=\mu\left[\varepsilon_{ij}\varepsilon_{ij}+\frac{\nu}{1-2\nu}(\varepsilon_{kk})^2\right]. \tag{3}$$

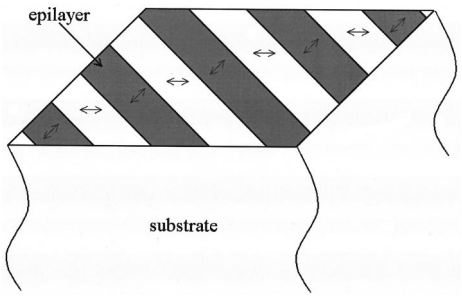


Fig. 3 A domain pattern in the epilayer constrained on the substrate

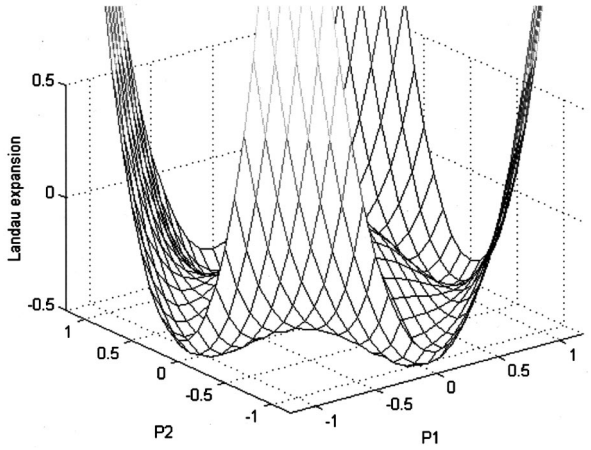


Fig. 4 The Landau expansion as a function of the order parameters. The four wells correspond to spontaneous states.

The stresses σ_{ij} in the bulk are the differential coefficients of the elastic energy density:

$$\delta W = \sigma_{ij}\delta\varepsilon_{ij}. \tag{4}$$

The repeated subscripts imply the summation convention.

We characterize the ferroelastic state of the epilayer by two order parameters, p_1 and p_2, which form the components of a vector lying in the plane of the epilayer. For example, when the layer is also ferroelectric, this vector coincides with the polarization. The excess surface energy density Γ is a function of the order parameters, the gradients of the order parameters, and the strains. Following the Ginzburg-Landau formalism, we expand Γ into a Taylor series to the lowest terms in $p_{\alpha,\beta}$ and $\varepsilon_{\alpha\beta}$:

$$\Gamma = f_L(p_\xi) + f_G(p_{\alpha,\beta}) + f_{\alpha\beta}(p_\xi)\varepsilon_{\alpha\beta}. \tag{5}$$

The Greek letters run from 1 to 2. The physical content of every term in (5) is interpreted as follows.

The first term in (5) is the Landau expansion. We assume that the epilayer has square symmetry in the high-temperature phase. The Landau expansion takes the following form ([15–17]):

$$f_L = a_0 + a_1(p_1^2+p_2^2) + a_{11}(p_1^4+p_2^4) + a_{12}p_1^2p_2^2 + a_{111}(p_1^6+p_2^6) + a_{112}(p_1^4p_2^2+p_1^2p_2^4). \tag{6}$$

Here a_0 is the surface energy density of the epilayer on the substrate when the epilayer is in the paraelastic phase ($p_\alpha=0$) and the substrate is unstrained ($\varepsilon_{\alpha\beta}=0$). The coefficient a_1 is proportional to $T-T_C$, i.e., the temperature difference from critical point, T_C. When $T<T_C$, $a_1<0$, and the function has four wells (Fig. 4). Each well corresponds to an unconstrained ferroelastic variant. There is no reason to regard the coefficients in (6) to be the same as those of bulk materials.

The second term in (5) stands for the gradient energy, which is taken to be quadratic in the gradients of the order parameters ([15–17]):

$$f_G(p_{\alpha,\beta}) = \tfrac{1}{2}g_{11}(p_{1,1}^2+p_{2,2}^2) + g_{12}p_{1,1}p_{2,2} + \tfrac{1}{2}g_{44}(p_{1,2}+p_{2,1})^2 + \tfrac{1}{2}g'_{44}(p_{1,2}-p_{2,1})^2. \tag{7}$$

The g parameters are positive constants. This is a continuum description of the domain wall energy.

The third term in (5) gives the strain dependence. We have only included the linear terms of the strains, neglecting the excess elastic constants of the epilayer relative to the substrate. The latter could also be included as a refinement of the theory. The coefficients $f_{\alpha\beta}$ are the components of the surface stress tensor. The surface stress tensor has been incorporated into the continuum elasticity theory by adding a strain-dependent surface energy to the free energy [10–14]. In this paper, we expand the surface stress $f_{\alpha\beta}$ in terms of the order parameters:

$$\begin{aligned} f_{11} &= f_0 - \tfrac{1}{2}\beta_{11}p_1^2 + \tfrac{1}{2}\beta_{12}p_2^2 \\ f_{22} &= f_0 + \tfrac{1}{2}\beta_{12}p_1^2 - \tfrac{1}{2}\beta_{11}p_2^2 \\ f_{12} &= -\beta_{44}p_1p_2. \end{aligned} \tag{8}$$

Here f_0 is the surface stress when the epilayer is in the paraelastic state. In the ferroelastic state, the surface stress is taken to be quadratic in the order parameters. The surface stress couples the domain pattern in the epilayer to the elastic deformation in the substrate.

The epilayer-substrate is a nonequilibrium thermodynamic system. The system can vary by two means: the elastic displacements and the order parameters. Characterize a virtual change of the system by δu_i and δp_α. The free-energy variation associated with this virtual change is

$$\delta G = \int\left[\frac{\partial f_L}{\partial p_\alpha} - \frac{\partial}{\partial x_\beta}\left(\frac{\partial f_G}{\partial p_{\alpha,\beta}}\right) + \frac{\partial f_{\xi\eta}}{\partial p_\alpha}\varepsilon_{\xi\eta}\right]\delta p_\alpha dA + \int\left[\left(\sigma_{3\alpha} - \frac{\partial f_{\alpha\beta}}{\partial x_\beta}\right)\delta u_\alpha + \sigma_{33}\delta u_3\right]dA + \int \sigma_{ij,j}\delta u_i dV. \tag{9}$$

We assume that elastic relaxation is much faster than domain switching. At a given time the system reaches elastic equilibrium instantaneously. Consequently, the energy variation associated with the elastic displacements vanishes, leading to the familiar equilibrium equations in elasticity:

$$\sigma_{ij,j} = 0, \tag{10}$$

and the boundary conditions:

$$\sigma_{3\alpha} = \frac{\partial f_{\alpha\beta}}{\partial x_\beta}, \quad \sigma_{33} = 0. \tag{11}$$

Associated with the virtual change of the order parameters, the variation of the free energy defines a thermodynamic force, F_α, namely,

$$\int F_\alpha \delta p_\alpha dA = -\delta G. \tag{12}$$

A comparison of (9) and (12) gives that

$$F_\alpha = -\left[\frac{\partial f_L}{\partial p_\alpha} - \frac{\partial}{\partial x_\beta}\left(\frac{\partial f_G}{\partial p_{\alpha,\beta}}\right) + \frac{\partial f_{\xi\eta}}{\partial p_\alpha}\varepsilon_{\xi\eta}\right]. \tag{13}$$

Following [12–14], we adopt a linear kinetic law that the rate of the change of the order parameters is proportional to the driving force, namely,

$$\frac{\partial p_\alpha}{\partial t} = LF_\alpha, \tag{14}$$

where L is the kinetic coefficient. Combining (13) and (14), we obtain the evolution equations for the order parameters:

$$\frac{\partial p_\alpha}{\partial t} = -L\left[\frac{\partial f_L}{\partial p_\alpha} - \frac{\partial}{\partial x_\beta}\left(\frac{\partial f_G}{\partial p_{\alpha,\beta}}\right) + \frac{\partial f_{\xi\eta}}{\partial p_\alpha}\varepsilon_{\xi\eta}\right]. \tag{15}$$

Assuming elastic equilibrium and using the divergence theorem, we obtain the free energy as follows:

$$G = \int [f_L + f_G - \tfrac{1}{2}\sigma_{3\alpha}u_\alpha]dA. \tag{16}$$

The integral extends over the surface area.

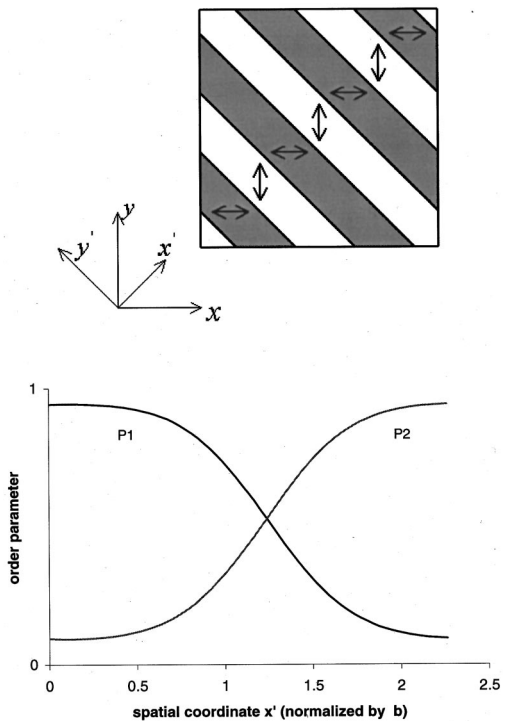


Fig. 5 Domain wall shape. The domain width is on the order of *b*.

3 Dimensionless Equations, Length Scales, and Time Scale

The four wells in the function f_L (Fig. 4) correspond to four states of spontaneous polarization. The value of the spontaneous polarization p_s satisfies

$$a_1+2a_{11}p^2+3a_{111}p^4=0, \tag{17}$$

giving

$$p_s=\left(\frac{-a_1}{a_{11}+\sqrt{a_{11}^2-3a_{111}a_1}}\right)^{1/2}. \tag{18}$$

We use p_s to normalize the order parameters p_1 and p_2.

Introduce dimensionless parameters

$$a'_{11}=\frac{a_{11}p_s^2}{|a_1|},\quad a'_{12}=\frac{a_{12}p_s^2}{|a_1|},\quad a'_{111}=\frac{a_{111}p_s^4}{|a_1|},\quad a'_{112}=\frac{a_{112}p_s^4}{|a_1|}, \tag{19}$$

$$\gamma_1=\frac{g_{12}+g_{44}-g'_{44}}{g_{11}},\quad \gamma_2=\frac{g_{44}+g'_{44}}{g_{11}}. \tag{20}$$

The driving force F_α in (13) contains three terms. The comparison of the terms involving f_L and f_G defines a length scale:

$$b=\sqrt{\frac{g_{11}}{-2a_1}}. \tag{21}$$

This length represents the width of the domain wall in the Ginzburg-Landau framework. We use b to normalize the spatial coordinates. Figure 5 shows the shape of a 90-deg domain wall. In [15], a boundary value problem was solved to obtain the analytical result for the 90-deg domain wall. In our work, the simulation from evolution equations shows similar shape, and the domain wall width is comparable to b.

In (13) a comparison between the gradient energy term and the surface stress term defines another length scale:

$$l=\frac{4\pi\mu g_{11}}{\beta_{11}^2p_s^2}. \tag{22}$$

This length scales the size of an individual domain in an equilibrium pattern. As stated in the Introduction, the equilibrium domain pattern is an outcome of the competition between refining due to elasticity and coarsening due to domain wall energy.

The time scale can be derived from the evolution Eqs. (15), giving

$$\tau=-\frac{1}{2a_1L}. \tag{23}$$

We will report time in the unit of τ.

In terms of the dimensionless quantities introduced above, the evolution equations become

$$\begin{aligned}\frac{\partial p_1}{\partial t}&=p_{1,11}+\gamma_1p_{2,21}+\gamma_2p_{1,22}\\&\quad-\frac{b}{l}\left(-p_1\varepsilon_{11}+\frac{\beta_{12}}{\beta_{11}}p_1\varepsilon_{22}-\frac{2\beta_{44}}{\beta_{11}}p_2\varepsilon_{12}\right)+p_1-2a'_{11}p_1^3\\&\quad-a'_{12}p_1p_2^2-3a'_{111}p_1^5-a'_{112}(2p_1^3p_2^2+p_1p_2^4)\end{aligned} \tag{24}$$

$$\begin{aligned}\frac{\partial p_2}{\partial t}&=p_{2,22}+\gamma_1p_{1,12}+\gamma_2p_{2,11}\\&\quad-\frac{b}{l}\left(\frac{\beta_{12}}{\beta_{11}}p_2\varepsilon_{11}-p_2\varepsilon_{22}-\frac{2\beta_{44}}{\beta_{11}}p_1\varepsilon_{12}\right)+p_2-2a'_{11}p_2^3\\&\quad-a'_{12}p_2p_1^2-3a'_{111}p_2^5-a'_{112}(2p_2^3p_1^2+p_2p_1^4).\end{aligned}$$

In the above the strains have to be determined by solving the elastic field in the semi-infinite solid subject to the following boundary conditions on the surface:

$$\begin{aligned}\sigma_{31}&=-p_1\frac{\partial p_1}{\partial x_1}+\frac{\beta_{12}}{\beta_{11}}p_2\frac{\partial p_2}{\partial x_1}-\frac{\beta_{44}}{\beta_{11}}\left(p_2\frac{\partial p_1}{\partial x_2}+p_1\frac{\partial p_2}{\partial x_2}\right)\\\sigma_{32}&=-p_2\frac{\partial p_2}{\partial x_2}+\frac{\beta_{12}}{\beta_{11}}p_1\frac{\partial p_1}{\partial x_2}-\frac{\beta_{44}}{\beta_{11}}\left(p_2\frac{\partial p_1}{\partial x_1}+p_1\frac{\partial p_2}{\partial x_1}\right)\\\sigma_{33}&=0\end{aligned} \tag{25}$$

where the stress is normalized by $\beta_{11}p_s^2/b$. The elastic field in a half-space due to a tangential point force on the surface is known as the Cerruti field [20]. A superposition gives the needed strain field:

$$\begin{aligned}\varepsilon_{11}&=\int\int\sigma_{31}\left\{-\frac{2(1-3\nu)(x-\xi)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}}\right.\\&\quad\left.-\frac{6\nu(x-\xi)^3}{[(x-\xi)^2+(y-\eta)^2]^{5/2}}\right\}d\xi d\eta\\&\quad+\int\int\sigma_{32}\left\{\frac{2\nu(y-\eta)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}}\right.\\&\quad\left.-\frac{6\nu(y-\eta)(x-\xi)^2}{[(x-\xi)^2+(y-\eta)^2]^{5/2}}\right\}d\xi d\eta\end{aligned}$$

$$\varepsilon_{22}=\int\int \sigma_{31}\left\{\frac{2\nu(x-\xi)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}} - \frac{6\nu(x-\xi)(y-\eta)^2}{[(x-\xi)^2+(y-\eta)^2]^{5/2}}\right\} d\xi d\eta + \int\int \sigma_{32}\left\{-\frac{2(1-3\nu)(y-\eta)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}} - \frac{6\nu(y-\eta)^3}{[(x-\xi)^2+(y-\eta)^2]^{5/2}}\right\} d\xi d\eta$$

$$\varepsilon_{12}=\frac{1}{2}\int\int \sigma_{31}\left\{-\frac{2(1-2\nu)(y-\eta)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}} - \frac{12\nu(x-\xi)^2(y-\eta)}{[(x-\xi)^2+(y-\eta)^2]^{5/2}}\right\} d\xi d\eta + \frac{1}{2}\int\int \sigma_{32}\left\{-\frac{2(1-2\nu)(x-\xi)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}} - \frac{12\nu(x-\xi)(y-\eta)^2}{[(x-\xi)^2+(y-\eta)^2]^{5/2}}\right\} d\xi d\eta. \tag{26}$$

The strains are normalized by $\beta_{11}p_s^2/4\pi\mu b$. The displacements, normalized by $\beta_{11}p_s^2/4\pi\mu$, are

$$u_1=\int\int \sigma_{31}\left\{\frac{2(1-\nu)}{[(x-\xi)^2+(y-\eta)^2]^{1/2}} + \frac{2\nu(x-\xi)^2}{[(x-\xi)^2+(y-\eta)^2]^{3/2}}\right\} d\xi d\eta + \int\int \sigma_{32}\left\{\frac{2\nu(x-\xi)(y-\eta)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}}\right\} d\xi d\eta$$

$$u_2=\int\int \sigma_{31}\left\{\frac{2\nu(x-\xi)(y-\eta)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}}\right\} d\xi d\eta + \int\int \sigma_{32}\left\{\frac{2(1-\nu)}{[(x-\xi)^2+(y-\eta)^2]^{1/2}} + \frac{2\nu(y-\eta)^2}{[(x-\xi)^2+(y-\eta)^2]^{3/2}}\right\} d\xi d\eta. \tag{27}$$

4 Computer Simulation

To illustrate the model, we now present several preliminary computer simulations. Referring to [16,17], we adopt the following parameters:

$$a'_{11}=0.5,\quad a'_{12}=1.74,\quad a'_{111}=0.063,\quad a'_{112}=0.148$$
$$g_{12}/g_{11}=1.1,\quad g_{44}/g_{11}=0.9,\quad g'_{44}/g_{11}=0.1$$
$$b/l=1.0,\quad \beta_{12}/\beta_{11}=0.05,\quad \beta_{44}/\beta_{11}=0.111,\quad \nu=0.3. \tag{28}$$

We have not carried out a comprehensive parametric study.

The evolution equations have both spatial and temporal derivatives. We discretize the infinite surface into an array of squares of size $N\times N$. All fields are assumed to be periodically replicated from one square to another. Each square is the computation unit cell, which is subdivided into grids of spacing Δx. Spatial derivatives are approximated with finite difference. At every time step, the elastic field is calculated by evaluating the double integrals, and the order parameter field is updated by using the Euler method. In our simulation, we use $N=128$, $\Delta x=1.0$, and adaptive time steps. The ratio of two length scales is $b/l=1.0$. The singular integrals in (26) and (27) are evaluated by adopting a technique in the Appendix. Since the double integrals must be calculated for every grid point, the simulation is rather slow. The evolution can also be simulated in reciprocal space ([12–17]), which we will implement for this model in the future.

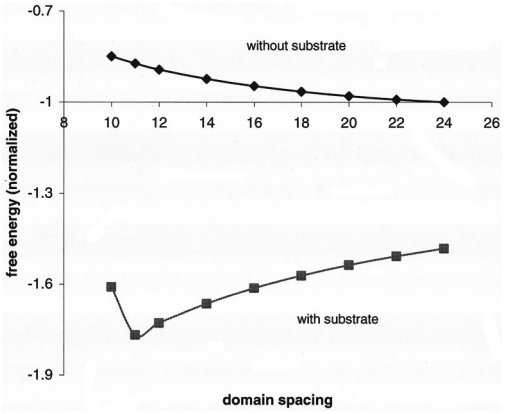


Fig. 6 Free energy versus domain spacing. Without substrate constraint, the energy decreases as the domain size increases. With substrate constraint, the energy reaches a minimum at a specific domain size.

For a given domain pattern, i.e., a prescribed field of the order parameters, the free energy can be computed from Eqs. (16), (25), and (27). In the first simulation, we specify a parallel domain pattern, fix the domain spacing, but allow the domain walls to relax locally. We calculate the free energy as a function of the domain spacing. Figure 6 illustrates the effect of substrate constraint. The domain size in this figure is the width of the domain along (x,y) coordinates of Fig. 5. As discussed in the Introduction, without the substrate, the total free energy of the epilayer decreases with the increase of the domain size, and the single domain has the lowest energy. With the substrate constraint, since the competition of elasticity and domain wall energy selects an equilibrium domain size, a valley exists on the curve of energy versus domain spacing. For the domains smaller than the equilibrium size, the relaxation of the elastic energy cannot accommodate the rapid increase of the domain wall length, so the more domains, the higher the free energy.

Figure 7 compares the domain evolution with substrate and without substrate constraint. Through every grid point we draw a short line segment, representing the magnitude and the direction of the order parameter. The left column shows the result with the substrate, and the right one without substrate. In both cases, we start with the initial condition of a small random perturbation of the order parameter field from zero, corresponding to the paraelastic state. When the layer is constrained, parallel domains form, and the domain size approaches to what has been shown in Fig. 5. When the layer is unconstrained, the domains coarsen, being limited only by the calculating cell size.

5 Summary

This paper presents a formalism to simulate the evolution of domain patterns in ferroelastic epilayers. The surface stress due to the spontaneous strains can be relaxed by the substrate mediated elastic interaction. Free energy is divided into two parts. One is the elastic energy of the semi-infinite substrate, and the other is the excess surface energy, which has the ingredients of the Landau expansion, the gradient energy, and the surface stress. Evolution equations are derived from energy variation. Computer simulation ascertains that the competition of coarsening and refining is responsible for equilibrium domain patterns. The formalism presented in this paper is flexible, and can be used to study other phenomena involving surface phase transition and pattern formation.

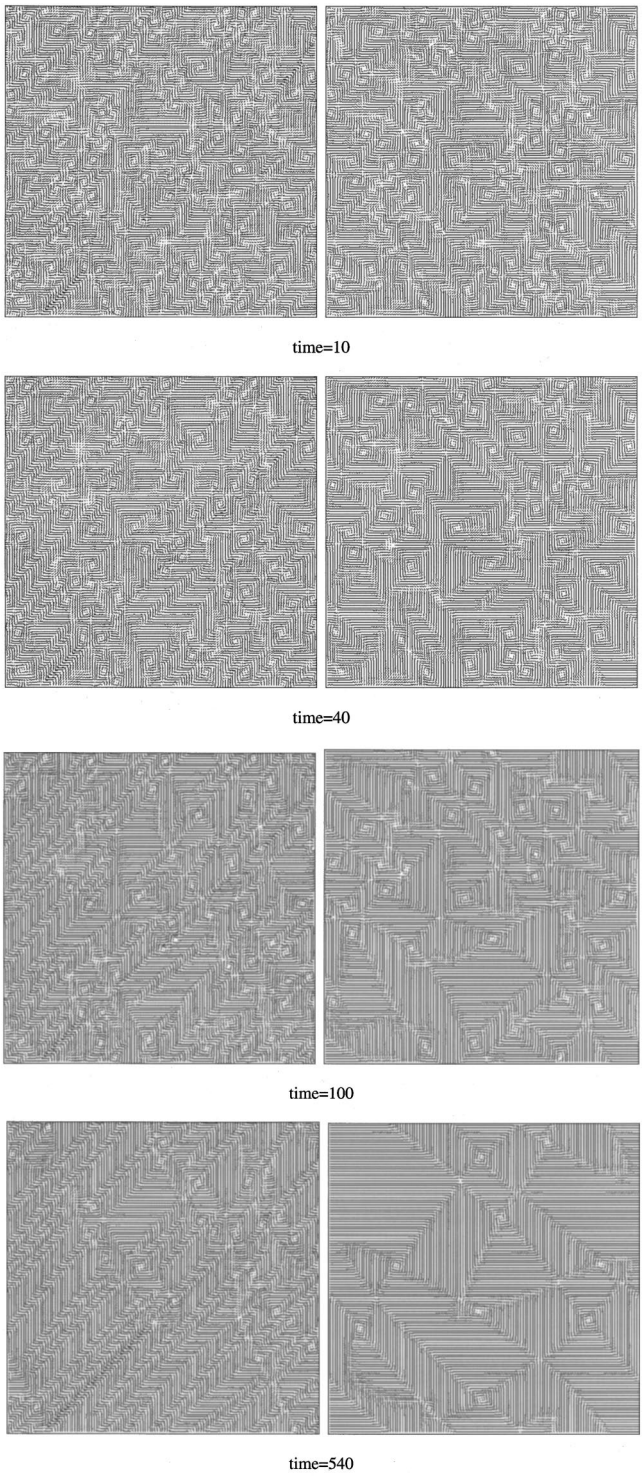


Fig. 7 Comparison of domain evolution in simulations with substrate constraint and without. The left column is the result with substrate, and parallel domains can be seen. The right column is the result without constraint, and the structure keeps coarsening.

Acknowledgments

This work is supported by the National Science Foundation through grant CMS-9820713, and by Princeton University through the Sir Gordon Wu Fellowship for Y. F. G. We would like to thank Dr. Wei Lu for helpful discussions.

Appendix

The integrals (26) and (27) extend over the entire surface. To save computation time, we only extend the integrals to a finite square of size $16b \times 16b$ in our simulation. The integrals are singular when $x=\xi$ and $y=\eta$. Let ε be a small number, say the grid spacing Δx. We treat the singularity as follows:

$$\int_{x-\varepsilon}^{x+\varepsilon}\int_{y-\varepsilon}^{y+\varepsilon}\frac{\sigma_{31}(\xi,\eta)(x-\xi)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}}\,d\xi\,d\eta$$

$$=\int_{x-\varepsilon}^{x+\varepsilon}\int_{y-\varepsilon}^{y+\varepsilon}\frac{[\sigma_{31}(\xi,\eta)-\sigma_{31}(x,y)](x-\xi)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}}\,d\xi\,d\eta$$

$$\approx\int_{x-\varepsilon}^{x+\varepsilon}\int_{y-\varepsilon}^{y+\varepsilon}\frac{\left[\frac{\partial\sigma_{31}}{\partial x}(\xi-x)+\frac{\partial\sigma_{31}}{\partial y}(\eta-y)\right](x-\xi)}{[(x-\xi)^2+(y-\eta)^2]^{3/2}}\,d\xi\,d\eta$$

$$=-\frac{\partial\sigma_{31}}{\partial x}\int_{x-\varepsilon}^{x+\varepsilon}\int_{y-\varepsilon}^{y+\varepsilon}\frac{(x-\xi)^2}{[(x-\xi)^2+(y-\eta)^2]^{3/2}}\,d\xi\,d\eta$$

$$=-\frac{\partial\sigma_{31}}{\partial x}\cdot\varepsilon\cdot\int_{-1}^{1}\int_{-1}^{1}\frac{s^2}{[s^2+t^2]^{3/2}}\,ds\,dt$$

$$=4\ln(\sqrt{2}-1)\,\varepsilon\,\frac{\partial\sigma_{31}}{\partial x}.$$

References

[1] Lines, M. E., and Glass, A. M., 1977, *Principles and Applications of Ferroelectrics and Related Materials*, Clarendon Press, Oxford.
[2] Pertsev, N. A., Zembilgotov, A. G., and Tagantsev, A. K., 1998, "Effect of Mechanical Boundary Conditions on Phase Diagrams of Epitaxial Ferroelectric Thin Films," Phys. Rev. Lett., **80**, pp. 1988–1991.
[3] Suo, Z., 1998, "Stress and Strain in Ferroelectrics," Curr. Opin. Solid State Mater. Sci., **3**, pp. 486–489.
[4] Pompe, W., Gong, X., Suo, Z., and Speck, J. S., 1993, "Elastic Energy Release due to Domain Formation in the Strained Epitaxy of Ferroelectric and Ferroelastic Films," J. Appl. Phys., **74**, pp. 6012–6019.
[5] Speck, J. S., and Pompe, W., 1994, "Domain Configurations due to Multiple Misfit Relaxation Mechanisms in Epitaxial Ferroelastic Thin Films: I Theory," J. Appl. Phys., **76**, pp. 466–476.
[6] Kwak, B. S., and Erbil, A., 1992, "Strain Relaxation by Domain Formation in Epitaxial Ferroelectric Thin Films," Phys. Rev. Lett., **68**, pp. 3733–3736.
[7] Sridhar, N., Rickman, J. M., and Srolovitz, D. J., 1996, "Twinning in Thin Films—I. Elastic Analysis," Acta Mater., **44**, pp. 4085–4096.
[8] Sridhar, N., Rickman, J. M., and Srolovitz, D. J., 1996, "Twinning in Thin Films—II, Equilibrium Microstructures," Acta Mater., **44**, pp. 4097–4113.
[9] Seul, M., and Andelman, D., 1995, "Domain Shapes and Patterns—The Phenomenology of Modulated Phases," Science, **267**, pp. 476–483.
[10] Ibach, H., 1997, "The Role of Surface Stress in Reconstruction, Epitaxial Growth and Stabilization of Mesoscopic Structures," Surf. Sci. Rep., **29**, pp. 193–263.
[11] Alerhand, O. L., Vanderbilt, D., Meade, R. D., and Joannopoulos, J. D., 1988, "Spontaneous Formation of Stress Domains on Crystal Surfaces," Phys. Rev. Lett., **61**, pp. 1973–1976.
[12] Lu, W., and Suo, Z., 1999, "Coarsening, Refining, and Pattern Emergence in Binary Epilayers," Z. Metallkd., **90**, pp. 956–960.
[13] Lu, W., and Suo, Z., 2001, "Dynamics of Nanoscale Pattern Formation of an Epitaxial Monolayer," J. Mech. Phys. Solids, **49**, pp. 1937–1950.
[14] Suo, Z., and Lu, W., 2000, "Composition Modulation and Nanophase Separation in a Binary Epilayer," J. Mech. Phys. Solids, **48**, pp. 211–232.
[15] Cao, W., and Cross, L. E., 1991, "Theory of Tetragonal Twin Structures in Ferroelectric Perovskites With a First-Order Phase Transition," Phys. Rev. B, **44**, pp. 5–12.
[16] Hu, H.-L., and Chen, L.-Q., 1998, "Three-Dimensional Computer Simulation of Ferroelectric Domain Formation," J. Am. Ceram. Soc., **81**, pp. 492–500.
[17] Nambu, S., and Sagala, D. A., 1994, "Domain Formation and Elastic Long-Range Interaction in Ferroelectric Perovkites," Phys. Rev. B, **50**, pp. 5838–5847.
[18] Cao, W. W., and Randall, C. A., 1996, "Grain Size and Domain Size Relations in Bulk Ceramic Ferroelectric Materials," J. Phys. Chem. Solids, **57**, pp. 1499–1505.
[19] Randall, C. A., Kim, N., Kucera, J. P., Cao, W. W., and Shrout, T. R., 1998, "Intrinsic and Extrinsic Size Effects in Fine-Grained Morphotoropic-Phase Boundary Lead Zirconate Titanate Ceramics," J. Am. Ceram. Soc., **81**, pp. 677–688.
[20] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK, p. 69.

S. P. A. Gill
Department of Engineering,
Leicester University,
University Road,
Leicester, LE1 7RH, UK

H. Gao
Mem. ASME
Division of Mechanics and Computation,
Department of Mechanical Engineering,
Stanford University,
Stanford, CA 94305

V. Ramaswamy

W. D. Nix

Department of Materials Science
and Engineering,
Stanford University,
Stanford, CA 94305

Confined Capillary Stresses During the Initial Growth of Thin Films on Amorphous Substrates

Changes in substrate curvature indicating the existence of compressive stress in isolated crystallites are commonly observed during the initial stages of thin film deposition of metals on glass or ceramic substrates. Following the suggestion of Abermann et al. (R. Abermann et al., 1978, Thin Solid Films, ***52****, p. 215), we attribute the origin of this compressive stress to the action of capillary forces during film growth. As new atomic layers are deposited, the capillary forces acting on atoms near the surface are stored as transformation strains in the bulk of the crystallites. To test this concept, we propose three models for evaluating the capillary strains and their induced compressive stresses in a crystalline. A finite element analysis is performed to show that the model predictions agree well with experimental data.* [DOI: 10.1115/1.1469001]

1 Introduction

Controlling stresses in deposited thin films has become an important challenge of microelectronic technology. It is well known that there are relationships between the film stress and the corresponding microstructure at various stages of growth. Better understanding of how film microstructures and stress develop and evolve may lead to better control of final stress states and microstructures in deposited thin films.

Film stress, obtained from substrate curvature, has been measured during growth of sputter-deposited Pt on amorphous substrates of SiO_2 ([1]) and is observed to be compressive at thicknesses less than about 5 Å. This can be seen in Fig. 1 where a negative curvature is evident due to the compressive stress. The stress changes to being tensile in thicker films, leading to a tensile maximum at about 33 Å, after which the stress becomes compressive again. The development of tensile stresses in the intermediate stage is thought to be due to crystallite coalescence ([2]) and those in the later compressive stage due to ''peening.'' Such stress behavior is commonly observed during the deposition of high mobility metal films by evaporation ([1,3,4,5]).

In this paper we focus on the origin of the compressive stress during the initial stage of film growth. This compressive stress is observed when the film is not continuous but consists of a number of isolated crystallites (or islands). Following the suggestion of Abermann et al. [3], we attribute this stress to the action of capillary forces during growth. We envision that, as new layers of atoms are deposited, the capillary forces acting on atoms near the free surface become stored as transformation strains in the bulk of the crystallites. Three models for evaluating the capillary strains and their induced compressive stresses in a crystallite are proposed in Section 2. The capillary strains stored in an island are related to the induced curvature measured in experiments in Section 3. The models are evaluated numerically and compared with experimental data in Section 4.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, March 15, 2001; final revision, December 10, 2001. Associate Editor: D. A. Kouris. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS

2 Models

The origin of the compressive stress in pre-coalescence crystallites is thought to be due to surface stress. Here we adopt the qualitative picture of these compressive stresses first given by Abermann et al. [3] and develop quantitative models to describe the curvature changes in this stage of film growth. The physical mechanism idealized here involves contributions to this compressive stress from two sources: (1) residual stresses built into the crystallites during growth by capillary forces and (2) stresses associated with the capillary forces exerted by the current configuration of the surface.

As discussed by Abermann et al. [3], during deposition an island is under compression due to surface stress. While the island is small the surface-induced stresses significantly deform the atomic lattice of the island. This deformation is considered to be locked into the system during growth as bonding at the island-substrate interface occurs while the island is in this compressive state. Assuming that no interfacial slipping or plastic relaxation occurs, the compression in the island due to its surface stress can be represented by a transformation stress. This stress depends on the growth history of the island and also on how one conceptualizes the growth process. As there is little fundamental information on these aspects of the problem, three models for the equivalent transformation stress will be proposed for comparison. Analysis of TEM micrographs of pre-coalescence of films indicates that the island heights are approximately half of the average base diameter ([1]). Hence we assume that the islands are hemispheres of radius R_0 as shown in Fig. 2(*a*). The x-y plane coincides with the substrate surface, the origin O is at the center of curvature of the island, and the vertical z-axis is an axis of rotational symmetry which acts out of the substrate.

For a solid surface, it has been pointed out ([6]) that one needs to distinguish between surface energy γ and surface stress f. While the surface energy defines the reversible work to form a new surface by a process such as cleavage, surface stress is the reversible work per unit area required to elastically deform a surface so that it is commensurate with the bulk. These two quantities differ

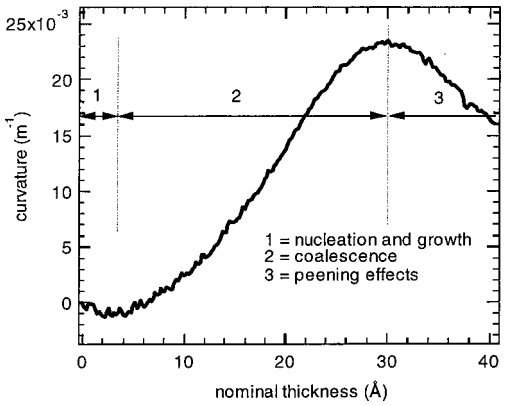


Fig. 1 *In-situ* curvature versus nominal thickness for a Pt film deposited on a SiO_2 substrate

in fundamental concept but are usually of the same order of magnitude. We will use surface stress to calculate the capillary forces.

2.1 Spherical Layer Transformation Models

2.1.1 Spherical Layer Model (Model I). Imagine that an island grows as a number of hemispherical monolayers. These layers are always complete and hence we can simulate the growth process by simply increasing the number of layers. Each layer forms under the action of the current surface capillary pressure due to surface stress

$$P(R)=\frac{2f}{R} \tag{2.1}$$

where f is the surface stress of the island and R is the radius of the spherical layer. Consequently the state of compression in the island will not be uniform as it grows but will decrease as the layer radius increases. Consideration of an infinitesimal element in a thin spherical shell under radial loading $P(R)$ yields a stress distribution of ([7])

$$\sigma_r=-P \quad \sigma_\Theta=\sigma_\Phi=-P-\frac{R}{2}\frac{dP}{dR} \tag{2.2}$$

where the (R,Θ,Φ) spherical coordinate system has been used and all shear transformation stresses are zero. The spherical radial distance R is the distance from the origin O, Θ is the horizontal angle between the x-axis and the projection of the radius into the x-y plane, and Φ is the vertical angle between the radius and the x-y plane. If the pressure is constant then a hydrostatic pressure distribution results as expected. As new atomic layers are deposited, atoms near the surface become embedded in the crystallite and attempt to change to their normal bulk state. This shift of reference for the stress-free state is represented by transformation stresses

$$\sigma_R^T=\frac{2f}{R} \quad \sigma_\Theta^T=\sigma_\Phi^T=\frac{f}{R}. \tag{2.3}$$

2.1.2 Cylindrical Layer Model (Model II). In this model the adatom attachment process during the growth of a new spherical monolayer is assumed to initiate at the perimeter of the island's base. In this case, it is imagined that a single hoop of atoms initially forms on the substrate around the base of the island and then the layer is completed by further deposition on top of these peripheral atoms. It is in the initial hoop formation stage that the bonding of the layer to the substrate occurs. In this situation the initial hoop of atoms can be considered to be subject to a surface stress in the plane of the island-substrate interface such that the magnitude of the capillary pressure experienced by the hoop is

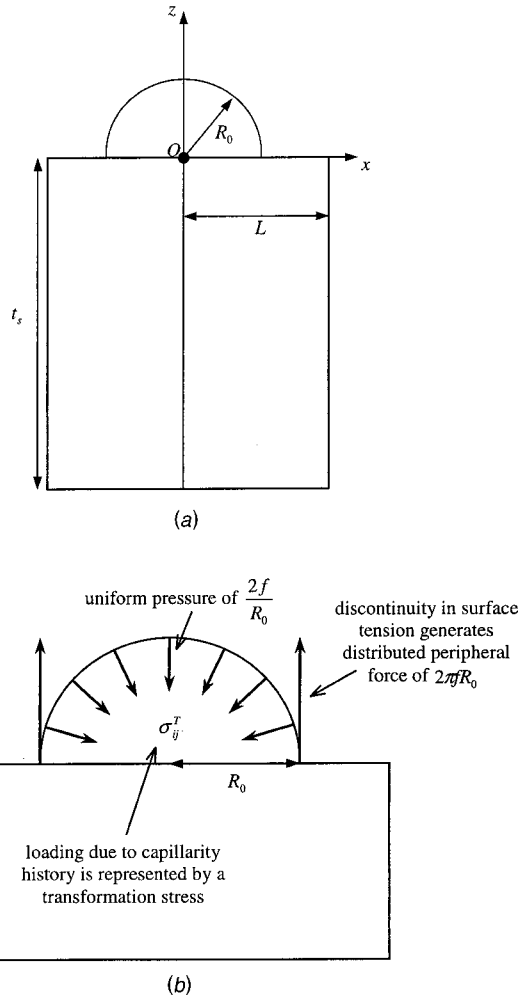


Fig. 2 (*a*) The problem geometry, (*b*) the loading on the island and substrate due to current capillarity and capillarity history

$$P(r)=\frac{f}{r} \tag{2.4}$$

where r is the radius of the hoop. As before, consideration of the stress state in a thin plane stress hoop element subject to radial loading ([7]) yields

$$\sigma_r=-P \quad \sigma_\theta=-P-r\frac{dP}{dr} \quad \sigma_z=0 \tag{2.5}$$

where the (r,θ,z) cylindrical coordinate system has been used and all shear transformation stresses are zero. The radius r is the horizontal distance from the origin and θ is the horizontal angle between the radius and the x-axis. As new atomic layers are deposited on a crystallite, the stress-free state is shifted from this compressed state to the normal bulk state. This shift is represented by transformation stresses

$$\sigma_r^T=\frac{f}{r} \quad \sigma_\theta^T=\sigma_z^T=0. \tag{2.6}$$

2.1.3 Current Capillary Forces. The above concentric layer models account for the residual stresses locked into a crystallite during growth due to its capillary history. This does not take into account the effect of the surface stress on the island in its current configuration. To do this it is necessary to apply a uniform pressure of $2f/R_0$ over the surface of a hemispherical island of radius R_0. The surface stress must also be balanced at the periphery of

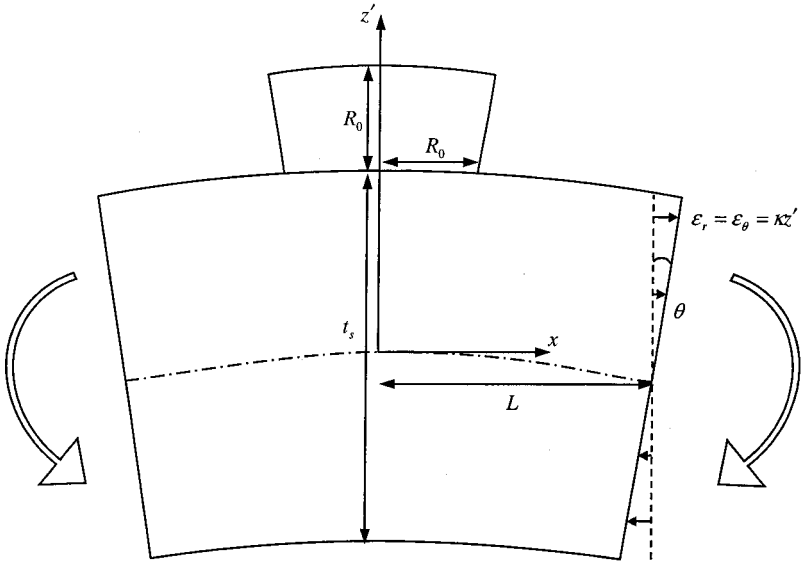


Fig. 3 A simplified model of the problem as a beam subject to a pure bending moment

the island where the substrate-island, island-vapor, and substrate-vapor interfaces meet. We assume that the substrate-vapor surface energy and the substrate-island energy are identical. This means that the surface tensions are automatically balanced in the plane of the substrate surface and that the island will adopt a (near) hemispherical shape. The island surface stress acts vertically upward from this point and must be included in the model. The resultant force will be balanced by deformation of the substrate.

The combined loading on the island due to the capillary history (σ_{ij}^T) and the current capillary is shown in Fig. 2(*b*).

2.2 Local Transformation Model (Model III). It is natural to assume that the island atoms in contact with the substrate are all in the same atomic plane and that there is no plastic deformation of the atomic structure of the island during growth, i.e., it is a perfect crystal. These two requirements imply that an island allowed to relax its residual stress state by removing it from the substrate should experience no other stresses besides the normal capillary pressure. This is not the case for the concentric layer transformation models described above, as the nonuniform nature of transformation stresses induces Poisson effects which distort the relaxed island. An implicit assumption in these models is that the current capillary pressure is always $2f/R$ or f/r, which is strictly true only for a free-standing spherical or cylindrical droplet free from the substrate. Furthermore, these models have the geometric restrictions to hemispherical or cylindrical crystallites.

In order to relieve the above limitations, we propose a local transformation model in the following. The origin of the capillary pressure is thought to be due to the difference between the natural lattice parameter of the surface atoms and that of the bulk atoms. We imagine that a crystallite would be stress-free if all the atoms were in the surface state. Consider a sphere of radius R representing the bulk material enveloped by a concentric spherical shell of thickness b representing the surface layer. Apply a uniform thermal expansion of ε^T to the sphere to transform its lattice spacing from the surface value to the bulk value. Lamé's equations for spherical symmetry ([7]) can be readily used to find the elasticity solution to this problem. It is found that a constant hydrostatic stress

$$\sigma_r^b=\sigma_\Theta^b=\sigma_\Phi^b=\frac{2E\varepsilon^T}{3(1-\nu)}\left[\frac{R^3}{(R+b)^3}-1\right]$$

is generated in the bulk (denoted by superscript *b*), and a variable hoop stress

$$\sigma_\Theta^s=\sigma_\Phi^s=\frac{2E\varepsilon^T}{3(1-\nu)}\left[\frac{R^3}{(R+b)^3}+\frac{R^3}{2r^3}\right]$$

is generated in the surface layer (denoted by the superscript *s*), where $R<r<R+b$. The surface hoop stress is equivalent to the surface stress induced by the surface tension, i.e., $\sigma_\Theta^s=\sigma_\Phi^s=f/b$. Assuming the surface layer thickness is much smaller than the bulk dimension ($b\ll R$), this is the case if $\varepsilon^T=(1-\nu)f/Eb$. Therefore the corresponding hydrostatic transformation stress required to transform from a surface atom to a bulk atom is

$$\sigma_r^T=\sigma_\Theta^T=\sigma_\Phi^T=\frac{E\varepsilon^T}{(1-2\nu)}=\frac{(1-\nu)}{(1-2\nu)}\frac{f}{b} \tag{2.7}$$

where b is taken to be the effective thickness of the surface monolayer. Imposing a uniform expansion in the bulk of a free droplet excluding the surface monolayer immediately generates the expected capillary pressure $\sigma_r^b=\sigma_\Theta^b=\sigma_\Phi^b=-2f/R$ within the droplet.

For a thin film deposited on a substrate, we assume that all atoms in a crystallite experience the same surface-to-bulk transformation given by (2.7) during the deposition process. We attribute the origin of the compressive stress in the crystallite to this transformation. Since adatoms are attached while they are in the surface state, the crystallite would be stress-free if all the atoms were in the surface state. The surface-to-bulk transformation causes a compressive stress to be developed within the crystallite. This model naturally returns to the state of free crystallites when the substrate is removed and is not limited to hemispherical or cylindrical geometries. There is no need to assume that the current capillary pressure remains in the form of $2f/R$ as the crystallite grows. In fact, due to the substrate constraint, the current capillary pressure is expected to be smaller than $2f/R$ because the crystallite will not be able to expand the same amount as a free droplet.

In the case of a free droplet, we could alternatively start with the statement that the droplet would be stress-free if all the atoms were in the bulk state. Imposing a uniform contraction given by (2.7) in the surface monolayer also generates the expected capillary pressure $2f/R$. However, this *bulk reference state* cannot be

directly applied to the film deposition process because adatoms are attached to the crystallite and the amorphous substrate while they are in the surface state.

3 Induced Curvature Model for an Isolated Island

Consider the simplified geometry shown in Fig. 3 where a hemispherical island of height and radius R_0 sits on top of the center of a cylindrical section of substrate of radius L and depth t_s. We assume that the relaxation strain in the island and substrate is of the form of that found in a plate subject to a uniform bending moment such that all the strain components are zero except

$$\varepsilon_r = \varepsilon_\theta = -\kappa z' \tag{3.1}$$

where z' coincides with the z-axis and is the vertical distance from the neutral plane of the island-substrate assembly and the curvature, κ, is the two-dimensional curvature in the r-z plane. This is taken to be positive if there is a tensile stress in the island, i.e., the island contracts and the assembly is concave. The island is assumed to be thin compared to the substrate so that the neutral plane can be taken to coincide with the neutral plane of the substrate in isolation, i.e., $z' = z - \frac{1}{2}t_s$. In reality the strain field in and around the island is quite complex but we can test the validity of this assumption later. It should apply near the base of the island where its free surfaces are sufficiently constrained by the substrate. This is the most important region for transmission of stress into the substrate. Near the top of the island the stresses will relax much more readily. This effect will be taken into account by the adoption of a numerically calculated geometrical parameter.

We use an energy method to determine the form of the relationship between loading, island size, and induced curvature. The total elastic energy in the system, U, has contributions from both the island (U_i) and the substrate (U_s) such that

$$U = U_s + U_i = \frac{1}{2}\int_{V_s+V_i} (\sigma_{ij} - \sigma_{ij}^T)(\varepsilon_{ij} - \varepsilon_{ij}^T) dV$$

where V_s and V_i are the substrate and island volumes, σ_{ij}^T and ε_{ij}^T are the transformation stress and strain components, and σ_{ij} and ε_{ij} are the stress and strain components taken with reference to this transformed state.

The transformation stresses are applied to the island only so the elastic stored energy in the substrate is simply

$$U_s = \frac{1}{2}\int_{V_s} \sigma_{ij}\varepsilon_{ij} dV = \frac{1}{2\overline{E}_s}\int_{-\frac{1}{2}t_s}^{\frac{1}{2}t_s}\int_0^L 2(\overline{E}_s \kappa z')^2 .2\pi r dr dz'$$

$$= \overline{E}_s \frac{\pi L^2 t_s^3}{12}\kappa^2 \tag{3.2}$$

where $\overline{E}_s = E_s/(1-\nu_s)$ is the biaxial modulus determined from the Young's modulus, E_s, and the Poisson ratio, ν_s of the substrate.

To illustrate the calculation of the elastic energy in the island we consider the transformation stress state of (2.2) proposed for model II. For the assumed cylindrical geometry this gives

$$U_i = \frac{1}{2\overline{E}_i}\int_{t_s/2}^{t_s/2+R_0}\int_0^{R_0}\left[\left(-\overline{E}_i\kappa z' - \frac{f}{r}\right)^2 + (-\overline{E}_i\kappa z')^2\right] 2\pi r dr dz'$$

$$= \pi\frac{\overline{E}_i t_s^2 R_0^3}{4}\kappa^2 + \pi f t_s R_0^2 \kappa + \text{constant}$$

where the biaxial modulus of the island, $\overline{E}_i = E_i/(1-\nu_i)$, is determined from the Young's modulus, E_i, and the Poisson ratio, ν_i, of the island material.

The total elastic stored energy

Table 1 The α parameter is equivalent to the curvature produced by a dense packing of pre-coalescence Pt islands with unit surface energy density on a SiO_2 substrate of unit thickness. It is given here for 13 different types of capillary loading. Where one transformation stress is shown, the other transformation stresses are zero. These contributions can be easily used to calculate the curvature due to combined loads as they are simply additive.

Parameter	Loading Type	Value
α_A	Current Capillarity Pressure	0.229
α_B	$\sigma_r = -R_0/r$	0.115
α_C	$\sigma_\theta = -R_0/r$	0.039
α_D	$\sigma_z = -R_0/r$	−0.058
α_E	$\sigma_R = -R_0/R$	0.014
α_F	$\sigma_\Theta = -R_0/R$	0.011
α_G	$\sigma_\Phi = -R_0/R$	0.020
α_H	$\sigma_r = -1$	0.056
α_I	$\sigma_\theta = -1$	0.016
α_J	$\sigma_z = -1$	−0.044
α_K	$\sigma_R = -1$	0.009
α_L	$\sigma_\Theta = -1$	0.016
α_M	$\sigma_\Phi = -1$	0.003

$$U = \pi\left(\frac{\overline{E}_s t_s^3 L^2}{12} + \frac{\overline{E}_i t_s^2 R_0^3}{4}\right)\kappa^2 + \pi\gamma t_s R_0^2 \kappa + \text{constant}$$

adopts the minimum energy configuration such that $\partial U/\partial\kappa = 0$ so

$$\kappa = -\frac{6\gamma}{\overline{E}_s t_s^2}\left(\frac{R_0}{L}\right)^2$$

assuming $R_0 \ll t_s$.[1] Given the unconservative assumption about the form of the strain field in this analysis it is proposed that the actual curvature for a hemispherical island subject to this type of transformation stress is

$$\kappa = -\alpha A_A \kappa_0 \tag{3.3}$$

where

$$\kappa_0 = \frac{6\gamma}{\overline{E}_s t_s^2},$$

α is a numerically determined geometrical parameter which depends on the type of loading, and $A_A = R_0^2/L^2$ is the areal fraction for this geometry. A distribution of islands can be considered to be a collection of these subproblems. Various α-parameters have been calculated (see Section 4.2) and are given for different types of loading in Table 1. For the case considered here $\alpha = \alpha_B$. By superposition of linear elastic solutions, the combined effect of more than one of these loadings is simply additive. For instance, from (2.1) we find for model I that $\alpha = 2\alpha_E + \alpha_F + \alpha_G$. The validity of Eq. (3.3) is tested in Fig. 4 for a loading of type B for Pt islands on SiO_2. The parameter α_B is found to be approximately invariant of the island size except for an edge-effect as the island approaches the edge of its catchment area. A weighted mean value of α over the growth life of the island is taken. Two examples of the other types of loading are also shown in Fig. 4. The current capillary pressure (loading type A) has a more pronounced edge-effect but in general α_A and α_H both remain approximately constant. Again, weighted mean average values for Pt on SiO_2 are quoted in Table 1.

Model III is fundamentally different from models I and II. In this case the curvature scales with the volume of the island rather than the surface area. In addition, the problem is no longer scale invariant as the ratio of the thickness of the surface layer to the

[1] It is useful to note that this method derives the Stoney formula ([3]) of Eq. (4.1) for a continuous film ($R_0 = L$) of thickness t_0 subject to a uniform tensile transformation stress σ_0.

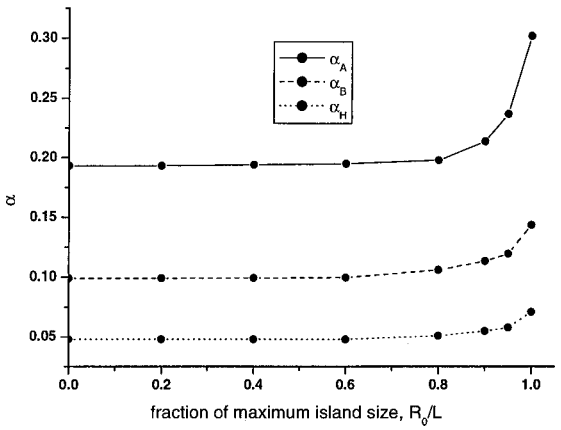


Fig. 4 Variation in α-parameters with the relative island size, R_0/L. The parameter remains approximately constant for small island sizes but there is a pronounced edge-effect as the island perimeter approaches the edge of the island catchment area. This effect is most obvious in the case of the current capillarity loading (type A). The effect with other loading types is similar to that shown for loading types B and H. The constant values quoted for α in Table 1 are weighted averages over the growth period of the island, $\int_0^1 \alpha(\eta)\eta^3 d\eta$, where $\eta = R_0/L$ and a constant volumetric growth rate is assumed.

island radius, b/R_0, depends on the chosen length scale. Therefore the induced curvature for model III of (2.7) can be expressed in the general form

$$\kappa_{III} = -\alpha_T(b/R_0)\frac{(1-\nu)}{(1-2\nu)}\frac{R_0}{b}A_A\kappa_0$$

where α_T is a dimensionless geometric parameter which is a function of b/R_0. Finite element calculations for this function are shown in Fig. 5. When the island is small ($R_0 < 5b$) it is less effective at bending the substrate as the transformed "bulk" volume is small compared to the initially stress free surface layer. As the island gets very large ($b/R_0 \to 0$) the effect of the surface layer becomes negligible and $\alpha_T \to \alpha_K + \alpha_L + \alpha_M$, the sum of the α-parameters representing a volumetric expansion without a surface layer. In general, the stress induced in the surface layer by the volumetric expansion is much less than the free surface stress f/b. This is because the expansion is heavily constrained by the underlying substrate. In practice the curvature is not sensitive to the behavior of α_T for small islands as the areal fraction is very small

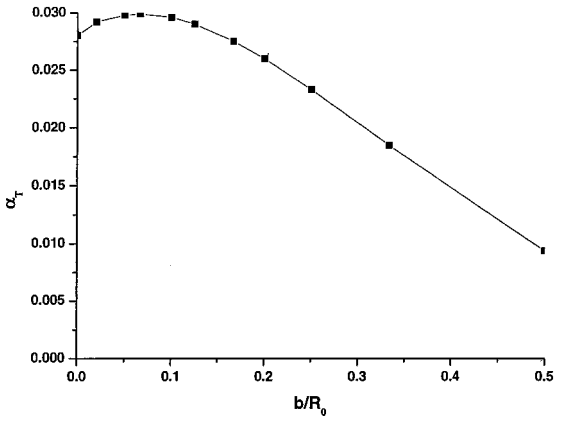


Fig. 5 The geometrical parameter α_T varies with island size R_0 given a constant surface layer thickness b

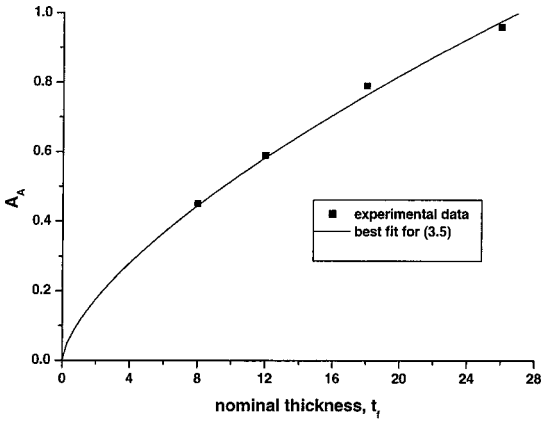


Fig. 6 A comparison of the areal fraction model of (3.5) with experimental data (1). The best correlation is obtained with a grain radius of $L \approx 40$ Å.

at this stage. Hence it is proposed that a reasonable model for this parameter is the large island asymptotic value of $\alpha_T = \alpha_K + \alpha_L + \alpha_M$.

3.1 Summary of Three Different Models. Based on the relationship (3.3) derived in the previous section, the three stress state models of Section 2 result in the following expressions for the induced substrate curvature:

$$\kappa_I = -(\alpha_A + 2\alpha_E + \alpha_F + \alpha_G)A_A\kappa_0$$

$$\kappa_{II} = -(\alpha_A + \alpha_B)A_A\kappa_0$$

$$\kappa_{III} = -(\alpha_K + \alpha_L + \alpha_M)\frac{(1-\nu)}{(1-2\nu)}\frac{R_0}{b}A_A\kappa_0$$

$$= -(\alpha_K + \alpha_L + \alpha_M)\frac{(1-\nu)}{(1-2\nu)}\frac{L}{b}A_A^{3/2}\kappa_0 \qquad (3.4)$$

where it is interesting to note that model III differs fundamentally from models I and II in the sense that the former predicts that the induced curvature depends on the island size whereas the latter predicts that it does not.

It is clear from the above that, apart from basic material and geometrical parameters, the only growth state variable required to know the current stress state is the areal fraction, A_A. Given the geometry in Fig. 2(a) the areal fraction is simply $A_A = \pi R_0^2/\pi L^2 = (R_0/L)^2$. Conservation of mass demands that the volume of the islands is equal to the volume of material deposited such that

$$3/2\pi R_0^3 = \pi L^2 \overline{t_f}$$

which implies that

$$A_A = \left(\frac{3\overline{t_f}}{2L}\right)^{2/3}. \qquad (3.5)$$

This is a very simple model which assumes that islands all nucleate a distance $2L$ apart at the start of deposition and that no further nucleation takes place during growth. However, the comparison with experimental measurements in Fig. 6 shows that it provides, coincidentally or otherwise, a very good description of the geometrical evolution of this system and therefore this expression will be utilized in calculating (3.4) in this paper.

4 Numerical Implementation

4.1 Finite Element Model. Due to the nonuniformity of the expected deformation field, the model is implemented numerically using the finite element method. The problem is assumed to have the geometry shown in Fig. 2(a). An individual island can there-

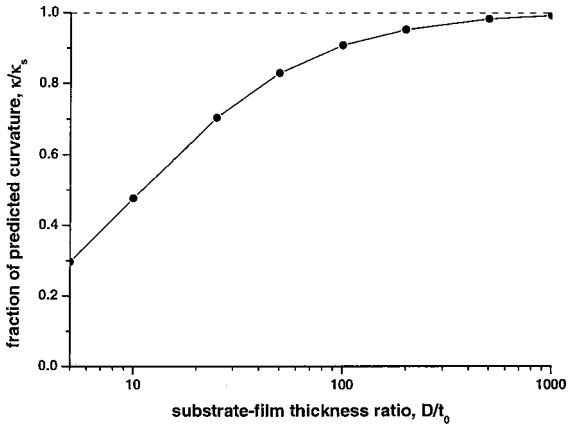


Fig. 7 Comparison between the substrate curvature obtained from the finite element analysis, κ, and that predicted by the Stoney formula, κ_s, with varying substrate thickness *D*. It can be seen that the two models converge for large substrate thicknesses above 100 times the film thickness t_0.

fore be analyzed in isolation from the others and the problem can be reduced to a two-dimensional axisymmetric one. The thickness of the substrate, t_s, is much greater than the radial span of the catchment area, L, by four to five orders of magnitude. Finite element analysis of a problem with such an extreme aspect ratio is numerically problematic and unnecessary. Any nonuniform deformation of the substrate will occur local to the surface and so only a small representative depth, D, of the substrate needs to be considered (where $D \ll t_s$). Results for a thick substrate can be extrapolated from these thin substrate calculations as it is known that globally the substrate is subject to a pure bending deformation. Geometric boundary constraints imply that the outer boundary of the substrate section ($r=L$) must remain linear and hence the radial displacements of this boundary must be a linear function of the axial coordinate, z, as shown in Fig. 3, as expressed in (3.1). These constraints on the boundary displacements are enforced by the introduction of appropriate Lagrangian terms into the variational functional of the elastic potential used in the finite element analysis. Six-noded triangular elements are used. The substrate mesh is quite coarse far from the surface as the displacement field is highly predictable in this region. The capillary history transformation stresses are calculated in the axisymmetric coordinate system at the elemental Gauss points and are incorporated into the nodal force matrix. The forces at each node on the island surface due to the current capillary pressure and the vertical surface tension at the edge of the island are also added into the nodal force matrix. The transformed island is allowed to expand and relax its internal elastic energy subject to the constraints placed on it by the substrate. The vertical angular rotation ϑ of the radial perimeter of the substrate is obtained and the curvature calculated from $\kappa = \vartheta/L$.

4.2 Testing the Thin Substrate Assumption. As discussed above, the aspect ratio of the problem using the actual substrate thickness is extreme and therefore it is necessary to extrapolate information from an analysis of a system with a substrate of significantly reduced thickness. The Stoney formula is used to test and validate this proposition. This formula demonstrates that, for a thin continuous film on a thick substrate, the induced curvature in a substrate due to a tensile stress σ_0 in a film of thickness t_0 is inversely proportional to the square of the substrate thickness, t_s, such that

$$\kappa_s = \frac{6\sigma_0 t_0}{\overline{E_s} t_s^2}. \tag{4.1}$$

Table 2 Properties of materials considered in this paper

Material	Young's Modulus E (GPa)	Poisson Ratio ν	Surface Energy γ (J/m^2)
SiO_2	74	0.23	-
Pt	168	0.38	2.5
Cu	130	0.34	1.75
Ag	83	0.37	1.25
Pd	121	0.39	1.95
Ni	200	0.31	2.3
Mo	329	0.31	2.85
W	411	0.28	3.2

By comparing this theoretically predicted value with the value calculated by the finite element model for a uniformly stressed continuous film on a relatively thin substrate of thickness, D, a representative small thickness can be chosen from which the thick substrate results can be satisfactorily extrapolated. The results of such a comparison are shown in Fig. 7. It is clear that the model works and that the correct curvature can be predicted to within 10 percent using a substrate that is 100 times the film thickness. Thicker substrates cause numerical problems as the finite element stiffness matrix becomes nearly singular and so we have taken $D = 100R_0$ in subsequent calculations.

4.3 A comparison of the Three Capillary History Models for Pt on SiO_2. The three curvature models defined in Eqs. (3.4) are evaluated for the Pt on SiO_2 system using the numerically calculated α-values in Table 1. The material parameters are given in Table 2. The results are compared with experimental measurements ([1]) in Fig. 8. The substrate thickness used is $t_s = 110\ \mu$m. The areal fraction model of (3.5) was used. The material parameter for model III in (3.4c) was taken to be $L/b = 14.5$ based on the best fit of (3.5) to the experimental areal fraction data for $L = 40$ Å and the lattice spacing for Pt for the surface layer thickness $b = 2.775$ Å. In this case (3.4) can be written

$$\begin{aligned} \kappa_I &= \kappa_C - 0.0008A_A = -0.0038A_A = -0.00043\overline{t_f^2}^{2/3} \\ \kappa_{II} &= \kappa_C - 0.0016A_A = -0.0046A_A = -0.00051\overline{t_f^2}^{2/3} \\ \kappa_{III} &= -0.0138A_A^{3/2} = -0.00052\overline{t_f} \end{aligned} \tag{4.2}$$

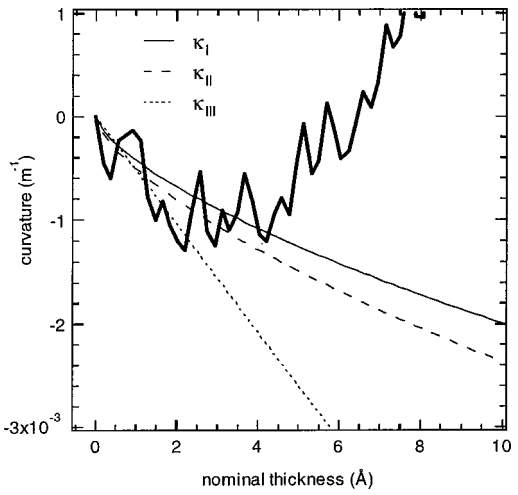


Fig. 8 Comparison between compressive stress curvature predictions of (4.2) and experimental results ([1]). All three models are of the correct order of magnitude. The models are only valid before an appreciable amount of island coalescence occurs and tensile stresses start developing in the film. The experimental data indicates that this is around 5 Å.

Table 3 α-parameters for materials with the elastic properties given in Table 2. The parameter $\gamma(\alpha_A+\alpha_B)$ gives an indication of the relative magnitude of the curvatures predicted for model II of (3.4). The α-parameters are not strongly dependent on the elastic properties of the materials and it can be seen from the rightmost column that the variation in their surface energies is more significant in determining the induced curvature.

Material	α_A	$2\alpha_E+\alpha_F+\alpha_G$	α_B	$\alpha_K+\alpha_L+\alpha_M$	$\gamma(\alpha_A+\alpha_B)$
Pt	0.229	0.059	0.115	0.028	0.860
Cu	0.196	0.087	0.133	0.041	0.575
Ag	0.163	0.088	0.155	0.042	0.398
Pd	0.212	0.064	0.131	0.030	0.669
Ni	0.215	0.082	0.111	0.038	0.750
Mo	0.222	0.061	0.085	0.030	0.875
W	0.212	0.061	0.076	0.029	0.950

where in models I and II the current capillary contributes a curvature of $\kappa_C=-\alpha_A A_A\kappa_0=-0.0030A_A$. All three models predict initial curvatures of the correct order of magnitude. The approximation made in deriving (4.2*c*) for model III slightly overestimates the curvature for $t_f<2$ Å. The scatter in the experimental results is sufficiently large to make any detailed correlation difficult. These models cease to accurately represent the stress state once the islands start to coalesce at a nominal thickness of about 5–6 Å as tensile stresses begin to develop. This is thought to be due to island coalescence ([3]).

4.4 Dependence of Curvature on Elastic Properties and Island Shape. The above results indicate that the compressive stresses present in islands on amorphous substrates in the early stages of growth are fairly insensitive to the model chosen to predict them. This has also been found to be the case with varying the island shape. Assuming semi-spherical islands, the dihedral angle between the island surface at its perimeter and the substrate surface was varied from 60 deg to 90 deg (hemisphere). The curvature was found to vary by only 10 percent over this range (despite a 30 percent change in volume) and no strong dependence on island shape was observed.

A number of other materials suitable to this type of experiment were considered. They and their material properties are listed in Table 2. The relevant α-parameters for the three models derived in this paper are given in Table 3. It is found that, although the elastic properties of the deposited material do affect the induced curvature, the surface stress is the more dominant factor. In general, the elastically softer materials such as Cu, Pd, and Ag have larger contributions from the capillary history. This is because the induced curvature depends on the transformation strain in the island which is inversely proportional to the Young's modulus of the material, as the transformation stress is not dependent on the elastic properties. The origin of a negative curvature from the current capillary contribution of models I and II requires some explanation as one might expect from examination of these forces in Fig. 2(*b*) that they could produce a positive curvature. The mechanism by which a negative curvature is produced is shown in Fig. 9. Essentially the capillary forces bend an island's edges upwards and push its center downwards. This downward central movement exerts a pressure on the substrate below, which is displaced radially outwards from the axis of symmetry generating a negative curvature. This effect is more dominant if the island has a high Poisson ratio and high Young's modulus. In fact, if the Young's modulus of the island material is reduced below 50 percent of that of the substrate material this mechanism is no longer apparent and the current capillary pressure generates a positive curvature.

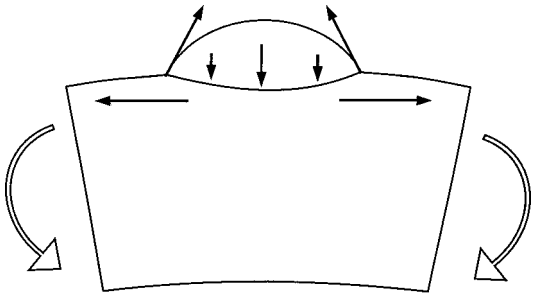


Fig. 9 The deformation mechanism (exaggerated) due to the current capillarity forces acting on an elastically stiff island generates a negative curvature. If the island is elastically much softer than the substrate then the current capillarity can generate a positive curvature.

For model III, the underlying physical arguments suggest that the substrate constrains the expansion of the bulk atoms to their equilibrium lattice spacing. This results in the induced surface stress being much less than that for an unconstrained crystallite. Finite element calculations indicate that the maximum induced surface stress in this case is less than $f/10$. This suggests that the values for f used in models I and II are excessively large. If this were the case, these models would not predict induced curvatures of the magnitude observed experimentally.

5 Conclusions

We have developed three models to explain the compressive stress in isolated crystallites during the initial growth of thin films on amorphous substrates. Models I and II account for contributions to this compressive stress curvature due to the effect of transformation strain accumulated inside crystallite layers during deposition and that due to the capillary forces acting on the surface of the crystallite. These contributions are found to be comparable for Pt on SiO_2 although this depends on the elastic properties of the system under consideration. Model III is a local transformation model in which we assume all atoms in a crystallite experience a surface-to-bulk transformation strain during the deposition process and attribute the origin of the compressive stress in the crystallite to this transformation. Model III naturally returns to the state of free crystallites when the substrate is removed and is not limited to a particular island geometry. Although model III differs fundamentally from models I and II, the induced curvatures predicted by all three models seem to agree equally well with the existing experimental results. As noted earlier, the curvature predicted by model III increases with the island size whereas models I and II do not. Hence it would be of some interest to conduct further experiments with different island sizes to clarify this distinction between these two approaches.

Acknowledgment

The work of S. Gill was supported by the EPSRC.

References

[1] Phillips, M. A., Ramaswamy, V., Clemens, B. M., and Nix, W. D., 2002, "Stress and Microstructure Evolution During Initial Growth of Pt on Amorphous Substrates," J. Mater. Res., accepted for publication.
[2] Nix, W. D., and Clemens, B. M., 1999, "Crystallite Coalescence: A Mechanism for Intrinsic Tensile Stresses in Thin Films," J. Mater. Res., **14**, p. 3467.
[3] Abermann, R., Kramer, R., and Maser, J., 1978, "Structure and Internal Stress

in Ultrathin Ag Films Deposited on MgF_2 and SiO Substrates,'' Thin Solid Films, **52**, p. 215.
[4] Abermann, R., and Koch, R., 1979, ''Internal Stress on Thin Silver and Gold Films and Its Dependence on Gas Absorption,'' Thin Solid Films, **62**, p. 195.
[5] Abermann, R., and Koch, R., 1980, ''In situ Determination of the Structure of Thin Metal Films by Internal Stress Measurements: Structure Dependence of Ag and Cu Films on O_2 Pressure During Deposition,'' Thin Solid Films, **66**, p. 217.
[6] Cammarata, R. C., and Sieradzki, K., 1994, ''Surface and Interface Stresses,'' Annu. Rev. Mater. Sci., **24**, p. 215.
[7] Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, McGraw-Hill, New York, pp. 392–395.

W. W. Gerberich

N. I. Tymiak

J. C. Grunlan

Department of Chemical Engineering
and Materials Science,
University of Minnesota,
421 Washington Avenue, S.E.,
Minneapolis, MN 55455

M. F. Horstemeyer

Center for Materials and Engineering Sciences,
Sandia National Laboratories,
MS 9404,
Livermore, CA 94551-0969

M. I. Baskes

Los Alamos National Laboratory,
MSG 755,
Los Alamos, NM 87545

Interpretations of Indentation Size Effects

For very shallow indentations in W, Al, Au, and Fe-3wt%Si single crystals, hardness decreased with increasing depth irrespective of increasing or decreasing strain gradients. As such, strain gradient theory appears insufficient to explain the indentation size effect (ISE) at depths less than several hundred nanometers. Present research links the ISE to a ratio between the energy of newly created surface and plastic strain energy dissipation. Also, the contact surface to plastic volume ratio was nearly constant for a range of shallow depths. Based on the above, an analytical model of hardness versus depth provides a satisfactory fit to the experimental data and correlates well with embedded atom simulations. [DOI: 10.1115/1.1469004]

Introduction

Material length scales have been a subject of great interest to nearly all engineering and science disciplines. Of large interest to the mechanics/materials community are those material length scales in the 0.1 to 10 μm regime, that have been investigated with small volume torsion wire ([1]) and nanoindentation ([2,3]) experiments. The principle theoretical treatment, with origins in gradient microstructure analysis ([4–7]), has been the use of strain gradient plasticity approaches ([1,3,8–10]). Most effective in tying the materials and mechanics approaches together is a microstructurally based strain gradient analysis propose by Gao et al. [9] and most clearly applied to the indentation size effect by Nix and Gao [3]. This utilized the time-honored strain gradients from geometrically necessary dislocation relationships ([11]) which have been repeatedly verified by experiment ([2,12]).

In the same time frame, propelled by the discoveries of scanning tunneling and atomic force microscopies, principally the physics and chemistry communities addressed much lower contact forces in the nanonewton regime and examined various nanotribology issues. As polymer surfaces came under scrutiny, such probes were elevated to larger μN forces by using stiffer stainless steel cantilevers and previous continuum models are being used to examine adhesive forces ([13]). These involve using the Johnson Kendall Roberts (JKR) ([14]) and Derjaguin Muller Toporov (DMT) ([15]) approaches, and later an incorporation of a Dugdale zone to smoothly obtain the JKR/DMT transition ([16]).

In the region of scale between these atomistic and gradient microstructure regimes, there are possibly one or more phenomena that may contribute to an indentation size effect. It is our intent here to review briefly a number of these effects which can be related to small volumes under contact. This is possible through some recent nanoindentation results ([17]). Specifically, because the previous experimental body of literature ([2,3]) dealing with the indentation size effect emphasized depths of penetration greater than 100 nm and atomic force microscopy had emphasized nm level contacts, we recently evaluated two materials, ⟨100⟩ crystals of aluminum and tungsten, in the regime of 10 to 500 nm ([17]). At the suggestion of Baskes and Horstemeyer [18] who considered the surface to volume ratio to be key, we both experimentally and theoretically determined plastic zone sizes and surfaces of contact to assess the importance of this parameter. At the time neither research group had a physical rationale of why this was important except that atomistic simulations on the one side ([19]) and an experimental evaluation on the other ([17]) strongly suggested that surface to volume, S/V, was the critical parameter. In the present study we group two more materials ⟨001⟩ Fe-3wt%Si and ⟨001⟩ Au with the previous two for critical evaluations of possible contributions to contact forces. First, however, it is appropriate to briefly review the background of two recent studies and a hierarchy of scales that may influence contact forces and therefore any indentation size effect (ISE).

Background

Interest in the indentation size effect (ISE) has resurfaced a number of times; e.g., in 1970 Gane and Cox [20] demonstrated that in Au single crystals that hardness could be increased by a factor of three by decreasing the contact diameter from 10^4 to 10^2 nm. A rekindled interest in the ISE was fostered by the availability of depth sensing instrumentation ([21]) and nearly two decades later, Stelmashenko et al. [22] showed a similar hardness increase at shallow depths in various orientations of single crystal tungsten. They also had a reasoned explanation in terms of the local dislocation hardening due to geometrically necessary dislocations. The hardness, H, was given by

$$H = A\alpha\mu\left[b^2\rho_0 + \frac{b\cot\beta}{d}\right]^{1/2} \tag{1}$$

with A and α nondimensional coefficients of constraint and Taylor hardening, μ the shear modulus, b the burgers vector, ρ_0 a background dislocation density, $\cot\beta$ the wedge shape, and d the diagonal of a Vicker's hardness diamond indenter. Using a reason-

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, March 15, 2001; final revision, January 8, 2001. Associate Editor: D. A. Kouris. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

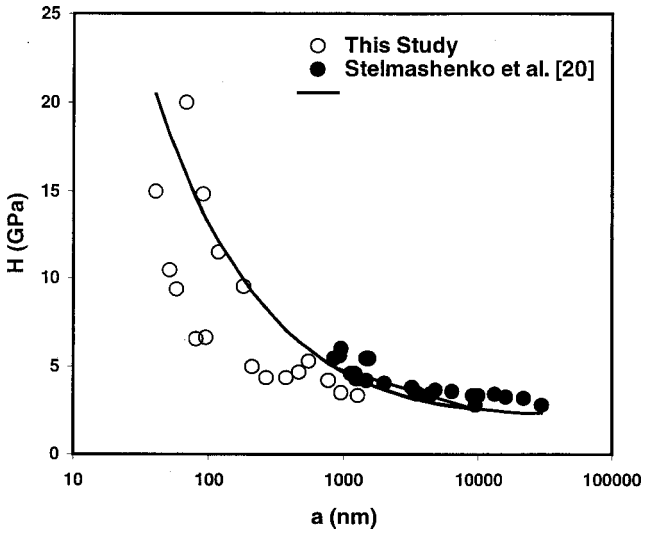


Fig. 1 Hardness as a function of contact dimension in ⟨100⟩ tungsten crystals showing the ISE

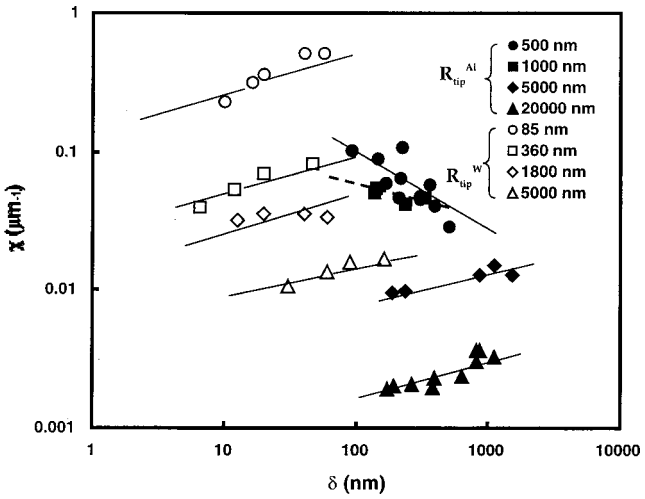


Fig. 2 Average strain gradients, χ, versus depth of penetration, δ, into ⟨100⟩ W (open symbols) and ⟨100⟩ Al (closed symbols) crystals. Four different diamond tip radii used in each case.

able value of $A\alpha \sim 1.5$, the fit to both their data and more recent data for spherical tips ([17,23,24]) is seen in Fig. 1 for single crystal tungsten. This agreement is remarkable considering that the spherical indenters ranged from 85 to 5000 nm in radius, the Vicker's indenter is a sharp, four-sided pyramid and Stelmashenko's data represents five crystal variants while the present data is only for ⟨001⟩. Note that $d/2.5$ is used for the comparison to make equivalent areas of contact for the two types of indenter tips. As some of the low data points for the spherical contact radii, a, represent contact depths of only 10 nm, the scatter observed could be partially a result of surface roughness ([25,26]).

Such findings along with other observed material scale effects led to a phenomenological theory of strain gradient plasticity by Fleck and Hutchinson [1,27] and somewhat later to a mechanism-based strain gradient plasticity theory by Gao et al. [9]. While the mechanism-based theory can ([9]), it is generally recognized that such strain-gradient plasticity theories should not be used at very shallow depths in the vicinity of 100 nm or less. Also, such small size scales become close to the realm of atomistic simulations where depths of penetration of 1 nm have easily been achieved ([28]). These two facts caused us recently to examine the ISE effect in some detail both experimentally ([17]) and computationally ([19]). Regarding the experiments, both average plastic strains, $\overline{\varepsilon_p}$, and strain gradients, $\overline{\chi} = d\varepsilon_p/dc$, were estimated from experimental measures of $\overline{\varepsilon_p}$[1] and plastic zone size, c. As summarized in Fig. 2, the average gradient for a given indentation, $\overline{\chi}$, increased slightly for both small and large spherical tip radii at shallow depths of penetration less than 100 nm. However, deeper penetration depths in the single crystal aluminum produced decreasing values of $\overline{\chi}$ with increasing depths for the sharpest cones but remained spherical-like for the bluntest ones. This can be rationalized partially from simple geometric arguments as Nix and Gao [3] have done for geometrically necessary dislocations emanating from a sharp wedge. The comparison in Fig. 3 suggests that for the spherical tip the average gradient for a spherical contact would be independent of the contact radius (or depth) while the value of χ would fall off as $1/a$ for a sharp wedge. This reinforced current views that continuum-based gradient plasticity models are not appropriate to very small plasticity scales. That is, hardnesses at the shallowest depths were accompanied by strain gradients which were increasing or nearly constant for indentations at greater depth. For that reason an *ad hoc* model was initially determined based upon the observed c/a ratio which was found to first order to be ([17])

$$\left(\frac{c}{a}\right)^2 \simeq \frac{\eta}{a}; \quad \eta = \text{material constant.} \tag{2}$$

Given that even in this small scale range that the plastic zone size is described well by Johnson's cavity model ([29,30]) one can use

$$c = \left[\frac{3P}{2\pi\sigma_f}\right]^{1/2} \tag{3}$$

where P is the applied indentation load and σ_f is an appropriate flow stress. Coupled with Eq. (2) it was shown that a first-order prediction of the ISE could be made ([17]). This still wasn't satisfying, however, as there are no principles of physics or mechanics involved that would explain the indentation size effect.

One can take another look at Eq. (2) and see that if the contact area, πa^2, is coupled to a hemispherical volume of deforming material, $(2/3)\pi c^3$, then a surface-to-volume ratio can be defined as

$$\frac{S}{V} = \frac{3a^2}{2c^3} \tag{4}$$

or

$$\frac{S}{V} = \frac{3}{2\eta} \cdot \frac{a}{c}, \tag{5}$$

Eq. (5) deriving from the size scale parameter of Eq. (2). Given that the a/c ratio was mildly varying by only a factor of two with contact radii up to about 1 μm, this suggests for shallow depths of penetration that the surface to volume ratio defined by Eq. (5) is nearly constant!

At the same time, molecular dynamics simulations using embedded atom method potentials ([19]) were performed on single crystal nickel. With simple shear boundary conditions, Horstemeyer and Baskes ([19]) showed a dependence of the yield stress on the specimen size. It was proposed that dislocation nucleation was a critical phenomena that determined the yield point as a function of volume-to-surface area. This initially caused some confusion between the two research groups until it was clarified as to how volume-to-surface area was being assessed. This is discussed in a note added in proof.

Since the surface-to-volume ratio appeared to play a pivotal role for shallow contacts, we reviewed contact forces over the

[1]These will be addressed in more detail in the experimental section which follows.

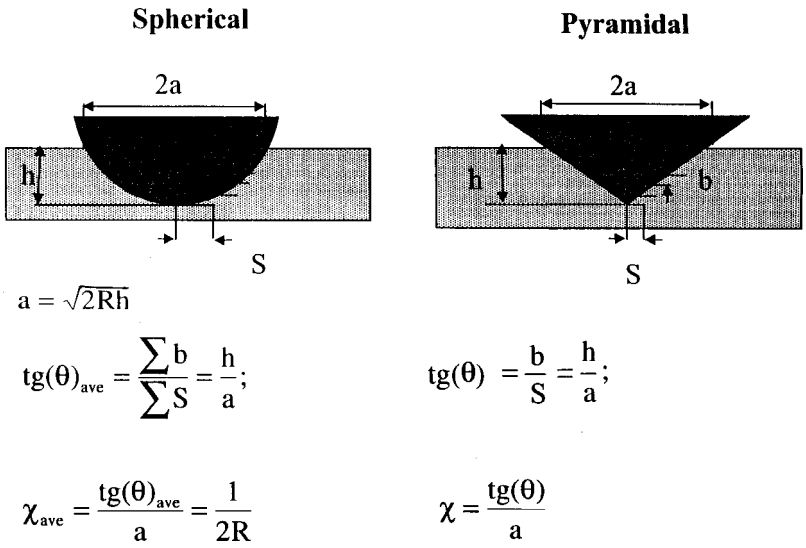


Fig. 3 Schematic of spherical and sharp wedge contacts showing difference in strain gradient dependence on contact shape

scale of interest, $\leqslant 10$ μm, as summarized in Table 1] ([2,3,12,31–41]). For the present study we believe the first three categories represent forces too small to be significant to an ISE. Clearly, adhesive forces can be important particularly to polymer contacts where pull-off forces can be in the 10 μN regime. However, for a number of metal/metal and diamond/metal contacts in laboratory air, the pull-off forces were less than 1 μN representing a force generally less than about three percent of the total.

Consider, next, the level of contact forces associated with surface energy and surface stress. While the forces in a narrow annulus around a surface contact have a limited area of influence, surface energies and excess surface stress may act over areas at least as large as the extent of the plastic zone around the contact. Consider first surface energy, γ_s, which is the work to create new surface as might be related to creating cracks in an oxide film below the contact or new surface as slip steps emerge around the indenter. In another type of experiment on small wires, finite loads are found where creep rates become zero because of the surface energy effect ([33]). Regarding surface stress, the work in elastically deforming surface atoms, this is sufficiently large to trigger surface reconstruction and lattice parameter changes in small spheres ([34–36]). For example, if one considers the surface work only associated with the surface stress, σ_s, this can be converted through the Laplace pressure on a sphere to a change in lattice parameter ([16]). (See Appendix A.) Such observations of lattice parameter changes by TEM electron diffraction have been found ([34–36]). Regarding the next scale level of forces and lengths, there has been a large effort ([38–40,42,43]) in attempting to describe both nucleation and yield forces associated with the onset of plasticity and the arrest of a displacement excursion.

Finally, the well-documented ISE for sharp-wedge tips driven into single and polycrystals and interpreted in terms of length scales associated with strain gradient plasticity needs little introduction here ([2,3]). Such indentation experiments can be described by

Table 1 Contact Forces: a Hierarchy of Scales

Phenomena	Observations	Forces/Energy	Scales
Jump to contact ([3])	STM/AFM	van der Waals	nM/Å
Film dewetting ([32])	AFM	van der Waals versus surface tension (Laplace pressure)	nN/nm
Adhesion ([16])	Surface forces apparatus AFM/nanoindentation	van der Waals, chemical bonding: JKR/DMT	μN/nm
Creep of small volume wires ([33])	Load at zero creep as a function of wire diameter—ISE?	Surface energy, γ_s (work to create a new boundary)	μN/nm
Lattice parameter (a_0) changes ([34–36])	TEM observations show a_0 scales as $1/R$ in spheres		
Thin film yielding under an electrolyte ([37])	Voltage driven surface stress affects nanoindentation induced yielding	Surface stress, σ_s (work to elastically deform surface atoms)	μN/nm
Yield initiation ([38–40])	Drumhead effect; oxide film effect in nanoindentation—ISE?	Film elasticity, image forces, film fracture	μN/nm $\downarrow$ mN
Yield arrest ([24,40])	Displacement excursion in indentation—ISE? TEM, SEM Rosettes	Dislocation array stress field	μN/nm $\downarrow$ mN
Small wire torsion ([41])	Torque versus wire size or	Strain-gradient modified	μN/nm
Indentation size effects (ISE) ([2,3,12])	hardness in sharp-wedge penetration (ISE)	Taylor hardening	$\downarrow$ $\downarrow$ mN μm

Table 2 Details of load (P), displacement (δ), contact radius (a), plastic zone radius (c), hardness (H), and surface to volume (S/V) at the end of yield excursions in ⟨100⟩ Fe-3wt%Si, R_{tip}=70 nm. (Note that only every other data point in Fig. 7(c) is detailed here.)

P μN	δ nm	a nm	c nm	H GPa	H/σ_{ys}	S/V $m^{-1}\times10^{-5}$	$\left(\frac{S}{V}\right)^2\left(\frac{H}{\sigma_{ys}}\right)^3$ $m^{-2}\times10^{-14}$
38	8	33.5	224	10.8	30	1.50	6.08
50	10.5	38.2	258	10.9	30.3	1.27	4.49
60	21	54.2	282	6.5	18.1	1.96	2.28
74	24	58.0	313	7.0	19.5	1.65	2.02
94	40	74.8	353	5.35	14.9	1.91	1.21
100	40.5	75.9	364	5.52	15.3	1.79	1.14
114	48	85.7	389	4.93	13.7	1.88	0.91
118	52	91.1	396	4.53	12.6	2.00	0.80
150	62	104	446	4.43	12.3	1.82	0.62
155	70	114	453	3.79	10.5	2.10	0.51
163	70	114	465	3.98	11.1	1.94	0.52
175	84	132	482	3.19	8.9	2.02	0.39
198	95	146	512	2.95	8.2	2.40	0.31
In ⟨100⟩ Au, R_{tip}=205 nm with only every third data point in Fig. 7(a) detailed:							
16	6.5	51.6	309	1.99	24.9	1.35	2.81
20	7.6	55.8	345	2.04	25.5	1.14	2.15
21	9.0	60.7	354	1.81	22.6	1.25	1.80
30	12	70.1	423	1.94	24.2	0.97	1.33
48	14	75.8	535	2.66	33.2	0.56	1.15
40	15.5	79.7	489	2.00	25	0.81	1.04
50	16.5	82.2	546	2.35	29.4	0.623	0.986
44	18	85.9	512	1.90	23.8	0.825	0.918
48	20	90.6	535	1.86	23.2	0.804	.807
47	22	95.0	530	1.66	20.7	0.989	.733
58	25	101	588	1.81	22.6	0.832	.799
62	27	105	608	1.79	22.4	0.759	.647
72	30	111	656	1.86	23.3	0.654	.541
70	33	116	646	1.65	20.6	0.749	.490
71	34	118	651	1.62	20.2	0.757	.472
88	40	128	725	1.71	21.4	0.645	.408
92	47	139	741	1.52	19	0.712	.348
98	52	146	765	1.46	18.2	0.714	.307

$$\frac{H}{H_0}=\sqrt{1+\frac{\delta^*}{\delta}} \tag{6}$$

where H_0 is the hardness in the absence of strain gradients, δ^* is a length scale parameter, and δ is the depth of penetration. We would like to point out two aspects. First, the form of Eqs. (1) and (6) are identical with $H_0=A\alpha\mu/c_1$ and $\delta^*=c_2b$ but the interpretation is different. Stemashenko [22] originally only considered hardening due to geometrically necessary dislocations while Nix and Gao [3] considered a flow stress given by

$$\sigma_f=\sigma_{ys}\sqrt{f^2(\varepsilon)+\delta^*\chi_{\rho G}} \tag{7}$$

with hardening a function of both strain and strain gradient. The second point is that both average strain and average strain gradients tend to increase with increasing depth for ''sharp tips'' at very shallow depths or for larger spherical tips at all depths. See Fig. 2. Since Eq. (7) implies larger flow stresses for deeper depths of penetrations with spherical tips, this did not seem to explain the shallow penetration data of Fig. 1. Some corroboration of this at light loads was found by Yasuda et al. [43] who documented that the dislocation density of the isotropic plastic zone increased with increasing depths of penetration. For that reason they rejected the model of Stelmashenko et al. [22] which only emphasized the dislocation density aspect. The above aspects led us to the conclusion that for very shallow indents that a model based upon surface, S, to volume relation, V, considerations might be promising.

Experimental

Details of the loads, displacements, plastic zones, contact radii, strains, and strain gradients are given elsewhere ([17]) for ⟨100⟩ Al and ⟨100⟩ W. For this study also analyzed in the same way were ⟨100⟩ Au [44] and ⟨100⟩ Fe-3wt%Si ([17–23]) single crystals. Justification in terms of geometrically necessary dislocations gave strains as δ/c and average gradients, $\overline{\chi}$, as δ/c^2. Tip radii of 500, 1000, 5000, and 20,000 nm were used for the ⟨100⟩ Al, 85, 360, 1800, and 5000 nm for the ⟨100⟩ W and single tips of 70 nm and 205 nm for ⟨100⟩ Fe-3wt%Si and Au, respectively. In one instance a 700 nm spherical diamond tip was used for tungsten as noted. Since the Fe-3wt%Si and Au were not published elsewhere, these are shown in detail in Table 2. The Fe-3wt%Si tip used to determine strain gradient was relatively sharp and since it was a three-sided Berkovich, an appropriate area function for the transition from spherical to triangular was utilized to define contact area and an effective radius of contact.[2] For further justification of the strain and strain gradients used, we had accomplished a numerical analysis of a very shallow contact into simulated nanocrystalline tungsten with a yield strength of 4 GPa and a modulus of 400 GPa. The 500 nm spherical tip at the end of a 90-deg cone was driven into the tungsten to a depth of 167 nm. The approach ([45]) used was an explicit, numerical formulation utilizing the finite difference method for a three-dimensional model based on large deformation, elastoplastic contact mechanisms. Moving boundaries and quasi-static states were handled with an updated Lagrangian approach. The resulting Mises strain as a function of distance from the tip is shown in Fig. 4 with strains up to about 0.5. The average strain in the zone was 0.154 whereas if we consider δ/c it is 0.209 for this 800 nm plastic zone. Further, one can curve fit and show the average gradient in the zone to be about 5×10^5 m^{-1} while δ/c^2 is 2.6×10^5 m^{-1}. The difference between

[2]Previously, we had shown for different shapes of indenters that Eq. (3) gave an appropriate account of measured plastic zones ([30]). While such a shape change at deeper depths will not affect the surface to volume ratio or mean pressure determinations, this may bias their relationship with displacement. The reader should be aware that such variance would cause a small shift in some of the data plots where depth to the one-third power is encountered.

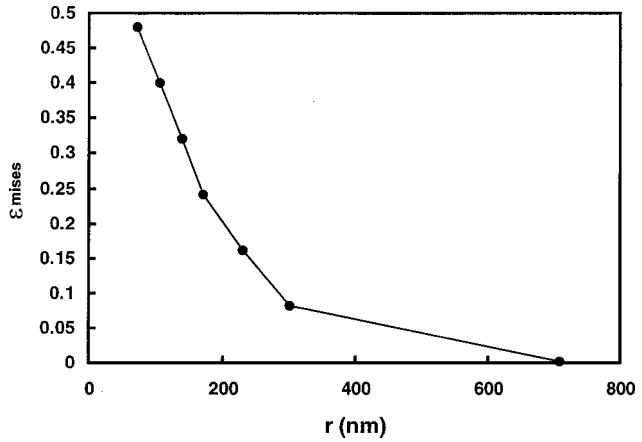


Fig. 4 Mises strain as a function of distance, *r*, from indenter tip for a three-dimensional finite difference numerical analysis

these and the $1/2R$ gradient of 10^6 m^{-1} indicated in Fig. 3 is due to the ''gage length'' basis of ''*c*'' and ''*a*'' used, the former being preferred.

Theoretical

To understand how important the surface to volume, S/V, ratio is, estimates of surface work and the volume work associated with plastic deformation under a contact are in order. Considering first the surface work, a number of contributions including both surface energy and surface stress are examined. For the general case the total work associated with creating either new area, dA, or new surface energy, $d\gamma$, can be written as

$$gdA = \gamma dA + Ad\gamma \tag{8}$$

using Maugis' terminology ([16]). He goes on to demonstrate that for a surface of N atoms with an area $A = Na_0$ that the surface work is due to both elastically stretched bonds, da_0, and new numbers of atoms exposed, dN, giving

$$gdA = \gamma a_0 dN + \sigma_s N da_0 \tag{9}$$

where γ is surface energy and σ_s is the surface stress. With the definition of strain dA/A being both plastic, ε_p, and elastic, ε_ε, it follows that ([46])

$$g = \gamma \frac{\varepsilon_p}{\varepsilon} + \sigma_s \frac{\varepsilon_E}{\varepsilon} \tag{10}$$

where ε is the total strain. Thus, the total surface work, W_s, is made up of new area associated with irreversible plastic work as well as elastically stretched bonds associated with reversible work. There are a number of possible contributions to indentation surface work as follows:

(i) creating new surface associated with straining material outside the contact. This could involve oxide fracture, γ_s^{ox}, oxide/metal interface fracture, γ_i^{m-ox}, and/or slip step emergence in the metal itself, γ_s^m:

$$W_s^i = \pi(c^2 - a^2)[f_1\gamma_s^{ox} + f_2 w^{m-ox} + f_3\gamma_s^m];$$
$$w^{m-ox} = \gamma_s^m + \gamma_s^{ox} - \gamma^m. \tag{11a}$$

(i) creating surface by fracture of oxide, γ_s^{ox}, or the metal/oxide interface, $m-ox$, giving

$$W_s^{ii} = \pi a^2[f_4\gamma_s^{ox} + f_5 w^{m-ox}]; \quad w^{m-ox} = \gamma_s^m + \gamma_s^{ox} - \gamma_s^{mox}. \tag{11b}$$

(i) formation of adhesion between the diamond indenter-tip and the oxide surface film:

$$W_s^{iii} = \pi a^2[f_6 w^{d-ox}]; \quad w^{d-ox} = \gamma_s^d + \gamma_s^{ox} - \gamma_s^{dox}. \tag{11c}$$

(i) surface stress deformation work outside the area of contact

$$W_s^{iv} = \pi(c^2 - a^2)[f_7\sigma_s]. \tag{11d}$$

Note here that the interfacial energies, w, are the Dupré works of adhesion and the f_i are the fraction of areas contributing. Since $f_1 = f_4$, $f_2 = f_5$, one can show to first order for an annulus of plastic deformation of $c \sim 2a$ that these sum to

$$W_s^{\text{total}} \sim \sum_i w_s^i \sim 4\pi a^2 \overline{f}[\gamma_s^{ox} + \gamma_s^m];$$
$$\gamma_s^{ox} \sim \gamma_s^m, \overline{f} \simeq 0.125$$
$$W_s^{\text{total}} \sim \pi a^2 \gamma_s^m. \tag{12}$$

Further simplifying assumptions were that we could ignore surface stress, that f_1 and f_2 for fractured oxide or metal/oxide interfaces were about a quarter of the contact area fraction under the indenter, f_6, and that $\gamma_s^d \sim \gamma_s^{ox} > \gamma^{dox}, \gamma^{mox}$. While one can easily argue the details, the simple relation we desired here is that to first order the total surface work is given by the product of the contact area and metal surface energy.

Next, consider the volume deformation associated with plastic work, W_p. From a continuum standpoint W_p can be defined in terms of the plastic volume, V, and the tensile yield stress σ_{ys}, for an elastic-perfectly plastic material giving

$$W_p = V\int_0^{\varepsilon_p} \sigma_{ys} d\varepsilon_p. \tag{13a}$$

Defining an incremental strain as before $d\varepsilon_p = d\delta/c$ and noting that the hemispherical plastic volume would be $(2/3)\pi c^3$ one sees that

$$W_p = V\int_0^{\delta} \frac{\sigma_{ys}}{c} d\delta = \frac{2}{3}\pi c^2 \sigma_{ys}\delta. \tag{13b}$$

There is a hidden assumption that ''*c*'' is constant but recent experiments ([47]) with direct AFM imaging demonstrate a substantial plastic zone prior to the yield excursion. An alternative way of examining this work is to define an incremental work similar to that proposed by Cheng and Cheng [48]. Using Eq. (3) with $\sigma_f \sim \sigma_{ys}$ gives

$$dW_p = Pd\delta = \frac{2\pi\sigma_{ys}c_i^2}{3} d\delta \tag{13c}$$

with c_i the instantaneous plastic zone size. If we assume that $c_i = c =$ constant, then

$$W_p = \frac{2\pi\sigma_{ys}c^2}{3}\int_0^{\delta} d\delta = \frac{2}{3}\pi c^2\sigma_{ys}\delta. \tag{13d}$$

It is seen that (13*b*) and (13*d*) are identical. Further corroboration is taken from the dislocation theory by assuming that concentric loops of length from near zero to c ($c/2$ on average) move down glide cylinders to produce work under a shear stress, τ_{ys}. This gives

$$W_p = \underset{\substack{\text{number}\\ \text{of loops}}}{\frac{\delta}{b}} \cdot \underset{\substack{\text{force per unit length}\\ \text{x average loop length}}}{\tau_{ys} b \pi c} \cdot \underset{\text{distance moved}}{c} = \pi c^2 \tau_{ys}\delta. \tag{13e}$$

Taking $\tau_{ys} \sim (2/3)\sigma_{ys}$ makes (13*b*) = (13*d*) = (13*e*). Still since $\overline{c} < c$ but only slightly and the displacement excursions have been shown to be about half the total displacement ([40]), it follows that a good estimate of the plastic work is

$$W_p \simeq c^2 \tau_{ys}\delta_{exc}. \tag{14}$$

Using the total displacement with Eqs. (12) and (14), the surface to volume work ratio is given by

$$\frac{W_s}{W_p} = \frac{2\pi\gamma_s^m}{\tau_{ys}\delta}\left(\frac{a}{c}\right)^2. \tag{15}$$

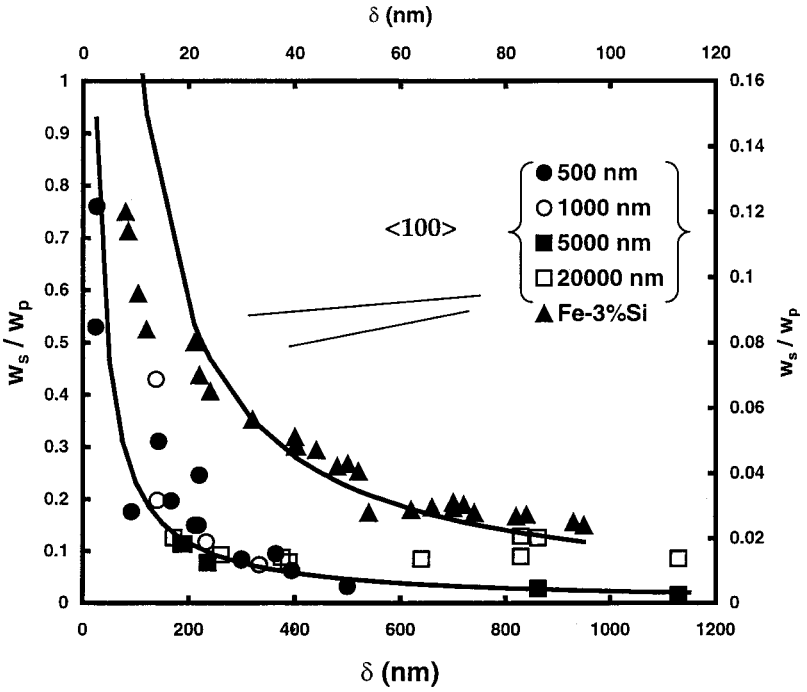


Fig. 5 Ratio of surface to volume works as a function of indentation depth into ⟨100⟩ Al and ⟨100⟩ Fe-3wt%Si single crystals

What this illustrates is that if the a/c ratio is nearly constant at shallow penetration, then the portion of surface work resisting penetration rapidly decreases with increasing depth of penetration as γ_s^m and τ_{ys} are constant. Back at Eq. (5) it was suggested that for relatively shallow depths of penetration that S/V may be nearly constant. We next examine this and the surface work argument as to how this gives rise to an indentation size effect.

Results and Discussion

First, both W_s and W_p were determined from Eqs. (12) and (14) using either calculated values of c and a as for Fe-3wt%Si or AFM imaged values for aluminum crystals. Values for shear yield stress were 180 MPa for Fe-3wt%Si and 30 MPa for aluminum. These data as shown in Fig. 5 illustrate two points. First, for very shallow penetrations the surface work can be a significant portion of the total work. While a surface work contribution of 10% may not seem significant we believe the values in Fig. 5 to be an underestimate of the true ratios. After these calculations were made it was suggested that the surface area could easily be larger than the πa^2 used if the vertical surface steps associated with both oxide fracture and slip band emergence were used, e.g., pile up. Consider just a native oxide fracture of 3 nm, and approximately half of the dislocation loops emerging at the free surface to fracture oxide. One can show that the surface steps as long traces through the plastic zone could produce a new area about three times larger than πa^2. Second, the solid curves are Eq. (15) with c/a a single value of three for both materials. Going back to Eq. (5) then this does strongly imply that if the size scale parameters were relatively unchanged, that the surface to volume ratios would also be constant and near to each other. From data of previous ([17,23,44]) and present investigations, S/V values were calculated and are shown as a function of depth in Fig. 6. Here it is remarkable to see how constant the S/V values are up to depths of about 200 nm (note the scale difference for aluminum). The slightly greater scatter for Al and W can be partially attributed to using four different tip radii in each case. As to the average S/V

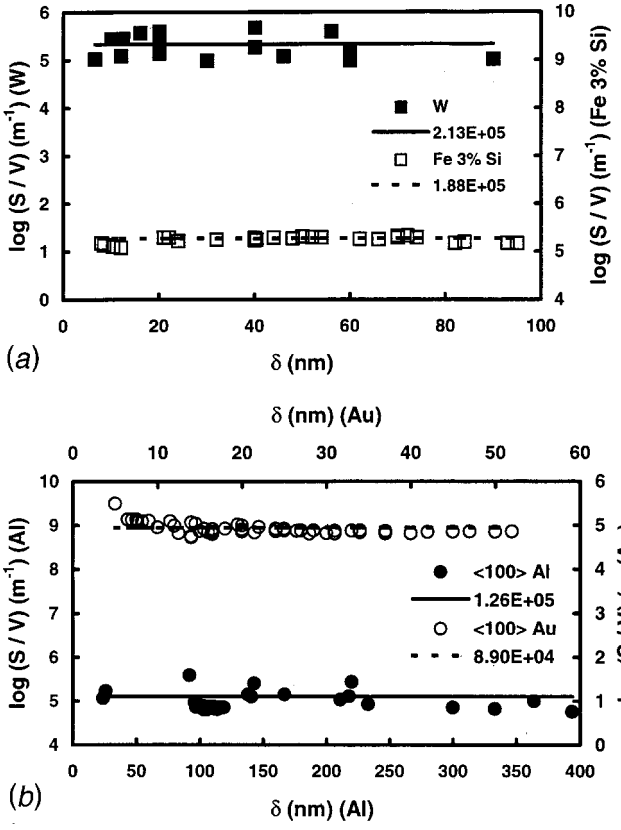


Fig. 6 Surface-to-volume ratio, defined by projected contact area to plastic volume, as a function of indentation depth for (*a*) ⟨100⟩ W and ⟨100⟩ Fe-3wt%Si; (*b*) ⟨100⟩ Au and ⟨100⟩ Al single crystals. Solid and dashed curves represent the mean *S/V* values for each material. Note the different scales for the two materials in (*b*).

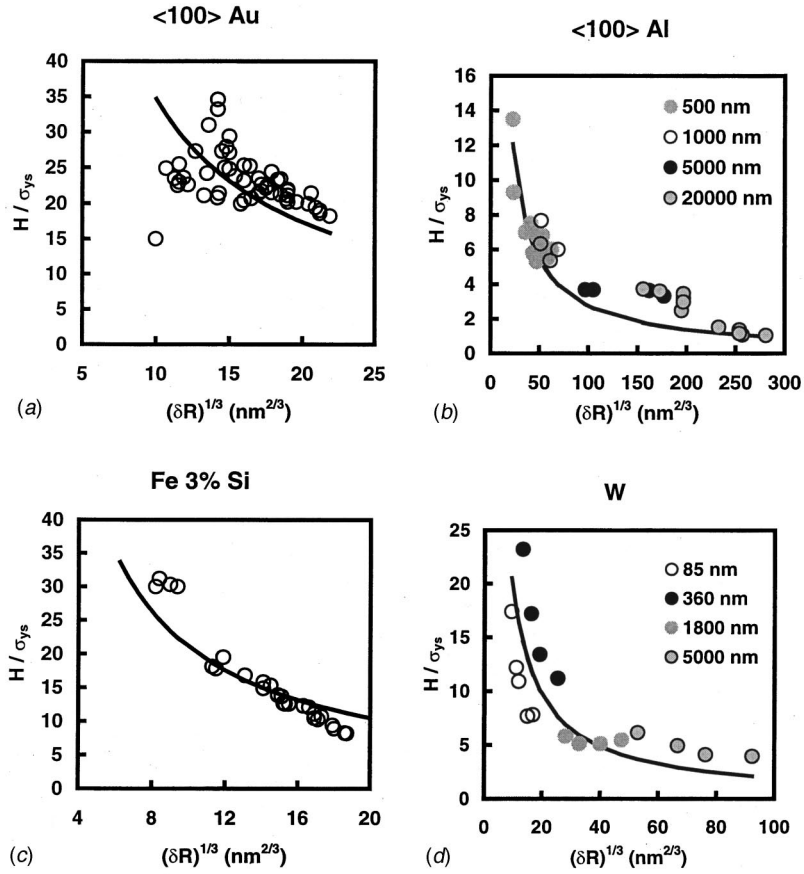


Fig. 7 Fit of the proposed model (Eq. (19)) for four ⟨100⟩ oriented single crystals. Single tips of 205 nm and 70 nm radii used in (*a*) and (*c*), respectively; multiple spherical tips with radii noted used in (*b*) and (*d*)

values, these tended to be in the $(1\text{-}2)\times 10^5$ m^{-1} regime for all four materials, the ranking of which we consider later.

With $S/V\sim$constant for a given material at shallow depths, it was straightforward to arrive at an ISE. Defining $S/V=3a^2/2c^3$, it is seen that

$$\left(\frac{S}{V}\right)^{2/3}=\frac{(3/2)^{2/3}a^{4/3}}{c^2}. \tag{16}$$

For plastic contacts of a spherical tip Johnson [29] takes $a^2\sim 2\delta R$ which is nearly the geometric value and combining this with the mean pressure defines hardness as

$$H\simeq\frac{P}{\pi a^2}\simeq\frac{P}{2\pi\delta R}. \tag{17}$$

But we had already demonstrated that the plastic zone could be given by Johnson's cavity model, Eq. (3), so that by eliminating P, (17) becomes

$$H=\frac{\sigma_f c^2}{3\delta R}. \tag{18}$$

It is seen with S/V constant, Eq. (16) and $a^2\sim 2\delta R$ can be used to eliminate c^2 in Eq. (18) giving

$$H\simeq\frac{\sigma_f}{\left(\frac{S}{V}\right)^{2/3}}\cdot\frac{1}{(3\delta R)^{1/3}}. \tag{19}$$

H/σ_f is shown versus $(\delta R)^{1/3}$ for all four materials in Fig. 7. With the average values of S/V from Fig. 6, it is also seen that the model fits the data both in terms of the ISE but also appropriately ranking the tip radius effect. It should be mentioned here that σ_f was taken as the yield stress for Au, Al, and Fe-3 wt%Si but was taken as the flow stress for W as discussed by Bahr et al. [23]. As seen from Eq. (19), $(S/V)^2\ (H/\sigma_{ys})^3$ should collapse all of the data when shown versus $(\delta R)^{-1}$. Such a master plot in Fig. 8

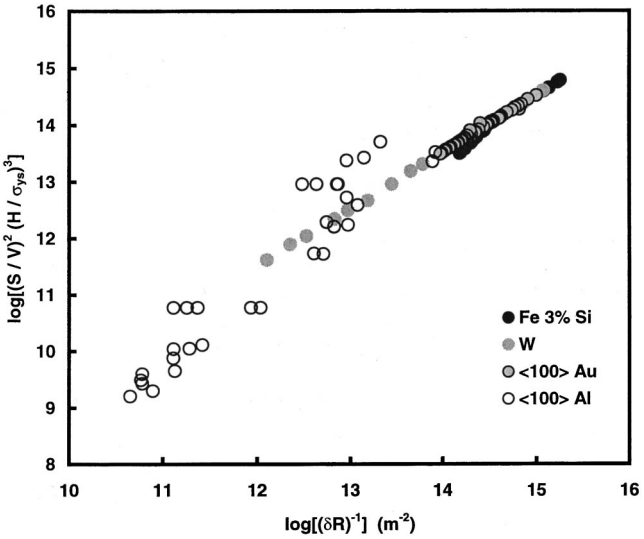


Fig. 8 Master plot of Eq. (19) for all materials

Table 3 Scaling of flow stress, surface energy, and elasticity to surface/volume ratio

	Au	Al	Fe	W
γ_s N/M†	1485	980	1950	2800
c_{11}, Pa$\times 10^{-10\dagger}$	18.6	10.8	24.2	52.1
σ_f, Pa$\times 10^{-6}$	80	60	360	860
$\sigma_f/c_{11}\times 10^4$	4.3	5.6	14.9	16.5
$\sigma_f^2/\gamma_s c_{11}$, m^{-1}	23	34	275	507
S/V, m$^{-1}\times 10^{-5}$	0.89	1.26	1.88	2.13

†Data from Hirth and Loethe, Ref. [49].

does this but it is seen that a break in the curve appears at a δR of about 3×10^{-13} m^2 for the aluminum data which involved the deeper penetrations. For the 1000 nm tip radius this represents a penetration depth of 300 nm, just about where the surface to volume ratio ceases to be constant in Fig. 6. It appears then that a different mechanism for the ISE may take place at deeper penetration depths where the surface work contribution is minimal.

One final discussion point is the apparent ranking of average S/V ratios in Fig. 6. One way is through the physical properties of the materials as given in Table 3 [49]. Here it is seen that the two lowest surface energy metals have the lowest S/V ratios which seems counterintuitive. On the other hand, the ratio of flow stress to the elastic stiffness constant, c_{11}, does seem to rank order the S/V values. In addition the elastic strain energy density clearly scales with σ_f^2/E so that the greater stored elastic energy about indents into the higher yield strength materials may be playing an additional role in requiring an increased surface-to-volume ratio. If that is the case then dimensional analysis would imply that $\sigma_f^2/\gamma_s c_{11}$ should scale with S/V as these both have m^{-1} units. Comparing the last two rows of Table 3 shows this to be the case on a semi-log basis.

Note added in proof: Sometime after this paper was 95 percent complete our respective research groups contacted each other and found that our surface to volume concepts didn't even fit on the same page. It became readily apparent that we had utilized two different definitions of volume wherein V_B represented the volume displaced by the indenter whereas V_G represented the plastic volume involved in volume work. To properly compare our data we needed to define $(V/S)_B$ for our spherical indentations. For a spherical segment indenting a planar surface, V and S are defined by geometry, giving

$$\left[\frac{V}{S}\right]_B = \frac{\frac{\pi\delta^2}{3}[3R-\delta]}{2\pi R\delta} = \frac{\delta}{2} - \frac{\delta^2}{6R} \approx \frac{\delta}{2}; \quad R \gg \delta \tag{20}$$

where for Baskes and Horstemeyer V_B is in terms of the indenter displaced volume. For comparison, we report yield strength as estimated from hardness ($\sigma_{ys}\sim H/3$) normalized on Young's modulus E. From Eq. (19), this combined with (20) gives

$$H = \frac{\sigma_{ys}^0}{\left(\frac{S}{V}\right)_G^{2/3}} \cdot \frac{1}{\left(6\left[\frac{V}{S}\right]_B R\right)^{1/3}}. \tag{21a}$$

Here, we distinguish the bulk yield strength, σ_{ys}^0, from the yield strength, σ_{ys}, that would be obtained from $H/3$, thereby involving the ISE. Taking $H/3$ as σ_{ys} and normalizing on modulus gives

$$\frac{\sigma_{ys}}{E} = \frac{\sigma_{ys}^0}{3E\left(\frac{S}{V}\right)_G^{2/3}} \cdot \frac{1}{\left(6R\left[\frac{V}{S}\right]_B\right)^{1/3}}. \tag{21b}$$

Note for a given material that σ_{ys}^0/E is constant and $[S/V]_G$ is approximately constant at shallow depths (Fig. 6). The result for a constant tip shape is that σ_{ys}/E should be proportional to $[V/S]^{-0.33}$ whereas Horstemeyer and Baskes originally reported an exponent of -0.38. This is further remarkable since in reexamining data left out of the original analysis, those data shown along with a new least-squares fit give a slope equal to -0.33 in near perfect agreement with the $-1/3$ slope predicted by (21*b*). As is seen in Fig. 9, there is now a reasonable qualitative agreement between the two sets of data.

Summary and Conclusions

From examination of average strains and strain gradients as a function of indentation depth, this reinforces previous views that gradient plasticity models do not apply at very shallow depths. An alternative model for depths up to several hundred nanometers is

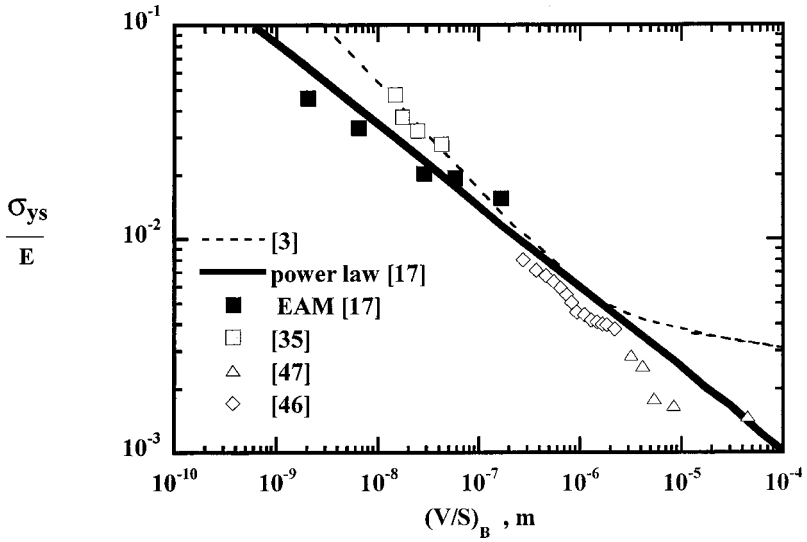


Fig. 9 Incorporation of the Baskes and Horstemeyer definition of volume to surface ratio, $(V/S)_B$, for comparison of atomistic simulations to the present data: solid line is a power-law fit with -0.38 slope similar to Eq. (21*b*)

proposed in terms of a surface work and plastic volume work concept. The critical feature is the $(S/V)_G$ where S is the projected contact area and V_G is the plastic volume under the contact. It is further concluded that

1 The functional dependence of both average strain and average strain gradient on indentation depth is even qualitatively different for spherical versus sharp wedge indenters.

2 At very shallow depths average strain gradients increase with increasing indentation depth for all spherical tip radii and even for larger spherical tip radii at depths deeper than 100 nm.

3 Assessment of surface and volume works demonstrate that both of these can be fundamentally important at shallow depths of indentation where size scales are small.

4 For shallow depths of indentation the $(S/V)_G$ is nearly independent of depth.

5 The proposed model based upon a constant $(S/V)_G$ model predicts normalized hardness to decrease as $(\delta R)^{-1/3}$. This fits $\langle 100 \rangle$ single crystal data for Au, Al, Fe-3wt%Si, and W.

Acknowledgments

We greatly appreciate use of a high-load transducer and the aluminum single crystal provided by Hysitron, Inc. Research support from NSF-DMI-99871863 for two of us (NIT and WWG) from a Graduate School Fellowship for one of us (JCG) and from a grant by Seagate Technology through MINT for one of us (WWG) is gratefully acknowledged.

Appendix

Surface Stress as a Work Term. Using the Laplace pressure to find the effect of surface stress on the lattice parameter change in nanometer scale particles, Maugis [16] finds

$$\sigma_s = \frac{R}{2}\Delta p = -\frac{R}{2\chi_c}\frac{\Delta V}{V} = -\frac{3}{2}\frac{R}{\chi_c}\frac{\Delta a}{a} \qquad (A.1)$$

where the pressure change is interpreted in terms of the relative volume change, $\Delta V/V$, and the compressibility, χ_c. Defining $\Delta a/a$ as a lattice strain, ε, and noting that the bulk modulus, K, is the inverse of compressibility, this gives the surface stress as

$$\sigma_s = \frac{3}{2}KR\varepsilon. \qquad (A.2)$$

By measuring lattice strains via electron diffraction in small spheres, Vermaak et al. [34–36] determined σ_s in Au, Ag, and Pt to range from about 1.2 to 2.6 J/m^2.

References

[1] Fleck, N. A., and Hutchinson, J. W., 1993, "A Phenomenological Theory for Strain Gradient Effects in Plasticity," J. Mech. Phys. Solids, **41**, pp. 1825–1857.
[2] Ma, Q., and Clarke, D. R., 1995, "Size Dependent Hardness of Silver Single Crystals," J. Mater. Res., **10**(4), pp. 853–863.
[3] Nix, W. D., and Gao, H., 1998, "Indentation Size Effects in Crystalline Materials: A Law for Strain Gradient Plasticity," J. Mech. Phys. Solids, **46**(3), pp. 411–425.
[4] Aifantis, E. C., 1984, "On the Microstructural Origin of Certain Inelastic Modes," ASME J. Eng. Mater. Technol., **106**, pp. 326–330.
[5] Zbib, H., and Aifantis, E. C., 1989, "On the Localization and Post Localization of Plastic Deformation—Part I: On the Initiation of Shear Bands," Res. Mech., pp. 261–277.
[6] Zbib, H., and Aifantis, E. C., 1989, "On the Localization and Post Localization of Plastic Deformation—Part II: On the Evolution and Thickness of Shear Bands," Res. Mech., pp. 279–292.
[7] Zbib, H., and Aifantis, E. C., 1989, "On the Localization and Post Localization of Plastic Deformation—Part III: On the Structure and Velocity of Postevin–Le Chatelier Bands," Res. Mech., pp. 293–305.
[8] Shu, J. Y., and Fleck, N. A., 1998, "The Prediction of a Size Effect in Microindentation," Int. J. Solids Struct., **35**(13), pp. 1363–1383.
[9] Gao, H., Huang, Y., Nix, W. D., and Hutchinson, J. W., 1999, "Mechanism-Based Strain Gradient Plasticity—I. Theory," J. Mech. Phys. Solids, **47**, pp. 1239–1263.
[10] Hutchinson, J. W., 2000, "Plasticity at the Micron Scale," Int. J. Solids Struct., **37**, pp. 225–238.
[11] Ashby, M. F., 1970, "The Deformation of Plastically Non-Homogeneous Alloys," Philos. Mag., **21**, pp. 399–424.
[12] Poole, W. J., Ashby, M. F., and Fleck, N. A., 1996, "Micro-Hardness Tests on Annealed and Work-Hardened Copper Polycrystals," Scr. Metall., **34**(4), pp. 559–564.
[13] Wahl, K., and Asif, S. A., 2000, Naval Research Laboratories, personal communication.
[14] Johnson, K. L., Kendall, K., and Roberts, A. D., 1971, "Surface Energy and the Contact of Elastic Solids," Proc. R. Soc. London, Ser. A, **A321**, pp. 301–313.
[15] Derjaguin, B. V., Muller, V. M., and Toporov, Yu. P., 1975, "Effect of Contact Deformations on the Adhesion of Particles," J. Colloid Interface Sci., **53**, pp. 314–326.
[16] Maugis, D., 1999, "Contact, Adhesion and Rupture of Elastic Solids," *Series in Solid State Sciences*, Springer, New York, pp. 62–66, 283–295.
[17] Tymiak, N. I., Kramer, D. E., Bahr, D. F., and Gerberich, W. W., 2001, "Plastic Strain and Strain Gradients at Very Small Penetration Depths," Acta Mater., **49**, pp. 1021–1034.
[18] Baskes, M., and Horstemeyer, M., 1999, Sandia National Labs, private communication.
[19] Horstemeyer, M. F., and Baskes, M. I., 1999, "Atomistic Finite Deformation Simulations: A Discussion on Length Scale Effects in Relation to Mechanical Stresses," ASME J. Eng. Mater. Technol., **121**, pp. 114–119.
[20] Gane, M., and Cox, J. M., 1970, "The Micro-Hardness of Metals at Very Low Loads," Philos. Mag., **22**(179), pp. 881–891.
[21] Oliver, W. C., and Pharr, G. M., 1992, "An Improved Technique for Determining Hardness and Elastic Modulus Using Load and Displacement Sensing Experiments," J. Mater. Res., **7**(6), pp. 1564–1583.
[22] Stelmashenko, N. A., Walls, M. G., Brown, L. M., and Milman, Yu. V., 1993, "Microindentations on W and Mo Oriented Single Crystals: An STM Study," Acta Metall. Mater., **41**(1), pp. 2855–2865.
[23] Bahr, D. F., Kramer, D. E., and Gerberich, W. W., 1998, "Non-Linear Deformation Mechanisms During Nanoindentation," Acta Mater., **46**(10), pp. 3605–3617.
[24] Gerberich, W. W., Kramer, D. E., Tymiak, N. I., Volinsky, A. A., Bahr, D. F., and Kriese, M. D., 1999, "Nanoindentation-Induced Defect-Interface Interactions: Phenomena, Methods and Limitations," Acta Mater., **47**(15), pp. 4115–4123.
[25] Bobji, M. S., Biswas, S. K., and Pethica, J. B., 1997, "Effect of Roughness on the Measurement of Nanohardness—A Computer Simulation Study," Appl. Phys. Lett., **71**(8), pp. 1059–1061.
[26] Gerberich, W. W., Yu, W., Kramer, D., Strojny, A., Bahr, D., Lilleodden, E., and Nelson, J., 1998, "Elastic Loading and Elastoplastic Unloading from Nanometer Level Indentations for Modulus Determinations," J. Mater. Res., **13**, pp. 421–439.
[27] Begley, M. R., and Hutchinson, J. W., 1998, "The Mechanics of Size-Dependent Indentation," J. Mech. Phys. Solids, **46**, pp. 2049–2068.
[28] Belak, J., and Stowers, I. F., 1992, "The Indentation and Scraping of a Metal Surface: A Molecular Dynamics Study," *Fundamentals of Friction*, I. L. Singer and H. M. Pollock, eds., Kluwer Academic, Dordrecht, pp. 511–520.
[29] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge Univ., Cambridge, UK, Press, pp. 153–184.
[30] Kramer, D., Huang, H., Kriese, M., Robach, J., Nelson, J., Wright, A., Bahr, D., and Gerberich, W. W., 1999, "Yield Strength Predictions from the Plastic Zone Around Nanocontacts," Acta Mater., **47**(1), pp. 333–343.
[31] Binnig, G., Quate, C. F., and Gerber, Ch., 1987, "Atomic Force Microscope," Phys. Rev. Lett., **56**, pp. 930–933.
[32] Kheshgi, H. S., and Scriven, L. E., 1991, "Dewetting, Nucleation and Growth of Dry Regions," Chem. Eng. Sci., **46**, pp. 519–526.
[33] Josell, D., and Spaepen, F., 1993, "Determination of the Interfacial Tension by Zero Creep Experiments on Multilayers," Acta Metall., **41**, pp. 3015–3027.
[34] Vermaak, J. S., and Kuhlman-Wilsdorf, D., 1968, "Measurement of the Average Surface Stress of Gold as a Function of Temperature in the Temperature Range 50–985°," J. Phys. Chem., **72**, pp. 4150–4154.
[35] Mays, C. W., Vermaak, J. S., and Kuhlman-Wilsdorf, D., 1968, "Surface Stress and Surface Tension. II. Determination of the Surface Stress of Gold," Surf. Sci., **12**, pp. 134–140.
[36] Wasserman, H. J., and Vermaak, J. S., 1970, "Determination of a Lattice Contraction in Very Small Silver Particles," Surf. Sci., **22**, pp. 164–172.
[37] Friesen, C., Dimitrov, N., Cammarata, R. C., and Sieradzki, K., 2000, "Surface Stress and the Electrocapilarity of Solid Electrodes," Surf. Sci., submitted for publication.
[38] Gerberich, W. W., Venkataraman, S. K., Huang, H., Harvey, S. E., and Kohlstedt, D. L., 1995, "The Injection of Plasticity by Millinewton Contacts," Acta Metall. Mater., **43**(4), pp. 1569–1576.
[39] Michalske, T. A., and Houston, J. E., 1998, "Dislocation Nucleation at Nano-Scale Mechanical Contacts," Acta Mater., **46**(2), pp. 391–396.
[40] Gerberich, W. W., Nelson, J. C., Lilleodden, E. T., Anderson, P., and Wyrobek, J. T., 1996, "Indentation Induced Dislocation Nucleation: The Initial Yield Point," Acta Mater., **44**(9), pp. 3585–3598.
[41] Fleck, N. A., Muller, G. M., Ashby, M. F., and Hutchinson, J. W., 1994, Acta Metall. Mater., **42**(2), pp. 475–487.
[42] Gouldstone, A., Koh, H.-J., Zeng, K.-Y., Giannakopolous, A. E., and Suresh, S., 2000, "Discrete and Continuous Deformation During Nanoindentation of Thin Films," Acta Mater., **48**, pp. 2277–2295.

[43] Yasuda, K., Shinohara, K., Kinoshita, C. and Arai, M., 1994, ''An Interpretation of the Indentation Size/Load Effect on Diamond Pyramid Hardness,'' *Strength of Materials*, Oikawa et al., eds., The Japan Institute of Metal, pp. 865–868.
[44] Corcoran, S. G., Colton, R. J., Lilleodden, E. T., and Gerberich, W. W., 1997, ''Anomalous Plastic Deformation of Surfaces: Nanoindentation of Gold Single Crystals,'' Phys. Rev. B, **55**(24), pp. 16057–16060.
[45] Cheng, L., 1996, ''Numerical Modeling of Indentation and Scratch Problems,'' Ph.D. Thesis, University of Minnesota.
[46] Couchman, P. R., Jesser, W. A., Kuhlmann-Wilsdorf, D., and Hirth, J. P., 1972, Surf. Sci., **33**, pp. 429–436.
[47] Kramer, D. E., Yoder, K. B., and Gerberich, W. W., 2001, ''Surface Constrained Plasticity: Oxide Rupture and the Yield Point Process,'' Philos. Mag. A, **81**(8), pp. 2033–2058.
[48] Cheng, Y.-T., and Cheng, C.-M., 1998, ''Relationships Between Hardness, Elastic Modulus, and the Work of Indentation,'' Appl. Phys. Lett., **73**(5), pp. 614–616.
[49] Hirth, J., and Loethe, J., 1982, *Theory of Dislocations*, 2nd Ed., John Wiley and Sons, pp. 837–839.

R. Kukta
Assistant Professor,
Department of Mechanical Engineering,
Assoc. Mem. ASME

A. Peralta
Assistant Professor,
Department of Mechanical Engineering,
Assoc. Mem. ASME

State University of New York at Stony Brook,
Stony Brook, NY 11794

D. Kouris
Professor,
Department of Mechanical Engineering,
University of Wyoming,
Laramie, WY 82071-3295
Mem. ASME

Surface Steps: From Atomistics to Continuum

The elastic interaction between surface steps biases growth and can have a dominant influence on growth modes of thin films. A new continuum solution is presented that takes into account the actual geometry of the steps. It is shown that even in the absence of external stress, the interaction energy contains a logarithmic dependence not previously reported in the literature. This additional term is of the same order of magnitude as the ones contained in the solution of Marchenko and Parshin that has been generally used during the past twenty years. Opposite and similar steps are investigated and closed-form expressions for the relevant elastic fields and interaction energies are derived. Results compare favorably with the ones obtained from embedded atom method (EAM) simulations and a recent elastic lattice model. [DOI: 10.1115/1.1468999]

1 Introduction

Until ten years ago, growth modes of thin films have been characterized solely on the basis of thermodynamic considerations. Consequently, two or three-dimensional growth was selected as a result of the relative magnitudes of three interfacial free energies: the surface free energy of the film, the surface free energy of the substrate, and the interfacial free energy between the growing film and the substrate. We now know that such an approach is often inaccurate (for details and references, please see Kouris et al. [1]).

Currently, it is well understood that during the growth of crystal surfaces, three-dimensional structures can develop due to the elastic interaction of surface defects. These defect interactions determine the kinetics of growth processes and often control the growth mode morphology. They decay slowly with defect separation and have longer range than the indirect electronic interactions (Stoneham [2]).

The geometric discontinuity at a surface step (Fig. 1) is responsible for a discontinuity of the surface stress, resulting in a force couple with a magnitude proportional to the surface stress. Marchenko and Parshin [3], in their continuum model of a surface step, proposed an equivalency between the couple mentioned above and two force dipoles (one with and another without moment), acting along the line of the step (Fig. 1). The ensuing elastic analysis for an isotropic substrate predicted that the interaction energy between two identical steps is inversely proportional to the square of the separation distance. Other relevant studies include the work of Stewart et al. [4] and Duport et al. [5].

In a recent paper, Kukta and Bhattacharya [6] determined the elastic field of a step up to second order. Their model demonstrated that second-order fields are necessary for obtaining accurate elastic fields and cannot be described simply by point sources on a half-plane.

In the present study, we determine the elastic field of interacting steps. Airy stress functions are derived for the cases of opposite and similar steps. Opposite steps are illustrated in Figs. 2(*a*, *b*) that correspond to a mesa, or a surface depression (pit). Similar steps (Fig. 2(*c*)) constitute the element of a vicinal surface.[1] In addition, expressions for the displacements and interaction energies are presented. In the case of opposite steps, the interaction energy includes a logarithmic dependence unlike the ones previously reported (Tersoff et al. [7] and Alerhand et al. [8]). Our solution proves that a logarithmic term is present even in unstrained (no external stress or lattice misfit stress) homo-epitaxial films. The results are compared with lattice solutions and embedded atom model (EAM) simulations previously developed by Kouris et al. [1,9].

The elastic field of surface steps is reconsidered in the first part of the discussion that follows. After a brief reference to the well-known point source model, Chapter 2 describes the Airy stress function of two interacting steps. In Chapter 3, the interaction energy between steps is calculated and compared with the Marchenko and Parshin [3] model. The last two chapters of the article include results for the elastic field and interaction energy, as well as some relevant conclusions.

[1]Substrates with orientations slightly off the crystallographic plane.

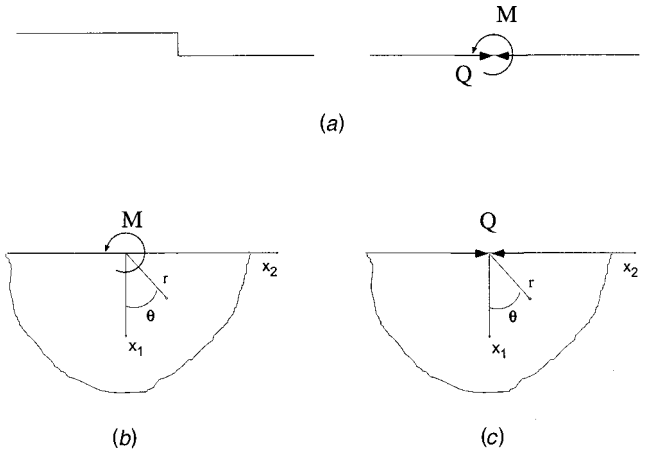


Fig. 1 (*a*) A straight step and the mechanically equivalent flat surface of [3], (*b*) a concentrated couple *M* acting at the free surface, (*c*) dipole of strength *Q* acting at the free surface

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, March 15, 2001; final revision, December 15, 2001. Editor: L. T. Wheeler. Discussion on the paper should be addressed to the Editor, Professor Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

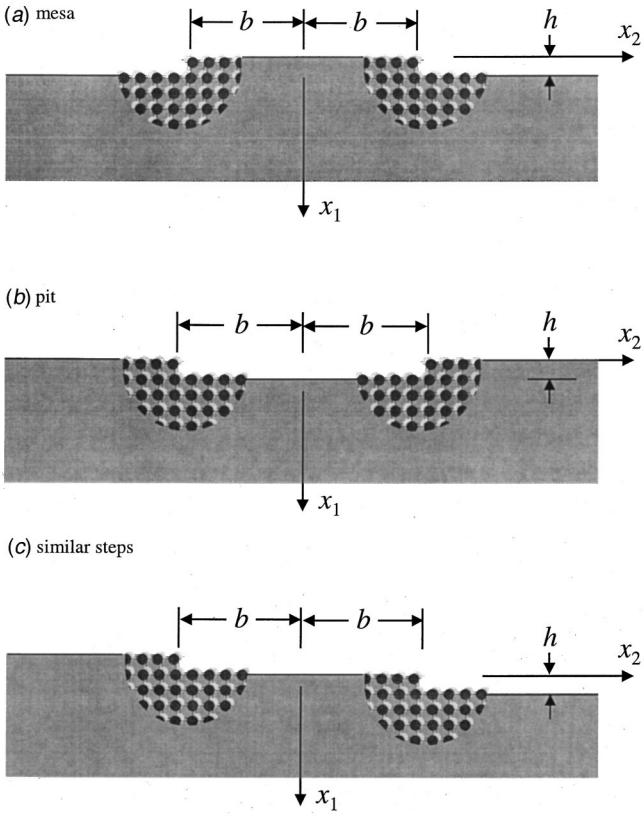


Fig. 2 (*a*) **Two opposite steps forming a mesa,** (*b*) **two opposite steps forming a pit,** (*c*) **two similar steps**

2 Elastic Field of Similar and Opposite Steps

2.1 Preliminary Considerations. Figure 1 illustrates the cross section of a surface step and the corresponding elastic model proposed by Marchenko and Parshin [3]. A straight step is infinite in the x_3-direction resulting in a two-dimensional (plane-strain) problem. Their continuum model involves a mechanically equivalent flat surface with two force dipoles acting along the line of the step. The first dipole is equivalent to a moment M, proportional to the surface stress,

$$M = fh \tag{1}$$

where f is the surface stress[2] and h denotes the height of the step. The strength Q of the dipole without a moment is generally unknown. Details of the resulting elastic field can be found in Kouris et al. [9].

2.2 A New Continuum Solution. By utilizing the solution for the elastic field of a single surface step provided by Kukta and Bhattacharya [6], one can construct the Airy stress functions for opposite (Figs. 2(*a*) and 2(*b*)) and similar steps (Fig. 2(*c*)). The term ''mesa'' will be used to describe the terrace of finite width illustrated in Fig. 2(*a*). The geometry of Fig. 2(*b*) indicates a surface depression and will be referred to as ''pit.''

The Airy stress function for the mesa (Fig. 3(*a*)) is given by

$$\begin{aligned} U = \frac{hf}{\pi}\Bigg\{ & \frac{x_1(x_2+b)}{r_2^2} - \frac{x_1(x_2-b)}{r_1^2} - \tan^{-1}\left(\frac{x_1}{x_2+b}\right) \\ & + \tan^{-1}\left(\frac{x_1}{x_2-b}\right)\Bigg\} - \frac{h\omega x_1^2}{\pi}\left\{\frac{1}{r_1^2}+\frac{1}{r_2^2}\right\} \\ & - \frac{2h^2\omega x_1}{\pi^2}\left\{\frac{x_1^2-(x_2-b)^2}{r_1^4}+\frac{x_1^2-(x_2+b)^2}{r_2^4}\right\} \\ & \times\left\{\pi+\tan^{-1}\left(\frac{x_2-b}{x_1}\right)-\tan^{-1}\left(\frac{x_2+b}{x_1}\right)\right\} + \frac{4h^2\omega x_1^2}{\pi^2}\left\{\frac{x_2-b}{r_1^4}\right. \\ & \left. + \frac{x_2+b}{r_2^4}\right\}\ln\left(\frac{r_1}{r_2}\right) - \frac{4h^2\omega x_1^2}{\pi^2}\left\{\frac{x_2-b}{r_1^4}-\frac{x_2+b}{r_2^4}\right\}\ln\left(\frac{\alpha h}{2b}\right) \end{aligned} \tag{2}$$

[2]A comprehensive discussion on the issue of surface stress can be found in Cammarata [10].

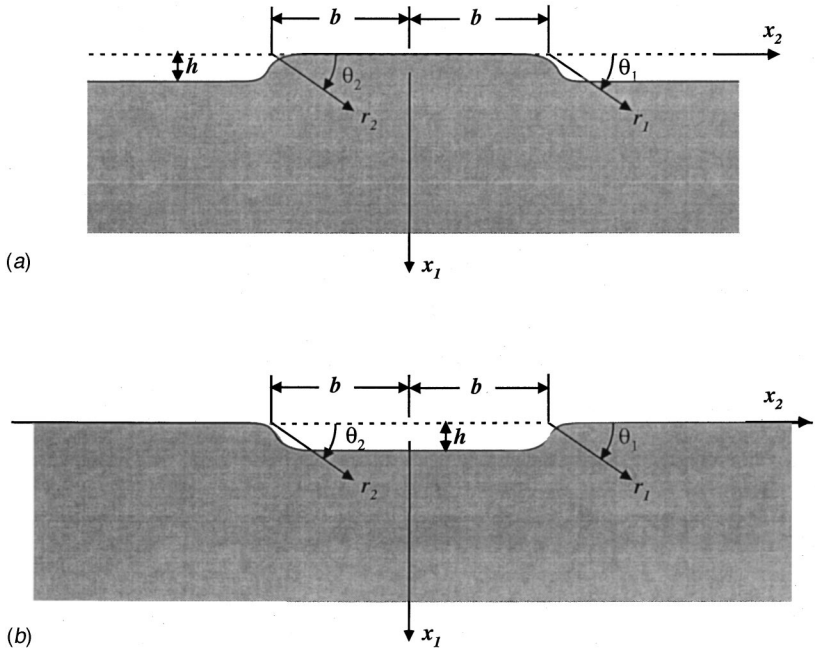


Fig. 3 (*a*) **The geometry and coordinate system of a mesa,** (*b*) **the geometry and coordinate system of a pit.**

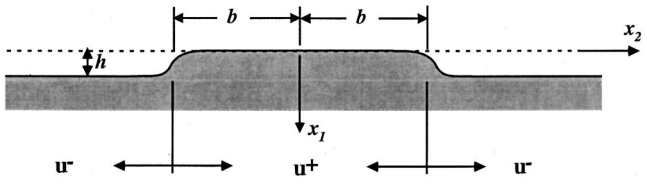


Fig. 4 Notation used in the expressions for the displacements

where $r_1=[x_1^2+(x_2-b)^2]^{1/2}$, $r_2=[x_1^2+(x_2+b)^2]^{1/2}$, α is a dimensionless constant and ω is a "strength" term defined through $Q=\omega h$. The Airy stress functions for the pit (Fig. 3(*b*)) and the similar steps are constructed using the same procedure. The corresponding expressions are given in the Appendix. Stresses and displacements are determined using the Airy stress functions. The expressions of the displacements have been obtained with respect to the geometry of Fig. 3 and the notation of Fig. 4 and are also provided in the Appendix.

3 Interaction Energy Between Steps

The point source model of Marchenko and Parshin [3] yields an expression for the interaction energy between steps which is equivalent to

$$E_{\text{int}}=\frac{1-\nu}{\pi\mu}[(\omega h)^2\pm(fh)^2]\frac{1}{d^2}, \tag{3}$$

where d is the distance between the two steps. The positive sign in the quantity $[(\omega h)^2\pm(fh)^2]$ corresponds to similar, while the negative sign refers to opposite steps. According to this solution, steps of similar sign (Fig. 2(*c*)) are repelled since $E_{\text{int}}>0$, while opposite steps (Figs. 2(*a*), (*b*)) may attract or repel each other, depending on the sign of the quantity $[(\omega h)^2-(fh)^2]$.

The solution introduced in Chapter 2 yields an interaction energy for two steps forming a mesa that is complete:

$$E_{\text{int}}^{\text{mesa}}=\frac{(\omega^2-f^2)(1-\nu)}{\pi\mu}\left(\frac{h}{d}\right)^2-\frac{8\omega^2(1-\nu)}{\pi^2\mu}\left(\frac{h}{d}\right)^3\ln\left(\frac{\alpha h}{d}\right)-\frac{2\omega^2(1-\nu)}{\pi^2\mu}\left(\frac{h}{d}\right)^3+O\left(\frac{h^4}{d^4}\ln^2\frac{h}{d}\right). \tag{4}$$

The first term is identical to (3). This expression can be rewritten in a form that allows for the direct comparison of the terms multiplying the "strength" ω:

$$E_{\text{int}}^{\text{mesa}}=-\frac{f^2(1-\nu)}{\pi\mu}\left(\frac{h}{d}\right)^2+\frac{\omega^2(1-\nu)}{\pi\mu}\left(\frac{h}{d}\right)^2\left\{1-\frac{2}{\pi}\left(\frac{h}{d}\right)-\frac{8}{\pi}\left(\frac{h}{d}\right)\ln\left(\frac{\alpha h}{d}\right)\right\}+O\left(\frac{h^4}{d^4}\ln^2\frac{h}{d}\right). \tag{5}$$

Similarly, the interaction energy of two steps forming a pit is

$$E_{\text{int}}^{\text{pit}}=\frac{(\omega^2-f^2)(1-\nu)}{\pi\mu}\left(\frac{h}{d}\right)^2+\frac{8\omega^2(1-\nu)}{\pi^2\mu}\left(\frac{h}{d}\right)^3\ln\left(\frac{\alpha h}{d}\right)+\frac{2\omega^2(1-\nu)}{\pi^2\mu}\left(\frac{h}{d}\right)^3+O\left(\frac{h^4}{d^4}\ln^2\frac{h}{d}\right) \tag{6}$$

or in a condensed form

$$E_{\text{int}}^{\text{pit}}=-\frac{f^2(1-\nu)}{\pi\mu}\left(\frac{h}{d}\right)^2+\frac{\omega^2(1-\nu)}{\pi\mu}\left(\frac{h}{d}\right)^2\left\{1+\frac{2}{\pi}\left(\frac{h}{d}\right)+\frac{8}{\pi}\left(\frac{h}{d}\right)\ln\left(\frac{\alpha h}{d}\right)\right\}+O\left(\frac{h^4}{d^4}\ln^2\frac{h}{d}\right). \tag{7}$$

It is important to realize that if one compares the magnitude of the bracketed terms that include the logarithms in Eqs. (5) and (7), one can appreciate the significant contribution of the new term, not only for the near-field but also for the intermediate range of the ratio (h/d).

In the case of similar steps of identical step height, the interaction energy is

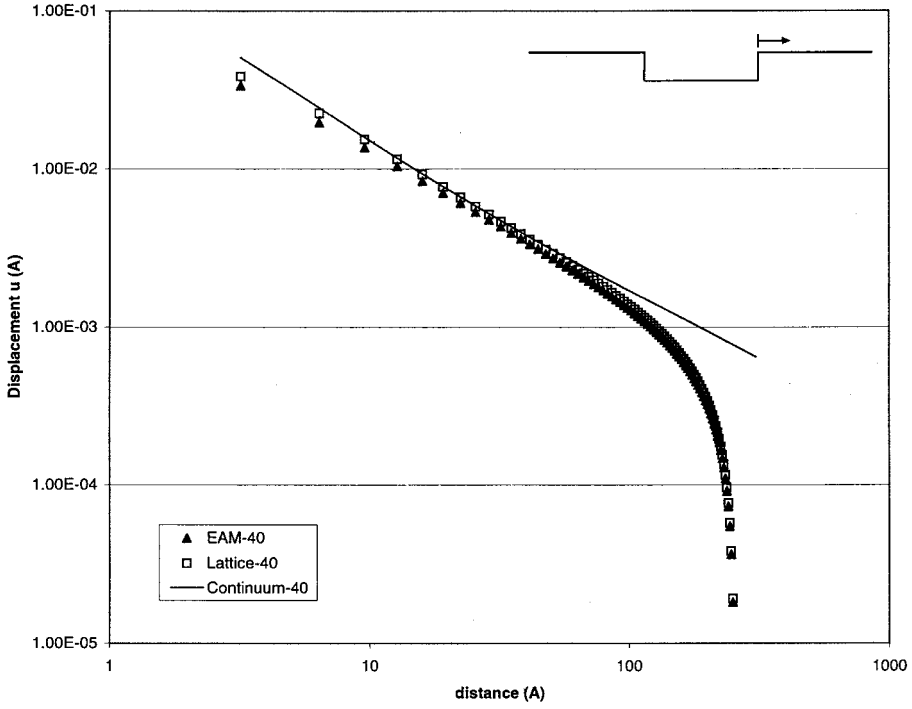


Fig. 5 Displacement u_2 along the top terrace of a pit, as a function of distance

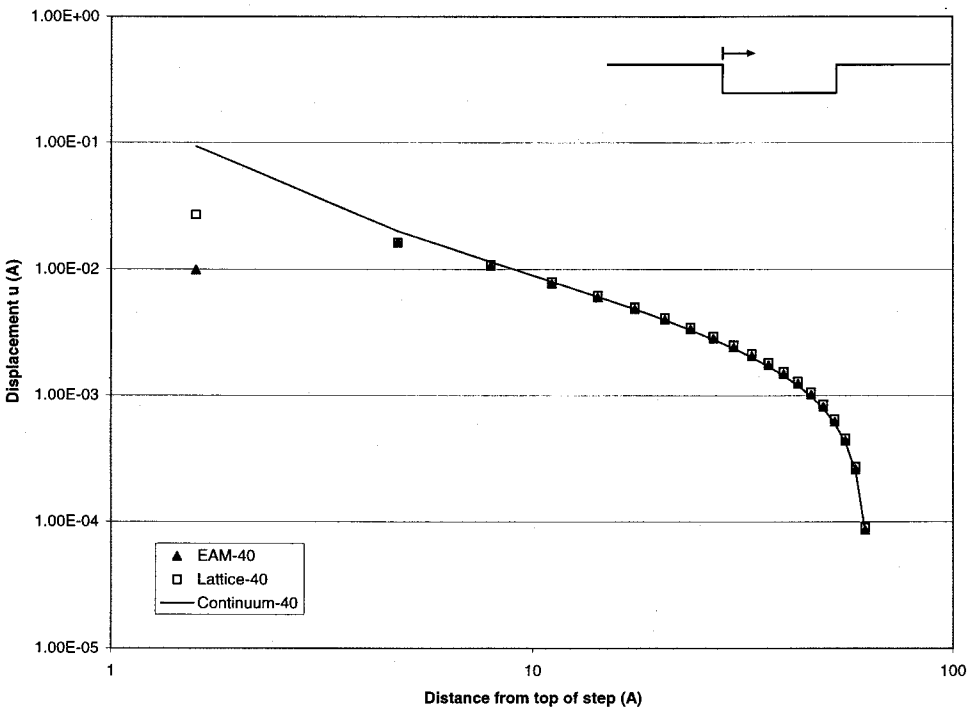


Fig. 6 Displacement u_2 along the bottom terrace of a pit, as a function of distance

$$E_{\text{int}}^{\text{similar}} = \frac{(\omega^2 + f^2)(1-\nu)}{\pi\mu}\left(\frac{h}{d}\right)^2 + O\left(\frac{h^4}{d^4}\ln^2\frac{h}{d}\right) \tag{8}$$

If the similar steps have different step heights, a logarithmic term similar to the one above, is introduced to Eq. (8).

4 Results and Discussion

A simple comparison between the Airy stress functions illustrates that Eq. (2) and the point source model presented in [3] share two common terms. The higher order terms in (2), however, yield stresses and displacements that are significant.

In a previous paper ([9]), we had observed that the surface displacement near a step exhibits loss of symmetry between the upper and lower terraces. Kukta and Bhattacharya [6] illustrated the origin of this phenomenon. We observe a similar behavior in the case of a mesa or a pit. Figures 5 and 6 illustrate the surface displacement u_2 at the top and bottom of a pit, respectively, as a function of the distance from the top of a step. The size of the pit

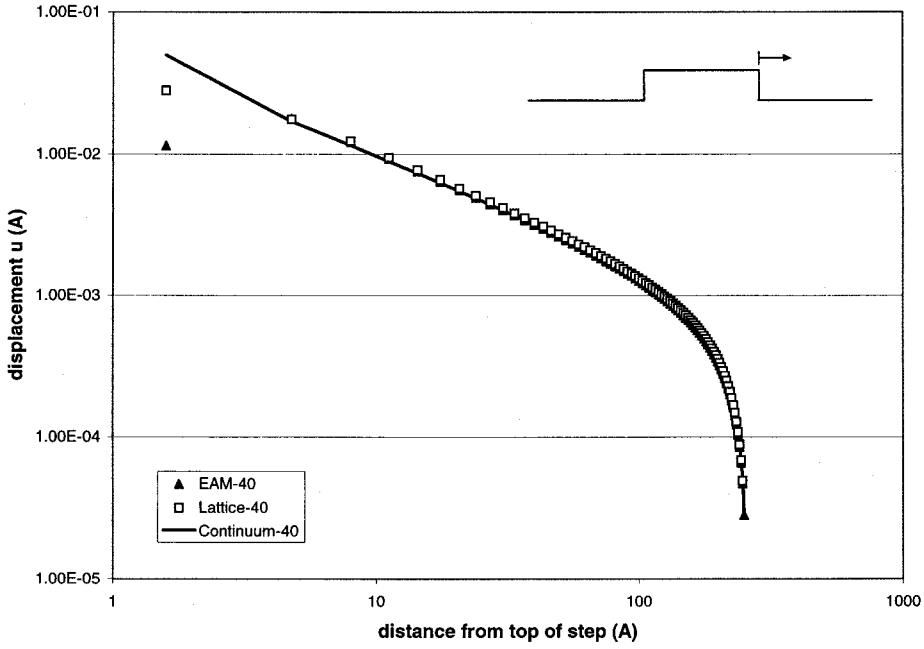


Fig. 7 Displacement u_2 along the bottom terrace of a mesa, as a function of distance

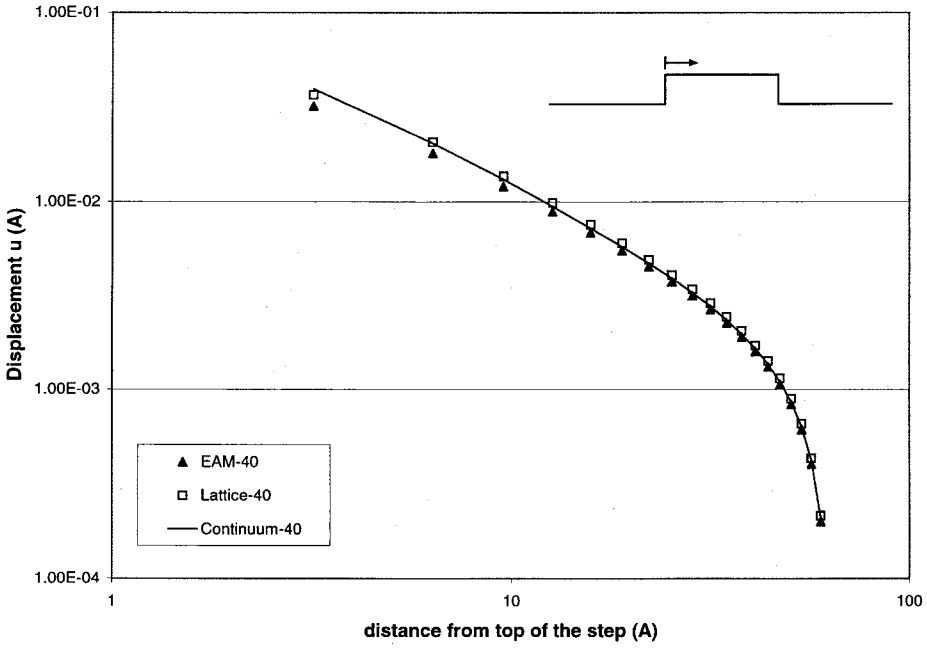


Fig. 8 Displacement u_2 along the top terrace of a mesa, as a function of distance

is equal to 40 lattice spaces and the continuum solution expressed in Eq. (*A*8) is compared with results from EAM simulations and our discrete lattice model. Values of $T=0.0432$ Angstroms and $\alpha=1$ were used. They correspond to a surface stress $f=3.032$ J/m^2, a step height $h=1.5826$ Angstroms, a shear modulus $\mu=160.6$ GPa, and a Poisson's ratio $\nu=0.28$ for tungsten. Ω was determined by comparing the continuum expression for u_2 with the lattice solution; $\Omega=0.073$ Angstroms (T and Ω are used in the displacement expressions and are defined in the Appendix). Figure 6 illustrates the clear deviation from the $1/r$ decay one would expect on the basis of the point source model in [3]. Similar behavior is observed in the case of the mesa (Figs. 7 and 8). It is worth noting that the agreement with EAM simulations is remarkable.

When comparing Eq. (3) of Marchenko and Parshin [3] with the interaction energies (4)–(8), one observes that the new term with the logarithmic dependence changes the $1/d^2$ decay rate. In addition, the magnitude of this term is comparable to the ones proportional to $1/d^2$ and cannot be ignored. It is substantial for both the short and intermediate range of h/d. The interaction energy of two steps forming a mesa, as a function of mesa size, is illustrated in Fig. 9. The continuum solution is compared with the lattice and EAM results. A careful atomistic study of Shilkrot and

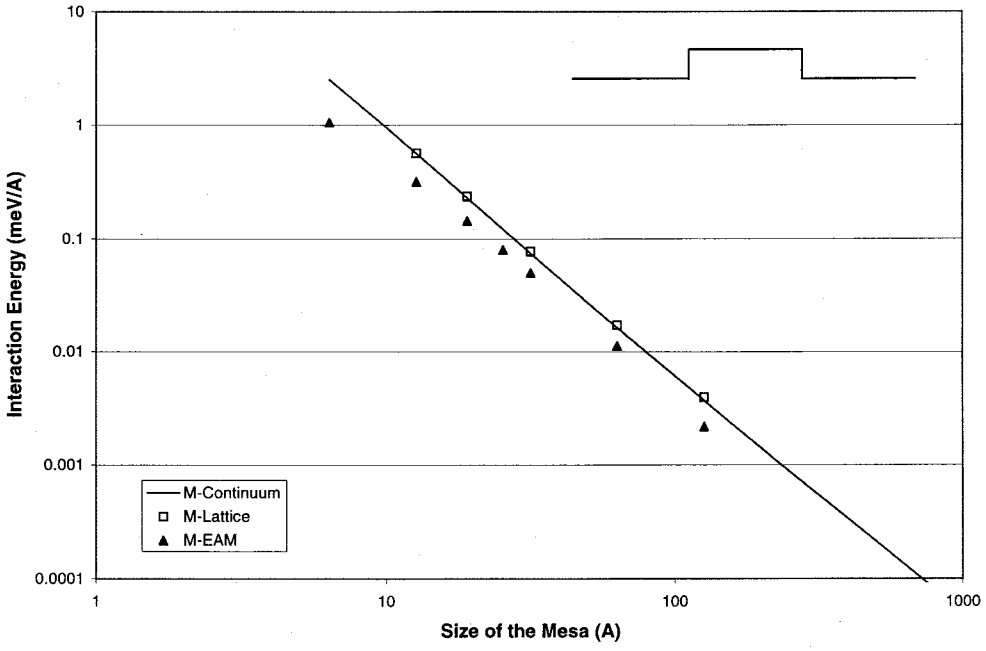


Fig. 9 Interaction energy between two steps forming a mesa, as a function of the mesa size

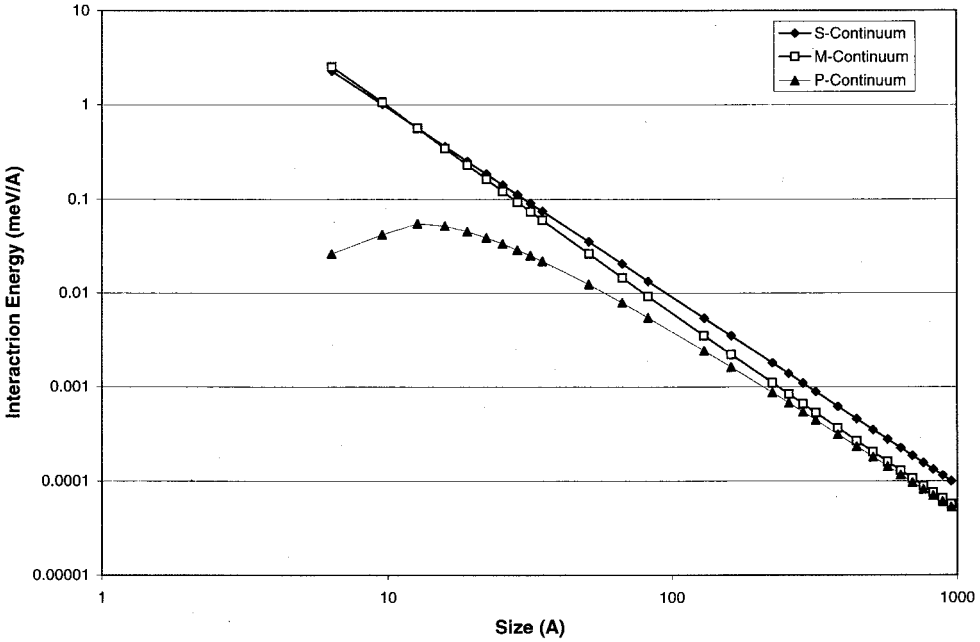


Fig. 10 Interaction energy as a function of size for similar steps, a mesa and a pit

Srolovitz [11] did not identify the correct dependence but attributed the variation from the model in [3] to terms of order $1/d^3$.

Using Eqs. (4), (6), and (8), one can plot the interaction energy as a function of size, for similar steps, a mesa and a pit. Figure 10 illustrates the results; the differences between a mesa and a pit are quite clear. Similar steps follow the expected $1/d^2$ decay. A comparison between the mesa and the pit reveals a sign difference in the logarithm. This was observed before in simulations but could not be explained by the Marchenko and Parshin model.

If one determines the magnitude of the term $2/\pi(h/d)+8/\pi(h/d)\ln(\alpha h/d)$ appearing in Eqs. (4)–(7), for realistic values of h/d, one finds that it is always negative with an absolute value less than unity (but not much less, for moderate values of the ratio!). This ensures that the bracketed term in Eq. (7) remains positive but could be significantly ''weakened'' when compared to the corresponding contribution in (5). The implications for the overall sign of the interaction energy and the corresponding repulsion-attraction between steps are significant.

Finally, an observation must be made regarding the interaction energy of similar steps (Fig. 2(*c*)). In this case, Eqs. (3) and (8) are identical and predict repulsion (at least in the isotropic case discussed herein). However, if the heights of the steps are different, a logarithmic term similar to the ones in (6) and (8) is introduced. This yields some interesting results when step-bunching issues are involved. We are currently examining this topic and we plan to discuss it in a future communication.

5 Conclusions

A closed-form solution for the isotropic elastic field of interacting surface steps has been presented. The interaction energies for similar and opposite sign steps reveal the presence of a logarithmic dependence, not previously reported. This term is *always* present in the case of opposite sign steps and can control the sign of the interaction energy. Unlike similar terms discussed in the literature, this contribution to the energy does not appear due to broken orientational symmetry or the presence of an external strain; it does not involve a force monopole. It is simply present because of the geometric nature of the steps. In the case of similar steps, the logarithmic term appears when the step heights are not equal.

This solution has significant implications and may require the re-examination of some well-established results in the area of thin film growth.

Acknowledgments

This work was partially supported by the National Science Foundation's CMS and DMR Divisions (Ken Chong and Bruce McDonald, Program Directors), under Grants CMS-9988597 and DMR-0090079.

Appendix

Airy Stress Functions. The Airy stress function for the pit (Fig. 4) is given as

$$U=\frac{hf}{\pi}\left\{\frac{x_1(x_2-b)}{r_1^2}-\frac{x_1(x_2+b)}{r_2^2}-\tan^{-1}\left(\frac{x_1}{x_2-b}\right)+\tan^{-1}\left(\frac{x_1}{x_2+b}\right)\right\}-\frac{h\omega x_1^2}{\pi}\left\{\frac{1}{r_1^2}+\frac{1}{r_2^2}\right\}+\frac{2h^2\omega x_1}{\pi^2}\left\{\frac{x_1^2-(x_2-b)^2}{r_1^4}+\frac{x_1^2-(x_2+b)^2}{r_2^4}\right\}\left\{\tan^{-1}\left(\frac{x_2-b}{x_1}\right)-\tan^{-1}\left(\frac{x_2+b}{x_1}\right)\right\}-\frac{4h^2\omega x_1^2}{\pi^2}\left\{\frac{x_2-b}{r_1^4}+\frac{x_2+b}{r_2^4}\right\}\ln\left(\frac{r_1}{r_2}\right)+\frac{4h^2\omega x_1^2}{\pi^2}\left\{\frac{x_2-b}{r_1^4}-\frac{x_2+b}{r_2^4}\right\}\ln\left(\frac{\alpha h}{2b}\right). \quad (A1)$$

The Airy stress function for the similar steps (Fig. 2(*c*)) is given by

$$U=\frac{hf}{\pi}\left\{-\frac{x_1(x_2-b)}{r_1^2}-\frac{x_1(x_2+b)}{r_2^2}+\tan^{-1}\left(\frac{x_1}{x_2-b}\right)+\tan^{-1}\left(\frac{x_1}{x_2+b}\right)\right\}-\frac{h\omega x_1^2}{\pi}\left\{\frac{1}{r_1^2}+\frac{1}{r_2^2}\right\}-\frac{2h^2\omega x_1}{\pi^2}\left\{\frac{x_1^2-(x_2-b)^2}{r_1^4}+\frac{x_1^2-(x_2+b)^2}{r_2^4}\right\}\left\{\tan^{-1}\left(\frac{x_2-b}{x_1}\right)+\tan^{-1}\left(\frac{x_2+b}{x_1}\right)\right\}-\frac{4h^2\omega x_1^2}{\pi^2}\left\{\frac{x_2-b}{r_1^4}+\frac{x_2+b}{r_2^4}\right\}\ln\left(\frac{2\alpha bh}{r_1r_2}\right). \tag{A2}$$

Displacements. The upper (+) and lower (−) terrace displacements (Fig. 4) for the mesa can be determined using Eq. (2). They are

$$u_1^{\pm}=hT\left(\frac{1}{x_2+b}-\frac{1}{x_2-b}\right)\mp\frac{h^2\Omega}{2}\left[\frac{1}{(x_2-b)^2}+\frac{1}{(x_2+b)^2}\right] \tag{A3}$$

$$u_2^{\pm}=\pm\frac{h^2T}{2}\left[\frac{1}{(x_2-b)^2}-\frac{1}{(x_2+b)^2}\right]+h\Omega\left(\frac{1}{x_2-b}+\frac{1}{x_2+b}\right)+\frac{h^2\Omega}{\pi}\left[\frac{1}{(x_2-b)^2}+\frac{1}{(x_2+b)^2}\right]\ln\left(\frac{x_2+b}{x_2-b}\right)^2+\frac{h^2\Omega}{\pi}\left[\frac{1}{(x_2-b)^2}+\frac{1}{(x_2+b)^2}\right]\ln\left(\frac{\alpha h}{2b}\right)^2 \tag{A4}$$

where

$$\Omega=\frac{\omega(1-\nu)}{\pi\mu}\ \text{ and }\ T=\frac{f(1-\nu)}{\pi\mu}.$$

The second term in (*A*3) and the third term in (*A*4) decay as $1/r^2$ from each step (quadrupole sources). While we cannot find quadrupoles that can be added to the steps to completely remove these terms, it is impossible to distinguish them from a quadrupole source. Furthermore, these two terms depend on the reference half-plane used in the solution while the other terms do not. If we neglect them, the displacements become

$$u_1=hT\left(\frac{1}{x_2+b}-\frac{1}{x_2-b}\right)+O\left[\frac{h^2}{(x_2-b)^2},\frac{h^2}{(x_2+b)^2}\right] \tag{A5}$$

$$u_2=h\Omega\left(\frac{1}{x_2-b}+\frac{1}{x_2+b}\right)+\frac{h^2\Omega}{\pi}\left[\frac{1}{(x_2-b)^2}+\frac{1}{(x_2+b)^2}\right]\ln\left(\frac{x_2+b}{x_2-b}\right)^2+\frac{h^2\Omega}{\pi}\left[\frac{1}{(x_2-b)^2}-\frac{1}{(x_2+b)^2}\right]\ln\left(\frac{\alpha h}{2b}\right)^2+O\left[\frac{h^2}{(x_2-b)^2},\frac{h^2}{(x_2+b)^2}\right]. \tag{A6}$$

In the case of the pit, the displacements are

$$u_1^{\pm}=hT\left(\frac{1}{x_2-b}-\frac{1}{x_2+b}\right)\mp\frac{h^2\Omega}{2}\left[\frac{1}{(x_2-b)^2}+\frac{1}{(x_2+b)^2}\right] \tag{A7}$$

$$u_2^{\pm}=\mp\frac{h^2T}{2}\left[\frac{1}{(x_2-b)^2}-\frac{1}{(x_2+b)^2}\right]+h\Omega\left(\frac{1}{x_2-b}+\frac{1}{x_2+b}\right)-\frac{h^2\Omega}{\pi}\left[\frac{1}{(x_2-b)^2}+\frac{1}{(x_2+b)^2}\right]\ln\left(\frac{x_2+b}{x_2-b}\right)^2-\frac{h^2\Omega}{\pi}\left[\frac{1}{(x_2-b)^2}-\frac{1}{(x_2+b)^2}\right]\ln\left(\frac{\alpha h}{2b}\right)^2. \tag{A8}$$

Dropping terms of order $(h/r)^2$ and higher we may write

$$u_1=hT\left(\frac{1}{x_2-b}-\frac{1}{x_2+b}\right) \tag{A9}$$

$$u_2=h\Omega\left(\frac{1}{x_2-b}+\frac{1}{x_2+b}\right)-\frac{h^2\Omega}{\pi}\left[\frac{1}{(x_2-b)^2}+\frac{1}{(x_2+b)^2}\right]\ln\left(\frac{x_2+b}{x_2-b}\right)^2-\frac{h^2\Omega}{\pi}\left[\frac{1}{(x_2-b)^2}-\frac{1}{(x_2+b)^2}\right]\ln\left(\frac{\alpha h}{2b}\right)^2. \tag{A10}$$

For the similar steps of Fig. 2(*c*), the displacements are simplified using similar arguments to yield

$$u_1=hT\left(\frac{1}{x_2-b}+\frac{1}{x_2+b}\right) \tag{A11}$$

$$u_2=h\Omega\left(\frac{1}{x_2-b}+\frac{1}{x_2+b}\right)+\frac{h^2\Omega}{\pi}\left[\frac{1}{(x_2-b)^2}+\frac{1}{(x_2+b)^2}\right]\ln\left[\frac{2\alpha hb}{(x_2+b)(x_2-b)}\right]^2. \tag{A12}$$

Being consistent with neglecting quadrupole sources, we may assume $\alpha=1$ in all the above expressions.

Force of Interaction. The force of interaction between steps can be determined by differentiating the interaction energy:

$$F=-\frac{\partial E_{\text{int}}}{\partial d}. \tag{A13}$$

For the mesa, the force becomes

$$F_{\text{int}}^{\text{mesa}}=-2\frac{1-\nu}{\pi\mu}f^2\frac{h^2}{d^3}+2\frac{1-\nu}{\pi\mu}\omega^2\frac{h^2}{d^3}-24\frac{1-\nu}{\pi^2\mu}\omega^2\frac{h^3}{d^4}\ln\left(\frac{h}{d}\right)+O\left(\frac{h^3}{d^4}\right). \tag{A14}$$

Similarly, for the pit

$$F_{\text{int}}^{\text{pit}}=-2\frac{1-\nu}{\pi\mu}f^2\frac{h^2}{d^3}+2\frac{1-\nu}{\pi\mu}\omega^2\frac{h^2}{d^3}+24\frac{1-\nu}{\pi^2\mu}\omega^2\frac{h^3}{d^4}\ln\left(\frac{h}{d}\right)+O\left(\frac{h^3}{d^4}\right), \tag{A15}$$

and finally for the similar steps

$$F_{\text{int}}^{\text{similar}}=+2\frac{1-\nu}{\pi\mu}f^2\frac{h^2}{d^3}+2\frac{1-\nu}{\pi\mu}\omega^2\frac{h^2}{d^3}+O\left(\frac{h^3}{d^4}\right). \tag{A16}$$

References

[1] Kouris, D., Peralta, A., and Sieradzki, K., 2000, "Surface Islands and Their Elastic Interaction With Adatoms," Surf. Sci., **445**, pp. 420–429.

[2] Stoneham, A. M., 1977, "Elastic Interactions between Surface Adatoms and Between Surface Clusters," Solid State Commun., **24**, pp. 425–428.

[3] Marchenko, V. I., and Parchin, A. Y., 1980, "Elastic Properties of Crystal Surfaces," Sov. Phys. JETP, **52**, pp. 129–131.

[4] Stewart, J., Pohland, O., and Gibson, M. J., 1994, "Elastic-Displacement Field of an Isolated Surface Step," Phys. Rev. B, **49**(19), pp. 13848–13858.

[5] Duport, C., Politi, P., and Villain, J., 1995, "Growth Instabilities Induced by Elasticity in a Vicinal Surface," J. Phys. I, France, **5**, pp. 1317–1350.

[6] Kukta, R. V., and Bhattacharya, K., 2001, "A Micromechanical Model of Surface Steps," J. Mech. Phys. Solids, **50**, pp. 615–649.

[7] Tersoff, J., Phang, Y. H., Zhang, Z., and Lagally, M. G., 1995, "Step-Bunching Instability of Vicinal Surfaces Under Stress," Phys. Rev. Lett., **75**, pp. 2730–2733.

[8] Alerhand, O. L., Vanderbilt, D., Meade, R. D., and Joannopoulos, J. D., 1988, ''Spontaneous Formation of Stress Domains on Crystal Surfaces,'' **61**(17), pp. 1973–1976.
[9] Kouris, D., Peralta, A., and Sieradzki, K., 1999, ''Elastic Interaction of Defects on Crystal Surfaces,'' J. Eng. Mater. Technol., **121**, pp. 129–135.
[10] Cammarata, R. C., 1994, ''Surface and Interface Stress Effects in Thin Films,'' Prog. Surf. Sci., **46**, pp. 1–38.
[11] Shilkrot, L. E., and Srolovitz, D. J., 1996, ''Elastic Field of a Surface Step: Atomistic Simulations and Anisotropic Elastic Theory,'' Phys. Rev. B, **53**(16), pp. 11120–11127.

S. Peng
Mem. ASME

K. Cho
Mem. ASME

Division of Mechanics and Computation,
Department of Mechanical Engineering,
Stanford University, Palo Alto, CA 94305

Nano Electro Mechanics of Semiconducting Carbon Nanotube

The effect of a flattening distortion on the electronic properties of a semiconducting carbon nanotube is investigated through first-principles calculations. As a function of the mechanical deformation, electronic bandgap is reduced leading to a semiconductor-metal transition. However, further deformation reopens the bandgap and induces a metal-semiconductor transition. The semiconductor–metal transitions take place as a result of curvature-induced hybridization effects, and this finding can be applied to develop novel nano electro mechanical systems. [DOI: 10.1115/1.1469003]

1 Introduction

A carbon nanotube (CNT) is a cylindrical shell of carbon atoms with a diameter as small as 1 nanometer and a length up to 1 to 100 micrometers. A single-wall CNT consists of only carbon atoms, and can essentially be described as a single layer of graphite sheet (graphene) wrapped into a cylinder ([1]). Nanotubes have recently received an increasing attention due to their unique physical properties (e.g., diverse electronic, chemical, and mechanical properties) ([2]), since these unique properties make them one of the most promising candidates for building blocks of molecule-scale machines and nanoelectronic devices ([3–6]). Carbon nanotubes can behave either as semiconductors or metals depending on atomic arrangement determined by chirality's vector ([7]).

Recently, it was predicted that the cross section of a carbon nanotube might be flattened based on the observation of a large volume reduction in bundles of single-wall nanotubes (SWNTs) in a high-pressure experiment ([8]). Experimentally, flattening of a nanotube can in principle be induced by an applied force of the tip of an atomic force microscope (AFM) or a scanning tunneling microscope (STM) positioned on the nanotube. A cross-section flattening can also take place naturally at the kink site of a carbon nanotube as a consequence of bending ([9,10]). A recent experiment has shown that mechanical deformation can significantly change CNT's electronic behavior ([11]). Reference [11] shows that the conductance of a metallic nanotube can be reduced by two orders of magnitude through AFM tip-induced deformations. In this work, Dai and co-workers have explained the large conductance change using sp^3 hybridization under the AFM tip. Calculations based on the extended-Hückel approach also indicate that the bending of an armchair SWNT may affect its conductance ([12]). Furthermore, first-principles calculations show that a bent semiconducting SWNT exhibits quantum dot behavior through electron localization at the kink sites ([9,13]), and that flattening of a (10, 0) nanotube induces bandgap closing ([14]). These experimental and theoretical findings illustrate an important potential of applying the electromechanical coupling of nanotubes to novel nano electro mechanical device applications.

In this paper, we explore the possibility of simulation-based engineering on how a mechanical deformation on a semiconducting SWNT leading to a cross-sectional flattening, changes its electronic structure. Specifically, as a representative semiconducting nanotube, we focus on (8, 0) SWNT, which has an LDA energy gap of 0.56 eV. We have performed first-principles calculations using the density functional theory (DFT) method, and theoretical investigations are explained in the following to answer several important fundamental questions regarding the relationship between nanotube flattening and induced electronic property change.

2 Simulation Method

In order to obtain a detailed understanding of the nanotube deformation effect, we have carried out first-principles total energy pseudo-potential calculations on flattened SWNT system using density functional theory (DFT) within the local density approximation (LDA) ([15]). The Kohn-Sham single-electron wavefunctions are expanded by 18,000 plane waves in a supercell

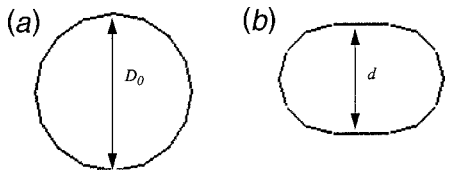


Fig. 1 Description of flatness. The degree of flattening is characterized by the parameter $\eta=(D_0-d)/D_0$.

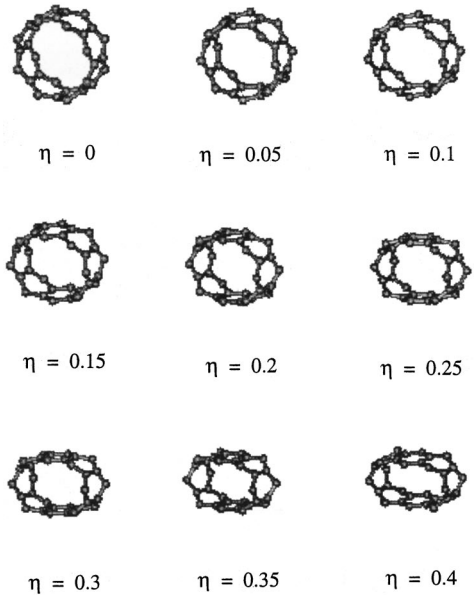


Fig. 2 Flattening of a (8, 0) carbon nanotube with different degrees of deformation up to 40% flattening

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, March 15, 2001; final revision, October 30, 2001. Associate Editor: D. A. Kouris. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

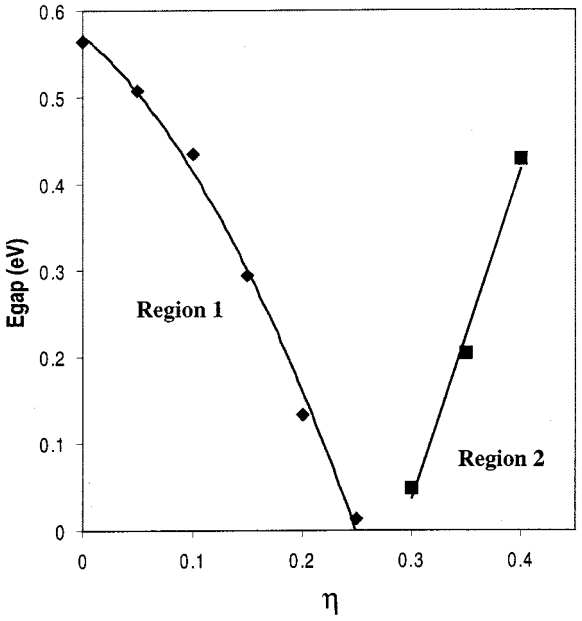


Fig. 3 Energy gap as a function of the flatness. Region 1 shows a band gap closing corresponding to a semiconductor-metal transition, and Region 2 shows bandgap reopening leading to metal-semiconductor transition.

of $12\times12\times4.29$ Å^3 corresponding to 40Ry cutoff energy. A semiconducting (8, 0) nanotube is placed in the supercell including one unit of the (8, 0) nanotube with the tube axis along *z*-direction. The Brillouin zone sampling is approximated by six *K*-points along the tube axis, which is shown to be a good approximation for (8, 0) and (10, 0) nanotubes ([16,17]). The geometry of each flattened nanotube is fully relaxed, while maintaining the flattened part, until the atomic forces become smaller than 0.1 eV/Å.

The geometry of a flattened nanotube is described in Fig. 1. The nanotube cross section is initially circular (Fig. 1(*a*)), and the deformed cross section can be described as two straight lines (top and bottom parts) joined together by two semicircles (Fig. 1(*b*)). In order to characterize the flatness, a dimensionless quantity is defined, $\eta=(D_0-d)/D_0$, where D_0 is the original diameter of an undeformed (8,0) nanotube (6.3 Å), and d is the distance between the two flattened parts (straight lines) of the cross section. The atomic structure is then relaxed with a constraint of freezing the carbon atoms in the flattened parts. Figure 2 shows the atomic structure of a (8, 0) nanotube unit undergoing up to 40% deformation.

3 Results

The dependence of the energy gap on the flatness is shown in Fig. 3. The semiconductor-metal transition can be clearly seen at $\eta=0.25$. It is interesting to observe that the energy gap shows the two distinct behaviors as a function of flattening, η. These two distinct regions are denoted in Fig. 3 as Region 1 and Region 2. In Region 1, the energy bandgap is quadratically decreasing from 0.56 eV at $\eta=0$ to 0.012 eV at $\eta=0.25$. In Region 2, the energy bandgap is linearly increasing again up to 0.4 eV at $\eta=0.4$. As far as we know, this finding represents the first discovery of a continuous semiconductor–metal–semiconductor (SMS) transition as a function of mechanical deformation in nanostructures. We expect that this SMS transition can be used to develop a nano electro mechanical feedback system that can maintain constant mechanical deformation or a bistable switch.

The origin of this band gap change can be seen from the band structures, shown in Fig. 4. Top panels in Fig. 4 show the band structures of the nanotube undergoing SM transition (Region 1), and bottom panels correspond to Region 2 of MS transition. In Region 1, the band structure shows that the lowest energy in conduction band moves down to close the band gap as the tube flattened from $\eta=0$ to $\eta=0.25$, while the two degenerate highest energy states in valence band do not change significantly. In the Region 2, the two degenerate highest energy states in valence band move down as the flatness further increases to $\eta>0.25$. In

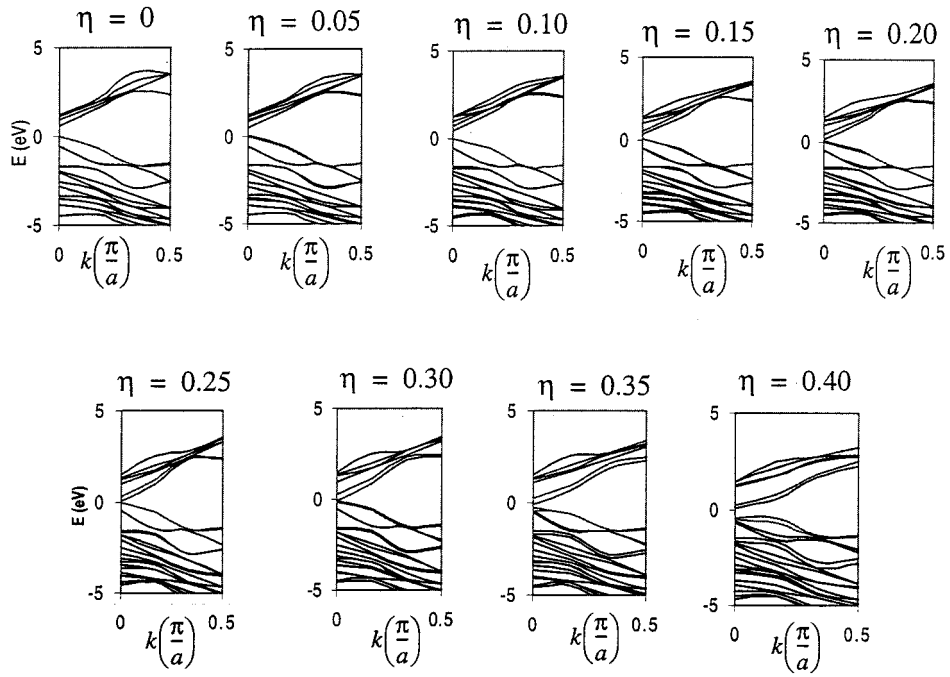


Fig. 4 Electronic band structures at different degrees of flattening deformation. Negative energies correspond to valence band state, and positive energies correspond to conduction band states.

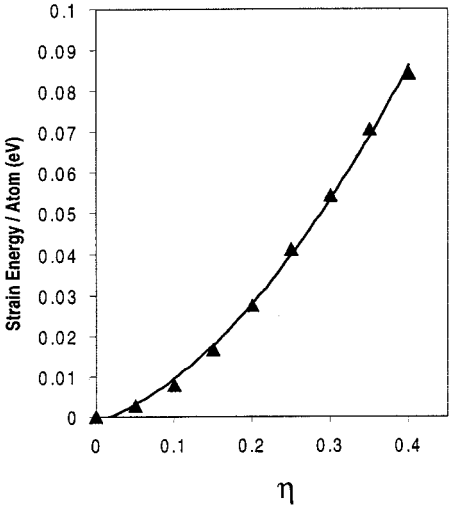


Fig. 5 Strain energy per atom as a function of flattening, η

this MS transition part, conduction bands are not changing so that the bandgap opens again. We are currently analyzing the nature of electronic structures to elucidate microscopic origin of band structure changes.

Experimentally, an AFM tip pushing on a nanotube can induce a flattened cross section, which may be related to the observed electronic property change in metallic nanotube ([11]). The force necessary to induce a semiconductor–metal transition can be calculated as a derivative of the strain energy w.r.t. the distortion. Figure 5 shows the strain energy as a function of η which can be considered as strain. The quadratic behavior is consistent with a linear elastic strain energy behavior: $E_s = 1/2K\eta^2$ with stiffness $K = 0.73$ eV.

4 Conclusions

In summary, we have studied the electronic structure change of a semiconducting carbon nanotube as a function of the flattening deformation, and discovered a novel semiconductor–metal–semiconductor transition effect. Our finding is consistent with a simulation study of (10, 0) nanotube flattening which shows a similar SM transition ([14]). Our finding on (8, 0) nanotube predicts that a further flattening of (10, 0) nanotube will reopen the bandgap that was not addressed by Ref. [14].

Acknowledgments

We thank Professor Dai for many helpful discussions and sharing experimental data before publication. The calculations are performed on Origin2000 at NCSA super computer lab allocated through the NPACI Grant SUA239 ''Nanoscale Material Simulations.''

References

[1] Saito, R., Dresselhaus, G., and Dresselhaus, M. S., 1998, *Physical Properties of Carbon Nanotubes*, Imperial College Press, London.
[2] Iijima, S., 1991, ''Helical Microtubules of Graphitic Carbon,'' Nature (London), **354**, pp. 56–58.
[3] Treacy, M. M. J., Ebbesen, T. W., and Gibson, J. M., 1996, ''Exceptionally High Yong's Modulus Observed for Individual Carbon Nanotubes,'' Nature (London), **381**, pp. 678–680.
[4] Falvo, M. R., Clary, G. J., Taylor, R. M., Chi, V., Brooks, F. P., Washburn, S., and Superfine, R., 1997, ''Bending and Buckling of Carbon Nanotubes Under Large Strain,'' Nature (London), **389**, pp. 582–584.
[5] Wong, E. W., Sheehan, P. E., and Lieber, C. M., 1997, ''Nanobeam Mechanics: Elasticity, Strength, and Toughness of Nanorods and Nanotubes,'' Science, **277**, pp. 1971–1974.
[6] Charlier, J. C., and Issi, J. P., 1998, ''Electronic Structure and Quantum Transport in Carbon Nanotubes,'' Appl. Phys. A: Mater. Sci. Process., **67**, pp. 79–87.
[7] Hamada, N., Sawada, S., and Oshiyama, A., 1992, ''New One-Dimensional Conductors: Graphitic Microtubules,'' Phys. Rev. Lett., **68**, pp. 1579–1581.
[8] Heyd, R., Charlier, A., and McRae, E., 1997, ''Uniaxial-Stress Effects on the Electronic Properties of Carbon Nanotubes,'' Phys. Rev. B, **55**, pp. 6820–6824.
[9] Lourie, O., Cox, D. M., and Wagner, H. D., 1998, ''Buckling and Collapse of Embedded Carbon Nanotubes,'' Phys. Rev. Lett., **81**, pp. 1638–1641.
[10] Yakobson, B. I., and Smalley, R. E., 1997, ''Fullerene Nanotubes: C 1,000,000 and Beyond,'' Am. Sci., **85**, pp. 324–337.
[11] Thomas, W. T., Zhou, C. W., Alexseyev, L., Kong, J., Dai, H. J., Lei, L., Jayanthi, C. S., Tang, M. J., and Wu, S. Y., 2000, ''Reversible Electromechanical Characteristics of Carbon Nanotubes Under Local-Probe Manipulation,'' Nature (London), **405**, pp. 769–772.
[12] Rochefort, A., Salahub, D. R., and Avouris, P., 1998, ''The Effect of Structural Distortions on the Electronic Structure of Carbon Nanotubes,'' Chem. Phys. Lett., **297**, pp. 45–50.
[13] Iijima, S., 1996, ''Structural Flexibility of Carbon Nanotubes,'' J. Chem. Phys., **104**, pp. 2089–2092.
[14] Mazzoni, M. S. C. and Chacham, H., 2000, ''Bandgap Closure of a Flattened Semiconductor Carbon Nanotubes: A First-Principle Study,'' Appl. Phys. Lett., **76**, pp. 1561–1563.
[15] Payne, M. C. Teter, M. P., Allan, D. C., and Joannopoulos, J. D., 1992, ''Iterative Minimization Techniques for *ab initio* Total-Energy Calculations: Molecular Dynamics and Conjugate Gradients,'' Rev. Mod. Phys., **64**, pp. 1045–1096.
[16] Srivestava, D., Menon, M., and Cho, K. J., 1999, ''Nanoplasticity of Single-Wall Carbon Nanotubes Under Uniaxial Compression,'' Phys. Rev. Lett., **83**, pp. 2973–2976.
[17] Peng, S., and Cho, K. J., 2000, ''Chemical Control of Nanotube Electronics,'' Nanotechnology, **11**, pp. 57–60.

P. Zhang

Y. Huang[1]
e-mail: huang9@uiuc.edu

Department of Mechanical and Industrial Engineering,
University of Illinois,
Urbana, IL 61801

H. Gao
Division of Mechanics and Computation,
Department of Mechanical Engineering,
Stanford University,
Palo Alto, CA 94305

K. C. Hwang
Department of Engineering Mechanics,
Tsinghua University,
Beijing 100084, P. R. China

Fracture Nucleation in Single-Wall Carbon Nanotubes Under Tension: A Continuum Analysis Incorporating Interatomic Potentials

Carbon nanotubes show great promise for applications ranging from nanocomposites, nanoelectronic components, nanosensors, to nanoscale mechanical probes. These materials exhibit very attractive mechanical properties with extraordinarily high stiffness and strength, and are of great interest to researchers from both atomistic and continuum points of view. In this paper, we intend to develop a continuum theory of fracture nucleation in single-walled carbon nanotubes by incorporating interatomic potentials between carbon atoms into a continuum constitutive model for the nanotube wall. In this theory, the fracture nucleation is viewed as a bifurcation instability of a homogeneously deformed nanotube at a critical strain. An eigenvalue problem is set up to determine the onset of fracture, with results in good agreement with those from atomistic studies. [DOI: 10.1115/1.1469002]

1 Introduction

Interest in carbon nanotubes continues to grow since their first discovery ([1,2]) and the establishment of new effective methods of producing them ([3]). Nanotubes have a single or multiple layers of atoms in the tube thickness direction. These single-wall or multiwall carbon nanotubes display superior mechanical properties. For example, both atomistic simulations and transmission electron microscopy and atomic force microscopy experiments have shown that the Young's moduli of carbon nanotubes are in the terapascal (TPa) range ([4–10]), orders of magnitude larger than their macroscopic counterparts (graphite). The strengths of carbon nanotubes, defined by the maximum force per unit area before failure, are also exceptionally high, on the order of 30 GPa in bending ([7,10]). Yakobson et al. [7] studied buckling instability of carbon nanotubes in compression and found that a 6-nm long nanotube with 1-nm diameter can sustain a large compressive strain of 5% prior to buckling. Even larger strain levels were reported for carbon nanotubes under torsion ([7]). Mechanisms responsible for these superior mechanical properties of carbon nanotubes at large strains have been identified by atomistic and quantum studies ([11–14]). The superior material behavior and light mass density of carbon nanotubes are the basis for many proposed applications of nanotubes ranging from nanocomposites to probe microscopy.

There are, however, very limited studies on the fracture behavior of carbon nanotubes. Lourie and Wagner [15] made transmission electron microscopy observations of fracture of single-wall carbon nanotubes embedded in an epoxy resin under uniaxial tension. Yakobson and Smalley [16] conducted atomistic studies of fracture of a zigzag (13,0) single-wall carbon nanotube and reported the fracture process to be similar to the unraveling of a sweater. Based on molecular dynamics simulations using a realistic interatomic potential for carbon ([17,18]), Yakobson et al. [19] studied fracture of carbon nanotubes under simple tension. They established that, before the strain reached a critical level, deformation in the nanotube remained uniform. At this critical strain level, which was named the *breaking strain* ([19]), a few carbon bonds broke almost simultaneously, and the resulting ''hole'' in the nanotube wall became a precursor of fracture. The atomic disorder propagated rapidly along the circumference of the nanotube, and a largely distorted and unstable neck formed. The *breaking strain*, which represented the critical strain for a nanotube to undergo nonuniform deformation, depended strongly on the temperature. It was approximately 55% at 50K, 50% at 100K, 42% at 300K, 34% at 600K, and 25% at 1200K. At later stages of fracture, the nanotube fragments were connected by one (or few) unravelling monoatomic chain which grew up to hundreds of atoms in length before its breakage.

There are virtually no continuum studies of nanotubes because it is generally thought that continuum mechanics theories are not applicable on the atomic or nanometer scale, and one must rely on atomistic studies. Friesecke and James [20] proposed an approach to pass the atomic information to a continuum theory for a nanostructure in which one or more dimensions are large relative to atomic scale. Yakobson et al. [7] introduced a continuum shell model for single-wall carbon nanotubes in order to determine the Young's modulus and effective nanotube thickness by fitting the tension and bending rigidity obtained from molecular dynamics simulations. The buckling strain and buckling mode of a compressed nanotube predicted by the shell model agreed well with the molecular dynamics simulations ([7,21]) and with the experimentally observed patterns ([21–23]). However, as Yakobson et al. [19] pointed out, the continuum shell model was not capable of predicting the *breaking strain* of a nanotube under tension, nor the associated fracture behavior of a nanotube.

In this paper we proposed a new continuum theory that incorporates the Tersoff-Brenner interatomic potential for carbon ([17,18]). A systematic approach is adopted to derive the continuum strain energy density by averaging the bonding energy in carbon bonds, which leads to a continuum constitutive model for

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, March 15, 2001; final revision, December 15, 2001. Associate Editor: D. A. Kouris. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

carbon nanotubes. We then use this continuum theory to study the *breaking strain*, at which the nanotube begins to deform nonuniformly, leading rapidly to the formation of a necked zone followed by the final stage of fracture. The *breaking strain* determined from molecular dynamics simulations is modeled as the critical strain of the nanotube in the continuum analysis at the onset of bifurcation, i.e., when nonuniform deformation starts.

2 Interatomic Potential for Carbon

Tersoff [17] and Brenner [18] provided an expression of the bonding energy between atoms i and j for carbon,

$$V(r_{ij}) = V_R(r_{ij}) - \overline{B}_{ij} V_A(r_{ij}), \tag{1}$$

where r_{ij} is the distance from atom i to atom j, V_R and V_A are the repulsive and attractive pair terms given by

$$V_R(r) = \frac{D^{(e)}}{S-1} e^{-\sqrt{2S}\beta(r-R^{(e)})} f_c(r), \tag{2}$$

$$V_A(r) = \frac{D^{(e)} S}{S-1} e^{-\sqrt{2/S}\beta(r-R^{(e)})} f_c(r); \tag{3}$$

where, for carbon, parameters $D^{(e)} = 6.325$ eV, $S = 1.29$, $\beta = 15\ \text{nm}^{-1}$, $R^{(e)} = 0.1315$ nm is the unstretched bond length at equilibrium, and the function f_c is merely a smooth cutoff function to limit the range of the potential, and is given by

$$f_c(r) = \begin{cases} 1 & r < R^{(1)}, \\ \dfrac{1}{2}\left\{1 + \cos\left[\dfrac{\pi(r-R^{(1)})}{R^{(2)}-R^{(1)}}\right]\right\} & R^{(1)} < r < R^{(2)}, \\ 0 & r > R^{(2)}, \end{cases} \tag{4}$$

which is continuous and has a cutoff of $R^{(2)} = 0.2$ nm and $R^{(1)} = 0.17$ nm to include only the first-neighbor shell for carbon.

The parameter $\overline{B}_{ij}$ in (1) represents a multibody coupling between the bond from atom i to atom j and the local environment of atom i, and is given by

$$\overline{B}_{ij} = \frac{1}{2}(B_{ij} + B_{ji}), \tag{5}$$

where

$$B_{ij} = \left[1 + \sum_{k(\neq i,j)} G(\theta_{ijk}) f_c(r_{ik})\right]^{-\delta}, \tag{6}$$

$\delta = 0.80469$, r_{ik} is the distance between atoms i and k, f_c is the cutoff function in (4), θ_{ijk} is the angle between bonds $i-j$ and $i-k$, and the function G is given by

$$G(\theta) = a_0\left[1 + \frac{c_0^2}{d_0^2} - \frac{c_0^2}{d_0^2 + (1+\cos\theta)^2}\right], \tag{7}$$

and the constants $a_0 = 0.011304$, $c_0 = 19$, $d_0 = 2.5$. It is straightforward to show that $\overline{B}_{ij}$ is very close to unity (one). For $\theta = 2/3\pi$, as in the equilibrium structure of a carbon nanotube, the coefficient $\overline{B}_{ij} = 0.95$.

3 A Continuum Theory Incorporating the Interatomic Potentials

Gao and Klein [24] proposed a systematic approach to incorporate a cohesive force law into the constitutive model of solids. An empirical cohesive force law was used to represent the interactions between material points in the solid. The constitutive law was then derived by homogenizing all cohesive bonds at each material point. This approach is adopted in the present study to incorporate the interatomic potential into a continuum model for carbon.

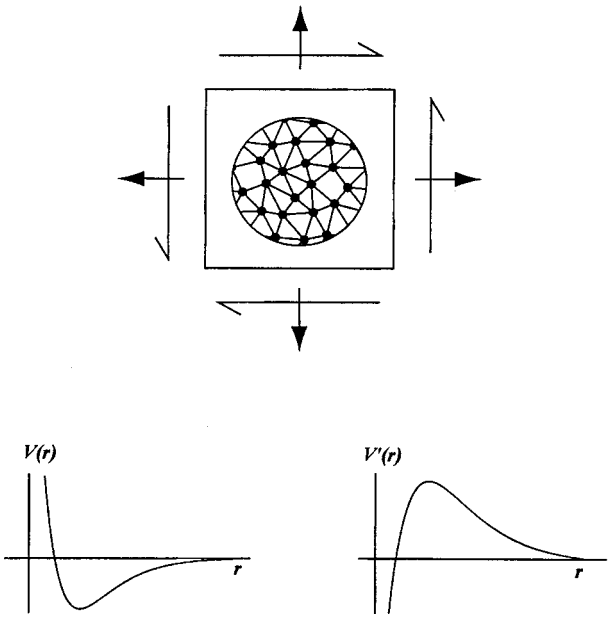


Fig. 1 A multiscale framework to establish a continuum theory incorporating the interatomic potential *V*

As shown in Fig. 1, a multiscale approach is used to link the strain energy density on the continuum level to the interatomic potentials on the atomic level. On the continuum level, a small representative cell is taken for each material point such that the deformation within the cell can be considered uniform. The strain energy stored in the representative cell on the continuum level is evaluated by the total energy of all atomic bonds within the cell.

Let $\mathbf{F}$ denote the deformation gradient at a material point on the continuum level. The Lagrangian strain tensor is

$$\mathbf{E} = \frac{1}{2}(\mathbf{F}^{\mathbf{T}} \cdot \mathbf{F} - \mathbf{I}), \tag{8}$$

where $\mathbf{F}^T$ is the transpose of $\mathbf{F}$ and $\mathbf{I}$ is the second-order identity tensor. The unstretched bond length between atoms i and j at equilibrium is denoted by $r_{ij}^{(0)}$, while the bond orientation is characterized by its unit vector $\mathbf{n}^{(0)}$ in the undeformed configuration. The stretched bond length after the deformation is imposed becomes

$$r_{ij} = r_{ij}^{(0)} \sqrt{1 + 2\mathbf{n}^{(0)} \cdot \mathbf{E} \cdot \mathbf{n}^{(0)}}. \tag{9}$$

The energy stored in the bond is $V(r_{ij})$, where V is the interatomic potential for carbon given in Eq. (1). Following the Cauchy-Born rule ([25,26]), the total strain energy stored in the representative cell is $\Sigma_{i<j} V(r_{ij})$, where the summation is over all atomic bonds within the cell. Therefore, the strain energy density at this material point on the continuum level is

$$W = \frac{\sum_{i<j} V(r_{ij})}{\Omega_\epsilon}, \tag{10}$$

where Ω_ϵ is the volume of the representative cell.

For a carbon nanotube, the interatomic potential in Eq. (1) accounts for the interaction among atoms only within the first-neighbor shell. Therefore, the unstretched bond length at equilibrium is $r_{ij}^{(0)} = R^{(e)} = 0.1315$ nm. The summation over all bonds in Eq. (10) can be effectively evaluated by an integration over the bond orientations,

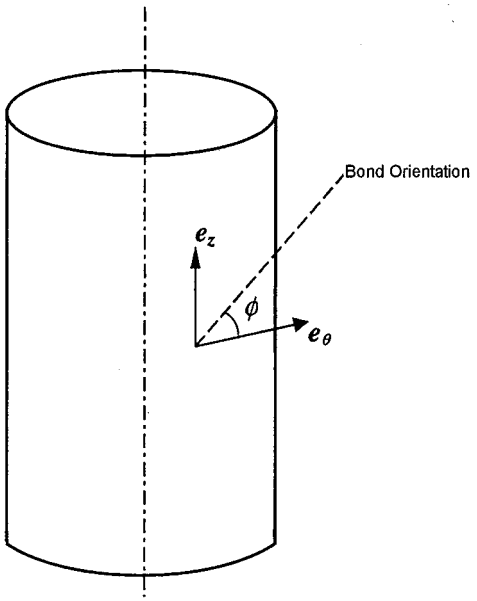


Fig. 2 A schematic diagram to show the coordinate system on the nanotube surface; $\mathbf{e}_\theta$ and $\mathbf{e}_Z$ are unit vectors in the circumferential and axial directions, respectively, and ϕ is the angle between the bond and the circumferential direction

$$W=\int_0^{2\pi} V(r)D(\phi)d\phi, \tag{11}$$

where, as shown in Fig. 2, ϕ is the angle between the bond and the circumferential direction $\mathbf{e}_\theta$ within the nanotube surface; the bond orientation $\mathbf{n}^{(0)}=\cos\phi\mathbf{e}_\theta+\sin\phi\mathbf{e}_Z$ and $\mathbf{e}_Z$ is the unit vector in the axial direction; the stretched bond length $r=R^{(e)}\sqrt{1+2\mathbf{n}^{(0)}\cdot\mathbf{E}\cdot\mathbf{n}^{(0)}}$; and $D(\phi)$ is the bond density function such that $D(\phi)d\phi$ is the number of bonds per unit area of the undeformed nanotube surface with bond orientation between ϕ and $\phi+d\phi$. For a single-wall carbon nanotube, the bond density function D takes the form

$$D=D_0\sum_{m=1}^{3}\delta(\phi-\phi_m), \tag{12}$$

where δ is the Dirac delta function, ϕ_1, ϕ_2, and ϕ_3 are discrete angles of carbon bonds of the hexagonal lattice within the nanotube surface, and the constant D_0 is the number of carbon bonds in the ϕ_m direction per unit area of the undeformed nanotube surface. Since a carbon nanotube has a hexagonal lattice structure, it can be shown that

$$D_0=\frac{2}{3\sqrt{3}}\frac{1}{R^{(e)^2}}, \tag{13}$$

where $R^{(e)}=0.1315$ nm is the unstretched bond length for carbon.

The second (symmetric) Piola-Kirchhoff stress $\mathbf{T}$ is the work conjugate of the Lagrangian strain $\mathbf{E}$, and is obtained from the strain energy density W as

$$\mathbf{T}=\frac{\partial W}{\partial\mathbf{E}}=\int_0^{2\pi}\frac{V'(r)}{r}R^{(e)^2}\mathbf{n}^{(0)}\mathbf{n}^{(0)}D(\phi)d\phi. \tag{14}$$

Its increment, $\dot{\mathbf{T}}$, is related to the increment of strain, $\dot{\mathbf{E}}$, via the incremental modulus $\mathbf{C}$,

$$\dot{\mathbf{T}}=\mathbf{C}:\dot{\mathbf{E}}, \tag{15}$$

where the incremental modulus for a single-wall carbon nanotube is given by

$$\mathbf{C}=\frac{\partial^2 W}{\partial\mathbf{E}\partial\mathbf{E}}=\int_0^{2\pi}\left(\frac{V''(r)}{r^2}-\frac{V'(r)}{r^3}\right)R^{(e)^4}\mathbf{n}^{(0)}\mathbf{n}^{(0)}\mathbf{n}^{(0)}\mathbf{n}^{(0)}D(\phi)d\phi. \tag{16}$$

The interatomic potential of carbon, given in Eq. (1), has been directly incorporated into the above constitutive relations. The stresses and incremental moduli, however, are given in terms of integration with respect to the spatial orientations of the atomic bonds. In order to further simplify the analysis, we make the following approximations:

(i) The coefficient $\overline{B}_{ij}$ in the interatomic potential for carbon, Eq. (1), represents a many-body coupling effect. As discussed in Section 2, the coefficient $\overline{B}_{ij}$ is very close to unity (one) for carbon nanotubes, and the difference between $\overline{B}_{ij}$ and one is always less than 5%. Since we are interested in the *breaking strain* of carbon nanotubes rather than the atomic arrangement of carbon bonds, we may approximate $\overline{B}_{ij}$ by one.

(ii) Rigorously speaking, the graphite structure of a carbon nanotube does not exactly satisfy the Cauchy-Born rule because the hexagonal graphite structure lacks the centrosymmetry required for a homogeneous deformation gradient. This means that a ''homogenously deformed'' nanotube is actually inhomogeneous on the unit cell level. We are interested in a bifurcation analysis of fracture nucleation in the spirit of Euler's classical analysis of beam buckling. For this purpose, we approximate the hexagonal graphite wall of a nanotube by a comparison medium having randomized bond structure but identical mass density and Young's modulus as the graphite. That is, we smear the bond density D into an isotropic function $D(\phi)=3/2\pi D_0=1/\sqrt{3}\pi 1/R^{(e)^2}$. This density function has the same number of atomic bonds per unit area as the graphite structure and satisfies the centrosymmetry required for the Cauchy-Born rule.

Based on the above approximations, the second Piola-Kirchhoff stress $\mathbf{T}$ in Eq. (14) and the incremental modulus $\mathbf{C}$ in Eq. (16) can then be simplified as

$$\mathbf{T}=\frac{3D_0}{2\pi}R^{(e)^2}\int_0^{2\pi}\frac{V'(r)}{r}\mathbf{n}^{(0)}\mathbf{n}^{(0)}d\phi, \tag{17}$$

$$\mathbf{C}=\frac{3D_0}{2\pi}R^{(e)^4}\int_0^{2\pi}\left(\frac{V''(r)}{r^2}-\frac{V'(r)}{\mathbf{r}^3}\right)\mathbf{n}^{(0)}\mathbf{n}^{(0)}\mathbf{n}^{(0)}\mathbf{n}^{(0)}d\phi. \tag{18}$$

4 The Bifurcation Analysis of a Carbon Nanotube Under Tension

We investigate a single-wall carbon nanotube subjected to simple tension along the axial direction Z. The nanotube undergoes uniform and axisymmetric deformation prior to bifurcation, and the only nonzero component of the second Piola-Kirchhoff stress is T_{ZZ}. At the onset of bifurcation, the deformation starts to become nonuniform.

Let (R,θ,Z) denote the cylindrical coordinate in the initial, undeformed configuration, and $(\mathbf{e}_\mathbf{R},\mathbf{e}_\theta,\mathbf{e}_\mathbf{Z})$ the corresponding unit vectors. The displacement is

$$\mathbf{U}=U_\mathbf{R}e_\mathbf{R}+U_\theta\mathbf{e}_\theta+U_Z\mathbf{e}_\mathbf{Z}. \tag{19}$$

The deformation gradient $\mathbf{F}$ is

$$\mathbf{F}=\mathbf{I}+\mathbf{U}\nabla, \tag{20}$$

where $\nabla=\mathbf{e}_\mathbf{R}\partial/\partial R+\mathbf{e}_\theta/R\partial/\partial\theta+\mathbf{e}_\mathbf{Z}\partial/\partial Z$ is the gradient operator. The Lagrangian strain $\mathbf{E}$, second Piola-Kirchhoff stress $\mathbf{T}$, and incremental moduli $\mathbf{C}$ are obtained in terms of $\mathbf{U}$ from Eqs. (8), (17), and (18), respectively.

The equilibrium equation is

$$(\mathbf{F}\cdot\mathbf{T})\cdot\nabla=0. \tag{21}$$

The traction-free conditions on the inner and outer surfaces of the nanotube are

$$\mathbf{F}\cdot\mathbf{T}\cdot\mathbf{e_R}=0. \tag{22}$$

Integration of equilibrium Eq. (21) over the tube thickness, in conjunction with boundary condition (22), gives the governing equations for displacements U_R, U_θ, and U_Z,

$$\frac{1}{R}\frac{\partial}{\partial\theta}(F_{R\alpha}T_{\alpha\theta})-\frac{1}{R}F_{\theta\alpha}T_{\alpha\theta}+\frac{\partial}{\partial Z}(F_{R\alpha}T_{\alpha Z})=0,$$

$$\frac{1}{R}F_{R\alpha}T_{\alpha\theta}+\frac{1}{R}\frac{\partial}{\partial\theta}(F_{\theta\alpha}T_{\alpha\theta})+\frac{\partial}{\partial Z}(F_{\theta\alpha}T_{\alpha Z})=0, \tag{23}$$

$$\frac{1}{R}\frac{\partial}{\partial\theta}(F_{Z\alpha}T_{\alpha\theta})+\frac{\partial}{\partial Z}(F_{Z\alpha}T_{\alpha Z})=0,$$

where R is the nanotube radius; the summation for α is over θ and Z; and $\mathbf{F}$ and $\mathbf{T}$ are the average deformation gradient and stress over the nanotube thickness.

At the onset of bifurcation, the deformation is axisymmetric ($U_\theta=0$), and the deformation gradient $\mathbf{F}$ is uniform and has nonzero components $F_{\theta\theta}=1+U_R/R$ and $F_{ZZ}=dU_Z/dZ$. The only nonzero component of the second Piola-Kirchhoff stress $\mathbf{T}$ at the onset of bifurcation is T_{ZZ}, and from Eq. (17),

$$T_{ZZ}=\frac{3D_0}{2\pi}R^{(e)^2}\int_0^{2\phi}\frac{V'(r)}{r}\sin^2\phi d\phi, \tag{24}$$

where $r=R^{(e)}\sqrt{F_{\theta\theta}^2\cos^2\phi+F_{ZZ}^2\sin^2\phi}$. The vanishing of $T_{\theta\theta}$ in uniaxial tension requires

$$0=T_{\theta\theta}=\frac{3D_0}{2\pi}R^{(e)^2}\int_0^{2\pi}\frac{V'(r)}{r}\cos^2\phi d\phi, \tag{25}$$

which governs the relation between the nonzero deformation gradients $F_{\theta\theta}$ and F_{ZZ}. Similarly, the incremental modulus, $\mathbf{C}$ in Eq. (18), is uniform at the onset of bifurcation and has nonzero components

$$C_{\theta\theta\theta\theta}=\frac{3D_0}{2\pi}R^{(e)^4}\int_0^{2\pi}\left[\frac{V''(r)}{r^2}-\frac{V'(r)}{r^3}\right]\cos^4\phi d\phi,$$

$$C_{ZZZZ}=\frac{3D_0}{2\pi}R^{(e)^4}\int_0^{2\pi}\left[\frac{V''(r)}{r^2}-\frac{V'(r)}{r^3}\right]\sin^4\phi d\phi, \tag{26}$$

$$C_{\theta\theta ZZ}=C_{\theta Z\theta Z}=\frac{3D_0}{2\pi}R^{(e)^4}\int_0^{2\pi}\left[\frac{V''(r)}{r^2}-\frac{V'(r)}{r^3}\right]\sin^2\phi\cos^2\phi d\phi.$$

It should be pointed out, however, that the increments $\dot{\mathbf{F}}$ and $\dot{\mathbf{T}}$ are nonuniform and not necessarily asixymmetric at the onset of bifurcation, and may have many nonzero components (e.g., $\dot{T}_{\theta\theta}$, $\dot{T}_{ZZ}$ and $\dot{T}_{\theta Z}$).

Making use of the incremental constitutive relation (14), the increments of the governing Eqs. (23) at the onset of bifurcation take the form

$$-\frac{1}{R}C_{\theta\theta\theta\theta}F_{\theta\theta}^2\dot{F}_{\theta\theta}-\frac{1}{R}C_{\theta\theta ZZ}F_{\theta\theta}F_{ZZ}\dot{F}_{ZZ}+T_{ZZ}\frac{\partial\dot{F}_{RZ}}{\partial Z}=0,$$

$$\frac{1}{R}C_{\theta\theta\theta\theta}F_{\theta\theta}^2\frac{\partial\dot{F}_{\theta\theta}}{\partial\theta}+\frac{1}{R}C_{\theta\theta ZZ}F_{\theta\theta}F_{ZZ}\frac{\partial\dot{F}_{ZZ}}{\partial\theta}+(T_{ZZ}$$

$$+C_{\theta\theta ZZ}F_{\theta\theta}^2)\frac{\partial\dot{F}_{\theta Z}}{\partial Z}+C_{\theta\theta ZZ}F_{\theta\theta}F_{ZZ}\frac{\partial\dot{F}_{Z\theta}}{\partial Z}=0, \tag{27}$$

$$\frac{1}{R}C_{\theta\theta ZZ}F_{\theta\theta}F_{ZZ}\frac{\partial\dot{F}_{\theta Z}}{\partial\theta}+\frac{1}{R}C_{\theta\theta ZZ}F_{ZZ}^2\frac{\partial\dot{F}_{Z\theta}}{\partial\theta}+C_{\theta\theta ZZ}F_{\theta\theta}F_{ZZ}\frac{\partial\dot{F}_{\theta\theta}}{\partial Z}$$

$$+(T_{ZZ}+C_{ZZZZ}F_{ZZ}^2)\frac{\partial\dot{F}_{ZZ}}{\partial Z}=0,$$

where

$$\dot{F}_{R\theta}=\frac{1}{R}\frac{\partial\dot{U}_R}{\partial\theta}-\frac{\dot{U}_\theta}{R},\quad \dot{F}_{\theta\theta}=\frac{\dot{U}_R}{R}+\frac{1}{R}\frac{\partial\dot{U}_\theta}{\partial\theta},\quad \dot{F}_{Z\theta}=\frac{1}{R}\frac{\partial\dot{U}_Z}{\partial\theta},$$

$$\dot{F}_{RZ}=\frac{\partial\dot{U}_R}{\partial Z},\quad \dot{F}_{\theta z}=\frac{\partial\dot{U}_\theta}{\partial Z},\quad \dot{F}_{ZZ}=\frac{\partial\dot{U}_Z}{\partial Z}. \tag{28}$$

The nanotube is subjected to an axial displacement and no-shear traction at the two ends. Therefore, at the onset of bifurcation, the boundary conditions for the bifurcation solution are

$$\dot{U}_Z=\dot{t}_R=\dot{t}_\theta=0\quad\text{at } Z=0 \text{ and } Z=L, \tag{29}$$

where L is the length of the nanotube, $\dot{t}_R$ and $\dot{t}_\theta$ are the shear traction increments in the radial and circumferential directions, respectively. It can be shown that the above boundary condition can be equivalently written as

$$\dot{U}_Z=\frac{\partial\dot{U}_R}{\partial Z}=\frac{\partial\dot{U}_\theta}{\partial Z}=0\quad\text{at } Z=0 \text{ and } Z=L. \tag{30}$$

The homogeneous governing Eqs. (27) and (28) and boundary conditions (30) form an eigenvalue problem with the axial strain E_{ZZ}, or equivalently, the deformation gradient F_{ZZ} as the eigenvalue. We first study the axisymmetic bifurcation mode, $\dot{U}_R=\dot{U}_R(Z)$, $\dot{U}_\theta=0$, $\dot{U}_Z=\dot{U}_Z(Z)$. The governing Eq. (27) become

$$T_{ZZ}\frac{d^2\dot{U}_R}{dZ^2}-\frac{1}{R}C_{\theta\theta ZZ}F_{\theta\theta}F_{ZZ}\frac{d\dot{U}_Z}{dZ}-C_{\theta\theta\theta\theta}F_{\theta\theta}^2\frac{\dot{U}_R}{R^2}=0,$$

$$(T_{ZZ}+C_{ZZZZ}F_{ZZ}^2)\frac{d^2\dot{U}_Z}{dZ^2}+\frac{1}{R}C_{\theta\theta ZZ}F_{\theta\theta}F_{ZZ}\frac{d\dot{U}_R}{dZ}=0. \tag{31}$$

Its solution, satisfying the homogeneous boundary conditions (30), has the form

$$(\dot{U}_R,\dot{U}_Z)=\left(\dot{U}_{R0}\cos\frac{m\pi Z}{L},\ \dot{U}_{Z0}\sin\frac{m\pi Z}{L}\right),$$

where $m=1,2,3,\cdots$ is the eigen mode number and $(\dot{U}_{R0},\dot{U}_{Z0})$ is the eigenvector. The governing Eq. (31) then become two linear, homogeneous algebraic equations for $\dot{U}_{R0}$ and $\dot{U}_{Z0}$. In order to have a nontrivial solution (i.e., bifurcation), the determinant of the coefficient matrix for the linear algebraic equations must vanish. This yields the critical condition for bifurcation as

$$F_{\theta\theta}^2[F_{ZZ}^2(C_{\theta\theta\theta\theta}C_{ZZZZ}-C_{\theta\theta ZZ}^2)+C_{\theta\theta\theta\theta}T_{ZZ}]+T_{ZZ}(T_{ZZ}$$

$$+C_{ZZZZ}F_{ZZ}^2)\left(\frac{m\pi R}{L}\right)^2=0. \tag{32}$$

This, in conjunction with $T_{\theta\theta}=0$ in Eq. (25), provides two equations to determine $F_{\theta\theta}$ and F_{ZZ} at the onset of bifurcation.

For each mode number $m=1,2,3,\cdots$, $F_{\theta\theta}$ and F_{ZZ} at the onset of bifurcation are obtained numerically. Specifically, $F_{\theta\theta}$ is solved in terms of F_{ZZ} from Eq. (25), and then F_{ZZ} is solved from Eq. (32). The nanotube radius and length are fixed at $R=0.5$ nm and $L=5$ nm in the present study, consistent with Yakobson et al.'s [19] molecular dynamics studies. The numerical solution has shown that the first bifurcation mode ($m=1$) gives the smallest F_{ZZ} for axisymmetric bifurcation, $(F_{ZZ})_{\text{critical}}=1.43$. The corresponding axial strain at the onset of bifurcation is

$$(E_{ZZ})_{\text{critical}}=52\%. \tag{33}$$

We have also analyzed the nonaxisymmetric bifurcation mode, $\dot{U}_R=\dot{U}_R(Z)\cos n\theta$, $\dot{U}_\theta=\dot{U}_\theta(Z)\sin n\theta$, $\dot{U}_Z=\dot{U}_Z(Z)\cos n\theta$, where n is the wave number in the circumferential direction. The governing Eqs. (27) then become three homogeneous ordinary differential equations with homogeneous boundary conditions (30). The solution of this eigenvalue problem gives the critical axial strain for nonaxisymmetic bifurcation. The numerical solution has

shown, however, that the critical strain for nonaxisymmetic bifurcation is always larger than the axisymmetric bifurcation strain in Eq. (33).

Without any parameter fitting, the bifurcation strain in Eq. (33) predicted by the continuum theory incorporating the interatomic potential agrees very well with the *breaking strain* of 55% calculated by molecular dynamics simulations at a low temperature, 50 K [19]. This excellent agreement at low temperature is because the interatomic potential [17,18] incorporated in the continuum theory has not accounted for the temperature effect, and is applicable at low temperature. In order for such a continuum theory to successfully predict the *breaking strain* at a finite temperature, an interatomic potential including the temperature effect (e.g., [27]) needs to be used instead.

5 Summary

We have proposed a systematic approach to establish a continuum theory incorporating the interatomic potentials. Specifically, we have derived the constitutive law for single-wall carbon nanotubes from the interatomic potential for carbon [17,18]. The proposed continuum theory is applied to study fracture nucleation in a single-wall carbon nanotube under uniaxial tension. The fracture nucleation is viewed as a bifurcation instability of a homogeneously deformed nanotube at a critical strain. The axial strain at the onset of bifurcation predicted by this continuum theory incorporating the interatomic potential is 52%. Without any parameter fitting, this is in excellent agreement with the *breaking strain* of 55% obtained from molecular dynamics simulations for a carbon nanotube in tension [19]. The excellent agreement between the atomistic and continuum studies demonstrates that the proposed continuum theory is capable of analyzing the deformation of nanotubes and other nanostructures.

Acknowledgment

P. Z. acknowledges a fellowship from the Computational Science and Engineering Program, the University of Illinois at Urbana-Champaign (UIUC). Y. H. acknowledges the support from NSF (grants #99-83739 and #00-99909) and from the M&IE Program of Exploratory Studies at UIUC. H. G. acknowledges the support from NSF (grants #EEC-0085569 and #CMS-9979717). The authors also acknowledge support from the NSF of China.

References

[1] Iijima, S., 1991, ''Helical Microtubules of Graphitic Carbon,'' Nature (London), **354**, pp. 56–58.
[2] Ebbesen, T. W., and Ajayan, P. M., 1992, ''Large-Scale Synthesis of Carbon Nanotubes,'' Nature (London), **358**, pp. 220–222.
[3] Thess, A., Lee, R., Nikolaev, P., Dai, H. J., Petit, P., Rotert, J., Xu, C. H., Lee, Y. H., Kim, S. G., Rinzler, A. G., Colbert, D. T., Scuseria, G. E., Tomanek, D., Fischer, J. E., and Smalley, R. E., 1996, ''Crystalline Ropes of Metallic Carbon Nanotubes,'' Science, **273**, pp. 483–487.
[4] Robertson, D. H., Brenner, D. W., and Mintmire, J. W., 1992, ''Energetics of Nanoscale Graphitic Tubules,'' Phys. Rev. B, **45**, pp. 12592–12595.
[5] Ruoff, R. S., and Lorents, D. C., 1995, ''Mechanical and Thermal Properties of Carbon Nanotubes,'' Carbon, **33**, pp. 925–930.
[6] Treacy, M. M. J., Ebbesen, T. W., and Gibson, J. M., 1996, ''Exceptionally High Young's Modulus Observed for Individual Carbon Nanotubes,'' Nature (London), **381**, pp. 678–680.
[7] Yakobson, B. I., Brabec, C. J., and Bernholc, J., 1996, ''Nanomechanics of Carbon Tubes: Instabilities Beyond Linear Response,'' Phys. Rev. Lett., **76**, pp. 2511–2514.
[8] Cornwell, C. F., and Wille, L. T., 1997, ''Elastic Properties of Single-Walled Carbon Nanotubes in Compression,'' Solid State Commun., **101**, pp. 555–558.
[9] Lu, J. P., 1997, ''Elastic Properties of Carbon Nanotubes and Nanoropes,'' Phys. Rev. Lett., **79**, pp. 1297–1300.
[10] Wong, E. W., Sheehan, P. E., and Lieber, C. M., 1997, ''Nanobeam Mechanics: Elasticity, Strength, and Toughness of Nanorods and Nanotubes,'' Science, **277**, pp. 1971–1975.
[11] Nardelli, M. B., Yakobson, B. I., and Bernholc, J., 1998, ''Brittle and Ductile Behavior in Carbon Nanotubes,'' Phys. Rev. Lett., **81**, pp. 4656–4659.
[12] Nardelli, M. B., Yakobson, B. I., and Bernholc, J., 1998, ''Mechanism of Strain Release in Carbon Nanotubes,'' Phys. Rev. B, **57**, pp. R4277–R4280.
[13] Yakobson, B. I., 1998, ''Mechanical Relaxation and 'Intramolecular Plasticity' in Carbon Nanotubes,'' Appl. Phys. Lett., **72**, pp. 918–920.
[14] Srivastara, D., Menon, M., and Cho, K. J., 1999, ''Nanoplasticity of Single-Wall Nanotubes Under Uniaxial Compression,'' Appl. Phys. Lett., **83**, pp. 2973–2976.
[15] Lourie, O., and Wagner, H. D., 1998, ''Transmission Electron Microscopy Observations of Fracture of Single-Wall Carbon Nanotubes Under Axial Tension,'' Appl. Phys. Lett., **73**, pp. 3527–3529.
[16] Yakobson, B. I., and Smalley, R. E., 1997, ''Fullerene Nanotubes: C-1000000 and Beyond,'' Am. Sci., **85**, pp. 324–337.
[17] Tersoff, J., 1988, ''New Empirical Approach for the Structure and Energy of Covalent Systems,'' Phys. Rev. B, **37**, pp. 6991–7000.
[18] Brenner, D. W., 1990, ''Empirical Potential for Hydrocarbons for Use in Simulating the Chemical Vapor Deposition of Diamond Films,'' Phys. Rev. B, **42**, pp. 9458–9471.
[19] Yakobson, B. I., Campbell, M. P., Brabec, C. J., and Bernholc, J., 1997, ''High Strain Rate Fracture and C-Chain Unraveling in Carbon Nanotubes,'' Comput. Mater. Sci., **8**, pp. 341–348.
[20] Friesecke, G., and James, R. D., 1999, ''A Scheme for the Passage From Atomic to Continuum Theory for Thin Films, Nanotubes and Nanorods,'' J. Mech. Phys. Solids, **48**, pp. 1519–1540.
[21] Iijima, S., Brabec, C., Maiti, A., and Bernholc, J., 1996, ''Structural Flexibility of Carbon Nanotubes,'' J. Chem. Phys., **104**, pp. 2089–2092.
[22] Chopra, N. G., Benedict, L. X., Crespi, V. H., Cohen, M. L., Louie, S. G., and Zettl, Z., 1995, ''Fully Collapsed Carbon Nanotubes,'' Nature (London), **377**, pp. 135–138.
[23] Despres, J. F., Daguerre, E., and Lafdi, K., 1995, ''Flexibility of Graphene Layers in Carbon Nanotubes,'' Carbon, **33**, pp. 87–89.
[24] Gao, H. J., and Klein, P., 1998, ''Numerical Simulation of Crack Growth in an Isotropic Solid With Randomized Internal Cohesive Bonds,'' J. Mech. Phys. Solids, **46**, pp. 187–218.
[25] Milstein, F., 1980, ''Review: Theoretical Elastic Behavior at Large Strains,'' J. Mater. Sci., **15**, pp. 1071–1084.
[26] Tadmor, E. B., Ortiz, M., and Phillips, R., 1996, ''Quasicontinuum Analysis of Defects in Solids,'' Philos. Mag. A, **73**, pp. 1529–1563.
[27] LeSar, R., Najafabadi, R., and Srolovitz, D. J., 1989, ''Finite-Temperature Defect Properties From Free-Energy Minimization,'' Phys. Rev. Lett., **63**, pp. 624–627.

Stress Relaxation in Prestressed Composite Laminates

A. P. Suvorov

G. J. Dvorak
Fellow ASME

Department of Mechanical Engineering,
Aeronautical
Engineering and Mechanics,
Rensselaer Polytechnic Institute,
Troy, NY 12180-3590

Viscoelastic deformation caused in symmetric laminated plates by release of fiber prestress and by uniform thermomechanical loads is analyzed on the constituent, ply and overall laminate scales with the Transformation Field Analysis (TFA) method (G. J. Dvorak, Proc. R. Soc. Lond., 1992, A437, pp. 311–327). Fiber prestress is applied in individual plies prior to matrix cure and released after matrix consolidation. Linear or nonlinear viscoelastic constitutive relations are used to evaluate the inelastic deformation rates in terms of current constituent stress averages. The TFA method regards both thermal and inelastic strains as piecewise uniform eigenstrains acting in superposition with mechanical loads and fiber prestress release on an elastic laminate. Interactions between the eigenstrains at the three different size scales are described by certain influence functions derived from micromechanical analysis of the plies and laminates. Applications describe stress relaxation in two carbon/epoxy laminates after cooling from the curing temperature and release of optimized fiber prestress, that allows maximum tensile load application while keeping both interior and free-edge stresses within prescribed strength limits. Subsequent viscoelastic deformation under constant rate loading, and stress relaxation caused by a sustained application of an elevated temperature to a laminate without prestress are also analyzed. Results are presented in the form of initial failure maps that identify overall stress states which may or may not initiate a specific damage mode in the laminate. [DOI: 10.1115/1.1460909]

1 Introduction

Analysis of the effect of fiber prestress release on overall response, local fields, and damage resistance of composite laminates is a relatively new subject that has been examined in several recent publications ([1–4]). Typically applied to reduce fiber waviness for improvement of compressive strength, fiber prestress may also create large residual stresses that either improve or diminish damage resistance of laminated composite plates and cylindrical shells. Since the forces required to cause significant prestress magnitudes are well within the capability of present equipment used in filament winding, fiber placement or pultrusion, it is likely that many composite structures do support potentially detrimental residual stress fields. Table 1 shows the force magnitudes needed to cause 1000 MPa prestress in commonly used filament tows.

Our earlier studies have focused on elastic laminates, however, release of fiber prestress represents a sustained compressive normal stress that may cause viscoelastic deformation in polymer matrix systems. Such response under various thermomechanical loads has been investigated by several authors in recent years ([5–10]), together with experimental characterization of time-dependent deformation of polymer matrices or plies ([8–16]). Some of these experimental results and constitutive equations are utilized herein. As expected, only matrix-dominated ply responses are affected by matrix viscosity, negligible creep rates are observed in the fiber direction.

This paper extends our previous work on fiber prestress to viscoelastic matrix laminates. Constitutive relations describing both linear and nonlinear viscoelastic response of polymer matrices are reviewed in Section 2, together with recent material data for the EPON 828 epoxy resin. Section 3 describes the transformation field analysis method ([17–19]), which regards inelastic and thermal deformations in the fiber and matrix constituents and individual plies of the laminate as piecewise uniform eigenstrains that are applied in superposition with external loads and fiber prestress release to an elastic composite structure. Interactions of the eigenstrains at the ply and laminate scales are described by certain influence functions that depend only on elastic moduli of the constituents, plies and laminates. The procedure used for evaluation of stress relaxation in the phases and plies is described in Section 4. Finally, response of two AS4/EPON 828 laminates subjected to release of optimized fiber prestress and different loading histories is analyzed and the results are presented, in part, in terms of changes to the initial failure maps that bracket damage-free laminate stress regions.

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, February 28, 2001; final revision, October 23, 2001. Associate Editor: M.-J. Pindera. 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.

2 Viscoelastic Constitutive Relations

2.1 Linear Viscoelastic Relations. The linear viscoelastic constitutive equations for phase r of the composite aggregate are taken in the form

$$\boldsymbol{\epsilon}_r(t)=\int_{0-}^{t}\boldsymbol{J}_r(t-\tau)\,\frac{d\boldsymbol{\sigma}_r(\tau)}{d\tau}\,d\tau \tag{1}$$

where $\boldsymbol{J}_r$ is the creep compliance function, $\boldsymbol{J}_r(0)$ is the elastic compliance. It is convenient to change the limits of integration in (1) from 0_- and t to 0_+ and t and rewrite (1) as

$$\boldsymbol{\epsilon}_r(t)=\boldsymbol{J}_r(t)\,\boldsymbol{\sigma}_r(0)+\int_{0_+}^{t}\boldsymbol{J}_r(t-\tau)\,\frac{d\boldsymbol{\sigma}_r(\tau)}{d\tau}\,d\tau. \tag{2}$$

Initial phase stresses $\boldsymbol{\sigma}_r(0)$ in (2) are evaluated from the elastic solution for the composite system under prescribed overall traction and displacement boundary conditions. Integration of (2) by parts changes this equation to

$$\boldsymbol{\epsilon}_r(t)=\boldsymbol{J}_r(0)\,\boldsymbol{\sigma}_r(t)+\int_{0_+}^{t}\dot{\boldsymbol{J}}_r(t-\tau)\,\boldsymbol{\sigma}_r(\tau)\,d\tau. \tag{3}$$

The phase strain rate is obtained by taking the time derivative of (3),

Table 1 Force needed to generate 1000 MPa stress in filaments/tow

Fiber	Diameter, μm	Filaments/tow	Force, (N)	Force, (lbs)
S-glass	14	2000	307.88	69.22
Kevlar	12	1000	113.12	25.43
Carbon AS4	8	12,000	603.19	135.61
Carbon P-100S	10	2000	157.08	35.31

$$\dot{\boldsymbol{\epsilon}}_r(t)=\boldsymbol{J}_r(0)\,\dot{\boldsymbol{\sigma}}_r(t)+\dot{\boldsymbol{J}}_r(0)\,\boldsymbol{\sigma}_r(t)+\int_{0_+}^{t}\ddot{\boldsymbol{J}}_r(t-\tau)\,\boldsymbol{\sigma}_r(\tau)\,d\tau. \quad (4)$$

The inelastic component of the strain can be obtained by subtracting from (3) the elastic strain $\boldsymbol{J}_r(0)\boldsymbol{\sigma}_r(t)$, leading to

$$\boldsymbol{\epsilon}_r^{in}(t)=\int_{0_+}^{t}\dot{\boldsymbol{J}}_r(t-\tau)\,\boldsymbol{\sigma}_r(\tau)\,d\tau. \quad (5)$$

Rate of the inelastic strain (5) is given in accordance with (4) as

$$\dot{\boldsymbol{\epsilon}}_r^{in}(t)=\dot{\boldsymbol{J}}_r(0)\,\boldsymbol{\sigma}_r(t)+\int_{0_+}^{t}\ddot{\boldsymbol{J}}_r(t-\tau)\,\boldsymbol{\sigma}_r(\tau)\,d\tau. \quad (6)$$

To provide a specific example, we consider a composite with a linear viscoelastic isotropic matrix which behaves like a standard linear solid under uniaxial tension and exhibits only elastic compressibility. For this material, the axial tensile creep compliance function is given by

$$J_x^m(t)=\frac{1}{E_m}+\frac{1}{E_v}\left(1-\exp\left(-\frac{E_v}{\eta_v}t\right)\right) \quad (7)$$

where E_m is the matrix Young's modulus, E_v is the Young's modulus of the elastic spring inside the Kelvin element, and η_v is the coefficient of viscosity for the Kelvin element.

The lateral creep compliance function, derived in Appendix A, is given by

$$J_y^m(t)=\frac{q_0^\nu}{q_0^E}\left[1+\left(\frac{E_v q_1^\nu}{\eta_v q_0^\nu}-1\right)\exp\left(-\frac{E_v}{\eta_v}t\right)\right] \quad (8)$$

where

$$q_0^E=\frac{E_m E_v}{E_m+E_v}\quad q_0^\nu=\frac{1}{2}-\frac{E_m E_v}{6K(E_m+E_v)}\quad q_1^\nu=\frac{3\eta_v K-E_m\eta_v}{6K(E_m+E_v)}. \quad (9)$$

and K is the matrix bulk modulus.

The matrix creep compliance for multiaxial state of stress then becomes

$$\boldsymbol{J}_m=\begin{pmatrix} J_x^m & -J_y^m & -J_y^m & 0 & 0 & 0\\ -J_y^m & J_x^m & -J_y^m & 0 & 0 & 0\\ -J_y^m & -J_y^m & J_x^m & 0 & 0 & 0\\ 0 & 0 & 0 & 2(J_x^m+J_y^m) & 0 & 0\\ 0 & 0 & 0 & 0 & 2(J_x^m+J_y^m) & 0\\ 0 & 0 & 0 & 0 & 0 & 2(J_x^m+J_y^m)\end{pmatrix}. \quad (10)$$

This form suggests that the Poisson's ratio, $\nu(t)=J_y^m/J_x^m$, is a time-dependent function. The initial elastic value is $\nu(0)=(3K-E_m)/6K$ and in the limit $t\to\infty$, $\nu(\infty)=q_0^\nu$, defined in (9_2).

2.2 Nonlinear Viscoelastic Relations. We consider a nonlinear viscoelastic model that consists of n nonlinear Kelvin elements and an elastic spring connected in series. The elastic component of the total strain is characterized by the spring Young's modulus, and the creep component by material constants of the Kelvin elements. Thus, the total strain is

$$\boldsymbol{\epsilon}=\boldsymbol{M}\boldsymbol{\sigma}+\boldsymbol{\epsilon}^{in} \quad (11)$$

where $\boldsymbol{M}$ is the elastic compliance matrix. For example, the model suggested by Ellyin [16] for nonlinear viscoelastic behavior of EPON 828 epoxy provides the inelastic strain rate $\dot{\boldsymbol{\epsilon}}^{in}$ in the form

$$\dot{\boldsymbol{\epsilon}}^{in}=\sum_{s=1}^{n}\dot{\boldsymbol{\epsilon}}^{cs}=\sum_{s=1}^{n}\left[a_s\sigma_{eq}^{\alpha_s-1}\boldsymbol{V}\boldsymbol{\sigma}-b_s\boldsymbol{\epsilon}^{cs}\right] \quad (12)$$

where $\boldsymbol{\epsilon}^{cs}$ is the creep strain of sth Kelvin element, a_s, α_s, b_s are material constants determined from experimental data, σ_{eq} is the equivalent stress, $\sigma_{eq}=(3/2s_{ij}s_{ij})^{1/2}$ and $s_{ij}=\sigma_{ij}-1/3\sigma_{kk}\delta_{ij}$ is the deviatoric stress tensor. The matrix $\boldsymbol{V}$ is given as

$$\boldsymbol{V}=\begin{pmatrix} 1 & -\nu & -\nu & 0 & 0 & 0\\ -\nu & 1 & -\nu & 0 & 0 & 0\\ -\nu & -\nu & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 2(1+\nu) & 0 & 0\\ 0 & 0 & 0 & 0 & 2(1+\nu) & 0\\ 0 & 0 & 0 & 0 & 0 & 2(1+\nu)\end{pmatrix}. \quad (13)$$

If all $\alpha_s=1$, the material is linear viscoelastic, and the form of matrix $\boldsymbol{V}$ implies that the Poisson's ratio for the current constitutive model remains constant and equal to that of the elastic matrix material. All Kelvin elements are assumed to have the same Poisson's ratio.

For the sth Kelvin element we have

$$\dot{\boldsymbol{\epsilon}}^{cs}=a_s\sigma_{eq}^{\alpha_s-1}\boldsymbol{V}\boldsymbol{\sigma}-b_s\boldsymbol{\epsilon}^{cs} \quad (14)$$

which can be rewritten as

$$e^{-b_s t}\,\frac{d}{dt}\left(\boldsymbol{\epsilon}^{cs}e^{b_s t}\right)=a_s\sigma_{eq}^{\alpha_s-1}\boldsymbol{V}\boldsymbol{\sigma}. \quad (15)$$

Therefore,

$$\boldsymbol{\epsilon}^{cs}=a_s e^{-b_s t}\int_0^t\sigma_{eq}^{\alpha_s-1}\boldsymbol{V}\boldsymbol{\sigma}e^{b_s\tau}d\tau. \quad (16)$$

For n Kelvin elements connected in series, the total creep (inelastic) strain is

$$\boldsymbol{\epsilon}^{in}=\sum_{s=1}^{n} a_s e^{-b_s t}\int_0^t \sigma_{eq}^{\alpha_s-1}\boldsymbol{V}\boldsymbol{\sigma}e^{b_s\tau}d\tau. \tag{17}$$

The inelastic strain rate is obtained either by direct differentiation of (17) with respect to time, or by substituting (16) into (12),

$$\dot{\boldsymbol{\epsilon}}^{in}=\sum_{s=1}^{n}\left[a_s\sigma_{eq}^{\alpha_s-1}\boldsymbol{V}\boldsymbol{\sigma}-b_s a_s e^{-b_s t}\int_0^t \sigma_{eq}^{\alpha_s-1}\boldsymbol{V}\boldsymbol{\sigma}e^{b_s\tau}d\tau\right]. \tag{18}$$

Note that in the present models the time-dependent inelastic strains (5) and (17) are reversible, i.e., after unloading the strains are completely recovered with time. The tests performed by Xia and Ellyin [13] on EPON 828 polymers support the fact that the epoxy resin is a viscoelastic material with no plastic permanent strains generated during loading process.

2.3 Temperature Dependence of Constitutive Relations. The creep compliance matrix in (1) and the constitutive relations (17) are defined at a fixed reference temperature θ_o, usually chosen as a room temperature. However, the temperature dependence of the constitutive relations may not always be neglected.

In thermorheologically simple materials, the time-temperature principle reflects the effect of a constant temperature change on all time-dependent material properties by a uniform shift in time scale. Thus, the creep compliance in (1) at a constant temperature θ is obtained from the creep compliance for the reference temperature θ_o by changing the time scale,

$$\boldsymbol{J}_r(t,\theta)=\boldsymbol{J}_r\left(\frac{t}{a_T(\theta)},\theta_o\right) \tag{19}$$

where $a_T(\theta)$ is a shift factor dependent on temperature. For a nonconstant temperature change the reduced time variable is introduced as

$$\xi(t)=\int_0^t\frac{dt'}{a_T[\theta(t')]} \tag{20}$$

and

$$\boldsymbol{J}_r(t,\theta)=\boldsymbol{J}_r(\xi(t),\theta_o). \tag{21}$$

For the nonlinear material considered in Section 2.2 a similar change of variables can be performed. For a constant temperature change θ, the inelastic strain (17) is obtained from the inelastic strain at reference temperature θ_o by replacing t with $t/a_T(\theta)$ in the exponential functions in (17) and by multiplying the result by $1/a_T(\theta)$,

$$\boldsymbol{\epsilon}^{in}=\sum_{s=1}^{n}\frac{a_s}{a_T}e^{-b_s t/a_T}\int_0^t\sigma_{eq}^{\alpha_s-1}\boldsymbol{V}\boldsymbol{\sigma}e^{b_s\tau/a_T}d\tau. \tag{22}$$

The inelastic strain rate then becomes,

$$\dot{\boldsymbol{\epsilon}}^{in}=\sum_{s=1}^{n}\left[\frac{a_s}{a_T}\sigma_{eq}^{\alpha_s-1}\boldsymbol{V}\boldsymbol{\sigma}-b_s\frac{a_s}{a_T^2}e^{-b_s t/a_T}\int_0^t\sigma_{eq}^{\alpha_s-1}\boldsymbol{V}\boldsymbol{\sigma}e^{b_s\tau/a_T}d\tau\right] \tag{23}$$

3 Transformation Field Analysis of Composite Systems

3.1 A Fibrous Ply. Consider a representative volume element consisting of many perfectly bonded phases $r=1,2,\ldots N$. The RVE is subjected to either uniform tractions $\boldsymbol{\sigma}(t)$, or to displacement boundary conditions derived from a uniform strain field $\boldsymbol{\epsilon}(t)$ on its external boundary, and to a temperature history $\theta(t)$. Response of the phases can be represented by various time-dependent constitutive relations, such as (1), (11), and (17). At any time t these relations admit the additive decomposition

$$\boldsymbol{\epsilon}_r(t)=\boldsymbol{M}_r\boldsymbol{\sigma}_r(t)+\boldsymbol{\mu}_r(t)\quad \boldsymbol{\sigma}_r(t)=\boldsymbol{L}_r\boldsymbol{\epsilon}_r(t)+\boldsymbol{\lambda}_r(t) \tag{24}$$

where $\boldsymbol{L}_r$ and $\boldsymbol{M}_r=\boldsymbol{L}_r^{-1}$ are the elastic phase stiffness and compliance matrices, $\boldsymbol{\mu}_r$ and $\boldsymbol{\lambda}_r$ are the average phase eigenstrain and eigenstress, referred to as transformation fields. These may include contributions of distinct physical origin, e.g., thermal strain (stress) and different kinds of inelastic strains, such as viscoelastic and anelastic deformations. More precisely, (24) implies that if at time t, a material point within the phase r were subjected to instantaneous unloading to zero stress, then the remaining strain would be equal to the local eigenstrain. If the total strains are small, then the local eigenstrains and relaxation stresses can be superimposed as

$$\boldsymbol{\mu}_r=\boldsymbol{m}_r\theta(t)+\boldsymbol{\epsilon}_r^{in}(t)+\ldots\qquad \boldsymbol{\lambda}_r=\boldsymbol{l}_r\theta(t)+\boldsymbol{\sigma}_r^{re}(t)+\ldots \tag{25}$$

where $\boldsymbol{m}_r$ is the vector of linear thermal expansion coefficients, and $\boldsymbol{l}_r$ is the thermal stress vector. The following interrelations hold:

$$\begin{aligned}&\boldsymbol{m}_r=-\boldsymbol{M}_r\boldsymbol{l}_r\quad \boldsymbol{\epsilon}_r^{in}(t)=-\boldsymbol{M}_r\boldsymbol{\sigma}_r^{re}(t)\\ &\boldsymbol{l}_r=-\boldsymbol{L}_r\boldsymbol{m}_r\quad \boldsymbol{\sigma}_r^{re}(t)=-\boldsymbol{L}_r\boldsymbol{\epsilon}_r^{in}(t).\end{aligned} \tag{26}$$

At any time t, the applied overall stress $\boldsymbol{\sigma}(t)$ or strain $\boldsymbol{\epsilon}(t)$, and the accumulated local transformation fields can be regarded as two separate loading agencies that both contribute to the local stresses in the composite system. This is expressed in the form

$$\boldsymbol{\sigma}_r(t)=\boldsymbol{B}_r\boldsymbol{\sigma}(t)+\sum_{s=1}^{N}\boldsymbol{F}_{rs}\boldsymbol{\lambda}_s(t)\qquad \boldsymbol{\epsilon}_r(t)=\boldsymbol{A}_r\boldsymbol{\epsilon}(t)+\sum_{s=1}^{N}\boldsymbol{D}_{rs}\boldsymbol{\mu}_s(t) \tag{27}$$

where $\boldsymbol{B}_r$, $\boldsymbol{A}_r$ are the elastic mechanical stress and strain concentration factors, and $\boldsymbol{F}_{rs}$, $\boldsymbol{D}_{rs}$ are the eigenstress and eigenstrain concentration factors. Particularly simple expressions are available for two-phase media. According to Dvorak [17], with phases denoted as $r=\alpha,\beta$,

$$\begin{aligned}&\boldsymbol{D}_{r\alpha}=(\boldsymbol{I}-\boldsymbol{A}_r)(\boldsymbol{L}_\alpha-\boldsymbol{L}_\beta)^{-1}\boldsymbol{L}_\alpha\quad \boldsymbol{D}_{r\beta}=-(\boldsymbol{I}-\boldsymbol{A}_r)(\boldsymbol{L}_\alpha-\boldsymbol{L}_\beta)^{-1}\boldsymbol{L}_\beta\\ &\boldsymbol{F}_{r\alpha}=(\boldsymbol{I}-\boldsymbol{B}_r)(\boldsymbol{M}_\alpha-\boldsymbol{M}_\beta)^{-1}\boldsymbol{M}_\alpha\\ &\boldsymbol{F}_{r\beta}=-(\boldsymbol{I}-\boldsymbol{B}_r)(\boldsymbol{M}_\alpha-\boldsymbol{M}_\beta)^{-1}\boldsymbol{M}_\beta.\end{aligned} \tag{28}$$

By taking a time derivative of (27_1) and using interrelations (26) one can obtain a system of equations for evaluation the local stresses in the composite system,

$$\dot{\boldsymbol{\sigma}}_r(t)=\boldsymbol{B}_r\dot{\boldsymbol{\sigma}}(t)-\sum_{s=1}^{N}\boldsymbol{F}_{rs}\boldsymbol{L}_s\{\boldsymbol{m}_s\dot{\theta}(t)+\dot{\boldsymbol{\epsilon}}_s^{in}(t)\}. \tag{29}$$

The inelastic strain rate $\dot{\boldsymbol{\epsilon}}_s^{in}(t)$ is connected to the stress history $\boldsymbol{\sigma}_s(t)$ through a constitutive relation given, for example, by (6) or (18). Appropriate replacement of variables should be made whenever the mechanical properties are dependent on temperature, in accordance with (21) or (23).

3.2 A Symmetric Laminated Plate. We now proceed to incorporate the inelastic phase constitutive relations into laminated plate analysis, and develop a system of equations for incremental evaluation of the local stresses in the fiber and matrix of each ply under a prescribed laminate loading history. The local stress fields then serve in evaluation of the deformation histories of the phases, plies, and laminate. For simplicity, we limit our present attention to symmetric laminates loaded by uniform in-plane tractions and temperature changes, although analogous TFA procedures can be relatively easily developed in the context of more general lamination theories or finite element programs ([19]). The ply and laminate are then analyzed only in plane stress.

We denote the (3×3) matrices describing the plane stress response of individual plies and laminate by boldface Roman capi-

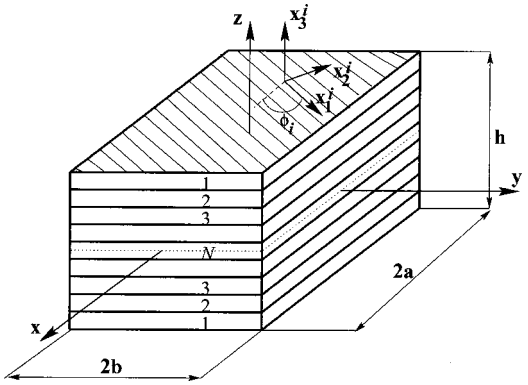


Fig. 1 Geometry of an element of a symmetric laminated plate

tals, e.g., **S**, **Q**, **T**, **R**, to distinguish them from the (6×6) matrices, ***L***, ***M***, ***D***, ***F*** describing phase and/or composite response in Sections 2 and 3.1. The (3×1) vectors are denoted by Greek boldface leters with the ply number subscript *i* or *j*. In constrast, except as noted, the (6×1) vectors, while also denoted by the Greek boldface letters, have a phase designation subscript, for which we use the letters *r*, *s*, *f*, or *m*. Quantities originally defined in ply coordinates and then transformed into global coordinates will be denoted by top bar.

Consider a composite laminate consisting of *N* pairs of unidirectional plies arranged in a symmetric layup with respect to the *xy*-midplane in the global coordinate system (*x*, *y*, *z*), Fig. 1. Assume that each ply is a two-phase composite consisting of a viscoelastic matrix and aligned fibers. Local coordinates of the *i*th pair of plies are denoted as (x_1^i, x_2^i, x_3^i); the x_1^i-direction coincides with that of the fiber axis, and $x_3^i \| z$. The composite laminate of total thickness *h* is subjected to overall in-plane stresses, defined in global coordinates as

$$\frac{1}{h}\mathbf{N}(t) = \frac{1}{h}\{N_x(t) N_y(t) N_{xy}(t)\}^T \tag{30}$$

and temperature history $\theta(t)$. Note that (30) represents the usual in-plane force resultant divided by the total laminate thickness.

Moreover, as suggested by (24) and (25), certain eigenstrains $\boldsymbol{\mu}_r^i$ and eigenstresses $\boldsymbol{\lambda}_r^i$ can be generated within each ply, for example, by inelastic deformation and thermal change. Their (6×1) ply averages are derived from the corresponding phase components and the elastic local field averages, using the generalized form ([18]) of the Levin's formula ([20])

$$\boldsymbol{\mu} = \sum_{r=1}^{M} c_r^i \boldsymbol{B}_r^{iT} \boldsymbol{\mu}_r^i \quad \boldsymbol{\lambda} = \sum_{r=1}^{M} c_r^i \boldsymbol{A}_r^{iT} \boldsymbol{\lambda}_r^i \tag{31}$$

where the $\boldsymbol{B}_r$ and $\boldsymbol{A}_r$ are the (6×6) mechanical concentration factors in (27); *M* is the number of phases in a ply, usually $M=2$. Quantities pertaining to the phase *r* of the *i*th pair of plies are assigned subscript *r* and superscript *i*, thus c_r^i is the volume fraction of phase *r* in the *i*th pair, and $\boldsymbol{\mu}_r^i$ is the eigenstrain generated in phase *r* of the *i*th pair of plies.

The summation in (31) is first carried out and the resulting (6×1) vectors are reduced by deleting rows 3, 4, and 5. In particular, the in-plane ply eigenstrains and eigenstresses in the *i*th pair of plies are obtained from (31) in the form

$$\boldsymbol{\mu}_i = \mathcal{U}^T \sum_{r=1}^{M} c_r^i \boldsymbol{B}_r^{iT} \boldsymbol{\mu}_r^i \quad \boldsymbol{\lambda}_i = \mathcal{U}^T \sum_{r=1}^{M} c_r^i \boldsymbol{A}_r^{iT} \boldsymbol{\lambda}_r^i \tag{32}$$

where

$$\mathcal{U}^T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{33}$$

All other (6×1) vectors derived from the RVE analysis in Section 3.1 and describing ply response are reduced as in (32).

Under the in-plane loading of a symmetric laminate, both the elastic and inelastic strains will be of the same magnitude in each pair of plies symmetric about the laminate midplane. Also, thermal strains induced by a uniform change in temperature will posses this symmetry. Therefore, the constitutive relations describing the response of a ply can be written either in the ply or laminate coordinate system as

$$\boldsymbol{\epsilon}_i = \mathbf{S}_i \boldsymbol{\sigma}_i + \boldsymbol{\mu}_i \quad \boldsymbol{\sigma}_i = \mathbf{Q}_i \boldsymbol{\epsilon}_i + \boldsymbol{\lambda}_i \tag{34}$$

$$\overline{\boldsymbol{\epsilon}}_i = \overline{\mathbf{S}}_i \overline{\boldsymbol{\sigma}}_i + \overline{\boldsymbol{\mu}}_i \quad \overline{\boldsymbol{\sigma}}_i = \overline{\mathbf{Q}}_i \overline{\boldsymbol{\epsilon}}_i + \overline{\boldsymbol{\lambda}}_i \tag{35}$$

where $\mathbf{Q}_i$ and $\mathbf{S}_i = \mathbf{Q}_i^{-1}$ are the (3×3) plane stress stiffness and compliance matrices of the *i*th pair of plies, and $\boldsymbol{\lambda}_i = -\mathbf{Q}_i \boldsymbol{\mu}_i$ is the ply eigenstress. Individual terms in (34) and (35) are connected by transformations between ply and global coordinate systems,

$$\begin{gathered} \overline{\mathbf{Q}}_i = \mathbf{R}_i^T \mathbf{Q}_i \mathbf{R}_i \quad \overline{\mathbf{S}}_i = \mathbf{T}_i^T \mathbf{S}_i \mathbf{T}_i \\ \mathbf{Q}_i = \mathbf{T}_i \overline{\mathbf{Q}}_i \mathbf{T}_i^T \quad \mathbf{S}_i = \mathbf{R}_i \overline{\mathbf{S}}_i \mathbf{R}_i^T \\ \overline{\boldsymbol{\epsilon}}_i = \mathbf{T}_i^T \boldsymbol{\epsilon}_i \quad \overline{\boldsymbol{\sigma}}_i = \mathbf{R}_i^T \boldsymbol{\sigma}_i \\ \boldsymbol{\epsilon}_i = \mathbf{R}_i \overline{\boldsymbol{\epsilon}}_i \quad \boldsymbol{\sigma}_i = \mathbf{T}_i \overline{\boldsymbol{\sigma}}_i . \end{gathered} \tag{36}$$

The transformation matrices are defined as

$$\mathbf{T}_i = \begin{pmatrix} c^2 & s^2 & 2cs \\ s^2 & c^2 & -2cs \\ -cs & cs & c^2 - s^2 \end{pmatrix} \quad \mathbf{R}_i = [\mathbf{T}_i^{-1}]^T$$
$$c = \cos(\phi_i) \quad s = \sin(\phi_i) \tag{37}$$

and the angle ϕ_i is measured counterclockwise from the laminate 0-deg direction *x* to the x_1^i fiber direction in the *i*th ply, Fig. 1.

Following the transformations (36) of the reduced vectors, the terms entering (35) are

$$\begin{gathered} \overline{\boldsymbol{\epsilon}}_i = \{\epsilon_x^i \ \ \epsilon_y^i \ \ \epsilon_{xy}^i\}^T \quad \overline{\boldsymbol{\sigma}}_i = \{\sigma_x^i \ \ \sigma_y^i \ \ \sigma_{xy}^i\}^T \\ \overline{\boldsymbol{\mu}}_i = \{\mu_x^i \ \ \mu_y^i \ \ \mu_{xy}^i\}^T \quad \overline{\boldsymbol{\lambda}}_i = \{\lambda_x^i \ \ \lambda_y^i \ \ \lambda_{xy}^i\}^T . \end{gathered} \tag{38}$$

As in (25), both thermal and inelastic components are included into the ply transformation fields,

$$\overline{\boldsymbol{\mu}}_i = \overline{\boldsymbol{\alpha}}_i \theta(t) + \overline{\boldsymbol{\epsilon}}_i^{in}(t) + \ldots \quad \overline{\boldsymbol{\lambda}}_i = \overline{\boldsymbol{\beta}}_i \theta(t) + \overline{\boldsymbol{\sigma}}_i^{e}(t) + \ldots \tag{39}$$

where $\boldsymbol{\alpha}_i = \{\alpha_1 \alpha_2 \alpha_{12}\}^T$ is a (3×1) vector of in-plane linear thermal expansion coefficients of the ply and $\boldsymbol{\beta}_i = -\mathbf{Q}_i \boldsymbol{\alpha}_i$ is the ply thermal stress vector. Also,

$$\overline{\boldsymbol{\mu}}_i = -\overline{\mathbf{S}}_i \overline{\boldsymbol{\lambda}}_i \quad \overline{\boldsymbol{\lambda}}_i = -\overline{\mathbf{Q}}_i \overline{\boldsymbol{\mu}}_i . \tag{40}$$

Since no through-the-thickness constraints are imposed on the laminate, there are out-of-plane ply strains caused by the in-plane stresses in each ply. These thickness strains can be evaluated for the transversely isotropic ply of elastic compliance M_{ij} as

$$\begin{gathered} \epsilon_3^i = M_{31}^i \sigma_1^i + M_{32}^i \sigma_2^i + \mu_3^i \\ \epsilon_4^i = \mu_4^i \quad \epsilon_5^i = \mu_5^i . \end{gathered} \tag{41}$$

If the laminate is subjected to the in-plane stress components in (30), then $\mu_4^i = \mu_5^i = 0$ and the ϵ_3^i in each ply can be integrated through the thickness to find the total $\boldsymbol{\epsilon}_3 = \overline{\boldsymbol{\epsilon}_3}$ component of the laminate strain.

The average ply stresses caused by mechanical loads $\mathbf{N}(t)$ (30) and ply transformation fields $\boldsymbol{\mu}_i$, $\boldsymbol{\lambda}_i$ (32) are evaluated in terms of certain stress distribution factors,

$$\overline{\boldsymbol{\sigma}}_i = \mathbf{H}_i \mathbf{N}/h + \sum_{j=1}^{N} \mathbf{K}_{ij} \overline{\boldsymbol{\lambda}}_j \tag{42}$$

where $\mathbf{H}_i$, $\mathbf{K}_{ij}$ are the stress and eigenstress distribution factors, respectively,

$$\mathbf{H}_i = \overline{\mathbf{Q}}_i \mathbf{Q}^{-1} \quad \mathbf{Q} = \sum_{i=1}^{N} c_i \overline{\mathbf{Q}}_i \quad \mathbf{K}_{ij} = \delta_{ij}\mathbf{I} - c_j \mathbf{H}_i \tag{43}$$

where c_i is the volume fraction of the ith pair of plies, and $\mathbf{Q}$ is the overall in-plane stiffness of the laminate. Derivation of the distribution factors is presented in Appendix B. Note that the first term in (42) is the ply stress in an elastic laminate, and the second term represents the contribution caused by both thermal and inelastic strains in all plies to the stresses in the ith pair of plies; the distribution factor $\mathbf{K}_{ij}$ includes the self-induced contribution by $\boldsymbol{\lambda}_i$.

This completes evaluation of the average ply stresses in the laminate. These stresses are now transformed into local coordinates in each ply. Also, in the last term of (42) the average ply eigenstress is replaced by ply eigenstrain, using (40_2), and the thermal parts are separated according to (39_1),

$$\overline{\boldsymbol{\sigma}}_i = \mathbf{H}_i \mathbf{N}/h - \sum_{j=1}^{N} \mathbf{K}_{ij} \overline{\mathbf{Q}}_j [\overline{\boldsymbol{\alpha}}_j \theta(t) + \overline{\boldsymbol{\epsilon}}_j^{in}(t)]. \tag{44}$$

Since ply response can be evaluated only in terms of the constituent responses, the ply stresses (44) are now augmented by zeros in rows 3, 4, and 5, and included as overall stress components acting on a representative volume of the composite material of the ply. This involves (27_1) and (29), and provides the following equations for phase stresses and their rates, caused in each ith pair of plies by the thermomechanical loading applied to the laminate,

$$\boldsymbol{\sigma}_r^i + \sum_{s=1}^{M} \boldsymbol{F}_{rs}^i \boldsymbol{L}_s^i \{\boldsymbol{\epsilon}_s^{in}\}^i + \boldsymbol{B}_r^i \mathcal{U} \mathbf{T}_i \sum_{j=1}^{N} \mathbf{K}_{ij} \overline{\mathbf{Q}}_j \mathbf{T}_j^T \boldsymbol{\epsilon}_j^{in}$$
$$= \boldsymbol{B}_r^i \mathcal{U} \mathbf{T}_i \left[\mathbf{H}_i \mathbf{N}/h - \sum_{j=1}^{N} \mathbf{K}_{ij} \overline{\mathbf{Q}}_j \mathbf{T}_j^T \boldsymbol{\alpha}_j \theta(t) \right] - \sum_{s=1}^{M} \boldsymbol{F}_{rs}^i \boldsymbol{L}_s^i \boldsymbol{m}_s^i \theta(t) \tag{45}$$

$$\dot{\boldsymbol{\sigma}}_r^i + \sum_{s=1}^{M} \boldsymbol{F}_{rs}^i \boldsymbol{L}_s^i \{\dot{\boldsymbol{\epsilon}}_s^{in}\}^i + \boldsymbol{B}_r^i \mathcal{U} \mathbf{T}_i \sum_{j=1}^{N} \mathbf{K}_{ij} \overline{\mathbf{Q}}_j \mathbf{T}_j^T \dot{\boldsymbol{\epsilon}}_j^{in}$$
$$= \boldsymbol{B}_r^i \mathcal{U} \mathbf{T}_i \left[\mathbf{H}_i \dot{\mathbf{N}}/h - \sum_{j=1}^{N} \mathbf{K}_{ij} \overline{\mathbf{Q}}_j \mathbf{T}_j^T \boldsymbol{\alpha}_j \dot{\theta}(t) \right] - \sum_{s=1}^{M} \boldsymbol{F}_{rs}^i \boldsymbol{L}_s^i \boldsymbol{m}_s^i \dot{\theta}(t). \tag{46}$$

The phase inelastic strains $\boldsymbol{\epsilon}_s^{in}$ in the second term are related to local phase stress averages by the constitutive relations (6) or (18). The inelastic strain $\boldsymbol{\epsilon}_j^{in}$ in the third term is a (3×1) vector obtained from (31), as a (6×1) vector in terms of local phase inelastic strains $\{\boldsymbol{\epsilon}_r^{in}\}^j$ in the jth pair of plies, and then reduced according to (32). The same reduction procedure needs to be followed in evaluating the $\boldsymbol{\alpha}_j$ in the second term on the right-hand side. Replacing the inelastic strain components by the local phase stresses, c.f., (5), (6), (17), (18), changes (45) and (46) into a system of equations for incremental evaluation of local phases stress averages in all plies.

Note that the mechanical and transformation concentration factors are connected to material constants only through phase elastic moduli which appear in (3) or (11), and are assumed to be time-independent. Note also that all eigenstrain contributions, both thermal and inelastic, originate in individual matrix and fiber phases of each ply of the laminate, and that they enter the phase stress equations in each ply both directly, in the second left-hand term of (45) and (46), and indirectly in the third term, which describes contribution to phase stresses in the ply i caused by all average ply eigenstrains.

4 Solution Procedure

Integration of (46) starts from known initial conditions at $t=0$. In the case of prestress release, these are evaluated by applying to the laminate the in-plane force resultant caused by prestress release in all plies. The average stress applied to the prestressed fibers of the ith pair of plies is specified in the local ply coordinates as

$$\boldsymbol{\sigma}_f^P(0) = \{(\sigma_{11}^f)_P^i(0) \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0\}^T. \tag{47}$$

Release of this prestress is equivalent to applying to the ply-free edges intersecting the fibers initial compressive stress,

$$\mathbf{N}^P(0)/h = -\sum_{i=1}^{N} c_i \mathbf{R}_i^T \{c_f^i (\sigma_{11}^f)_P^i(0) \;\; 0 \;\; 0\}^T \tag{48}$$

where c_i is the volume fraction of the ith pair of plies, c_f^i is the fiber volume fraction in that pair of plies. The stresses in the matrix and fibers immediately after prestress release are evaluated in the ply coordinates as

$$\boldsymbol{\sigma}_m^i(0) = \boldsymbol{B}_m^i \mathcal{U} \mathbf{T}_i \mathbf{H}_i \mathbf{N}^P(0)/h \quad \boldsymbol{\sigma}_f^i(0) = \boldsymbol{B}_f^i \mathcal{U} \mathbf{T}_i \mathbf{H}_i \mathbf{N}^P(0)/h + \boldsymbol{\sigma}_f^P(0). \tag{49}$$

The total ply stresses follow by superposition of the stresses caused by application of $\mathbf{N}^P(0)$ and initial ply tractions that support fiber prestress (47),

$$\boldsymbol{\sigma}_i(0) = \mathbf{T}_i \mathbf{H}_i \mathbf{N}^P(0)/h + \{c_f^i (\sigma_{11}^f)_P^i(0) \;\; 0 \;\; 0\}^T. \tag{50}$$

In the absence of direct mechanical loading, Eqs. (47)–(50) are the initial conditions of the stress relaxation process driven by prestress release. Equations (49) are readily recognized as the leading right-hand term in (45) at $t=0$, when there are no inelastic strains. Depending on the details of the fabrication procedure, (49) can be augmented to reflect a rapid change from the processing temperature, by including there the thermal loading terms from (27).

We adopt the method of numerical integration of Eqs. (46). If the phase stresses are known up to the current time t, the local fields at time $t+\Delta t$ are found by integration of (46), using for example the Runge-Kutta formula of order two. This involves evaluation of the integrands in (6) or (18) denoted symbolically as $\mathbf{f}(t-\tau, \boldsymbol{\sigma}_r(\tau))$.

The algorithm for evaluation of phase stresses can be summarized as follows:

Step 1: Select the number of time intervals, n, and compute time increment $\Delta t = (t_{n+1} - t_1)/n$, where $t_1 = 0$ and t_{n+1} defines the time interval in which the histories $\mathbf{N}(t)$ and $\theta(t)$ are defined.

Step 2: Find the initial values of the local stress fields $\boldsymbol{\sigma}_r^i(0)$.

Step 3: For $k = 1,2,\ldots,n$, do steps 4, 5, and 6.

Step 4: Evaluate $\dot{\boldsymbol{\sigma}}_r^i(t_k)$ from (46). If $k=1$ the integrals in (6) and (18) are zeros, for any other k they can be found using the trapezoidal rule,

$$\int_{0_+}^{t} \mathbf{f}(t-\tau, \boldsymbol{\sigma}_r(\tau)) d\tau = \frac{\Delta t}{2} \sum_{j=1}^{k-1} [\mathbf{f}(t_k - t_j, \boldsymbol{\sigma}_r(t_j)) + \mathbf{f}(t_k - t_{j+1}, \boldsymbol{\sigma}_r(t_{j+1}))]. \tag{51}$$

Step 5: Compute trial stresses at t_{k+1} as

$$\boldsymbol{\sigma}_r^{i^{tr}}(t_{k+1}) = \boldsymbol{\sigma}_r^i(t_k) + \Delta t \dot{\boldsymbol{\sigma}}_r^i(t_k) \tag{52}$$

and the stress rate from (46) at time t_{k+1} using adjusted stresses (52), denoted as

Table 2 Properties of EPON 828 matrix (Ellyin [16]) and AS4 carbon fibers (Blackketter and Upadhyaya [23])

Elastic				Viscoelastic			
E_m (GPa)	2.75	ν_{12}^f	0.2	b_1 (1/s)	0.015	α_1	4
ν_m	0.4	G_{23}^f (GPa)	5.5	b_2 (1/s)	0.0015	α_2	2
α_m (10^{-6}/°C) [23]	48	α_1^f (10^{-6}/°C)	−0.36	b_3 (1/s)	0.00015	α_3	2
E_1^f (GPa)	235	α_2^f (10^{-6}/°C)	18	a_1 (MPa^{-4}/s)	$2.1*10^{-10}$		
E_2^f (GPa)	14			a_2 (MPa^{-2}/s)	$4.7*10^{-9}$		
G_{12}^f (GPa)	28			a_3 (MPa^{-2}/s)	$5.6*10^{-9}$		

$$\dot{\boldsymbol{\sigma}}_r^{i*}(t_k)=\dot{\boldsymbol{\sigma}}_r^i(t_{k+1},\boldsymbol{\sigma}_r^{itr}(t_{k+1})). \tag{53}$$

Step 6: Finally, compute the stresses at time t_{k+1},

$$\boldsymbol{\sigma}_r^i(t_{k+1})=\boldsymbol{\sigma}_r^i(t_k)+\frac{\Delta t}{2}\{\dot{\boldsymbol{\sigma}}_r^i(t_k)+\dot{\boldsymbol{\sigma}}_r^{i*}(t_k)\}. \tag{54}$$

With the knowledge of local phase stresses $\boldsymbol{\sigma}_f^i$ and $\boldsymbol{\sigma}_m^i$, the phase inelastic strain $(\boldsymbol{\epsilon}_r^{in})^i$ in each pair of plies i can be evaluated from the constitutive relations (5) or (17). The inelastic ply strain $\boldsymbol{\epsilon}_i^{in}$ is obtained from Levin's formula (32_1). Total in-plane strains in each pair of plies at time t can be found from the constitutive relations (35_1) and decomposition of ply eigenstrains (39_1),

$$\overline{\boldsymbol{\epsilon}}_i(t)=\overline{\mathbf{S}}_i\overline{\boldsymbol{\sigma}}_i(t)+\overline{\boldsymbol{\alpha}}_i\theta(t)+\overline{\boldsymbol{\epsilon}}_i^{in}(t) \tag{55}$$

where the ply stress $\overline{\boldsymbol{\sigma}}_i$ is given by

$$\overline{\boldsymbol{\sigma}}_i(t)=\mathbf{R}_i^T\mathcal{U}^T[c_f^i\boldsymbol{\sigma}_f^i(t)+c_m^i\boldsymbol{\sigma}_m^i(t)] \tag{56}$$

and c_f^i, c_m^i are the phase volume fractions. Separating the ply eigenstrains from the total strains in (55) and substituting these into $(B7)$ provides the overall laminate eigenstrain. Together with the strains caused by the in-plane mechanical tractions, this yields the overall laminate strain caused by the combined thermomechanical loads and inelastic deformation in the form

$$\boldsymbol{\epsilon}(t)=\mathbf{SN}(t)/h+\sum_{i=1}^{N} c_i\mathbf{H}_i^T[\overline{\boldsymbol{\alpha}}_i\theta(t)+\overline{\boldsymbol{\epsilon}}_i^{in}(t)]. \tag{57}$$

The out-of-plane strain ϵ_3^i in each pair of plies can be evaluated from the ply constitutive relations and known in-plane stress components, as in (41). The overall out-of-plane strain of the laminate is found as

$$\overline{\epsilon_3}=\epsilon_3=\sum_{i=1}^{N} c_i\epsilon_3^i \tag{58}$$

where c_i is the ply volume fraction. Integration of the out-of-plane strains through the thickness of the plate appears to be adequate at some distance away from the free edges where strains and stresses become functions only of thickness coordinate z.

5 Applications

In the illustrative examples that follow, we examine the effect of viscoelastic matrix deformation on redistribution of ply stresses under three loading conditions. These include release of the optimized fiber prestress found in elastic laminates ([2]) followed by a relaxation period; subsequent loading of these laminates by proportional in-plane stresses at a prescribed rate; and sustained application of a uniform thermal change to a laminate without prestress. In all cases, fiber prestress is applied prior to and released after matrix cure and cooling to room temperature.

The material selected is the AS4/EPON 828 carbon/epoxy system, in two symmetric nine-layer laminate layups, a cross-ply $((0/90)_2/\overline{0})_s$ and a quasi-isotropic $(0/45/-45/90/\overline{0})_s$. The EPON 828 matrix is regarded as a nonlinearly viscoelastic solid conforming with (12)–(18); its selection was motivated by availability of experimentally based constitutive relations and material parameters ([16]), Table 2. The fiber is assumed to be elastic.

All results are displayed in the form of initial failure maps forming an internal envelope of ply strength branches, plotted in laminate biaxial normal stress plane. Procedures used to construct the failure maps were described by Dvorak and Suvorov [1]. The individual branches are derived from the critical stress criterion, hence they indicate the overall stress combinations that cause at least one of the ply stress components to reach the respective ply strength value, Table 3. $F_1^{t,c}$, $F_2^{t,c}$ denote axial and transverse ply strength in tension (t) or compression (c), and F_6–the strength in shear. Ply strength values used in plotting the initial damage envelopes are assumed to be independent of time and temperature. Ply elastic properties are estimated by the Mori-Tanaka method ([21]), Table 3. However, if experimentally measured elastic moduli of the AS4/EPON 828 composite were available, a better fit could be obtained with one of the more general methods ([22]). Evaluation of an estimate that conforms with available experimental data for an S-glass/epoxy composite is shown in [2].

In all laminates considered here, the initial elastic stress states are created by certain combinations of tension stress N_x/h, temperature change, and release of optimized fiber prestress that keeps both the interior and free-edge stresses at $y=\pm b$ within the ply strength limits, Table 3. The fiber prestress was obtained with the procedure described in [2], which minimizes an objective function representing the differences between the free-edge tractions applied to individual plies by prestress release and the corresponding ply stress components in the interior of the laminate. The resulting prestress magnitudes are listed in Table 4.

A simple failure map of a cross-ply AS4/EPON 828 laminate that has been cooled from curing temperature but not subjected to release of fiber prestress and any mechanical loads is shown by the thick black solid lines in Fig. 2. The branch labels $\sigma_{22}^{0(T)}$,

Table 3 Properties of AS4/EPON 828 system (elastic moduli are estimated by Mori-Tanaka method [21])

Property		Property	
c_f [24]	0.61	α_1 (10^{-6}/°C)	0.054
E_1 (GPa)	144.45	α_2 (10^{-6}/°C)	36.13
E_2 (GPa)	7.0	F_2^t (MPa) [24]	22.1
G_{12} (GPa)	3.57	F_6 (MPa) [24]	71.4
ν_{12}	0.274	F_2^c (MPa)	55.0
G_{23} (GPa)	2.33	F_1^t (MPa) [23]	2189.9
		F_1^c (MPa)	800.0

Table 4 Optimized fiber prestress magnitudes, $(\sigma_{11}^f)_P^i$

$((0/90)_2/\overline{0})_s$ carbon/epoxy at $N_x/h\|_{t=0}=836.1$ MPa and $\Delta\theta=-45$°C (a)	$(0/45/-45/90/\overline{0})_s$ carbon/epoxy at $N_x/h\|_{t=0}=508.1$ MPa and $\Delta\theta=-45$°C (b)
0° Plies: 2100 MPa 90° Plies: 470.8 MPa	0° Plies: 1950 MPa 90° Plies: 77.9 MPa 45° Plies: 62.7 MPa −45° Plies: 62.7 MPa

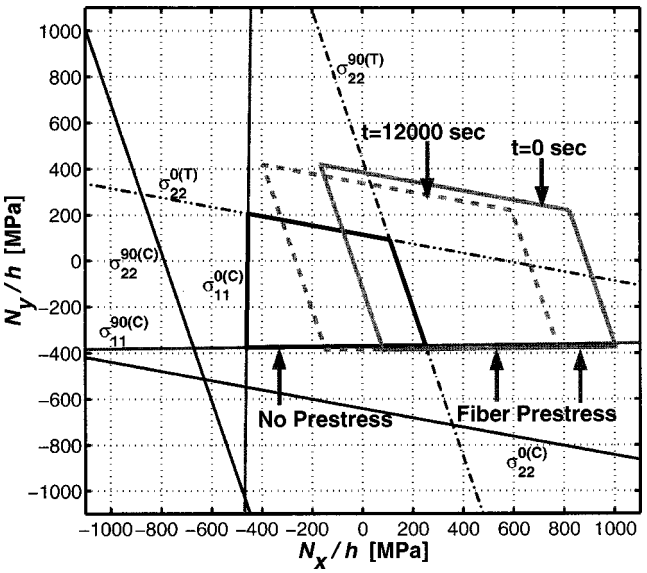


Fig. 2 Initial damage envelope for the $((0/90)_2/\bar{0})_s$ carbon/epoxy viscoelastic laminate before and after 12,000 sec stress relaxation. Initial fiber prestress is 2100 MPa in the 0-deg plies and 470.8 MPa in the 90-deg plies. Thermal change of $\Delta\theta = -45°C$ is applied at $t \geq 0$.

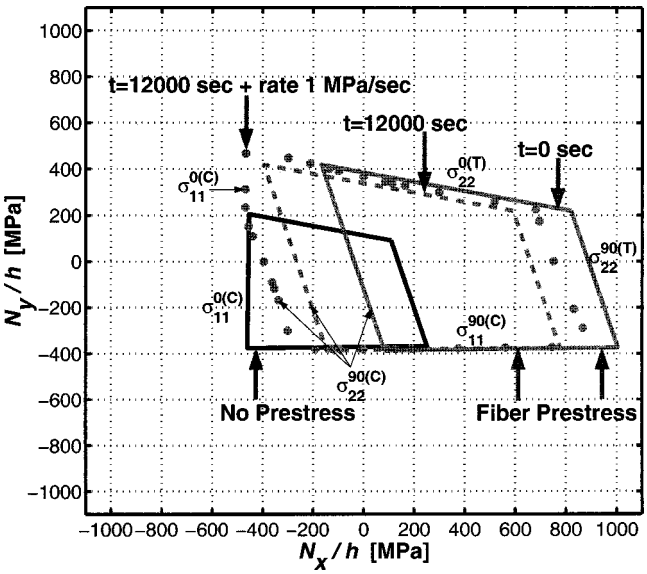


Fig. 4 Initial damage envelope for the $((0/90)_2/\bar{0})_s$ carbon/epoxy viscoelastic laminate before and after 12,000 sec stress relaxation. The overall stress is applied at the rate 1 MPa/sec. Optimized fiber prestress is 2100 MPa in the 0-deg plies and 470.8 MPa in the 90-deg plies. Thermal change of $\Delta\theta = -45°C$ is applied at $t \geq 0$.

$\sigma_{22}^{90(T)}$ denote laminate stress combinations causing onset of transverse cracking in the 0-deg or 90-deg plies; $\sigma_{11}^{0(C)}$ and $\sigma_{11}^{90(C)}$ indicate combinations causing onset of axial ply failure in compression. Additional $\sigma_{12}^{\pm 45}$ branches appear in Fig. 3 for a quasi-isotropic laminate, these indicate loading conditions causing onset of shear failure in the ± 45-deg plies.

In addition to the individual ply strength branches and the initial failure envelopes of laminates subjected only to cooling by $\Delta\theta = -45°C$ from curing temperature, Figs. 2 and 3 show initial failure envelopes of laminates that have been subjected to release

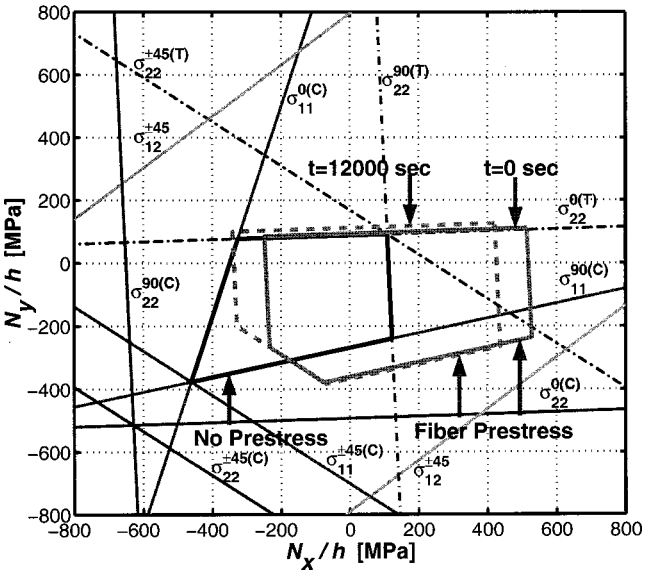


Fig. 3 Initial damage envelope for the $(0/45/-45/90/\bar{0})_s$ carbon/epoxy viscoelastic laminate before and after 12,000 sec stress relaxation. Optimized fiber prestress is 1950 MPa in the 0-deg plies, 77.9 MPa in the 90-deg plies, 62.7 MPa in the 45-deg plies and 62.7 MPa in the −45-deg plies. Thermal change of $\Delta\theta = -45°C$ is applied at $t \geq 0$.

of optimized fiber prestress, applied prior and during matrix cure ([2]). These are drawn in thick gray solid lines and labeled $t = 0$ sec, indicating absence of viscoelastic deformation. Relatively high prestress is applied in the 0-deg plies to maximize the overall normal stress N_x/h. The prestress applied to the $((0/90)_2/\bar{0})_s$ laminate (Table 4(*a*)) cancels all interlaminar stresses at the free edges when the laminate is subjected to the axial stress $N_x/h = 836.1$ MPa at time $t = 0$. The prestress indicated for the $(0/45/-45/90/\bar{0})_s$ laminate (Table 4(*b*)) minimizes interlaminar stresses at free edges when the axial stress $N_x/h = 508.1$ MPa is applied at $t = 0$.

The dashed thick gray lines in Figs. 2 and 3 show positions of the envelopes after stress relaxation lasting 12,000 sec. Numerical experiments have shown that this time is sufficient for revealing the effect of residual stress relaxation in the fiber and matrix on the final position of the failure envelopes in this material system; relaxation rates reached insignificant values within this period. The failure envelopes at 12,000 sec can be used in applications involving rapid mechanical loading of the prestressed laminate. Although cancelation/minimization of free-edge stresses by optimized prestress indicated in Table 4 is guaranteed only for $N_x/h = 836.1/508.1$ MPa applied at time $t = 0$, it is believed that this prestress may be beneficial in reduction of interlaminar stresses at time $t > 0$ as long as the loading vector stays within the relaxation-translated damage envelope.

The results suggest that both release of fiber prestress and subsequent stress relaxation cause translations of individual strength branches, in the overall stress direction that is parallel to the applied fiber prestress. The translation magnitudes are different for each branch, since prestress release and stress relaxation cause dissimilar changes in the ply stress components associated with specific failure modes. These differences are more pronounced in laminates made of the very strongly anisotropic plies employed herein, than in the S-glass/epoxy laminates studied by Dvorak and Suvorov [1]. These translations are seen to cause exchanges of the branches forming the internal envelopes, often leading to expansion of the failure envelopes after prestress release. Stress relaxation in the selected matrix material affects primarily the transverse and longitudinal shear response of the plies, as normal creep

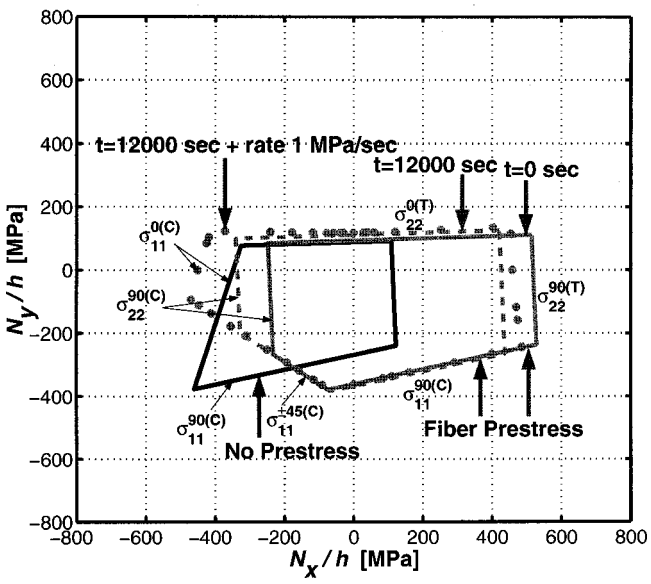


Fig. 5 Initial damage envelope for the $(0/45/-45/90/\bar{0})_s$ carbon/epoxy viscoelastic laminate before and after 12,000 sec stress relaxation. Overall stress is applied at the rate 1 MPa/sec. Optimized fiber prestress is 1950 MPa in the 0-deg plies, 77.9 MPa in the 90-deg plies, 62.7 MPa in the 45-deg plies and 62.7 MPa in the −45-deg plies. Thermal change of $\Delta\theta=-45°C$ is applied at $t\geqslant 0$.

rates are extremely small in the direction of the fibers. While we do not label the branches forming the prestress and relaxation-translated failure envelopes, their association to specific failure modes can be inferred by comparing their direction to those in the laminate without prestress.

Additional translation of the damage envelopes can also result from matrix stress relaxation under mechanical loading. This is shown in Figs. 4 and 5, where the envelopes are represented by end points of many applied proportional overall stress combinations, from zero initial values $N_x(12{,}000)=N_y(12{,}000)=0$, i.e., after the 12,000 sec stress relaxation following prestress release, at a constant rate $R=1$ MPa/sec. The displayed points indicate values of the applied stresses at initial ply failure; the failure mode is indicated by the proximate, parallel initial failure branches.

The positions reached by the newly translated branches can be better explained by reference to Fig. 6 that shows stress relaxation in the neat matrix subjected first to constant normal compressive strain for 12,000 sec, and then to either tensile or compressive constant strain rate. The chosen strain rate corresponds to the stress rate of about 1 MPa/sec. The initial compression apparently promotes faster stress relaxation under subsequent compressive strain rate, and nearly elastic response under a tensile rate.

This behavior helps in understanding the response of the laminates loaded by constant rate following the initial compression caused by prestress release, Figs. 4 and 5. In the initial part of a tensile loading path, the applied overall axial tension reduces the compressive transverse residual stress left in the 90-deg plies by prestress release. After that is reduced to zero, the overall tension generates transverse tension stresses in the 90-deg plies, until they reach the transverse ply strength. This results in a small translation of the strength branch $\sigma_{22}^{90(T)}$ toward higher laminate tensile stresses at initial failure of these plies. However, the residual and applied stresses are added under 0-deg compression, promoting greater relaxation of the transverse stress in the 90-deg plies, and thus causing a greater translation of the transverse compressive strength branch $\sigma_{22}^{90(C)}$ of the 90-deg ply. In fact, this branch has apparently translated so far (Fig. 4) that it has been partially replaced by the $\sigma_{11}^{0(C)}$ branch of the failure envelope. This effect is even more pronounced in Fig. 5, where the 90-deg transverse compression branch is no longer a part of the dotted failure envelope.

Figures 7 and 8 show first the initial envelopes of the laminates not subjected to any mechanical loads or temperature change. These are drawn in black solid lines. The figures also illustrate the effect of elevated temperature $\Delta\theta=100°C$ at time $t=0$ and after stress relaxation at $t=12{,}000$ sec on the position of the envelopes. It is assumed that the matrix material is thermorheologically simple and a shift factor $a_T\ll 1$ at $\Delta\theta=100°C$. Therefore, the time for thermal stress relaxation can be selected shorter than $t=12{,}000$ sec without producing noticeable changes in the position

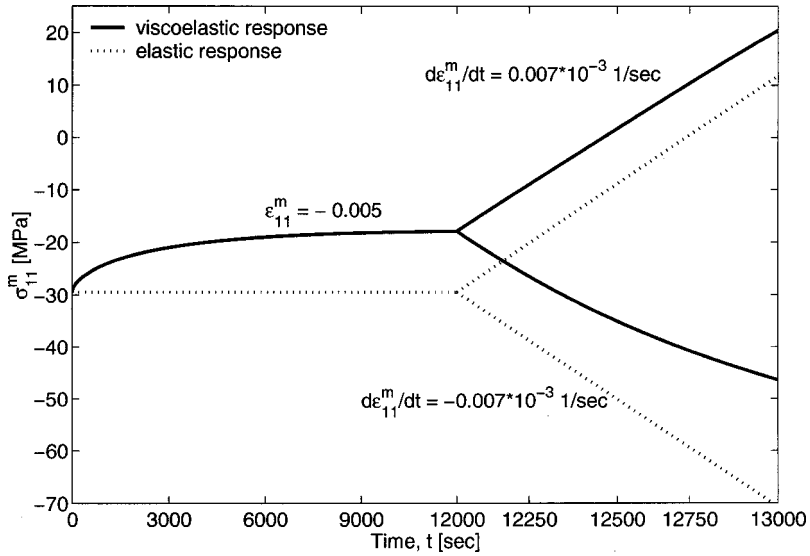


Fig. 6 Stress relaxation of the EPON 828 matrix material. The matrix is first subjected to constant axial strain $\epsilon_{11}^m=-0.005$. After 12,000 sec of stress relaxation strain is applied at a rate either $0.007*10^{-3}$ or $-0.007*10^{-3}$/sec, so that at time $t=13{,}000$ sec, the magnitude of axial strain is 0.002 or −0.012, respectively.

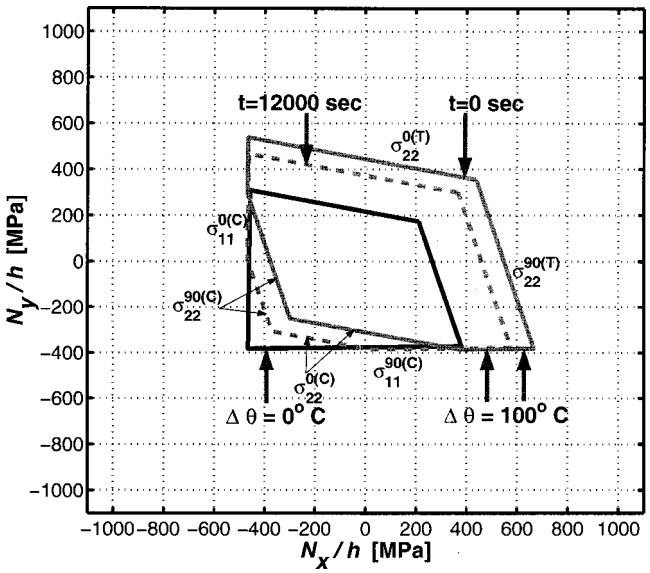


Fig. 7 Initial damage envelope for the $((0/90)_2/\bar{0})_s$ carbon/epoxy viscoelastic laminate before and after 12,000 sec stress relaxation. Thermal change of $\Delta\theta = 100°C$ is applied at $t \geqslant 0$.

of the envelopes. At time $t=0$, immediately after the temperature change, the translated envelope is drawn in a gray solid line, and after $t=12{,}000$ sec by a dashed gray line. No fiber prestress was applied in this case. Note the significant expansion of the failure envelope caused by dissimilar branch translations and interior position exchanges caused by the thermal stresses. Stress relaxation has a relatively small effect in this case.

Tables 5 and 6 summarize the magnitudes of important stress components reached in the plies and constituents of the two laminates at the several stages of loading history applied also in Figs. 4 and 5. The first two columns describe thermal and thermal plus residual stresses due to prestress release in elastic systems. The third column indicates changes in the second column values after stress relaxation lasting 12,000 sec; no mechanical or thermal loads are applied. Finally, the fourth columns presents stresses reached at the end of the single constant rate loading path $\dot{N}_x/h = 1$ MPa/sec.

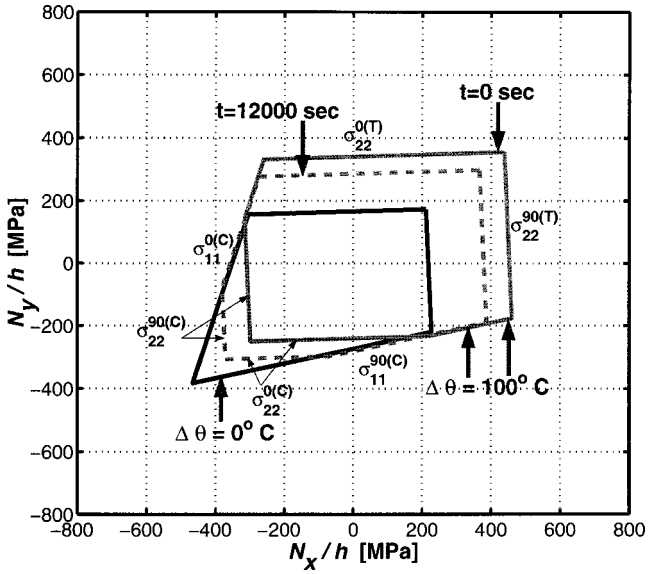


Fig. 8 Initial damage envelope for the $(0/45/-45/90/\bar{0})_s$ carbon/epoxy viscoelastic laminate before and after 12,000 sec stress relaxation. Thermal change of $\Delta\theta = 100°C$ is applied at $t \geqslant 0$.

Table 5 Average ply and fiber stresses in the $((0/90)_2/\bar{0})_s$ laminate

	$\Delta\theta$ $t=0$	$P_i+\Delta\theta$ $t=0$	$P_i+\Delta\theta$ $t=12{,}000$ sec	$N_x/h+P_i+\Delta\theta$ $t=12{,}752$ sec
0 ply: σ_1	−8.51	41.56	26.31	1335.9
0 ply: σ_{11}^f	−21.3	83.29	51.1	2179.5
0 ply: σ_2	10.47	−17.25	−12.52	2.13
90 ply: σ_1	−13.09	21.57	15.65	−2.66
90 ply: σ_{11}^f	−28.79	46.24	31.39	−11.2
90 ply: σ_2	10.63	−51.96	−32.89	22.1

$\Delta\theta = -45°C$, $N_x(12000)/h = 0$ MPa, $\dot{N}_x/h = 1$ MPa/sec
$P_i = (\sigma_{11}^f)_P^i$ 0 deg plies: 2100 MPa, 90 deg plies: 470.8 MPa

6 Closure

The theoretical part of the paper outlines the procedure to be followed in transformation field analysis of overall response and local ply and phase stress averages in viscoelastic composite laminates subjected to fiber prestress release, thermal and inelastic strains in the matrix and fiber in each ply, and to uniform thermomechanical loads applied as functions of time. This procedure takes into account interactions of the elastic, inelastic, and thermal deformations of individual phases within all plies of the laminate, and also interactions of the elastic strain and eigenstrain ply averages within the laminate. Two systems of equations, one for the time-dependent local stresses in the phases of all plies, and another one for their rates, are obtained for loading by overall mechanical tractions, release of fiber prestress, and local thermal and inelastic eigenstrains. Hierarchical interactions of these effects on the ply and laminate scales are analyzed with certain influence functions that depend only on the elastic properties of the fiber, matrix, plies, and laminate.

Results of several applications to specific loading programs are displayed by their effect on the position of ply initial failure branches derived from the critical stress criterion. It is shown that viscoelastic matrix deformation enables a similar response in the matrix-dominated deformation modes of a composite ply. These include transverse normal and longitudinal shear ply strains, but mostly excludes longitudinal normal strain associated with fiber-dominated deformation. The beneficial effect that fiber prestress

Table 6 Average ply and fiber stresses in the $(0/45/-45/90/\bar{0})_s$ laminate

	$\Delta\theta$ $t=0$	$P_i+\Delta\theta$ $t=0$	$P_i+\Delta\theta$ $t=12{,}000$ sec	$N_x/h+P_i+\Delta\theta$ $t=12{,}459$ sec
0 ply: σ_1	−7.1	274.7	253.93	1305.3
0 ply: σ_{11}^f	−19.0	454.84	418.25	2129.7
0 ply: σ_2	10.4	9.68	7.31	5.21
0 ply: σ_{12}	0.0	0.0	0.0	0.0
45 ply: σ_1	−10.95	−315.36	−315.8	48.9
45 ply: σ_{11}^f	−25.3	−514.91	−516.85	72.8
45 ply: σ_2	10.54	−10.32	−6.94	12.8
45 ply: σ_{12}	−0.19	28.02	18.58	−11.8
−45 ply: σ_1	−10.95	−315.36	−315.8	48.9
−45 ply: σ_{11}^f	−25.3	−514.91	−516.85	72.8
−45 ply: σ_2	10.54	−10.32	−6.94	12.8
−45 ply: σ_{12}	0.19	−28.02	−18.58	11.8
90 ply: σ_1	−14.81	255.12	274.58	−45.9
90 ply: σ_{11}^f	−31.6	417.83	450.58	−81.4
90 ply: σ_2	10.68	−30.32	−20.99	22.1
90 ply: σ_{12}	0.0	0.0	0.0	0.0

$\Delta\theta = -45°C$ $N_x(12000)/h = 0$ MPa $\dot{N}_x/h = 1$ MPa/sec
$P_i = (\sigma_{11}^f)_P^i$ 0 deg plies: 1950 MPa, 90 deg plies: 77.9 MPa,
+45 deg plies: 62.7 MPa, −45 deg plies: 62.7 MPa

may have on the position of initial failure envelope of a composite laminate is thus modified, and usually reduced. However, the reduction is relatively small, even in the present system with a fairly viscous matrix. The applied prestress was optimized with regard to allowable stresses supported by the free edges of elastic plies under simple tension applied to the laminate, as described in a companion paper ([2]). Therefore, the initial failure envelopes constructed at $t=0$, immediately after prestress release, define a laminate stress region which guarantees absence of damage both at and away from the free edges. Short-duration cyclic loads may also be applied within the elastic damage envelopes, providing that ply strength remains constant and viscous effect are limited. Since viscoelastic deformation promotes stress relaxation, one may speculate that the selected prestress may be admissible even within the strength constraints on free edge stresses at $t>0$.

In the normal stress plane, fiber prestress causes a substantial expansion of the failure envelopes, as well as their translation in the dominant prestress direction. The expansion appears to be present only in laminates with strongly anisotropic plies; it was negligible in the S-glass/epoxy systems studied in our earlier paper ([1]). Application of fiber prestress can therefore be used as an effective tool for damage prevention in this and others composite laminates.

Acknowledgment

This research was supported in part by the Solid Mechanics Program of the Engineering and Environmental Sciences Division of the U.S. Army Research Office under Agreement No. DAAG55-98-0464, and by the Ship Structures Division of the Office of Naval Research under Grant No. N0001140010090. Drs. Bruce LaMattina and Yapa D. S. Rajapakse served as program monitors.

Appendix A

If the material behaves like a standard linear solid under uniaxial tension and exhibits only elastic compressibility, its axial tensile creep compliance function is given by

$$J_x^m(t)=\frac{1}{E_m}+\frac{1}{E_v}\left(1-\exp\left(-\frac{E_v}{\eta_v}t\right)\right) \tag{A1}$$

where E_m is the matrix Young's modulus, E_v is the Young's modulus of the elastic spring inside the Kelvin element, and η_v is the coefficient of viscosity for the Kelvin element.

For the constant stress σ_{xx}^0 applied to the matrix, the axial and lateral strains are given as

$$\epsilon_{xx}(t)=\sigma_{xx}^o J_x^m(t) \quad \epsilon_{yy}(t)=-\sigma_{xx}^o J_y^m(t) \tag{A2}$$

where $J_y^m(t)$ is the lateral tensile creep compliance function. The creep Poisson function $\nu(t)$ is defined as the negative ratio of $\epsilon_{xx}(t)$ and $\epsilon_{yy}(t)$, i.e.,

$$\nu(t)=\frac{J_y^m(t)}{J_x^m(t)}. \tag{A3}$$

To find $J_y^m(t)$, we write the differential equations for the viscoelastic solid as

$$P^E\sigma_{xx}=Q^E\epsilon_{xx} \quad P^\nu\epsilon_{yy}=-Q^\nu\epsilon_{xx} \tag{A4}$$

where P^E, Q^E, P^ν, Q^ν are differential operators. From (*A*4) it follows that

$$Q^E P^\nu \epsilon_{yy}=-Q^\nu P^E\sigma_{xx}. \tag{A5}$$

Solution of differential equation (*A*5) for a constant stress σ_{xx}^0 yields the lateral creep compliance function in $(A2_2)$.

For a standard linear solid, the operators in $(A4_1)$ are

$$P^E=1+p_1^E D \quad Q^E=q_0^E+q_1^E D$$
$$p_1^E=\frac{\eta_v}{E_m+E_v} \quad q_0^E=\frac{E_m E_v}{E_m+E_v} \quad q_1^E=\frac{E_m\eta_v}{E_m+E_v} \tag{A6}$$

where D is the time derivative. The operators P^ν and Q^ν in $(A4_2)$ remain to be determined. Since the matrix exhibits elastic compressibility,

$$\overline{\sigma}(t)=3K\overline{\epsilon}(t) \tag{A7}$$

where $\overline{\sigma}$, $\overline{\epsilon}$ denote mean stress and strain fields, one can define operators

$$P^K=1 \quad Q^K=3K. \tag{A8}$$

and then use the connection

$$\nu=\frac{3K-E}{6K} \tag{A9}$$

which together with (*A*4) provides

$$\frac{Q^\nu}{P^\nu}=\frac{3K-\dfrac{Q^E}{P^E}}{6K}=\frac{\dfrac{1}{2}P^E-\dfrac{1}{6K}Q^E}{P^E}. \tag{A10}$$

Substituting (*A*6) into (*A*10),

$$P^\nu=P^E=1+p_1^\nu D \quad Q^\nu=q_0^\nu+q_1^\nu D$$
$$p_1^\nu=\frac{\eta_v}{E_m+E_v} \quad q_0^\nu=\frac{1}{2}-\frac{E_m E_v}{6K(E_m+E_v)} \quad q_1^\nu=\frac{3\eta_v K-E_m\eta_v}{6K(E_m+E_v)} \tag{A11}$$

allows (*A*5) to be written as

$$(q_0^E+q_1^E D)\epsilon_{yy}=-(q_0^\nu+q_1^\nu D)\sigma_{xx} \tag{A12}$$

since $P^E=P^\nu$. For the step load $\sigma_{xx}^0 H(t)$, solution of this equation gives the lateral creep compliance function in $(A2_2)$,

$$J_y^m(t)=\frac{q_0^\nu}{q_0^E}\left[1+\left(\frac{q_0^E q_1^\nu}{q_1^E q_0^\nu}-1\right)\exp\left(-\frac{q_0^E}{q_1^E}t\right)\right]. \tag{A13}$$

Note that according to (*A*6), $q_0^E/q_1^E=E_v/\eta_v$, hence the exponents in (*A*1) and (*A*13) are identical.

Appendix B

Here we derive the stress and eigenstress distribution factors appearing in (42) and (43). Displacement continuity at ply interfaces requires that the in-plane strains (38_1) are equal to the overall strain $\boldsymbol{\epsilon}$,

$$\overline{\boldsymbol{\epsilon}}_i=\boldsymbol{\epsilon}. \tag{B1}$$

The laminate constitutive equations are written in a form analogous to (35),

$$\boldsymbol{\epsilon}=\mathbf{S}\mathbf{N}/h+\boldsymbol{\mu} \quad \mathbf{N}/h=\mathbf{Q}\boldsymbol{\epsilon}+\boldsymbol{\lambda} \tag{B2}$$

where $\mathbf{Q}$, $\mathbf{S}$ are the (3×3) overall in-plane stiffness and compliance of the laminate, $\mathbf{S}=\mathbf{Q}^{-1}$. The $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ are the overall transformation fields; since the mechanical load (30) and the average ply transformation fields (31) are independent, these overall fields are caused solely by the ply averages (31).

The part of stress averages caused in the ith pair of plies by the mechanical loads (30) follows from (35_2), (*B*1), and $(B2_1)$ as

$$\overline{\boldsymbol{\sigma}}_i=\overline{\mathbf{Q}}_i\overline{\boldsymbol{\epsilon}}_i=\overline{\mathbf{Q}}_i\mathbf{Q}^{-1}\mathbf{N}/h. \tag{B3}$$

Comparing this with the first term in (42) yields

$$\mathbf{H}_i=\overline{\mathbf{Q}}_i\mathbf{Q}^{-1}. \tag{B4}$$

The eigenstress distribution factor $\mathbf{K}_{ij}$ evaluates the stress in the ith pair of plies caused by a uniform eigenstress $\overline{\boldsymbol{\lambda}}_j=-\overline{\mathbf{Q}}_j\overline{\boldsymbol{\mu}}_j$ that is induced by the eigenstrain in the jth pair of plies. The derivation starts by introducing the eigenstrain $\overline{\boldsymbol{\mu}}_j$ into the jth ply of a

fully constrained laminate at $\boldsymbol{\epsilon}=\mathbf{0}$. This causes the stress $\overline{\boldsymbol{\sigma}}_i = \delta_{ij}\overline{\boldsymbol{\lambda}}_j = -\delta_{ij}\overline{\mathbf{Q}}_j\overline{\boldsymbol{\mu}}_j$, present only in the jth pair of plies, and overall reaction stress $\mathbf{N}^R/h = -c_j\overline{\mathbf{Q}}_j\overline{\boldsymbol{\mu}}_j$; δ_{ij} is the Kronecker symbol, but no summation is indicated by repeated subscripts. Next, the constraint is removed by applying the stress $\mathbf{N}/h = -\mathbf{N}^R/h$ to the entire laminate, which reduces the overall stress to zero. The resulting ply stress is obtained by superposition of the initial and unloading contributions,

$$\overline{\boldsymbol{\sigma}}_i = -\delta_{ij}\overline{\mathbf{Q}}_j\overline{\boldsymbol{\mu}}_j - \mathbf{H}_i\mathbf{N}^R/h. \tag{B5}$$

Comparing (*B*5) with the eigenstress term in (42_2) yields the eigenstress distribution factor,

$$\mathbf{K}_{ij} = \delta_{ij}\mathbf{I} - c_j\mathbf{H}_i\,. \tag{B6}$$

Moreover, removing the constraint changes the overall strain $(B2_1)$ from zero to the value of $-\mathbf{S}\mathbf{N}^R/h = c_j\mathbf{S}\overline{\mathbf{Q}}_j\overline{\boldsymbol{\mu}}_j$. Generated in the absence of any mechanical loads, this strain represents the contribution of the jth pair of plies to the overall laminate eigenstrain. Adding such contributions from all plies and letting $\mathbf{S} = \mathbf{Q}^{-1}$ yields the overall eigenstrain in the form

$$\boldsymbol{\mu} = \sum_{j=1}^{N} c_j\mathbf{H}_j^T\overline{\boldsymbol{\mu}}_j \tag{B7}$$

which is seen to be the laminate-level equivalent of the Levin's formula (31_1).

References

[1] Dvorak, G. J., and Suvorov, A. P., 2000, "Effect of Fiber Prestress on Residual Stresses and Onset of Damage in Symmetric Laminates," Compos. Sci. Technol., **60**, (8), pp. 1129–1139.

[2] Suvorov, A. P., and Dvorak, G. J., 2001, "Optimized Fiber Prestress for Reduction of Free Edge Stresses in Composite Laminates," Int. J. Solids Struct., **38**, pp. 6751–6786.

[3] Dvorak, G. J., Prochazka, P., and Srinivas, M. V., 1999, "Design and Fabrication of Submerged Cylindrical Laminates I," Int. J. Solids Struct., **36**, pp. 3917–3943.

[4] Srinivas, M. V., Dvorak, G. J., and Prochazka, P., 1999, "Design and Fabrication of Submerged Cylindrical Laminates II. Effect of Fiber Prestress," Int. J. Solids Struct., **36**, pp. 3945–3976.

[5] Schapery, R. A., 1967, "Stress Analysis of Viscoelastic Composite Materials," J. Compos. Mater., **1**, pp. 228–267.

[6] Lou, Y. C., and Schapery, R. A., 1971, "Viscoelastic Characterization of a Nonlinear Fiber-Reinforced Plastic," J. Compos. Mater., **5**, pp. 208–234.

[7] Mignery, L. A., and Schapery, R. A., 1991, "Viscoelastic and Nonlinear Adherent Effects in Bonded Composite Joints," J. Adhes., **34**, pp. 17–40.

[8] Schapery, R. A., 1993, "Compressive Strength and Failure Time Based on Local Buckling in Viscoelastic Composites," Appl. Mech. Rev., **46**, pp. 221–228.

[9] Sims, D. F., and Halpin, J. C., 1974, "Methods for Determining the Elastic and Viscoelastic Response of Composite Materials," *Composite Materials: Testing and Design* (Third Conference), ASTM STP, Philadelphia, PA, **546**, pp. 46–66.

[10] Wang, T. M., and Daniel, I. M., 1992, "Thermoviscoelastic Analysis of Residual Stresses and Warpage in Composite Laminates," J. Compos. Mater., **26**, pp. 883–899.

[11] Skudra, A. M., and Auzukalns, Y. V., 1973, "Creep and Long-Term Strength of Unidirectional Reinforced Plastics in Compression," Polym. Mech., **6**, pp. 718–722.

[12] Ferry, J. D., 1980, *Viscoelastic Properties of Polymers*, John Wiley and Sons, New York.

[13] Xia, Z., and Ellyin, F., 1998, "Time-Dependent Behavior and Viscoelastic Constitutive Modelling of an Epoxy Polymer," Polymer and Polymer Comp., **6**, pp. 75–83.

[14] Hi, Y., Xia, Z., and Ellyin, F., 2000, "Mechanical Behavior of an Epoxy Resin Under Multiaxial Loadings. Part I: Experimental Study," Polym. Polym. Comp., **8**, pp. 11–18.

[15] Hi, Y., Xia, Z., and Ellyin, F., 2000, "Mechanical Behavior of an Epoxy Resin Under Multiaxial Loadings. Part II: Comparison of Viscoelastic Constitutive Model Prediction," Polym. Polym. Comp., **8**, pp. 157–166.

[16] Ellyin, F., Hu, Y., and Xia, Z., 2000, "Multiaxial Behavior and Viscoelastic Constitutive Modeling of Epoxy Polymers," *Recent Trends in Constitutive Modeling of Advanced Materials*, AMD Vol. **239**, pp. 13–25.

[17] Dvorak, G. J., 1992, "Transformation Field Analysis of Inelastic Composite Materials," Proc. R. Soc. London, Ser. A, **437**, pp. 311–327.

[18] Dvorak, G. J., and Benveniste, Y., 1992, "On Transformation Strains and Uniform Fields in Multiphase Elastic Media," Proc. R. Soc. London, Ser. A, **A437**, pp. 291–310.

[19] Dvorak, G. J., Bahei-El-Din, Y. A., and Wafa, A. M., 1994, "Implementation of the Transformation Field Analysis for Inelastic Composite Materials," Comput. Mech., **14**, pp. 201–228.

[20] Levin, V. M., 1967, "Thermal Expansion Coefficients of Heterogeneous Materials," Izv. AN SSSR, Mekhanika Tverdogo Tela, **2**, pp. 88–94 (English Translation, Mechanics of Solids, 11, pp. 58–61).

[21] Mori, T., and Tanaka, K., 1973, "Average Stress in Matrix and Average Elastic Energy of Materials With Misfitting Inclusions," Acta Metall., **21**, pp. 571–574.

[22] Dvorak, G. J., and Srinivas, M. V., 1999, "New Estimates of Overall Properties of Heterogeneous Solids," J. Mech. Phys. Solids, **47**, pp. 899–920.

[23] Blackketter, D. M., and Upadhyaya, D., 1993, "Micromechanics Predictions of the Transverse Tensile Strength of Carbon Fiber/Epoxy Composites: the Influence of the Matrix and Interface," Polym. Compos., **14**, pp. 437–446.

[24] Adams, D. F., King, T. R., and Blackketter, D. M., 1990, "Evaluation of the Transverse Flexure Test Method for Composite Materials," Compos. Sci. Technol., **39**, pp. 341–353.

J. P. Briggs
Charles Schwab & Company,
SF211MN-8-223,
101 Montgomery Street,
San Francisco, CA 94104

P. Ponte Castaneda
Department of Mechanical Engineering and
Applied Mechanics,
University of Pennsylvania,
Philadelphia, PA 19104-6315

Variational Estimates for the Effective Response of Shape Memory Alloy Actuated Fiber Composites

The homogenization procedure of Ponte Castañeda is used to estimate the effective behavior of active composite materials consisting of aligned shape memory alloy (SMA) fibers embedded in a linear elastic matrix. Results are presented for thermal activation of the SMA with various applied tractions on the composite. While increasing stiffness of the matrix phase inhibits the contraction of the SMA, the simulations indicate that the use of a prestress in the manufacturing of the composite may provide an increase in the response time of the system without reducing performance. [DOI: 10.1115/1.1464873]

1 Introduction

Smart composites can be loosely defined as composites with one or more constitutive materials having embedded sensing and/or actuation. The ability either to sense an internal state, or to actively control its behavior, makes such materials highly desirable. Applications to actuation devices have shown exciting possibilities, and in particular, shape memory alloys (SMAs) have opened the door to the possibility of truly biomimetic forms of locomotion for robots. As such, there has been a great deal of research activity surrounding smart materials, ranging from materials development to modeling and active control.

Sensing is critical to control and in many respects can be used to great effect in concert with very simple control algorithms. However, as a system becomes increasingly more complicated, less stable or sensing more costly, good mathematical models can provide a means of minimizing the ''effort'' required to produce a desirable output. In the case of SMA-fiber actuators, simple controllers become less and less reliable as hysteresis begins to accumulate and degrade performance.

Hysteresis and sensing issues are only amplified in an SMA composite. It is certainly possible to consider controlling individual wires within the composite, but the cost of such an approach is prohibitive, and raises issues of global convergence for control. Furthermore, when displacement or strain in a continuum is the primary objective of the control system, there is no simple method of gathering good data for feedback to the controller. In fact, the majority of the states associated with SMA are not even observable. Some type of alternative means of counteracting the effects of hysteresis and lack of real-time data is required. We believe that by estimating effective behavior, homogenization techniques provide a critical link between understanding SMA's as individual actuators and SMA's as components in smart composites. In the future, it is our goal to use homogenization as a means of hysteresis inversion within a control system.

The theory of homogenization for linear composites is fairly well developed. Appropriate references dealing with linear theories are provided by the review articles of Willis [1,2] and the monograph of Christensen [3]. With new variational techniques, nonlinear composites have received increased attention in recent years. Willis [2] and Talbot and Willis [4] extended the Hashin-Shtrikman variational principles to include nonlinear constitutive behavior. Ponte Castañeda [5] proposed a new variational principle which can be used to obtain general types of bounds and estimates for the properties of nonlinear composites. The procedure makes use of arbitrary bounds and estimates for suitable classes of linear comparison composites. More recently, a ''second order'' procedure has been developed by Ponte Castañeda [6] that delivers estimates for nonlinear composites at finite contrast exact to second order. In this work, use will be made of the original ''linear comparison composite'' variational procedure of Ponte Castañeda.

Most of the original one-dimensional macroscopic SMA models simulate phase transformation effects without considering general thermomechanical activation conditions ([7–10]), where either temperature or stress are assumed constant. Often surfaces in temperature-stress space define the conditions for termination and commencement of phase transformation, commonly referred to as transformation plasticity ([11]) because of the conceptual similarity between transformation surfaces and plastic yield surfaces. To simplify the analysis, only complete transformations were considered. Notable exceptions include the works of Ivshin and Pence [12] and Falk [13,14] who used embedded hysteresis models to capture the effects of incomplete transformations (*minor loops*). Due to the apparent success of these general mathematical hysteresis models, many subsequent models addressed incomplete transformations by incorporating Duhem-Madelung [15–17] and Landau-Devonshire [18] hysteresis models.

More recent macroscopic models have begun to appear in the literature in the past couple of years. Building upon earlier, simpler models, combined with experimental work on SMA wires, Bo and Lagoudas [19,20] derived a sophisticated three-dimensional ''incremental'' model capable of capturing both general thermomechanical loading conditions and minor loops ([21]). The complexity of the model arising from its incremental form makes it more difficult to implement within any type of homogenization-based control scheme. Bekker and Brinson [17] introduced a macroscale kinetic law related to the Duhem-Madelung differential models with dependence on a transformation phase diagram. While their model is also capable of simulating the effects of general thermomechanical loading conditions and minor loops, and appears to be simple enough to implement within the framework of a control algorithm, it is not clear whether analytical

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 10, 2001; final revision, September 27, 2001. Associate Editor: M.-J. Pindera. 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.

expressions can be generated for the stress energy. This makes it difficult to implement within the linear comparison composite homogenization procedure ([5]).

In the context of composites and structures, there has been relatively little work in studying the effects of embedded SMA fibers. The paucity of studies involving SMA composites using continuum mechanics is most probably a direct result of SMA model complexity. Boyd and Lagoudas [19,22] used an incremental version of the Mori-Tanaka method to determine effective properties of an SMA fiber composite and compared their results to a full finite element method simulation for a periodic system. Their model for the SMA was based upon the earlier work of Tanaka, and hypothesized the existence of an isotropic strain-rate yield function coupled with linear transformation surfaces. Only isothermal stress-activated major loops were considered. On the other hand, the homogenization methods used in this pioneering study are known to be computationally intensive. It is one of our primary goals in this work to make use of more efficient homogenization methods to decrease the computational resources in calculating the effective response of the SMA fiber-activated composites for the purpose of real-time control.

Brinson [23] made efforts to include both complex stress-temperature states and minor loop behavior for a one-dimensional SMA wire attached to a polymeric rod, where linear and nonlinear beam theory were used to couple the states of stress and strain in the rod and SMA wire. Similarly, de Blonk and Lagoudas [24] modeled deformation of an elastomeric rod with embedded SMA fibers distributed as one-dimensional elements. Fiber-matrix interactions and effective behavior were not considered. While it is possible, at least in principle, to make use of more sophisticated SMA models with more accurate homogenization estimates for SMA-actuated composites, our objective here is to develop an efficient model, yielding a good balance of accuracy versus complexity, which can be implemented for better control than has been possible to date. A recent article by Bernardini [25] describing energy functions for SMA's may provide just such an example of an alternate constitutive model to incorporate within the context of this formulation.

2 SMA Constitutive Model

The SMA constitutive behavior used in this work is a modification by Briggs and Ostrowski [26] of the earlier Boyd and Lagoudas, isotropic strain rate model ([19]), given by

$$\boldsymbol{\varepsilon}=[\mathbf{M}^A+\xi(\mathbf{M}^M-\mathbf{M}^A)]\boldsymbol{\sigma}+\boldsymbol{\Lambda}. \tag{1}$$

Although Boyd and Lagoudas made use of an ''incremental'' version of the relation in Eq. (1), as will become evident in the discussion to follow, the ''total'' form is more consistent with the homogenization scheme to be used in this work (even if it is possibly less accurate). The first set of terms in this relation represents a proportional mixture of the elastic moduli of the two phases, where ξ is the martensitic volume fraction. To reduce implementation complexity, this mixture is replaced with a single linear isotropic term, $\mathbf{M}$. The second term $\boldsymbol{\Lambda}$ is the strain induced by phase transformation, which evolves as a function of the instantaneous state of temperature and effective stress τ_e ([27]), as determined by

$$\boldsymbol{\Lambda}=\frac{3}{2}\frac{\omega}{D}\frac{\boldsymbol{\sigma}'}{\tau_e}(\xi[\theta(t),\tau_e(t)]-1). \tag{2}$$

The effective stress is defined to be $\tau_e=\sqrt{\frac{1}{2}\boldsymbol{\sigma}'\cdot\boldsymbol{\sigma}'}$, where $\boldsymbol{\sigma}'=\boldsymbol{\sigma}-\sigma_m\boldsymbol{\delta}$, $\sigma_m=\frac{1}{3}\mathrm{tr}(\boldsymbol{\sigma})$, $\boldsymbol{\delta}$ is the second order identity tensor, and (in keeping with the notation of Boyd and Lagoudas) ω and D are material parameters specified in Table 1. It should be noted that the martensitic volume fraction of the SMA varies between zero and one. This implies that in the purely martensitic state there will exist a large positive strain. In general applications the SMA fibers will be embedded within a surrounding matrix in a purely martensitic state and then subsequently heated to force transition to austenite with the associated strain recovery. Without the -1 in Eq. (2) this strain recovery in the SMA fiber would result in positive strains in the fibers and negative strains in the surrounding matrix. In order to make the strains in both the fiber and the matrix consistent and more intuitive, the -1 is introduced into the martensite term $(\xi[\theta(t),\tau_e(t)]-1)$ as a simple offset. The result for the SMA is then zero strain in pure martensite and negative strain during strain recovery.

Table 1 SMA material parameters

$E^{sma}=70.0\times10^3\ MPa$,	$D=50.0\times10^3\ MPa$
$v^{sma}=0.33$	$\omega=1.750\times10^3\ MPa$
$C=35\ \frac{MPa}{^oC}$	$a=0.12\ \frac{1}{^oC}$
$A^{om}=55^oC$	$M^{om}=45^oC$

While not fully explored here, the model proposed by Briggs and Ostrowski [26] is entirely capable of simulating *minor loop* behavior. As with previous work by Bekker and Brinson [17] and Ivshin and Pence [12,28], the ability to capture minor loop behavior is a result of the inclusion of a Duhem-Madelung type hysteresis model in the function $\xi[\theta(t),\tau_e(t)]$. A Duhem-Madelung formulation as described in Eqs. (3) and (4) requires two independent branches, where each branch corresponds to the increasing or decreasing of the dependent variable, martensite.

$$\text{for } \dot{\xi}>0: \tag{3}$$

$$\xi[\theta(t),\tau_e(t)]=\xi(t_i)+[1-\xi(t_i)]\times\left\{\frac{F^M[\theta(t),\tau_e(t)]-F^M[\theta(t_i),\tau_e(t_i)]}{1-F^M[\theta(t_i),\tau_e(t_i)]}\right\}$$

$$\text{for } \dot{\xi}<0: \tag{4}$$

$$\xi[\theta(t),\tau_e(t)]=\xi(t_i)-\xi(t_i)\left\{\frac{F^A[\theta(t),\tau_e(t)]-F^A[\theta(t_i),\tau_e(t_i)]}{-1-F^A[\theta(t_i),\tau_e(t_i)]}\right\}$$

The subscript i indicates the value of the appropriate quantity when $\dot{\xi}(t)$ last equaled zero or changed sign. The use of switching states is a method of enforcing continuity of the constitutive relations in the two branch model. After each change in the sign of $\dot{\xi}(t)$ the initial conditions in temperature and stress for the evolution of $\xi[\theta(t),\tau_e(t)]$ are determined by the termination states of the previous branch.

Two envelope functions, F^M, F^A, in stress-temperature space are chosen (one for each branch) as boundary surfaces within which $\xi[\theta(t),\tau_e(t)]$ is confined. The addition of general envelope functions F^A, F^M is similar in concept to that presented by Lagoudas and Bo [29] in their unified constitutive model. The choice of function can be made using experimental data or thermodynamic arguments. In the particular case of the modified model of Briggs and Ostrowski [26], the functions are based upon experiments using a high-temperature 0.015-inch nickel-titanium wire from Dynalloy, Inc.:

$$F^M[\theta,\tau_e]=\tanh[a^M(\theta^{dM}-\theta)+b^M\tau_e] \tag{5}$$

$$F^A[\theta,\tau_e]=\tanh[a^A(\theta^{dA}-\theta)+b^A\tau_e]. \tag{6}$$

The parameters internal to F^M and F^A are defined as

$$a^M=a^A=a,\quad C^M=C^A=C,$$

$$b^M=\frac{a^M}{C^M},\quad b^A=b^M,$$

$$\theta^{dM}=M^{om}-\tfrac{1}{2}(A^{om}-M^{om})\tanh^{-1}[\xi(t_i)],$$

$$\theta^{dA}=A^{om}+\tfrac{1}{2}(A^{om}-M^{om})\tanh^{-1}[1-\xi(t_i)],$$

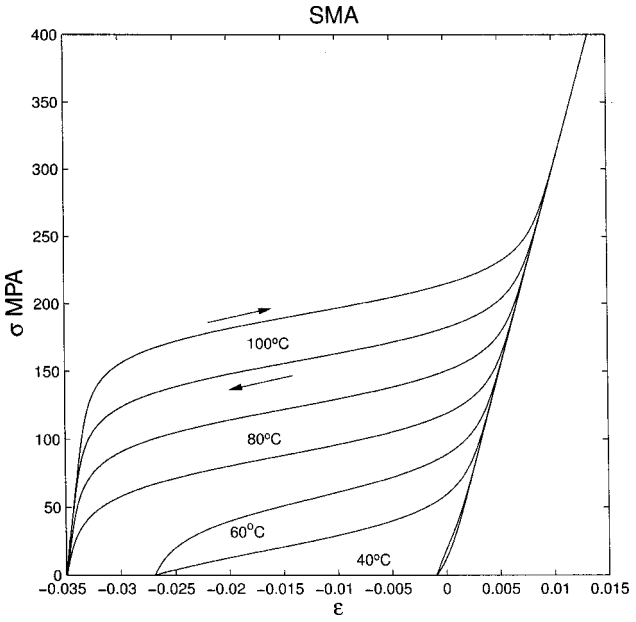


Fig. 1 Longitudinal, isothermal stress cycling for an SMA fiber where the initial phase mixture and strain are determined by temperature

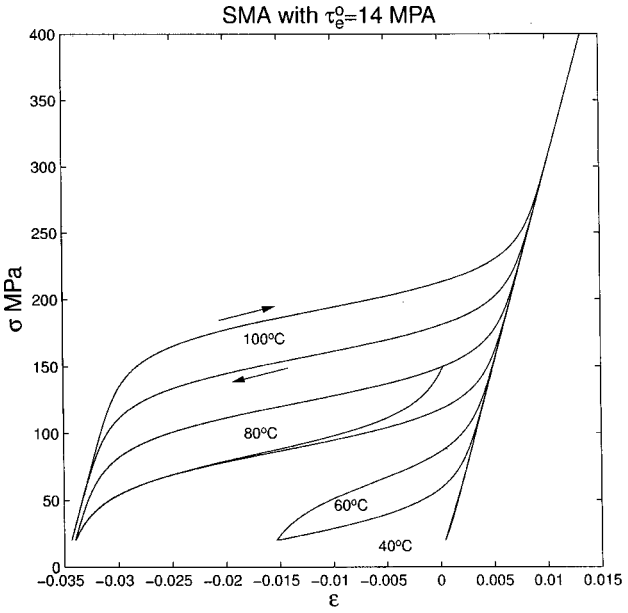


Fig. 3 Longitudinal, isothermal stress cycling of an SMA fiber with a prestress

a, C, A^{om}, M^{om} are included in Table 1, where E, ν, D were taken from earlier the work by Lagoudas [22]. For a *trained* SMA actuator, proportional loading cycles can form closed paths in stress-strain space as in Fig. 1. Heating and cooling shift the location of the stress-free strain state. Notice also that for very low and very high temperature the SMA is linear. These temperature extremes correspond to pure martensite and austenite phases. The presence of an interior loop at 80°C is included simply as a demonstration of the minor loop behavior. In Fig. 2 we have included compressive stress in the SMA model. While there is certainly some interesting behavior for this loading condition, we have chosen to focus our analysis on tension and shear primarily because of the potential applications of such a composite. We foresee the use of active composites in roles where interaction between the composite and its environment will result in stresses that resist the formation of transformation strains. The most obvious analogy can be found in the muscles of biological systems.

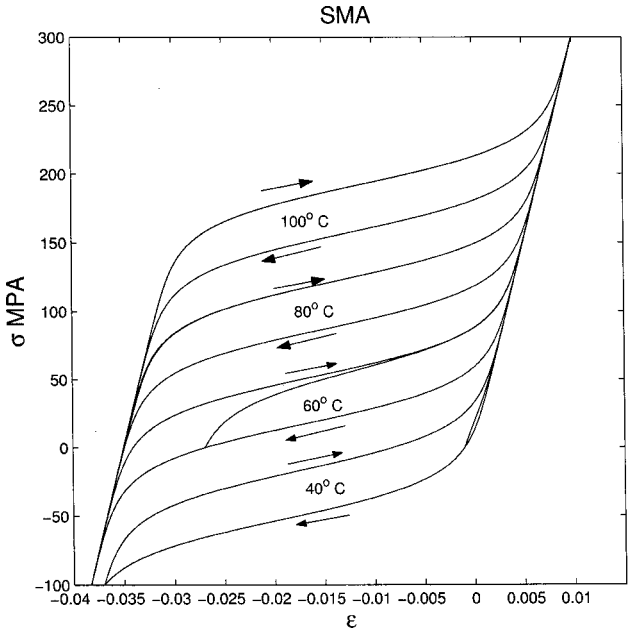


Fig. 2 Longitudinal, isothermal stress cycling with compression for an SMA fiber where the initial phase mixture and strain are determined by temperature

If the SMA fiber has a prestress, the initial tension shifts the equilibrium configuration to a higher stress state as in Fig. 3, while retaining the two branches corresponding to forward and reverse transformations. The configuration from which the strain evolves is now one that has a nonzero stress, but maintains the closed loop characteristic. Figure 3 is just such an example with four different isothermal stress cycles and an initial fiber tension corresponding to $\tau_e = 14$ MPa.

One can also consider iso-stress thermal activation. Figure 4 demonstrates thermal activation for an SMA wire initially in a state of pure martensite with longitudinal tension. As the prestress increases, the subsequent hysteresis loop in temperature-strain space shifts to a higher temperature range with an offset in strain. The size and shape of the hysteresis loop remains consistent between the different states of prestress, implying that the change in temperature required for a complete loop is constant for a constant stress.

In the homogenization analysis below, use will be made of the stress energy function $u(\boldsymbol{\sigma}, \theta, \xi_o, \mathrm{sgn}(\dot{\xi}))$ for the SMA, which is obtained by integrating Eq. (1) with respect to stress. Note that because we are treating the hysteresis in the SMA as a pseudo-elastic material with separate constitutive branches, two potentials, one for loading and one for unloading, are required. The SMA is assumed to be isotropic and the energy function may be written in terms of its isotropic stress invariants (τ_e, σ_m) such that

$$\text{for } \dot{\xi} > 0: \tag{7}$$

$$u(\tau_e, \sigma_m, \theta) = \frac{3}{2}\frac{\omega}{D}\left\{\xi(t_i) - \frac{[1-\xi(t_i)]F^M[\theta(t_i), \tau_e(t_i)]}{1-F^M[\theta(t_i), \tau_e(t_i)]} - 1\right\}\tau_e + \frac{3}{2}\frac{\omega}{D}\left\{\frac{[1-\xi(t_i)]}{1-F^M[\theta(t_i), \tau_e(t_i)]}\right\}G^M[\theta, \tau_e] + \frac{1}{2k}\sigma_m^2 + \frac{1}{2\mu}\tau_e^2,$$

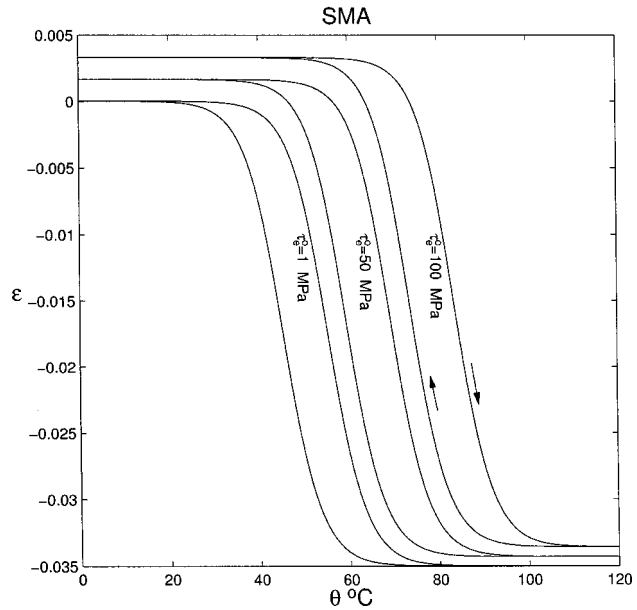


Fig. 4 Thermal cycling of an SMA fiber with a constant tensile prestress

$$\text{for } \dot{\xi}<0: \tag{8}$$

$$u(\tau_e,\sigma_m,\theta)=\frac{3}{2}\frac{\omega}{D}\left\{\xi(t_i)-\frac{[\xi(t_i)]F^A[\theta(t_i),\tau_e(t_i)]}{1+F^A[\theta(t_i),\tau_e(t_i)]}-1\right\}\tau_e+\frac{3}{2}\frac{\omega}{D}\left\{\frac{[\xi(t_i)]}{1+F^A[\theta(t_i),\tau_e(t_i)]}\right\}G^A[\theta,\tau_e]+\frac{1}{2k}\sigma_m^2+\frac{1}{2\mu}\tau_e^2,$$

where the functions G correspond to the following integrals:

$$G^M[\theta,\tau_e]=\int F^M[\theta,\tau_e]\left(\frac{\boldsymbol{\sigma}'}{\tau_e}\right)\cdot d\boldsymbol{\sigma} \tag{9}$$

$$=\frac{2}{b^M}\ln(|\cosh[a^M(\theta^{dM}-\theta)+b^M\tau_e]|) \tag{10}$$

$$G^A[\theta,\tau_e]=\int F^A[\theta,\tau_e]\left(\frac{\boldsymbol{\sigma}'}{\tau_e}\right)\cdot d\boldsymbol{\sigma} \tag{11}$$

$$=\frac{2}{b^A}\ln(|\cosh[a^A(\theta^{dA}-\theta)+b^A\tau_e]|). \tag{12}$$

3 Homogenization Estimates for Nonlinear Composites

Consider a representative volume element (RVE) occupying the volume Ω comprised of N distinct phases occupying subregions $\Omega^{(r)}$. Consistent with the above discussion for SMA materials, the constitutive behaviors of the individual components of the composite are assumed to be governed by a potential, or stress-energy function, u, in such a way that the infinitesimal stress and strain fields are related by

$$\boldsymbol{\varepsilon}(\mathbf{x})=\frac{\partial u}{\partial\boldsymbol{\sigma}}(\mathbf{x},\boldsymbol{\sigma}). \tag{13}$$

The local potentials of the N phases, $u^{(r)}$ are assumed to be homogeneous, and the phases perfectly bonded at the interfaces, so that the potential for the composite can be written as

$$u(\mathbf{x},\boldsymbol{\sigma})=\sum_{r=1}^{N}\chi^{(r)}(\mathbf{x})u^{(r)}(\boldsymbol{\sigma}) \tag{14}$$

where the characteristic functions $\chi^{(r)}$ are equal to one if the position vector $\mathbf{x}$ is inside phase r and zero otherwise. Spatial averages over Ω and $\Omega^{(r)}$ are denoted by $\langle\cdot\rangle$ and $\langle\cdot\rangle^{(r)}$, respectively. For example, the volume fraction $c^{(r)}$ are defined by $c^{(r)}=\langle\chi^{(r)}(\mathbf{x})\rangle$. Also, the mean stress and strain are the averages of the local stress and strain fields: $\overline{\boldsymbol{\varepsilon}}=\Sigma_{r=1}^{N}c^{(r)}\overline{\boldsymbol{\varepsilon}}^{(r)}$ and $\overline{\boldsymbol{\sigma}}=\Sigma_{r=1}^{N}c^{(r)}\overline{\boldsymbol{\sigma}}^{(r)}$, where the notations $\overline{\boldsymbol{\varepsilon}}^{(r)}=\langle\boldsymbol{\varepsilon}\rangle^{(r)}$ and $\overline{\boldsymbol{\sigma}}^{(r)}=\langle\boldsymbol{\sigma}\rangle^{(r)}$ have been used.

Now, assuming that the size of the typical inclusion is small relative to the size of the specimen, and that the boundary conditions vary slowly with respect to inclusion size, the effective strain of the composite may be written in the form (see, e.g., Ponte Castañeda, and Suquet [30]):

$$\overline{\boldsymbol{\varepsilon}}=\frac{\partial\tilde{U}(\overline{\boldsymbol{\sigma}})}{\partial\overline{\boldsymbol{\sigma}}} \tag{15}$$

where $\tilde{U}(\overline{\boldsymbol{\sigma}})$ is the effective potential given by

$$\tilde{U}(\overline{\boldsymbol{\sigma}})=\min_{\boldsymbol{\sigma}\in S(\overline{\boldsymbol{\sigma}})}\langle u(\mathbf{x},\boldsymbol{\sigma})\rangle. \tag{16}$$

Here,

$$S(\overline{\boldsymbol{\sigma}})=\{\boldsymbol{\sigma}|\nabla\cdot\boldsymbol{\sigma}=0 \text{ in } \Omega,\quad \boldsymbol{\sigma}\mathbf{n}=\overline{\boldsymbol{\sigma}}\mathbf{n},\quad \text{on } \partial\Omega\} \tag{17}$$

is the set of statically admissible stresses with uniform traction on the boundary. The effective behavior of the composite is fully described by Eqs. (16) and (17). However, the determination of $\tilde{U}(\overline{\boldsymbol{\sigma}})$ requires the solution of a nonlinear boundary value problem with random microstructure.

To overcome some of the difficulties associated with the determination of the effective behavior of the nonlinear composite, Ponte Castañeda [5] introduced a variational procedure that can be used to generate estimates for the potential functions of nonlinear composites in terms of the effective potential of a linear comparison composite with identical microstructure.

Here, isotropy of the phases is assumed, so that the potential functions may be written in the form

$$u^{(r)}(\boldsymbol{\sigma})=\phi^{(r)}(\tau_e,\sigma_m) \tag{18}$$

where τ_e and σ_m have been defined previously.
Then, the potential functions of the phases may be expressed in terms of an optimization problem:

$$u^{(r)}(\boldsymbol{\sigma})=\max_{\mu_o^{(r)},k_o^{(r)}\geqslant 0}\{u_o^{(r)}(\boldsymbol{\sigma})-V^{(r)}(\mu_o^{(r)},k_o^{(r)})\}, \tag{19}$$

where $u_o^{(r)}$ is the potential function of an isotropic linear-elastic ''comparison'' material with shear and bulk modulus $(\mu_o^{(r)},k_o^{(r)})$, such that

$$u_o^{(r)}(\boldsymbol{\sigma})=\frac{1}{2\mu_o^{(r)}}\tau_e^2+\frac{1}{2k_o^{(r)}}\sigma_m^2 \tag{20}$$

and

$$V^{(r)}(\mu_o^{(r)},k_o^{(r)})=\max_{\boldsymbol{\sigma}}\{u_o^{(r)}(\boldsymbol{\sigma})-u^{(r)}(\boldsymbol{\sigma})\}. \tag{21}$$

Using relation (19) in expression (16), and interchanging the order of the minimum and the maxima, the following estimate is generated ([5]) for the effective potential of the composite

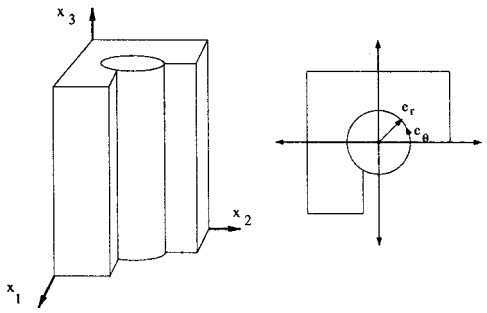


Fig. 5 Fiber composite with local coordinate system

$$\tilde{U}(\overline{\boldsymbol{\sigma}}) \geqslant \max_{\mu_o^{(r)},\kappa_o^{(r)}} \left\{ \tilde{U}_o(\overline{\boldsymbol{\sigma}}) - \sum_{r=1}^{n} c^{(r)} V^{(r)}(\mu_o^{(r)},\kappa_o^{(r)}) \right\} \tag{22}$$

where $\tilde{U}_o(\overline{\boldsymbol{\sigma}}) = \frac{1}{2}\overline{\boldsymbol{\sigma}} \cdot \tilde{\mathbf{M}}\overline{\boldsymbol{\sigma}}$ is the effective potential of a linear comparison composite with precisely the same microstructure as the original, given nonlinear composite. In this last relation, $\tilde{\mathbf{M}}$ denotes the effective compliance tensor of the elastic linear comparison composite with phase moduli tensor $\mathbf{M}^{(1)}$ and $\mathbf{M}^{(2)}$, and the same microstructure as the nonlinear composite.

It is possible to generalize the above estimate to include the case where there are uniform prestrains $\boldsymbol{\varepsilon}_i^{(r)}$ present in each phase r. In this case, the local potential functions can be written in the form

$$u^{(r)}(\boldsymbol{\sigma}) = \phi^{(r)}(\tau_e, \sigma_m) + \boldsymbol{\varepsilon}_i^{(r)} \cdot \boldsymbol{\sigma}. \tag{23}$$

Then, the effective potential function for the corresponding nonlinear composite with prestrains becomes

$$\tilde{U}(\overline{\boldsymbol{\sigma}}) \geqslant \max_{\mu_o^{(r)},\kappa_o^{(r)}} \left\{ \tilde{U}_T(\overline{\boldsymbol{\sigma}}) - \sum_{r=1}^{2} c^{(r)} V^{(r)}(\mu_o^{(r)},\kappa_o^{(r)}) \right\} \tag{24}$$

with

$$V^{(r)}(\mu_o^{(r)},\kappa_o^{(r)}) = \max_{\tau_e,\sigma_m} \{ u_o^{(r)}(\boldsymbol{\sigma}) - \phi^{(r)}(\tau_e,\sigma_m) \} \tag{25}$$

and where

$$\tilde{U}_T(\overline{\boldsymbol{\sigma}}) = \min_{\boldsymbol{\sigma} \in S(\overline{\boldsymbol{\sigma}})} \langle u_o^{(r)}(\boldsymbol{\sigma}) + \boldsymbol{\varepsilon}_i^{(r)} \cdot \boldsymbol{\sigma} \rangle \tag{26}$$

is the effective energy of a linear ''thermoelastic'' composite whose components have constitutive behaviors given by

$$\boldsymbol{\varepsilon}^{(r)} = \mathbf{M}^{(r)}\boldsymbol{\sigma}^{(r)} + \boldsymbol{\varepsilon}_i^{(r)}. \tag{27}$$

Now, for a two-phase system, like the systems of interest in this work, the variational problem posed in Eq. (26) for the effective thermoelastic potential is known to have the solution (Levin [31]):

$$\tilde{U}_T(\overline{\boldsymbol{\sigma}}) = \tfrac{1}{2}\overline{\boldsymbol{\sigma}} \cdot \overline{\mathbf{M}}\overline{\boldsymbol{\sigma}} + \overline{\boldsymbol{\varepsilon}}_i \cdot \overline{\boldsymbol{\sigma}} + \tfrac{1}{2}[\overline{\boldsymbol{\sigma}} + (\mathbf{M}^{(1)} - \mathbf{M}^{(2)})^{-1}(\boldsymbol{\varepsilon}_i^{(1)} - \boldsymbol{\varepsilon}_i^{(2)})] \cdot (\tilde{\mathbf{M}} - \overline{\mathbf{M}})[\overline{\boldsymbol{\sigma}} + (\mathbf{M}^{(1)} - \mathbf{M}^{(2)})^{-1}(\boldsymbol{\varepsilon}_i^{(1)} - \boldsymbol{\varepsilon}_i^{(2)})] \tag{28}$$

where the notation $\overline{\boldsymbol{\varepsilon}}_i$ has been used to denote the mean value of the prestrains $\boldsymbol{\varepsilon}_i^{(r)}$. Using relation (15), the average strain for the nonlinear (and linear comparison) composite is then found to be

$$\overline{\boldsymbol{\varepsilon}} = \tilde{\mathbf{M}}\overline{\boldsymbol{\sigma}} + \overline{\boldsymbol{\varepsilon}}_i + (\tilde{\mathbf{M}} - \overline{\mathbf{M}})(\mathbf{M}^{(1)} - \mathbf{M}^{(2)})^{-1}(\boldsymbol{\varepsilon}_i^{(1)} - \boldsymbol{\varepsilon}_i^{(2)}). \tag{29}$$

Also, the average stress in phase r is given by relation

$$\overline{\boldsymbol{\sigma}}^{(1)} = \overline{\boldsymbol{\sigma}} + \frac{1}{c^{(1)}}(\mathbf{M}^{(1)} - \mathbf{M}^{(2)})^{-1}(\tilde{\mathbf{M}} - \overline{\mathbf{M}}) \times [\overline{\boldsymbol{\sigma}} + (\mathbf{M}^{(1)} - \mathbf{M}^{(2)})^{-1}(\boldsymbol{\varepsilon}_i^{(1)} - \boldsymbol{\varepsilon}_i^{(2)})] \tag{30}$$

$$\overline{\boldsymbol{\sigma}}^{(2)} = \overline{\boldsymbol{\sigma}} - \frac{1}{c^{(2)}}(\mathbf{M}^{(1)} - \mathbf{M}^{(2)})^{-1}(\tilde{\mathbf{M}} - \overline{\mathbf{M}}) \times [\overline{\boldsymbol{\sigma}} + (\mathbf{M}^{(1)} - \mathbf{M}^{(2)})^{-1}(\boldsymbol{\varepsilon}_i^{(1)} - \boldsymbol{\varepsilon}_i^{(2)})].$$

These equations for the average stress in the phases are used in the following section to estimate the stress in the SMA fibers to initialize the hysteresis functions from Eqs. (3) and (4).

In the above relations, $\tilde{\mathbf{M}}$ is the effective compliance tensor of an elastic linear comparison composite with phase moduli tensor $\mathbf{M}^{(1)}$ and $\mathbf{M}^{(2)}$, distributed with the same microstructure as the nonlinear composite. In this work, use will be made of the Hashin-Shtrikman estimates of Walpole [32–34] for fiber composites:

$$\tilde{\mathbf{M}} = \{c^{(1)}\mathbf{M}^{(1)} + c^{(2)}\mathbf{M}^{(2)}[\mathbf{I} + \mathbf{Q}(\mathbf{M}^{(2)} - \mathbf{M}^{(1)})]^{-1}\} \{c^{(1)}\mathbf{I} + c^{(2)}[\mathbf{I} + \mathbf{Q}(\mathbf{M}^{(2)} - \mathbf{M}^{(1)})]^{-1}\}^{-1}. \tag{31}$$

In these relations, the microstructural tensor $\mathbf{Q}$ is related to the well-known Eshelby tensor, $\mathbf{S}$, through the $\mathbf{P}$ tensor, such that $\mathbf{P} = \mathbf{S}\mathbf{M}^{(1)}$ and $\mathbf{Q}^{-1} = \mathbf{M}^{(1)}(\mathbf{M}^{(1)} - \mathbf{P})^{-1}\mathbf{M}^{(1)}$. Using Walpole's [34] notation the $\mathbf{P}$ tensor may be expressed in the form

$$\mathbf{P} = \left[\frac{3}{6k + 8\mu}, 0, 0, 0, \frac{3k + 7\mu}{12k\mu + 16\mu^2}, \frac{1}{4\mu} \right]. \tag{32}$$

4 Applications and Results

In this section, the nonlinear homogenization method discussed in Section 3 is applied to the SMA-actuated fiber composites depicted in Fig. 5. Here, the matrix phase is taken to be isotropic, incompressible, linear elastic with given shear modulus $\mu^{(1)}$, and the SMA fibers, which are assumed to be aligned in the x_3-direction with transverse isotropy in the (x_1, x_2)-plane, exhibit the constitutive behavior described in some detail in Section 2. Therefore, the only source of nonlinearity in the composite arises from the SMA fibers. On the other hand, prestrains may exist in both the matrix and fiber phases, depending on the fabrication conditions and thermomechanical loading history. The phase fraction of SMA, $c^{(2)}$, is considered to be 0.33, and the phase fraction of the matrix, $c^{(1)}$, is 0.77.

Without considering minor loops, the constitutive behavior for unconstrained SMA fibers at constant temperature evolves from a single, fixed strain state. This implies that a complete proportional loading and unloading cycle will form closed loops. Alternatively, unconstrained traction-free SMA will remain stress free during thermal cycling. However, for constrained SMA fibers embedded within a composite varying the temperature will cause internal stresses to develop and evolve. Once the temperature has been fixed, despite the presence of prestresses in the phases, the constitutive behavior will again evolve from a single fixed point in stress-strain space. However, the fixed point in strain (the $\boldsymbol{\varepsilon}_i$) now happens to occur at nonzero stress for both phases.

The general solution procedure requires determining the location of the aforementioned fixed strain points for the SMA constitutive behavior and the surrounding matrix. These fixed points are a function of the fiber-matrix interaction stresses created during thermal phase transformation under traction-free conditions and are determined by solving the appropriate expressions for jump conditions between phases as described in the Appendix. Assuming that the evolution direction of ξ is known, the correct consti-

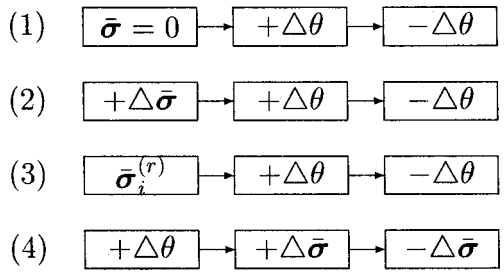


Fig. 6 Loading and heating scenarios for SMA composite

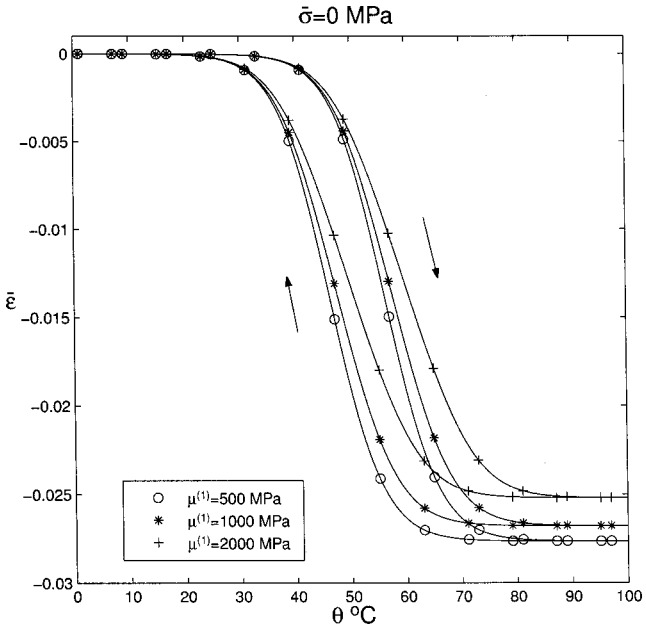


Fig. 7 Traction-free temperature cycle using three different matrix stiffnesses

tutive branch for the SMA is carefully selected. If temperature and applied stress are driving the SMA towards different transformation directions, the constitutive branch is determined by whether applied stress or temperature will dominate the evolution of the martensite in the envelope functions in Eqs. (5) and (6). Once the correct branch is identified, the initial states of temperature and stress in the SMA fiber are determined from Eqs. (30) and the traction-free $\boldsymbol{\varepsilon}_i$ are calculated using Eqs. (36) through (38). Of course, if temperature is fixed, the $\boldsymbol{\varepsilon}_i$ do not need to be recalculated. The homogenization procedure is then completed with the final applied stress and the appropriate $\boldsymbol{\varepsilon}_i$.

The homogenization equations can be shown to reduce to two coupled nonlinear algebraic equations in τ_e and $\mu_o^{(2)}$ which are solved using *Matlab* scripts with a Levenberg-Marquardt method.

Four general stress/temperature loading conditions will be examined in this section as illustrated in Fig. 6, where the $\pm\Delta$ indicates a change in the appropriate quantity. While these four cases do not represent all possible combinations of temperature and stress, they are representative of likely applications of the controlled SMA composite.

4.1 Case #1. Case 1 demonstrates heating and cooling of the SMA under traction-free boundary conditions for the composite. In Fig. 7 we have include three shear moduli for the surrounding matrix to illustrate the reduction in transformation strain as the matrix becomes increasingly stiff. Not only is the maximum transformation strain reduced, but the temperature required to achieve a particular strain becomes higher with the stiffness of the surrounding matrix. For purposes of controlling a system where *maximum* strain is the sole metric for control performance, these results indicate that embedding within a more compliant matrix is preferable. Maximum contraction is greater and the energy requirements to achieve a particular state of strain are less.

4.2 Case #2. Case 2 demonstrates uniform tractions on the composite followed by a single heating-cooling cycle of the embedded SMA. In Figs. 8 and 9, we have applied longitudinal and transverse tractions, respectively, where subsequent heating-cooling cycles cause strains to develop and recede. The hysteresis

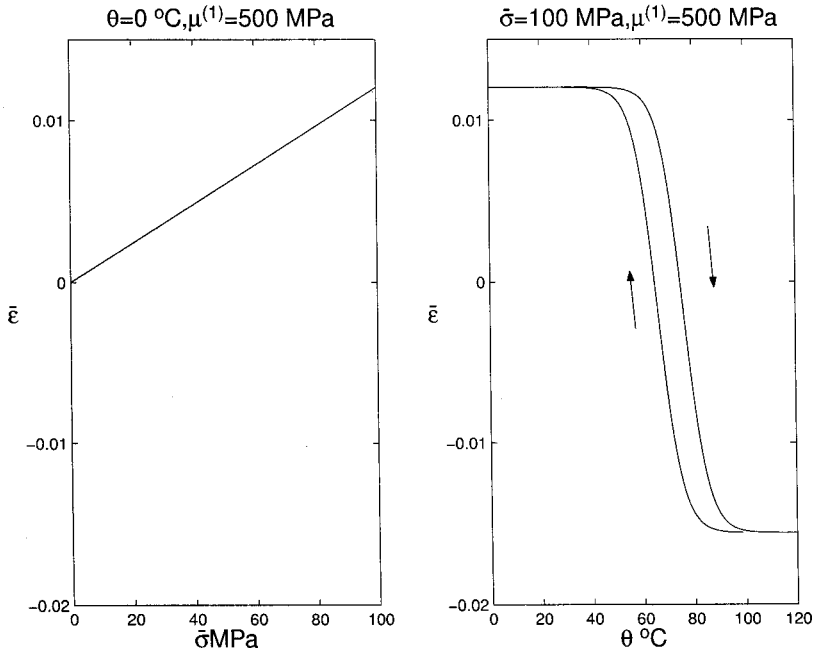


Fig. 8 Initial application of longitudinal traction of 100 MPa on the composite with subsequent temperature cycle

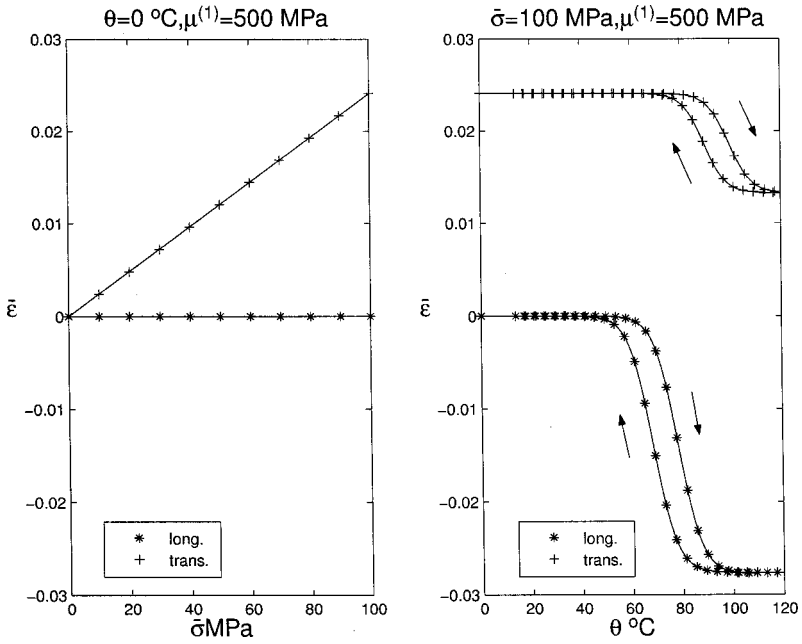


Fig. 9 Initial application of transverse shear traction of 100 MPa on the composite with subsequent temperature cycle

in the composite is biased by the combination of applied stress and internal fiber-matrix interaction stresses. In Fig. 8 where the applied stress and the fiber-matrix interactions are primarily in alignment and the fibers carry the larger portion of the load, the shape of the hysteresis remains similar to the previous case. Whereas, in Fig. 9, where most of the load is carried by the matrix in shear, hysteresis development is dominated by stress due to fiber-matrix interaction. However, a much smaller hysteresis loop now forms in the transverse direction. This result is consistent with stress development in the fibers of fiber-reinforced composites.

4.3 Case #3. Case 3 considers embedding prestressed SMA fibers in a matrix prior to attempting a heating cycle. This case differs slightly from that of Case 1 in that jump conditions are enforced with prestresses in each phase. The analysis of this case is discussed in more detail in the Appendix. The hysteresis loops in temperature-strain space in Fig. 10 have shifted towards higher temperatures without altering size or shape. We have seen in Fig. 7 that the size and shape of the hysteresis is dramatically affected by the stiffness of the surrounding matrix; however, this does not appear to be true for a pretensioning of the fibers. The shape and maximum attainable strains in Figs. 7, 8, and 10 are identical. Shape preservation of hysteresis loops under these conditions is a function of the linearity of the surrounding matrix combined with the fact that the transformation temperature is simply offset by the effective stress as seen in the definitions of F^A and F^M in Eqs. (5) and (6).

The fact that the maximum attainable thermal transformation strain is not retarded by fiber prestress has implications for control of the composite. Without using forced convection, there is no active means of controlling SMA cooling rate and its associated low temperature shape recovery. However, by manufacturing the composite with prestress in the wires, there appears to be a means of accelerating shape recovery simply by using thermal gradients. With a shift of the operational temperature range of the composite, the relative difference between ambient temperature and temperature of the composite can be increased. This increase provides a thermodynamic enhancement in convective heat loss. A higher flux out of the composite correlates to faster rebound of the SMA and the composite as a whole.

4.4 Case #4. Case 4 demonstrates three cyclic loading conditions for the composite after the SMA fibers have been preheated. The three cyclic loading simulations include longitudinal tractions, transverse tractions, and both tractions applied simulta-

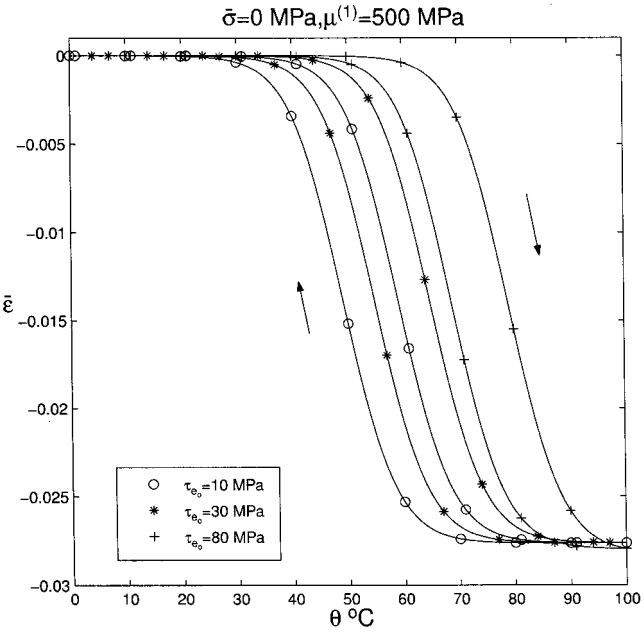


Fig. 10 Initial longitudinal prestress in the fiber with subsequent temperature cycle

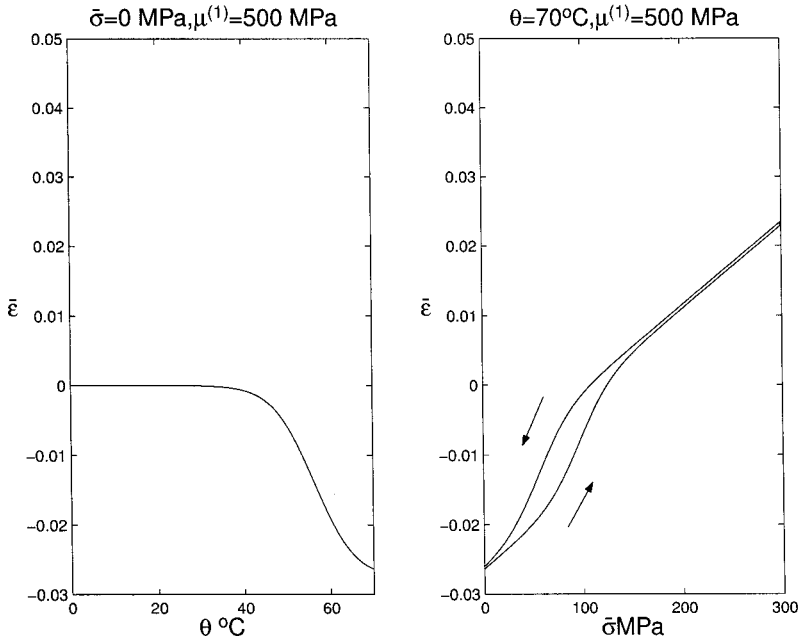


Fig. 11 Initial thermal transformation strain with a single longitudinal traction of 300 MPa on the composite

neously. The initial heating causes a transformation to austenite, and the subsequent cyclic tractions will increase and decrease the quantity of martensite.

In Fig. 11 the composite exhibits an initial heating-induced contraction followed by a single isothermal longitudinal loading cycle. Figure 12 is a slight variation on Fig. 11. In this plot the SMA fiber is heated to a variety of different temperatures followed by a single traction cycle on the composite. Notice that the composite assumes many of the characteristics of pure SMA under longitudinal tension. As the initial temperature is reduced prior to loading, hysteresis becomes less and less dramatic. Figure 13 is an application of a transverse shear load to the composite after heating the SMA. Notice the longitudinal strain of the composite caused by heating the SMA is unchanged while a hysteresis loop has emerged in the transverse shear direction. The magnitude (enclosed area) of the transverse hysteresis loop is less than that found in Fig. 11. The final loading condition in Fig. 14 is the combination of longitudinal and transverse tractions. Hysteresis loops appear in both transverse and longitudinal strain components, where the hysteresis magnitude of each cycle is less than that of either Fig. 11 or 13. Individual hysteresis loops in each stress state become smaller as the number of stress states increases.

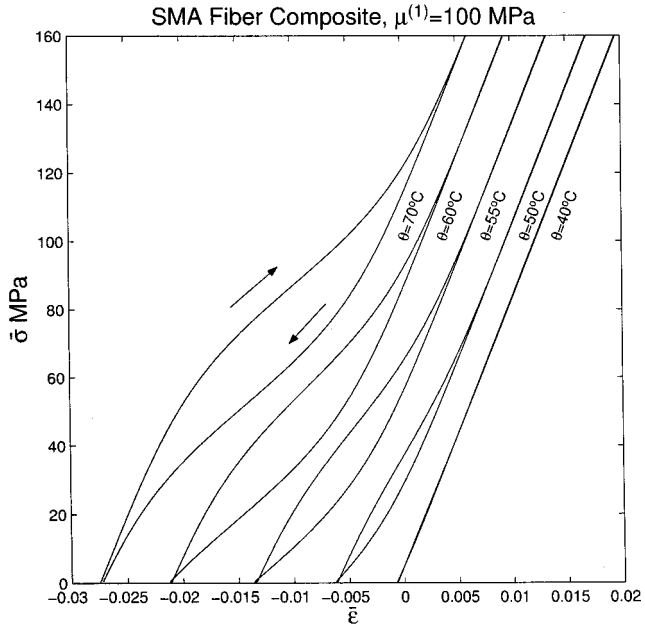


Fig. 12 Initial thermal transformation strain with a single longitudinal stress cycle of 200 MPa on the composite

5 Concluding Remarks

The objectives in this work were twofold. The first objective was to explore possible performance optimization—performance being best characterized by maximum contraction and largest bandwidth of the system response as a whole—by varying the possible ''configurations'' of SMA composites. Bandwidth here is a concept from control theory describing a systems ability to track sinusoidal inputs, where the larger the bandwidth in general implies the system can successfully track higher frequency sinusoids. However, these were found to not be complementary characteristics. Some type of tradeoff would necessarily be expected for any SMA composite. The second purpose was to determine the feasibility of including such homogenization techniques in a control-type environment. For homogenization to be a successful component of a controller it must be both relatively accurate and be computationally efficient.

In addressing the first objective, as would be expected, increasing the matrix stiffness reduced the overall contraction of the composite and required higher activation temperatures. In the absence of any criteria other than percentage contraction, a soft matrix is more desirable. However, a soft matrix will suffer in bandwidth performance for two reasons. First, the low activation temperature reduces natural conduction of heat away from the composite and slows transformation of the SMA back to martensite. Second, a soft matrix will have little restoring force. Interestingly, where tractions were applied to the composite or prestresses

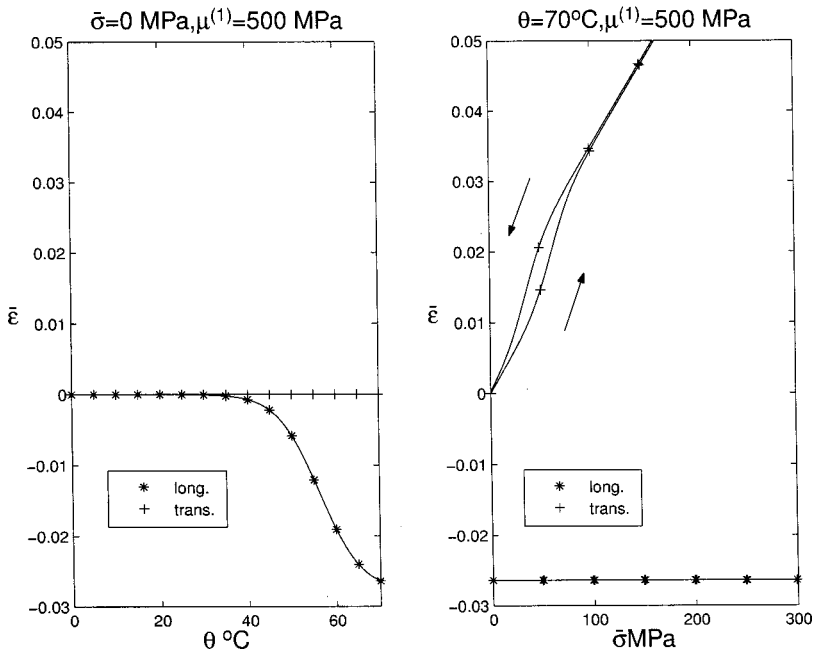


Fig. 13 Initial thermal transformation strain with a single transverse shear stress cycle of 300 MPa on the composite

created during manufacturing, the *shape* of the hysteresis remained consistent with previous thermal activations where the phases were initially stress-free. The position of the hysteresis loops simply shifted to higher temperature regimes. Because conduction is driven by temperature differential, in choosing a soft matrix but including a pretension on the fibers, it may be possible to have large percentage contraction and better response time. Of course, this improvement has an associated cost in energy requirements for heating.

In addressing the second objective, a simple three-dimensional SMA constitutive law with an embedded Duhem-Madelung type hysteresis model was used in conjunction with homogenization. A generalization of Ponte Castañeda's homogenization procedure to include prestrains was formulated to account for the complicated behavior of the SMA. This particular constitutive model successfully demonstrated the ability of a simple model to efficiently capture many complicated behaviors of an embedded SMA. In the

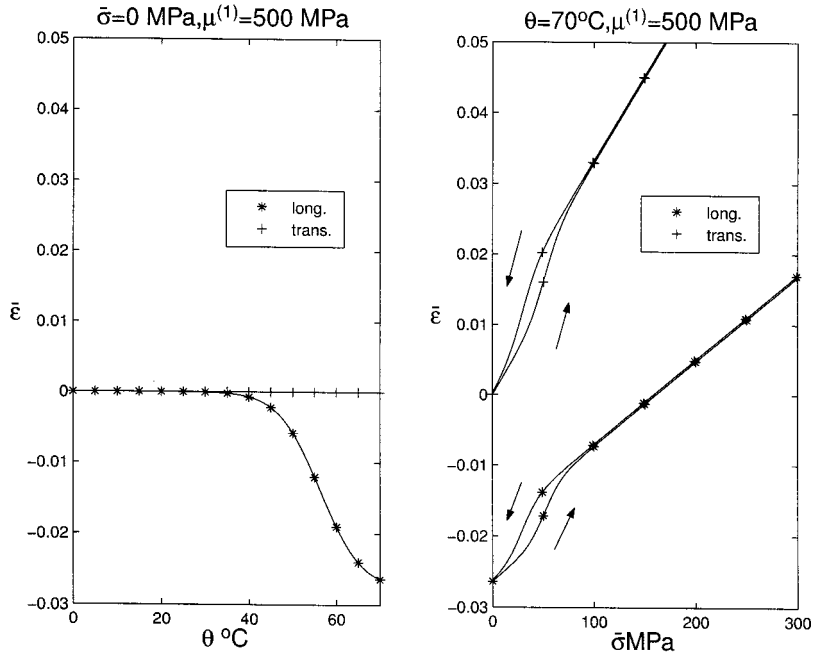


Fig. 14 Initial thermal transformation strain with a single longitudinal and transverse shear stress cycle of 300 MPa on the composite

future, better phenomenological models can be incorporated within the procedures developed in this article. By calculating average stress and strain fields for the composite, homogenization provides significant advantages over more traditional numerical techniques, where inclusion definition and nonlinear constitutive behavior significantly increase computational time.

Acknowledgments

The first author was supported by a DOE Integrated Manufacturing fellowship. The work was also supported by the Research Foundation of the University of Pennsylvania and by NSF grant DMS-99-71958.

Appendix A

Eigenstrains due to Transformation of Constrained Continuous Fibers. The strains due to thermal activation and/or pretensioning of the fibers are functions of the stresses within each phase. Fortunately, with the average stress $\overline{\boldsymbol{\sigma}}$ a known quantity there are a sufficient number of equations to solve for these stresses (and thus obtain the strains). Under the assumption that temperature distribution varies slowly over the composite body, the temperature within a RVE can be considered a constant. Jump conditions for the states of stress and strain require the normal components ($\mathbf{n}$) of stress and tangential components ($\mathbf{t}$) of strains in the phases to be continuous: $[\boldsymbol{\varepsilon}\mathbf{t}]=0$, $[\boldsymbol{\sigma}\mathbf{n}]=0$. The strains are determined by using the state of average stress $\overline{\boldsymbol{\sigma}}=c^{(1)}\boldsymbol{\sigma}^{(1)}+c^{(2)}\boldsymbol{\sigma}^{(2)}$, combined with continuity of the traction stresses and Fig. 5.

$$\sigma_{rr}^{(1)}=\overline{\sigma}_{rr}=\sigma_{rr}^{(2)}, \quad \sigma_{r3}^{(1)}=\overline{\sigma}_{r3}=\sigma_{r3}^{(2)}, \quad \sigma_{\theta r}^{(1)}=\overline{\sigma}_{\theta r}=\sigma_{\theta r}^{(2)} \tag{33}$$

and

$$\sigma_{33}^{(1)}=\left(\frac{\overline{\sigma}_{33}-c^{(2)}\sigma_{33}^{(2)}}{c^{(1)}}\right), \quad \sigma_{\theta z}^{(1)}=\left(\frac{\overline{\sigma}_{\theta z}-c^{(2)}\sigma_{\theta z}^{(2)}}{c^{(1)}}\right) \tag{34}$$

$$\sigma_{\theta\theta}^{(1)}=\left(\frac{\overline{\sigma}_{\theta\theta}-c^{(2)}\sigma_{\theta\theta}^{(2)}}{c^{(1)}}\right)$$

Likewise, tangential strain continuity implies

$$\varepsilon_{33}^{(1)}=\varepsilon_{33}^{(2)}, \quad \varepsilon_{\theta 3}^{(1)}=\varepsilon_{\theta 3}^{(2)}, \quad \varepsilon_{\theta\theta}^{(1)}=\varepsilon_{\theta\theta}^{(2)}. \tag{35}$$

Without prestresses the jump conditions between fibers and matrix with an average stress $\overline{\boldsymbol{\sigma}}$ are reduced to the following equations for the stresses in the fiber phase ($^{(2)}$), where we define $\gamma[\theta(t),\tau_e(t)]$ to be $3/2\omega/D(\xi[\theta(t),\tau_e(t)]-1)$ to simplify notation:

$$0=-\frac{2c^{(2)}\sigma_{\theta\theta}-c^{(2)}\sigma_{33}-2\overline{\sigma}_{\theta\theta}+c^{(1)}\overline{\sigma}_{rr}+\overline{\sigma}_{33}}{6c^{(1)}\mu^{(1)}}-\frac{2(\sigma_{\theta\theta}+\sigma_{33}+\overline{\sigma}_{rr})\tau_e\mu^{(2)}+3k^{(2)}(2\sigma_{\theta\theta}-\sigma_{33}-\overline{\sigma}_{rr})(\tau_e+2\gamma[\theta(t),\tau_e(t)]\mu^{(2)})}{18k^{(2)}\tau_e\mu^{(2)}} \tag{36}$$

$$0=-\frac{-c^{(2)}\sigma_{\theta\theta}+2c^{(2)}\sigma_{33}+\overline{\sigma}_{\theta\theta}+c^{(1)}\overline{\sigma}_{rr}-2\overline{\sigma}_{33}}{6c^{(1)}\mu^{(1)}}+\frac{-2(\sigma_{\theta\theta}+\sigma_{33}+\overline{\sigma}_{rr})\tau_e\mu^{(2)}+3k^{(2)}(\sigma_{\theta\theta}-2\sigma_{33}+\overline{\sigma}_{rr})(\tau_e+2\gamma[\theta(t),\tau_e(t)]\mu^{(2)})}{18k^{(2)}\tau_e\mu^{(2)}} \tag{37}$$

$$0=-\frac{\gamma[\theta(t),\tau_e(t)]\sigma_{\theta 3}}{\tau_e}+\frac{-(c^{(2)}\sigma_{\theta 3})+\overline{\sigma}_{\theta 3}}{2c^{(1)}\mu^{(1)}}-\frac{\sigma_{\theta 3}}{2\mu^{(2)}} \tag{38}$$

where

$$\tau_e=\frac{\sqrt{2\sigma_{\theta\theta}^2+3\sigma_{\theta 3}^2+2\sigma_{33}^2+3\overline{\sigma}_{\theta r}^2-2\sigma_{33}\overline{\sigma}_{rr}+2\overline{\sigma}_{rr}^2-2\sigma_{\theta\theta}(\sigma_{33}+\overline{\sigma}_{rr})+3\overline{\sigma}_{r3}^2}}{\sqrt{6}}.$$

If a uniform initial tension $\boldsymbol{\sigma}_o^{(2)}$ is placed on trained SMA fibers at low temperature and the surrounding matrix material is subsequently poured and set around them, the strain in the fibers will be equal to $\mathbf{M}\boldsymbol{\sigma}_o^{(2)}$, while the strain in the matrix will be equal to zero. Once the initial tension is removed from the fibers, the composite as a whole reaches some new state of equilibrium, where each component will have a prestress. Equations (33) through (38) remain valid, but the stress in each phase now must include the prestresses $\overline{\boldsymbol{\sigma}}_{\text{pre}}^{(r)}$. The relationships in Eq. (30) used with a linear Hashin-Shtrikman estimate for $\tilde{\mathbf{M}}$ (both fiber and matrix are linear at low temperature) can determine $\overline{\boldsymbol{\sigma}}_{\text{pre}}^{(r)}$, where the prestrains are $\mathbf{M}\boldsymbol{\sigma}_o^{(2)}$ and zero for the fibers and matrix, respectively.

For $\overline{\boldsymbol{\sigma}}=0$, Eq. (33) defines six of the twelve states of stress. Equation (34) reduces the remaining six unknown states of stress to three. The three relationships in Eq. (35) combined with the constitutive behaviors allow us to solve for the remaining states of stress in three coupled nonlinear equations defined by Eqs. (36) through (38). These equations were solved with Matlab scripts using a Levenberg-Marquardt [35] method.

References

[1] Willis, J. R., 1981, "Variational and Related Methods for the Overall Properties of Composites," *Advances in Applied Mechanics*, C. Y. Yih, ed., Academic Press, San Diego, CA, pp. 1–78.
[2] Willis, J. R., 1983, "The Overall Elastic Response of Composite-Materials," ASME J. Appl. Mech., **50**, pp. 1202–1209.
[3] Christenesen, R. M., 1979, *Mechanics of Composite Materials*, John Wiley and Sons, New York.
[4] Talbot, D. R. S., and Willis, J. R., 1985, "Variational Principles for Nonlinear Inhomogeneous Media," IMA J. Appl. Math., **35**, pp. 39–54.
[5] Ponte Castañeda, P., 1991, "The Effective Mechanical Properties of Nonlinear Isotropic Composites," J. Mech. Phys. Solids, **39**(1), pp. 45–71.
[6] Ponte Castañeda, P., 1996, "Exact Second-Order Estimates for the Effective Mechanical Properties of Nonlinear Composite Materials," J. Mech. Phys. Solids, **44**, pp. 827–862.
[7] Abeyaratne, R., and Knowles, J. K., 1993, "A Continuum Model for Thermoelastic Solid Capable of Undergoing Phase Transitions," J. Mech. Phys. Solids, **41**, pp. 541–571.
[8] Tanaka, K., 1886, "A Thermomechanical Sketch of Shape Memory Effect: One Dimensional Tensile Behavior," Res. Mech., **18**, pp. 251–263.
[9] Tanaka, K., Kobayashi, S., and Sato, Y., 1986, "Thermomechanics of Transformation Pseudoelasticity and Shape Memory Effect in Alloys," Int. J. Plast., **2**, 59–72.
[10] Liang, C., and Rogers, C. A., 1990, "One-Dimensional Thermomechanical Constitutive Relations of Shape Memory Materials," J. Intell. Mater. Syst. Struct., **1**, pp. 207–234.

[11] Leblond, J. B., Devaux, J. D., and Devaux, J. C., 1989, ''Mathematical Modeling of Transformation Plasticity in Steels, I: Case of Ideal-Plastic Phases,'' Int. J. Plast., **5**, pp. 551–572.
[12] Ivshin, Y., and Pence, T. J., 1994, ''A Constitutive Model for Hysteretic Phase Transition Behavior,'' Int. J. Eng. Sci., **32**(4), pp. 681–704.
[13] Falk, F., 1980, ''Model Free Energy, Mechanics, and Thermodynamics of Shape Memory Alloys,'' Acta Metall. Mater., **28**, pp. 1773–1780.
[14] Falk, F., 1983, ''Ginzburg-Landau Theory of Static Domain Walls in Shape-Memory Alloys,'' Z. Phys. B: Condens. Matter, **51**, pp. 177–185.
[15] Brinson, L. C., and Huang, M. S., 1996, ''Simplifications and Comparisons of Shape Memory Alloy Constitutive Models,'' J. Intell. Mater. Syst. Struct., **7**, pp. 108–104.
[16] Boyd, J. G., and Lagoudas, D. C., 1995, ''A Thermodynamic Constitutive Model for Shape Memory Materials, Part I. The Monolithic Shape Memory Alloys,'' Int. J. Plast., **12**, pp. 805–841.
[17] Bekker, A., and Brinson, L. C., 1997, ''Temperature-Induced Phase Transformation in a Shape Memory Alloy: Phase Diagram Based Kinetics Approach,'' J. Mech. Phys. Solids, **45**(6), pp. 949–988.
[18] Muller, I., and Xu, H., 1991, ''On the Pseudo-Elastic Hysteresis,'' Acta Metall. Mater., **39**(3), pp. 263–271.
[19] Boyd, J. G., and Lagoudas, D. C., 1994, ''Thermomechanical Response of Shape Memory Composites,'' J. Intell. Mater. Syst. Struct., **4**, pp. 333–346.
[20] Bo, Z., Lagoudas, D. C., and Miller, D., 1999, ''Material Characterization of SMA Actuators Under Nonproportional Thermomechanical Loading,'' ASME J. Eng. Mater. Technol., **121**, pp. 75–85.
[21] Bo, Z., and Lagoudas, D. C., 1999, ''Thermomechanical Modeling of Polycrystalline SMAs Under Cyclic Loading, Part I: Theoretical Derivations,'' Int. J. Eng. Sci., **37**, July.
[22] Lagoudas, D. C., Boyd, J. G., and Bo, Z., 1994, ''Micromechanics of Active Composites With SMA Fibers,'' ASME J. Eng. Mater. Technol., **116**, pp. 337–347.
[23] Brinson, L. C. et al., 1997, ''Analysis of Controlled Beam Using SMA Wires,'' J. Intell. Mater. Syst. Struct., **8**(1), pp. 12–25.
[24] de Blonk, B. J., and Lagoudas, D. C., 1997, ''Actuation of Elastomeric Rods With Embedded Two-Way Shape Memory Alloy Actuators,'' Smart Mater. Struct., **6**, pp. 265–277.
[25] Bernardini, D., 2001, ''On the Macroscopic Free Energy Functions for Shape Memory Alloys,'' J. Mech. Phys. Solids, pp. 813–837.
[26] Briggs, J., and Ostrowski, J., 2002, ''Experimental Feedforward and Feedback Control of a 1-D SMA Composite,'' Smart Mater. Struct. **11**(1), pp. 9–23.
[27] Bondaryev, E. N., and Wayman, C. M., 1998, ''Some Stress-Strain-Temperature Relationships for Shape Memory Alloys,'' Metall. Trans. A, **19A**, pp. 2407–2413.
[28] Ivshin, Y., and Pence, T. J., 1994, ''A Thermomechanical Model for a One Variant Shape Memory Material,'' J. Intell. Mater. Syst. Struct., **5**, pp. 455–473.
[29] Lagoudas, D. C., Bo, Z., and Qidwai, M. A., 1996, ''A Unified Thermodynamic Constitutive Model for SMA and Finite Element Analysis of Active Matrix Composites,'' Mech. Compos. Mat. Struct., **3**, pp. 153–179.
[30] Ponte Castañeda, P., and Suquet, P., 1998, ''Nonlinear Composites,'' Adv. Appl. Mech., **34**, pp. 171–302.
[31] Levin, V. M., 1968, ''On the Coefficients of Thermal Expansion of Heterogeneous Materials,'' Mekh. Tver. Tela, p. 88.
[32] Walpole, L. J., 1966, ''On the Bounds for the Overall Elastic Moduli of Inhomogeneous Systems—I,'' J. Mech. Phys. Solids, **14**, pp. 151–262.
[33] Walpole, L. J., 1966, ''On the Bounds for the Overall Elastic Moduli of Inhomogeneous Systems—II,'' J. Mech. Phys. Solids, **14**, pp. 289–301.
[34] Walpole, L. J., 1969, ''On the Overall Elastic Moduli of Composite Materials,'' J. Mech. Phys. Solids, **17**, pp. 235–251.
[35] Marquardt, D. W., 1963, J. Soc. Ind. Appl. Math., **11**, pp. 431–441.

Antiplane Crack Problem in Functionally Graded Piezoelectric Materials

Chunyu Li

G. J. Weng[1]
Professor
e-mail: weng@jove.rutgers.edu
Fellow ASME

Department of Mechanical and Aerospace Engineering,
Rutgers University,
New Brunswick, NJ 08903

In this paper the problem of a finite crack in a strip of functionally graded piezoelectric material (FGPM) is studied. It is assumed that the elastic stiffness, piezoelectric constant, and dielectric permitivity of the FGPM vary continuously along the thickness of the strip, and that the strip is under an antiplane mechanical loading and in-plane electric loading. By using the Fourier transform, the problem is first reduced to two pairs of dual integral equations and then into Fredholm integral equations of the second kind. The near-tip singular stress and electric fields are obtained from the asymptotic expansion of the stresses and electric fields around the crack tip. It is found that the singular stresses and electric displacements at the tip of the crack in the functionally graded piezoelectric material carry the same forms as those in a homogeneous piezoelectric material but that the magnitudes of the intensity factors are dependent upon the gradient of the FGPM properties. The investigation on the influences of the FGPM graded properties shows that an increase in the gradient of the material properties can reduce the magnitude of the stress intensity factor. [DOI: 10.1115/1.1467091]

1 Introduction

Since the Curie brothers announced their pioneering discovery of the piezoelectric effect in 1880, there have been a number of studies devoted to the theoretical analysis and engineering application of the piezoelectric materials ([1–3]). Up to now, piezoelectric materials have been found to have wide applications in the smart systems of aerospace, automotive, medical, and electronic fields due to the intrinsic coupling characteristics between their electric and mechanical fields. However, as the piezoelectric materials are being extensively used as actuators or transducers in the technologies of smart and adaptive systems, the mechanical reliability and durability of these materials become increasingly important. For example, piezoelectric bimorphs are a particular type of piezoelectric devices ([4,5]). They usually consist of two long and thin piezoelectric elements, which are bonded over their long faces by using adhesive epoxy resin, and suitably covered with electrodes. The principal disadvantage of this kind of materials is that the bonding agent may crack at low temperature and creep at high temperature. These drawbacks may lead to lifetime limitation and restrict the utility of piezoelectric actuator in the area of measured devices that require high reliability ([6]). To meet the demand of advanced piezoelectric materials in lifetime and reliability and with the help of the development in modern material processing technology, the concept of functionally graded materials has recently been extended into the piezoelectric materials ([6–10]). These new kind of materials with continuously varying properties may be called functionally graded piezoelectric materials (FGPM). The advantages of using a device wholly made of the FGPM or using an FGPM as a transit layer instead of the bonding agent are that no discernible internal seams or boundaries exist, and that no internal stress peaks are caused when voltage is applied and thus failure from internal debonding or from stress peaks developed in conventional bimorphs can be avoided ([7,11]). The preliminary results of fatigue tests indicate that the lifetime of the actuator can be significantly increased ([11]).

Because of the brittle nature of piezoelectric ceramics the problems on fracture mechanics of piezoelectric materials have received much attention in recent decades. Some important contributions include Parton [12], Deeg [13], Sosa and Pak [14], Pak [15], Suo et al. [16], Wang [17], Dunn [18], Park [19], Park and Sun [20], Zhang and Tong [21], and Shindo et al. [22], among others. These literatures examined the general solutions of the mechanical and electric coupling and systematically studied the crack problems in piezoelectric materials. However, because the FGPM are just an emerging class of piezoelectric materials, researches on their mechanic behaviors are still very few ([23,24]). To our knowledge, there is still no article considering the fracture problem in an FGPM.

The main objective of this paper is to explore the fracture mechanics of FGPM. We consider a finite crack problem in a strip of FGPM under antiplane mechanical loadings and in-plane electric loadings. We assume that the elastic stiffness, piezoelectric constant, and dielectric permittivity of the FGPM vary continuously along the thickness of the strip. But in order to overcome the complexity of the mathematics involved, we shall focus in this initial study on a special class of FGPMs in which the variations of these properties are in the same proportion. These would allow us to shed some light on the influence of the material gradient upon the stress and electric intensity factors. Instead of using the approximate impermeable boundary conditions along the crack surfaces as many studies have done, we shall study the crack problem using the exact boundary conditions ([21]). Zhang and Tong [21] have found by analyzing an elliptical cavity that the two commonly used boundary conditions (i.e., impermeable and permeable) are actually two limiting cases of the exact boundary conditions when the cavity approaches a slit crack. In our work, the crack problem is first reduced into two pairs of dual integral equations by using the Fourier transforms. They are then reduced into Fredholm integral equations of the second kind. Shindo et al. [22] have developed this approach for crack problems in a homogeneous piezoelectric ceramic strip, and it is now extended to the crack problems in FGPMs. The closed forms of singular stress and electric fields around the crack tip are obtained. The influences of

[1]Author 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 MECHANICSASME . Manuscript received by the ASME Applied Mechanics Division, May 30, 2001; final revision, November 21, 2001. Associate Editor: M.-J. Pindera. 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.

the FGPM graded properties on the stress intensity factor and electric displacement intensity factor are investigated. It is observed that an increase in the gradient of the material properties can reduce the magnitudes of the intensity factors.

2 Basic Equations of Functionally Graded Piezoelectric Materials

Consider a linear piezoelectric material. The governing equations in the Cartesian coordinates $x_i (i=1,2,33)$ are given by

$$\sigma_{ij,j}=0 \quad \text{and} \quad D_{i,i}=0, \tag{1}$$

where σ_{ij} is the stress tensor, D_i the electric displacement vector, a comma denotes partial differentiation with respect to the coordinate x_i, and the Einstein summation convention over repeated indices is used.

For an anisotropic functionally graded piezoelectric material, the constitutive relation is

$$\sigma_{ij}=c_{ijkl}\varepsilon_{kl}-e_{kij}E_k, \tag{2}$$

$$D_i=e_{ikl}\varepsilon_{kl}+\in_{ik}E_k, \tag{3}$$

where ε_{kl} is the strain tensor, E_k the electric field intensity, and c_{ijkl}, e_{kij}, $\in_{ik}$ are the elastic stiffness tensor, the piezoelectric tensor, and the dielectric tensor, respectively. Unlike in a homogeneous piezoelectric material, the c_{ijkl}, e_{kij}, and $\in_{ik}$ are now functions of the coordinates $x_i (i=1,2,3)$.

From the restrictions set by the crystal symmetry, the interchange symmetries of the material property tensors are

$$c_{ijkl}=c_{ijlk}=c_{jikl}=c_{klij}, \quad e_{kij}=e_{kji}, \quad \in_{ik}=\in_{ki}. \tag{4}$$

The relation between the strain tensor and the displacement is given by

$$\varepsilon_{ij}=\tfrac{1}{2}(u_{i,j}+u_{j,i}), \tag{5}$$

where u_i is the displacement component and the electric field intensity is

$$E_i=-\phi_{,i}, \tag{6}$$

where ϕ is the electric potential.

Suppose that the crack is filled with air, then the electric potential inside the crack should satisfy the following equation:

$$\nabla^2\phi=0, \tag{7}$$

and the relation between the electric displacement and the electric field intensity is given by

$$D_i=\in_0 E_i, \tag{8}$$

where $\in_0$ is the permittivity of vacuum.

It is well known that the piezoelectric effect exists only in crystals that do not have a center of symmetry. What has been commonly studied in the literature and also one of the most useful types is the transversely isotropic piezoelectric material belonging to the hexagonal crystal class 6 mm. For this kind of material, by taking direction 3 to be the axis of symmetry, the nonzero material parameters are

$$\begin{gathered} c_{1111}=c_{2222}=c_{11}, \quad c_{1122}=c_{12}, \\ c_{1133}=c_{2233}=c_{13}, \quad c_{3333}=c_{33}, \\ c_{2323}=c_{3131}=c_{44}, \quad c_{1212}=\tfrac{1}{2}(c_{11}-c_{12})=c_{66}, \\ e_{311}=e_{322}=e_{31}, \quad {}_{333}=e_{33}, \quad e_{113}=e_{223}=e_{15}, \\ \in_{22}=\in_{11}. \end{gathered} \tag{9}$$

3 Antiplane Crack Problem

To obtain the singular stress and electric fields, we consider the simplest crack problem, i.e., the antiplane Griffith crack. The finite crack of length $2c$ is embedded in an infinite long strip of a transversely isotropic functionally graded piezoelectric material. The electric and antiplane mechanical loadings are shown as Fig. 1. A set of Cartesian coordinates (x,y,z) is attached at the center of the strip and the principal axes are taken to coincide with the reference axes x, y, and z, in which z is the poling axis. It is also assumed that the strip is thick enough in the z-direction to allow a state of antiplane shear. The crack is situated along the plane $y=0$.

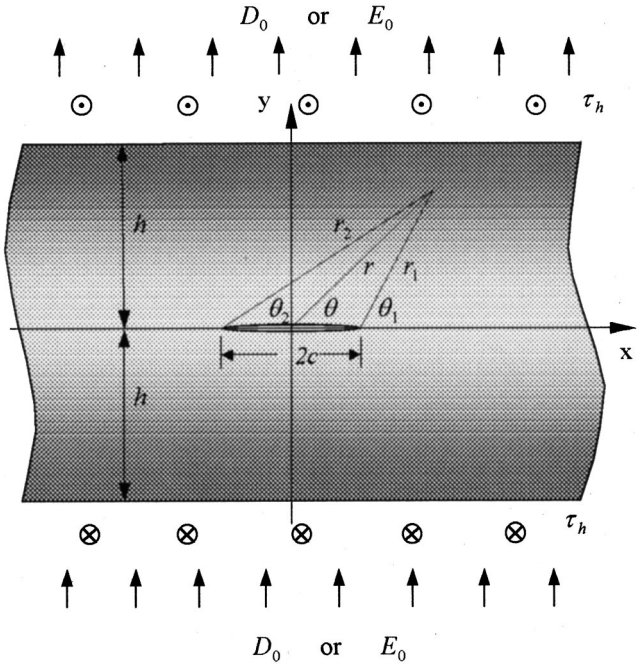


Fig. 1 An antiplane crack problem of functionally graded piezoelectric materials

Under the antiplane mechanical loading and in-plane electric loading, only the out-of-plane displacement and in-plane electric fields need to be considered. We take the notations as follows:

$$\begin{gathered} u_x=u_y=0, \quad u_z=u_z(x,y), \\ E_x=E_x(x,y), \quad E_y=E_y(x,y), \quad E_z=0 \end{gathered} \tag{10}$$

where (u_x,u_y,u_z) are components of displacements and (E_x, E_y, E_z) the components of the electric field vector. In this case, the constitutive relations Eqs. (2) and (3) can be simplified as

$$\begin{gathered} \sigma_{xz}=c_{44}u_{z,x}-e_{15}E_x, \\ \sigma_{yz}=c_{44}u_{z,y}-e_{15}E_y, \\ D_x=e_{15}u_{z,x}+\in_{11}E_x, \\ D_y=e_{15}u_{z,y}+\in_{11}E_y \end{gathered} \tag{11}$$

where $(\sigma_{xz}, \sigma_{yz})$ are the stress components and (D_x, D_y) the components of the electric displacement vector. The subscript comma denotes the partial derivative with respect to the coordinates.

As shown in Fig. 1 with the shaded region showing an increasing intensity, the material property parameters are taken to vary continuously along the y-direction inside the strip. To achieve our objective of obtaining the stress and electric fields in the FGPM, we consider the following distribution to simplify the problem:

$$\begin{gathered} c_{44}(y)=c_{44}^0(1+\alpha|y|)^k, \quad e_{15}(y)=e_{15}^0(1+\alpha|y|)^k, \\ \in_{11}(y)=\in_{11}^0(1+\alpha|y|)^k, \end{gathered} \tag{12}$$

where k is a constant and α can be determined by the material property parameters. If we use the notations $(c_{44}^0, e_{15}^0, \in_{11}^0)$ and

$(c_{44}^h, e_{15}^h, \in_{11}^h)$ to stand for the values at the plane $y=0$ and the surfaces $y=\pm h$ of the strip, respectively, the expression of α will be

$$\alpha=\frac{1}{h}\left(\sqrt[k]{c_{44}^h/c_{44}^0}-1\right). \tag{13}$$

Due to the assumed proportionality of the property variations in Eq. (12), this gradient parameter α also serves to characterize the variation of other material constants.

Substituting Eqs. (11) and (12) into Eq. (1), we can obtain the governing equations in the following forms

$$\begin{aligned} c_{44}^0\left(\nabla^2 u_z+\frac{k\alpha}{1+\alpha y}u_{z,y}\right)+e_{15}^0\left(\nabla^2\phi+\frac{k\alpha}{1+\alpha y}\phi_{,y}\right)=0,\\ e_{15}^0\left(\nabla^2 u_z+\frac{k\alpha}{1+\alpha y}u_{z,y}\right)-\in_{11}^0\left(\nabla^2\phi+\frac{k\alpha}{1+\alpha y}\phi_{,y}\right)=0, \end{aligned} \tag{14}$$

where $\nabla^2=\partial^2/\partial x^2+\partial^2/\partial y^2$ is the two-dimensional Laplacian operator in the variables x and y.

Because of the assumed symmetry about the y-axis in geometry and the applied loading, it is sufficient to consider the problem for $x\geqslant 0, y\geqslant 0$ part only. Thus the boundary conditions for this problem become

$$\sigma_{yz}(x,0)=0, \quad 0\leqslant x<c, \tag{15}$$

$$u_z(x,0)=0, \quad c\leqslant x<\infty, \tag{16}$$

$$E_x(x,0^+)=E_x^c(x,0^-), \quad 0\leqslant x<c, \tag{17}$$

$$\phi(x,0)=0, \quad c\leqslant x<\infty, \tag{18}$$

$$D_y(x,0^+)=D_y^c(x,0^-), \quad 0\leqslant x<c, \tag{19}$$

$$\sigma_{yz}(x,0^+)=\sigma_{yz}(x,0^-), \quad c\leqslant x<\infty, \tag{20}$$

$$\sigma_{yz}(x,h)=\tau_h, \quad 0\leqslant x<\infty. \tag{21}$$

Case 1:

$$D_y(x,h)=D_0, \quad 0\leqslant x<\infty. \tag{22}$$

Case 2:

$$E_y(x,h)=E_0, \quad 0\leqslant x<\infty \tag{23}$$

where the superscript c stands for the electric quantities in the void inside the crack, D_0 is a uniform electric displacement, and E_0 a uniform electric field applied externally. The shear stress can be expressed as

$$\tau_h=\begin{cases} \dfrac{\bar{c}_{44}^h}{c_{44}^h}\tau_0-\dfrac{e_{15}^h}{\in_{11}^h}D_0, & \text{(Case 1)},\\ \tau_0-e_{15}^h E_0, & \text{(Case 2)}, \end{cases} \tag{24}$$

where τ_0 is a uniform shear stress at zero electric loading and $\bar{c}_{44}^h=c_{44}^h+(e_{15}^h)^2/\in_{11}^h$ is the piezoelectric stiffened elastic constant.

4 Solution of the Problem

Since $(e_{15}^0)^2+c_{44}^0\in_{11}^0\neq 0$ in general, it can be obtained from Eq. (14) that

$$\nabla^2 u_z+\frac{k\alpha}{1+\alpha y}u_{z,y}=0, \tag{25}$$

$$\nabla^2\phi+\frac{k\alpha}{1+\alpha y}\phi_{,y}=0. \tag{26}$$

Now the problem has been converted to finding the displacement field and the electric potential that satisfy Eqs. (25) and (26) and meet the boundary conditions Eqs. (15)–(23).

Considering the symmetry of the problem about the y-axis, we introduce the Fourier cosine transforms to Eqs. (25) and (26). The solutions are obtained as

$$\begin{aligned} u_z(x,y)=\frac{2}{\pi}\int_0^\infty (1+\alpha y)^{-\beta}\Bigg\{A_1(s)I_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\\ +A_2(s)K_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\Bigg\}\cos(sx)ds+a_1 y, \end{aligned} \tag{27}$$

$$\begin{aligned} \phi(x,y)=\frac{2}{\pi}\int_0^\infty (1+\alpha y)^{-\beta}\Bigg\{B_1(s)I_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\\ +B_2(s)K_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\Bigg\}\cos(sx)ds-b_1 y, \end{aligned} \tag{28}$$

where $\beta=(k-1)/2$, I_β, and K_β are the first and second kind of modified Bessel's functions, respectively, $A_i(s)$, $B_i(s)(i=1,2)$ are the unknowns to be solved, and a_1, b_1 are real constants, which will be determined from the edge loading conditions.

The stress, electric field, and electric displacement expressions can be derived from Eqs. (27) and (28) along with Eq. (11) as

$$\begin{aligned} \sigma_{xz}(x,y)=-\frac{2}{\pi}\int_0^\infty s(1+\alpha y)^{-\beta}\Bigg\{[c_{44}(y)A_1(s)\\ +e_{15}(y)B_1(s)]I_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]+[c_{44}(y)A_2(s)\\ +e_{15}(y)B_2(s)]K_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\Bigg\}\sin(sx)ds, \end{aligned} \tag{29}$$

$$\begin{aligned} \sigma_{yz}(x,y)=-\frac{2}{\pi}\int_0^\infty\Bigg(\Bigg\{\beta\alpha(1+\alpha y)^{-\beta-1}I_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\\ -s(1+\alpha y)^{-\beta}I_\beta'\left[(1+\alpha y)\frac{s}{\alpha}\right]\Bigg\}[c_{44}(y)A_1(s)\\ +e_{15}(y)B_1(s)]+\Bigg\{\beta\alpha(1+\alpha y)^{-\beta-1}K_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\\ -s(1+\alpha y)^{-\beta}K_\beta'\left[(1+\alpha y)\frac{s}{\alpha}\right]\Bigg\}[c_{44}(y)A_2(s)\\ +e_{15}(y)B_2(s)]\Bigg)\cos(sx)ds+c_{44}(y)a_1-e_{15}(y)b_1, \end{aligned} \tag{30}$$

$$\begin{aligned} E_x(x,y)=\frac{2}{\pi}\int_0^\infty s(1+\alpha y)^{-\beta}\Bigg\{B_1(s)I_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\\ +B_2(s)K_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\Bigg\}\sin(sx)ds, \end{aligned} \tag{31}$$

$$\begin{aligned} E_y(x,y)=\frac{2}{\pi}\int_0^\infty\Bigg(\Bigg\{\beta\alpha(1+\alpha y)^{-\beta-1}I_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\\ -s(1+\alpha y)^{-\beta}I_\beta'\left[(1+\alpha y)\frac{s}{\alpha}\right]\Bigg\}B_1(s)\\ +\Bigg\{\beta\alpha(1+\alpha y)^{-\beta-1}K_\beta\left[(1+\alpha y)\frac{s}{\alpha}\right]\\ -s(1+\alpha y)^{-\beta}K_\beta'\left[(1+\alpha y)\frac{s}{\alpha}\right]\Bigg\}B_2(s)\Bigg)\\ \times\cos(sx)ds+b_1, \end{aligned} \tag{32}$$

$$D_x(x,y)=-\frac{2}{\pi}\int_0^\infty s(1+\alpha y)^{-\beta}\Big\{[e_{15}(y)A_1(s) - \in_{11}(y)B_1(s)]I_\beta\Big[(1+\alpha y)\frac{s}{\alpha}\Big]+[e_{15}(y)A_2(s) - \in_{11}(y)B_2(s)]K_\beta\Big[(1+\alpha y)\frac{s}{\alpha}\Big]\Big\}\sin(sx)ds, \quad (33)$$

$$D_y(x,y)=-\frac{2}{\pi}\int_0^\infty\Big(\Big\{\beta\alpha(1+\alpha y)^{-\beta-1}I_\beta\Big[(1+\alpha y)\frac{s}{\alpha}\Big] - s(1+\alpha y)^{-\beta}I'_\beta\Big[(1+\alpha y)\frac{s}{\alpha}\Big]\Big\}[e_{15}(y)A_1(s) - \in_{11}(y)B_1(s)]+\Big\{\beta\alpha(1+\alpha y)^{-\beta-1}K_\beta\Big[(1+\alpha y)\frac{s}{\alpha}\Big] - s(1+\alpha y)^{-\beta}K'_\beta\Big[(1+\alpha y)\frac{s}{\alpha}\Big]\Big\}[e_{15}(y)A_2(s) - \in_{11}(y)B_2(s)]\Big)\cos(sx)ds+e_{15}(y)a_1 + \in_{11}(y)b_1, \quad (34)$$

where I'_β and K'_β are derivatives of I_β and K_β, respectively.

On the other hand, in the vacuum between the crack surfaces, the electric potential that satisfies Eq. (7) is

$$\phi^c(x,y)=\frac{2}{\pi}\int_0^\infty C(s)\sinh(sy)\cos(sx)ds, \quad 0\leqslant x<c \quad (35)$$

where $C(s)$ is also an unknown. The electric fields and electric displacements in the void inside the crack can be obtained as

$$E_x^c(x,0)=0, \quad 0\leqslant x<c, \quad (36)$$

$$E_y^c(x,0)=-\frac{2}{\pi}\int_0^\infty sC(s)\cos(sx)ds, \quad 0\leqslant x<c, \quad (37)$$

$$D_x^c(x,0)=0, \quad 0\leqslant x<c, \quad (39)$$

$$D_y^c(x,0)=-\frac{2}{\pi}\int_0^\infty \in_0 sC(s)\cos(sx)ds, \quad 0\leqslant x<c. \quad (40)$$

By applying the edge loading conditions of Eqs. (21)–(24), the following relations between the unknown functions can be found:

$$A_2(s)=R_{21}A_1(s), \quad (41)$$

$$B_2(s)=R_{21}B_1(s), \quad (42)$$

where

$$R_{21}=-\frac{\beta\alpha(1+\alpha h)^{-1}I_\beta[(1+\alpha h)s/\alpha]-sI'_\beta[(1+\alpha h)s/\alpha]}{\beta\alpha(1+\alpha h)^{-1}K_\beta[(1+\alpha h)s/\alpha]-sK'_\beta[(1+\alpha h)s/\alpha]}. \quad (43)$$

Meanwhile the constants a_1, b_1 can be determined as follows:

$$a_1=\frac{\in^h_{11}\tau_h+e^h_{15}D_0}{c^h_{44}\in^h_{11}+(e^h_{15})^2}, \quad b_1=\frac{c^h_{44}D_0-e^h_{15}\tau_h}{c^h_{44}\in^h_{11}+(e^h_{15})^2}, \quad \text{(Case 1)}, \quad (43)$$

$$a_1=\frac{\tau_h+e^h_{15}E_0}{c^h_{44}}, \quad b_1=E_0, \quad \text{(Case 2)}. \quad (44)$$

Substitutions of Eqs. (31), (36), and (28) into the boundary conditions of Eqs. (17) and (18) yield a pair of dual integral equations of the following forms:

$$\int_0^\infty sB_1(s)[I_\beta(s/\alpha)+R_{21}K_\beta(s/\alpha)]\sin(sx)ds=0, \quad 0\leqslant x<c, \quad (45)$$

$$\int_0^\infty B_1(s)[I_\beta(s/\alpha)+R_{21}K_\beta(s/\alpha)]\cos(sx)ds=0, \quad c\leqslant x<\infty. \quad (46)$$

If we define

$$B_1(s)[I_\beta(s/\alpha)+R_{21}K_\beta(s/\alpha)]=\frac{\pi c^2}{2}\int_0^1\sqrt{\xi}\Phi(\xi)J_0(sc\xi)d\xi, \quad (47)$$

where $J_0(\ )$ is the zero-order Bessel function of the first kind, then Eq. (46) is satisfied automatically and the satisfaction of Eq. (45) requires that the function $\Phi(\xi)=0$. Therefore we can determine that the unknown function $B_1(s)=0$ and then straightly $B_2(s)=0$.

Noting above results of the unknown functions $B_1(s)$, $B_2(s)$, substitutions of Eqs. (27) and (32) into the boundary conditions (15) and (16) yield another pair of dual integral equations

$$\int_0^\infty sF(s)A(s)\cos(sx)ds=\frac{\pi}{2}\frac{c^0_{44}a_1-e^0_{15}b_1}{c^0_{44}}, \quad 0\leqslant x<c, \quad (48)$$

$$\int_0^\infty A(s)\cos(sx)ds=0, \quad c\leqslant x<\infty, \quad (49)$$

where

$$A(s)=A_1(s)[I_\beta(s/\alpha)+R_{21}K_\beta(s/\alpha)], \quad (50)$$

$$F(s)=\frac{[\beta\alpha I_\beta(s/\alpha)-sI'_\beta(s/\alpha)]+R_{21}[\beta\alpha K_\beta(s/\alpha)-sK'_\beta(s/\alpha)]}{s[I_\beta(s/\alpha)+R_{21}K_\beta(s/\alpha)]}. \quad (51)$$

Following the method proposed by Copson [25], the solution that satisfies Eqs. (48), (49) can be written as

$$A(s)=\frac{\pi c^2}{2}\frac{c^0_{44}a_1-e^0_{15}b_1}{c^0_{44}}\int_0^1\sqrt{\xi}\Psi(\xi)J_0(sc\xi)d\xi, \quad (52)$$

which satisfies Eq. (49) automatically. The satisfaction of Eq. (50) requires that function $\Psi(\xi)$ be governed by the following standard Fredholm integral equation of the second kind:

$$\Psi(\xi)+\int_0^1\Psi(\xi)G(\xi,\eta)d\eta=\sqrt{\xi}. \quad (53)$$

The kernel function $G(\xi,\eta)$ in Eq. (53) is

$$G(\xi,\eta)=\sqrt{\xi\eta}\int_0^\infty s[F(s/c)-1]J_0(s\xi)J_0(s\eta)ds. \quad (54)$$

5 Singular Stress and Electric Fields Around the Crack Tip

In order to obtain the stress and electric fields around the crack tip we follow the analysis of Sih and Embley [26] on the stress singularity at the crack tip to consider the singular behaviors of the stress, electric field, and electric displacement by setting $s\to\infty$.

Integration of $A(s)$ in Eq. (52) by parts yields

$$A(s)=\frac{\pi c}{2}\frac{c^0_{44}a_1-e^0_{15}b_1}{c^0_{44}}\frac{1}{s}\Big\{\Psi(1)J_1(sc) - \int_0^1\xi J_1(sc\xi)\frac{d}{d\xi}\Big[\frac{\Psi(\xi)}{\sqrt{\xi}}\Big]d\xi\Big\}, \quad (55)$$

where $J_1(\ )$ is the first-order Bessel function of the first kind. The integral in Eq. (55) is bounded at the crack tip $x=\pm c$. Thus the singular behaviors of the stress and electric fields are governed by the leading term containing $\Psi(1)$. From Eqs. (29)–(34) and considering the following asymptotic behavior of the modified Bessel functions when $s\to\infty$,

$$I_\beta(s)=\frac{1}{\sqrt{2\pi s}}e^s\left[1+O\left(\frac{1}{s}\right)\right],\quad K_\beta(s)=\sqrt{\frac{\pi}{2s}}e^{-s}\left[1+O\left(\frac{1}{s}\right)\right], \tag{56}$$

$$I'_\beta(s)=\frac{1}{\sqrt{2\pi s}}e^s\left[1+O\left(\frac{1}{s}\right)\right],\quad K'_\beta(s)=-\sqrt{\frac{\pi}{2s}}e^{-s}\left[1+O\left(\frac{1}{s}\right)\right], \tag{57}$$

and, after complicated algebra, we obtain

$$E_x=0,\quad E_y=E_0, \tag{58}$$

$$\sigma_{xz}(x,y)=-(c_{44}^0a_1-e_{15}^0b_1)c\Psi(1)(1+\alpha y)^{k/2}\times\int_0^\infty J_1(sc)\exp(-sy)\sin(sx)ds+\ldots, \tag{59}$$

$$\sigma_{yz}(x,y)=-(c_{44}^0a_1-e_{15}^0b_1)c\Psi(1)(1+\alpha y)^{k/2}\times\int_0^\infty J_1(sc)\exp(-sy)\cos(sx)ds+\ldots, \tag{60}$$

$$D_x(x,y)=-\frac{e_{15}^0(c_{44}^0a_1-e_{15}^0b_1)}{c_{44}^0}c\Psi(1)(1+\alpha y)^{k/2}\times\int_0^\infty J_1(sc)\exp(-sy)\sin(sx)ds+\ldots, \tag{61}$$

$$D_y(x,y)=-\frac{e_{15}^0(c_{44}^0a_1-e_{15}^0b_1)}{c_{44}^0}c\Psi(1)(1+\alpha y)^{k/2}\times\int_0^\infty J_1(sc)\exp(-sy)\cos(sx)ds+\ldots. \tag{62}$$

Evaluation of the integrals in Eqs. (59)–(62) yields

$$\int_0^\infty J_1(sc)\exp(-sy)\cos(sx)ds=\frac{1}{c}-\frac{1}{\sqrt{r_1r_2}}\frac{r}{c}\cos\left(\theta-\frac{1}{2}\theta_1-\frac{1}{2}\theta_2\right), \tag{63}$$

$$\int_0^\infty J_1(sc)\exp(-sy)\sin(sx)ds=-\frac{1}{\sqrt{r_1r_2}}\frac{r}{c}\sin\left(\theta-\frac{1}{2}\theta_1-\frac{1}{2}\theta_2\right). \tag{64}$$

The polar coordinates r, r_1, r_2, θ, θ_1, and θ_2 are defined in Fig. 1.

Substituting Eqs. (63), (64) into Eqs. (59)–(62) and considering $\theta\to0$, $\theta_2\to0$ and $r_2\to2c$ when $r\to c$, we obtain the following near-fields of stress and electric displacement

$$\sigma_{xz}(r_1,\theta_1)=-\frac{K_{\mathrm{III}}}{\sqrt{2\pi r_1}}\sin\left(\frac{\theta_1}{2}\right)+O(r_1^0), \tag{65}$$

$$\sigma_{yz}(r_1,\theta_1)=\frac{K_{\mathrm{III}}}{\sqrt{2\pi r_1}}\cos\left(\frac{\theta_1}{2}\right)+O(r_1^0), \tag{66}$$

$$D_x(r_1,\theta_1)=-\frac{e_{15}^0}{c_{44}^0}\frac{K_{\mathrm{III}}}{\sqrt{2\pi r_1}}\sin\left(\frac{\theta_1}{2}\right)+O(r_1^0), \tag{67}$$

$$D_y(r_1,\theta_1)=\frac{e_{15}^0}{c_{44}^0}\frac{K_{\mathrm{III}}}{\sqrt{2\pi r_1}}\cos\left(\frac{\theta_1}{2}\right)+O(r_1^0). \tag{68}$$

The stress intensity factor K_{III} in Eqs. (62) and (63) is

$$K_{\mathrm{III}}=(c_{44}^0a_1-e_{15}^0b_1)\sqrt{\pi c}\Psi(1), \tag{69}$$

where $\Psi(1)$ is the value of $\Psi(\xi)$ evaluated at the crack tip corresponding to $\xi=1$.

As for the electric fields, it has been observed from Eq. (58) that no singularities exist in them. This conclusion is the same as that for homogeneous piezoelectric materials (see, e.g., [22]).

If we express all the near-fields as in the traditional linear fracture mechanics, we have

$$E_x(r_1,\theta_1)=-\frac{K_{\mathrm{III}}^E}{\sqrt{2\pi r_1}}\sin\left(\frac{\theta_1}{2}\right),\quad E_y(r_1,\theta_1)=\frac{K_{\mathrm{III}}^E}{\sqrt{2\pi r_1}}\cos\left(\frac{\theta_1}{2}\right), \tag{70}$$

$$\sigma_{xz}(r_1,\theta_1)=-\frac{K_{\mathrm{III}}^\sigma}{\sqrt{2\pi r_1}}\sin\left(\frac{\theta_1}{2}\right),\quad \sigma_{yz}(r_1,\theta_1)=\frac{K_{\mathrm{III}}^\sigma}{\sqrt{2\pi r_1}}\cos\left(\frac{\theta_1}{2}\right), \tag{71}$$

$$D_x(r_1,\theta_1)=-\frac{K_{\mathrm{III}}^D}{\sqrt{2\pi r_1}}\sin\left(\frac{\theta_1}{2}\right),\quad D_y(r_1,\theta_1)=\frac{K_{\mathrm{III}}^D}{\sqrt{2\pi r_1}}\cos\left(\frac{\theta_1}{2}\right), \tag{72}$$

where K_{III}^E is the electric field intensity factor, K_{III}^σ is the stress intensity factor, and K_{III}^D is the electric displacement intensity factor. These field intensity factors are then given by

$$K_{\mathrm{III}}^E=0, \tag{73}$$

$$K_{\mathrm{III}}^\sigma=K_{\mathrm{III}}=(c_{44}^0a_1-e_{15}^0b_1)\sqrt{\pi c}\Psi(1), \tag{74}$$

$$K_{\mathrm{III}}^D=\frac{e_{15}^0}{c_{44}^0}K_{\mathrm{III}}=\frac{e_{15}^0}{c_{44}^0}(c_{44}^0a_1-e_{15}^0b_1)\sqrt{\pi c}\Psi(1). \tag{75}$$

6 Results and Discussion

The expressions in Eqs. (70)–(72) for the near-tip stress and electric fields show that the stresses and the electric displacements at the crack tip in functionally graded piezoelectric materials still possess the inverse square root singularity in terms of the local coordinate at the crack tip and the angular distribution functions are also the same as the cases of homogeneous piezoelectric solids with cracks. This is a very important conclusion. Because the complexity of electromechanical coupling terms in piezoelectric materials plus the nonlinear terms brought by the graded properties to obtain the analytical solution of the crack problem of functionally graded piezoelectric materials is rather difficult, many practical engineering fracture problems related to functionally graded piezoelectric materials should be solved with the help of some numerical methods. The conclusion deduced from this paper provides a theoretical base for these numerical methods in their simulation of the crack-tip stress and electric fields in functionally graded piezoelectric materials.

It is observed from Eqs. (74) and (75) that the forms of the stress intensity factor and the electric displacement intensity factor in the FGPM are similar to those in the homogeneous piezoelectric material. Their values, however, are different because a_1 and b_1 depend on the properties at the outer surfaces $y=\pm h$. Since the electric displacement intensity factor K_{III}^D is proportional to the stress intensity factor K_{III}^σ, we shall illustrate the variation of K_{III}^σ

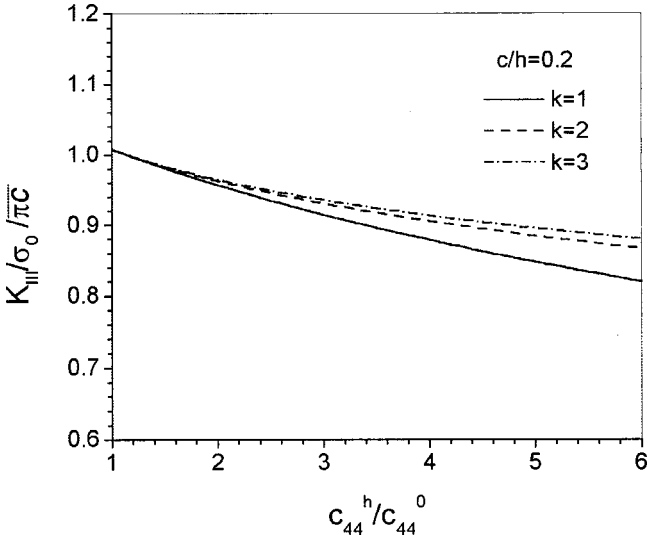


Fig. 2 The variations of the normalized stress intensity factor with the ratio of c_{44}^h/c_{44}^0

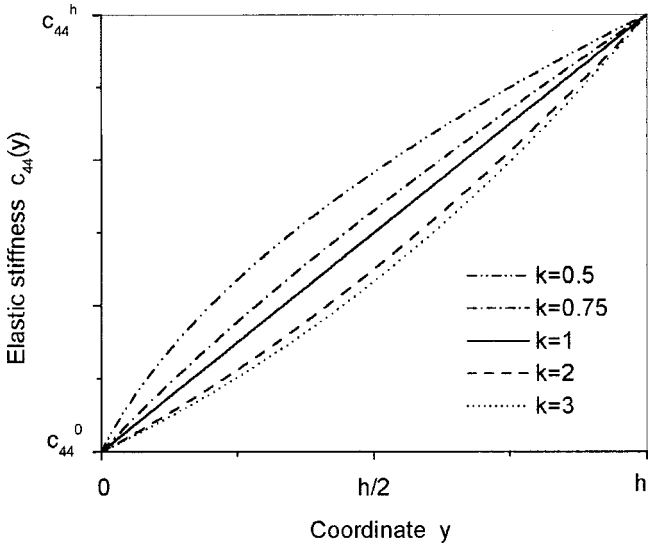


Fig. 4 The variations of the normalized stress intensity factor with the ratio of c/h

only in what follows. The variation of K_{III}^D can be obtained straightly through multiplying the K_{III}^σ by the proportional constant e_{15}^0/c_{44}^0.

Figure 2 displays the variations of the normalized stress intensity factor $K_{\mathrm{III}}/\sigma_0\sqrt{\pi c}$ (where $\sigma_0=c_{44}^0a_1-e_{15}^0b_1$) with the ratio c_{44}^h/c_{44}^0 at $c/h=0.2$. It is evident that the normalized stress intensity factor decreases with the increase of c_{44}^h/c_{44}^0. This means that the stress and electric displacement intensity factors can be reduced by increasing the material property gradient of functionally graded piezoelectric materials. Figure 2 also shows that the influence of the gradient exponent k tends to increase with increasing material gradient c_{44}^h/c_{44}^0. In addition, when the ratio of c_{44}^h/c_{44}^0 approaches to 1, which is the case of homogeneous piezoelectric materials, our results approach the result of corresponding problem in a homogeneous piezoelectric material ([22]). This implies the correctness and accuracy of our results.

Figure 3 illustrates the effect of the gradient exponent k on the variations of the normalized stress intensity factor $K_{\mathrm{III}}/\sigma_0\sqrt{\pi c}$. It can be seen that the greater the ratio of c_{44}^h/c_{44}^0, the stronger the influence of the gradient exponent k. However, the influence is dominant only when k is in the range between 0 and 3. When the value of k is beyond 3, its influence becomes weak. It is also observed that for a definite c_{44}^h/c_{44}^0 and a definite c/h the normalized stress intensity factor decreases when the value of k reduces. If we define the gradient index α as the local gradient of the FGPM, this means that the stress and electric displacement intensity factors can also be reduced by increasing the local material property gradient of functionally graded piezoelectric materials, because we know from Eq. (13) that the reduction of the value k results in the increase of the local gradient α for a definite c_{44}^h/c_{44}^0 and a definite c/h. For the sake of explanation, the influence of k on the gradient of the material property $c_{44}(y)$ is shown in Fig. 4.

The variation of the normalized stress intensity factor $K_{\mathrm{III}}/\sigma_0\sqrt{\pi c}$ with the ratio of crack length to strip thickness is shown in Fig. 5. Different from the case for homogeneous piezoelectric materials, in which the normalized stress intensity factor increases monotonically with increasing c/h, the normalized stress intensity factors initially decrease and then gradually in-

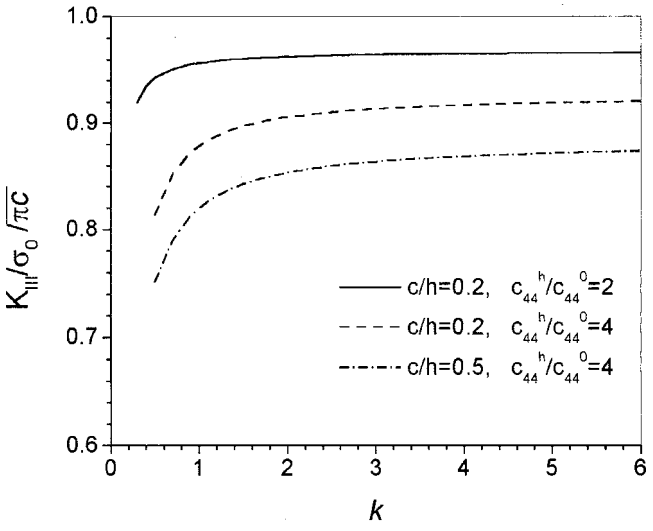


Fig. 3 The effect of the gradient exponent k on the normalized stress intensity factor

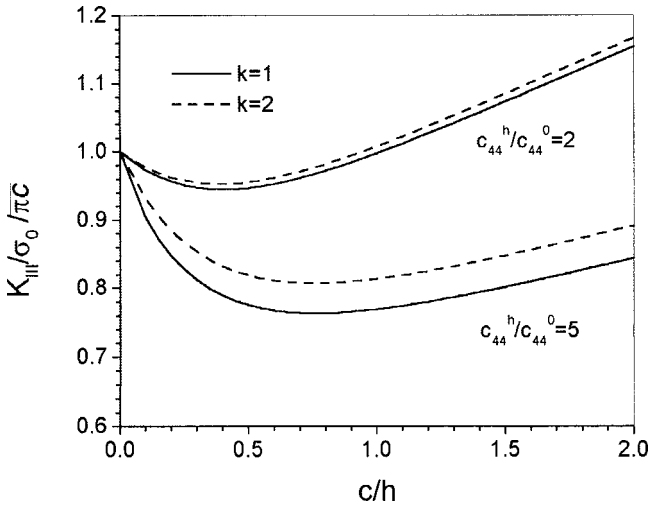


Fig. 5 The variation of the gradient with the exponent k

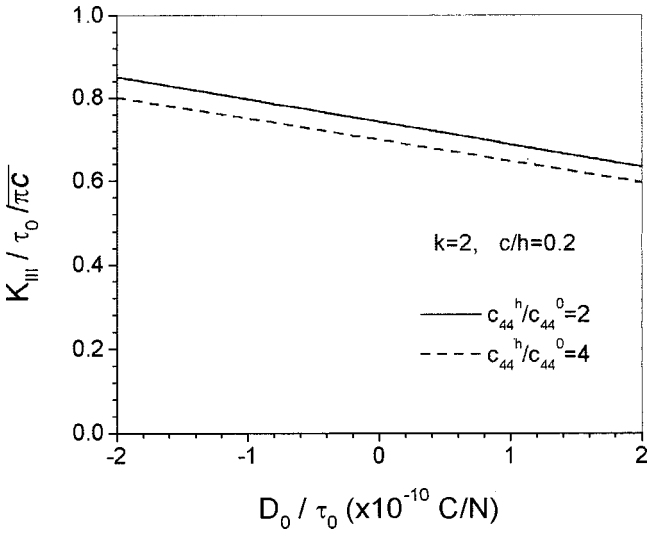


Fig. 6 The effect of the electric displacement on the normalized stress intensity factor

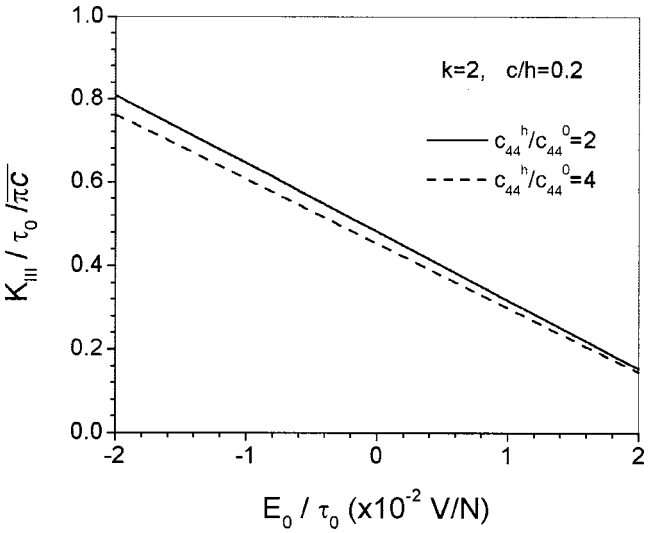


Fig. 7 The effect of the electric intensification on the normalized stress intensity factor

crease with increasing c/h. The greater the gradient of the material property is, the steeper the initial decrease of the normalized stress intensity factor. This phenomenon at first glance may seem to be unreasonable, but it can be explained. Suppose the crack length was definite, and then the strip thickness h would decrease with the increase of c/h. On the one hand, from Eq. (13) we know that the decrease of h implies the increase of the gradient α. Thus the material property gradient would increase with the increase of c/h. In the above two paragraphs, we have deduced the conclusion that the normalized stress intensity factor decreases with the increase of the material property gradient. So the phenomenon of the normalized stress intensity factors initially decreasing with the increase of c/h is thus explained. On the other hand, the upper and lower free boundaries of the FGPM strip usually have the effect of increasing the stress intensity factor. With the decrease of h, the effect of the free boundaries becomes stronger. When c/h increases to a certain point, i.e., when the value of h decreases to a certain value, the effect of the free boundaries starts to offset the influence of the gradient, and then the stress intensity factor gradually increases.

In order to uncover the effects of the electric intensification and the electric displacement on the stress intensity factor, we redefine the normalized stress intensity factor as $K_{\mathrm{III}}^{\sigma}/\tau_0\sqrt{\pi c}$ and derive the explicit expressions for the two cases of free boundary conditions as follows:

$$\text{Case 1:}\ \frac{K_{\mathrm{III}}^{\sigma}}{\tau_0\sqrt{\pi c}} = \frac{1}{\gamma}\left(1 + \frac{(e_{15}^0)^2}{c_{44}^0 \in_{11}^0} - \frac{e_{15}^0}{\in_{11}^0}\frac{D_0}{\tau_0}\right)\Psi(1), \quad (76)$$

$$\text{Case 2:}\ \frac{K_{\mathrm{III}}^{\sigma}}{\tau_0\sqrt{\pi c}} = \left(\frac{1}{\gamma} - e_{15}^0\frac{E_0}{\tau_0}\right)\Psi(1), \quad (77)$$

where $\gamma = c_{44}^h/c_{44}^0$. We take $c_{44}^0 = 3.53\times10^{10}$ N/m^2, $e_{15}^0 = 17.0$ C/m^2, and $\in_{11}^0 = 151\times10^{-10}$ C/Vm for PZT-5H. The effects of the applied electric displacement D_0 and the applied electric field E_0 are displayed in Fig. 6 and Fig. 7, respectively. It reveals that the positive external electric displacement or positive external electric field helps reduce the stress intensity factor and the negative external electric displacement or negative external electric field increases the stress intensity factor. The effect of external electric displacement or external electric field on the stress intensity factor is linear.

It should be pointed out that the finite permeability in the crack cavity has effects on the electric field intensities and electric displacements. From Eqs. (73)–(75), however, it can be seen that the finite permeability in the crack is not related to the singular parts of these variables. As we are mainly concerned with the fractural behavior (the singular parts) of FGPMs, the effect of the finite permeability is not an issue here.

7 Conclusions

The problem of a finite crack in a strip of a functionally graded piezoelectric material is studied. It is assumed that the variations of the elastic stiffness, piezoelectric constant, and dielectric permittivity be continuously varying along the thickness of the strip and the strip is under antiplane mechanical loading and in-plane electric loading. By using the Fourier transform, the problem is first reduced into two pairs of dual integral equations and then into Fredholm integral equations of the second kind. The near-tip singular stress and electric fields are obtained from the asymptotic expansion of the stresses and electric fields around the crack tip. It is found that the forms of the singular stress and electric fields at the tip of a crack in the FGPM can be cast in the same forms as those in a homogeneous piezoelectric material. This result is very important in that one can use conventional numerical methods of fracture mechanics developed for homogeneous piezoelectric solids to study the crack problems in such a class of functionally graded piezoelectric materials.

The computed results show that the gradient of the material property has a considerable effect on the fracture behavior of an FGPM. It is found that increasing the gradient of the material properties is helpful to the reduction of the stress and electric displacement intensity factors.

Acknowledgment

This work was supported by the China Scholarship Council and the National Science Foundation of the USA under CMS-9625304 and CMS-0114801.

References

[1] Cady, W. G., 1946, *Piezoelectricity*, McGraw-Hill, New York.
[2] Ikeda, T., 1996, *Fundamentals of Piezoelectricity*, Oxford University Press. Oxford, UK.
[3] Jaffe, B., Cook, Jr., W. R., and Jaffe, H., 1971, *Piezoelectric Ceramics*, Academic Press, London.
[4] Ono, T., 1990, "Optical Beam Deflector Using a Piezoelectric Bimorph Actuator," Sens. Actuators, **A22**, pp. 726–728
[5] Smits, J. G., Dalke, S. I., and Cookey, T. K., 1991, "Constituent Equations of Piezoelectric Bimorphs," Sens. Actuators, **A28**, pp. 41–61.

[6] Zhu, X., Wang, Q., and Meng, Z., 1995, ''A Functionally Gradient Piezoelectric Actuator Prepared by Power Metallurgical Process in PNN-PZ-PT System,'' J. Mater. Sci. Lett., **14**, pp. 516–518.
[7] Wu, C. M., Kahn, M., and Moy, W., 1996, ''Piezoelectric Ceramics With Functional Gradients: A New Application in Material Design,'' J. Am. Ceram. Soc., **79**(3), pp. 809–812.
[8] Yamada, K., Sakamura, J., and Nakamura, K., 1998, ''Broadband Ultrasound Transducers Using Effectively Graded Piezoelectric Materials,'' *Proceedings of the IEEE Ultrasonic Symposium*, **2**, IEEE, Piscataway, NJ, pp. 1085–1089.
[9] Shelley, W. F., Wan, S., and Bowman, K. J., 1999, ''Functionally Graded Piezoelectric Ceramics,'' Mater. Sci. Forum, **308–311**, pp. 515–520.
[10] Hudnut, S., Almajid, A., and Taya, M., 2000, ''Functionally Gradient Piezoelectric Bimorph Type Actuator,'' Proc. SPIE, C. S. Lynch, ed., **3992**, pp. 376–386.
[11] Zhu, X., Xu, J., and Meng, Z., et al., 2000, ''Microdisplacement characteristics and microstructures of functionally gradient piezoelectric ceramic actuator,'' Mater. Des. **21**, pp. 561–566.
[12] Parton, V. Z., 1976, ''Fracture Behavior of Piezoelectric Materials,'' Ph.D. thesis, Purdue University.
[13] Deeg, W. F., 1980, ''The Analysis of Dislocation, Crack and Inclusion Problems in Piezoelectric Solids,'' Ph.D. thesis, Stanford University.
[14] Sosa, H. A., and Pak, Y. E., 1990, ''Three-Dimesional Eigenfunction Analysis of a Crack in a Piezoelectric Material,'' Int. J. Solids Struct., **26**, pp. 1–15.
[15] Pak, Y. E., 1992, ''Linear Electroelastic Fracture Mechanics of Piezoelectric Materials,'' Int. J. Fract., **54**, pp. 79–100.
[16] Suo, Z., Kuo, C. M., Barnett, D. M., and Willis, J. R., 1992, ''Fracture Mechanics for Piezoelectric Ceramics,'' J. Mech. Phys. Solids, **40**, pp. 739–765.
[17] Wang, B., 1992, ''Three-Dimesional Analysis of a Flat Elliptical Crack in a Piezoelectric Medium,'' Int. J. Eng. Sci., **30**, pp. 781–791.
[18] Dunn, M., 1994, ''The Effects of Crack Face Boundary Conditions on the Fracture Mechanics of Piezoelectric Solids,'' Eng. Fract. Mech., **48**, pp. 25–39.
[19] Park, S. B., 1994, ''Fracture Behavior of Piezoelectric Materials,'' Ph.D. thesis, Purdue University.
[20] Park, S. B., and Sun, C. T., 1995, ''Effect of Electric Field on Fracture of Piezoelectric Ceramic,'' Int. J. Fatigue, **70**, pp. 203–216.
[21] Zhang, T., and Tong, P., 1996, ''Fracture Mechanics for a Mode III Crack in a Piezoelectric Material,'' Int. J. Solids Struct., **33**(3), pp. 343–359.
[22] Shindo, Y., Tanaka, K., and Narita, F., 1997, ''Singular Stress and Electric Fields of a Piezoelectric Ceramic Strip With a Finite Crack Under Longitudinal Shear,'' Acta Mech., **120**, pp. 31–45.
[23] Lee, P. C. Y., and Yu, J. D., 1998, ''Governing Equations for a Piezoelectric Plate With Graded Properties Across the Thickness,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control, **45**(1), pp. 236–250.
[24] Yamada, K., Sakamura, J., and Nakamura, K., 2000, ''Equivalent Network Representation for Thickness Vibration Modes in Piezoelectric Plates With an Exponentially Graded Parameter,'' Jpn. J. Appl. Phys., Part 2, **39**(1), A/B, pp. L34–L37.
[25] Copson, E. T., 1961, ''On Certain Dual Integral Equations,'' Proc. Glasgow Math. Assoc., **5**, pp. 21–24.
[26] Sih, G. C., and Embley, G. T., 1972, ''Sudden Twisting of a Penny-Shaped Crack,'' ASME J. Appl. Mech., **39**, pp. 395–400.

A Local Theory of Elastoplastic Deformation of Two-Phase Metal Matrix Random Structure Composites

V. A. Buryachenko[1]
Visiting Professor

F. G. Rammerstorfer
Professor

A. F. Plankensteiner
Research Assistant

Christian Doppler Laboratory for
Micromechanics of Materials,
Institute of Light Weight Structures and
Aerospace Engineering,
TU Vienna,
A-1040 Vienna, Austria

A two-phase material is considered, which consists of a homogeneous elastoplastic matrix containing a homogeneous statistically uniform random set of ellipsoidal inclusions with the same form, orientation, and mechanical properties. The multiparticle effective field method (used in this paper) in the original form assumes constant plastic strains in the matrix. This assumption is replaced by the following micromechanical model: Each inclusion consists of an elastic core and a thin coating. The mechanical properties of both the matrix and the coating are the same but with different plastic strains. Homogeneous plastic strains are assumed inside the matrix and in each of separate subdomains of the coating. A general theory of plasticity is developed for arbitrary loading based on incremental elastoplastic analysis. The consideration of inhomogeneity of plastic strains in the coating enables to obtain some principally new effects of elastoplastic deformation.
[DOI: 10.1115/1.1479697]

1 Introduction

Significant contributions to the micromechanics of metal matrix composites (MMCs) have been given by Dvorak and co-workers [1–3] by the use of particular mean field methods. In this framework the distribution of the plastic part of the strains as well as the stresses within the components are assumed to be homogeneous. For a deterministic, periodic arrangement of inclusions this approach was extended by Dvorak et al. [3] to mean fields varying of over subdomains. For random packing in the framework of the original mean field assumption the average stresses inside the components can be estimated by different methods. Most popular and widely used methods are variants of effective medium ([4–6]) and mean field methods ([2,7,8]). Since the widely used mean-field methods are capable of estimating only the average stresses in the components, its use for the averaging of functions describing nonlinear effects, e.g., strength ([9,10]) yielding ([11]), hardening ([12]), and creep ([13]) may be limited, because in reality significant inhomogeneities of the stress fields in the components (especially in the matrix) can be obtained.

When plastic deformations occur, the homogeneity of the mechanical properties of the components is lost and the local properties of the phases become position-dependent. Exact solutions by analytical or numerical methods can be obtained only for model composites having deterministic phase arrangements, such as composites with regular micro structure (see, e.g., [2,14–17]). Alternatively, nonlinear multiphase materials may be described by self-consistent schemes. For example, Duva [18] as well as Lee and Mear [19] studied a special case of infinitely small concentrations of heterogeneities in nonlinear matrix composites and used differential self-consistent schemes for obtaining approximate constitutive relations applicable for a large range of inclusion concentrations.

The three-phase spherical model (see, e.g., [20]) and its generalizations were used by Zhu and Weng [21] to predict the creep behavior of metal-matrix composites and by Olsson et al. [22] for elastoplastic analysis of thermal cycling of particulate composites. Taggart and Bassani [23] as well as Herrmann and Mihovsky [24] showed a principle role of the concentration of plastic strains in the vicinity of the inclusions in the overall deformation properties of composite materials.

The classical mean field method uses the first-order statistical moment of stresses in each individual component. A modified method, based on the average second moment of stresses in the components, was proposed by Buryachenko and Lipanov [25], Qiu and Weng [26], Ding and Weng [27] (see [28,29]) for the estimations of different nonlinear effective properties such as elasticity, plasticity, and viscosity. Other theories basing on a similar principle are connected with generalizations of the Hashin and Shtrikman [30] variational principles to nonlinear materials, as proposed by Talbot and Willis [31] (see [32]).

The present contribution is aimed at investigating the elastoplastic behavior of two-phase materials in the framework of flow theory and small elastoplastic strains. Ellipsoidal coated inclusions with the same shape, orientation, and mechanical properties are statistically uniformly distributed inside the space. Each inclusion consists of an elastic core and a thin coating. The mechanical properties of the coating are the same as that of the matrix. Homogeneity of the plastic strains is assumed inside the matrix and in individual subdomains of the coating, which are considered as individual components. By this means the proposed method can be considered as a logical extension of the transformation field analysis by Dvorak [1] to random arrangement, when the phases with inhomogeneous stress states are subdivided into a finite number of subdomains with homogeneous stress states. The changes of the average stresses within the components are estimated with the help of the multiparticle effective field method (MEFM) proposed by one of the authors for linear problems (see [28,29]). For a single inclusion the thin-layer hypothesis and the assumption of a homogeneous stress state in the core are used. For a dilute concentration of the inclusions the proposed model is tested by the

[1]Present address: Air Force Research Laboratory, AFRL MLBC, Oh 45433-7750. e-mail: buryach@aol.com.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, December 17, 1996; final revision, September 15, 2000. Associate Editor: J. L. Bassani. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

use of finite element analysis. The employment of the proposed theory for predicting the elastoplastic deformation behavior of a model material with a finite concentration of inclusions is shown.

2 Mechanical Properties and Geometrical Structure of the Components

2.1 Leading Equations. Let the total local small strain $\boldsymbol{\varepsilon}$ be decomposed into an elastic ε^e and plastic $\boldsymbol{\varepsilon}^p$ contribution $\boldsymbol{\varepsilon}=\varepsilon^e+\varepsilon^p$. The local constitutive equation, which connects the local stress tensor $\boldsymbol{\sigma}(x)$ and the local elastic strain tensor $\boldsymbol{\varepsilon}^e$ is given in the form $\boldsymbol{\sigma}(x)=\mathbf{L}(\mathbf{x})\boldsymbol{\varepsilon}^e(\mathbf{x})$, where $\mathbf{L}(\mathbf{x})$ is the fourth-order anisotropic tensor of elasticity; in particular for isotropic material $\mathbf{L}=(3k(\mathbf{x}),2\mu(\mathbf{x}))\equiv 3k(\mathbf{x})\mathbf{N}_1+2\mu(\mathbf{x})\mathbf{N}_2$, where $\mathbf{N}_1=\boldsymbol{\delta}\otimes\boldsymbol{\delta}/3$; $\mathbf{N}_2=\mathbf{I}-\mathbf{N}_1$, k and μ are the bulk and shear moduli, $\boldsymbol{\delta}$ and $\mathbf{I}$ are the unit second-order and fourth-order tensors. At the macrodomain w uniform boundary traction conditions generating the homogeneous macroscopic stress $\boldsymbol{\sigma}^0$ are given, the phases are assumed to be perfectly bonded.

For the description of the behavior of the matrix and the elastic-plastic coating the so-called J_2-flow theory with combined isotropic-kinematic hardening is used. The von Mises form of the yield surface is given by

$$f\equiv\tau-F(\gamma)=0,\quad F(0)=\tau_0 \tag{1}$$

in terms of the effective stress and plastic strain increments, as defined by $\tau=(1.5s^a_{ij}s^a_{ij})^{1/2}$, $d\gamma=(2d\varepsilon^p_{ij}d\varepsilon^p_{ij}/3)^{1/2}$. Here F is a nonlinear function describing the hardening effect, for example $F(\gamma)=\tau_0+h\gamma^n$, where h and n are the strength coefficient and the work-hardening exponent, respectively. $\mathbf{s}^a\equiv\mathbf{N}_2(\boldsymbol{\sigma}-\mathbf{a}^p)$ is the deviator of the active stresses; $\mathbf{a}^p$ is a symmetric second-order tensor corresponding to the ''back-stress'' defining the location of the center of the yield surface in the deviatoric stress space. For the evaluation of the back stress $\mathbf{a}^p$ Ziegler's hardening rule is used:

$$d\mathbf{a}^p=d\gamma A\mathbf{s}^a,\quad A\equiv A(\gamma), \tag{2}$$

where $\mathbf{a}^p=0$ at $\gamma=0$. The material behaves elastically if $f<0$, or if $f=0$ and $(\partial f/\partial\boldsymbol{\sigma}):d\boldsymbol{\sigma}\leqslant 0$; elastic-plastic deformations take place under active loading, when $f=0$ and $(\partial f/\partial\boldsymbol{\sigma}):d\boldsymbol{\sigma}>0$. By using the associated flow rule, the yield function f is taken as plastic potential function from which the incremental plastic strains can be derived as

$$d\boldsymbol{\varepsilon}^p=d\lambda\frac{\partial f}{\partial\boldsymbol{\sigma}},\quad \text{for}\quad \frac{\partial f}{\partial\boldsymbol{\sigma}}:d\boldsymbol{\sigma}>0, \tag{3}$$

where $d\lambda$ is a proportionality factor derived from the condition of continuous variation of the yield surface $df=0$.

2.2 Geometrical Structure of Composites. The paper discusses a mesodomain w with a characteristic function W containing a set $X=(V_i,\mathbf{x}_i,\omega_i)$, $(i=1,2,\dots)$ of coated ellipsoidal inclusions v_i with characteristic functions V_i, centers x_i (that forms a Poisson set), semi-axes a^c_j $(j=1,2,3)$ and aggregate of Euler angles ω. It is assumed that the inclusions v_i have identical mechanical and geometrical properties. Each inclusion consists of an elastic ellipsoidal core $v^s_i\subset v_i$ with semi-axes a^s_j $(j=1,2,3)$, characteristic function $V^s_i(\mathbf{x})$, and a thin coating $v^c_i\equiv v_i\backslash v^s_i$ bounded by a homothetic ellipsoidal surface ∂v^c_i with semi-axes $a^c_j=a^s_j(1+\xi)$ $(j=1,2,3, 0<\xi<1)$ and a characteristic function $V^c_i(\mathbf{x})\equiv V_i(\mathbf{x})-V^s_i$; here ξ is the relative thickness of the coating. The mechanical properties $\mathbf{g}(\mathbf{x})=\mathbf{g}^{(0)}$ $(\mathbf{g}=\mathbf{L},\tau^0,h,n)$ are the same for both the matrix $v^{(0)}\equiv w\backslash\cup v_i$ and the coating v^c_i. The plastic strains are constant in the matrix and are an inhomogeneous function along the surface ∂v^s_i. The volume v^c_i is subdivided along ∂v^s_i into several local volumes v^c_{ij} $(\text{mes}\, v^c_i\ll\text{mes}\, v^c_i, j=1,\dots n^c)$, such that the plastic strains $\boldsymbol{\varepsilon}^p(\mathbf{x})$, $(\mathbf{x}\in v^c_{ij})$ are constant inside the each individual subdomain v^c_{ij}. In the core of the inclusions we assume $\mathbf{g}(\mathbf{x})=\mathbf{g}^{(0)}+\mathbf{g}_1(\mathbf{x})=\mathbf{g}^{(0)}+\mathbf{g}_1$; $\mathbf{g}_1(\mathbf{x})=\text{const}$ for $\mathbf{x}\in v^{(1)}$. The upper index of the material properties tensor, written in parentheses, shows the number of the respective component: $^{(0)}$ correspond to the matrix, $^{(1)}$ to the core of the inclusions and $^{(k)}$ $(k=2,n^c+1)$ to the coating. It is assumed that the representative domain w contains a statistically large number of inclusions v_i; all random quantities under discussion are described by statistically homogeneous ergodic random fields and, hence, the ensemble averaging could be replaced by volume averaging $\langle(.)\rangle$ and $\langle(.)\rangle^{(k)}$ $(k=0,\dots,n^c+1)$ for the overall average and for phase averages, respectively. The bar appearing above the region represents its measure, $\overline{v}\equiv\text{mes}\, v$. $V^{(k)}$ is the characteristic function of $v^{(k)}$. The average over an individual inclusion $v_i\in v^{(1)}$ $(i=1,2,\dots)$: is defined by $\langle(.)\rangle_i=\langle(.)\rangle^{(1)}$ if $v_i\in v^{(1)}$. $c^{(k)}$ is the concentration, i.e., volume fraction, of component $v^{(k)}$: $c^{(k)}=\langle V^{(k)}\rangle$; $c^{(1)}=\overline{v}^s_m n^{(1)}$, $c^{(0)}=1-\Sigma c^{(l)}$ $(k=0,\dots,n^c+1; l=1,\dots,n^c+1, m=1,2,\dots)$. In the following the notation $\langle(.)(\mathbf{x})|\mathbf{x}_1;\dots;\mathbf{x}_m\rangle$ is used for the conditional average taken for the ensemble of a statistically homogeneous ergodic field $X=(v_i)$, on the condition that there are inclusions at the points $\mathbf{x}_1,\dots,\mathbf{x}_m$ and $\mathbf{x}_1\neq\dots\neq\mathbf{x}_m$. The notation $\langle(.)\times(\mathbf{y})|;\mathbf{x}_1;\dots;\mathbf{x}_m\rangle$ means the case $\mathbf{y}\notin v_1,\dots,v_m$.

3 Average Stresses Inside the Components and Overall Elastic Moduli

A general infinite system of integral equations can be derived (see for references and details Willis [33] and Buryachenko [29])

$$\langle\boldsymbol{\sigma}(\mathbf{x})|\mathbf{x}_1;\dots;\mathbf{x}_n\rangle-\sum_{i=1}^{n}\int\boldsymbol{\Gamma}(\mathbf{x}-\mathbf{y})\langle V_i(\mathbf{y})\boldsymbol{\eta}(\mathbf{y})|\mathbf{x}_1;\dots;\mathbf{x}_n\rangle\mathbf{dy}$$

$$=\boldsymbol{\sigma}^0+\int\boldsymbol{\Gamma}(\mathbf{x}-\mathbf{y})\{\langle\boldsymbol{\eta}(\mathbf{y})|;\mathbf{x}_1;\dots;\mathbf{x}_n\rangle-\langle\boldsymbol{\eta}\rangle\}d\mathbf{y}, \tag{4}$$

where $\mathbf{x}\in v_1,\dots,v_n$ $(n=1,2,\dots)$ in the nth line of the system, and $\boldsymbol{\eta}(\mathbf{y})=\mathbf{M}_1(\mathbf{y})\boldsymbol{\sigma})(\mathbf{y})+\boldsymbol{\varepsilon}^p_1(\mathbf{y})$ is the strain polarization tensor. $\mathbf{M}_1(\mathbf{y})$ is the jump of the compliance $\mathbf{M}(\mathbf{y})\equiv\mathbf{L}(\mathbf{y})^{-1}$ inside the core of the inclusion v^s with respect to the matrix $v^{(0)}$. The integral operator kernel $\boldsymbol{\Gamma}(\mathbf{x}-\mathbf{y})\equiv-\mathbf{L}^{(0)}[\mathbf{I}\delta(\mathbf{x}-\mathbf{y})+\nabla\nabla\mathbf{G}(\mathbf{x}-\mathbf{y})\mathbf{L}^{(0)}]$, is defined by the Green tensor $\mathbf{G}$ of the Lame' equation of an infinite homogeneous medium with elastic modulus tensor $\mathbf{L}^{(0)}$: $\nabla\{\mathbf{L}^{(0)}[\nabla\otimes\mathbf{G}(\mathbf{x})+(\nabla\otimes\mathbf{G}(\mathbf{x}))^{\mathrm{T}}]/2\}=-\boldsymbol{\delta}\delta(\mathbf{x})$, $\delta(\mathbf{x})$ is the Dirac delta function.

In order to simplify the exact system (4) we now apply the main hypothesis of many micromechanical methods, called the effective field hypothesis (see for references and details Buryachenko [29])

H1 *Each inclusion v_i has an ellipsoidal form and is located in the field $\overline{\boldsymbol{\sigma}}_i(\mathbf{y})\equiv\overline{\boldsymbol{\sigma}}(\mathbf{x}_i)$, $(\mathbf{y}\in v_i)$ which is homogeneous over the inclusion v_i, and the perturbation introduced by the inclusion v_i at the point $\mathbf{y}\notin v_i$ is defined by the relation*

$$\int\boldsymbol{\Gamma}(\mathbf{y}-\mathbf{x})V_i(\mathbf{x})\boldsymbol{\eta}(\mathbf{x})d\mathbf{x}=\overline{v}_i\mathbf{T}_i(\mathbf{y}-\mathbf{x}_i)\boldsymbol{\eta}_i. \tag{5}$$

Here $\boldsymbol{\eta}_i\equiv\langle\boldsymbol{\eta}(\mathbf{x})V_i(\mathbf{x})\rangle_{(i)}$ is an average over the volume of the inclusion v_i (but not over the ensemble), $\langle(.)\rangle_i\equiv\langle\langle(.)\rangle_{(i)}\rangle$, and $\mathbf{T}_i(\mathbf{y}-\mathbf{x}_i)=(\overline{v}_i)^{-1}\int\boldsymbol{\Gamma}(\mathbf{y}-\mathbf{x})V_i(\mathbf{x})d\mathbf{x}$, $\mathbf{y}\notin v_i$.

In view of the linearity of the problem there exist constant fourth and second-rank tensors $\mathbf{B}^{(j)}$, $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{C}^{(jk)}$, $\mathbf{C}$, $\mathbf{F}$, respectively, such that

$$\boldsymbol{\sigma}(\mathbf{x})=\mathbf{B}^{(j)}\overline{\sigma}_i+\sum_{k=1}^{n^c+1}\mathbf{C}^{(jk)}\boldsymbol{\varepsilon}_1^{p(k)}, \tag{6}$$

$$\langle\boldsymbol{\sigma}\rangle_i=\mathbf{B}\langle\overline{\boldsymbol{\sigma}}\rangle_i+\mathbf{C},\quad \overline{v}_1\langle\boldsymbol{\eta}\rangle_i=\mathbf{R}\langle\overline{\boldsymbol{\sigma}}\rangle_i+\mathbf{F}, \tag{7}$$

($\mathbf{x}\in v^{(j)}$), where $\mathbf{C}^{(j1)}\equiv\mathbf{0}$ and the tensors $\mathbf{B}$, $\mathbf{C}$, $\mathbf{B}^{(j)}$, $\mathbf{C}^{(jk)}$, ($j,k=1,\ldots,n^c+1$) are represented in the Appendix. The tensors $\mathbf{R}$ and $\mathbf{F}$ are found by the use of the Eshelby theorem $\mathbf{R}=\overline{v_i}\mathbf{Q}^{-1}(\mathbf{I}-\mathbf{B})$, $\mathbf{F}=-\overline{v_i}\mathbf{Q}^{-1}\mathbf{C}$; the tensor $\mathbf{Q}$ is associated with the well-known Eshelby tensor $\mathbf{S}$ by $\mathbf{S}=\mathbf{I}-\mathbf{M}^{(0)}\mathbf{Q}$.

Buryachenko and Rammerstorfer [34] proved that in the framework of the effective field hypothesis **H1** only the effective parameters $\mathbf{M}^*$, $\boldsymbol{\varepsilon}^p_*$ governing the overall constitutive relation $\langle\boldsymbol{\varepsilon}\rangle=\mathbf{M}^*\langle\boldsymbol{\sigma}\rangle+\boldsymbol{\varepsilon}^p_*$ as well as statistical average of the local stresses inside the inclusions have the general representation. In particular, from Eq. (6) we get

$$\mathbf{M}^*=\mathbf{M}^{(0)}+\mathbf{YR}n^{(1)},\quad \boldsymbol{\varepsilon}^p_*=\boldsymbol{\varepsilon}^{p(0)}+\mathbf{YF}n^{(1)}, \tag{8}$$

$$\langle\boldsymbol{\sigma}\rangle^{(j)}=\mathbf{B}^{*(j)}\boldsymbol{\sigma}^0+\sum_{k=0}^{n^c+1}\mathbf{C}^{*(jk)}\boldsymbol{\varepsilon}_1^{p(k)}, \tag{9}$$

where

$$\mathbf{C}^{*(j0)},\quad \mathbf{C}^{*(j1)}\equiv 0,\quad (j=0,\ldots,n^c+1)$$

and

$$\mathbf{B}^{*(l)}=\mathbf{B}^{(l)}\mathbf{R}^{-1}\mathbf{YR},\quad (l=1,\ldots,n^c+1),$$

$$\mathbf{B}^{*(0)}=\frac{1}{c^{(0)}}\left[\mathbf{I}-\sum_{l=1}^{n^c+1}c^{(l)}\mathbf{B}^{*(l)}\right],$$

$$\mathbf{C}^{*(jl)}=\mathbf{C}^{(jl)}+\mathbf{B}^{(j)}\mathbf{R}^{-1}(\mathbf{Y}-\mathbf{I})\left[\overline{v_i^s}\mathbf{M}_1^{(1)}\mathbf{C}^{(1l)}+\frac{c^{(l)}\mathbf{I}}{n^{(1)}}\right],$$

$$\mathbf{C}^{*(0l)}=-\frac{1}{c^{(0)}}\sum_{k=1}^{n^c+1}c^{(k)}\mathbf{C}^{*(kl)}. \tag{10}$$

The tensor $\mathbf{Y}$ is determined by the purely elastic action (with $\mathbf{F}\equiv\mathbf{0}$) of the surrounding inclusions on the separated one. For a dilute concentration of the inclusions, i.e., $c^{(1)}\to 0$, we have $\mathbf{Y}\to\mathbf{I}$. The actual form of the tensor $\mathbf{Y}$, used in the analysis as an approximation, depends on additional assumptions for closing of the infinite system (4). For the purely elastic case (with $\mathbf{F}\equiv 0$) such relations are represented by Buryachenko and Rammerstorfer [34].

4 Onset of Yielding

Most of the mean field methods (references can be found, e.g., in the survey papers by Dvorak, [2] and Buryachenko [28]) use as initial yield condition

$$\tau^{(0)}=\tau_0^{(0)},\quad \tau^{(0)}\equiv\sqrt{1.5\langle\boldsymbol{\sigma}\rangle^{(0)}:\mathbf{N}_2:\langle\boldsymbol{\sigma}\rangle^{(0)}}. \tag{11}$$

The hypothesis (11) is, in the above-mentioned sense, inconsistent. This becomes obvious if we consider a macroscopically isotropic porous material with a matrix which is described by von Mises plasticity. In this case under hydrostatic loading condition $\langle\sigma_{ij}\rangle=\langle\sigma_{nn}\rangle\delta_{ij}/3$, irrespective of the microstructure of the pores and the method of calculation of $\langle\boldsymbol{\sigma}\rangle^{(0)}$ (for example, by Eq. (9) or any other formula), we obtain that $\langle\mathbf{s}\rangle^{(0)}\equiv\mathbf{0}$, where $\mathbf{s}\equiv\mathbf{N}_2\boldsymbol{\sigma}$. This would lead to the condition that the plastically incompressible matrix would never yield under hydrostatic loading, what is in contradiction with experimental observations. The above-mentioned inconsistency can be avoided if the same general scheme of mean field methods with subdivisions of the components in separate subdomains are considered. So in our case for varied loading onset of yielding appears in the individual subdomain of the coating v^c_{ij} or in the remaining matrix $v^{(0)}$, respectively, if the corresponding yield criterion

$$\tilde{\tau}^{(j)}=\tau_0^{(0)},\quad \tilde{\tau}^{(j)}\equiv\sqrt{1.5\langle\boldsymbol{\sigma}\rangle^{(j)}:\mathbf{N}_2:\langle\boldsymbol{\sigma}\rangle^{(j)}}, \tag{12}$$

($j=0,1,\ldots,n^c+1$) is used. The composite starts to deform plastically, when $\max_j\tilde{\tau}^{(j)}-\tau_0^{(0)}=0$ ($j=0,\ldots,n^c+1$) is fulfilled. Up to the moment when yielding starts, the problem of estimating $\tilde{\tau}^{(j)}$ is linear, which allows the use of the relations (9) for calculating the yield surface

$$\tilde{f}(\langle\boldsymbol{\sigma}\rangle)=\max_j\sqrt{1.5(\mathbf{B}^{*(j)}\boldsymbol{\sigma}^0)^T:\mathbf{N}_2:(\mathbf{B}^{*(j)}\boldsymbol{\sigma}^{(0)})}-\tau_0^{(0)}=0, \tag{13}$$

($j=0,\ldots,n^c+1$). The initial yield surfaces of the components (i.e., the individual subdomains) (12) may be embedded within each other or they may intersect in the space of macrostresses $\langle\boldsymbol{\sigma}\rangle$.

It should be mentioned that special so-called localized models of plasticity of particulate composites have been developed. So Hervé and Zaoui (1993) proposed a generalization of the elastic three-phase spherical model (see e.g. Hashin, 1986) towards the elastoplastic case in the framework of a secant modulus concept. We will not discuss here in general the known advantages of flow theory over deformation theory of plasticity in the case of nonradial loading, that takes place usually at the local level near the inclusions even if the overall stress path is radial. We just look on the method of the calculation of onset of yielding in the layer model by Hervé and Zaoui [35], who use the criterion (in our notations, see the Appendix)

$$\sqrt{1.5\langle\boldsymbol{\sigma}^c(\mathbf{s})\rangle_{s^i}:\mathbf{N}_2:\langle\boldsymbol{\sigma}^c(s)\rangle_{s^i}}-\tau_0^{(0)}=0, \tag{14}$$

where $\langle\boldsymbol{\sigma}^c(\mathbf{s})\rangle_{s^i}$ is the surface average of the stresses $\boldsymbol{\sigma}^c(\mathbf{s})$ in the coating over the surface of the core s^i. As a consequence of the properties of the interface operator $\overline{\boldsymbol{\Gamma}}(\mathbf{n})$ this average is given by $\langle\boldsymbol{\sigma}^c(\mathbf{s})\rangle_{s^i}\equiv\boldsymbol{\sigma}^{(1)}$ and, therefore, for porous materials one would obtain an infinite overall initial yield stress for pure hydrostatic loading σ_0^0. More recently a model has been proposed by Bornet et al. [36] which assumes a yield criterion by

$$\sqrt{1.5\langle\{\langle\boldsymbol{\sigma}^c\rangle(\mathbf{s}):\mathbf{N}_2:\langle\boldsymbol{\sigma}^c\rangle(\mathbf{s})\}\rangle_{s^i}}-\tau_0^{(0)}=0. \tag{15}$$

Here $\langle\boldsymbol{\sigma}^c\rangle(\mathbf{s})$ is a statistical average of the stresses in the point $\mathbf{s}\in s^i$. Obviously the criterion (12) described in this paper is a generalization of (15) and, the yield surface (12) is inside the elliptical curve (15) in a dimensionless coordinate system $X_\tau=\sigma_0^0/\tau_0^{(0)}$, $Y_\tau=(1.5\mathbf{s}^0:\mathbf{s}^0)^{1/2}/\tau_0^{(0)}$, where $\sigma_0^0\equiv\boldsymbol{\delta}:\boldsymbol{\sigma}^0/3$, $\mathbf{s}^0\equiv\mathbf{N}_2\boldsymbol{\sigma}^0$.

Zhu and Weng [21] used a criterion, equivalent to (12), in that approach at each increment of external loading the increment of homogeneous inelastic strains in the matrix was assumed to be the average of the local inelastic strain increments over the matrix volume. This, however, would lead to zero increments of inelastic strains in the case of hydrostatic loading of macroisotropic media.

Instead of the assumptions of mean field methods (11) Buryachenko and Lipanov [25] (for more details see also [28,29]) proposed the criterion

$$\sqrt{1.5(\boldsymbol{\sigma}:\mathbf{N}_2:\boldsymbol{\sigma})^{(0)}}-\tau_0^{(0)}=0, \tag{16}$$

which is based on the estimation of the second moment of the stresses in the matrix. For the composite with the isotropic incompressible matrix containing identical spherical voids, Buryachenko and Rammerstorfer [37] presented the yield surfaces (11), (13), and (16) in the dimensionless coordinate system X_τ, Y_τ. The yield surface for stress states at the boundary of the voids (13) has a nonelliptical form and lies inside the elliptical yield surface (16); the solution of linear elastic problem (8) and (9) by the two-particle approximation of MEFM was used. The criterions (13) and (16) show that yielding will also take place under purely hydrostatic loading, which is in contrast to the results for onset of yielding obtained by conventional analysis of composites based on the assumption of homogeneity of the microstress fields in the matrix, Eq. (11). Comparisons of the yield surfaces proposed with

other known results are considered by Ponte Castañeda and Suquet [32] and by Buryachenko [28,29] in more detail.

5 Elastoplastic Deformation

Now we shall deal with the elastoplastic state, when under successive external loading elastic-plastic deformations will take place for the components $i_1, \ldots, i_r$

$$\tilde{f}^{(i_q)}=0, \quad \frac{\partial \tilde{f}^{(i_q)}}{\partial \langle \boldsymbol{\sigma} \rangle^{(i_q)}} : d\langle \boldsymbol{\sigma} \rangle^{(i_q)} > 0, \tag{17}$$

where the index i_q, $(q=1,\ldots,r)$ passes through the number $(i_1,\ldots,i_r)$ of plastically deformed components $v^{(i_q)}$, $(0 \leqslant i_q \leqslant n^c+1, i_q \neq 1)$ and where a homogeneous yield criterion inside the volume $v^{(i_q)}$ is assumed:

$$\tilde{f}^{(i_q)} \equiv \tilde{\tau}^{(i_q)} - F^{(i_q)}(\gamma^{(i_q)}) = 0, \tag{18}$$

$$\tilde{\tau}^{(i_q)} = \sqrt{1.5 \langle \mathbf{s}^a \rangle^{(i_q)} : \langle \mathbf{s}^a \rangle^{(i_q)}}. \tag{19}$$

By using the associated flow rule, the yield function $\tilde{f}^{(i_q)}$ is taken as plastic potential function of the matrix from which the incremental plastic strains inside the matrix can be determined as

$$d\boldsymbol{\varepsilon}^{p(i_q)} = d\lambda^{(i_q)} \frac{\partial \tilde{f}^{(i_q)}}{\partial \langle \boldsymbol{\sigma} \rangle^{(i_q)}}. \tag{20}$$

At each incremental step of the external stresses homogeneous plastic strains $\boldsymbol{\varepsilon}^{p(i_q)}$ and increments of hardening parameters are assumed within each subdomain. With Ziegler's rule (2) we obtain

$$d\mathbf{a}^{p(i_q)} = d\gamma^{(i_q)} A \langle \mathbf{s}^a \rangle^{(i_q)}, \quad A = \frac{H^{(i_q)}}{F^{(i_q)}(\gamma^{(i_q)})}, \tag{21}$$

where $H^{(i_q)}$ is the plastic tangent modulus, derived from the uniaxial stress-plastic strain curve of the matrix material,

$$H^{(i_q)} = \frac{\partial \tilde{\tau}^{a(i_q)}}{\partial \gamma^{(i_q)}}, \quad d\gamma^{(i_q)} = \sqrt{2 d\boldsymbol{\varepsilon}^{p(i_q)} : d\boldsymbol{\varepsilon}^{p(i_q)}/3}, \tag{22}$$

$$d\tilde{\tau}^{a(i_q)} \equiv \sqrt{\langle d\mathbf{s}^a \rangle^{(i_q)} : \langle d\mathbf{s}^a \rangle^{(i_q)}}. \tag{23}$$

From the requirement of consistency of the plastic deformation process we come up with

$$\frac{\partial \tilde{f}^{(i_q)}}{\partial \langle \boldsymbol{\sigma} \rangle^{(i_q)}} : d\langle \boldsymbol{\sigma} \rangle^{(i_q)} + \frac{\partial \tilde{f}^{(i_q)}}{\partial \mathbf{a}^{p(i_q)}} : d\mathbf{a}^{p(i_q)} + \frac{\partial \tilde{f}^{(i_q)}}{\partial \gamma^{(i_q)}} d\gamma^{(i_q)} = 0. \tag{24}$$

The partial derivatives in (24) are found under the following assumptions: $\boldsymbol{\varepsilon}^{p(i_q)}$, $\langle \boldsymbol{\sigma} \rangle^{p(i_q)} = \text{const} (q=1,\ldots,r)$. From (9), (20) and (21), (22) the relations for the differential of average stresses and hardening parameters inside the component $v^{(q)}$ can be derived as

$$d\langle \boldsymbol{\sigma} \rangle^{(i_q)} = \mathbf{B}^{*(i_q)} d\boldsymbol{\sigma}^0 - \sum_{k=1}^{r} \mathbf{C}^{*(i_q i_k)} : \left[\lambda^{(i_k)} \frac{\partial \tilde{f}^{(i_k)}}{\partial \langle \boldsymbol{\sigma} \rangle^{(i_k)}} - \lambda^{(0)} \frac{\partial \tilde{f}^{(0)}}{\partial \langle \boldsymbol{\sigma} \rangle^{(0)}} \right], \tag{25}$$

$$d\gamma^{(i_q)} = d\lambda^{(i_q)}, \quad d\mathbf{a}^{p(i_q)} = d\lambda^{(i_q)} A(\gamma^{(i_q)}) \langle \mathbf{s}^a \rangle^{(i_q)}. \tag{26}$$

In Eq. (26) $d\lambda^{(0)} \equiv 0$ and $\partial \tilde{f}^{(0)} / \partial \langle \boldsymbol{\sigma} \rangle^{(0)} \equiv 0$ if the conditions (17) are not met for the component $v^{(0)}$.

Substitution of the Eqs. (25), (26) in (24) leads to the following relation for the proportionality factor:

$$d\lambda^{(i_q)} = \sum_{s=1}^{r} \beta_{qs} b_s, \quad b_s = \frac{\partial \tilde{f}^{(i_s)}}{\langle \boldsymbol{\sigma} \rangle^{(i_s)}} : \mathbf{B}^{*(i_s)} : \boldsymbol{\sigma}^0. \tag{27}$$

The elements of the inverse of the matrix $(\boldsymbol{\beta})$ are given by

$$(\boldsymbol{\beta}^{-1})_{qs} = \frac{\partial \tilde{f}^{(i_q)}}{\partial \langle \boldsymbol{\sigma} \rangle^{(i_q)}} : \left[\mathbf{C}^{*(i_q i_s)} (\delta_{i_s 0} - 1) + \delta_{i_s 0} \sum_{n=1}^{r} \mathbf{C}^{*(i_q i_n)} \right] : \frac{\partial \tilde{f}^{(i_s)}}{\partial \langle \boldsymbol{\sigma} \rangle^{(i_s)}} + 2 \frac{\partial F^{(i_q)}}{\partial \gamma^{(i_q)}} \delta_{qs}. \tag{28}$$

Plastic deformations in the component $v^{(i_q)}$ take place, if the condition $d\lambda^{(i_q)} > 0$ is satisfied together with the local yield criterion (17), (18). By this means for a prescribed loading path $\boldsymbol{\sigma}^0(t)$, (where t is a monotonically varying parameter) the system of Eq. (25) becomes a system of six ordinary differential equations for the determination of all components of the tensor $\langle \boldsymbol{\sigma} \rangle^{(q)}$. In so doing only five from these six equations are independent by virtue of the yield condition (17). After integration of the system (25) the averaged plastic strains inside each subdomain are found and the averaged strains are given by $\langle \boldsymbol{\varepsilon} \rangle^{(i_q)} = \mathbf{M}^{(0)} \langle \boldsymbol{\sigma} \rangle^{(i_q)} + \boldsymbol{\varepsilon}^{p(i_q)}$. The overall plastic strains $\boldsymbol{\varepsilon}^p_*$ and the total overall strains are defined by the relations $\boldsymbol{\varepsilon}^p_* = \Sigma_j (\mathbf{B}^{*(j)})^T \boldsymbol{\varepsilon}^{p(j)} c^{(j)}$, $(j=0,\ldots,n^c+1)$, $\langle \boldsymbol{\varepsilon} \rangle = \mathbf{M}^* \boldsymbol{\sigma}^0 + \boldsymbol{\varepsilon}^p_*$. The numerical integration of Eq. (25) under the restriction (18) can be carried out by different integration schemes analogous to the treatment of homogeneous materials with vertex yield surfaces as for example described in Ortiz and Popov [38], Ray and Utku [39], Nemat-Nasser and Hori [40], and Papadopoulos and Taylor [41]. The same methodology usually is applied in particular versions of mean field methods in the mechanics of composite materials by Lagoudas et al. [42], Dvorak et al. [3], and Isupov [43]. In this work a backward difference method with a variable integration step has been used as the integration scheme for Eq. (20). An iterative procedure is used, starting with the elastic predictor (27) which provides a first estimate of the stress increment in the component. This means that at each step an Euler-backward scheme is used in combination with a Newton method, with the initial guess being determined by the elastic predictor.

6 Numerical Results

We consider an isotropic composite consisting of a steel matrix and identical spherical carbidic inclusions with elastoplastic parameters (ν is the Poisson's ratio):

Phase	k (GPa)	ν	τ_0 (GPa)	h (GPa)	n
matrix	175	0.3	2.75	1.5	0.5
inclusion	300	0.25	∞	-	-

At first we consider the case of a dilute concentration of the inclusions $c^{(1)} \ll 1$. This leads to $\mathbf{Y} \equiv \mathbf{I}$, see (8), and we have the possibility to check the quality of our local plasticity model by the use of finite element analysis (FEA). Comparison of the accumulated effective plastic strains $\gamma(\mathbf{x})/c^{(1)}$ $(\mathbf{x} \in v^c_i)$, calculated by the proposed analytical method with results obtained by finite element analysis (FEA), are represented in Fig. 1 for hydrostatic loading $\boldsymbol{\sigma}^0 = \sigma^0_0 \boldsymbol{\delta}$. FEA results are presented for two points in the coating v^c_i, near the boundaries with the core v^s_i $(|\mathbf{x}| = a+0)$ and with the matrix $v^{(0)}$ $(|\mathbf{x}| = a(1+\xi)-0)$, respectively; here ξ is the relative thickness of the coating, and three thicknesses are considered in Fig. 1: $\xi = 0.033$, $\xi = 0.1$. In the local plasticity theory the subdivision of the coating v^c_i into one layer of individual subdomains v^c_{ij} $(j=1,\ldots,n^c)$ is employed by the use of sections $\theta = \pi/n^c$ in a spherical coordinate system (θ, ϕ, r), the origin of which coincides with the center of the inclusion. Hereafter we will restrict our problems to axisymmetric loading and, therefore, the plastic strains in the ribbon $a < r < a(1+\xi)$, $k\pi/n^c < \phi < (k+1)\pi/n^c$ $(k=0,\ldots,n^c-1)$ can be obtained by the rotation around the axis $\theta = 0$. As we see in Fig. 1 for the considered case of $n^c = 11$ and for a sufficiently thin coating $(\xi \leqslant 0.1)$ the proposed analytical model provides satisfactory exactness and, as a consequence, from here on we will consider the thickness $\xi = 0.1$ only. In Fig. 2 the

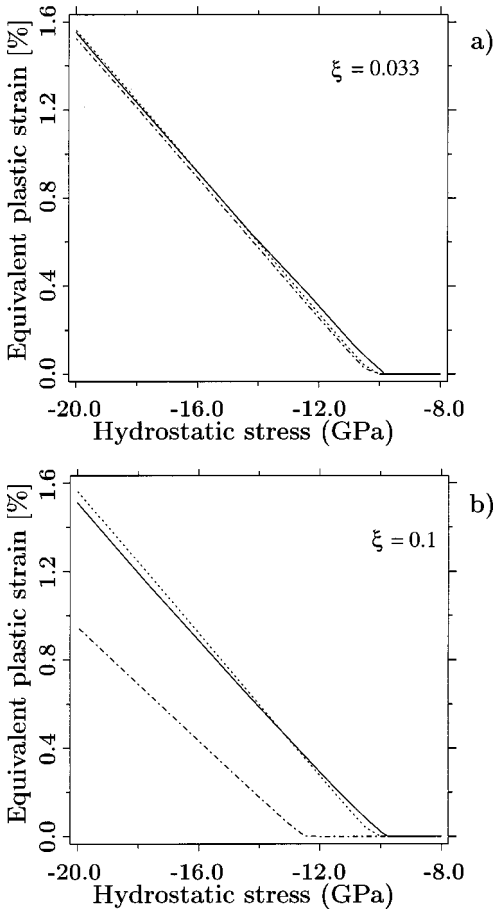


Fig. 1 Accumulated effective plastic strains $\gamma(\mathbf{x})/c^{(1)}$ as a function of hydrostatic loading σ_0^0 calculated by FEA (dashed curve for $|x|=a+0$, dot-dashed curve for $|x|=a(1+\xi)-0$) and by the proposed model (solid curve). (*a*) $\xi=0.033$, (b) $\xi=0.1$

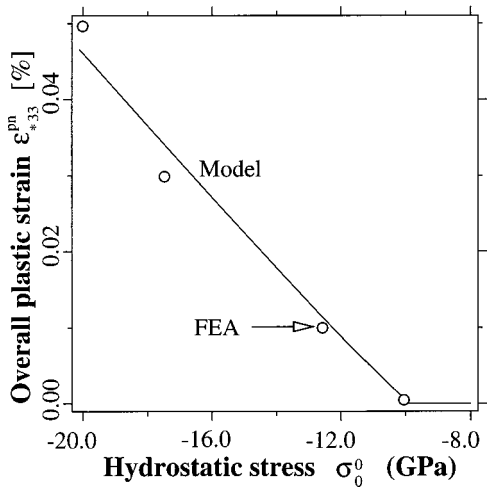


Fig. 2 Normalized overall plastic strain ϵ_{*33}^{pn} as a function of hydrostatic loading σ_0^0 calculated for $\xi=0.1$ by the proposed model (12), (17)–(19) (solid curve) and by FEA (circles)

normalized overall plastic strains $\varepsilon_*^{pn}\equiv\varepsilon_*^p/c^{(1)}$ as functions of the loading σ_0^0 are calculated by the use of both the proposed model and of FEA. For uniaxial loading $\sigma_{ij}^0=\sigma_{33}^0\delta_{i3}\delta_{j3}$, $\sigma_{33}^0\sim t$ the influence of the mesh width of the uniform subdivision (π/n^c) of the coating on the local values of the von Mises effective plastic strains $\gamma^{(i)}$ $(i=2,\ldots,n^c+1)$ (23) is studied; here t is the time, i.e., a monotonically increasing parameter. The results shown in Fig. 3 were calculated for both $n^c=11$ and for $n^c=33$. In Fig. 3 the values $\gamma(\mathbf{x})$ as a function of the polar angle $0<\theta<90$ deg are shown for $\sigma_{33}^0=2.8$ GPa, whereas onset of yielding takes place at $\sigma_{33}^0=2.2$ GPa. As can be seen for the considered case both subdivisions of the coating ($n^c=11$ and $n^c=33$) lead to similar results for the local values of the von Mises effective plastic strains $\gamma^{(i)}$ $(i=2,\ldots,n^c+1)$ (23). The increase of the degree of subdivision, $n^c=33$, leads to an increase in the calculated values ε_*^{pn} which is however smaller than 0.5 percent. This leads to the conclusion that even a crude subdivision ($n^c=11$) provides good estimates for the overall plastic strains ε_*^{pc}. Because of this in the following we will consider the fixed thickness $\xi=0.1$ and the fixed degree of subdivision $n^c=11$. Let us now study the finite concentration of carbidic inclusions ($c^{(1)}=0.25$), for which the FEA-unit cell analysis is not able to capture real random arrangements of the inclusions. In this case the effective field $\langle\overline{\boldsymbol{\sigma}}\rangle_i\neq\boldsymbol{\sigma}^0$ but is defined by particle interaction and by accumulated plastic strains (both in the coating and in the remaining matrix). Whereas the solution of the corresponding linear problem can be found by any known method (see, e.g., [29,40]), in the current paper we will use only the solution of the linear elastic problem (8), (9) by the two-particle approximation of MEFM (see, e.g., [29,34]). Let us assume uniaxial tensional loading $\sigma_{ij}^0=\sigma_{33}^0\delta_{i3}\delta_{j3}$, $\sigma_{33}^0\sim t$. Figure 4 shows the comparison of the overall plastic strains ε_{*33}^p once calculated by the use of traditional mean field method assumptions (11) and once by the proposed assumptions (12). As can be seen the employment of the proposed criterion (12) tends to decrease the overall initial yield stress of the composite. Furthermore, the concentration of plastic strains in the coating leads to a ''softening'' effect of the coated inclusions. This ''softening'' results in a redistribution of the stresses from the coated inclusions to the matrix, which causes an increase of plastic deformations of the matrix. The limiting case of the ''softening'' of the inclusions is the replacement of them by voids, resulting in a significant increase of the overall plastic strains (such a result is presented in Fig. 4 by the use of the traditional scheme (11), (17)–(19)). For the qualitative comparison in Fig. 4 the overall plastic strains are calculated also by a FEA-unit cell model for face-centered-cubic (FCC) packing of the inclusions, when the orientation of unit cell coincides with the global coordinate system. For the considered

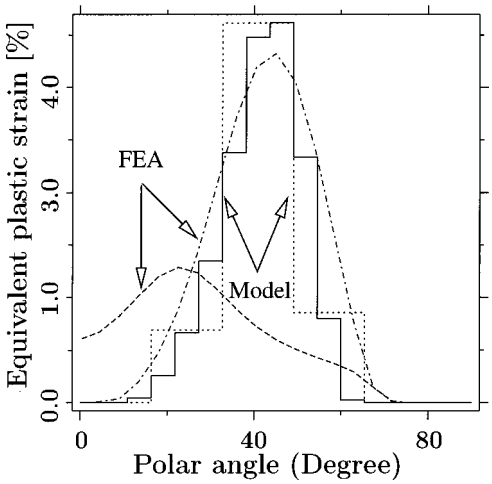


Fig. 3 Accumulated effective plastic strains $\gamma(\mathbf{x})$ as a function of the polar angle θ, calculated by FEA (dot-dashed curve for $|\mathbf{x}|=a+0$, dotted curve for $|\mathbf{x}|=a(1+\xi)-0$) and by the proposed model (dashed curve for $n^c=11$, solid curve for $n^c=33$)

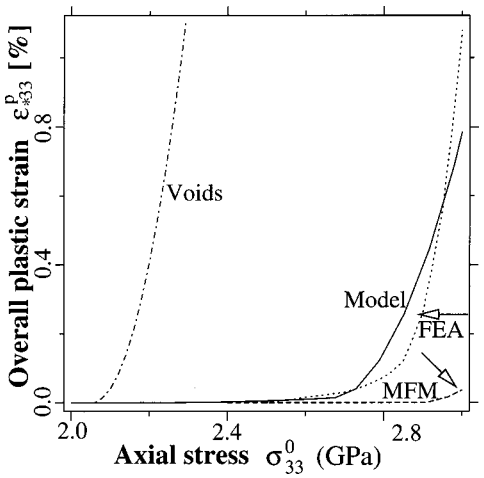


Fig. 4 Overall plastic strain ϵ^p_{*33} as a function of uniaxial loading σ^0_{33} calculated by the proposed model (12) (solid curve), by FEA (dotted curve) and by the traditional mean field method (11) (dashed curve). Overall plastic strain ϵ^p_{*33} for model material with replacement of the inclusions by voids (dot-dashed curve).

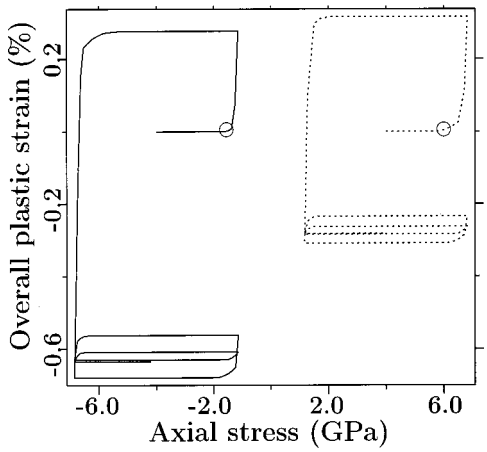


Fig. 6 Overall plastic strain ϵ^p_{*33} as a function of uniaxial cyclic loading σ^0_{33} calculated for different constant hydrostatic contributions $\sigma^{fix}_0 = -4$ GPa (solid curve), 4 GPa (dotted curve); $\sigma^{max}_{33} = 2.8$ GPa, (○)–onset of yielding

small concentration of the inclusions ($c = 0.25$) the difference of the effective elastic moduli for FCC packing and for random arrangement should be rather small. As can be seen in Fig. 4 the error of the proposed model (in comparison with FEA) is much smaller than that of the traditional mean field method (11). This significant difference between the use of yield functions (11) and (12) are explained by the possibility of the proposed model to pick up the effect of plastic strain inhomogeneity inside the coating, i.e., around the inclusions. Let us now compare the estimations obtained by the proposed model and by other popular methods. There are different versions of mean field methods employing flow theory (Dvorak [2]) and secant concept method (e.g., Qiu and Weng, [26]). In Fig. 5 the curves 4 and 5 are calculated in the framework of mean field method (11) (at $\xi = 0$) by the secant concept method and by the flow theory, respectively. Modified

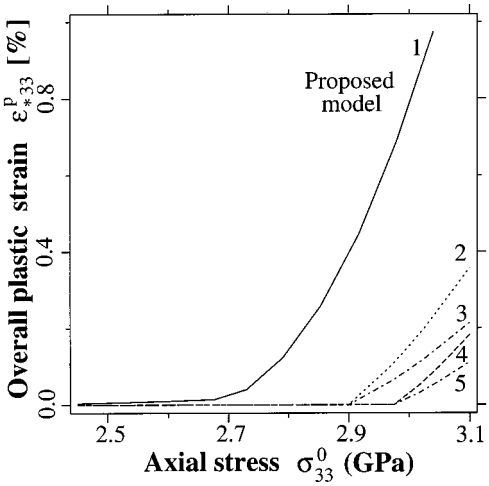


Fig. 5 Overall plastic strain ϵ^p_{*33} as a function of uniaxial loading σ^0_{33} calculated by the proposed model (12) (solid curve 1), by a modified approach based on the the estimations of second moment of stresses (dotted curve 2—secant concept method, dot-dashed curve 3—flow theory), and by mean field method (dashed curve 4—secant concept method, dot-dashed curve 5—flow theory)

improved methods utilize the estimations of statistical averages in each component of either the stress potential or the second moment of stresses (16). Such estimations take into account (in the average sense) the measure of stress inhomogeneity in the matrix and can be found by perturbation methods and the method of integral equations (see for references Buryachenko [29]) as well as by variational methods (see for references Ponte Castañeda and Suquet [32]). For our elastoplastic analysis in the framework of homogeneity of plastic strains in the matrix ($\xi = 0$) we will use the exact perturbation method in combination with the secant concept method by Buryachenko [28] as well as the method of integral equations in combination with flow theory (see for references Buryachenko [29]). The corresponding curves 2 and 3, respectively, are presented in Fig. 5. Finally, the proposed local plasticity model takes into account the inhomogeneity of plastic strains in some matrix layer around the inclusions ($\xi = 0.1$) as well as the inhomogeneity of the first-order moment of the stresses in the coating (curve 1). As can be seen from Fig. 5 the mean field method approach (curves 4 and 5) yields predictions which are stiffer than the predictions of the modified approach (curves 2 and 3). Secant concept model (curves 2 and 4) seems softer by comparison with flow theory (curves 3 and 5) both for classical and for modified methods, based on the first-order moments and second-order moments of the stresses in each individual phase, respectively. The known question regarding the correctness of the secant concept model is not discussed here. It is just mentioned that even for radial external loading the stress path at the local level is not a radial one. At the same time in the considered case ($\sigma^0_{33} = 3$ GPa) the modified secant approach (curve 2) predicts values of the overall plastic strain ε^p_{*33} which are considerably smaller than those calculated by the use of both FEA-unit cell model and the local model of plasticity for random packing of the inclusions (curve 1). As can be seen in Fig. 5 the better discretization of the stress-plastic strain state in the proposed model (curve 1) leads to significantly improved predictions of the overall plastic strains (at least with respect to the comparison of the presented estimations for random structure composites with the FEA-unit cell model).

As a further example we consider a cyclic external loading, described $\boldsymbol{\sigma}^0 = \boldsymbol{\sigma}^{fix} + \boldsymbol{\sigma}^{\max} + (\boldsymbol{\sigma}^{\min} - \boldsymbol{\sigma}^{\max}))(t-1)[H(t-1) - H(1-t)]$, which is a combination of a constant hydrostatic loading $\sigma^{fix}_{ij} = \sigma^{fix}_0 \delta_{ij}$ and a uniaxial cyclic stress with zero mean-stress and an amplitude $\sigma^{am} = \sigma^{\max}_{33}$, $\sigma^{\max} = -\sigma^{\min}$, $\sigma^{\max}_{ij} = \sigma^{\max}_{33} \delta_{i3} \delta_{j3}$; here H is the Heaviside step function and t is the time. From Fig. 6 we see that the overall plastic strain component ϵ^p_{*33} after the

first cycle (with $\sigma_{33}^{\max}=2.8$ GPa) converges to a steady-state response (or steady cycle), i.e., a closed loop is received. From Fig. 6 we see a significant role of the hydrostatic loading which leads to a considerable variation of the closed loop of the overall plastic strains (plastic shakedown). It should be mentioned that the use of the criterion (11) tends to an overprediction of the initial yield stress of the composite, $\sigma_{33}^{0Y}=2.92$ GPa instead of $\sigma_{33}^{0Y}=2.34$ GPa obtained by the proposed method, and to independence of plastic deformations on hydrostatical loading $\boldsymbol{\sigma}^{fix}=\sigma_0^{fix}\boldsymbol{\delta}$. Therefore, in the considered range of external loading $\sigma_{33}^{\max}<2.92$ GPa by using (11) the composite material would deform elastically, and for $\sigma_{33}^{\max}>2.92$ GPa after the first cycle of elastoplastic deformation the process would result in elastic shakedown.

7 Concluding Remarks

Let us discuss the main hypotheses as well as the limitations of the proposed estimations and their possible generalizations. The possible constitutive relations are not limited to the von Mises yield criterion (1) assumed in this study, and modification of the present method to accommodate general yield criteria and general hardening laws can be performed. It is only important that response of a ''coated'' inclusion is defined by the Eqs. (6) and (7), notwithstanding the inclusion can be considered as some sort of a ''black box.'' The local model proposed can be applied to a wide class of nonlinear problems for which the local properties of the components become location-dependent: nonlinear elasticity and conductivity, viscosity and creeping, and viscoplasticity. Moreover, the model proposed can be generalized easily to any number of thin coating layers of nonellipsoidal inclusions (see Buryachenko and Rammerstorfer [34]). The next step of the improvement of the proposed model is a consideration of statistical averages of the second moments of stresses in both the matrix and the coating estimated by the method of integral equations by (see, e.g., Buryachenko [34]). The model proposed for composites with identical inclusions is valid for any averaging method of micromechanics using the hypothesis **H1**, and can be generalized in the framework of the MEFM by Buryachenko [34] to composites with any number of different components containing inclusions with a distribution of size, shape, orientation, and properties. The principal limitation of this paper is due to the assumption of statistical homogeneity of the composite microstructure. Nevertheless, in the light of the researches on functionally graded composites performed by the MEFM (see [34]) obviously, consideration of the local model of elastoplastic deformation of graded materials are also possible.

Acknowledgments

Parts of this work were financially supported by the Österreichisches Bundesministerium für Wissenschaft, Forschung und Kunst (under grant GZ 49.935/3-II/4/94) and by the AFOSR of USA. The first author expresses his appreciation to Prof. G. J. Dvorak for helpful discussions and to MSc. E. Gavrilova for preparation of the manuscript.

Appendix

Single Ellipsoidal Inclusion. Buryachenko and Rammerstorfer [34] (see there additional references also) obtained an approximate solution for a single ellipsoidal inclusion with thin coating under the approximative assumption of a homogeneous stress state in the core, $\boldsymbol{\sigma}^i$, and at the infinity, $\boldsymbol{\sigma}^0:\boldsymbol{\sigma}(\mathbf{x})\equiv\boldsymbol{\sigma}^i=\text{const}$, $\mathbf{x}\in v^i$, $\boldsymbol{\sigma}^0=\text{const}$. Then the stresses in the coating $\boldsymbol{\sigma}^c(\mathbf{s})$ are found by the relation $\boldsymbol{\sigma}^c(\mathbf{s})=\boldsymbol{\sigma}^i+\tilde{\boldsymbol{\Gamma}}(\mathbf{n})\,[\mathbf{M}_1^{(1)}(\mathbf{s})\boldsymbol{\sigma}^i-\boldsymbol{\varepsilon}^p(\mathbf{s})]$, where the interface operator $\tilde{\boldsymbol{\Gamma}}(\mathbf{n})$ is defined by $\tilde{\boldsymbol{\Gamma}}(\mathbf{n})=\mathbf{L}^{(0)}-\mathbf{L}^{(0)}\mathbf{U}(\mathbf{n})\mathbf{L}^{(0)}$, $\mathbf{U}(\mathbf{n})=\mathbf{n}\otimes\mathbf{G}(\mathbf{n})\otimes\mathbf{n}$, $\mathbf{G}(\mathbf{n})=\mathbf{L}(\mathbf{n})^{-1}$, $\mathbf{L}(\mathbf{n})=\mathbf{L}^{(0)}(\mathbf{n}\otimes\mathbf{n})$, $\mathbf{n}$ is the unit outward normal vector on s^i in the point $\mathbf{s}\in s^i$, s^i is the outer surface of the boundary ∂v_i^s.

Let the volume v_i^c be subdivided into one layer of n^c individual local volumes v_{ij}^c ($\text{mes}\, v_{ij}^c \ll \text{mes}\, v_i^c, j=1,\dots n^c$), such that the plastic strain $\boldsymbol{\varepsilon}^p(\mathbf{x})$ and $\boldsymbol{\sigma}^c(\mathbf{x})$ ($\mathbf{x}\in v_{ij}^c$) are constant inside each subdomain v_{ij}^c. For example in a spherical coordinates system (θ,φ,r) coinciding with the semi-axes a_j^s this subdivision can be done by the use of cutting the coating v_i^c along surfaces $\theta=\text{const}$, $\varphi=\text{const}$. Estimations for the tensors in Eqs. (6) and (7) can be obtained

$$\mathbf{B}^{(1)}=(\mathbf{I}+\mathbf{Q}\mathbf{M}_1^{(1)})^{-1},\quad \mathbf{B}^{(l)}=[\mathbf{I}+\tilde{\boldsymbol{\Gamma}}(\mathbf{n}_l)\mathbf{M}_1^{(1)}]\mathbf{B}^{(1)},$$

$$\mathbf{C}^{(1l)}=-\mathbf{B}^{(1)}\mathbf{Q}\left[\delta_{1l}+\frac{\overline{v_{il}}}{\overline{v_i}}(1-\delta_{1l})\right]+\frac{\overline{v_{il}}}{\overline{v_i}}(1-\delta_{1l})\mathbf{B}^{(1)}\boldsymbol{\Gamma}(\mathbf{x}_i-\mathbf{y}_{il}),$$

$$\mathbf{C}^{(ml)}=[\mathbf{I}+\tilde{\boldsymbol{\Gamma}}(\mathbf{n}_m)\mathbf{M}_1^{(1)}]\mathbf{C}^{(1l)}+\tilde{\boldsymbol{\Gamma}}(\mathbf{n}_m)(\delta_{1l}-\delta_{ml}),$$

$$\mathbf{B}\equiv\frac{1}{1-c^{(0)}}\sum_{j=1}^{n^c+1}c^{(j)}\mathbf{B}^{(j)}=\mathbf{I}+\frac{\overline{v_i^s}}{\overline{v_i}}(\mathbf{B}^{(1)}-\mathbf{I}),$$

$$\mathbf{C}=\frac{1}{1-c^{(0)}}\sum_{j,k=1}^{n^c+1}c^{(j)}\mathbf{C}^{(jk)}\boldsymbol{\varepsilon}_1^{p(k)},$$

where $l,\ m=2,\dots,n^c+1$; $\mathbf{x}_i$ and $\mathbf{y}_{il}$ are the centers of the domains v_i and v_{il}^c, respectively; $\mathbf{n}_l$ is the unit outward normal vector on s^i in the point $\mathbf{y}_{il}$.

References

[1] Dvorak, G. J., 1992, ''Transformation Field Analysis of In-Elastic Composite Materials,'' Proc. R. Soc. London, Ser. A, **437**, pp. 311–330.

[2] Dvorak, G. J., 1993, ''Nadai Lecture—Micromechanics of In-Elastic Composite Materials: Theory and Experiment,'' ASME J. Eng. Mater. Technol., **115**, pp. 330–338.

[3] Dvorak, G. J., Bahei-El-Din, Y. A., and Wafa, A. M., 1994, ''The Modeling of Inelastic Composite Materials With the Transformation Field Analysis,'' Modell. Simul. Mater. Sci. Eng., **2**, pp. 571–585.

[4] Hutchinson, J. W., 1976, ''Bounds and Self-Consistent Estimates for Creep of Polycrystalline Materials,'' Proc. R. Soc. London, Ser. A, **348**, pp. 101–130.

[5] Nemat-Nasser, S., and Obata, M., 1986, ''Rate-Dependent, Finite Deformation of Polycristals,'' Proc. R. Soc. London, Ser. A, **A407**, pp. 377–404.

[6] Corvasce, F., Lipinski, P., and Berveiller, M., 1990, ''The Effects of Thermal, Plastic and Elastic Stress Concentration on the Overall Behavior of Metal Matrix Composites,'' *Inelastic Deformation of Composite Materials*, ed. G. J. Dvorak, ed., Springer-Verlag, New York, pp. 389–408.

[7] Tandon, G. P., and Weng, G. J., 1988, ''A Theory of Particle-Reinforced Plasticity,'' ASME J. Appl. Mech., **55**, pp. 126–135.

[8] Wakashima, K., and Tsukamoto, H., 1988, ''A Unified Micromechanical Approach Toward Thermomechanical Tailoring of Metal Matrix Composites,'' ISIJ Int., **32**, pp. 883–892.

[9] Arsenault, R. J., and Taya, M., 1987, ''Thermal Residual Stress in Metal Matrix Composite,'' Acta Metall., **35**, pp. 651–659.

[10] Reifsnider, K. L., and Gao, Z., 1991, ''A Micromechanics Model for Composites Under Fatigue Loading,'' Int. J. Fatigue, **13**, pp. 149–156.

[11] Lin, S. C., Yang, C., Mura, T., and Iwakuma, T., 1992, ''Average Elastic-Plastic Behavior of Composite Materials,'' Int. J. Solids Struct., (George Hermann 70th Anniversary Issue), **28**, pp. 1859–1872.

[12] Qiu, Y. P., and Weng, G. J., 1991, ''The Influence of Inclusion Shape on the Overall Elastoplastic Behavior of a Two-Phase Isotropic Composite,'' Int. J. Solids Struct., **27**, pp. 1537–1550.

[13] Zhu, Z. G., and Weng, G. J., 1989, ''Creep Deformation of Particle-Strengthened Metal-Matrix Composites,'' J. Eng. Mater. Technol., **111**, pp. 99–105.

[14] Accorsi, M. L., and Nemat-Nasser, S., 1986, ''Bounds on the Overall Elastic and Instantaneous Elastoplastic Moduli of Periodic Composites,'' Mech. Mater., **5**, pp. 209–220.

[15] Weissenbek, E., Böhm, H. J., and Rammerstorfer, F. G., 1994, ''Micromechanical Investigations of Arrangement Effects in Particle Reinforced Metal Matrix Composites,'' Comput. Mater. Sci., **3**, pp. 263–278.

[16] Needleman, A., and Tvergard, V., 1993, ''Comparison of Crystal Plasticity and Isotropic Hardening Predictions for Metal-Matrix Composites,'' ASME J. Appl. Mech., **60**, pp. 70–76.

[17] Nakamura, T., and Suresh, S., 1993, ''Effect of Thermal Residual Stresses and Fiber Packing on Deformation of Metal-Matrix Composites,'' Acta Metall. Mater., **41**, pp. 1665–1681.

[18] Duva, J. M., 1988, ''A Constitutive Description of Nonlinear Materials Containing Voids,'' Mech. Mater., **5**, pp. 137–144.
[19] Lee, B. J., and Mear, M. E., 1991, ''Effect of Inclusion Shape on the Stiffness of Non-Linear Two-Phase Composites,'' J. Mech. Phys. Solids, **39**, pp. 630–649.
[20] Hashin, Z., 1983, ''Analysis of Composite Materials—A Survey,'' ASME J. Appl. Mech., **50**, pp. 481–505.
[21] Zhu, Z. G., and Weng, G. J., 1990, ''A Local Theory for the Calculation of Overall Creep Strain of Particle-Reinforced Composites,'' Int. J. Plast., **6**, pp. 449–469.
[22] Olsson, M., Giannakopoulos, A. E., and Suresh, S., 1995, ''Elastoplastic Analysis of Thermal Cycling: Ceramic Particles in a Metallic Matrix,'' J. Mech. Phys. Solids, **43**, pp. 1639–1671.
[23] Taggart, D. G., and Bassani, J. L., 1991, ''Elastic-Plastic Behavior of Particle Reinforced Composites—Influence of Residual Stresses,'' Mech. Mater., **12**, pp. 63–80.
[24] Herrmann, K. P., and Mihovsky, I. M., 1994, ''On the Modeling of the Inelastic Thermomechanical Behavior and the Failure of Fibre-Reinforced Composites—A Unified Approach,'' *Advances in Applied Mechanics*, K. Z. Markov, ed, World Scientific, Singapore, pp. 141–191.
[25] Buryachenko, V. A., and Lipanov, A. M., 1989, ''Effective Field Method in the Theory of Perfect Plasticity of Composite Materials,'' Prikl. Mekh. i Tekh. Fiz., (3), pp. 149/N155. (in Russian. Engl. Transl.: J. Appl. Mech. Tech. Phys., 1989, **30**, pp. 482–487.
[26] Qiu, Y. P., and Weng, G. J., 1992, ''A Theory of Plasticity for Porous Materials and Particle-Reinforced Composites,'' ASME J. Appl. Mech., **59**, pp. 261–268.
[27] Ding, K., and Weng, G. J., 1998, ''Plasticity of Particle-Reinforced Composites With a Ductile Interface,'' ASME J. Appl. Mech., **65**, pp. 596–601.
[28] Buryachenko, V. A., 1996, ''The Overall Elastoplastic Behavior of Multiphase Materials With Isotropic Components,'' Acta Mech., **119**, pp. 93–117.
[29] Buryachenko, V. A., 2001, ''Multiparticle Effective Field Method and Related Methods in Micromechanics of Composite Materials,'' Appl. Mech. Rev., **54**(1), pp. 1–47.
[30] Hashin, Z., and Shtrikman, S., 1963, ''A Variational Approach to the Theory of the Behavior of Multiphase Materials,'' J. Mech. Phys. Solids, **11**, pp. 130–140.
[31] Talbot, D. R. S., and Willis, J. R., 1985, ''Variational Principles for Nonlinear Inhomogeneous Media,'' IMA J. Appl. Math., **35**, pp. 39–54.
[32] Ponte Castañeda, P., and Suquet, P., 1998, ''Nonlinear Composites,'' Adv. Appl. Mech., **34**, pp. 171–302.
[33] Willis, J. R., 1981, ''Variational and Related Methods for the Overall Properties of Composites,'' Adv. Appl. Mech., **21**, pp. 1–78.
[34] Buryachenko, V. A., and Rammerstorfer, F. G., 2001, ''On the Thermostatics of Composites With Coated Random Distributed Inclusions,'' Int. J. Solids Struct., **37**, pp. 3177–3200.
[35] Hervé, E., and Zaoui, A., 1993, ''N-Layered Inclusion-Based Micromechanical Modeling,'' Int. J. Eng. Sci., **31**, pp. 1–10.
[36] Bornet, M., Hervé, E., Stolz, C., and Zaoui, A., 1994, ''Self-Consistent Approaches and Strain Heterogeneities in Two-Phase Elastoplastic Materials,'' Appl. Mech. Rev., **47**(1), pp. S66–S76.
[37] Buryachenko, V. A., and Rammerstorfer, F. G., 1999, ''On the Thermoelasticity of Random Structure Particulate Composites,'' ZAMP, **50**, pp. 934–947.
[38] Ortiz, M., and Popov, E. P., 1985, ''Accuracy and Stability of Integration Algorithms for Elastoplastic Constitutive Relations,'' Int. J. Numer. Methods Eng., **21**, pp. 1561–1576.
[39] Ray, S. K., and Utku, S., 1989, ''A Numerical Model for the Thermo-elastoplastic Behavior of a Material,'' Int. J. Numer. Methods Eng., **28**, pp. 1103–1114.
[40] Nemat-Nasser, S., and Hori, M., 1993, *Micromechanics: Overall Properties of Heterogeneous Materials*. North-Holland, Amsterdam.
[41] Papadopoulos, P., and Taylor, R. L., 1994, ''On the Application of Multi-step Integration Methods to Infinitesimal Elastoplasticity,'' Int. J. Numer. Methods Eng., **37**, pp. 3169–3184.
[42] Lagoudas, D. S., Gavazzi, A. C., and Nigam, H., 1991, ''Elastoplastic Behavior of Metal Matrix Composites Based on Incremental Plasticity and the Mori-Tanaka Averaging Scheme,'' Comput. Mech., **8**, pp. 193–203.
[43] Isupov, L. P., 1996, ''Constitutive Equations of Plastic Anisotropic Composite Medium,'' *IUTAM Symp. Micromech. of Plasticity and Damage of Multiphase Materials*, A. Pineau and Z. Zaoui, eds., Kluwer Academic Publ., Dordrecht, pp. 91–98.

Z. C. Feng

B. He

Department of Mechanical and Aerospace Engineering, University of Missouri–Columbia, Columbia, MO 65211

S. J. Lombardo

Department of Chemical Engineering, University of Missouri–Columbia, Columbia, MO 65211

Stress Distribution in Porous Ceramic Bodies During Binder Burnout

A model has been developed for describing the stresses that arise during binder burnout in three-dimensional porous bodies. The pressure gradient that arises from the decomposition of binder in the pore space is treated as an equivalent body force. For input into the mechanics model, the pressure distribution is obtained from the analytical solution for three-dimensional porous bodies with anisotropic permeability. The normal and shear stresses are then calculated from finite element analysis for bodies of parallelepiped geometry. In general, the normal stresses occur at the center of the body and are an order of magnitude larger than the shear stresses. Both the normal and shear stresses depend on the body size, the body geometry, and on the permeability. [DOI: 10.1115/1.1460908]

1 Introduction

In the fabrication of ceramic components by powder processing routes, the strength of formed bodies is often enhanced by the addition of polymeric binders ([1]). These polymers are then removed in later processing steps by decomposing them at elevated temperature into gas-phase products. Two types of models ([2–20]) have appeared for describing how the degradation products exit the body. For bodies having large volume fractions of binder and thus low initial porosity, diffusion of the decomposition products through the remaining binder in the nearly filled space has been treated as the rate-limiting step ([2–11]). For bodies prepared with lower volume fractions of binder, gas-phase flow is the faster transport mechanism ([12–20]). In both cases, however, a distribution of pressure arises within the pore space of the body.

Depending on the processing parameters and the size of the component, the pressure increase inside the ceramic body during binder burnout can be so large as to decrease the product yield ([7,19,21]). In Liau et al. [19], it was demonstrated that the yield depended on the heating cycle and on the dimensions of square-parallelepiped multilayer ceramic capacitors (MLCs). The authors also noted that the flow of the gas-phase decomposition products in the MLCs was enhanced in the direction parallel to the planes containing the metal electrodes, which suggests that the permeability is anisotropic within the body. To describe these observations, an analytical model ([20]) has been developed for describing the pressure distribution during binder burnout in three-dimensional porous ceramic bodies with anisotropic permeability.

Although the pressure is known to increase within the body during binder burnout, the mechanical stresses arising during this process are ultimately what cause the ceramic component to fail. When the pressure distribution is known, the distribution of stresses can be calculated by adopting mechanical models for the ceramic body. Stangle and Aksay [14] used the stress formulation developed for partially saturated granular media ([22]) to calculate the stress distribution in spherical bodies. They considered both capillary stresses due to liquid-phase binders and stresses due to the increase in pressure arising from binder degradation products. Tsai [15] considered binder burnout in axisymmetric porous bodies. Both the radial and tangential stresses, which satisfy the equilibrium equations, were considered to consist of contributions from the gas phase and the solid phase (skeleton) by an equal amount. An isotropic constitutive law was used for the stress-strain relationship for the skeleton and the void fraction does not appear explicitly in the stress formulation. This pair of models take as their starting point that the local pressure and the local stress are linearly additive quantities.

In this paper, we propose a general three-dimensional model for the stresses caused by the buildup of internal pressure in a ceramic body during binder removal. Instead of calculating average stresses in porous media as is conventionally done ([23]), we model the ceramic body as a solid skeleton permeated by a gaseous phase whose pressure follows a known distribution. We first derive the equilibrium equations for the stresses on the skeleton and use an isotropic constitutive law. The gradient of the internal gas pressure is shown to be responsible for the stresses, and we develop a model formulation in which the pressure gradient is treated as an ''equivalent body force.'' Numerical methods are then used to solve for the stress distribution in rectangular parallelepipeds, a common geometry of ceramic capacitors.

2 The Mechanics Model

We make the following assumptions in the development of the model:

- The porous ceramic during the burnout process consists of a void fraction, ε, and a solid skeleton fraction, 1-ε.
- The void space is occupied by gas of internal pressure, p, which varies continuously in the pore space of the body. The gas phase cannot support shear stresses and both viscous and inertial stresses accompanying fluid flow are neglected.
- The solid skeleton can be modeled as an isotropic linear-elastic solid.

In Fig. 1, we show a free-body diagram of a two-dimensional infinitesimal volume. Force and moment equilibrium leads to the following equations:

$$(1-\varepsilon)\left(\frac{\partial \sigma_{xx}}{\partial x}+\frac{\partial \sigma_{xy}}{\partial y}\right)=\varepsilon\frac{\partial p}{\partial x} \tag{1}$$

$$(1-\varepsilon)\left(\frac{\partial \sigma_{xy}}{\partial x}+\frac{\partial \sigma_{yy}}{\partial y}\right)=\varepsilon\frac{\partial p}{\partial y} \tag{2}$$

where σ_{xx}, σ_{xy}, and σ_{yy} are the components of true stress on the skeleton. Obviously, an increase in internal pressure causes stress, an effect that is proportional to the void fraction, ε, appearing on the right-hand side of Eqs. (1) and (2). The factor of 1-ε on the

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 19, 2001; final revision, Oct. 16, 2001. Associate Editor: H. Gao. 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.

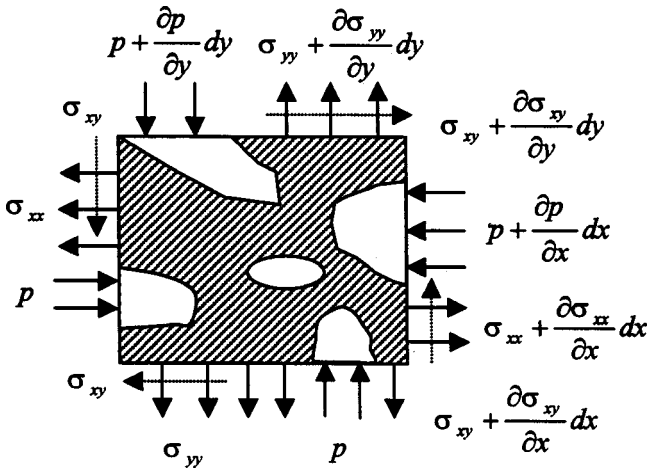


Fig. 1 Free-body diagram of an infinitesimal two-dimensional element showing the solid skeleton (hatched) and the continuous porous network. The shear stresses are denoted by the dashed arrows.

left-hand side of Eqs. (1) and (2) arises because only the skeletal part of the body bears the stress. The occurrence of a pressure gradient is thus equivalent to a body force throughout the continuum.

The stresses are often defined as the force per unit area without subtracting the nonload bearing voids. These nominal stresses, $\tilde{\sigma}$, are smaller than the true stresses as given by

$$\tilde{\sigma}_{ij}=(1-\varepsilon)\sigma_{ij} \tag{3}$$

where i and j correspond to x and y. In terms of the nominal stresses, the equilibrium equations become

$$\frac{\partial \tilde{\sigma}_{xx}}{\partial x}+\frac{\partial \tilde{\sigma}_{xy}}{\partial y}=\varepsilon\frac{\partial p}{\partial x} \tag{4}$$

$$\frac{\partial \tilde{\sigma}_{xy}}{\partial x}+\frac{\partial \tilde{\sigma}_{yy}}{\partial y}=\varepsilon\frac{\partial p}{\partial y}. \tag{5}$$

Note that the nominal stresses differ from the true stresses by a constant factor, which is only dependent on the void fraction.

The equilibrium equations in two dimensions can be generalized into three dimensions as

$$\frac{\partial \tilde{\sigma}_{xx}}{\partial x}+\frac{\partial \tilde{\sigma}_{xy}}{\partial y}+\frac{\partial \tilde{\sigma}_{xz}}{\partial z}=\varepsilon\frac{\partial p}{\partial x} \tag{6}$$

$$\frac{\partial \tilde{\sigma}_{xy}}{\partial x}+\frac{\partial \tilde{\sigma}_{yy}}{\partial y}+\frac{\partial \tilde{\sigma}_{yz}}{\partial z}=\varepsilon\frac{\partial p}{\partial y} \tag{7}$$

$$\frac{\partial \tilde{\sigma}_{xz}}{\partial x}+\frac{\partial \tilde{\sigma}_{yz}}{\partial y}+\frac{\partial \tilde{\sigma}_{zz}}{\partial z}=\varepsilon\frac{\partial p}{\partial z}. \tag{8}$$

To complete the description of the problem, the skeleton is assumed to follow isotropic elastic constitutive laws ([24]) with Young's modulus, E, and the Poisson's ratio, ν. An alternative approach ([25]) leading to the identical model is to use the theory of interacting continua as proposed by Green and Naghdi [26] and applied by Ortiz [27].

The relationship between the internal pressure and stress distributions described by Eqs. (6)–(8) can be clarified for the following special case. We consider two-dimensional problems (plane stress or plane strain) with body forces derivable from the gradient of a potential, which is the pressure in our case. The Airy stress function, Φ, can then be invoked by letting ([24]):

$$\tilde{\sigma}_{xx}=\varepsilon p+\frac{\partial^2\Phi}{\partial y^2},\quad \tilde{\sigma}_{xy}=-\frac{\partial^2\Phi}{\partial x\partial y},\quad \tilde{\sigma}_{yy}=\varepsilon p+\frac{\partial^2\Phi}{\partial x^2}. \tag{9}$$

The Airy stress function is then determined from solution to the nonhomogeneous biharmonic equation

$$\frac{\partial^4\Phi}{\partial x^4}+2\frac{\partial^4\Phi}{\partial x^2\partial y^2}+\frac{\partial^4\Phi}{\partial y^4}=D\left(\frac{\partial^2 p}{\partial x^2}+\frac{\partial^2 p}{\partial y^2}\right) \tag{10}$$

where the constant D depends on whether the problem is plane stress or plane strain. The plane-stress and plane-strain results can be generalized to sheet-like geometries (length in one direction is much smaller than the other two) or rod-like geometries (length in one direction is much larger than the other). Although numerical methods are required even to solve for Φ with these geometries, great simplification for two-dimensional problems can be achieved if the internal pressure satisfies the Laplace equation. Under these circumstances, the right-hand side of Eq. (10) becomes zero, and the biharmonic equation is thus homogeneous. If we further assume that no external loads are applied on the ceramic body, $\Phi=0$ is a solution which satisfies both the biharmonic equation and the boundary conditions. By uniqueness, we conclude that $\Phi=0$ everywhere within the body and thus the stresses inside the body are isotropic tensors and are proportional to the internal pressure as given by Eq. (9). This conclusion only holds for the two-dimensional case. In summary, stresses and internal pressures are related through differential equations; only in rare circumstances are they related by simple algebraic relations.

3 Stresses in Rectangular Parallelepipeds

To calculate the stresses in a three-dimensional ceramic body, we first need to know how the internal pressure varies with position. The starting point is to use Darcy's law for flow in porous media ([28]) when a source term is present. Applying conservation of mass then leads to a partial differential equation, the solution ([20]) to which describes the pressure distribution in a parallelepiped of dimensions L_x, L_y, and L_z, as

$$p=P_0\sqrt{1+\sum_{i=1,3,5,\ldots}\ \sum_{j=1,3,5,\ldots}\ \sum_{k=1,3,5,\ldots}A_{ijk}\cos\left(\frac{i\pi x}{L_x}\right)\cos\left(\frac{j\pi y}{L_y}\right)\cos\left(\frac{k\pi z}{L_z}\right)} \tag{11}$$

for $-L_x/2<x<L_x/2$, $-L_y/2<y<L_y/2$, $-L_z/2<z<L_z/2$, where P_0 is the ambient pressure surrounding the porous body, and

$$A_{ijk}=8C\left(\frac{2}{\pi}\right)^5\frac{1}{ijk[i^2+(j/W)^2+(k/H)^2]}(-1)^{\frac{i+j+k-3}{2}}. \tag{12}$$

The source term C in Eq. (12) is dependent on the reaction rate, the length scale of the body and the permeability, κ; W, and H, which are defined in terms of the permeability in different directions, κ_x, κ_y, and κ_z, are the effective dimensionless width and height of the body:

$$W=\sqrt{\frac{\kappa_x}{\kappa_y}}\frac{L_y}{L_x},\quad H=\sqrt{\frac{\kappa_x}{\kappa_z}}\frac{L_z}{L_x}. \tag{13}$$

After determining the pressure distribution from Eq. (11), we compute the three components of the body forces appearing in Eqs. (6)–(8) as

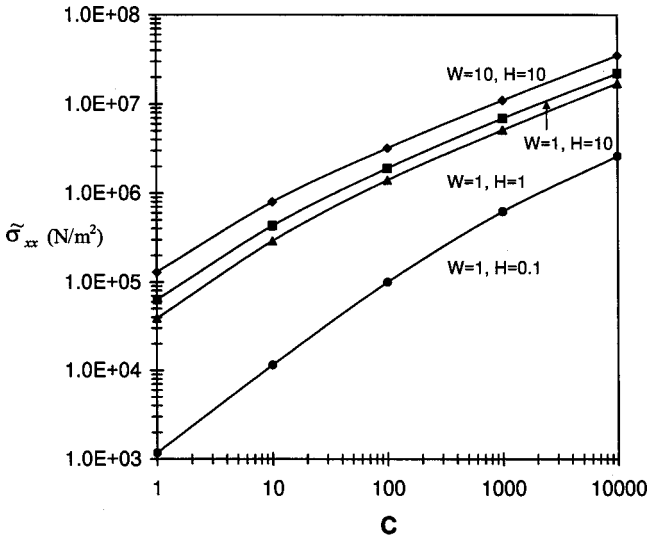


Fig. 2 The dependence of the maximum stress $\tilde{\sigma}_{xx}$ with *C* for parallelipiped bodies of different *W* and *H* with isotropic permeability

$$f_x = -\varepsilon \frac{\partial p}{\partial x}, \quad f_y = -\varepsilon \frac{\partial p}{\partial y}, \quad f_z = -\varepsilon \frac{\partial p}{\partial z}. \tag{14}$$

These body forces are then distributed equally among all nodes in a commercial finite element analysis program, Algor. For "brick" elements in Algor, the body forces of each element are thus equally distributed among the eight nodes in each element. For stress calculations in a rectangular parallelepiped with overall dimensions 2 cm×4 cm×1 cm, we use 125 "brick" elements to represent one-eighth of the body. The product of the void fraction and ambient pressure is taken as $\varepsilon P_0 = 10^5$ N/m^2; because of the linear elasticity assumption, the stresses are proportional to P_0. The void fraction, ε, is therefore left to be arbitrary, and the stresses that are calculated by finite element analysis are thus the nominal stresses which are $(1-\varepsilon)$ times the true stresses on the skeleton. The nominal stresses presented here are due to the pressure gradient alone. For the skeleton material, we use the same material properties as Tsai [15]: $\lambda = 116.8\times10^5$ N/m^2 and $G = 77.9\times10^5$ N/m^2 which corresponds to Poisson's ratio $\nu = 0.3$ and Young's modulus $E = 2\times10^7$ N/m^2.

The stress components $\tilde{\sigma}_{xx}$, $\tilde{\sigma}_{yy}$, $\tilde{\sigma}_{zz}$, and $\tilde{\sigma}_{xy}$ are calculated for a rectangular parallelepiped of overall dimensions of 2 cm×4 cm×1 cm for C=100 with equal permeability ($W=2$, $H=0.5$) in all directions. The maximum pressure in the center of the body is 3.55×10^5 N/m^2 as determined by Eq. (11). Corresponding to this pressure, we find that the maximum normal stresses $\tilde{\sigma}_{xx}$, $\tilde{\sigma}_{yy}$, and $\tilde{\sigma}_{zz}$ are 8.5×10^5 N/m^2, 8×10^5 N/m^2, and 9.9×10^5 N/m^2, respectively, and these maxima all occur at the center of the body as well. The three components of normal stress all decrease monotonically in a nonlinear manner as the edges of the body are approached. The occurrence of the maximum stress in the center of the body agrees with the commonly observed failure mode in that the ceramic parts fracture along surfaces of symmetry corresponding to the normal stresses in the *x* and *y*-directions.

For the case treated above, the maximum shear stresses $\tilde{\sigma}_{xy}$, $\tilde{\sigma}_{yz}$, and $\tilde{\sigma}_{xz}$ are 0.33×10^5 N/m^2, 0.39×10^5 N/m^2, and 0.50×10^5 N/m^2, respectively, and do not occur at the center of the body. We note, in particular, that the maximum shear stresses are at least an order of magnitude smaller than the maximum normal stresses. Calculations on rectangular parallelepipeds with other aspect ratios lead to the same general observation. For this reason, our computational results will be given primarily in terms of the three normal stress components $\tilde{\sigma}_{xx}$, $\tilde{\sigma}_{yy}$, and $\tilde{\sigma}_{zz}$.

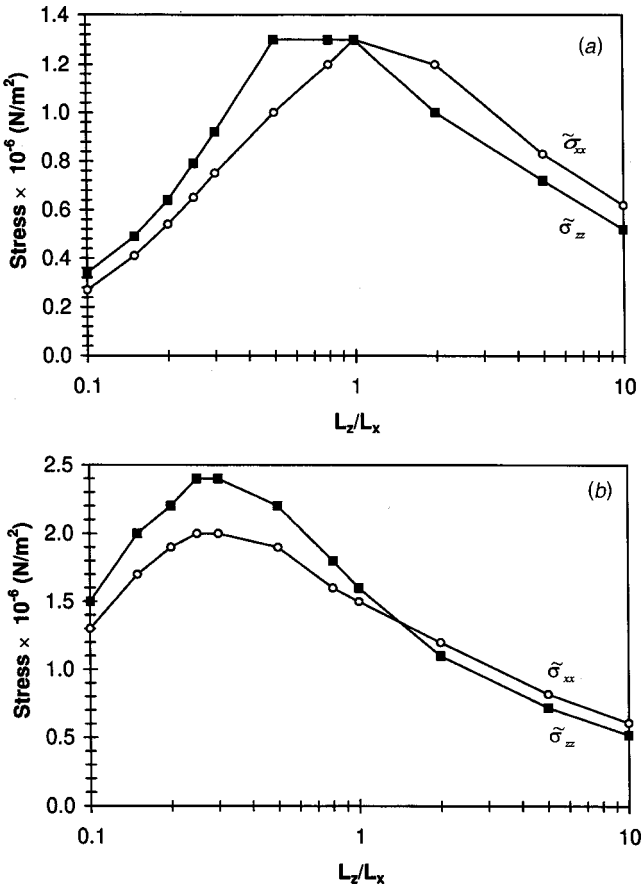


Fig. 3 The dependence of the maximum stresses $\tilde{\sigma}_{xx}$ and $\tilde{\sigma}_{zz}$ with aspect ratio for a square parallepiped body of fixed volume $V=8$ cm^3 with (*a*) equal permeability of $\kappa_x=\kappa_y=\kappa_z=10^{-15}$ m^2 and (*b*) unequal permeability of $\kappa_x=\kappa_y=10^{-15}$, $\kappa_z=10^{-16}$ m^2

Equations (11) and (12) describe the dependence of the pressure distribution on the constant *C*, which is related to the rate of reaction of binder decomposition, the length scale, and the permeability. As a result of Eqs. (6)–(8), the stresses are also dependent on *C*. Figure 2 shows that the maximum normal stress $\tilde{\sigma}_{xx}$ increases with *C* for bodies of isotropic permeability for different values of *W* and *H*. In our previous work [20], we have found that the maximum internal pressure is mainly controlled by the flow of decomposition products across the smallest length of the body; the differences in pressure for the top three cases listed in Fig. 2 are small since the shortest length of these cases is the same. We see from Fig. 2 that this limiting behavior is true for the maximum normal stress $\tilde{\sigma}_{xx}$ as well.

In order to investigate the differences in maximum stress in bodies with anisotropic permeability, we have examined the following two cases: (a) $\kappa_x=\kappa_y=\kappa_z=10^{-15}$ m^2 and (b) $\kappa_x=\kappa_y=10^{-15}$, $\kappa_z=10^{-16}$ m^2. The normal stresses are calculated for bodies of identical volume, $V=8$ cm^3 with $L_x=L_y$, and $L_z=8$ cm$^3/L_x^2$. We also let $C=25$ cm$^{-2}L_x^2$ to correctly account for its dependence on the length scale ([20]). Figures 3(*a*) and 3(*b*) show the maximum normal stresses, and since $\tilde{\sigma}_{xx}=\tilde{\sigma}_{yy}$ for square parallelepipeds, only $\tilde{\sigma}_{xx}$ and $\tilde{\sigma}_{zz}$ are shown. From Fig. 3, we see that in each case, the difference between the two normal stress components is small. In case (a), the maximum normal stress occurs at $L_z/L_x=1$, i.e., the body with cubic geometry. In case (b), the maximum normal stress occurs at L_z/L_x near 0.3. When the different permeability is taken into account in calculating the "effective" aspect ratio *H* by Eq. (12), however, we find

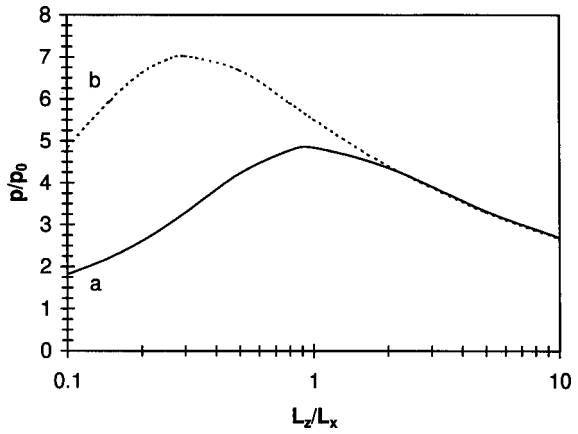


Fig. 4 The dependence of the maximum internal pressure with aspect ratio for a square parallepiped body of fixed volume $V=8\ \text{cm}^3$ with (*a*) equal permeability of $\kappa_x=\kappa_y=\kappa_z=10^{-15}\ \text{m}^2$ and (*b*) unequal permeability of $\kappa_x=\kappa_y=10^{-15}$, $\kappa_z=10^{-16}\ \text{m}^2$

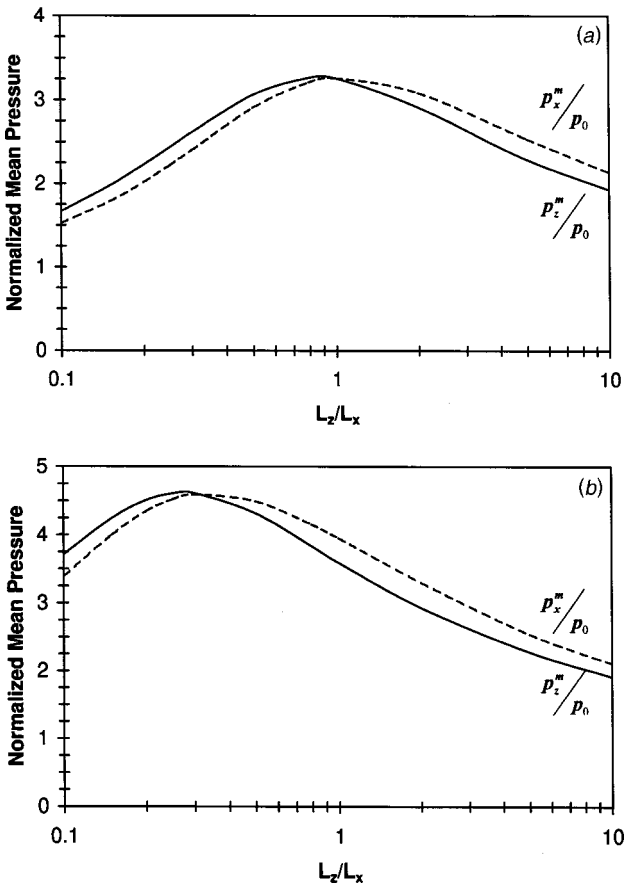


Fig. 5 The dependence of the maximum mean pressure with aspect ratio for a square parallepiped body of fixed volume $V=8\ \text{cm}^3$ with (*a*) equal permeability of $\kappa_x=\kappa_y=\kappa_z=10^{-15}\ \text{m}^2$ and (*b*) unequal permeability of $\kappa_x=\kappa_y=10^{-15}$, $\kappa_z=10^{-16}\ \text{m}^2$

that the maximum normal stress occurs at $H=1$. The fact that the stresses vary with L_z/L_x should not be surprising since we are considering parallelepipeds with fixed volume; bodies with either small or large aspect ratios will have one short length scale by which the pressure is relieved, thereby lowering the stress.

The maximum internal pressure corresponding to cases (a) and (b) are shown in Fig. 4. We see that maximum pressure varies with aspect ratio in manner which is very similar to how the stresses in Fig. 3 vary with aspect ratio and permeability. Note especially that the maximum internal pressure occurs at L_z/L_x near 0.3 ($H=1$) for the case of anisotropic permeability.

The normal stresses shown in Fig. 3(*a*) indicate that $\tilde{\sigma}_{zz}$ is greater than $\tilde{\sigma}_{xx}$ for $L_z<L_x$ and this ordering is reversed for $L_z>L_x$. In other words, the normal stress corresponding to the shorter length is larger than the normal stress corresponding to the longer length. To understand this particular ordering of the normal stresses, we calculate on the three symmetric planes of the body the mean pressures, which are defined as

$$p_x^m=\frac{4}{L_yL_z}\int_0^{L_y/2}\int_0^{L_z/2}p(x=0,y,z)dydz \tag{15}$$

$$p_y^m=\frac{4}{L_xL_z}\int_0^{L_x/2}\int_0^{L_z/2}p(x,y=0,z)dxdz \tag{16}$$

$$p_z^m=\frac{4}{L_xL_y}\int_0^{L_x/2}\int_0^{L_y/2}p(x,y,z=0)dxdy. \tag{17}$$

Figure 5(*a*) corresponds to case (a) of isotropic permeability, where we see that the ordering of the mean pressures is the same as that of the normal stresses. Thus, for bodies of isotropic permeability, the maximum normal stress and mean pressure both act along the shortest direction.

For case (b) of anisotropic permeability, the mean pressures are plotted in Fig. 5(*b*), where we see that the two curves cross each other at L_z/L_x near 0.3 ($H=1$). That is, the maximum mean pressure occurs along the shortest direction if the lengths are adjusted by the permeability following Eq. (13). Comparing Fig. 5(*b*) with Fig. 3(*b*), we find that the same cannot be said about the normal stresses; the two curves in Fig. 3(*b*) cross each other when L_z/L_x is near one. In summary, the maximum normal stress and maximum mean pressure occur where the effective aspect ratio H is unity; the aspect ratio at which the ordering of the stress changes, however, occurs at the true aspect ratio L_z/L_x of unity.

Conclusions

We have developed a mechanics model that allows us to calculate the stress distribution in porous three-dimensional bodies based on a known internal pressure distribution. This model, when combined with finite element analysis, allows us to study the stress distribution in ceramic components during binder burnout. This combined approach thus allows one to design bodies based on consideration of the effect of mechanical stresses on product yield.

For the simple geometry we have considered, the normal stresses, which occur at the center of the body, are an order of magnitude larger than the shear stresses. In general, the normal stresses follow the same trends as the internal pressure. This is especially so for bodies with isotropic permeability. Therefore, instead of carrying out stress calculations, which require numerical solution, the internal pressure, for which an analytical solution exists, can often be used for development of the binder burnout cycle and for product design.

During the binder burnout process, very little is known about how the material properties such as the Young's modulus and Poisson's ratio vary spatially and temporally. In addition, the failure criteria have not been identified. In light of these limitations, the mean pressure on different material surfaces can be used as a convenient compromise between the full finite element stress calculations and the internal pressure calculations. Although the mean pressure is calculated solely from the internal pressure, it captures the effect that geometry plays on the distribution of the internal load within the body.

Acknowledgment

One of the authors has benefited greatly from the discussions with Prof. Rohan Abeyaratne.

References

[1] Lewis, J. A., 1997, ''Binder Removal From Ceramics,'' Annu. Rev. Mater. Sci., **27**, pp. 147–173.
[2] Spronson, D. W., and Messing, G. L., 1988, ''Organic Binder Removal Processes in Closed Pore Powder-Binder Systems,'' *Ceramic Transactions*, Vol. 1 (*Ceramic Powder Science IIa*), G. L. Messing, E. Fuller, and H. Hausner, eds., Am. Ceram. Soc., Westerville, OH, pp. 528–537.
[3] Barone, M. R., and Ulicny, J. C., 1990, ''Liquid-Phase Transport During Removal of Organic Binders in Injection-Molded Ceramics,'' J. Am. Ceram. Soc., **73**, pp. 3323–3333.
[4] Lewis, J. A., and Cima, M. J., 1990, ''Diffusivities of Dialkyl Phthalates in Plasticized Poly(Vinyl Butyral): Impact on Binder Thermolysis,'' J. Am. Ceram. Soc., **73**, pp. 2702–2707.
[5] Calvert, P., and Cima, M., 1989, ''Theoretical Models for Binder Burnout,'' J. Am. Ceram. Soc., **73**, pp. 575–579.
[6] Cima, M. J., Lewis, J. A., and Devoe, A. D., 1989, ''Binder Distribution in Ceramic Greenware During Thermolysis,'' J. Am. Ceram. Soc., **72**, pp. 1192–1199.
[7] Evans, J. R. G., Edirisinghe, M. J., Wright, J. K., and Crank, J., 1991, ''On the Removal of Organic Vehicle From Moulded Ceramic Bodies,'' Proc. R. Soc. London, Ser. A, **A432**, pp. 321–340.
[8] Matar, S. A., Edirisinghe, M. J., Evans, J. R. G., and Twizell, E. H., 1993, ''Effect of Porosity Development on the Removal of Organic Vehicle from Ceramic or Metal Moldings,'' J. Mater. Res., **8**, pp. 617–625.
[9] Matar, S. A., Edirisinghe, M. J., Evans, J. R. G., Twizell, E. H., and Song, J. H., 1995, ''Modeling the Removal of Organic Vehicle from Ceramic or Metal Moldings: The Effect of Gas Permeation on the Incidence of Defects,'' J. Mater. Sci., **30**, pp. 3805–3810.
[10] Song, J. H., Edirisinghe, M. J., Evans, J. R. G., and Twizell, E. H., 1996, ''Modeling the Effect of Gas Transport on the Formation of Defects During Thermolysis of Powder Moldings,'' J. Mater. Res., **11**, pp. 830–840.
[11] Matar, S. A., Edirisinghe, M. J., Evans, J. R. G., and Twizell, E. H., 1996, ''Diffusion of Degradation Products in Ceramic Moldings During Thermal Pyrolysis: Effect of Geometry,'' J. Am. Ceram. Soc., **79**, pp. 749–755.
[12] German, R. M., 1987, ''Theory of Thermal Debinding,'' Int. J. Powder Metall., **23**, pp. 237–245.
[13] Lograsso, B. K., and German, R. M., 1990, ''Thermal Debinding of Injection Molded Powder Compacts,'' Powder Metallurgy International, **22**, pp. 17–22.
[14] Stangle, G. Y., and Aksay, I. A., 1990, ''Simultaneous Momentum, Heat and Mass Transfer With Chemical Reaction in a Disordered Porous Medium: Application to Binder Removal From A Ceramic Green Body,'' Chem. Eng. Sci., **45**, pp. 1719–1731.
[15] Tsai, D.-S., 1991, ''Pressure Buildup and Internal Stresses During Binder Burnout: Numerical Analysis,'' AIChE J., **37**, pp. 547–554.
[16] Manguin-Fritsch, A., Burlet, H., Fourt P. M., and Abouaf, M., 1992, ''Modélisation de la Pyrolyse d'un Liant Organique de Mise en Forme de Pièces Céramiques Injectées,'' *L'Industrie Céramique & Verrière*, **887**, pp. 744–749.
[17] West, A. C., and Lombardo, S. J., 1998, ''The Role of Thermal and Transport Properties on the Binder Burnout of Injection Molded Ceramic Components,'' Chem. Eng. J., **71**, pp. 243–252.
[18] Shivashankar, T. S., and German, R. M., 1990, ''Effective Length Scale for Predicting Solvent-Debinding Times of Components Produced by Powder Injection Molding,'' J. Am. Ceram. Soc., **82**, pp. 1146–1152.
[19] Liau, L. C.-K., Peters, B., Krueger, D. S., Gordon, A., Viswanath, D. S., and Lombardo, S. J., 2000, ''The Role of Length Scale on Pressure Increase and Yield of Poly(Vinyl-Butyral)-Barium Titanate-Platinum Multi-Layer Ceramic Capacitors During Binder Burnout,'' J. Am. Ceram. Soc., **83**, pp. 2645–2653.
[20] Lombardo, S. J., and Feng, Z. C., 2002, ''A Model for the Pressure Distribution During Binder Burnout in Three-Dimensional Porous Ceramic Bodies With Anisotropic Permeability,'' J. Mater. Res., accepted for publication.
[21] Peters, B., and Lombardo, S. J., 2001, ''Optimization of Multi-Layer Ceramic Capacitor Geometry for Maximum Yield During Binder Burnout,'' J. Mater. Sci., **12**, pp. 403–409.
[22] McTigue, D. F., Wilson, R. K., and Nunziato, J. W., 1983, ''An Effective Stress Principle for Partially Saturated Granular Media,'' *Mechanics of Granular Materials: New Models and Constitutive Relations* J. T. Jenkins, and M. Satake, eds., Elsevier, Amsterdam, pp. 195–210.
[23] Coussy, O., 1995, *Mechanics of Porous Continua*, John Wiley and Sons, New York.
[24] Fung, Y. C., 1965, *Foundations of Solid Mechanics*. Prentice-Hall, Englewood Cliffs, NJ.
[25] suggested by a reviewer.
[26] Green, A. E., and Naghdi, P. M., 1965, ''A Dynamical Theory of Interacting Continua,'' Int. J. Eng. Sci., **3**, pp. 231–241.
[27] Ortiz, M., 1985, ''A Constitutive Theory for the Inelastic Behavior of Concrete,'' Mech. Mater., **4**, pp. 67–93.
[28] Harr, M. E., 1997, *Mechanics of Particulate Media*, McGraw-Hill, New York.

Isoparametric Graded Finite Elements for Nonhomogeneous Isotropic and Orthotropic Materials

Jeong-Ho Kim

G. H. Paulino[1]
Mem. ASME,
e-mail: paulino@uiuc.edu

Department of Civil and Environmental Engineering,
University of Illinois at Urbana-Champaign,
Newmark Laboratory,
205 North Mathews Avenue,
Urbana, IL 61801

Graded finite elements are presented within the framework of a generalized isoparametric formulation. Such elements possess a spatially varying material property field, e.g. Young's modulus (E) and Poisson's ratio (ν) for isotropic materials; and principal Young's moduli (E_{11},E_{22}), in-plane shear modulus (G_{12}), and Poisson's ratio (ν_{12}) for orthotropic materials. To investigate the influence of material property variation, both exponentially and linearly graded materials are considered and compared. Several boundary value problems involving continuously nonhomogeneous isotropic and orthotropic materials are solved, and the performance of graded elements is compared to that of conventional homogeneous elements with reference to analytical solutions. Such solutions are obtained for an orthotropic plate of infinite length and finite width subjected to various loading conditions. The corresponding solutions for an isotropic plate are obtained from those for the orthotropic plate. In general, graded finite elements provide more accurate local stress than conventional homogeneous elements, however, such may not be the case for four-node quadrilateral (Q4) elements. The framework described here can serve as the basis for further investigations such as thermal and dynamic problems in functionally graded materials. [DOI: 10.1115/1.1467094]

1 Introduction

Recent advances in material processing have allowed manufacturing a wide diversity of functionally graded materials (FGMs) ([1–3]). Such materials possess continuously graded properties with gradual change in microstructure ([4,5]). The materials are made to take advantage of desirable features of its constituent phases. For instance, in a thermal protection system, FGMs take advantage of heat and corrosion resistance typical of ceramics, and mechanical strength and toughness typical of metals.

FGMs are nonhomogeneous with regard to thermomechanical and strength related properties. Depending on the processing technique, they may exhibit either isotropic or anisotropic material properties. For instance, large bulk FGMs produced by spark plasma sintering (SPS) technique may be modeled as isotropic materials ([6]). On the other hand, materials processed using plasma spray technique have generally a lamellar structure ([7]), while materials processed by electron beam physical vapor deposition (PVD) may have a columnar structure ([8]). Thus, in studying the mechanics of the former class of materials (fabricated by SPS), a nonhomogeneous isotropic model may be appropriate; and for the latter class of materials (fabricated by plasma spraying or PVD), a nonhomogeneous orthotropic model may suffice as a first approximation. Thus, both types of material models, i.e., isotropic and orthotropic, are investigated here.

As the manufacturing of FGMs advances, new modeling techniques are also developed for such materials ([3,9]). Here, we focus on the finite element method for nonhomogeneous materials using a generalized isoparametric formulation. The graded elements obtained with this formulation are compared with conventional homogeneous elements, as illustrated by Fig. 1. Notice that the graded element incorporates the material property gradient at the size scale of the element, while the homogeneous element produces a stepwise constant approximation to a continuous material property field such as the one shown in Fig. 1.

This paper discusses and compares the behavior of graded versus conventional homogeneous elements under various loading conditions in both isotropic and orthotropic FGMs with respect to analytical solutions which are either available in the literature or derived in this work. The manuscript is organized as follows. The next subsection presents an example which serves as a motivation to this work. In this example, the FGM leads to a stress redistribution with lower stress concentration factor (SCF) than the corresponding problem with homogeneous material. Next, a brief literature survey and comments on previous related work are given. Section 2 presents some exact solutions for displacements and stresses in orthotropic FGMs. The exact solutions for isotropic FGMs are obtained as particular instances of those for orthotropic FGMs. Section 3 reviews finite element formulations. Section 4 addresses the generalized isoparametric graded finite element formulation. Sections 5 and 6 present finite element results for stresses in isotropic and orthotropic FGMs, respectively, which are compared with analytical solutions. Finally, Section 7 provides some concluding remarks.

1.1 Motivation. Functionally graded composites, with smooth variation of volume fractions, offer various advantages such as reduction of residual stress ([10]) and increased bonding strength ([11]). Moreover, if properly used, such materials may also lead to reduction of stress concentration or stress intensity factors ([12]). For example, Hasselman and Youngblood ([13]) found that the maximum tensile thermal stresses in brittle ceramics can be reduced significantly by spatially varying thermal conductivity in a hollow circular cylinder subjected to radially inward or outward steady-state heat flow, and Horgan and Chan ([14]) investigated the effect of material nonhomogeneity on the response of linearly elastic isotropic hollow circular cylinders or

[1]To whom correspondence should be addressed.

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, July 2, 2001; final revision Nov. 14, 2001. Associate Editor: M.-J. Pindera. 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 THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS JOURNAL OF APPLIED MECHANICS.

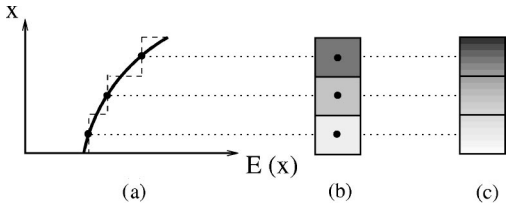


Fig. 1 Homogeneous versus graded finite elements. (*a*) Property variation along one coordinate axis; (*b*) homogeneous elements; (*c*) graded elements. Notice that the property of the homogeneous element corresponds to the property at the centroid of the graded element.

disks under uniform internal or external pressure by varying Young's modulus with respect to the radial direction and found that the maximum hoop stress in a nonhomogeneous material does not, in general, occur on the inner surface in contrast with a homogeneous material.

To further motivate the use of FGMs in engineering applications, consider the isotropic FGM link bar of Fig. 2(*a*). The bar has unit thickness, it is subjected to unit axial tension load at the right end, and it is considered in a state of generalized plane stress. The basic FGM constituents are titanium monoboride (TiB) and commercially pure titanium (CP Ti) as illustrated by Fig. 2(*b*). The elastic properties of the base materials are ([15])

$$E_{TiB}=375\ \text{GPa}, \quad \nu_{TiB}=0.14$$

$$E_{Ti}=107\ \text{GPa}, \quad \nu_{Ti}=0.34.$$

The graded region is incorporated with an exponential material variation. Thus Young's modulus and Poisson's ratio are functions of the Cartesian coordinate y (see Fig. 2), i.e.,

$$E(y)=E_{Ti}e^{\beta_E y}, \quad \nu(y)=\nu_{Ti}e^{\beta_\nu y}, \tag{1}$$

respectively, where $1/\beta_E$ and $1/\beta_\nu$ are the length scales of nonhomogeneity which are given by

$$\beta_E=\frac{1}{W}\log(E_{TiB}/E_{Ti}), \quad \beta_\nu=\frac{1}{W}\log(\nu_{TiB}/\nu_{Ti}), \tag{2}$$

respectively, where W is the width of the symmetric model as shown in Fig. 2(*b*). Figure 3(*a*) shows the finite element mesh for the symmetric portion of the link bar with 1000 quadrilateral elements of eight nodes (Q8). These elements are graded finite elements as illustrated by Fig. 1 and explained subsequently in this paper. Figure 3(*b*) shows the σ_{xx} stress contour for the homogeneous link bar (either TiB or Ti) and Fig. 3(*c*) shows the σ_{xx} stress contour for the FGM link bar (TiB/Ti). The main stress values (nodal average) are summarized in Table 1. Notice that the maximum stress location in the FGM bar is different from that in the homogeneous bar—the maximum stress occurs in A' (see Fig. 2(*b*)) for the homogeneous bar, while it occurs in B' (see Fig. 2(*b*)) for the FGM bar. Moreover, the maximum stress is lower in the FGM than in the homogeneous bar. Thus, the FGM leads to stress redistribution with a lower SCF as illustrated by Table 1 and Fig. 3. In summary, this example shows, by means of elastic finite element analysis, that *the stress response of (inhomogeneous) FGMs differ substantially from those of their homogeneous counterparts*.

1.2 Related Numerical Work. Several numerical models have been used to investigate FGMs, including integral equations (e.g., [16–18]), the higher order model (e.g., [19,20]), boundary elements (e.g., [21,22]), and finite elements (e.g., [10,23–30]). This work concentrates on the finite element method for FGMs using the isoparametric concept for graded elements.

A few additional comments about the related work by Santare and Lambros ([30]) are in order. They have also published in this journal a graded finite element model for nonhomogeneous materials. However, their work differs from ours in the sense that they sample the material properties directly at the Gauss points of the element, while we adopt a generalized isoparametric formulation. Although the two methods are different, they are equivalent for fine mesh discretizations ([31]). They investigated the behavior of

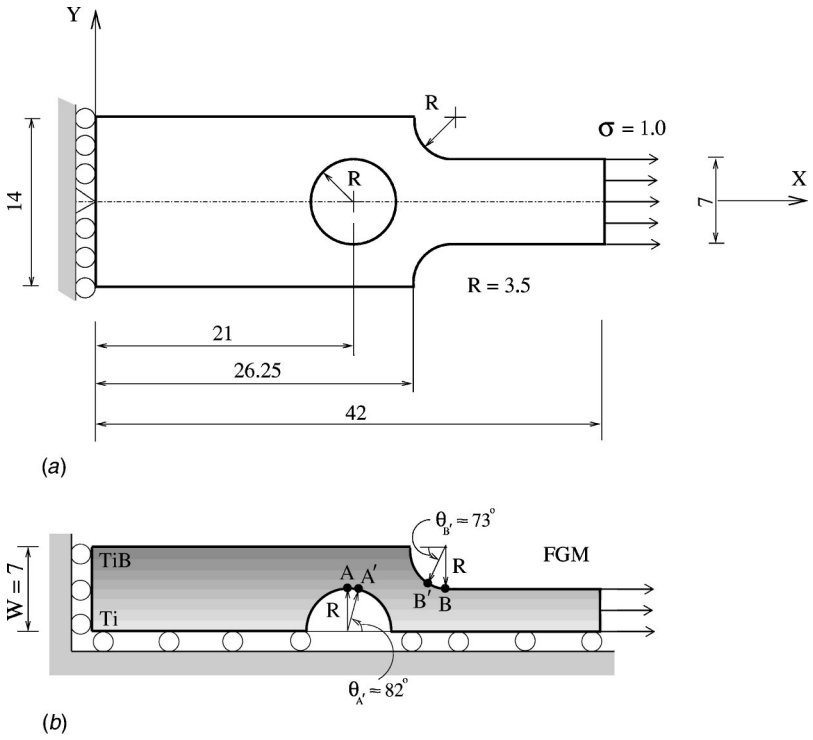


Fig. 2 FGM link bar (units: N, mm): (*a*) geometry and boundary conditions; (*b*) symmetric model

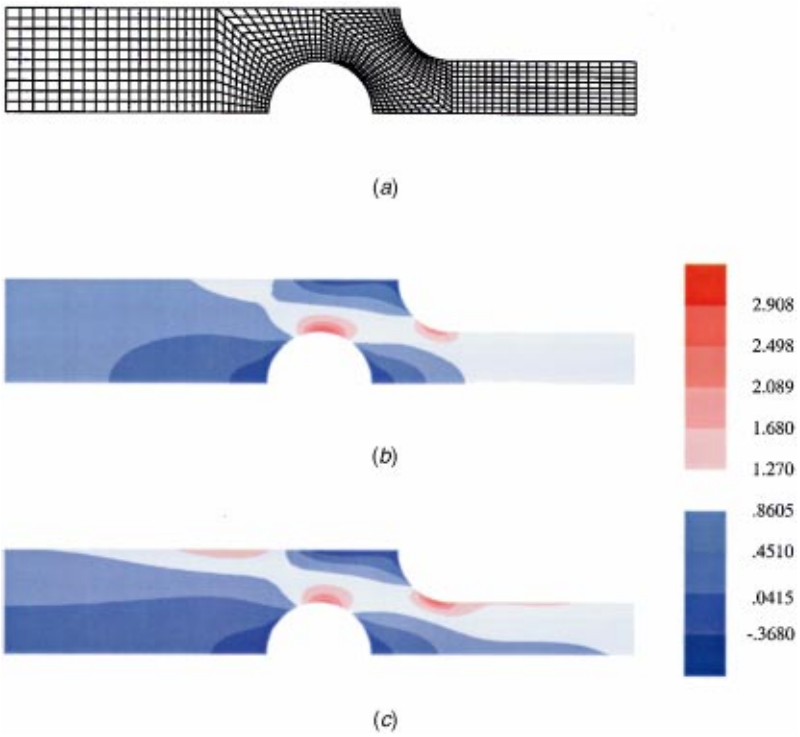


Fig. 3 FGM link bar under unit axial tension (units: N/mm^2): (*a*) mesh configuration with 1000 graded Q8 elements; (*b*) σ_{xx} stress distributions for homogeneous link bar (both TiB and Ti); (*c*) σ_{xx} stress distributions for FGM bar (TiB/Ti)

Table 1 Representative σ_{xx} stress values (N/mm^2) for the link bar of Figs. 2 and 3

	Location	
σ_{xx}	A'	B'
Homogeneous	2.908	2.137
FGM	2.369	2.601

four-node quadrilaterals (Q4) for isotropic FGMs only. In addition to the bilinear element (Q4), we also investigate the behavior of eight-node quadrilaterals (Q8) for both isotropic and orthotropic FGMs. They investigated exponential material variation only. Here, we compare both exponential and linear material variations. Finally, we believe that the generalized isoparametric formulation is more natural to the finite element method than the Gauss point sampling of material properties because the generalized formulation embraces the important isoparametric concept—the same shape functions are used to interpolate the unknown displacements, the geometry, and the material parameters.

2 Some Exact Solutions for Nonhomogeneous Elasticity

Exact solutions for both isotropic and orthotropic functionally graded materials (FGMs) will be used as reference solutions for the numerical examples that follow. We consider an *orthotropic* functionally graded plate of infinite length and finite width subjected to various loading conditions such as remote fixed grip, tension, and bending, as shown in Fig. 4. Both exponential and linear material variations are considered. First, analytical solutions for stresses and displacements are developed for orthotropic FGMs and, afterwards, they are particularized (e.g., in the limit) for isotropic FGMs. The analytical solutions for exponentially graded isotropic FGMs coincide with those of Erdogan and Wu

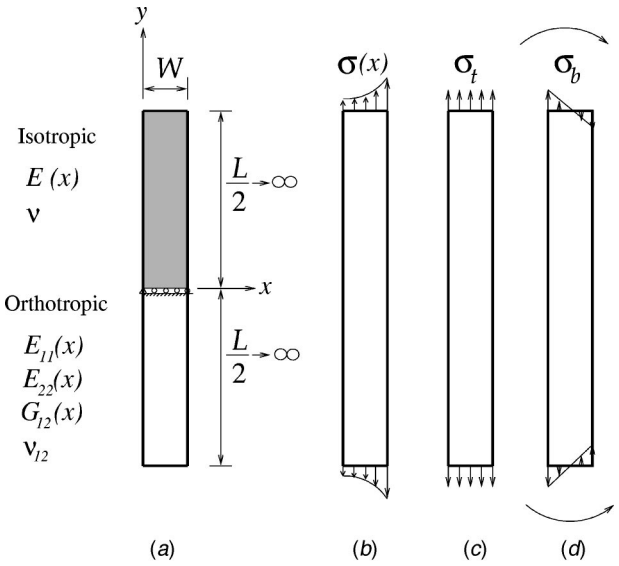


Fig. 4 An isotropic or orthotropic functionally graded plate: (*a*) geometry and material properties—the shaded portion indicates the symmetric region of the plate used in this analysis; (*b*) fixed grip loading with a schematic of the corresponding stresses at the end points of the plate; (*c*) tension loading; (*d*) bending loading

([32]) and Paulino and Kim ([33]). The analytical solutions for linearly graded isotropic and orthotropic FGMs are new solutions derived in this work.

2.1 Exponential Material Variation. Consider a plate under generalized plane stress conditions (see Fig. 4) made of a nonhomogeneous orthotropic material. Assume the Poisson's ratio

(ν_{12}) constant, and the Young's moduli and in-plane shear modulus with variations given by the following expressions:

$$\begin{aligned} E_{11}(x) &= E_{11}^0 e^{\beta_{11}x} \\ E_{22}(x) &= E_{22}^0 e^{\beta_{22}x} \\ G_{12}(x) &= G_{12}^0 e^{\beta_{12}x} \\ \nu_{12}(x) &= \text{constant}, \end{aligned} \tag{3}$$

where $E_{11}^0 = E_{11}(0)$, $E_{22}^0 = E_{22}(0)$, and $G_{12}^0 = G_{12}(0)$ are the material properties at the $x=0$ line (see Fig. 4(a)), and the coefficients β_{ij} above are independent nonhomogeneity parameters characterized by

$$\begin{aligned} \beta_{11} &= \frac{1}{W}\log\left[\frac{E_{11}(W)}{E_{11}(0)}\right] \\ \beta_{22} &= \frac{1}{W}\log\left[\frac{E_{22}(W)}{E_{22}(0)}\right] \\ \beta_{12} &= \frac{1}{W}\log\left[\frac{G_{12}(W)}{G_{12}(0)}\right], \end{aligned} \tag{4}$$

where W is the width of the FGM plate as shown in Fig. 4. Notice that in this case the β_{ij} parameters have units $[Length]^{-1}$.

For a corresponding nonhomogeneous isotropic material ($E = E_{11} = E_{22}$, $G = G_{12}$, $\nu_{12} = \nu$), the Poisson's ratio is assumed constant and the Young's modulus varies exponentially, i.e.,

$$\begin{aligned} E(x) &= E^0 e^{\beta x} \\ \nu(x) &= \text{constant} \end{aligned} \tag{5}$$

where $E^0 = E(0)$. The nonhomogeneity parameter β is given by

$$\beta = \frac{1}{W}\log\left[\frac{E(W)}{E(0)}\right] \tag{6}$$

which has units $[Length]^{-1}$.

2.1.1 Fixed Grip Loading. For fixed grip loading (see Fig. 4(b)) with $\varepsilon_{yy}(x, \pm\infty) = \varepsilon_0$, the stress distribution becomes

$$\sigma_{yy}(x) = E_{22}^0 \varepsilon_0 e^{\beta_{22}x}. \tag{7}$$

Using strain-displacement relations and the boundary conditions

$$u_x(0,0) = 0, \quad u_y(x,0) = 0, \tag{8}$$

one obtains the displacements

$$\begin{aligned} u_x(x,y) &= -\nu_{12}\varepsilon_0 \frac{E_{22}^0}{E_{11}^0}\frac{1}{\beta_{22}-\beta_{11}}\left[e^{(\beta_{22}-\beta_{11})x}-1\right] \\ u_y(x,y) &= \varepsilon_0 y. \end{aligned} \tag{9}$$

Notice that for isotropic materials ($E = E_{11} = E_{22}$, $G_{12} = G$, and $\nu_{12} = \nu$), the stress distribution (7) becomes ([32])

$$\sigma_{yy}(x) = E^0 \varepsilon_0 e^{\beta x} \tag{10}$$

and the displacements are obtained in the limit of Eq. (9) as $(\beta_{22}-\beta_{11}) \to 0$. Thus ([33])

$$\begin{aligned} u_x(x,y) &= -\nu\varepsilon_0 x \\ u_y(x,y) &= \varepsilon_0 y. \end{aligned} \tag{11}$$

2.1.2 Tension and Bending. For tension and bending loads (see Figs. 4(c) and 4(d), respectively), the applied stresses are defined by

$$N = \sigma_t W, \quad M = \frac{\sigma_b W^2}{6}, \tag{12}$$

where N is a membrane resultant along the $x = W/2$ line (see Fig. 4(a)), and M is the bending moment. For these two loading cases, the compatibility condition $\partial^2 \varepsilon_{yy}/\partial x^2 = 0$ gives $\varepsilon_{yy} = Ax + B$ and thus

$$\sigma_{yy}(x) = E_{22}^0 e^{\beta_{22}x}(Ax + B) \tag{13}$$

where the constants A (with unit $[Length]^{-1}$) and B (dimensionless) are determined from

$$\int_0^W \sigma_{yy}(x)\,dx = N, \quad \int_0^W \sigma_{yy}(x)\,x\,dx = M \tag{14}$$

by assuming

$$\begin{aligned} M &= NW/2 \quad \text{for tension} \\ N &= 0 \quad \text{for bending.} \end{aligned} \tag{15}$$

Thus, for tension load, the stress distribution is given by Eq. (13) with

$$\begin{aligned} A &= \frac{\beta_{22}N}{2E_{22}^0}\left(\frac{W\beta_{22}^2 e^{\beta_{22}W} - 2\beta_{22}e^{\beta_{22}W} + W\beta_{22}^2 + 2\beta_{22}}{e^{\beta_{22}W}\beta_{22}^2W^2 - e^{2\beta_{22}W} + 2e^{\beta_{22}W} - 1}\right), \\ B &= \frac{\beta_{22}N}{2E_{22}^0}\left(\frac{e^{\beta_{22}W}[e^{\beta_{22}W}(-W^2\beta_{22}^2 + 3\beta_{22}W - 4) + W^2\beta_{22}^2 - 2\beta_{22}W + 8] - \beta_{22}W - 4}{(e^{\beta_{22}W}-1)(e^{\beta_{22}W}\beta_{22}^2W^2 - e^{2\beta_{22}W} + 2e^{\beta_{22}W} - 1)}\right). \end{aligned} \tag{16}$$

For bending load, the stress distribution is also given by Eq. (13), however, the coefficients A and B for this case are

$$\begin{aligned} A &= \frac{\beta_{22}^2 M}{E_{22}^0}\left(\frac{\beta_{22}(1-e^{\beta_{22}W})}{e^{\beta_{22}W}\beta_{22}^2W^2 - e^{2\beta_{22}W} + 2e^{\beta_{22}W} - 1}\right), \\ B &= \frac{\beta_{22}^2 M}{E_{22}^0}\left(\frac{\beta_{22}We^{\beta_{22}W} - e^{\beta_{22}W} + 1}{e^{\beta_{22}W}\beta_{22}^2W^2 - e^{2\beta_{22}W} + 2e^{\beta_{22}W} - 1}\right), \end{aligned} \tag{17}$$

respectively. For both tension and bending loads, using the strain-displacement relations and the boundary conditions (8), one obtains the displacements

$$\begin{aligned} u_x(x,y) &= -\nu_{12}\frac{E_{22}^0}{E_{11}^0}\left\{\frac{\left(Ax - \dfrac{A}{\beta_{22}-\beta_{11}} + B\right)}{\beta_{22}-\beta_{11}}e^{(\beta_{22}-\beta_{11})x}\right. \\ &\quad \left. + \frac{A - B(\beta_{22}-\beta_{11})}{(\beta_{22}-\beta_{11})^2}\right\} - \frac{A}{2}y^2 \\ u_y(x,y) &= (Ax + B)y. \end{aligned} \tag{18}$$

The constants A and B refer to the appropriate loading case above, either tension (Eq. (16)) or bending (Eq. (17)).

For the isotropic case ($E=E_{11}=E_{22}$, $G_{12}=G$, and $\nu_{12}=\nu$), the stress distribution is obtained by Eqs. (16) and (17) (for tension and bending loads, respectively) with β_{22} replaced by β, which agree with Erdogan and Wu's [32] solution. The displacements are obtained in the limit of Eq. (18) as $(\beta_{22}-\beta_{11})\to 0$. Thus ([33])

$$u_x(x,y)=\nu\left(\frac{A}{2}x^2+Bx\right)-\frac{A}{2}y^2$$
$$u_y(x,y)=(Ax+B)y. \tag{19}$$

2.2 Linear Material Variation. Once again, consider a plate under generalized plane stress conditions, as illustrated by Fig. 4. Assume the Poisson's ratio (ν_{12}) is constant, and the Young's moduli and in-plane shear modulus with variations given by the following expressions (cf. Eq. (3)):

$$E_{11}(x)=E_{11}^0+\gamma_{11}x$$
$$E_{22}(x)=E_{22}^0+\gamma_{22}x$$
$$G_{12}(x)=G_{12}^0+\gamma_{12}x \tag{20}$$
$$\nu_{12}(x)=\text{constant},$$

where $E_{11}^0=E_{11}(0)$, $E_{22}^0=E_{22}(0)$, and $G_{12}^0=G_{12}(0)$ are the material properties at the $x=0$ line (see Fig. 4(a)) and the coefficients γ_{ij} are independent nonhomogeneity parameters characterized by

$$\gamma_{11}=\frac{E_{11}(W)-E_{11}(0)}{W}$$
$$\gamma_{22}=\frac{E_{22}(W)-E_{22}(0)}{W} \tag{21}$$
$$\gamma_{12}=\frac{G_{12}(W)-G_{12}(0)}{W}.$$

Notice that in this case the γ_{ij} parameters have units $[Force]/[Length]^3$.

For a corresponding nonhomogeneous isotropic material ($E=E_{11}=E_{22}$, $G=G_{12}$, $\nu_{12}=\nu$), the Poisson's ratio is assumed constant and the Young's modulus varies linearly, i.e.,

$$E(x)=E^0+\gamma x$$
$$\nu(x)=\text{constant} \tag{22}$$

where $E^0=E(0)$. The nonhomogeneity parameter γ is given by

$$\gamma=\frac{E(W)-E(0)}{W} \tag{23}$$

which has units $[Force]/[Length]^3$.

2.2.1 Fixed Grip Loading. For fixed grip loading (see Fig. 4(b)) with $\varepsilon_{yy}(x,\pm\infty)=\varepsilon_0$, the stress distribution becomes

$$\sigma_{yy}(x)=\varepsilon_0(E_{22}^0+\gamma_{22}x). \tag{24}$$

Using strain-displacement relations and the boundary conditions.

$$u_x(0,0)=0,\quad u_y(x,0)=0,$$

one obtains the displacements

$$u_x(x,y)=-\nu_{12}\varepsilon_0\left\{\frac{\gamma_{22}}{\gamma_{11}}x+\frac{E_{22}^0\ln(E_{11}^0+\gamma_{22}x)}{\gamma_{11}}\right.$$
$$-\frac{\gamma_{22}E_{11}^0\ln(E_{11}^0+\gamma_{11}x)}{\gamma_{11}^2}$$
$$\left.-\left(\frac{E_{22}^0}{\gamma_{11}}-\frac{\gamma_{22}E_{11}^0}{\gamma_{11}^2}\right)\ln(E_{11}^0)\right\}$$
$$u_y(x,y)=\varepsilon_0 y. \tag{25}$$

For isotropic materials ($E=E_{11}=E_{22}$, $G_{12}=G$, and $\nu_{12}=\nu$), the stress distribution (24) becomes

$$\sigma_{yy}(x)=\varepsilon_0(E^0+\gamma x), \tag{26}$$

and the displacements are obtained from Eq. (25) as

$$u_x(x,y)=-\nu\varepsilon_0 x$$
$$u_y(x,y)=\varepsilon_0 y. \tag{27}$$

2.2.2 Tension and Bending. For tension and bending loads (see Fig. 4(c) and 4(d), respectively), the applied stresses are defined by Eq. (12), i.e.,

$$N=\sigma_t W,\quad M=\frac{\sigma_b W^2}{6},$$

where N is a membrane resultant applied along the $x=W/2$ line (see Fig. 4(a)), and M is the bending moment. For these two loading cases, the compatibility condition $\partial^2\varepsilon_{yy}/\partial x^2=0$ gives $\varepsilon_{yy}=Ax+B$ and thus

$$\sigma_{yy}(x)=(E_{22}^0+\gamma_{22}x)(Ax+B), \tag{28}$$

where the constants A (with unit $[Length]^{-1}$) and B (dimensionless) are determined from Eq. (14), i.e.,

$$\int_0^W\sigma_{yy}(x)dx=N,\quad \int_0^W\sigma_{yy}(x)x\,dx=M,$$

by assuming (see Eq. (15))

$$M=NW/2\quad \text{for tension}$$
$$N=0\quad \text{for bending}. \tag{29}$$

Thus, for tension load, the stress distribution is given by Eq. (28) with

$$A=\frac{-\gamma_{22}N}{\frac{1}{6}\gamma_{22}^2W^3+\gamma_{22}E_{22}^0W^2+(E_{22}^0)^2W},$$
$$B=\frac{N(E_{22}^0+\gamma_{22}W)}{\frac{1}{6}\gamma_{22}^2W^3+\gamma_{22}E_{22}^0W^2+(E_{22}^0)^2W}. \tag{30}$$

For bending load, the stress distribution is also given by Eq. (28) with

$$A=\frac{-36M(2E_{22}^0+\gamma_{22}W)}{\gamma_{22}^2W^5+6E_{22}^0\gamma_{22}W^4+6(E_{22}^0)^2W^3},$$
$$B=\frac{36M(2E_{22}^0+\gamma_{22}W)}{\gamma_{22}^2W^5+6E_{22}^0\gamma_{22}W^4+6(E_{22}^0)^2W^3}\frac{3\gamma_{22}W^2+3E_{22}^0W}{2\gamma_{22}W+6E_{22}^0}. \tag{31}$$

For both tension and bending loads, using the strain-displacement relations and the boundary conditions.

$$u_x(0,0)=0,\ u_y(x,0)=0,$$

one obtains the displacements in closed form, which are given by

$$u_x(x,y) = -\nu_{12}\left\{ -\left(\frac{E_{11}^0 \gamma_{22} A}{\gamma_{11}^2} - \frac{\gamma_{22} B}{\gamma_{11}} - \frac{E_{22}^0 A}{\gamma_{11}} \right) x + \frac{\gamma_{22} A}{2\gamma_{11}} x^2 + \left(\frac{E_{22}^0 B}{\gamma_{11}} - \frac{E_{11}^0 \gamma_{22} B}{\gamma_{11}^2} - \frac{E_{11}^0 E_{22}^0 A}{\gamma_{11}^2} + \frac{(E_{11}^0)^2 \gamma_{22} A}{\gamma_{11}^3} \right) \ln(\gamma_{11} x + E_{11}^0) \right.$$
$$\left. - \left(\frac{E_{22}^0 B \gamma_{11}^2 - E_{11}^0 \gamma_{11} \gamma_{22} B - E_{11}^0 E_{22}^0 \gamma_{11} A + (E_{11}^0)^2 \gamma_{22} A}{\gamma_{11}^3} \right) \ln(E_{11}^0) \right\} - \frac{A}{2} y^2,$$
$$u_y(x,y) = (Ax + B) y. \tag{32}$$

For the isotropic case ($E = E_{11} = E_{22}$, $G_{12} = G$, and $\nu_{12} = \nu$), the stress distribution is obtained by Eqs. (30) and (31) (for tension and bending loads, respectively) with γ_{22} and E_{22}^0, replaced by γ and E^0, respectively. The displacements are obtained from Eq. (32) as

$$u_x(x,y) = \nu\left(\frac{A}{2} x^2 + Bx \right) - \frac{A}{2} y^2$$
$$u_y(x,y) = (Ax + B) y. \tag{33}$$

Notice that the form of the exact solutions for displacements in orthotropic FGMs differs significantly from that for isotropic FGMs because the former case depend on two principal Young's moduli, while in the latter case the explicit moduli dependence is absent.

3 Basic Finite Element Formulation

Displacements for an isoparametric finite element can be written as

$$\boldsymbol{u}^e = \sum_{i=1}^{m} N_i \boldsymbol{u}_i^e \tag{34}$$

where N_i are shape functions, $\boldsymbol{u}_i$ is the nodal displacement corresponding to node i, and m is the number of nodal points in the element. For example, for a Q4 element, the standard shape functions are

$$N_i = (1 + \xi\xi_i)(1 + \eta\eta_i)/4, \quad i = 1, \ldots, 4 \tag{35}$$

where (ξ, η) denote intrinsic coordinates in the interval $[-1,1]$ and (ξ_i, η_i) denote the local coordinates of node i. As usual, strains are obtained from displacements by differentiation as.

$$\boldsymbol{\varepsilon}^e = \boldsymbol{B}^e \boldsymbol{u}^e \tag{36}$$

where $\boldsymbol{B}^e$ is the strain-displacement matrix of shape function derivatives, and $\boldsymbol{u}^e$ is the nodal displacement vector. Thus strain-stress relations are given by

$$\boldsymbol{\sigma}^e = \boldsymbol{D}^e(\boldsymbol{x})\, \boldsymbol{\varepsilon}^e \tag{37}$$

where $\boldsymbol{D}^e(\boldsymbol{x})$ is the constitutive matrix, which is a function of position for nonhomogeneous materials, i.e., $\boldsymbol{D}^e(\boldsymbol{x}) = \boldsymbol{D}^e(x,y)$. The principle of virtual work (PVW) yields the following finite element stiffness equations ([34])

$$\boldsymbol{k}^e \boldsymbol{u}^e = \boldsymbol{F}^e \tag{38}$$

where $\boldsymbol{F}^e$ is the load vector and the element stiffness matrix is

$$\boldsymbol{k}^e = \int_{\Omega_e} \boldsymbol{B}^{e^T} \boldsymbol{D}^e(\boldsymbol{x}) \boldsymbol{B}^e d\Omega_e \tag{39}$$

in which Ω_e is the domain of element (e), and T denotes transpose. The reasoning above, at the element level, can be readily extended to the whole domain, which leads to a system of algebraic equations for the unknown displacements ([34]).

4 Generalized Isoparametric Graded Finite Elements

For simplicity of notation, the superscript (e), denoting the element, is dropped in this section. Material properties (e.g., at each Gaussian integration point) can be interpolated from the nodal material properties of the element using isoparametric shape functions which are the same for spatial coordinates (x,y):

$$x = \sum_{i=1}^{m} N_i x_i, \quad y = \sum_{i=1}^{m} N_i y_i \tag{40}$$

and displacements (u,v):

$$u = \sum_{i=1}^{m} N_i u_i, \quad v = \sum_{i=1}^{m} N_i v_i. \tag{41}$$

Thus, by generalization of the isoparametric concept, the Young's modulus $E = E(\boldsymbol{x})$ and Poisson's ratio $\nu = \nu(\boldsymbol{x})$ are interpolated as

$$E = \sum_{i=1}^{m} N_i E_i, \quad \nu = \sum_{i=1}^{m} N_i \nu_i \tag{42}$$

respectively, as illustrated by Fig. 5. Similar expansions can also be made to two-dimensional orthotropic materials where the four independent engineering elastic parameters are the principal Young's moduli, $E_{11} \equiv E_{11}(\boldsymbol{x})$, $E_{22} \equiv E_{22}(\boldsymbol{x})$, in-plane shear modulus $G_{12} \equiv G_{12}(\boldsymbol{x})$; and Poisson's ratio $\nu_{12} = \nu_{12}(\boldsymbol{x})$, i.e.,

$$E_{11} = \sum_{i=1}^{m} N_i (E_{11})_i, \quad E_{22} = \sum_{i=1}^{m} N_i (E_{22})_i,$$
$$G_{12} = \sum_{i=1}^{m} N_i (G_{12})_i, \quad \nu_{12} = \sum_{i=1}^{m} N_i (\nu_{12})_i, \tag{43}$$

as illustrated by Fig. 5.

Some material models may be given in terms of the volume fraction (V) of a material phase, "p," e.g., the metal phase in a ceramic/metal FGM ([35]). In this case, the generalized isoparametric formulation consists of approximating V^p by the standard interpolation

$$V^p = \sum_{i=1}^{m} N_i V_i^p \tag{44}$$

where V_i^p $(i = 1,2,\ldots,m)$ are the values of V^p at the nodal points. This approach offers a convenient framework to couple the finite

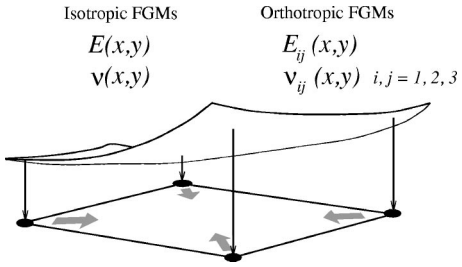


Fig. 5 Generalized isoparametric formulation for isotropic or orthotropic FGMs

element method with micromechanics-based models, e.g., self-consistent scheme.

The above framework allows development of a fully isoparametric formulation in the sense that the same shape functions are used to interpolate the unknown displacements, the geometry, and the material properties. Thus, the actual variation of the material properties may be approximated by the element interpolation functions (e.g., a certain degree of polynomial functions).

5 Numerical Examples

Although the finite element method offers a lot of flexibility in terms of modeling material property variation, the actual choice of properties and boundary value problems in this section was dictated by the analytical solutions derived in Section 2 for the plate configuration of Fig. 4. Here the analytical solutions are compared with the numerical ones. The examples are divided into two groups:

1. isotropic FGM plate
2. orthotropic FGM plate

For each group, two material variations along the Cartesian direction x are examined:

1. exponentially graded materials
2. linearly graded materials

and also the following loading conditions are considered:

1. fixed grip
2. tension loading
3. bending loading

The relevant stress values obtained numerically by the finite element method are compared with the analytical results. For fixed grip loading (see Fig. 4(b)), the stress σ_{yy} is considered. For tension applied parallel to the material gradation, the stress σ_{xx} is the quantity of interest, while for tension and bending loads applied perpendicular to the material gradation (see Figs. 4(c) and 4(d), respectively), the stress σ_{yy} is the relevant quantity. Moreover, for a few of the examples, the displacements computed numerically are also compared with the analytical results.

The finite element meshes consist of square elements (Q4 or Q8) with edges of unit length. For all the examples, 2×2 Gauss quadrature was employed. All the numerical stress values reported here are nodal values extrapolated directly from the Gauss points and without any averaging. The finite element program developed in this work was implemented by the authors in a simple code using MATLAB.

5.1 Isotropic Functionally Graded Plate. Figure 6 illustrates an isotropic FGM plate with material variation in the Cartesian direction x subjected to various loading conditions. Figure 6(a) shows the basic geometry, boundary conditions and properties. The finite element mesh consists of 9×9 Q4 or Q8 elements (either graded or homogeneous) as illustrated in Figs. 6(b) to 6(d). The Young's modulus varies from

$$E_1=E^0=E(0) \quad \text{to} \quad E_2=E(W) \tag{45}$$

either exponentially as given by Eq. (5) or linearly as given by Eq. (22) with $E_1=1.0$ and $E_2=8.0$. The independent nonhomogeneity parameters are given by Eqs. (6) and (23) for the exponential and linear material variations, respectively, with

$$\beta=(\ln(8/1))/9 \quad \text{and} \quad \gamma=7/9. \tag{46}$$

Consistent units are employed here. The Poisson's ratio is constant and it is selected as follows:

$\nu=0.3$ for tension and bending applied perpendicular to material gradation (Figs. 6(b) and 6(c), respectively)

$\nu=0.0$ for tension load parallel to material gradation (Fig. 6(d)).

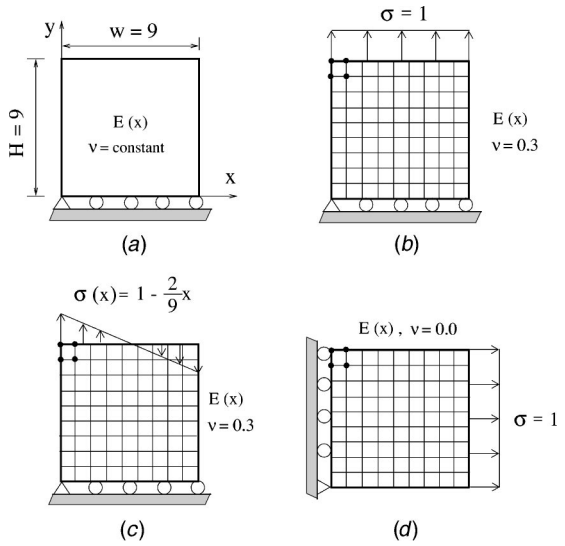


Fig. 6 Isotropic FGM plate with material variation in the x-direction: (*a*) geometry, boundary conditions and material properties; (*b*) tension load perpendicular to material gradation; (*c*) bending load; (*d*) tension load parallel to material gradation. The finite element mesh (9×9 quads: either Q4 or Q8) is illustrated in parts (*b*) through (*d*) with a representative Q4 element at the upper left hand corner

The behavior of the elements (homogeneous versus graded) is as follows. Figure 7 shows the stress σ_{yy} versus x for an exponentially graded isotropic plate subjected to a uniform displacement in the y direction with $\varepsilon_0=\Delta/H$. According to Eq. (10), the stress σ_{yy} is uniform in the y-direction and thus the graph of Fig. 7 is applicable to the entire range of y coordinates, i.e., $0\leqslant y\leqslant H$ (see Fig. 6(a)). In this case, the solution obtained with graded Q4 elements matches the exact solution. This is expected because the exact displacement field is linear (see Eq. (11) and Fig. 8), which is captured by linear isoparametric elements such as Q4. Moreover, because of the linearity of the analytical solution (Eq. (11)), a single Q4 element could be used to predict the exact solution. Figure 7 also shows that the stress obtained with homogeneous Q4 elements is piecewise constant due to the fact that these elements have a single value for each material property, which leads to a piecewise constant material property approximation as illustrated by Fig. 1. Therefore such homogeneous elements predict the actual stress values only at their centroids where

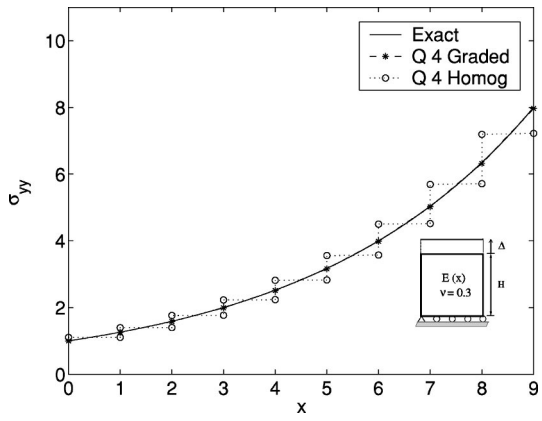


Fig. 7 Stress distribution (σ_{yy}) using Q4 elements for *fixed grip* ($\varepsilon_0=\Delta/H$) load applied perpendicular to the *exponential* material gradation

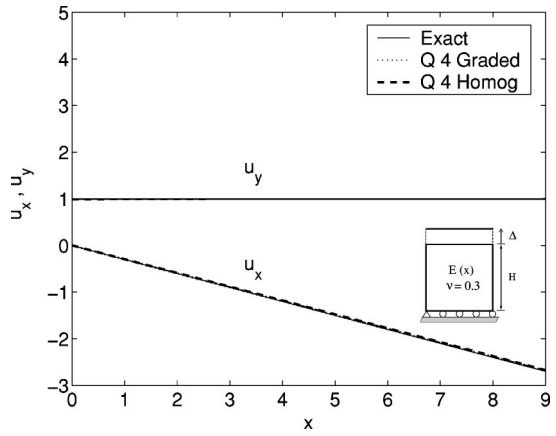


Fig. 8 Displacements (u_x and u_y) using Q4 elements for *fixed grip* load applied perpendicular to the *exponential* material gradation in isotropic FGMs

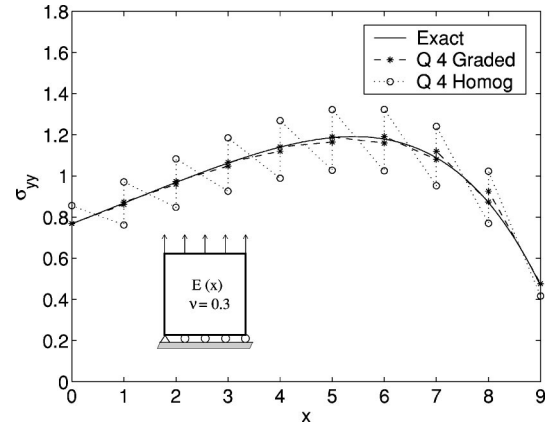


Fig. 9 Stress distribution (σ_{yy}) using Q4 elements for *tension* load applied perpendicular to the *exponential* material gradation

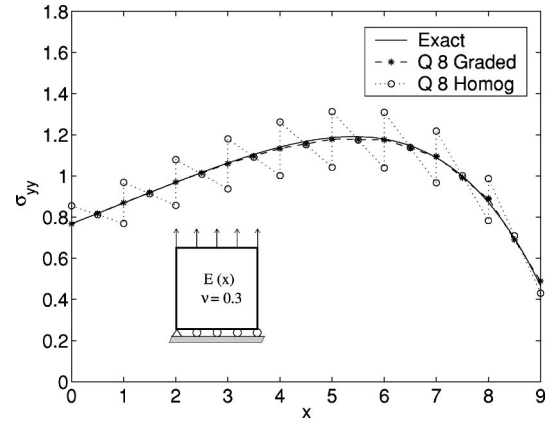


Fig. 10 Stress distribution (σ_{yy}) using Q8 elements for *tension* load applied perpendicular to the *exponential* material gradation

the properties match the material gradation. Moreover, the amplitude of the nodal stress jumps for homogeneous Q4 elements increases with coordinate x in a nearly exponential fashion, as illustrated by Fig. 7. These observations are consistent with those by Santare and Lambros ([30]). Of course, the exact solution is also recovered with higher-order graded elements, e.g., Q8. The homogeneous Q8 elements also lead to a piecewise constant nodal

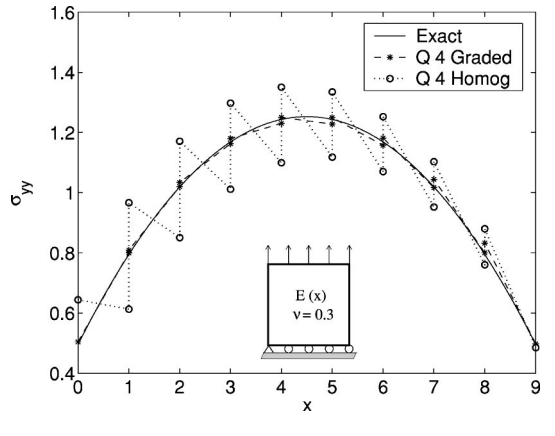


Fig. 11 Stress distribution (σ_{yy}) using Q4 elements for *tension* loading applied perpendicular to the *linear* material gradation

stress profile with the stress at the midnode location along the x direction matching the exact stress value, which occurs because the material properties at the mid-nodes match the actual material properties.

Figures 9 and 10 compare nodal stresses interpolated from stresses at Gauss integration points using graded and homogeneous Q4 and Q8 elements, respectively, which are subjected to tension loading applied perpendicular to the material gradation. Figures 11 and 12 show such comparison considering linear material variation. On the left side of the domain in Figs. 9–12, the exact solution shows an increasing trend of σ_{yy} with x, while the homogeneous elements (either Q4 or Q8) give σ_{yy} as a decreasing function of x in each individual element. Notice that this problem does not occur with the graded elements. In this case, the exact solution for displacements is quadratic (see Eqs. (19) and (33) for exponential and linear material variations, repectively), which coincides with the order of interpolation for the Q8 element. Moreover, the material variation for the linear case is captured by the element shape functions. The stress results for the Q8 element considering exponential and linear material variations are shown in Figs. 10 and 12, respectively. As expected, the homogeneous Q8 element shows piecewise variation while the graded Q8 element approaches the analytical solution quite well. The relatively small differences observed between the analytical and graded Q8 solutions may be attributed to the finite plate length (length/

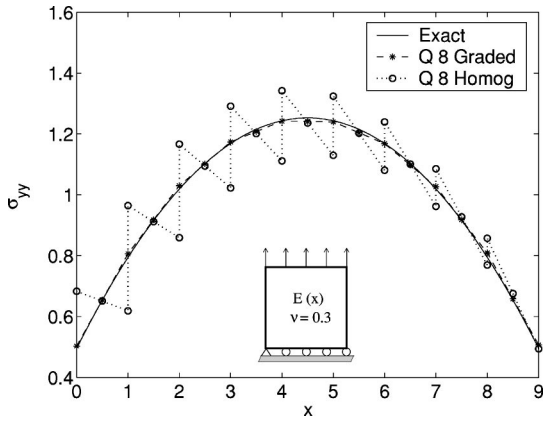


Fig. 12 Stress distribution (σ_{yy}) using Q8 elements for *tension* load applied perpendicular to the *linear* material gradation

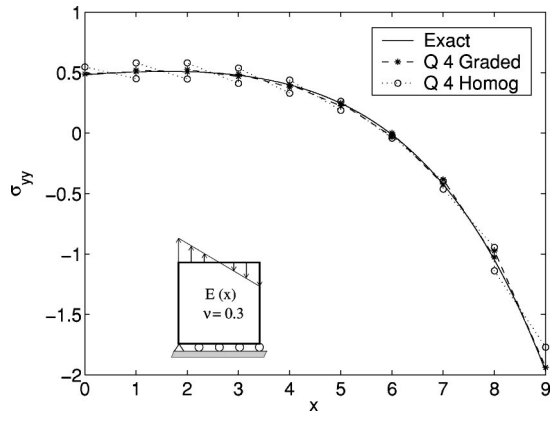


Fig. 13 Stress distribution (σ_{yy}) using Q4 elements for *bending* load applied perpendicular to the *exponential* material gradation

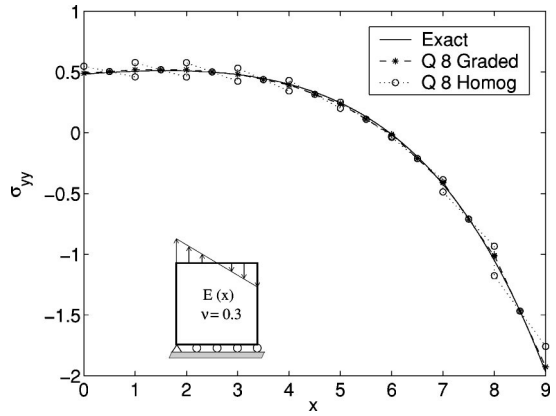


Fig. 14 Stress distribution (σ_{yy}) using Q8 elements for *bending* load applied perpendicular to the *exponential* material gradation

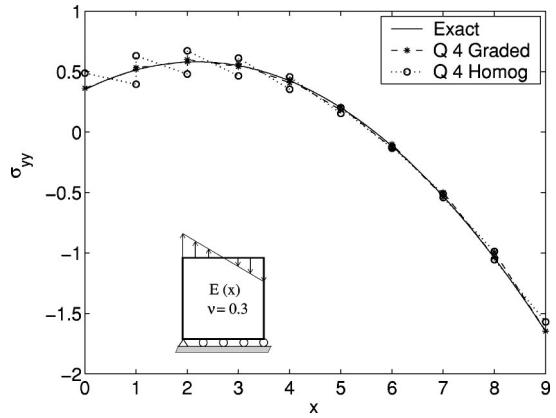


Fig. 15 Stress distribution (σ_{yy}) using Q4 elements for *bending* load applied perpendicular to the *linear* material gradation

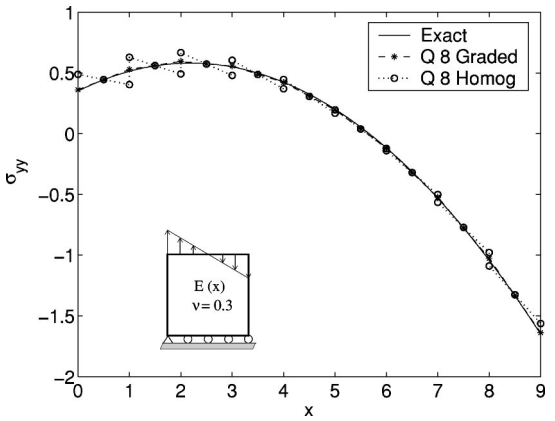


Fig. 16 Stress distribution (σ_{yy}) using Q8 elements for *bending* load applied perpendicular to the *linear* material gradation

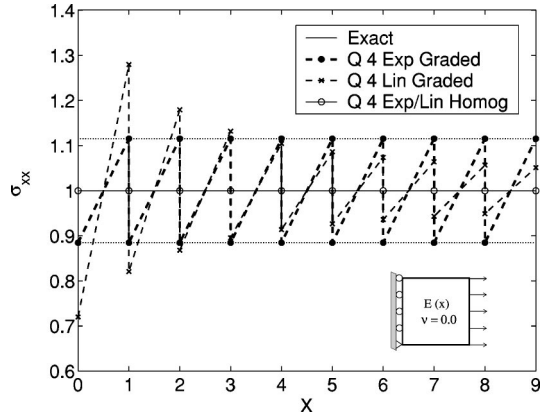


Fig. 17 Stress distribution (σ_{xx}) using Q4 elements (9×9 mesh) for *tension* load applied parallel to the *exponential* or *linear* material gradation

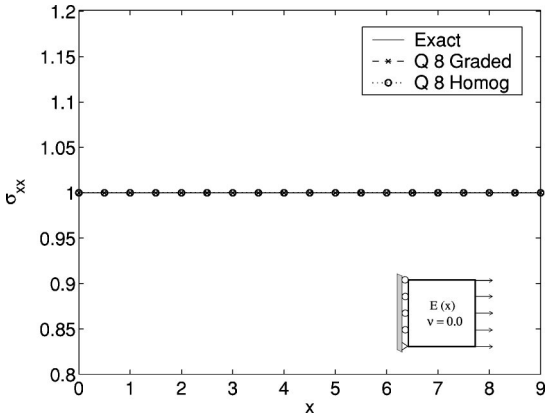


Fig. 18 Stress distribution (σ_{xx}) using Q8 elements (9×9 mesh) for *tension* load applied parallel to the *material* gradation

width=1 as shown in Fig. 6) utilized in the numerical calculation—the analytical solution was derived for an *infinitely long* plate of finite width.

A similar comparison is also made for a different loading case consisting of bending applied perpendicular to the material gradation. Figures 13 and 14 show the behavior of the Q4 and Q8 elements, respectively, for the exponential variation. Figures 15 and 16 show such comparison for the bending case considering linear material variation. The stress results for the Q8 element considering exponential and linear material variations are shown in Figs. 14 and 16, respectively. Similar comments to those made comparing the Q8 (homogeneous versus graded) and analytical solutions for the tension load case also hold for the present bending load case.

The above results lead to the following observations. The variation of stress with position x is larger for linear than with exponential material variations (cf. Figs. 9 and 11, 10 and 12, 13 and 15, and 14 and 16). In general, the amplitude of stress jumps between Q4 elements is larger than between Q8 elements, especially for conventional homogeneous elements (cf. Figs. 9 and 10,

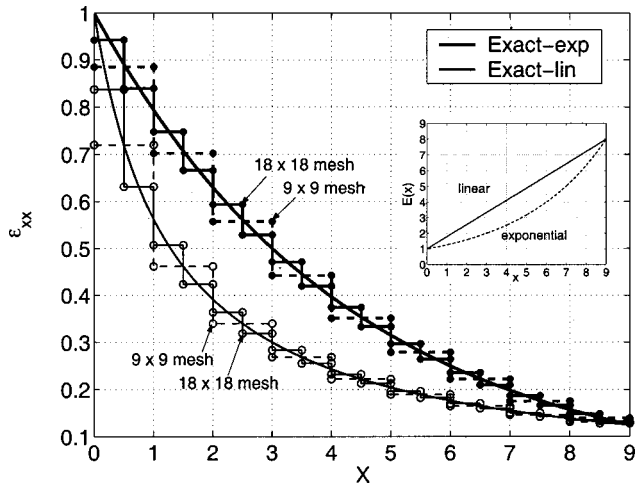


Fig. 19 Strain distribution (ε_{xx}) using Q4 elements (either 9×9 and 18×18) for *tension* load applied parallel to the material gradation (either exponential or linear)

11 and 12, 13 and 14, 15 and 16, and 17 and 18). As expected, the graded elements show superior performance to homogeneous elements, i.e., the graded elements provide a better approximation to the exact solution in every element. Essentially, the graded elements show good performance in terms of actual (i.e., no averaging) nodal stress (σ_{yy}) along the $y=0$ line and the homogeneous elements behave well in terms of the averaged nodal stresses.

Figures 17 and 18 compare nodal stresses of graded versus homogeneous Q4 and Q8 elements (9×9 mesh), respectively, which are subjected to tension applied parallel to the material gradation (see Fig. 6(*d*)). The exact solution is $\sigma_{xx}=1.0$. Different from the observation above, it is interesting to observe in Fig. 17 that the Q4 graded element shows poor performance when compared to Q4 homogeneous elements for both material variations (i.e., exponential and linear). Although mesh refinement (for a fixed material gradient) increases the accuracy of the solution, the same trend of Fig. 17 is observed for a finer mesh, e.g., 18×18. Figure 17 shows that the Q4 graded elements provide piecewise continuous solutions to the nodal stresses (σ_{xx}), while the homogeneous Q4 elements do recover the exact solution. This is the reverse of the effect seen in the previous load cases. However, a higher order element such as Q8 (either graded or homogeneous) is able to capture the exact solution in this case, as shown in Fig. 18.

A few additional remarks, regarding the behavior of Q4 elements observed in Fig. 17, are in order. Both graded and homogeneous elements lead to the same displacements at all nodes and the same constant strains for each element. Notice that along the $y=0$ line, the nodal stress range has constant amplitude for the exponential material case, while it has decreasing amplitude for the linear material case (see Fig. 17). The reason for this behavior is illustrated by Fig. 19 by investigating the strain distribution for two mesh discretizations (9×9 and 18×18 meshes). For instance, for the exponential material case, the nodal strains decrease exponentially while the Young's modulus increase exponentially. Thus the multiplication of these two factors cancel each other to give a constant stress amplitude at the nodal points, as shown in Fig. 17.

5.2 Orthotropic Functionally Graded Plate. Figure 20 shows orthotropic FGM plates, with material variation in the Cartesian direction *x*, subjected to various loading conditions. Figure 20(*a*) shows the basic geometry, boundary conditions and material property variation. The two principal Young's moduli and in-plane shear modulus vary proportionally either with an exponential function of *x* as given by Eq. (3) or with a linear function of *x* as

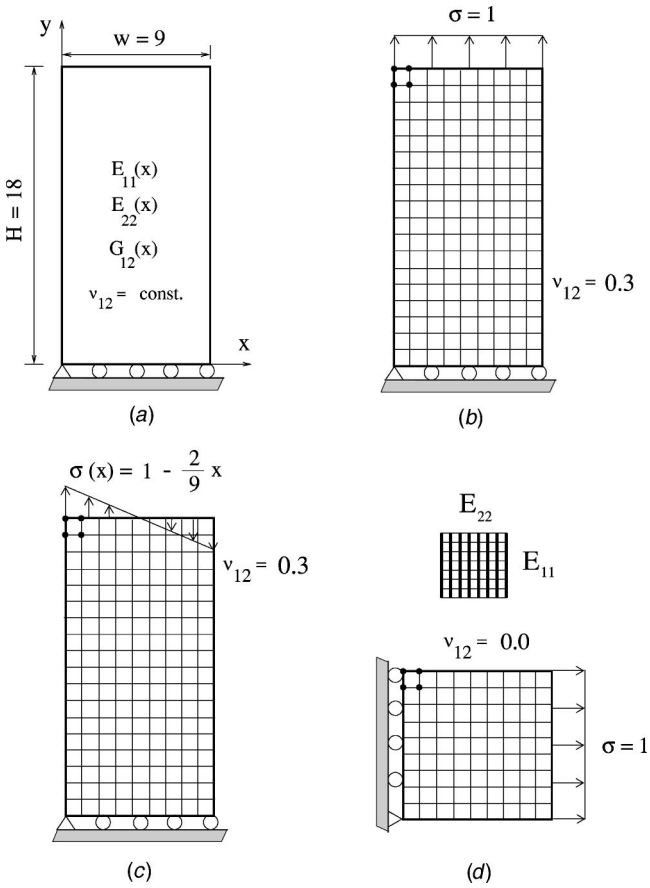


Fig. 20 Orthotropic FGM plate with material variation in the *x*-direction: (*a*) geometry, boundary conditions and material properties; (*b*) tension load perpendicular to material gradation; (*c*) bending load; (*d*) tension load parallel to the material gradation. The finite element mesh (Q4 or Q8 elements) is illustrated in parts (*b*) through (*d*) with a representative Q4 element at the upper left hand corner

given by Eq. (20). The independent nonhomogeneity parameters (β_{ij} and γ_{ij}) are given by Eqs. (4) and (21) for the exponential and linear material variations, respectively. The Poisson's ratio is assumed constant.

For the examples in Fig. 20, the finite element mesh consists of either Q4 or Q8 (graded or homogeneous) elements under generalized plane stress. The mesh for the geometry of Figs. 20(*b*) and 20(*c*) consists of 9×18 elements. For the sake of completeness, all the properties used in the numerical analyses are given as follows. However, due to space limitations, not all the results are shown here, but they are reported elsewhere ([36]). For the fixed grip case and for tension and bending perpendicular to the material gradation, the following data were used for the finite element analysis:

$$E_{11}^0=1, \quad E_{22}^0=0.1, \quad G_{12}^0=0.5, \quad \nu_{12}=0.3$$

in which consistent units are employed. For tension parallel to the material gradation, the following data were used for the finite element analysis:

$$E_{11}^0=1, \quad E_{22}^0=0.1, \quad G_{12}^0=0.5, \quad \nu_{12}=0.0.$$

For the single case of fixed grip loading, only exponential material variation was considered. In this case, the β_{ij} parameters are

$$\beta_{22}=(\ln 8)/9=\beta, \quad \beta_{11}=\beta/2, \quad \beta_{12}=\beta/3$$

so that the range of properties is the following

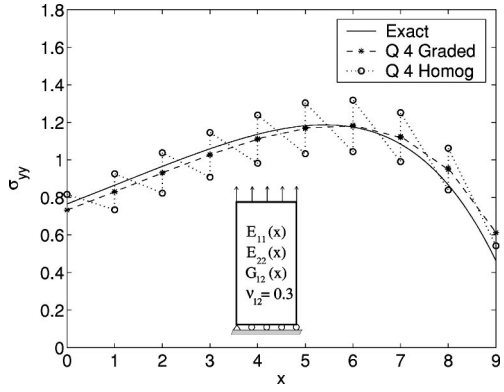


Fig. 21 Stress distribution (σ_{yy}) using Q4 elements for *tension* loading applied perpendicular to the *exponential* material gradation in orthotropic FGMs ($E_{11}^0=1$, $E_{22}^0=0.1$, $G_{12}^0=0.5$, $\nu_{12}=0.3$)

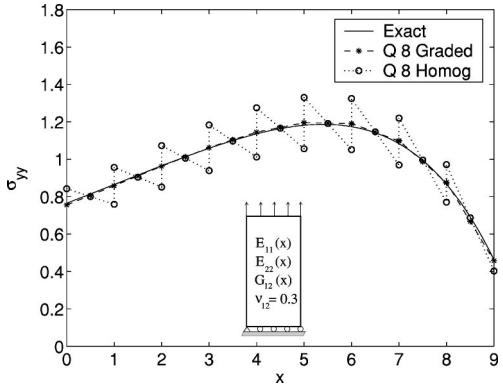


Fig. 22 Stress distribution (σ_{yy}) using Q8 elements for *tension* loading applied perpendicular to the *exponential* material gradation in orthotropic FGMs ($E_{11}^0=1$, $E_{22}^0=0.1$, $G_{12}^0=0.5$, $\nu_{12}=0.3$)

$$E_{11}=[1,2.828], \quad E_{22}=[0.1,0.8], \quad G_{12}=[0.5,1.0].$$

For all other loading cases (i.e., tension and bending perpendicular to the material gradation, and tension parallel to the material gradation), the β_{ij} parameters, characteristic of exponential material variation, are chosen so that the variations of E_{11}, E_{22}, and G_{12} are proportional ([16,17]), i.e.,

$$\beta_{11}=\beta_{22}=\beta_{12}=(\ln 8)/9=\beta, \tag{47}$$

and the γ_{ij} parameters, characteristic of linear material variation, are given by

$$\gamma_{11}=7/9=\gamma, \quad \gamma_{22}=0.7/9, \quad \gamma_{12}=3.5/9$$

so that the range of properties is the following:

$$E_{11}=[1.0,8.0], \quad E_{22}=[0.1,0.8], \quad G_{12}=[0.5,4.0].$$

Regarding the element behavior (homogeneous versus graded), several of the observations made for isotropic materials in the previous section also hold for orthotropic materials. Thus rather than repeating those common observations, this section focuses on new observations and insights. Moreover, the analytical solutions of Section 2 show that, for exponential material gradation (Section 2.1), the relevant stress quantity only depends on the nonhomogeneous parameter β_{22}, and the displacements depend on both β_{11} and β_{22}. For linear material gradation (Section 2.2), the relevant stress depends on γ_{22}, and the displacements depend on both γ_{11} and γ_{22}. This information will be helpful to understand the examples reported below.

For proportional variation of material properties (see Eq. (47)), the change of u_x with x is linear (rather than the nonlinear function of Eq. (9)), which is similar to the behavior of the isotropic plate under the same boundary conditions, i.e., fixed grip (see Fig. 8). This behavior can be seen by the following limit:

$$\lim_{(\beta_{22}-\beta_{11})\to 0} u_x = \lim_{(\beta_{22}-\beta_{11})\to 0} \left\{ -\nu_{12}\varepsilon_0 \frac{E_{22}^0}{E_{11}^0} \frac{1}{\beta_{22}-\beta_{11}} \times [e^{(\beta_{22}-\beta_{11})}-1] \right\}$$

$$= -\nu_{12}\varepsilon_0 \frac{E_{22}^0}{E_{11}^0} x. \tag{48}$$

Figures 21 and 22 compare nodal stresses interpolated from stresses at Gauss points using graded and homogeneous Q4 and Q8 elements, respectively, which are subjected to tension load applied perpendicular to the exponential material gradation. Figure 23 shows a comparison of the displacements (u_x and u_y) computed numerically with those obtained by means of Eq. (18) for all the element types investigated in the present loading case. The curves for u_y indicate that the best elements in terms of matching the analytical solution (Eq. (18)) are Q8 graded, Q8 homogeneous, Q4 graded and Q4 homogeneous, which is somehow expected. Qualitatively, the nodal stress plots considering linear material variation are somewhat similar to those of Figs. 21 and 22 and are not given here ([36]).

A similar comparison is also made for a different loading case consisting of bending applied perpendicular to the material gradation. Figure 24 shows a comparison of the displacements (u_x and u_y) computed numerically with those obtained by Eq. (18) for all the element types investigated in the present loading case. As expected, the Q8 elements capture the analytical solution (Eq. (18)) for u_y better than the Q4 elements. For the sake of brevity, the nodal stress plots are not given here ([36]).

Finally, a few comments regarding the case of tension loading applied parallel to the material gradation in orthotropic FGMs (Fig. 20(*d*)) are in order. Qualitatively, the counterintuitive behavior of homogeneous versus graded Q4 elements is similar to the case involving isotropic nonhomogeneous materials illustrated by Figs. 17 and 18. Thus, for orthotropic case, the Q4 graded element also shows poor performance when compared to the Q4 homogeneous elements for both material variations (i.e., exponential or linear). The reasons for such behavior are given in the last two

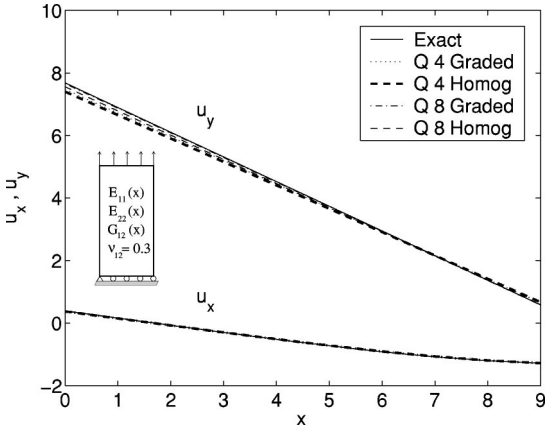


Fig. 23 Displacements (u_x and u_y) along $y=1$ using Q4 and Q8 elements for *tension* load applied perpendicular to the *exponential* material gradation in orthotropic FGMs ($E_{11}^0=1$, $E_{22}^0=0.1$, $G_{12}^0=0.5$, $\nu_{12}=0.3$)

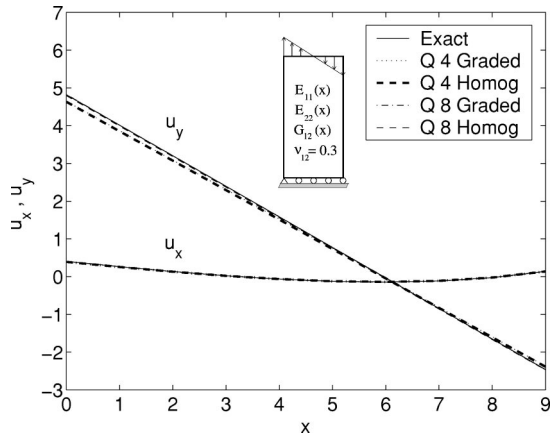


Fig. 24 Displacements (u_x and u_y) along $y=1$ using Q4 and Q8 elements for *bending* load applied perpendicular to the *exponential* material gradation in orthotropic FGMs ($E_{11}^0=1$, $E_{22}^0=0.1$, $G_{12}^0=0.5$, $\nu_{12}=0.3$)

paragraphs of the previous section and will not be repeated here. This is the reverse of the effect seen in the previous load cases for graded orthotropic materials where the graded elements show superior behavior to the corresponding homogeneous elements. Similarly to the isotropic case, a higher-order element such as Q8 (either graded or homogeneous) with 2×2 Gauss quadrature is able to capture the exact solution for this loading case.

5.3 Discussion. This study leads to the following remarks. The isotropic FGM plate (see Fig. 6) has length over width ratio equal to 1 and the orthotropic FGM plate (see Fig. 20) has ratio 2 (for tension and bending loading cases). Because the analytical solution (Section 2) was derived for an infinitely long plate, the higher the aspect ratio (within limits) the better the numerical solution (with respect to the analytical one). For some load cases, e.g., tension and bending perpendicular to the material gradation, the *homogeneous elements* give σ_{yy} as a decreasing function of x in each individual element on the left side of the domain, while the exact solution shows an increasing trend of σ_{yy} with x for this portion of the domain. However, the *graded elements* show the same trend as the exact solution in each element (see, for example, Figs. 11 and 12). The stress plots show that the graded Q8 element gives a smoother stress profile than the graded Q4 element (cf. Figs. 9 and 10). For each loading case, the numerical values of the stress components other than the relevant normal stress quantity should approach zero. Thus the remainder of this paragraph focus on the maximum magnitude of these stress values (obtained numerically) which are theoretically zero. In general, these stress magnitudes are lower with Q8 than with Q4 elements. For tension parallel to the material gradation, the numerical values of the stress σ_{xy} and σ_{yy} are exactly zero for all cases investigated. For the fixed grip case, the largest magnitude of σ_{xy} is $O(10^{-2})$ and occurs for the orthotropic plate with Q4 elements. The largest magnitude of σ_{xx} and σ_{xy} is $O(10^{-3})$ or less for all other analyses for this loading case. For tension and bending perpendicular to the exponential material gradation, the Q4 element leads to spurious shear stresses of $O(10^{0})$ for the orthotropic plate and of $O(10^{-1})$ for the isotropic plate. Smaller magnitudes for the maximum shear stresses are obtained considering linear material gradation. The stress σ_{xx} is of $O(10^{-2})$ or less for all the analyses involving these two loading cases.

6 Concluding Remarks

Graded finite elements, which incorporate the material property gradient at the size scale of the element, have been presented using a *generalized isoparametric formulation*. Both linear (Q4) and quadratic (Q8) quadrilateral elements have been investigated in detail. To address the influence of material property variation, both exponentially and linearly graded elements have been considered and compared. Several plates with continuously nonhomogeneous isotropic and orthotropic materials were considered under fixed grip, tension, and bending conditions. The performance of graded elements was compared to that of conventional elements with respect to analytical solutions.

Higher-order graded elements (e.g., quadratic and higher) are superior to conventional homogeneous elements based on the same shape functions. One should be careful when using graded elements with linear shape functions (e.g., Q4) as it may lose accuracy in certain situations such as uniform traction parallel to the material gradient direction. When using this element, we recommend to average the nodal properties of the element, which would convert it to a regular homogeneous element. Thus the value of material properties at the integration points used to compute stresses depend on whether first-order or higher-order elements are used. This simple procedure leads to a more robust element. A similar procedure is used in the finite element code ABAQUS ([37]) for heat transfer analysis and also in the WARP3D code ([38]).

Acknowledgments

We gratefully acknowledge the support from the National Science Foundation (NSF) under grant No. CMS-0115954 (Mechanics and Materials Program) and from the NASA Ames Research Center (NAG 2-1424) to the University of Illinois at Urbana-Champaign.

References

[1] Hirai, T., 1996, "Functionally Graded Materials," *Materials Science and Technology: Processing of Ceramics, Part 2*, R. J. Brook, ed., VCH Verlagsgesellschaft mbH, Weinheim, Germany, **17B**, pp. 292–341.
[2] Miyamoto, Y., Kaysser, W. A., Rabin, B. H., Kawasaki, A., and Ford, R. G., 1999, *Functionally Graded Materials: Design, Processing, and Applications*, Kluwer, MA.
[3] Suresh, S., and Mortensen, A., 1998, *Fundamentals of Functionally Graded Materials*, IOM Communications, London.
[4] Koizumi, M., 1993, "The concept of FGM," *Proceedings of the Second International Symposium on Functionally Gradient Materials, Ceramic Transactions*, J. B. Holt et al., eds., Westerville, Ohio, The American Ceramic Society, **34**, pp. 3–10.
[5] Pindera, M.-J., Aboudi, J., and Arnold, S. M., 1998, "Thermomechanical Analysis of Functionally Graded Thermal Barrier Coatings With Different Microstructural Scales," J. Am. Ceram. Soc., **81**(6), pp. 1525–1536.
[6] Tokita, M., 1999, "Development of Large-Size Ceramic/Metal Bulk FGM Fabricated by Spark Plasma Sintering," Mater. Sci. Forum, **308–311**, pp. 83–88.
[7] Sampath, S., Hermann, H., Shimoda, N., and Saito, T., 1995, "Thermal Spray Processing of FGMs," M.R.S. Bull., **20**(1), pp. 27–31.
[8] Kaysser, W. A., and Ilschner, B., 1995, "FGM Research Activities in Europe," M.R.S. Bull., **20**(1), pp. 22–26.
[9] Markworth, A. J., Ramesh, K. S., and Parks, Jr., W. P., 1995, "Modelling Studies Applied to Functionally Graded Materials," J. Mater. Sci., **30**, pp. 2183–2193.
[10] Lee, Y. D., and Erdogan, F., 1995, "Residual/Thermal Stresses in FGM and Laminated Thermal Barrier Coatings," Int. J. Fract., **69**, pp. 145–165.
[11] Kurihara, K., Sasaki, K., and Kawarada, M., 1990, "Adhesion Improvement of Diamond Films," *Proceedings of the First International Symposium on Functionally Gradient Materials*, M. Yamanouchi et al., eds., Tokyo, Japan.
[12] Jin, Z.-H., and Paulino, G.H., 2001, "Transient Thermal Stress Analysis of an Edge Crack in a Functionally Graded Material," Int. J. Fract., **107**(1), pp. 73–98.
[13] Hasselman, D. P. H., and Youngblood, G. E., 1978, "Enhanced Thermal Stress Resistance of Structural Ceramics With Thermal Conductivity Gradient," J. Am. Ceram. Soc., **61**(1–2), pp. 49–52.
[14] Horgan, C. O., and Chan, A. M., 1999, "Pressurized Hollow Cylinder or Disk Problem for Functionally Graded Isotropic Linearly Elastic Materials," J. Elast., **55**(1), pp. 43–59.
[15] Carpenter, R. D., Liang, W. W., Paulino, G. H., Gibeling, J. C., and Munir, Z. A., 1999, "Fracture Testing and Analysis of a Layered Functionally Graded Ti/TiB Beam in 3-Point Bending," Mater. Sci. Forum, **308–311**, pp. 837–842.
[16] Ozturk, M., and Erdogan, F., 1997, "Mode I Crack Problem in an Inhomogeneous Orthotropic Medium," Int. J. Eng. Sci., **35**(9), pp. 869–883.
[17] Ozturk, M., and Erdogan, F., 1999, "The Mixed Mode Crack Problem in an Inhomogeneous Orthotropic Medium," Int. J. Fract., **98**, pp. 243–261.
[18] Paulino, G. H., and Jin, Z.-H., 2000, "Viscoelastic Functionally Graded Ma-

terials Subjected to Antiplane Shear Fracture,'' ASME J. Appl. Mech., **68**(1), pp. 284–293.
[19] Aboudi, J., Pindera, M.-J., and Arnold, S. M., 1999, ''Higher-Order Theory for Functionally Graded Materials,'' Composites, Part B, **30**, pp. 777–832.
[20] Pindera, M.-J., and Dunn, P., 1997, ''Evaluation of the Higher-Order Theory for Functionally Graded Materials Via the Finite-Element Method,'' Composites, Part B, **28**(1/2), pp. 109–119.
[21] Goldberg, R. K., and Hopkins, D. A., 1995, ''Thermal Analysis of a Functionally Graded Material Subject to a Thermal Gradient Using the Boundary Element Method,'' Compos. Methods Appl. Mech. Eng., **5**(7), pp. 793–806.
[22] Sutradhar, S., Paulino, G. H., and Gray, L. J., 2002, ''Transient Heat Conduction in Homogeneous and Non-Homogeneous Materials by the Laplace Transform Galerkin Boundary Element Method,'' Eng. Anal. Bound. Elem., **26**(2), pp. 119–132.
[23] Eischen, J. W., 1987, ''Fracture of Nonhomogeneous Materials,'' Int. J. Fract., **34**, pp. 3–22.
[24] Williamson, R. L., Rabin, B. H., and Drake, J. T., 1993, ''Finite Element Analysis of Thermal Residual Stresses at Graded Ceramic-Metal Interfaces, Part I: Model Description and Geometric Effects,'' J. Appl. Phys., **74**(2), pp. 1310–1320.
[25] Drake, J. T., Williamson, R. L., and Rabin, B. H., 1993, ''Finite Element Analysis of Thermal Residual Stresses at Graded Ceramic-Metal Interfaces, Part II: Interface Optimization for Residual Stress Reduction,'' J. Appl. Phys., **74**(2), pp. 1321–1326.
[26] Giannakopoulos, A. E., Suresh, S., Finot, M., and Olsson, M., 1995, ''Elastoplastic Analysis of Thermal Cycling: Layered Materials With Compositional Gradients,'' Acta Mater., **43**(4), pp. 1335–1354.
[27] Gu, P., Dao, M., and Asaro, R. J., 1999, ''A Simplified Method for Calculating the Crack-Tip Field of Functionally Graded Materials Using the Domain Integral,'' ASME J. Appl. Mech., **66**(1), pp. 101–108.
[28] Dao, M., Gu, P., Maewal, A., and Asaro, R. J., 1997, ''A Micromechanical Study of Residual Stresses in Functionally Graded Materials,'' Acta Mater., **45**(8), pp. 3265–3276.
[29] Anlas, G., Santare, M. H., and Lambros, J., 2000, ''Numerical Calculation of Stress Intensity Factors in Functionally Graded Materials,'' Int. J. Fract., **104**, pp. 131–143.
[30] Santare, M. H., and Lambros, J., 2000, ''Use of Graded Finite Elements to Model the Behavior of Nonhomogeneous Materials,'' ASME J. Appl. Mech., **67**, pp. 819–822.
[31] Kim, J.-H., and Paulino, G. H., 2002, ''Finite Element Evaluation of Mixed Mode Stress Intensity Factors in Functionally Graded Materials,'' Int. J. Numer. Methods Eng., **53**(8), pp. 1903–1935.
[32] Erdogan, F., and Wu, B. H., 1997, ''The Surface Crack Problem for a Plate with Functionally Graded Properties,'' ASME J. Appl. Mech., **64**, pp. 449–456.
[33] Paulino, G. H., and Kim, J.-H., ''The Weak Patch Test for Nonhomogeneous Materials Modeled With Graded Finite Elements'' (submitted for publication).
[34] Hughes, T. J. R., 1987, *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Prentice-Hall, Englewood Cliffs, NJ.
[35] Jin, Z.-H., Paulino, G. H., and Dodds, Jr., R. H., 2002, ''Finite Element Evaluation of Quasi-Static Crack Growth in Functionally Graded Materials Using a Novel Cohesive Zone Fracture Model,'' ASME J. Appl. Mech., **69**, pp. 370-379.
[36] Kim, J.-H., ''Quasi-Static Crack Propagation in Functionally Graded Materials,'' Ph.D. thesis, University of Illinois at Urbana-Champaign, Urbana, IL.
[37] Hibbitt, Karlson, & Sorensen, Inc., 2000, *ABAQUS/Standard User's Manual*, Vol. II, Pawtucket, RI, Version 6.1 (p. 14.1.1-14).
[38] Gullerud, A. S., Koppenhoefer, K. C., Roy, A., Roychowdhury, S., and Dodds, Jr., R. H., 2001, *WARP3D-Release 13.11*, University of Illinois, UILU-ENG-95-2012.

B. Karp
Department of Mechanical Engineering,
Ben-Gurion University of the Negev,
P. O. B. 653,
Beer-Sheva 84105, Israel

D. Durban
Faculty of Aerospace Engineering,
Technion-Israel Institute of Technology,
Haifa 32000, Israel

Influence of Boundary Conditions on Decay Rates in a Prestrained Plate[1]

Decay of end perturbations imposed on a prestrained semi-infinite rectangular plate is investigated in the context of plane-strain incremental finite elasticity. A separation of variables eigenfunction formulation is used for the perturbed field within the plate. Numerical results for the leading decay exponent are given for three hyperelastic materials with various boundary conditions at the long faces of the plate. The study exposes a considerable sensitivity of axial decay rates to boundary data, to initial strain and to constitutive behavior. It is suggested that the results are relevant to the applicability of Saint-Venant's principle even though the eigenfunctions are not always self-equilibrating.
[DOI: 10.1115/1.1435365]

1 Introduction

Studies of axial decay rates of incremental end disturbances, in prestretched plates, are restricted to free boundaries (Durban and Stronge [1,2], and Durban and Karp [3]) with an earlier paper by Abeyaratne, Horgan, and Chung [4]. Here, we consider several new cases of boundary conditions (clamped, sliding, inextensional) to enhance earlier work in the context of estimating the validity of Saint-Venant's principle. Such boundary data can be viewed as modeling limit moduli ratios in composite multilayered plates.

The governing equations, extracted from ([3]), are summarized in the next section. The mathematical formulation leads to an eigensystem, for the two velocity components, that admits a separation of variables solution for the eigenfields. These are briefly discussed in Section 3 with the useful distinction between symmetric and antisymmetric perturbed fields. Axial decay is exponential with the decay rates obtained as eigenvalues of transcendental equations.

Numerical solutions for the lowest (exponential) decay rate—the leading decay exponent K—are detailed in Section 4. Calculations have been performed for three hyperelastic solids (foam rubbers) over a range of prestrain (up to an initial stretch of $\lambda=2.3$) and with different boundary conditions. The latter include free (FR), clamped (CL), sliding (SL), and inextensional (IN) walls. Further results can be deduced for mixed boundary conditions upon utilizing symmetrical and antisymmetrical properties of the eigenfields.

The leading decay exponent displays considerable sensitivity to initial stretch, constitutive behavior, and boundary data. Some of the results resemble those for anisotropic and multilayered composites in linear elasticity (Choi and Horgan [5,6] and Horgan and Simmonds [7]). This follows from the instantaneous anisotropy induced by initial stretch of the plate. The paper concludes with a discussion of the relevance of the present study to assessing the validity of Saint-Venant's principle in the incremental sense. The eigensystem generate infinite sets of eigenfunctions, each decaying in the axial direction but producing an eigenperturbation at the edge. It is suggested that though the eigenfields are, in general, not self-equilibrating over the perturbed end, they still provide useful bounds on axial decay rate of incremental end disturbances.

2 Perturbed Field Equations

Consider a semi-infinite plate (Fig. 1) uniformly stretched, under plane-strain conditions, in the z-direction, by an axial stress σ applied at the ends. Assume now that an incremental load is superposed on the uniform stress σ at the end $z=0$, thus inducing a quasi-static perturbed velocity field $\mathbf{V}$ within the body. We wish to examine the (plane-strain) instantaneous response of the prestrained plate to that incremental end disturbance under various boundary constraints, applied to the long faces $x=\pm h$ simultaneously with the end disturbance. In this setting, the plate is free to contract uniformly in the x-direction during initial stretching but is constrained at the onset of end loading. That formulation facilitates an eigenfunction analysis, which provides a simple assessment of the influence of boundary conditions on axial decay rates of prestrained plates.

The formulation follows an earlier work by the authors ([3]) where the standard case of free faces along the $x=\pm h$ has been studied, in the traditional spirit of investigating ([1,2]) the validity of Saint-Venant's principle. Accordingly, we shall just recapitulate a few selected results from ([3]) for future use.

The perturbed velocity vector is written as (see Fig. 1 for the unit triad $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$)

$$\mathbf{V}=u\mathbf{i}+w\mathbf{k} \tag{2.1}$$

where both velocity components (u,w) depend only on x and z. The Eulerian strain rate components follow in the form

$$\varepsilon_x=u_{,x} \quad \gamma_{xz}=\frac{1}{2}(u_{,z}+w_{,x}) \quad \varepsilon_z=w_{,z}\,. \tag{2.2}$$

Material response is governed by three in-plane relations in the form suggested by Hill [8], for incremental plane-strain response,

$$\overset{\nabla}{\sigma}_x=a\varepsilon_x+c\varepsilon_z \quad \overset{\nabla}{\tau}_{xz}=2\mu\gamma_{xz} \quad \overset{\nabla}{\sigma}_z=(c-\sigma)\varepsilon_x+b\varepsilon_z\,, \tag{2.3}$$

where $(\overset{\nabla}{\sigma}_x,\overset{\nabla}{\tau}_{xz},\overset{\nabla}{\sigma}_z)$ are the objective Jaumann stress rates and (a,b,c,μ) denote instantaneous material moduli. Notice that stress and strain rates can be interpreted also as objective increments in stress and strain since no physical time appears in our formulation. The instantaneous moduli in (2.3) are later derived from hyperelastic strain energy functions and are dependable on the uniform strain level.

[1]This work is based in part on a D.Sc. thesis submitted to the Technion.

Contributed by the Applied Mechanics Division in THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 14, 2001; final revision, Sept. 21, 2000. Editor: R. C. Benson. 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.

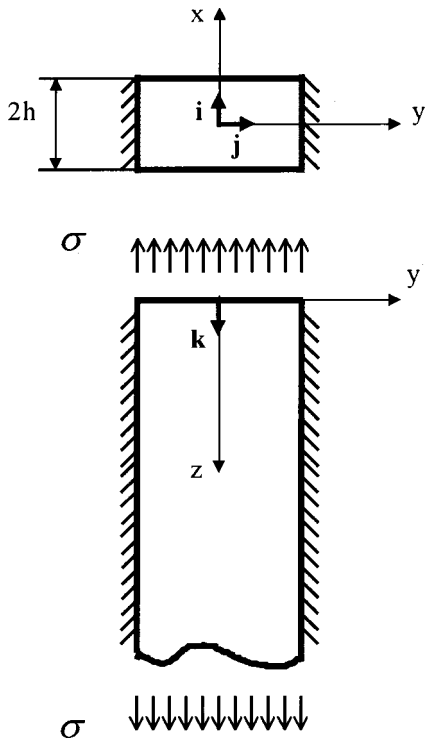


Fig. 1 Notation for a semi-infinite plate under uniform tension stress σ. Current thickness is $2h$.

Equilibrium of stress rates requires, in view of (2.3) and (2.1), that the velocities (u,w) satisfy the two equations ([8])

$$au_{,xx}+\beta u_{,zz}+(c+\alpha)w_{,xz}=0 \tag{2.4a}$$

$$(c+\alpha)u_{,xz}+\alpha w_{,xx}+bw_{,zz}=0 \tag{2.4b}$$

with

$$\alpha=\mu-\frac{1}{2}\sigma \quad \beta=\mu+\frac{1}{2}\sigma. \tag{2.5}$$

The boundary conditions which supplement Eqs. (2.4) are four different constraints imposed on the velocity vector (2.1) and on the traction rate vector (Durban and Stronge [9])

$$\pm\overset{\nabla}{\mathbf{t}}_x=\sigma_x\mathbf{i}+(\tau_{xz}-\sigma\gamma_{xz})\mathbf{k} \tag{2.6}$$

along the faces $x=\pm h$, respectively. All types of boundary data employed in this study can be expressed in terms of velocity components as detailed by the following classification:

Free boundaries (FR)

$$\begin{aligned}&\overset{\nabla}{\mathbf{t}}_x=\mathbf{0}\Rightarrow\sigma_x=0 \qquad au_{,x}+cw_{,z}=0\\ &\qquad\qquad\Rightarrow \qquad\qquad \text{at } x=\pm h\\ &\overset{\nabla}{\tau}_{xz}-\sigma\gamma_{xz}=0 \qquad u_{,z}+w_{,x}=0\end{aligned} \tag{2.7}$$

Clamped boundaries (CL)

$$\begin{aligned}&\mathbf{V}=\mathbf{0}\Rightarrow u=0\\ &\qquad\qquad\text{at } x=\pm h\\ &\qquad w=0\end{aligned} \tag{2.8}$$

Sliding boundaries (SL)

$$\begin{aligned}&\mathbf{V}\cdot\mathbf{i}=0 \qquad\qquad u=0\\ &\qquad\qquad\Rightarrow \qquad\qquad \text{at } x=\pm h\\ &\overset{\nabla}{\mathbf{t}}_x\cdot\mathbf{k}=0\Rightarrow\tau_{xz}-\sigma\gamma_{xz}=0 \qquad u_{,z}+w_{,x}=0\end{aligned} \tag{2.9}$$

Inextensional boundaries (IN)

$$\begin{aligned}&\overset{\nabla}{\mathbf{t}}_x=\mathbf{0}\Rightarrow\sigma_x=0 \qquad au_{,x}+cw_{,z}=0\\ &\qquad\qquad\Rightarrow \qquad\qquad \text{at } x=\pm h\\ &\mathbf{V}\cdot\mathbf{k}=0 \qquad\qquad w=0\end{aligned} \tag{2.10}$$

Solutions to Eqs. (2.4) along with each boundary data (2.7)–(2.10) generate eigenfunctions with associated eigenvalues that determine axial decay rates of imposed end disturbances. Emphasis in this study is placed on the influence of boundary data, when coupled with initial strain of the plate, on the decay rate of end disturbances. For each boundary data there is an infinite set of eigenfunctions decaying exponentially in the z-direction. Of these, the most important is the eigenfunction with the lowest rate of axial decay, as it may provide a bound on the decay rate of arbitrary incremental end loads.

3 Eigenfunctions and Eigenvalues

Separation of variables solutions of (2.4) are sought via the representation

$$u=U(x)\exp\left(\frac{i\pi kz}{2h}\right) \quad w=W(x)\exp\left(\frac{i\pi kz}{2h}\right) \tag{3.1}$$

where $U(x)$ and $W(x)$ are the transverse profiles, to be determined later, of the perturbed velocity field and k stands for the associated eigenvalue. The notation in (3.1) is slightly different from the one employed in common studies due to the factor $i\pi/2$ in the exponential argument. The present formulation, however, is in agreement with the notation in the literature of wave guides and serves as a basis for a follow-up study on the dynamic response of the prestrained plate.

Inserting (3.1) in Eqs. (2.4) results in the two ordinary differential equations

$$aU''-\beta\left(\frac{\pi k}{2h}\right)^2U+(c+\alpha)i\left(\frac{\pi k}{2h}\right)W'=0 \tag{3.2a}$$

$$(c+\alpha)i\left(\frac{\pi k}{2h}\right)U'+\alpha W''-b\left(\frac{\pi k}{2h}\right)^2W=0 \tag{3.2b}$$

where the prime denotes differentiation with respect to x. The solutions of (3.2) can be conveniently separated into symmetric and antisymmetric fields (indicated, respectively, by subscripts s and a) given by

$$U_s=A_1\sinh\left(\Gamma_1\frac{\pi kx}{2h}\right)+A_2\sinh\left(\Gamma_2\frac{\pi kx}{2h}\right) \tag{3.3a}$$

$$W_s=A_1\eta_1\cosh\left(\Gamma_1\frac{\pi kx}{2h}\right)+A_2\eta_2\cosh\left(\Gamma_2\frac{\pi kx}{2h}\right) \tag{3.3b}$$

and

$$U_a=A_3\cosh\left(\Gamma_1\frac{\pi kx}{2h}\right)+A_4\cosh\left(\Gamma_2\frac{\pi kx}{2h}\right) \tag{3.4a}$$

$$W_a=A_3\eta_1\sinh\left(\Gamma_1\frac{\pi kx}{2h}\right)+A_4\eta_2\sinh\left(\Gamma_2\frac{\pi kx}{2h}\right). \tag{3.4b}$$

Here (A_1,A_2,A_3,A_4) are integration constants, Γ_1 and Γ_2 are the roots of the characteristic equation of (3.2),

$$a\alpha\Gamma^4-d\Gamma^2+b\beta=0, \tag{3.5}$$

expressed as

$$\Gamma_1=\sqrt{\frac{d-\sqrt{d^2-4ab\alpha\beta}}{2a\alpha}} \quad \Gamma_2=\sqrt{\frac{d+\sqrt{d^2-4ab\alpha\beta}}{2a\alpha}} \tag{3.6}$$

where

$$d=ab+\alpha\beta-(c+\alpha)^2 \tag{3.7}$$

and

$$\eta_p=\frac{i(a\Gamma_p^2-\beta)}{\Gamma_p(c+\alpha)} \quad p=1,2. \tag{3.8}$$

Compliance with the boundary data (2.7)–(2.10) generates the transcendental equations for the eigenvalues k. A straightforward derivation leads to the following set of equations:

Free boundaries (FR)

$$\tanh\left(\Gamma_1\frac{\pi k}{2}\right)-\left(\frac{Q_1}{Q_2}\right)^{\pm 1}\tanh\left(\Gamma_2\frac{\pi k}{2}\right)=0 \tag{3.9}$$

where the $(+)$ and $(-)$ signs correspond to symmetric and antisymmetric modes, respectively, and

$$Q_p=\frac{a\Gamma_p+ic\eta_p}{\eta_p\Gamma_p+i} \quad p=1,2. \tag{3.10}$$

Clamped boundaries (CL)

$$\tanh\left(\Gamma_1\frac{\pi k}{2}\right)-\left(\frac{\eta_1}{\eta_2}\right)^{\pm 1}\tanh\left(\Gamma_2\frac{\pi k}{2}\right)=0 \tag{3.11}$$

where again the $(+)$ and $(-)$ signs correspond to symmetric and antisymmetric fields, respectively.

Sliding boundaries (SL)

In this case there are two equations for each field, namely

$$\sinh\left(\Gamma_1\frac{\pi k}{2}\right)=0 \ \text{ or } \ \sinh\left(\Gamma_2\frac{\pi k}{2}\right)=0 \tag{3.12a}$$

for symmetric modes, and

$$\cosh\left(\Gamma_1\frac{\pi k}{2}\right)=0 \ \text{ or } \ \cosh\left(\Gamma_2\frac{\pi k}{2}\right)=0 \tag{3.12b}$$

for antisymmetric modes.

Inextensional boundaries (IN)

Eigenvalue equations are here as in the (SL) case apart from the exchange of symmetric modes with antisymmetric modes. Thus

$$\cosh\left(\Gamma_1\frac{\pi k}{2}\right)=0 \ \text{ or } \ \cosh\left(\Gamma_2\frac{\pi k}{2}\right)=0 \tag{3.13a}$$

for symmetric modes, and

$$\sinh\left(\Gamma_1\frac{\pi k}{2}\right)=0 \ \text{ or } \ \sinh\left(\Gamma_2\frac{\pi k}{2}\right)=0 \tag{3.13b}$$

for antisymmetric modes.

The similarity between cases (SL) and (IN) is expected since for any symmetric mode the middle plane $x=0$ behaves as a sliding boundary, while for any antisymmetric mode this plane behaves as an inextensional boundary. In fact, it is a matter of ease to show that the solutions of (3.12) are the simple trigonometric roots

$$k=\pm i\frac{N}{\Gamma_1} \ \text{ and } \ k=\pm i\frac{N}{\Gamma_2} \tag{3.14}$$

where odd integers $N=1,3,5,\ldots$ correspond to antisymmetric modes while even integer $N=2,4,6,\ldots$ correspond to symmetric modes. The solution of (3.13) is identical with (3.14) but with odd integers in symmetric modes and even integers for antisymmetric modes. Each of the associated eigenfunctions induces an artificial boundary at $x=0,\pm 2h,\pm 3h,\ldots$ which is either a sliding (SL) plane or an inextensional (IN) plane, thus generating a periodicity which explains the similarity between conditions (SL) and (IN).

Each set of the transcendental equations generates infinite eigenvalues—as is apparent for sliding (inextensional) boundaries from (3.14)—which determine the decay rates of the associated eigenfunctions. Attention is focused here on eigenvalues k with positive imaginary parts that determine axial decay rates of eigenfunctions (3.1).

4 Material Model and Numerical Results

Lowest decay rates have been calculated for hyperelastic solids with the strain energy function (Hill [10])

$$W=\sum_j\frac{C_j}{m_j}\left[\lambda_1^{m_j}+\lambda_2^{m_j}+\lambda_3^{m_j}-3+\frac{1}{n}(J^{-nm_j}-1)\right] \tag{4.1}$$

where $(\lambda_1,\lambda_2,\lambda_3)$ are the principle stretches, $J=\lambda_1\lambda_2\lambda_3$ is the volume ratio, the summation is carried over pairs (C_j,m_j) which—like n—are known material parameters (notice that parameters C_j have dimensions of 10^{-2} Nmm^{-2}). The specific models considered in this study are the Blatz-Ko (BK) foam rubber (Blatz and Ko [11]) with the single term representation

$$\text{(BK)} \quad m_1=-2 \quad C_1=-22 \quad n=0.5 \tag{4.2}$$

and two vulcanized foam rubbers due to Storåkers [12]: a highly compressible natural rubber (S1) with

$$\text{(S1)} \quad m_1=-m_2=4.5 \quad C_1=1.85 \quad C_2=-9.20 \quad n=0.92 \tag{4.3}$$

and a nearly incompressible synthetic rubber (S2) with

$$\text{(S2)} \quad m_1=-m_2=3.6 \quad C_1=2.04 \quad C_2=-0.51 \quad n=2.5 \tag{4.4}$$

The instantaneous moduli of the plane-strain constitutive relations (2.3) follows as [3]

$$a=\frac{\lambda_x}{\lambda_z}\frac{\partial^2 W}{\partial\lambda_x^2}=(n+1)\sum_j m_jC_j\lambda^{-(nm_j+1)/(n+1)} \tag{4.5a}$$

$$b=\frac{\lambda_z}{\lambda_x}\frac{\partial^2 W}{\partial\lambda_z^2}=\sum_j C_j[(m_j-1)\lambda^{[(n+1)m_j-1]/(n+1)}+(nm_j+1)\lambda^{-(nm_j+1)/(n+1)}] \tag{4.5b}$$

$$c=\frac{\partial^2 W}{\partial\lambda_x\partial\lambda_z}=n\sum_j m_jC_j\lambda^{-(nm_j+1)/(n+1)} \tag{4.5c}$$

$$2\mu=\frac{\lambda_x^2+\lambda_z^2}{\lambda_x^2-\lambda_z^2}(\sigma_x-\sigma_z)=\frac{\lambda^{2(2n+1)/(n+1)}+1}{\lambda^{2(2n+1)/(n+1)}-1}\sigma \tag{4.5d}$$

where the principal stretches are identified as $\lambda_1=\lambda_z=\lambda$, $\lambda_2=\lambda_y=1$ and $\lambda_3=\lambda_x=\lambda^{-n/(n+1)}$, for all materials described by (4.1), and $\sigma_x=0$, $\sigma_z=\sigma$ are the true stress components. The uniaxial stress-stretch relation, under the plane-strain constraint, is given by

$$\sigma=\lambda^{-1(n+1)}\sum_j C_j(\lambda^{m_j}-\lambda^{-nm_j/(n+1)}). \tag{4.6}$$

The perturbed system (2.4) remains elliptic, for the three materials in (4.2)–(4.4), over the range of prestretch considered in this study, see ([3]). It is therefore permissible to restrict the space of eigenfunctions (3.1) to diffuse models with continuous derivatives.

With the known expressions for the instantaneous moduli it is possible to evaluate the eigenvalues for each case of boundary conditions from Eqs. (3.9), (3.11), and (3.14). The eigenvalues k are in general complex numbers which are labeled in increasing order of the (positive) imaginary part, namely

$$0<\mathrm{Im}(k_1)<\mathrm{Im}(k_2)<\mathrm{Im}(k_3)<\cdots. \tag{4.7}$$

It is apparent, by simple inspection of the transcendental eigenvalue equations that the roots appear on conjugate pairs and with opposite signs $(\pm k)$. Here, however, we concentrate on eigenvalues with $\mathrm{Im}(k)>0$, especially on k_1 in (4.7), which has the smallest imaginary part. That particular eigenvalue provides, via (3.1), a bound on the axial decay rate of end disturbances. The implications of the lowest decay rates to the validity of Saint-Venant's principle will be discussed in the next section.

In view of (4.7) we define the quantities (J,K) by

$$J+iK=\frac{\pi}{2}k_1\Leftrightarrow\frac{i\pi}{2}k_1=-K+iJ \tag{4.8}$$

where

$$K=\frac{\pi}{2}\mathrm{Im}(k_1)\quad J=\frac{\pi}{2}\mathrm{Re}(k_1) \tag{4.9}$$

with K denoting the leading axial decay exponent and J reflecting the axial oscillatory pattern of decay. Thus, the leading term of eigenfunctions (3.1) will decay as

$$\sim e^{-K(z/h)}\left(\cos J\frac{z}{h}+i\ \sin\ J\frac{z}{h}\right) \tag{4.10}$$

becoming dominant, among all eigenfunctions, at large distance from the perturbed edge.

Figure 2 shows the variation of the leading decay rate K with initial stretch λ, for the Blatz-Ko solid (4.2), with different boundary conditions. Similar results for the vulcanized foam rubbers (4.3)–(4.4) are displayed in Figs. 3 and 4. An available numerical procedure has been used to solve the eigenvalue equations, for both symmetric and antisymmetric modes, and to trace the first eigenvalue among all competing modes. It is apparent from Figs. 2–4 that the leading decay exponent K is highly sensitive to initial stretch level λ and to boundary conditions. Nevertheless, there are a few common characteristic patterns that emerge from the numerical results and deserve further elaboration.

To begin with, at small prestrain ($\lambda=1+\Delta$ with $\Delta\ll1$) the lowest decay rate, for all materials, is obtained with an antisymmetric mode and free faces, abbreviated as FR(asy). The asymptotic behavior of the leading decay exponent is simply ([3])

$$K\sim\sqrt{3\Delta}\quad J=0\ \text{ as }\ \Delta\to0 \tag{4.11}$$

reflecting very low rates of decay near the stress-free state. In fact, with that particular eigenfunction k_1 remains a purely imaginary number $(J=0)$ with increasing stretch for the vulcanized rubbers (Figs. 3–4). For the BK solid it reaches a sharp peak (Fig. 2) beyond which k_1 becomes a complex number and K decreases with further stretching. For the vulcanized rubbers K reaches a smooth maximum, and then decreases with increasing stretch, with no associated oscillatory branch $(J=0)$, over the range of stretch considered here.

The first symmetric eigenvalue for free faces (i.e., the one corresponding to symmetric fields FR(sym), with smallest positive imaginary part) gives much higher rates of decay for all three

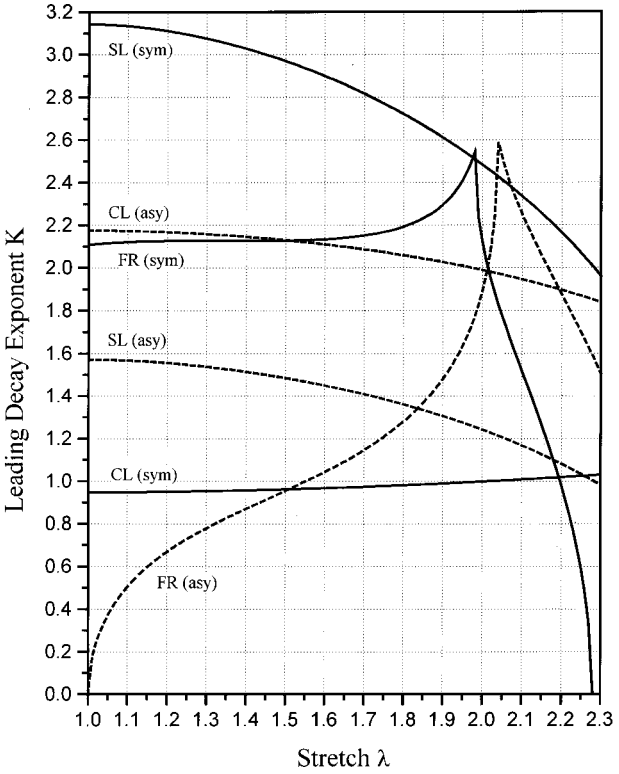


Fig. 2 Leading decay exponent K for the BK foam rubber, with different boundary conditions. Symmetric modes are shown with solid lines and indicated as (sym). Antisymmetric modes are shown with broken lines and indicated as (asy).

materials (Figs. 2–4). For vanishing initial stretch ($\lambda=1$) we recover the Papkovich-Fadle equation (the limit version of the symmetric equation in (3.9) when $\lambda\to1$)

$$\sin(2K-i2J)+(2K-i2J)=0 \tag{4.12}$$

with the leading solution

$$K=2.1061\quad J=1.1254. \tag{4.13}$$

The value of K associated with the first symmetric eigenfunction, for free faces, increases initially with stretch for the BK solid (Fig. 2), but decreases with stretch for the S1-S2 solids (Figs. 3–4). For all materials, however, there exists a definite stretch (where J vanishes) beyond which K decreases monotonically with λ. For the BK rubber that branch vanishes at the necking stretch ($\lambda=2.28$) indicating ([3]) very low decay near the onset of bifurcation in tension. Notice that the S1-S2 materials do not admit a point of instability in tension.

Thus, bounds on axial decay of end disturbances for the FR case are provided by the first antisymmetric eigenfunction for the vulcanized rubbers and by the envelope formed by both antisymmetric and symmetric solutions for the BK rubber. The intersection of those branches is at a stretch slightly above $\lambda=2$ (Fig. 2). The existence of a dominant antisymmetric mode near the stress free state could not be detected from the corresponding linear elastic equation

$$\sin(2K-i2J)-(2K-i2J)=0 \tag{4.14}$$

and has been exposed ([2,3]) only with the full nonlinear analysis.

For clamped boundaries (CL) the leading decay exponent corresponds to the first symmetric eigenfunction, CL(sym), for all three materials (Figs. 2–4), with no imaginary counterpart $(J=0)$. For the first antisymmetric mode, CL(asy), for the BK material K moderately decreases with stretch (Fig. 2) and is accompanied by an imaginary part (i.e., $J\neq0$). For the S1-S2 materials

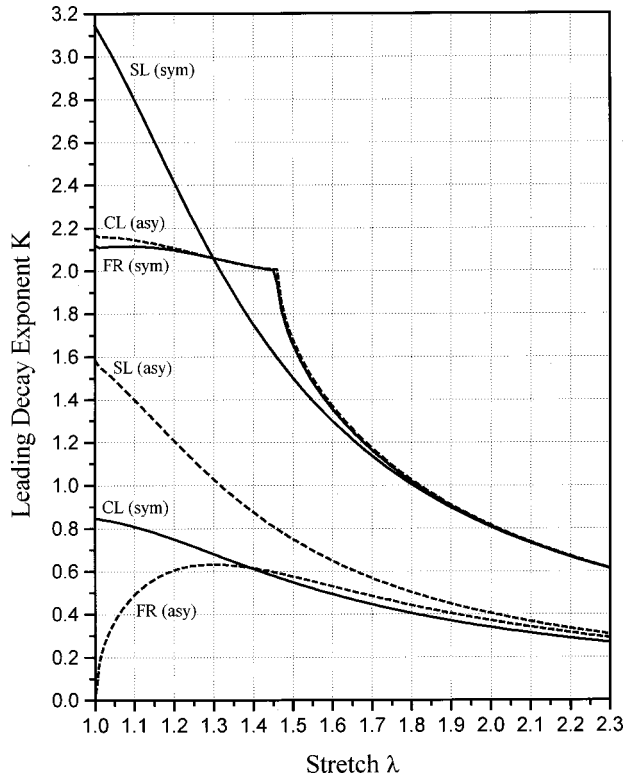


Fig. 3 Leading decay exponent *K* for the S1 highly compressible natural rubber, with different boundary conditions. Symmetric modes are shown with solid lines and indicated as (sym). Antisymmetric modes are shown with broken lines and indicated as (asy).

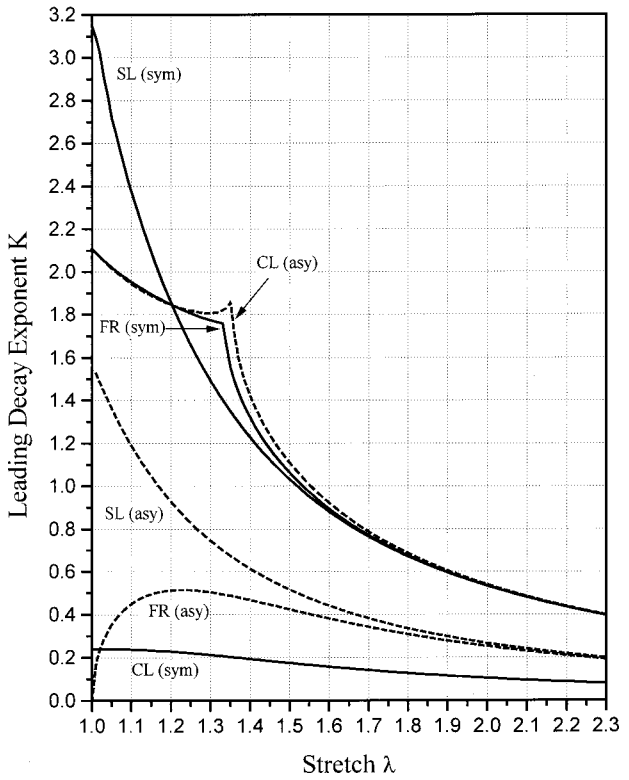


Fig. 4 Leading decay exponent *K* for the S2 nearly incompressible synthetic rubber, with different boundary conditions. Symmetric modes are shown with solid lines and indicated as (sym). Antisymmetric modes are shown with broken lines and indicated as (asy).

J vanishes at a definite stretch, characterized by a sharp corner (Figs. 3–4), beyond which K decreases sharply with λ, and decay is purely exponential.

The striking similarity between the curves for K in cases FR(sym) and CL(asy) in Figs. 3–4 for materials S1-S2, and in Figs. 2—up to moderate stretches—for the BK material, is supported by a simple analysis for small strains. Taking Eq. (3.11) to the limit of $\lambda=1$ we arrive at the linear elastic equations

$$\sin(2K-i2J)\mp\frac{2n+1}{2n+3}(2K-i2J)=0 \tag{4.15}$$

where the $(-)$ and $(+)$ signs correspond to symmetric and antisymmetric fields, respectively. Recalling, however, that the background, stress free, Poisson ratio of the entire hyperelastic family (4.1) is given by

$$\nu_0=\frac{n}{2n+1}\Rightarrow n=\frac{\nu_0}{1-2\nu_0}, \tag{4.16}$$

we find that (4.15) are exactly the clamped strip eigenvalue equations obtained by Little [13], for linear elastic materials, with the coefficients transformation

$$f=\frac{2n+1}{2n+3}=\frac{1}{3-4\nu_0}. \tag{4.17}$$

Now, for the nearly incompressible solid S2 we have that $n=25$ implying a background Poisson ratio of $\nu_0=0.4902$ with $f=0.9651$. The antisymmetric branch of (4.15) is then nearly identical with the symmetric FR branch in (4.12). That similarity increases with n as illustrated in Figs. 2–4. For the BK material with $n=0.5$ ($\nu_0=0.25$, $f=0.5$) we have at $\lambda=1$ that $K=2.1752$, for S1 with $n=0.92$ ($\nu_0=0.3239$, $f=0.5868$) we find the value of $K=2.160$, and for the S2 solid mentioned above we find $K=2.102$. These values should be compared with $K=2.1061$ in (4.13). A similar likeness exists for the imaginary part J of the CL(asy) leading eigenfunction.

By comparison, the first symmetric mode CL(sym) provides the smallest rates of decay for all three materials, when both walls are clamped, and over the full range of prestretch investigated here. The linear elastic results, obtained from the symmetric branch of (4.15) agree with those in ([13]) with $K=0.9447$ for the BK material, $K=0.845$ for the S1 material and $K=0.2399$ for the nearly incompressible solid S2. As n increases the asymptotic solution for K from (4.15) reads

$$K\sim\left(\frac{3}{2n+3}\right)^{1/2} \tag{4.18}$$

indicating vanishingly low decay rates for large values of n. For the S2 solid, with $n=25$, we find from (4.18) that $K=2.2379$, which is less than 1 percent below the exact value. Even for a relatively low value of $n=0.92$ for the S1 material we have from (4.18) the value $K=0.7873$ which is just about 7 percent below the exact results.

For sliding boundaries (SL) we have the closed form expressions (3.14) which for the BK material admit the simple form

$$K-iJ=\frac{\pi}{2}N\sqrt{\frac{(-R^2+8R-1)\pm(R-1)\sqrt{R^2-14R+1}}{6R}} \tag{4.19}$$

where $R=\lambda^{-(8/3)}$. Near the stress free state, with $\lambda=1+\Delta$ and $\Delta\ll 1$, we obtain from (4.19) the asymptotic behavior

$$K+iJ\sim\frac{\pi}{2}N\left(1+i\frac{4\sqrt{3}}{9}\Delta\right)\quad\text{as}\quad\Delta\to 0. \tag{4.20}$$

Thus, for sliding boundaries, the lowest rate of decay near the stress free state of a BK solid is dominated by the first antisymmetric function SL(asy) with $N=1$ and $K=\pi/2$. That function dominates also for finite prestrain (Fig. 2). A similar behavior has been observed in Fig. 3 for the S1 material, and in Fig. 4 for the S2 material, only that in these cases $J=0$ for both symmetric and antisymmetric leading eigenmodes.

Decay rates for inextensional boundaries (IN) are identical with those of the SL case except for the interchange of symmetric and antisymmetric modes. In symmetric fields we may regard the middle plane $x=0$ as a sliding boundary (SL) while in antisymmetric fields the middle plane behaves as an inextensional boundary (IN). Hence, with due account for the appropriate width $((1/2)h)$ in each half-field, we have the leading decay exponents

$$K_{\mathrm{IN(sym)}}=K_{\mathrm{SL(asy)}}\qquad K_{\mathrm{IN(asy)}}=K_{\mathrm{SL(sym)}} \tag{4.21}$$

while for mixed boundary conditions, with one face slipping and the other inextensional (SLIN), we find

$$K_{\mathrm{SLIN}}=\frac{1}{2}K_{\mathrm{SL(asy)}} \tag{4.22}$$

as the leading decay exponent.

5 Discussion

Initially, for small yet finite strains, all three materials admit the same order of leading decay rates for the boundary constraints considered here. These may be arranged in ascending magnitudes of K as

$$\begin{aligned}&\text{FR(asy), CL(sym), SL(asy) and IN(sym),}\\&\text{FR(sym), CL(asy), SL(sym) and IN(asy).}\end{aligned} \tag{5.1}$$

However, the range of validity of that sequence varies from about $\lambda<1.5$ for the BK material to $\lambda<1.3$ for the S1 material and $\lambda<1.02$ for the S2 material. This variation in range of stretch reflects the coupled influence of compressibility and prestretch near the stress free state.

With increasing stretch, there is a considerable sensitivity of the leading decay exponent to the level of prestrain, apparently due to the instantaneous anisotropy induced by initial stretch. The influence of boundary data for the BK material is noticeable also at high prestrain (Fig. 2). For the S1-S2 materials there is less influence on boundary conditions with increasing λ (Figs. 3–4) with the six curves grouping together closely around two leading decay exponents.

Values of the leading decay exponent for mixed modes can be deduced from the numerical results obtained here by utilizing the symmetry and antisymmetry of the eigenfunctions. Thus, the FR(asy) mode is also the eigenmode of a strip, having instantaneous thickness $(1/2)h$, with one boundary free and the other inextensional (FRIN). Hence, with an obvious notation,

$$K_{\mathrm{FRIN}}=\frac{1}{2}F_{\mathrm{FR(asy)}} \tag{5.2}$$

which for small strain equals half the value (4.11). A similar reasoning generates the mixed-modes decay exponents

$$K_{\mathrm{FRSL}}=\frac{1}{2}K_{\mathrm{FR(sym)}},\quad K_{\mathrm{CLIN}}=\frac{1}{2}K_{\mathrm{CL(asy)}},\quad K_{\mathrm{CLSL}}=\frac{1}{2}K_{\mathrm{CL(sym)}}. \tag{5.3}$$

These leading decay exponents are related to the current deformed thickness h at the instance of applying the end disturbance. The initial thickness h_0 is given by

$$h_0=h\lambda^{n/(n+1)} \tag{5.4}$$

for the entire hyperelastic family (4.1).

The leading decay exponent K is commonly accepted as a bound on the validity of Saint-Venant's principle when the eigenfunctions are self-equilibrating over the perturbed end $z=0$. This happens here in the case of free boundaries and for symmetric modes in the case of sliding boundaries. In other cases of boundary constraints there is no self-equilibrium of the eigenfunction at the end $z=0$. It may be argued, however, that if an arbitrary self-equilibrating end disturbance can be represented as a combination of all corresponding eigenfunctions then the leading decay exponent does provide a bound on axial decay rate. This issue should be resolved in conjunction with proof of completeness of the eigenfunctions for each boundary condition.

The boundary data examined in this study is ideal in the sense that no material properties are considered beyond the boundaries $x=\pm h$. In reality, multilayered composites ([6]) exhibit decay rates that depend on the relative moduli ratio of adjacent layers. It is conceivable, however, that at limit ratios of moduli in multilayered composites the leading decay exponent will behave essentially as exposed in this work.

Acknowledgments

Part of this study was supported by the fund for the promotion of research at the Technion. D. D. is grateful for the kind assistance of the Sydney Goldstein Chair in Aeronautical Engineering.

References

[1] Durban, D., and Stronge, W. J., 1988, "Diffusion of Self-Equilibrating End Loads in Plane Strain Plasticity," J. Mech. Phys. Solids, **36**, pp. 459–476.
[2] Durban, D., and Stronge, W. J., 1988, "Diffusion of Self-Equilibrating End Loads in Elastic Solids," ASME J. Appl. Mech., **55**, pp. 492–495.
[3] Durban, D., and Karp, B., 1992, "Axial Decay of Self-Equilibrating End Loads in Compressible Solids," ASME J. Appl. Mech., **59**, pp. 738–743.
[4] Abeyaratne, R., Horgan, C. O., and Chung, D.-T., 1985, "Saint-Venant End Effects for Incremental Plane Deformations of Incompressible Nonlinearly Elastic Materials," ASME J. Appl. Mech., **52**, pp. 847–852.
[5] Choi, I., and Horgan, C. O., 1977, "Saint-Venant's Principle and End Effects in Anisotropic Elasticity," ASME J. Appl. Mech., **44**, pp. 424–430.
[6] Choi, I., and Horgan, C. O., 1978, "Saint-Venant End Effects for Deformation of Sandwich Strips," Int. J. Solids Struct., **14**, pp. 187–195.
[7] Horgan, C. O., and Simmonds, J. G., 1994, "Saint-Venant End Effects in Composite Structures," Composites Eng., **3**, pp. 279–286.
[8] Hill, R., 1979, "On the Theory of Plane Strain in Finitely Deformed Compressible Materials," Math. Proc. Cambridge Philos. Soc., **86**, pp. 161–178.
[9] Durban, D., and Stronge, W. J., 1992, "Diffusion of Incremental Loads in Prestrained Bars," Proc. R. Soc. London, Ser. A, **439**, pp. 583–600.
[10] Hill, R., 1978, "Aspects of Invariance in Solid Mechanics," Adv. Appl. Mech., **18**, pp. 1–75.
[11] Blatz, P. J., and Ko, W. L., 1962, "Application of Finite Elastic Theory to the Deformation of Rubbery Materials," Trans. Soc. Rheol., **6**, pp. 223–251.
[12] Storåkers, B., 1986, "On Material Representation and Constitutive Branching in Finite Compressible Elasticity," J. Mech. Phys. Solids, **34**, pp. 125–145.
[13] Little, R. W., 1969, "Semi-Infinite Strip Problem with Built-In Edges," ASME J. Appl. Mech., **36**, pp. 320–323.

T.-H. Hao
State Key Lab for Modification of Polymer Materials and Chemical Fibers,
China Textile University,
P.O. Box 220,
Shanghai 200051, P. R. China

Displacement Boundary Value Problem of Piezoelectric Material Plane With an Elliptic Hole

In this paper, the displacement boundary value problem of a piezoelectric material plane with an elliptic hole is studied. As the permittivity of the medium in the hole is far less than that of the piezoelectric material, the electric induction of the medium is negligible. An exact solution is obtained. Its application to the inclusion problem has been given. The components of the stress tensor are discussed. [DOI: 10.1115/1.1464875]

Introduction

The mechanical and electric coupling of piezoelectric and ferroelectrics materials holds the key for mechatronic devices extensively used in electric engineering and information technology. It has been 24 years since Cherepanov [1] introduced the *J*-integral into the study of electromagnetic fields. A large amount of research data from various research groups (e.g., [2–6]) have been accumulated on the behavior of piezoelectric and ferroelectric materials. Yang [7] has given a comprehensive summary of the related studies before the year 2001. As an extension of these earlier results, the displacement boundary value problem of a piezoelectric material plane with an elliptic hole is dealt with in this paper. As the permittivity of the medium in the hole is far less than that of the piezoelectric material, on the surface of the hole, the normal component of electric displacement should diminish. With the conformal mapping method, this problem is solved. Lastly, its application to the inclusion problem has been given and the components of the stress tensor are calculated.

Basic Equations

In this section it is assumed that the piezoelectric material is a PZT-4 ceramic with material constants that can be found in Berlincourt et al. [8].

Let's consider the problem of a transversely isotropic piezoelectric material under a plane-strain deformation. Sosa [9] has dealt with it in detail. The result can be easily generalized to the anisotropic piezoelectric material as considered by Stroh [10], Barnett and Lothe [11], and Hao et al. [12]. In the study of Sosa [9] the mechanical field in such materials can be decoupled into a plane-strain problem and an antiplane shear problem. The constitutive relation for linear piezoelectric materials of this type can be written as

$$\varepsilon_i = a_{ij}\sigma_j + b_{hi}D_h, \quad E_f = -b_{fj}\sigma_j + \delta_{fh}D_h \tag{1}$$

where indices i and j have a range from 1 to 3, whereas indices f and h may only vary from 1 to 2. A column notation is anticipated in (1) for stress and strain components so that $\epsilon_3 = 2\epsilon_{12}$ and $\sigma_3 = \sigma_{12}$. The transverse isotropy makes $a_{31} = a_{13} = 0$, $a_{23} = a_{32} = 0$ and $b_{11} = b_{12} = b_{23} = 0$.

The decoupling from the antiplane case makes it possible to represent all field quantities by three pairs of analytic functions $f_k(z_k)$, where $z_k(x_1 + \mu_k x_2, k = 1,2,3)$ are three complex arguments. Symbols μ_1, μ_2, and μ_3 denote three complex roots of the characteristic equation (Appendix A) with positive imaginary parts. The linear combinations of these functions can represent stresses and electric displacements. According to Sosa [9], one has

$$\sigma_i = 2\,\mathrm{Re}\sum_{k=1}^{3} k_{ik} f_k''(z_k) \quad \text{and} \quad D_f = 2\,\mathrm{Re}\sum_{k=1}^{3} d_{fk} f_k''(z_k) \tag{2}$$

where $k_{1k} = \mu_k^2$, $k_{2k} = 1$, $k_{3k} = -\mu_k$, $d_{1k} = \lambda_k \mu_k$, $d_{2k} = -\lambda_k$ and $\lambda_k = [(b_{21} + b_{13})\mu_k^2 + b_{22}]/(\delta_{11}\mu_k^2 + \delta_{22})$.

The components of the displacement vector are expressed the same as given in Sosa [9].

$$u_1 = 2\,\mathrm{Re}\sum_{k=1}^{3} p_k f_k'(z_k) + \omega_1 x_2 + u_0 \quad u_2 = 2\,\mathrm{Re}\sum_{k=1}^{3} q_k f_k'(z_k) - \omega_1 x_1 + v_0 \tag{3}$$

where u_1 and u_2 are the components of displacement, p_k equals $a_{11}\mu_k^2 + a_{12} - b_{21}\lambda_k$, q_k equals $(a_{12}\mu_k^2 + a_{22} - b_{22}\lambda_k)/\mu_k$ and ω_1, u_0, v_0 are constants.

Boundary Conditions

Piezoelectric material occupies the plane except the area with an elliptic hole. On the surface of the hole, there are

$$u_1 = f_1(x_1, x_2) \quad \text{and} \quad u_2 = f_2(x_1, x_2) \tag{4}$$

where $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$ are known functions.

On the same surface, one also has

$$D_n = 0 \tag{5}$$

where D_n is the normal component of electric displacement.

At infinity

$$\sigma_i = \sigma_i^{\infty}; \quad D_i = D_i^{\infty}; (\partial u_2/\partial x_1 - \partial u_1/\partial x_2)/2 = \omega^{\infty} \tag{6}$$

where ω^{∞} is the rotation at infinity. When the displacement boundary value problem is studied, the rotation at infinity is necessary to be considered.

Solution to the Problem

Three complex analytical functions have different arguments in the form of $x_1 + \mu_k x_2$. These arguments generally have different values at the same physical point except infinity ∞ and points on Ox_1. In order to solve this question, according to Savin [13], one can find some mapping functions $\varpi_k(\varsigma)$. These functions map the outer region of the elliptic hole to the outer region of the unit circle. For different mapping functions $\varpi_k(\varsigma)$, the point $\varsigma = e^{i\theta}$ corresponds to the same point on physical plane. At infinity ∞, the point $\varsigma = \infty$ also corresponds to the same point on physical plane

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 27, 2001; final revision, December 5, 2001. 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.

(∞). As all the boundary conditions of this problem are defined at these points, one can deal with $f_k''[\varpi_k(\varsigma)]$ to replace $f_k''(z_k)$. This problem can be solved using the theory of a complex variable.

Now, how to choose the mapping functions $\varpi_k(\varsigma)$ is discussed. A nature choice is the mapping functions $\varpi_k(\varsigma)$ with following form:

$$\varpi_k(\zeta)=(a+i\mu_k b)/2\zeta+(a-i\mu_k b)\zeta/2, \quad k=0,1,2,3 \tag{7}$$

where μ_0 equals i, a, and b are the axes of the ellipse and μ_1, μ_2, μ_3 are the three roots of the characteristic equation (Appendix A) with positive imaginary parts.

Then, the boundary conditions have become the boundary values of these functions $f_k''[\varpi_k(\varsigma)]$.

On the surface of the hole (on the circumference of the unit circle), one has

$$u_1=2\ \mathrm{Re}\sum_{k=1}^{3} p_k f_k'[\varpi_k(\sigma)]=\varphi_1(\theta) \tag{8}$$

$$u_2=2\ \mathrm{Re}\sum_{k=1}^{3} q_k f_k'[\varpi_k(\sigma)]=\varphi_2(\theta) \tag{9}$$

where the variable σ equals $e^{i\theta}$, the variable θ can be found from $\sigma=e^{i\theta}$, and $\varphi_1(\theta);\varphi_2(\theta)$ are the known displacement functions.

In order to consider the electric displacement, a flow function ψ has been introduced by Sosa [9]. The flow function ψ has following differential relations:

$$D_1=\partial\psi/\partial x_2 \quad \text{and} \quad D_2=-\partial\psi/\partial x_1. \tag{10}$$

Therefore, the continuity equation (no charge) $\partial D_1/\partial x_1+\partial D_2/\partial x_2=0$ can be satisfied automatically (Appendix B). Applying the flow function ψ, the boundary condition (5) can be simplified. It is known that on the surface of the hole, one has (Appendix B)

$$n_1=-\partial x_2/\partial s, \quad n_2=\partial x_1/\partial s \tag{11}$$

where n_1 and n_2 are the components of the normal vector and s is the curve length along the surface of the hole in the counter clockwise direction.

On the basis of Eqs. (2) and (11), Eq. (5) becomes (Appendix B)

$$\psi=2\ \mathrm{Re}\sum_{k=1}^{3} \lambda_k f_k'[\varpi_k(\sigma)]=0. \tag{12}$$

At infinity

$$2\ \mathrm{Re}\sum_{k=1}^{3} k_{ik} f_k''(\infty)=\sigma_i^{\infty}; \quad 2\ \mathrm{Re}\sum_{a=1}^{3} d_{ak} f_k''(\infty)=D_i^{\infty};$$
$$\mathrm{Re}\sum_{k=1}^{3} (q_k-\mu_k p_k) f_k''(\infty)=\omega^{\infty}. \tag{13}$$

When $f_k''(\infty)=(B_k+iC_k)+O(1/z_k)$, one has

$$2\ \mathrm{Re}\sum_{k=1}^{3} k_{ik}(B_k+iC_k)=\sigma_i^{\infty}; \quad 2\ \mathrm{Re}\sum_{a=1}^{3} d_{ak}(B_k+iC_k)=D_i^{\infty};$$
$$\mathrm{Re}\sum_{k=1}^{3} (q_k-\mu_k p_k)(B_k+iC_k)=\omega^{\infty}. \tag{14}$$

There are three unknown complex constants (B_k+iC_k) and one has six real equations. The equations are sufficient to determine the unknown constants.

Now, the problem can be studied with two steps. First, it is considered that the whole plane is subjected to the homogeneous stress tensor σ_i^{∞}, homogeneous electric displacement vector D_i^{∞}, and homogeneous rotation ω^{∞}. On the unit circle and at infinity, they have caused the displacement vector and the flow function ψ as follows:

$$u_1=2\ \mathrm{Re}\sum_{k=1}^{3} p_k(B_k+iC_k)\varpi_k(\varsigma),$$
$$u_2=2\ \mathrm{Re}\sum_{k=1}^{3} q_k(B_k+iC_k)\varpi_k(\varsigma), \tag{15}$$
$$\psi=2\ \mathrm{Re}\sum_{k=1}^{3} \lambda_k(B_k+iC_k)\varpi_k(\varsigma)$$

Then, another case is considered, in which at infinity

$$\sigma_i^{\infty}=D_i^{\infty}=\omega^{\infty}=0, \quad \text{i.e.,} \quad B_k=C_k=0, \quad k=1,2,3 \tag{16}$$

and on the surface of the elliptic hole

$$u_1=2\ \mathrm{Re}\sum_{k=1}^{3} p_k f_k'[\varpi_k(\sigma)]=\varphi_{11}(\theta),$$
$$u_2=2\ \mathrm{Re}\sum_{k=1}^{3} q_k f_k'[\varpi_k(\sigma)]=\varphi_{12}(\theta) \tag{17}$$
$$\psi=2\ \mathrm{Re}\sum_{k=1}^{3} \lambda_k f_k'[\varpi_k(\sigma)]=\varphi_{13}(\theta)$$

where

$$\varphi_{11}(\theta)=\varphi_1(\theta)-2\ \mathrm{Re}\sum_{k=1}^{3} p_k(B_k+iC_k)\varpi_k(\sigma),$$
$$\varphi_{12}(\theta)=\varphi_2(\theta)-2\ \mathrm{Re}\sum_{k=1}^{3} q_k(B_k+iC_k)\varpi_k(\sigma),$$

and

$$\varphi_{13}(\theta)=-2\ \mathrm{Re}\sum_{k=1}^{3} \lambda_k(B_k+iC_k)\varpi_k(\sigma).$$

The functions $\varphi_{ij}(\theta)$ are known and discussed in detail in the numerical example.

On the basis of the equation in Muskhelishvili [14], one has (Appendix B)

$$2\sum_{k=1}^{3} p_k f_k'[\varpi_k(\varsigma)] = i\beta_1-(1/2\pi i)\int_{\gamma}\{\varphi_{11}(\theta)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]\}d\sigma$$
$$2\sum_{k=1}^{3} q_k f_k'[\varpi_k(\varsigma)] = i\beta_2-(1/2\pi i)\int_{\gamma}\{\varphi_{12}(\theta)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]\}d\sigma \tag{18}$$
$$2\sum_{k=1}^{3} \lambda_k f_k'[\varpi_k(\varsigma)] = i\beta_3-(1/2\pi i)\int_{\gamma}\{\varphi_{13}(\theta)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]\}d\sigma$$

where γ is the circumference of the unit circle in the counterclockwise direction.

For the boundary conditions, β_1, β_2, and β_3 are of no use; therefore, $i\beta_1=i\beta_2=i\beta_3=0$.

Considering two useful integrals $\int_\gamma (1/\sigma)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]d\sigma=-4\pi i/\varsigma$ and $\int_\gamma \sigma(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]d\sigma=0$ for $|\varsigma|>1$ (Appendix B), one obtains

$$\sum_{k=1}^{3} p_k f_k'[\varpi_k(\varsigma)]=d_1/\varsigma$$

$$\sum_{k=1}^{3} q_k f_k'[\varpi_k(\varsigma)]=d_2/\varsigma \tag{19}$$

$$\sum_{k=1}^{3} \lambda_k f_k'[\varpi_k(\varsigma)]=d_3/\varsigma$$

where

$$d_1=-(\varsigma/4\pi i)\int_\gamma \{\varphi_1(\theta)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]\}d\sigma-\sum_{k=1}^{3}\{[p_k(B_k+iC_k)(a+i\mu_k b)+\overline{p_k}(B_k-iC_k)(a+i\overline{\mu_k}b)]/2,$$

$$d_2=-(\varsigma/4\pi i)\int_\gamma \{\varphi_2(\theta)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]\}d\sigma-\sum_{k=1}^{3}\{[q_k(B_k+iC_k)(a+i\mu_k b)+\overline{q_k}(B_k-iC_k)(a+i\overline{\mu_k}b)]/2$$

and

$$d_3=-\sum_{k=1}^{3}\{[\lambda_k(B_k+iC_k)(a+i\mu_k b)+\overline{\lambda_k}(B_k-iC_k)(a+i\overline{\mu_k}b)]/2.$$

Using the linear algebra manipulations, one can obtain the following functions:

$$\begin{aligned}
f_1'[\varpi_1(\varsigma)]&=\{[-q_3d_1\lambda_2-d_2(\lambda_3p_2-p_3\lambda_2)+q_3p_2d_3+q_2(\lambda_3d_1-p_3d_3)]/[-q_3p_1\lambda_2-q_1(\lambda_3p_2-p_3\lambda_2)+q_3p_2\lambda_1+q_2(\lambda_3p_1-p_3\lambda_1)]\}/\varsigma\\
f_2'[\varpi_2(\varsigma)]&=\{[-q_3p_1d_3-q_1(\lambda_3d_1-p_3d_3)+q_3d_1\lambda_1+d_2(\lambda_3p_1-p_3\lambda_1)]/[-q_3p_1\lambda_2-q_1(\lambda_3p_2-p_3\lambda_2)+q_3p_2\lambda_1+q_2(\lambda_3p_1-p_3\lambda_1)]\}/\varsigma\\
f_3'[\varpi_3(\varsigma)]&=\{[-d_2p_1\lambda_2-q_1(d_3p_2-d_1\lambda_2)+d_2p_2\lambda_1+q_2(d_3p_1-d_1\lambda_1)]/[-q_3p_1\lambda_2-q_1(\lambda_3p_2-p_3\lambda_2)+q_3p_2\lambda_1+q_2(\lambda_3p_1-p_3\lambda_1)]\}/\varsigma
\end{aligned} \tag{20}$$

Substituting $[z_k+(z_k^2-a^2-\mu_k^2b^2)^{1/2}]/(a-i\mu_k b)$ for ς [or z_k for $\varpi_k(\varsigma)$], the three functions $f_k(z_k)$ have been determined.

Superposing the solutions on top of the homogeneous case, one obtains the final results are achieved. The physical quantities of every point can be found.

Numerical Example

According to Sosa [10], the material constants are

$$\begin{gathered}
a_{11}=8.205\times10^{-12},\quad a_{12}=-3.144\times10^{-12},\\
a_{22}=7.495\times10^{-12},\quad a_{33}=19.30\times10^{-12}\ (\mathrm{m^2N^{-1}})\\
b_{21}=-16.62\times10^{-3},\quad b_{22}=23.96\times10^{-3},\\
b_{13}=39.40\times10^{-3}\ (\mathrm{m^2C^{-1}})\\
\delta_{11}=7.66\times10^{7},\quad \delta_{22}=9.82\times10^{7}\ (\mathrm{V^2N^{-1}}),
\end{gathered} \tag{21}$$

These generalized complex variables $\mu_k(z_k=x_1+\mu_k x_2)$ are

$$\mu_1=1.228i,\quad \mu_2=0.203+1.067i,\quad \mu_3=-0.203+1.067i, \tag{22}$$

Now, a rigid elliptic inclusion is studied. As the permittivity of the inclusion is far less than that of the piezoelectric material, the electric induction of the inclusion is neglected.

On the surface of the hole, the displacement components are

$$u_1=-\omega_1x_2,\quad u_2=\omega_1x_1 \tag{23}$$

where ω_1 is the rotation of the inclusion and the rigid displacement vector is neglected.

At infinity, one has

$$\begin{gathered}
\sigma_1^\infty=-3.01\times10^6(\mathrm{N/m^2}),\quad \sigma_2^\infty=2\times10^6,\quad \sigma_3^\infty=0,\\
D_1^\infty=0(\mathrm{Cm^{-1}}),\\
D_2^\infty=-12.4\times10^{-4},\quad \omega^\infty=0.\ B_1=10^6,\\
C_1=B_2=C_2=B_3=C_3=0.
\end{gathered} \tag{24}$$

On the basis of the Eq. (23), considering $x_1=a\cos\theta$, $x_2=b\sin\theta$, $\cos\theta=(\sigma+1/\sigma)/2$ and $\sin\theta=(\sigma-1/\sigma)/2i$, one has

$$\begin{aligned}
u_1=\varphi_1(\theta)&=-\omega_1b(\sigma-1/\sigma)/2i\\
u_2=\varphi_2(\theta)&=\omega_1a(\sigma+1/\sigma)/2.
\end{aligned} \tag{25}$$

Substituting them into $\varphi_{11}(\theta)$, $\varphi_{12}(\theta)$, and $\varphi_{13}(\theta)$ in Eq. (17), it is obtained that

$$\begin{aligned}
\varphi_{11}(\theta)&=-\omega_1b(\sigma-1/\sigma)2i-2\ \mathrm{Re}\sum_{k=1}^{3}p_k(B_k+C_k)\varpi_k(\sigma)\\
\varphi_{12}(\theta)&=\omega_1a(\sigma+1/\sigma)/2-2\ \mathrm{Re}\sum_{k=1}^{3}q_k(B_k+C_k)\varpi_k(\sigma)\\
\varphi_{13}(\theta)&=-2\ \mathrm{Re}\sum_{k=1}^{3}\lambda_k(B_k+C_k)\varpi_k(\sigma).
\end{aligned} \tag{26}$$

Considering the symmetry of this case, one can deem that $\omega_1=0$. Substituting this conclusion and Eq. (26) into Eq. (19), one has

$$\sum_{k=1}^{3} p_k f_k'[\varpi_k(\varsigma)]=d_1/\varsigma$$

$$\sum_{k=1}^{3} q_k f_k'[\varpi_k(\varsigma)]=d_2/\varsigma \tag{27}$$

$$\sum_{k=1}^{3} \lambda_k f_k'[\varpi_k(\varsigma)]=d_3/\varsigma$$

where $d_1=-p_1B_1a$, $d_2=-iq_1B_1\mu_1b$ and $d_3=-\lambda_1B_1a$.

Using the linear algebra method, one has obtained the following functions $f_k'[\varpi_k(\varsigma)]$:

$$\begin{aligned}
f_1'[\varpi_1(\varsigma)]&=\{[-q_3d_1\lambda_2-d_2(\lambda_3p_2-p_3\lambda_2)+q_3p_2d_3+q_2(\lambda_3d_1-p_3d_3)]/[-q_3p_1\lambda_2-q_1(\lambda_3p_2-p_3\lambda_2)+q_3p_2\lambda_1+q_2(\lambda_3p_1-p_3\lambda_1)]\}/\varsigma\\
f_2'[\varpi_2(\varsigma)]&=\{[-q_3p_1d_3-q_1(\lambda_3d_1-p_3d_3)+q_3d_1\lambda_1+d_2(\lambda_3p_1-p_3\lambda_1)]/[-q_3p_1\lambda_2-q_1(\lambda_3p_2-p_3\lambda_2)+q_3p_2\lambda_1+q_2(\lambda_3p_1-p_3\lambda_1)]\}/\varsigma\\
f_3'[\varpi_3(\varsigma)]&=\{[-d_2p_1\lambda_2-q_1(d_3p_2-d_1\lambda_2)+d_2p_2\lambda_1+q_2(d_3p_1-d_1\lambda_1)]/[-q_3p_1\lambda_2-q_1(\lambda_3p_2-p_3\lambda_2)+q_3p_2\lambda_1+q_2(\lambda_3p_1-p_3\lambda_1)]\}/\varsigma.
\end{aligned} \tag{28}$$

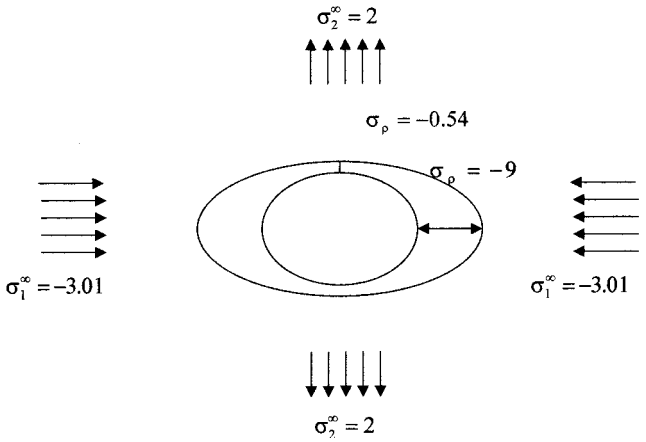


Fig. 1 The normal stress σ_ρ on the surface of the hold (10^6 N/m^2)

Substituting $[z_k+(z_k^2-a^2-\mu_k^2b^2)^{1/2}]/(a-i\mu_k b)$ for ς [or z_k for $\varpi_k(\varsigma)$], the three functions $f_k'(z_k)$ have been determined. Superposing the solutions on top of the homogeneous case, one obtains the final results. The physical quantities of every point can be found.

Now, let $a=2$, $b=1$. The components of the stress tensor are calculated. On the surface of the hole, the normal stress σ_ρ and the annular stress σ_θ are given in Figs. 1 and 2. The shear stress $\tau_{\rho\theta}$ is too small to be shown.

Concluding Remarks

An exact solution of the displacement boundary value problem of a piezoelectric material plane with an elliptic hole is given. Its application to a rigid inclusion is studied. The components of the stress tensor are calculated. Further study must focus on a more general method that considers the permittivity of the medium in the hole. It is also necessary to discuss the influence of the large deformation.

Acknowledgment

This work is supported by the National Natural Science Foundation of China.

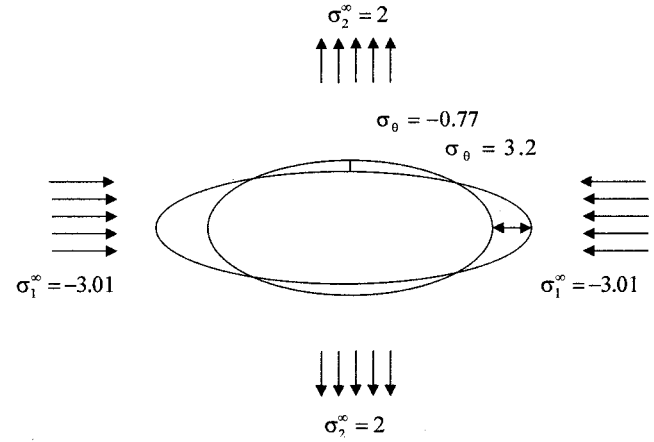


Fig. 2 The annular stress σ_θ on the surface on the hole (10^6 N/m^2)

Appendix A

The Characteristic Equation. According to Sosa [10], the stress function U and the flow function ψ are introduced so that the equations of elastic equilibrium and the continuity equation of electric displacement (no charge) can be satisfied. On the basis of the constitutive equations, the strains ε_i and the electric field intensities E_i can be obtained. Substituting them into the compatibility equations of strains and electric field intensity ($\partial E_1/x_2 - \partial E_2/x_1 = 0$), one obtains two linear differential equations. Eliminating the function ψ, the linear differential equation of the function U has been found. Assuming $U(x_1,x_2)=U(x_1+\mu x_2)$ and substituting it into the linear differential equation of the function U, the characteristic equation of μ has been obtained. That has been discussed in detail by Sosa [9].

Appendix B

The Deduction of Eqs. (12) and (18) and Two Useful Integrals. *1 The deduction of the Eq. (12):* The continuity equation (no charge) is $\partial D_1/\partial x_1+\partial D_2/\partial x_2=0$. According to Sosa [9], one has introduced a function ψ so that $D_1=\partial\psi/\partial x_2$ and $D_2=-\partial\psi/\partial x_1$.

$$\partial D_1/\partial x_1+\partial D_2/\partial x_2=\partial^2\psi/\partial x_2\partial x_1-\partial^2\psi/\partial x_1\partial x_2=0 \quad (A1)$$

The continuity equation (no charge) $\partial D_1/\partial x_1+\partial D_2/\partial x_2=0$ is satisfies automatically. For the curve of the surface of the hole, introducing the normal vector $\mathbf{n}$ (n_1,n_2) and the tangent vector $\mathbf{s}$ (s_1,s_2) (with counterclockwise direction), one has

$$n_1=-s_2=-\partial x_2/\partial s, \quad n_2=s_1=\partial x_1/\partial s. \quad (A2)$$

Substituting (A2) into the equation of D_n, one obtains

$$\begin{aligned} D_n &= D_1n_1+D_2n_2=-D_1s_2+D_2s_1 \\ &= -\partial\psi/\partial x_2\partial x_2/\partial s-\partial\psi/\partial x_1\partial x_1/\partial s \\ &= -d\psi/ds=0. \end{aligned}$$

$$\psi=\text{const.} \quad (A3)$$

As the function ψ is determined with an arbitrary constant, it can be deemed that the function ψ equals zero on the surface of the hole. Now, the expression of the function ψ is dealt with. From the relation between the function ψ and the electric displacement D_2 ($D_2=-\partial\psi/\partial x_1$) and the complex representation of the electric displacement D_2 $[D_2=-2\,\mathrm{Re}\sum_{k=1}^{3}\lambda_k f_k''(z_k)]$, one can find that $\psi=2\,\mathrm{Re}\sum_{k=1}^{3}\lambda_k f_k'(z_k)$,

2 The deduction of the Eq. (18): Considering the formula in Muskhelishvili [14] (his formula belongs to $|\varsigma|<1$ but our formula belongs to $|\varsigma|>1$), one deals with the function $F(\varsigma)$, analytic outside the unit circle. Therefore, the function $F(\varsigma)$ equals $\alpha+i\beta+\sum_{n=1}^{\infty}a_n/\varsigma^n$.

$$F(\infty)=\alpha+i\beta, \quad \overline{F}(1/\varsigma)=\alpha-i\beta+\sum_{n=1}^{\infty}\overline{a}_n\varsigma^n,$$

$$\lim_{\varsigma\to 0}\overline{F}(1/\varsigma)=\alpha-i\beta, \quad 1/\sigma=e^{-i\theta}=\overline{\sigma} \quad (A4)$$

Considering the Cauchy formula for $F(\varsigma)$, analytic outside the unit circle, one has

$$F(\varsigma)=-(1/2\pi i)\int_\gamma[F(\sigma)/(\sigma-\varsigma)]d\sigma+\alpha+i\beta, \quad |\varsigma|>1 \quad (A5)$$

$$0=(1/2\pi i)\int_\gamma [\overline{F}(1/\sigma)(\sigma-\varsigma)]d\sigma=(1/2\pi i)\int_\gamma [\overline{F}(\sigma)/(\sigma-\varsigma)]d\sigma \quad |\varsigma|>1 \tag{A6}$$

where γ is the circumference of the unit circle with counterclockwise direction.

Minus Eq. (*A*6), Eq. (*A*5) becomes

$$F(\varsigma)=-(1/2\pi i)\int_\gamma [F(\sigma)+\overline{F}(\overline{\sigma})]/(\sigma-\varsigma)d\sigma+\alpha+i\beta$$
$$=-(1/2\pi i)\left\{2\int_\gamma \mathrm{Re}\,F(\sigma)/(\sigma-\varsigma)d\sigma-2\pi i\alpha\right\}+i\beta \quad |\varsigma|>1. \tag{A7}$$

Now, the coefficient α is considered. For $|\varsigma|<1$, one has

$$(1/2\pi i)\int_\gamma [F(\sigma)/(\sigma-\varsigma)]d\sigma=\alpha+i\beta,$$
$$\overline{F}(1/\varsigma)=(1/2\pi i)\int_\gamma [\overline{F}(1/\sigma)/(\sigma-\varsigma)]d\sigma. \tag{A8}$$

When the variable ς equals zero, one knows

$$\alpha+i\beta=(1/2\pi i)\int_\gamma [F(\sigma)/\sigma]d\sigma,$$

$$\lim_{\varsigma\to 0}\overline{F}(1/\varsigma)=(1/2\pi i)\int_\gamma [\overline{F}(1/\sigma)/\sigma]d\sigma=\alpha-i\beta. \tag{A9}$$

Therefore, one obtains

$$(1/2\pi i)\int_\gamma [F(\sigma)+\overline{F}(\overline{\sigma})]/\sigma d\sigma$$
$$=(1/2\pi i)\int_\gamma [F(\sigma)+\overline{F}(1/\sigma)]/\sigma d\sigma$$
$$=\alpha+i\beta+\alpha-i\beta=2\alpha. \tag{A10}$$

Multiplying the coefficient πi, the following is obtained:

$$2\pi i\alpha=\pi i(1/2\pi i)\int_\gamma [F(\sigma)+\overline{F}(\overline{\sigma})]/\sigma d\sigma$$
$$=\pi i(1/2\pi i)2\int_\gamma \mathrm{Re}\,F(\sigma)/\sigma d\sigma=\int_\gamma \mathrm{Re}\,F(\sigma)/\sigma d\sigma. \tag{A11}$$

Substituting Eq. (*A*11) into Eq. (*A*7), one knows

$$F(\varsigma)=-(1/2\pi i)\left\{2\int_\gamma \mathrm{Re}\,F(\sigma)/(\sigma-\varsigma)d\sigma-2\pi i\alpha\right\}+i\beta$$
$$|\varsigma>1|=-(1/2\pi i)\left\{2\int_\gamma \mathrm{Re}\,F(\sigma)/(\sigma-\varsigma)d\sigma-\int_\gamma \mathrm{Re}\,F(\sigma)/\sigma d\sigma\right\}+i\beta \tag{A12}$$
$$|\varsigma|>1=-(1/2\pi i)\left\{\int_\gamma \mathrm{Re}\,F(\sigma)[2/(\sigma-\varsigma)-1/\sigma]d\sigma\right\}+i\beta \quad |\varsigma|>1=-(1/2\pi i)\left\{\int_\gamma \mathrm{Re}\,F(\sigma)[(\sigma+\varsigma)/\sigma(\sigma-\varsigma)]d\sigma\right\}+i\beta \quad |\varsigma|>1.$$

Letting $\mathrm{Re}\,F(\sigma)=f(\theta)$, one obtains

$$F(\varsigma)=i\beta-(1/2\pi i)\int_\gamma \{f(\theta)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]\}d\sigma, \quad |\varsigma|>1. \tag{A13}$$

When $F(\sigma)=2\sum_{k=1}^{3} p_k f_k'[\varpi_k(\sigma)]$ and $f(\theta)=\varphi_{11}(\theta)$, one obtains

$$2\sum_{k=1}^{3} p_k f_k'[\varpi_k(\varsigma)]$$
$$=i\beta_1-(1/2\pi i)\int_\gamma \{\varphi_{11}(\theta)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]\}d\sigma. \tag{A14}$$

That is the first part of the Eq. (18).

3 *Two useful integrals*: As $(\sigma+\varsigma)/\sigma(\sigma-\varsigma)=2/(\sigma-\varsigma)-1/\sigma$, one has

$$\int_\gamma (1/\sigma)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]d\sigma$$
$$=\int_\gamma (1/\sigma)[2/(\sigma-\varsigma)-1/\sigma]d\sigma$$
$$=\int_\gamma (2/\sigma)/(\sigma-\varsigma)d\sigma-\int_\gamma (1/\sigma)/\sigma d\sigma$$
$$=-2\pi i(2/\varsigma)-0=-4\pi i/\varsigma \quad |\varsigma|>1 \tag{A15}$$

$$\int_\gamma \sigma(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]d\sigma=\int_\gamma \sigma[2/(\sigma-\varsigma)-1/\sigma]d\sigma$$
$$=\int_\gamma 2\sigma/(\sigma-\varsigma)d\sigma-\int_\gamma d\sigma$$
$$=0-0=0 \quad |\varsigma|>1. \tag{A16}$$

4 The deduction of the Eq. (19): On the basis of Eq. (*A*14), it is known that

$$2\sum_{k=1}^{3} p_k f_k'[\varpi_k(\varsigma)]$$

$$= i\beta_1 - (1/2\pi i)\int_\gamma \{\varphi_{11}(\theta)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]\}d\sigma. \quad (A17)$$

Considering $i\beta_1 = 0$ and $\varphi_{11}(\theta) = \varphi_1(\theta) - 2\,\mathrm{Re}\sum_{k=1}^{3} p_k\,(B_k + iC_k) \times \varpi_k(\sigma)$, (A17) becomes

$$\sum_{k=1}^{3} p_k f_k'[\varpi_k(\varsigma)] = -(1/4\pi i)\int_\gamma \left\{\left[\varphi_1(\theta) - 2\,\mathrm{Re}\sum_{k=1}^{3} p_k(B_k + iC_k)\varpi_k(\sigma)\right](\sigma+\varsigma)/[\sigma(\sigma\varsigma)]\right\}d\sigma. \quad (A18)$$

Now, $2\,\mathrm{Re}\sum_{k=1}^{3} p_k\,(B_k + iC_k)\varpi_k(\sigma)$ is dealt with.

$$2\,\mathrm{Re}\sum_{k=1}^{3} p_k(B_k + iC_k)\varpi_k(\sigma)$$

$$= \sum_{k=1}^{3} [p_k(B_k + iC_k)\varpi_k(\sigma) + \overline{p_k}(B_k - iC_k)\overline{\varpi_k}(\overline{\sigma})]$$

$$= \sum_{k=1}^{3} \{p_k(B_k + iC_k)[(a + i\mu_k b)/\sigma + (a - i\mu_k b)\sigma]$$

$$+ \overline{p_k}(B_k - iC_k)[(a - i\overline{\mu}_k b)\sigma + (a + i\overline{\mu}_k b)/\sigma]\}/2$$

$$= \sum_{k=1}^{3} \{[p_k(B_k + iC_k)(a + i\mu_k b)$$

$$+ \overline{p_k}(B_k - iC_k)(a + i\overline{\mu}_k b)]/2\sigma] + (\ldots)\sigma\} \quad (A19)$$

Considering the two integrals (A15) and (A16), the following integral is

$$(1/2\pi i)\int_\gamma \mathrm{Re}\sum_{k=1}^{3} p_k(B_k + iC_k)\varpi_k(\sigma)(\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]d\sigma$$

$$= \sum_{k=1}^{3} \{[p_k(B_k + iC_k)(a + i\mu_k b) + \overline{p_k}(B_k - iC_k)(a + i\overline{\mu}_k b)]/\varsigma. \quad (A20)$$

Considering this result, one obtains

$$\sum_{k=1}^{3} p_k f_k'[\varpi_k(\varsigma)]$$

$$= -(1/4\pi i)\int_\gamma \{[\varphi_1(\theta)](\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]\}d\sigma$$

$$- \sum_{k=1}^{3} \{[p_k(B_k + iC_k)(a + i\mu_k b)$$

$$+ \overline{p_k}(B_k - iC_k)(a + i\overline{\mu}_k b)]/2\varsigma \quad (A21)$$

$$d_1 = -(\varsigma/4\pi i)\int_\gamma \{[\varphi_1(\theta)](\sigma+\varsigma)/[\sigma(\sigma-\varsigma)]\}d\sigma$$

$$- \sum_{k=1}^{3} \{[p_k(B_k + iC_k)(a + i\mu_k b)$$

$$+ \overline{p_k}(B_k - iC_k)(a + i\overline{\mu}_k b)]/2 \quad (A22)$$

$$\sum_{k=1}^{3} p_k f_k'[\varpi_k(\varsigma)] = d_1/\varsigma \quad (A23)$$

That is the first part of Eq. (19). Substituting $q_k; r_k$ for p_k, one obtains the whole Eq. (19).

5 The deduction of the Eq. (20): Using the matrix algebra, one can rewrite Eq. (19) with the following form:

$$\begin{bmatrix} p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{bmatrix}\begin{bmatrix} f_1'[\varpi_1(\varsigma)] \\ f_2'[\varpi_2(\varsigma)] \\ f_3'[\varpi_3(\varsigma)] \end{bmatrix} = \begin{bmatrix} d_1/\varsigma \\ d_2/\varsigma \\ d_3/\varsigma \end{bmatrix}.$$

One can find that

$$\Delta_0 = \det\begin{bmatrix} p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{bmatrix}, \quad \Delta_1 = \det\begin{bmatrix} d_1/\varsigma & p_2 & p_3 \\ d_2/\varsigma & q_2 & q_3 \\ d_3/\varsigma & \lambda_2 & \lambda_3 \end{bmatrix},$$

$$\Delta_2 = \det\begin{bmatrix} p_1 & d_1/\varsigma & p_3 \\ q_1 & d_2/\varsigma & q_3 \\ \lambda_1 & d_3/\varsigma & \lambda_3 \end{bmatrix} \text{ and } \Delta_3 = \det\begin{bmatrix} p_1 & p_2 & d_1/\varsigma \\ q_1 & q_2 & d_2/\varsigma \\ \lambda_1 & \lambda_2 & d_3/\varsigma \end{bmatrix}$$

$$f_1'[\varpi_1(\varsigma)] = \Delta_1/\Delta_0 = \{[-q_3 d_1\lambda_2 - d_2(\lambda_3 p_2 - p_3\lambda_2) + q_3 p_2 d_3 + q_2(\lambda_3 d_1 - p_3 d_3)]/[-q_3 p_1\lambda_2 - q_1(\lambda_3 p_2 - p_3\lambda_2) + q_3 p_2\lambda_1 + q_2(\lambda_3 p_1 - p_3\lambda_1)]\}/\varsigma$$

$$f_2'[\varpi_2(\varsigma)] = \Delta_2/\Delta_0, \quad f_3'[\varpi_3(\varsigma)] = \Delta_3/\Delta_0.$$

References

[1] Cherepanov, G. P., 1977, *Mechanics of Brittle Fracture*, McGraw-Hill, New York.

[2] Gao, H., Zhang, T.-Y., and Tong, P., 1997, "Local and Global Energy Release Rates for an Electrically Yielded Crack in a Piezoelectric Ceramics," J. Mech. Phys. Solids, **45**, pp. 491–509.

[3] Hao, T. H., et al., 1996, "Fracture Mechanics for the Design of Ceramics Multilayer Actuators," J. Mech. Phys. Solids, **44**, pp. 23–48.

[4] Sosa, H., 1992, "On The Fracture Mechanics of Piezoelectric Solids," Int. J. Solids Struct., **29**, pp. 2613–2622.

[5] Suo, Z., Kuo, C. M., Barnett, D. M., and Willis, J. R., 1992, "Fracture Mechanics for Piezoelectric Ceramics," J. Mech. Phys. Solids, **40**, pp. 739–765.

[6] Yang, W., and Suo, Z., 1994, "Cracks in Ceramic Actuators Caused by Electrostriction," J. Mech. Phys. Solids, **42**, pp. 649–663.

[7] Yang W., 2001, *Mechatronic Reliability*. Tsinghua University Press, Springer-Verlag, Beijing.

[8] Berlincourt, D. A., Curran, D. R., and Jaffe, H., 1964, "Piezoelectric and Piezoceramic Materials and Their Function in Transducers," Physical acoustics, W. P. Mason, ed., Academic Press, New York, **1A**, p. 177.

[9] Sosa, H., 1991, "Plane Problems in Piezoelectric Media With Defects," Int. J. Solids Struct., **28**, pp. 491–505.

[10] Stroh, A. N., 1958, "Dislocation and Cracks in Anisotropic Elasticity," Philos. Mag., **7**, pp. 625–646.

[11] Barnett, D. M., and Lothe, J., 1975, "Dislocations and Line Charges in Anisotropic Piezoelectric Insulators," Phys. Status Solidi B, **67**, pp. 105–111.

[12] Hao, T. H., et al., 1994, "A New Electric Boundary Conditions of Electric Fracture Mechanics and Its Applications," Eng. Fract. Mech., **47**, pp. 793–802.

[13] Savin, G. N., 1961, *Stress Concentration Around Holes*, Pergamon Press, New York.

[14] Muskhelishvili, N. I., 1963, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff, Groningen, Holland.

An Edge Dislocation in a Three-Phase Composite Cylinder Model With a Sliding Interface

X. Wang

Y.-p. Shen[1]
e-mail: ypsen@xjtu.edu.cn

Department of Engineering Mechanics,
Xi'an Jiotong University,
Xi'an,
Shaanxi Province 710049, P.R. China

An exact elastic solution is derived in a decoupled manner for the interaction problem between an edge dislocation and a three-phase circular inclusion with circumferentially homogeneous sliding interface. In the three-phase composite cylinder model, the inner inclusion and the intermediate matrix phase form a circumferentially homogeneous sliding interface, while the matrix and the outer composite phase form a perfect interface. An edge dislocation acts at an arbitrary point in the intermediate matrix. This three-phase cylinder model can simultaneously take into account the damage taking place in the circumferential direction at the inclusion-matrix interface and the interaction effect between the inclusions. As an application, we then investigate a crack interacting with the slipping interface. [DOI: 10.1115/1.1467090]

1 Introduction

The three-phase composite constitutive model plays an important role in the micromechanical analysis of composite materials. This three-phase model can be employed to simulate the interaction effects between neighboring inclusions ([1,2]) and also the interphase layer formed between the inclusion and the matrix ([3–6]). In addition, the methodology adopted in the three-phase model can also be extended without added difficulties to analyze a multiphase composite system (Ru [7]). The problem of dislocation in the three-phase model has also received many investigators' attention since the obtained solution can be used to study crack growth in composites, as well as strengthening and hardening mechanisms in alloyed materials. Luo and Chen [2] presented an exact solution to the problem of two concentric circles in an unbounded medium with a dislocation in the intermediate matrix region and no applied loads. In their model, all interfaces are assumed to be perfectly bonded; they found that in comparison with the two-phase model, the so-called trapping mechanism of dislocations is more likely to take place in the three-phase model; their analysis and calculation showed that in the three-phase model the orientation of Burgers vector has only limited influence on the stability of dislocation. Qaissaunee and Santare [3] considered the problem of an edge dislocation interacting with a three-phase elliptical inclusion, all interfaces are also assumed to be perfectly bonded in their analysis; their results showed that the presence of the interfacial zone can have a profound effect upon the stress field in the matrix and in the inclusion. Xiao and Chen [5,6] investigated the interaction between a dislocation and a coated circular inclusion. Their results showed that when the coating layer is thick, the elastic property of the inclusion has no significant influence on the force of the dislocation; while when the coating layer is thin, the elastic properties of both the inclusion and the coating can affect and change greatly the equilibrium position and the stability of the dislocation.

From another point of view, a study of elastic inclusions with sliding interfaces is important for modeling certain features of material behavior, such as the grain-boundary sliding in polycrystalline materials and the damage taking place in the circumferential direction at the inclusion-matrix interface. We will list some works on this line of research. Dundurs and Gangadharan [8] considered the interaction between an edge dislocation and a circular inclusion with a slipping interface. Mura and his collaborators [9–11] investigated the stress fields of ellipsoidal inclusions with sliding interfaces under various types of eigenstrain or remote loading. Lee et al. [12] presented an exact elastic solution for a circular sliding inclusion embedded in a half-plane in order to investigate the joint effect of a traction-free boundary and an interface. Gharpuray et al. [13] investigated an edge crack that terminates at a slipping interface with a different material. Ru [14] presented a general method to obtain the rigorous solution for a circular inclusion embedded within an infinite matrix with a circumferentially inhomogeneous sliding interface in plane elastostatics. By virtue of analytic continuation, the basic boundary value problem was reduced to a first-order differential equation for a single analytic function inside the circular inclusion, and the obtained differential equation was rigorously solved to obtain a finite form solution. Lubarda and Markenscoff [15] presented an energy study of sliding circular inclusions. They derived simple relationships between the energies of inclusions with sliding and bonded interfaces. Their results are of interest for the evaluation of average elastic properties of composites with sliding interfaces.

The limitations of the aforementioned works lie in that some investigators only considered three-phase inclusion with perfect interfaces and ignored the damage taking place on the inclusion-matrix interface, while other investigators considered the interface damage but ignored interaction between neighboring inclusions. The present investigation discusses the interaction problem between an edge dislocation and a three-phase circular inclusion with a sliding interface. In the three-phase cylinder model, the inclusion and the matrix form the inner circumferentially homogeneous sliding interface; while the matrix and the composite phase form the outer perfect bonding interface. It is of pragmatic and theoretical interest to examine the mobility of the dislocation when different types of interfaces coexist. The analysis demonstrates that all of the unknown coefficients can be determined in a decoupled way even though different types of interfaces coexist. As a result, mechanical analysis of this kind of composite system can be simplified considerably. It shall be addressed that the problem studied here could provide a fundamental solution for the dislocation density method for the interaction between an arbitrary crack and a three-phase circular inclusion with a sliding interface.

[1]Author 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, May 23, 2001; final revision, Oct. 15, 2001. Associate Editor: M.-J. Pindera. 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 demonstrate the importance of the obtained fundamental solution for edge dislocation, we then investigate matrix cracking in the three-phase composite system.

2 Formulation and Problem Statement

For plane deformations, the stresses and the corresponding boundary conditions can be expressed in terms of the two well-known Muskhelishvili's complex potentials $\phi(z)$ and $\psi(z)$ as (Muskhelishvili [16])

$$\begin{aligned}\sigma_{rr}+\sigma_{\theta\theta}&=2[\phi'(z)+\overline{\phi'(z)}]\\ \sigma_{rr}-i\sigma_{r\theta}&=\phi'(z)+\overline{\phi'(z)}-e^{2i\theta}[\overline{z}\phi''(z)+\psi'(z)]\end{aligned} \tag{1}$$

$$i\int (p_x+ip_y)ds=\phi(z)+z\overline{\phi'(z)}+\overline{\psi(z)} \tag{2}$$

$$2G(u_r+iu_\theta)=e^{-i\theta}[\kappa\phi(z)-z\overline{\phi'(z)}-\overline{\psi(z)}] \tag{3}$$

where $z=x+iy=re^{i\theta}$ is the complex coordinate, G is the in-plane shear modulus. For the plane deformation of isotropic materials, $\kappa=3-4\nu$ where ν is Poisson's ratio; for the plane deformation of transversely isotropic materials, $\kappa=1+2G_{12}/K_{12}$, where K_{12} and G_{12} are the plane-strain bulk modulus and in-plane shear modulus.

We now consider the problem as shown in Fig. 1. The circular region $|z|<a$ is occupied by inclusion phase S_1 whose elastic properties are κ_1 and G_1; the annulus region $a<|z|<b$ is occupied by the matrix phase S_2 whose elastic properties are κ_2 and G_2; the outer region $|z|>b$ is occupied by the composite phase S_3 whose elastic properties are κ_3 and G_3. In the following analysis, we will use the subscripts 1, 2, and 3 to identify the corresponding holomorphic functions defined in the regions S_1, S_2, and S_3.

• The inclusion and the matrix phase form a sliding interface Γ_a. The shear stress is proportional to the tangential displacement discontinuity on the sliding interface, and in addition both the normal stress and normal displacement are continuous across the interface, i.e.,

$$\sigma_{r\theta}^{(1)}=\sigma_{r\theta}^{(2)}=\lambda[u_\theta^{(2)}-u_\theta^{(1)}]\quad \sigma_{rr}^{(1)}=\sigma_{rr}^{(2)}\quad u_r^{(1)}=u_r^{(2)}\quad |\tau|=a \tag{4}$$

where λ is a non-negative constant interface parameter. When $\lambda=0$, shear stress on the interface Γ_a becomes zero and the interface is smooth; when $\lambda=\infty$, the interface Γ_a becomes a perfect interface.

The matrix and the composite phase are ideally bonded on their common interface Γ_b, i.e., both tractions and displacements are continuous.

$$\sigma_{r\theta}^{(2)}=\sigma_{r\theta}^{(3)}\quad \sigma_{rr}^{(2)}=\sigma_{rr}^{(3)}\quad u_r^{(2)}=u_r^{(3)}\quad u_\theta^{(2)}=u_\theta^{(3)}\quad |\xi|=b \tag{5}$$

An edge dislocation with Burgers vector $\hat{b}=\hat{b}_x+i\hat{b}_y$ is located at an arbitrary point $e=\hat{x}+i\hat{y}$ in the matrix S_2. In addition, the three-phase composite system is subjected to remote uniform loadings σ_{xx}^{∞}, σ_{yy}^{∞}, σ_{xy}^{∞}.

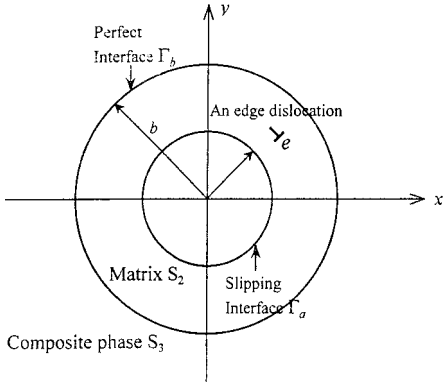


Fig. 1 An edge dislocation in the three-phase cylinder composite model with a slipping interface

3 Analysis

In order to treat the boundary conditions on the inner circle Γ_a, we introduce the analytical continuation defined by (see England [17])

$$\begin{cases}\phi_1(z)=-z\overline{\phi_1'}(a^2/z)-\overline{\psi_1}(a^2/z) & |z|>a\\ \phi_2(z)=-z\overline{\phi_2'}(a^2/z)-\overline{\psi_2}(a^2/z) & a^2/b<|z|<a\end{cases}. \tag{6}$$

Accordingly, we can introduce the following defined analytical continuation to treat the boundary conditions on the outer circle Γ_b

$$\begin{cases}\phi_2(z)=-z\overline{\phi_2'}(b^2/z)-\overline{\psi_2}(b^2/z) & b<|z|<b^2/a\\ \phi_3(z)=-z\overline{\phi_3'}(b^2/z)-\overline{\psi_3}(b^2/z) & |z|<b\end{cases}. \tag{7}$$

3.1 Satisfaction of Boundary Conditions on the Sliding Interface Γ_a. In view of (6), the continuity condition of tractions across Γ_a can be expressed as

$$[\phi_1(\tau)+\phi_2(\tau)]^+=[\phi_1(\tau)+\phi_2(\tau)]^-\quad |\tau|=a \tag{8}$$

where the superscripts "+" and "−" denote the inner and the outer sides of the contour being considered.

It can be found from Eq. (8) that the function $f(z)=\phi_1(z)+\phi_2(z)$ is holomorphic in the annulus $a^2/b<|z|<b$, except at the poles $z=e$ and $z=R^2/\overline{e}$. The principal part $f_0(z)$ of $f(z)$ is

$$f_0(z)=A\ln(z-e)-A\ln\frac{z-a^2/\overline{e}}{z}+\frac{\gamma_1\overline{A}}{z-a^2/\overline{e}}\quad (a^2/b<|z|<b) \tag{9}$$

where

$$A=\frac{G_2(\hat{b}_x+i\hat{b}_y)}{\pi i(1+\kappa_2)}\qquad \gamma_1=\frac{a^2(a^2-|e|^2)}{\overline{e}^3}. \tag{10}$$

Therefore, we can express

$$\begin{aligned}f(z)&=\phi_1(z)+\phi_2(z)\\ &=\sum_{n=1}^{+\infty}(A_nz^n+A_{-n}z^{-n})+A\ln(z-e)\\ &\quad -A\ln\frac{z-a^2/\overline{e}}{z}+\frac{\gamma_1\overline{A}}{z-a^2/\overline{e}}\quad (a^2/b<|z|<b),\end{aligned} \tag{11}$$

and the Laurent series whose coefficients $A_{\pm n}$ are to be determined is convergent in the annulus $a^2/b<|z|<b$.

The continuity condition of normal displacement across Γ_a can be expressed as

$$\begin{aligned}&\frac{1}{2G_1}\mathrm{Re}\{\tau^{-1}[\kappa_1\phi_1^+(\tau)+\phi_1^-(\tau)]\}\\ &\quad=\frac{1}{2G_2}\mathrm{Re}\{\tau^{-1}[\kappa_2\phi_2^-(\tau)+\phi_2^+(\tau)]\}\quad |\tau|=a.\end{aligned} \tag{12}$$

Substituting (11) into the above and eliminating $\phi_2^+(\tau)$ and $\phi_2^-(\tau)$ will yield

$$\begin{aligned}&\phi_1^+(\tau)+\frac{\tau^2}{a^2}\eta_1\overline{\phi_1^+}(a^2/\tau)+\eta_1\phi_1^-(\tau)+\frac{\tau^2}{a^2}\overline{\phi_1^-}(a^2/\tau)\\ &\quad=\beta_1\left[f(\tau)+\frac{\tau^2}{a^2}\overline{f}(a^2/\tau)\right]\quad |\tau|=a\end{aligned} \tag{13}$$

where the dimensionless parameters η_1 and β_1 are defined by

$$\eta_1=\frac{\kappa_2G_1+G_2}{\kappa_1G_2+G_1}\qquad \beta_1=\frac{\kappa_2G_1+G_1}{\kappa_1G_2+G_1}. \tag{14}$$

Now consider the function $\Omega(z)$ defined by

$$\Omega(z)=\begin{cases}\phi_1(z)+\dfrac{z^2}{a^2}\eta_1\overline{\phi}_1(a^2/z)+[(1-\eta_1)\overline{\phi_1'(0)}+\beta_1\overline{A/e}]z-\beta_1 \\ \quad\times\left[\displaystyle\sum_{n=1}^{+\infty}A_n z^n+\sum_{n=1}^{+\infty}\overline{A_{-(n-2)}}a^{-2n+2}z^n+\left(A-\overline{A}\frac{z^2}{a^2}\right)\ln(z-e)-\frac{e^2\overline{\gamma_1}Az^2}{a^4(z-e)}\right] & |z|<a \\ -\eta_1\phi_1(z)-\dfrac{z^2}{a^2}\overline{\phi}_1(a^2/z)+[(1-\eta_1)\overline{\phi_1'(0)}+\beta_1\overline{A/e}]z+\beta_1 \\ \quad\times\left[\displaystyle\sum_{n=1}^{+\infty}A_{-n}z^{-n}+\sum_{n=1}^{+\infty}\overline{A_{n+2}}a^{2n+2}z^{-n}-\left(A-\overline{A}\frac{z^2}{a^2}\right)\ln\frac{z-a^2/\overline{e}}{z}+\frac{\gamma_1\overline{A}}{z-a^2/\overline{e}}\right] & |z|>a.\end{cases} \tag{15}$$

It is seen from Eq. (13) that $\Omega(z)$ is analytic and single-valued in the whole complex plane including the points at infinity. According to the Liouville's theorem, $\Omega(z)$ should be a constant. Due to the fact that the constant only represents rigid-body translation ([16]), then it can be set zero and the following condition can be arrived at

$$\Omega(z)\equiv 0. \tag{16}$$

From the above condition, we obtain the following two equations:

$$\begin{aligned}&\phi_1(z)+\frac{z^2}{a^2}\eta_1\overline{\phi}_1(a^2/z)\\ &\quad=-[(1-\eta_1)\overline{\phi_1'(0)}+\beta_1\overline{A/e}]z+\beta_1\\ &\qquad\times\left[\sum_{n=1}^{+\infty}A_n z^n+\sum_{n=1}^{+\infty}\overline{A_{-(n-2)}}a^{-2n+2}z^n\right.\\ &\qquad\left.+\left(A-\overline{A}\frac{z^2}{a^2}\right)\ln(z-e)-\frac{e^2\overline{\gamma_1}Az^2}{a^4(z-e)}\right]\quad |z|<a\end{aligned} \tag{17}$$

$$\begin{aligned}&\frac{z^2}{a^2}\overline{\phi}_1(a^2/z)+\eta_1\phi_1(z)\\ &\quad=[(1-\eta_1)\overline{\phi_1'(0)}+\beta_1\overline{A/e}]z+\beta_1\\ &\qquad\times\left[\sum_{n=1}^{+\infty}A_{-n}z^{-n}+\sum_{n=1}^{+\infty}\overline{A_{n+2}}a^{2n+2}z^{-n}\right.\\ &\qquad\left.+\left(\overline{A}\frac{z^2}{a^2}-A\right)\ln\frac{z-a^2/\overline{e}}{z}+\frac{\gamma_1\overline{A}}{z-a^2/\overline{e}}\right]\quad |z|>a.\end{aligned} \tag{18}$$

Then the first compatibility condition can be obtained from the compatibility between (17) and (18) as follows:

$$\beta_1\,\mathrm{Re}\{A_1\}+(\eta_1-1)\mathrm{Re}\{\phi_1'(0)\}=\beta_1\,\mathrm{Re}\left\{\frac{A}{e}\right\}. \tag{19}$$

For convenience, we introduce the following dimensionless parameter χ

$$\chi=\frac{\lambda a\left(\dfrac{\kappa_1}{G_1}+\dfrac{1}{G_2}\right)}{2}. \tag{20}$$

The tangential interface condition on Γ_a can be expressed as

$$\begin{aligned}\mathrm{Im}\{\phi_1'^{+}(\tau)-\phi_1'^{-}(\tau)\}&=\frac{\lambda a}{2G_2}\mathrm{Im}\{\tau^{-1}[\kappa_2\phi_2^-(\tau)+\phi_2^+(\tau)]\}\\ &\quad-\frac{\lambda a}{2G_1}\mathrm{Im}\{\tau^{-1}[\kappa_1\phi_1^+(\tau)+\phi_1^-(\tau)]\}.\end{aligned} \tag{21}$$

Substituting (11) into the above equation and eliminating $\phi_2^+(\tau)$ and $\phi_2^-(\tau)$ will lead to

$$\begin{aligned}&\frac{1}{\chi}\tau[\phi_1'(\tau)+\overline{\phi_1'}(a^2/\tau)]^+-\frac{1}{\chi}\tau[\phi_1'(\tau)+\overline{\phi_1'}(a^2/\tau)]^-+\phi_1^+(\tau)\\ &\quad-\frac{\tau^2}{a^2}\eta_1\overline{\phi}_1^+(a^2/\tau)+\eta_1\phi_1^-(\tau)-\frac{\tau^2}{a^2}\overline{\phi}_1^-(a^2/\tau)\\ &\quad=\beta_1\left[f(\tau)-\frac{\tau^2}{a^2}\overline{f}(a^2/\tau)\right]\quad |\tau|=a.\end{aligned} \tag{22}$$

In view of the above equation, we introduce the function $\Pi(z)$ defined by

$$\Pi(z)=\begin{cases}\phi_1(z)-\dfrac{z^2}{a^2}\eta_1\overline{\phi}_1(a^2/z)+\dfrac{1}{\chi}z[\phi_1'(z)+\overline{\phi_1'}(a^2/z)]-[(1+\eta_1)\overline{\phi_1'(0)}+\beta_1\overline{A/e}]z-\beta_1 \\ \quad\times\left[\displaystyle\sum_{n=1}^{+\infty}A_n z^n-\sum_{n=1}^{+\infty}\overline{A_{-(n-2)}}a^{-2n+2}z^n+\left(A+\overline{A}\frac{z^2}{a^2}\right)\ln(z-e)+\frac{e^2\overline{\gamma_1}Az^2}{a^4(z-e)}\right] & |z|<a \\ -\eta_1\phi_1(z)+\dfrac{z^2}{a^2}\overline{\phi}_1(a^2/z)+\dfrac{1}{\chi}z[\phi_1'(z)+\overline{\phi_1'}(a^2/z)]-[(1+\eta_1)\overline{\phi_1'(0)}+\beta_1\overline{A/e}]z+\beta_1 \\ \quad\times\left[\displaystyle\sum_{n=1}^{+\infty}A_{-n}z^{-n}-\sum_{n=1}^{+\infty}\overline{A_{n+2}}a^{2n+2}z^{-n}-\left(A+\overline{A}\frac{z^2}{a^2}\right)\ln\frac{z-a^2/\overline{e}}{z}+\frac{\gamma_1\overline{A}}{z-a^2/\overline{e}}\right] & |z|>a.\end{cases} \tag{23}$$

We can find that $\Pi(z)$ is analytic and single-valued in the whole complex plane including the points at infinity. By Liouville's theorem, $\Pi(z)$ is identically zero. We can then obtain the following two equations:

$$\begin{aligned}
&\phi_1(z)-\frac{z^2}{a^2}\,\eta_1\overline{\phi}_1(a^2/z)+\frac{1}{\chi}z[\phi_1'(z)+\overline{\phi_1'}(a^2/z)]\\
&\quad=[(1+\eta_1)\overline{\phi_1'(0)}+\beta_1\overline{A/e}]z+\beta_1\\
&\qquad\times\left[\sum_{n=1}^{+\infty}A_nz^n-\sum_{n=1}^{+\infty}\overline{A_{-(n-2)}}a^{-2n+2}z^n\right.\\
&\qquad\left.+\left(A+\overline{A}\frac{z^2}{a^2}\right)\ln(z-e)+\frac{e^2\overline{\gamma_1}Az^2}{a^4(z-e)}\right]\quad |z|<a \qquad (24)\\
&\eta_1\phi_1(z)-\frac{z^2}{a^2}\,\overline{\phi}_1(a^2/z)-\frac{1}{\chi}z[\phi_1'(z)+\overline{\phi_1'}(a^2/z)]\\
&\quad=-[(1+\eta_1)\overline{\phi_1'(0)}+\beta_1\overline{A/e}]z+\beta_1\\
&\qquad\times\left[\sum_{n=1}^{+\infty}A_{-n}z^{-n}-\sum_{n=1}^{+\infty}\overline{A_{n+2}}a^{2n+2}z^{-n}\right.\\
&\qquad\left.-\left(A+\overline{A}\frac{z^2}{a^2}\right)\ln\frac{z-a^2/\overline{e}}{z}+\frac{\gamma_1\overline{A}}{z-a^2/\overline{e}}\right]\quad |z|>a. \qquad (25)
\end{aligned}$$

The compatibility between (24) and (25) gives the second compatibility condition

$$\beta_1\,\mathrm{Im}\{A_1\}-(1+\eta_1)\mathrm{Im}\{\phi_1'(0)\}=\beta_1\,\mathrm{Im}\left\{\frac{A}{e}\right\}. \qquad (26)$$

Noting that A_1 in (19) and (26) equals $-\overline{\Phi_{(1)}(0)}+A_0$ in Luo and Chen [2], and that $\phi_1'(0)$ equals $\Phi_{(1)}(0)$ in Luo and Chen [2], then it can be easily proved that the two compatibility conditions (19) and (26) are equivalent to the result obtained by Luo and Chen ([2], Eq. (14)) for a perfectly bonded interface.

We can get the following first-order partial differential equations for $\phi_1(z)$ from Eqs. (17), (18), (24), and (25)

$$\begin{aligned}
&2(\chi\eta_1-1)\phi_1(z)+(1+\eta_1)z\phi_1'(z)\\
&\quad=(\eta_1-1+2\chi\eta_1)\phi_1'(0)z+(n-2+2\chi\eta_1)\beta_1\\
&\qquad\times\sum_{n=2}^{+\infty}\left(A_n-A\frac{e^{-n}}{n}\right)z^n+\beta_1\sum_{n=3}^{+\infty}(n-2)\left[a^{-2n+2}\overline{A_{-(n-2)}}\right.\\
&\qquad\left.+\frac{e^{-n+3}}{a^4}\overline{\gamma_1}A+\frac{e^{-n+2}}{a^2(n-2)}\overline{A}\right]z^n\quad (|z|<a) \qquad (27)\\
&-2(1+\chi)\eta_1\phi_1(z)+(1+\eta_1)z\phi_1'(z)\\
&\quad=(\eta_1-1+2\chi\eta_1)\overline{\phi_1'(0)}z-(n+2+2\chi)\beta_1\\
&\qquad\times\sum_{n=1}^{+\infty}\left(A_{-n}+A\frac{a^{2n}\overline{e}^{-n}}{n}+\gamma_1\overline{A}a^{2(n-1)}\overline{e}^{-n+1}\right)z^{-n}\\
&\qquad-\beta_1\sum_{n=1}^{+\infty}(n+2)\left(\overline{A_{n+2}}a^{2n+2}-\frac{a^{2n+2}\overline{e}^{-n-2}}{n+2}\overline{A}\right)z^{-n}\\
&\qquad(|z|>a). \qquad (28)
\end{aligned}$$

It can be easily checked that the left-hand side of Eq. (27) is in agreement with the result derived by Ru [14] when setting $f(z)=0$ (homogeneous interface) in Eq. (4.17) of his paper. In Eqs. (27) and (28), we have expanded all of the terms into power series to facilitate the ensuing analysis.

From (27), we can get the expression for $\phi_1(z)$ in its definition region $|z|<a$ as follows:

$$\begin{aligned}
\phi_1(z)&=\phi_1'(0)z+\beta_1\sum_{n=2}^{+\infty}\frac{n-2+2\chi\eta_1}{(n+2\chi)\eta_1+n-2}\left(A_n-A\frac{e^{-n}}{n}\right)z^n\\
&\quad+\beta_1\sum_{n=3}^{+\infty}\frac{n-2}{(n+2\chi)\eta_1+n-2}\left(a^{-2n+2}\overline{A_{-(n-2)}}\right.\\
&\quad\left.+\frac{e^{-n+3}}{a^4}\overline{\gamma_1}A+\frac{e^{-n+2}}{a^2(n-2)}\overline{A}\right)z^n\quad (|z|<a). \qquad (29)
\end{aligned}$$

From (28), we can similarly get the expression for $\phi_1(z)$ in its continuation region $|z|>a$ as follows:

$$\begin{aligned}
\phi_1(z)&=-\overline{\phi_1'(0)}z+\beta_1\sum_{n=1}^{+\infty}\frac{n+2+2\chi}{n+(n+2+2\chi)\eta_1}\left(A_{-n}+A\frac{a^{2n}\overline{e}^{-n}}{n}\right.\\
&\quad\left.+\gamma_1\overline{A}a^{2(n-1)}\overline{e}^{-n+1}\right)z^{-n}+\beta_1\sum_{n=1}^{+\infty}\frac{n+2}{n+(n+2+2\chi)\eta_1}\\
&\quad\times\left(\overline{A_{n+2}}a^{2n+2}-\frac{a^{2n+2}\overline{e}^{-n-2}}{n+2}\overline{A}\right)z^{-n}\quad (|z|>a). \qquad (30)
\end{aligned}$$

Substituting (29) and (30) into (11) will result in expressions for $\phi_2(z)$ as follows:

$$\begin{aligned}
\phi_2(z)&=\sum_{n=1}^{+\infty}(A_nz^n+A_{-n}z^{-n})+A\ln(z-e)-A\ln\frac{z-a^2/\overline{e}}{z}\\
&\quad+\frac{\gamma_1\overline{A}}{z-a^2/\overline{e}}-\phi_1'(0)z-\beta_1\sum_{n=2}^{+\infty}\frac{n-2+2\chi\eta_1}{(n+2\chi)\eta_1+n-2}\\
&\quad\times\left(A_n-A\frac{e^{-n}}{n}\right)z^n-\beta_1\sum_{n=3}^{+\infty}\frac{n-2}{(n+2\chi)\eta_1+n-2}\\
&\quad\times\left(a^{-2n+2}\overline{A_{-(n-2)}}+\frac{e^{-n+3}}{a^4}\overline{\gamma_1}A+\frac{e^{-n+2}}{a^2(n-2)}\overline{A}\right)z^n\\
&\quad(a^2/b<|z|<a) \qquad (31)\\
\phi_2(z)&=\sum_{n=1}^{+\infty}(A_nz^n+A_{-n}z^{-n})+A\ln(z-e)-A\ln\frac{z-a^2/\overline{e}}{z}\\
&\quad+\frac{\gamma_1\overline{A}}{z-a^2/\overline{e}}+\overline{\phi_1'(0)}z-\beta_1\sum_{n=1}^{+\infty}\frac{n+2+2\chi}{n+(n+2+2\chi)\eta_1}\\
&\quad\times\left(A_{-n}+A\frac{a^{2n}\overline{e}^{-n}}{n}+\gamma_1\overline{A}a^{2(n-1)}\overline{e}^{-n+1}\right)z^{-n}\\
&\quad-\beta_1\sum_{n=1}^{+\infty}\frac{n+2}{n+(n+2+2\chi)\eta_1}\left(\overline{A_{n+2}}a^{2n+2}\right.\\
&\quad\left.-\frac{a^{2n+2}\overline{e}^{-n-2}}{n+2}\overline{A}\right)z^{-n}\quad (a<|z|<b). \qquad (32)
\end{aligned}$$

Up to now, the boundary conditions (4) on the sliding interface Γ_a have been completely satisfied.

3.2 Satisfaction of Boundary Conditions on the Perfect Interface Γ_b. The continuity condition of traction across interface Γ_b can be expressed as

$$[\phi_2(\xi)+\phi_3(\xi)]^+=[\phi_2(\xi)+\phi_3(\xi)]^-\quad |\xi|=b. \qquad (33)$$

The following expression can be obtained from the above equation:

$$g(z)=\phi_2(z)+\phi_3(z)$$
$$=\sum_{n=1}^{+\infty}(B_n z^n+B_{-n}z^{-n})+A\ln(z-e)-A\ln(z-b^2/\overline{e})$$
$$+pA\ln z+\frac{\gamma_2\overline{A}}{z-b^2/\overline{e}}\quad a<|z|<b^2/a \tag{34}$$

where

$$\gamma_2=\frac{b^2(b^2-|e|^2)}{\overline{e}^3}\quad p=\frac{G_3}{G_2}\frac{1+k_2}{1+\kappa_3}, \tag{35}$$

and the Laurent series whose coefficients $B_{\pm n}$ are to be determined is convergent in the annulus region $a<|z|<b^2/a$.

The continuity condition of displacements across Γ_b can be expressed as

$$\frac{1}{2G_2}[\kappa_2\phi_2^+(\xi)+\phi_2^-(\xi)]=\frac{1}{2G_3}[\kappa_3\phi_3^-(\xi)+\phi_3^+(\xi)]\quad |\xi|=b. \tag{36}$$

Substituting (34) into the above and eliminating $\phi_2^+(\xi)$ and $\phi_2^-(\xi)$ will yield

$$\phi_3^+(\xi)+\eta_2\phi_3^-(\xi)=\beta_2 g(\xi)\quad |\xi|=b \tag{37}$$

where the two dimensionless parameters η_2 and β_2 are defined by

$$\eta_2=\frac{\kappa_3G_2+G_3}{\kappa_2G_3+G_2}\quad \beta_2=\frac{\kappa_2G_3+G_3}{\kappa_2G_3+G_2}. \tag{38}$$

Consider the function $\Delta(z)$ defined by

$$\Delta(z)=\begin{cases}\phi_3(z)-\beta_2\left[\sum_{n=1}^{+\infty}B_n z^n-A\ln(z-b^2/\overline{e})+\frac{\gamma_2\overline{A}}{z-b^2/\overline{e}}\right]-pA\ln z+\eta_2\Pi z+\frac{\overline{\Pi}'b^2}{z} & |z|<b\\ -\eta_2\phi_3(z)+\beta_2\left[\sum_{n=1}^{+\infty}B_{-n}z^{-n}+A\ln\frac{z-e}{z}\right]+\eta_2pA\ln z+\eta_2\Pi z+\frac{\overline{\Pi}'b^2}{z} & |z|>b\end{cases} \tag{39}$$

where

$$\Pi=\frac{1}{4}(\sigma_{xx}^{\infty}+\sigma_{yy}^{\infty}),\quad \Pi'=\frac{1}{2}(\sigma_{yy}^{\infty}-\sigma_{xx}^{\infty})+i\sigma_{xy}^{\infty}. \tag{40}$$

It is apparent that $\Delta(z)$ is analytic and single-valued in the whole complex plane including the points at infinity. By Liouville's theorem, $\Delta(z)$ is identically zero, i.e.,

$$\Delta(z)\equiv 0. \tag{41}$$

Hence we can obtain the following two expressions for $\phi_3(z)$:

$$\phi_3(z)=\beta_2\left[\sum_{n=1}^{+\infty}B_n z^n-A\ln(z-b^2/\overline{e})+\frac{\gamma_2\overline{A}}{z-b^2/\overline{e}}\right]+pA\ln z$$
$$-\eta_2\Pi z-\frac{\overline{\Pi}'b^2}{z}\quad |z|<b \tag{42}$$

$$\phi_3(z)=\frac{\beta_2}{\eta_2}\left[\sum_{n=1}^{+\infty}B_{-n}z^{-n}+A\ln\frac{z-e}{z}\right]+pA\ln z+\Pi z$$
$$+\frac{\overline{\Pi}'b^2}{\eta_2 z}\quad |z|>b. \tag{43}$$

We have checked carefully our results with those obtained by Luo and Chen (2, Eq. (19)) and found a discrepancy between the two in the singular behavior of $\phi_3(z)$ at $z=0$. It can be proved that their result can not guarantee the dislocation condition along the circle when traversed counterclockwise a full circuit. Furthermore, our result is in agreement with that derived by Honein and Herrmann [18].

Remark. It can be deduced from Eq. (19), in [2] that

$$\psi_3(z)=\frac{iG_2(\hat{b}_x-i\hat{b}_y)}{\pi(1+\kappa_2)}\ln z+0(1)\quad \text{when}\quad z\to\infty, \tag{44}$$

while according to the definition for an edge dislocation with Burgers vector $\hat{b}_x+i\hat{b}_y$, the following asymptotic behavior for $\psi_3(z)$ shall establish ([18])

$$\psi_3(z)=\frac{iG_3(\hat{b}_x-i\hat{b}_y)}{\pi(1+\kappa_3)}\ln z+0(1)\quad \text{when}\quad z\to\infty. \tag{45}$$

Apparently, Eq. (44) is not in agreement with (45), while our results (42) and (43) are in accordance with (45).

Substituting the above two expressions (42) and (43) into (34) will result in the following expressions for $\phi_2(z)$:

$$\phi_2(z)=(1-\beta_2)\left[\sum_{n=1}^{+\infty}B_n z^n-A\ln(z-b^2/\overline{e})+\frac{\gamma_2\overline{A}}{z-b^2/\overline{e}}\right]$$
$$+\sum_{n=1}^{+\infty}B_{-n}z^{-n}+A\ln(z-e)+\eta_2\Pi z+\frac{\overline{\Pi}'b^2}{z}$$
$$a<|z|<b \tag{46}$$

$$\phi_2(z)=\sum_{n=1}^{+\infty}B_n z^n-A\ln(z-b^2/\overline{e})+\frac{\gamma_2\overline{A}}{z-b^2/\overline{e}}+\left(1-\frac{\beta_2}{\eta_2}\right)$$
$$\times\left[\sum_{n=1}^{+\infty}B_{-n}z^{-n}+A\ln\frac{z-e}{z}\right]+A\ln z-\Pi z-\frac{\overline{\Pi}'b^2}{\eta_2 z}$$
$$b<|z|<b^2/a. \tag{47}$$

The boundary conditions (5) on the perfect interface Γ_b have been completely satisfied in this subsection.

3.3 Determination of the Unknown Coefficients in the Laurent Series. In order to simultaneously satisfy all of the boundary conditions on interfaces Γ_a and Γ_b, the two expressions for $\phi_2(z)$ and $\psi_2(z)$ obtained in the above two subsections must be compatible to each other (see Luo and Chen [2], England [17], Worden and Keer [19], and Chao and Tan [20]). Physically, the compatibility conditions for $\phi_2(z)$ and $\psi_2(z)$ mean that the stress field and displacement field in the intermediate matrix are unique.

The compatibility condition for $\phi_2(z)$ in the annulus $a<|z|<b$ will result in

$$\sum_{n=1}^{+\infty} A_n z^n + \overline{\phi_1'(0)} z = (1-\beta_2)\left[\sum_{n=1}^{+\infty} B_n z^n - A\ln(z-b^2/\overline{e}) + \frac{\gamma_2 \overline{A}}{z-b^2/\overline{e}}\right] + \eta_2 \Pi z \tag{48}$$

$$\sum_{n=1}^{+\infty} B_{-n} z^{-n} + \frac{\overline{\Pi}' b^2}{z} = \sum_{n=1}^{+\infty} A_{-n} z^{-n} - A\ln\frac{z-a^2/\overline{e}}{z} + \frac{\gamma_1 \overline{A}}{z-a^2/\overline{e}} - \beta_1 \sum_{n=1}^{+\infty} \frac{n+2+2\chi}{n+(n+2+2\chi)\eta_1}\left(A_{-n} + A\frac{a^{2n}\overline{e}^{-n}}{n} + \gamma_1 \overline{A} a^{2(n-1)}\overline{e}^{-n+1}\right) z^{-n}$$

$$-\beta_1 \sum_{n=1}^{+\infty} \frac{n+2}{n+(n+2+2\chi)\eta_1}\left(\overline{A_{n+2}} a^{2n+2} - \frac{a^{2n+2}\overline{e}^{-n-2}}{n+2}\overline{A}\right) z^{-n}. \tag{49}$$

The compatibility condition for $\psi_2(z)$ in $a<|z|<b$ will result in the identity

$$\overline{\phi}_2(a^2/z) - \overline{\phi}_2(b^2/z) + \frac{a^2-b^2}{z}\phi_2'(z) = 0 \quad a<|z|<b. \tag{50}$$

The above identity is equivalent to the compatibility condition obtained by Worden and Keer [19]. Substituting (31) and (47) into the above equation and also applying (48) and (49) will result in the following equation:

$$\sum_{n=1}^{+\infty} (a^{2n}\overline{A_n} z^{-n} + a^{-2n}\overline{A_{-n}} z^n) + \overline{A}\ln\frac{z-a^2/\overline{e}}{z} - \beta_1 \sum_{n=2}^{+\infty} \frac{(n-2+2\chi\eta_1)a^{2n}}{n\eta_1+n-2+2\chi}\left(\overline{A_n} - \overline{A}\frac{\overline{e}^{-n}}{n}\right) z^{-n} - \overline{\phi_1'(0)} a^2 z^{-1}$$

$$-\beta_1 \sum_{n=3}^{+\infty} \frac{(n-2)a^{2n}}{n\eta_1+n-2+2\chi}\left[a^{-2n+2}A_{-(n-2)} + \frac{\overline{e}^{-n+3}}{a^4}\gamma_1\overline{A} + \frac{\overline{e}^{-n+2}}{a^2(n-2)}A\right] z^{-n}$$

$$-\sum_{n=1}^{+\infty} b^{2n}\overline{B_n} z^{-n} + \left(\frac{\beta_2}{\eta_2} - 1\right)\left[\sum_{n=1}^{+\infty} b^{-2n}\overline{B_{-n}} z^n + \overline{A}\ln(z-b^2/\overline{e})\right] + \Pi b^2 z^{-1} + \frac{\Pi'}{\eta_2} z + (a^2-b^2)$$

$$\times\left[\overline{\phi_1'(0)} z^{-1} + \sum_{n=-1}^{+\infty} (n+2)A_{n+2} z^n - \sum_{n=3}^{+\infty} (n-2)B_{-(n-2)} z^{-n} - Ae^{-1}z^{-1} - \overline{\Pi}' b^2 z^{-3}\right] = 0 \quad a<|z|<b. \tag{51}$$

The above Eq. (51) can be decomposed into the following two equations:

$$\sum_{n=1}^{+\infty} a^{-2n}\overline{A_{-n}} z^n + \left(\frac{\beta_2}{\eta_2} - 1\right)\left[\sum_{n=1}^{+\infty} b^{-2n}\overline{B_{-n}} z^n + \overline{A}\ln(z-b^2/\overline{e})\right] + \frac{\Pi'}{\eta_2} z + (b^2-a^2)\sum_{n=1}^{+\infty} (n+2)A_{n+2} z^n = 0 \tag{52}$$

$$\sum_{n=1}^{+\infty} a^{2n}\overline{A_n} z^{-n} + \overline{A}\ln\frac{z-a^2/\overline{e}}{z} - \overline{\phi_1'(0)} b^2 z^{-1} - \beta_1 \sum_{n=2}^{+\infty} \frac{(n-2+2\chi\eta_1)a^{2n}}{n\eta_1+n-2+2\chi}\left(\overline{A_n} - \overline{A}\frac{\overline{e}^{-n}}{n}\right) z^{-n}$$

$$-\beta_1 \sum_{n=3}^{+\infty} \frac{(n-2)a^{2n}}{n\eta_1+n-2+2\chi}\left[a^{-2n+2}A_{-(n-2)} + \frac{\overline{e}^{-n+3}}{a^4}\gamma_1\overline{A} + \frac{\overline{e}^{-n+2}}{a^2(n-2)}A\right] z^{-n} + \Pi b^2 z^{-1}$$

$$-\sum_{n=1}^{+\infty} b^{2n}\overline{B_n} z^{-n} + (a^2-b^2)\left[A_1 z^{-1} - \sum_{n=3}^{+\infty} (n-2)B_{-(n-2)} z^{-n} - Ae^{-1}z^{-1} - \overline{\Pi}' b^2 z^{-3}\right] = 0. \tag{53}$$

Expanding all of the terms in Eqs. (48), (49), (52), and (53) and equating the same power of z will result in the following set of linear algebraic equations:

$$\begin{cases} A_1 + (\beta_2-1)B_1 + \overline{\phi_1'(0)} = (1-\beta_2)\left(\dfrac{\overline{e}}{b^2}A - \dfrac{\overline{e}^2}{b^4}\gamma_2\overline{A}\right) + \eta_2\Pi \\ \rho A_1 - B_1 - \phi_1'(0) + (\rho-1)\overline{A_1} = \dfrac{\rho}{e}A + \dfrac{\rho-1}{\overline{e}}\overline{A} - \Pi \end{cases} \tag{54}$$

$$\begin{cases} A_2 + (\beta_2-1)B_2 = (1-\beta_2)\left(\dfrac{\overline{e}^2}{2b^4}A - \dfrac{\overline{e}^3}{b^6}\gamma_2\overline{A}\right) \\ A_2 - \rho^{-2}B_2 - \dfrac{\chi\beta_1}{(1+\chi)}\overline{A_2} = \dfrac{A}{2e^2} - \dfrac{\chi\beta_1}{(1+\chi)}\dfrac{\overline{A}}{2e^2} \end{cases} \tag{55}$$

$$A_n + (\beta_2-1)B_n = (1-\beta_2)\left(\frac{\overline{e}^n}{nb^{2n}}A - \frac{\overline{e}^{n+1}}{b^{2n+2}}\gamma_2\overline{A}\right), \tag{56a}$$

$$n\beta_1 a^{2n-2}\overline{A_n} - [n-2+(n+2\chi)(\eta_1-\beta_1)]A_{-(n-2)} + [n-2+(n+2\chi)\eta_1]B_{-(n-2)}$$

$$= [n-2+(n+2\chi)(\eta_1-\beta_1)]\left[\frac{a^{2n-4}}{(n-2)\overline{e}^{n-2}}A + \frac{a^{2n-6}}{\overline{e}^{n-3}}\gamma_1\overline{A}\right] + \beta_1\frac{a^{2n-2}}{\overline{e}^n}\overline{A} - [1+(1+2\chi)\eta_1]\overline{\Pi}' b^2\delta_{n3}, \tag{56b}$$

$$(b^2-a^2)nA_n + a^{-2n+4}\overline{A_{-(n-2)}} + \left(\frac{\beta_2}{\eta_2}-1\right)b^{-2n+4}\overline{B_{-(n-2)}} = \left(\frac{\beta_2}{\eta_2}-1\right)\frac{\overline{e}^{n-2}}{(n-2)b^{2n-4}}\overline{A} - \frac{\Pi'}{\eta_2}\delta_{n3}, \tag{56c}$$

$$\left[1-\frac{\beta_1(n-2+2\chi\eta_1)}{(n+2\chi)\eta_1+n-2}\right]a^{2n}\overline{A_n}-\frac{(n-2)\beta_1 a^2}{(n+2\chi)\eta_1+n-2}A_{-(n-2)}-b^{2n}\overline{B_n}+(b^2-a^2)(n-2)B_{-(n-2)}$$
$$=\frac{(n-2)\beta_1 a^{2n}}{(n+2\chi)\eta_1+n-2}\left[\frac{\overline{e}^{-n+3}}{a^4}\gamma_1\overline{A}+\frac{\overline{e}^{-n+2}}{a^2(n-2)}A\right]+\left[1-\frac{\beta_1(n-2+2\chi\eta_1)}{(n+2\chi)\eta_1+n-2}\right]\frac{a^{2n}}{n\overline{e}^n}\overline{A}+(a^2-b^2)\overline{\Pi}'b^2\delta_{n3}\quad\text{for } n=3,4,5,\ldots,+\infty \tag{56d}$$

where δ_{ij} is the Kronecker delta and $\rho=(a/b)^2$ is the volume concentration of the inclusion.

Associating Eqs. (54), (55), (56), (19), and (26), then all of the unknowns $\phi_1'^*(0)=a\phi_1'(0)$, $A^*_{\pm n}=A_{\pm n}a^{\pm n}$ and $B^*_{\pm n}=B_{\pm n}a^{\pm n}$ can be uniquely determined in the following decoupled way:

$$\begin{pmatrix} A_1^* \\ B_1^* \\ \phi_1'^*(0)\end{pmatrix}=\mathbf{K}_1^{-1}\,\mathrm{Re}\{\mathbf{V}_1+\mathbf{U}_1\}+i\mathbf{T}_1^{-1}\,\mathrm{Im}\{\mathbf{V}_1\} \tag{57}$$

$$\begin{pmatrix} A_2^* \\ B_2^*\end{pmatrix}=\mathbf{K}_2^{-1}\,\mathrm{Re}\{\mathbf{V}_2\}+i\mathbf{T}_2^{-1}\,\mathrm{Im}\{\mathbf{V}_2\} \tag{58}$$

$$\begin{pmatrix} A_n^* \\ A_{-(n-2)}^* \\ B_n^* \\ B_{-(n-2)}^*\end{pmatrix}=\mathbf{K}_n^{-1}\,\mathrm{Re}\{\mathbf{V}_n+\mathbf{U}_3\delta_{n3}\}+i\mathbf{T}_n^{-1}\,\mathrm{Im}\{\mathbf{V}_n+\mathbf{U}_3\delta_{n3}\}\quad(n=3,4,5,\ldots,+\infty) \tag{59}$$

where

$$\mathbf{K}_1=\begin{pmatrix} 1 & \beta_2-1 & 1 \\ 2\rho-1 & -1 & -1 \\ \beta_1 & 0 & \eta_1-1\end{pmatrix},$$

$$\mathbf{T}_1=\begin{pmatrix} 1 & \beta_2-1 & -1 \\ 1 & -1 & -1 \\ \beta_1 & 0 & -\eta_1-1\end{pmatrix}$$

$$\mathbf{K}_2=\begin{pmatrix} 1 & \beta_2-1 \\ 1-\dfrac{\chi\beta_1}{1+\chi} & -\rho^{-2}\end{pmatrix},\quad \mathbf{T}_2=\begin{pmatrix} 1 & \beta_2-1 \\ 1+\dfrac{\chi\beta_1}{1+\chi} & -\rho^{-2}\end{pmatrix}$$

$$\mathbf{K}_n=\begin{pmatrix} 1 & 0 & (\beta_2-1) & 0 \\ n\beta_1 & -[n-2+(n+2\chi)(\eta_1-\beta_1)] & 0 & [n-2+(n+2\chi)\eta_1] \\ (\rho^{-1}-1)n & 1 & 0 & \left(\dfrac{\beta_2}{\eta_2}-1\right)\rho^{n-2} \\ \left[1-\dfrac{\beta_1(n-2+2\chi\eta_1)}{(n+2\chi)\eta_1+n-2}\right] & -\dfrac{(n-2)\beta_1}{(n+2\chi)\eta_1+n-2} & -\rho^{-n} & (\rho^{-1}-1)(n-2)\end{pmatrix}$$

$$\mathbf{T}_n=\begin{pmatrix} 1 & 0 & (\beta_2-1) & 0 \\ -n\beta_1 & -[n-2+(n+2\chi)(\eta_1-\beta_1)] & 0 & [n-2+(n+2\chi)\eta_1] \\ (\rho^{-1}-1)n & -1 & 0 & \left(1-\dfrac{\beta_2}{\eta_2}\right)\rho^{n-2} \\ \left[\dfrac{\beta_1(n-2+2\chi\eta_1)}{(n+2\chi)\eta_1+n-2}-1\right] & -\dfrac{(n-2)\beta_1}{(n+2\chi)\eta_1+n-2} & \rho^{-n} & (\rho^{-1}-1)(n-2)\end{pmatrix}$$

$$\mathbf{V}_1=\begin{pmatrix}(1-\beta_2)\overline{\beta}\rho \\ \beta^{-1}\rho \\ \beta_1\beta^{-1}\end{pmatrix}A+\begin{pmatrix}(\beta_2-1)(\overline{\beta}^{-1}-\beta\rho) \\ (\rho-1)\overline{\beta}^{-1} \\ 0\end{pmatrix}\overline{A}, \tag{60}$$

$$\mathbf{V}_2=\begin{pmatrix}0.5(1-\beta_2)\overline{\beta}^2\rho^2 \\ 0.5\beta^{-2}\end{pmatrix}A+\begin{pmatrix}(\beta_2-1)\rho(1-\rho\beta\overline{\beta}) \\ -0.5\beta^{-2}\dfrac{\chi\beta_1}{1+\chi}\end{pmatrix}\overline{A},$$

$$\mathbf{V}_n=\begin{pmatrix}(1-\beta_2)\dfrac{\overline{\beta}^n\rho^n}{n} \\ \dfrac{[n-2+(n+2\chi)(\eta_1-\beta_1)]}{(n-2)\beta^{n-2}} \\ 0 \\ \dfrac{\beta_1}{(n-2+n\eta_1+2\chi\eta_1)\beta^{n-2}}\end{pmatrix}A+\begin{pmatrix}(\beta_2-1)\overline{\beta}^{n-2}\rho^n(\rho-\beta\overline{\beta}) \\ [n-2+(n+2\chi)(\eta_1-\beta_1)]\dfrac{1-\beta\overline{\beta}}{\beta^n}+\beta_1\dfrac{1}{\beta^n} \\ \left(\dfrac{\beta_2}{\eta_2}-1\right)\dfrac{\overline{\beta}^{n-2}\rho^{n-2}}{n-2} \\ \dfrac{(n-2)\beta_1(1-\beta\overline{\beta})}{(n-2+n\eta_1+2\chi\eta_1)\beta^n}+\left[1-\dfrac{\beta_1(n-2+2\chi\eta_1)}{(n+2\chi)\eta_1+n-2}\right]\dfrac{1}{n\beta^n}\end{pmatrix}\overline{A}$$

$$\mathbf{U}_1 = a\Pi \begin{bmatrix} \eta_2 \\ -1 \\ 0 \end{bmatrix}, \quad \mathbf{U}_3 = -a \begin{bmatrix} 0 \\ [1+(1+2\chi)\eta_1]\rho^{-1}\overline{\Pi}' \\ \dfrac{1}{\eta_2}\Pi' \\ (\rho^{-2}-\rho^{-1})\overline{\Pi}' \end{bmatrix},$$

$$\beta = \frac{e}{a}.$$

We can observe from the above that the three unknowns A_1, B_1, and $\phi'(0)$ are independent of the interface parameter χ, while all the other unknown coefficients in the Laurent series will depend on the interface parameter χ. It can be seen that the structure of the resulting linear algebraic equations obtained previously is simpler than that obtained by Luo and Chen [2] even the more complicated sliding interface is considered in our discussion. It suffices to solve independently a set of linear Eqs. (58) of second order, a set of linear Eqs. (57) of third order, and sets of linear Eqs. (59) of fourth order to uniquely determine all of the unknown coefficients in the Laurent series. The problem of an edge dislocation interacting with arbitrary number of concentric circular inclusions with sliding interfaces can also be solved through applying the decoupling relationship. In addition, the advantage of this methodology will become more apparent when the number of the interphase layers is increased. Here we point out that Ru [7] also observed a similar decoupling relationship when solving a circular inhomogeneity with stepwise graded interphase under thermomechanical loadings. Since the fact that he only considered remote uniform loadings, then it is only needed to treat a 2×2 real matrix and a 4×4 real matrix to obtain stress distribution within the circular inclusion. It is just due to the existence of this decoupling relationship in the concentric circular inclusion that the novel methodology, whose main feature is the transfer matrix method, proposed by Ru [7] can be utilized. If the circumferentially inhomogeneous sliding interface model (Ru [14]) is adopted, the decoupling strategy will become invalid due to the fact that the power series method will lead to a coupled infinite system of algebraic equations for the unknown coefficients.

3.4 Explicit Expressions for Muskhelishvili's Potentials $\phi_i(z)$ and $\psi_i(z)$. Up to now, all of the coefficients have been uniquely determined. Then explicit expressions for the holomorphic functions can be expressed as

$$\phi_1(z) = \phi_1'(0)z + \beta_1 \sum_{n=2}^{+\infty} \frac{n-2+2\chi\eta_1}{(n+2\chi)\eta_1+n-2}\left(A_n - A\frac{e^{-n}}{n}\right)z^n + \beta_1 \sum_{n=3}^{+\infty} \frac{n-2}{(n+2\chi)\eta_1+n-2}\left(a^{-2n+2}\overline{A_{-(n-2)}} + \frac{e^{-n+3}}{a^4}\overline{\gamma_1}A + \frac{e^{-n+2}}{a^2(n-2)}\overline{A}\right)z^n \quad z \in S_1 \tag{61a}$$

$$\psi_1(z) = -\beta_1 \sum_{n=1}^{+\infty} \frac{(n+2)(n+1+2\chi\eta_1)a^2}{(n+2+2\chi)\eta_1+n}\left(A_{n+2} - A\frac{e^{-n-2}}{n+2}\right)z^n - \beta_1 \sum_{n=1}^{+\infty} \frac{(n+1)(n+2)+2\chi}{(n+2+2\chi)\eta_1+n}\left(a^{-2n}\overline{A_{-n}} + \frac{e^{-n+1}}{a^2}\overline{\gamma_1}A + \frac{e^{-n}}{n}\overline{A}\right)z^n \quad z \in S_1 \tag{61b}$$

$$\phi_2(z) = \sum_{n=1}^{+\infty} B_{-n}z^{-n} + (1-\beta_2)\left[\sum_{n=1}^{+\infty} B_n z^n - A\ln(z-b^2/\overline{e}) + \frac{\gamma_2\overline{A}}{z-b^2/\overline{e}}\right] + \eta_2\Pi z + \frac{\overline{\Pi}'b^2}{z} + A\ln(z-e) \quad z \in S_2 \tag{62a}$$

$$\psi_2(z) = b^2\sum_{n=1}^{+\infty} nB_{-n}z^{-n-2} + (\beta_2-1)b^2 \times\left[\sum_{n=1}^{+\infty} nB_n z^{n-2} - \frac{A}{z(z-b^2/\overline{e})} - \frac{\gamma_2\overline{A}}{z(z-b^2/\overline{e})^2}\right] - \sum_{n=1}^{+\infty} b^{2n}\overline{B_n}z^{-n} + \left(\frac{\beta_2}{\eta_2}-1\right) \times\left[\sum_{n=1}^{+\infty} b^{-2n}\overline{B_{-n}}z^n + \overline{A}\ln(z-b^2/e)\right] + \overline{A}\ln(z-e) - \frac{\overline{e}\overline{A}}{z-e} + [(1-\eta_2)\Pi + Ae^{-1}]b^2z^{-1} + \overline{\Pi}'b^4z^{-3} + \frac{\Pi'}{\eta_2}z \quad z \in S_2 \tag{62b}$$

$$\phi_3(z) = \frac{\beta_2}{\eta_2}\left[\sum_{n=1}^{+\infty} B_{-n}z^{-n} + A\ln\frac{z-e}{z}\right] + pA\ln z + \Pi z + \frac{\overline{\Pi}'b^2}{\eta_2 z} \quad z \in S_3 \tag{63a}$$

$$\psi_3(z) = \frac{\beta_2}{\eta_2}b^2\left[\sum_{n=1}^{+\infty} nB_{-n}z^{-n-2} - \frac{eA}{z^2(z-e)}\right] + \beta_2\left[-\sum_{n=1}^{+\infty} b^{2n}\overline{B_n}z^{-n} + \overline{A}\ln\frac{z-e}{z} + \frac{\overline{\gamma_2}Ae^2}{b^2(z-e)}\right] + (\eta_2-1)\Pi b^2z^{-1} - pAb^2z^{-2} + \frac{\overline{\Pi}'b^4}{\eta_2}z^{-3} + \Pi' z + p\overline{A}\ln z \quad z \in S_3. \tag{63b}$$

The stress field can be obtained by substituting the above three pairs of complex potentials into (1). Or, equivalently,

$$\sigma_{xx}^{(k)} = \mathrm{Re}\{2\phi_k'(z) - \overline{z}\phi_k''(z) - \psi_k'(z)\}$$
$$\sigma_{yy}^{(k)} = \mathrm{Re}\{2\phi_k'(z) + \overline{z}\phi_k''(z) + \psi_k'(z)\} \quad \text{with} \quad k = 1,2,3 \tag{64}$$
$$\sigma_{xy}^{(k)} = \mathrm{Im}\{\overline{z}\phi_k''(z) + \psi_k'(z)\}.$$

The Peach-Koehler force, which is a measure of the force acting on the dislocation due to its interaction with the inclusion, can also be derived in terms of the above obtained complex potentials following the method adopted by Luo and Chen [2], Qaissaunee and Santare [3]. It is of interest to point out that when the three-phase composite system is only subjected to remote uniform loadings, then stress fields within the circular inclusion S_1 are coincident with the expressions for the stress fields derived with a single sharp interface described by an imperfect interface model (see Ru [14]), and are also coincident with the expressions for the stress fields for a circular inhomogeneity with an arbitrary *N*-layered perfectly bonded interphase (see Ru [7]).

4 Application

In this section, we assume that a crack lies on the real axis in the interval $[t_1, t_2]$ as shown in Fig. 2. In addition we assume that

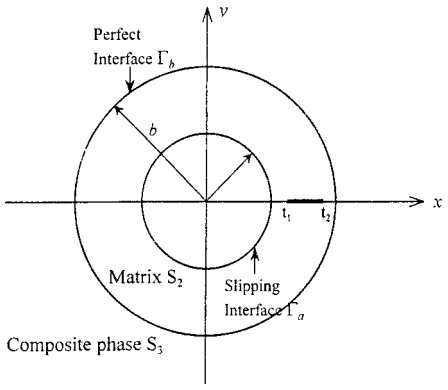


Fig. 2 A matrix crack in the three-phase cylinder composite model

the remote principal stresses are parallel to the two coordinate axes, then it suffices to only consider mode I crack. The resulting singular integral equations for unknown dislocation density $f(t)$ are

$$\int_{t_1}^{t_2} \frac{f(t)}{t-x}\,dt + \int_{t_1}^{t_2} K(t,x)f(t)dt = \frac{\kappa_2+1}{2G_2}\pi p(x) \quad t_1<x<t_2 \tag{65}$$

and the single-valuedness condition

$$\int_{t_1}^{t_2} f(t)dt = 0. \tag{66}$$

For brevity, we omit the detailed expressions for $K(t,x)$ and $p(x)$ in Eq. (65). The above singular integral equations are numerically solved based on Gaussian-Chebyshev integration formulas (Erdogan and Gupta, [21]). Figure 3 illustrates the variation of normalized stress intensity factor (SIF) $k_1^*=k_1/(\sigma_{yy}\sqrt{s})$, where s is the half-length of the crack $s=(t_2-t_1)/2$, at t_1 versus G_3/G_2 and imperfect parameter χ with $G_1=20G_2$, $\nu_1=\nu_2=\nu_3=0.3$, $a/b=1/8$, $t_1=3.5a$, $t_2=4.5a$ under uniaxial tension σ_{yy}^{∞}. Figure 4 shows the variation of $k_1^*(t_1)$ versus b/a and χ with $G_1=20G_2$, $G_3=10G_2$, $\nu_1=\nu_2=\nu_3=0.3$, $t_1=3.5a$, $t_2=4.5a$ under uniaxial tension σ_{yy}^{∞}. Figure 5 demonstrates the variation of $k_1^*(t_1)$ versus $\sigma_{xx}^{\infty}/\sigma_{yy}^{\infty}$ and χ with $G_1=20G_2$, $G_3=10G_2$, $\nu_1=\nu_2=\nu_3=0.3$, $a/b=1/8$, $t_1=3.5a$, $t_2=4.5a$ under biaxial tension $\sigma_{xx}^{\infty},\sigma_{yy}^{\infty}$. For comparison, the results calculated based on the two-phase model with a perfect interface (Erdogan et al. [22]) are also depicted in Fig. 3. A comparison of Fig. 3 with Fig. 3 in Luo and Chen [23] reveals that when $\chi\to\infty$ (i.e., perfect interface) our numerical results are in accordance with those obtained by Luo and Chen [23]. Similarly, a comparison of Fig. 4 with Fig. 4 in Luo and Chen [23] reveals that when $\chi\to\infty$ our numerical results are also in agreement with those obtained by Luo and Chen [23]. Then the analytical solution derived in this paper is verified from one aspect. From Figs. 3 to 5, we can readily draw a conclusion that the degree of imperfection has a significant influence on the stress intensity factor. It can also be found from Fig. 4 that the SIF is more sensitive to b/a when χ is decreased. It can be observed from Fig. 5 that there appears to be a linear relationship between SIF and ratio $\sigma_{xx}^{\infty}/\sigma_{yy}^{\infty}$ for a fixed χ; in addition, we find an interesting phenomenon that σ_{xx}^{∞} will exert no influence on SIF when $\chi=13.1$. Figure 6 illustrates the variations of $k_1^*(t_1)$ versus crack location t_1 and χ with $G_1=20G_2$, $G_3=10G_2$, $\nu_1=\nu_2=\nu_3=0.3$, $a/b=1/8$, $t_2=a+t_1$ under uniaxial tension σ_{yy}^{∞}. We find that the introduction of the sliding interface will increase the potential for the crack to extend when the crack is far away from the two interfaces Γ_a and Γ_b; the introduction of the sliding interface will retard the crack growth when the crack approaches one of the two circles. We illustrate in Fig. 7 the variations of vertical opening displacement $\Delta u_y G_2/(a\sigma_{yy}^{\infty})$ versus χ with $G_1=20G_2$, $G_3=10G_2$, $\nu_1=\nu_2=\nu_3=0.3$, $a/b=1/8$, $t_1=2.5a$, $t_2=5.5a$ under uniaxial tension σ_{yy}^{∞}. We find that Δu_y at the points on the left portion of the crack will decline with the increment of the degree of damage in the tangential direction on the interface Γ_a; while Δu_y at the points on the right portion of the crack will grow with the increment of the degree of damage in the tangential direction on the interface Γ_a.

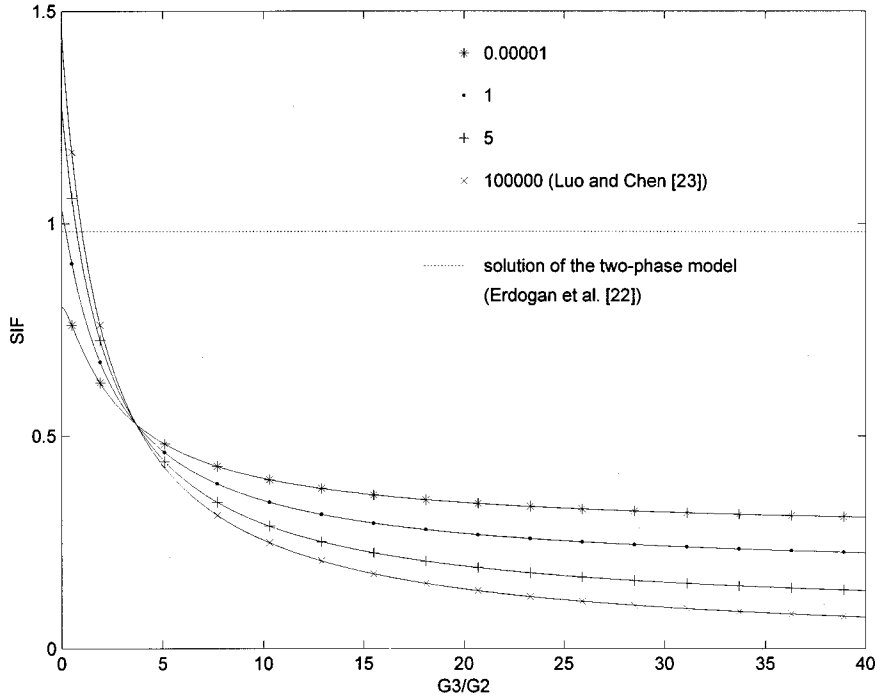


Fig. 3 Effects of G_3/G_2 and χ on stress intensity factor

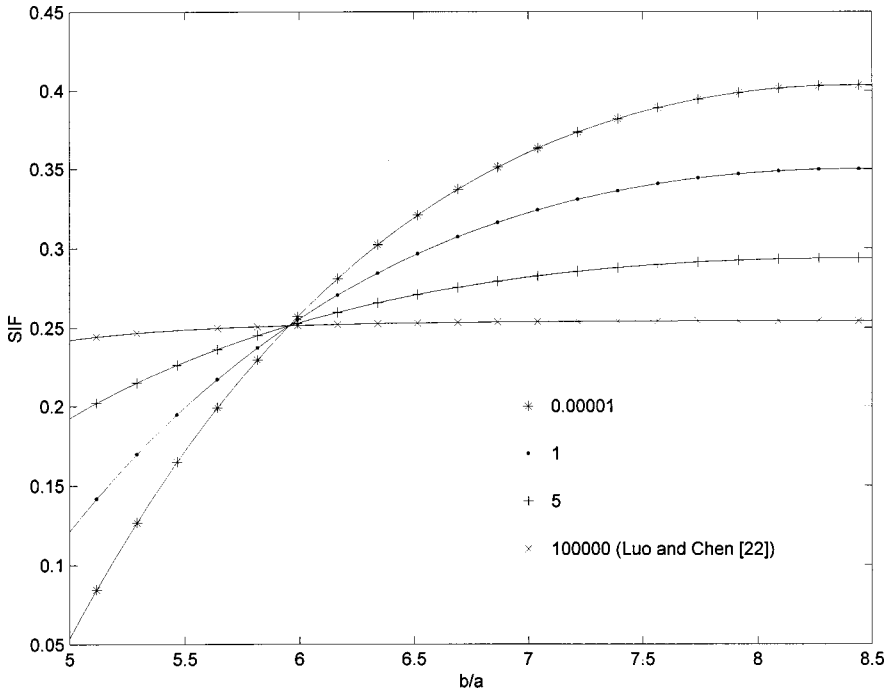


Fig. 4 Effects of b/a and χ on stress intensity factor

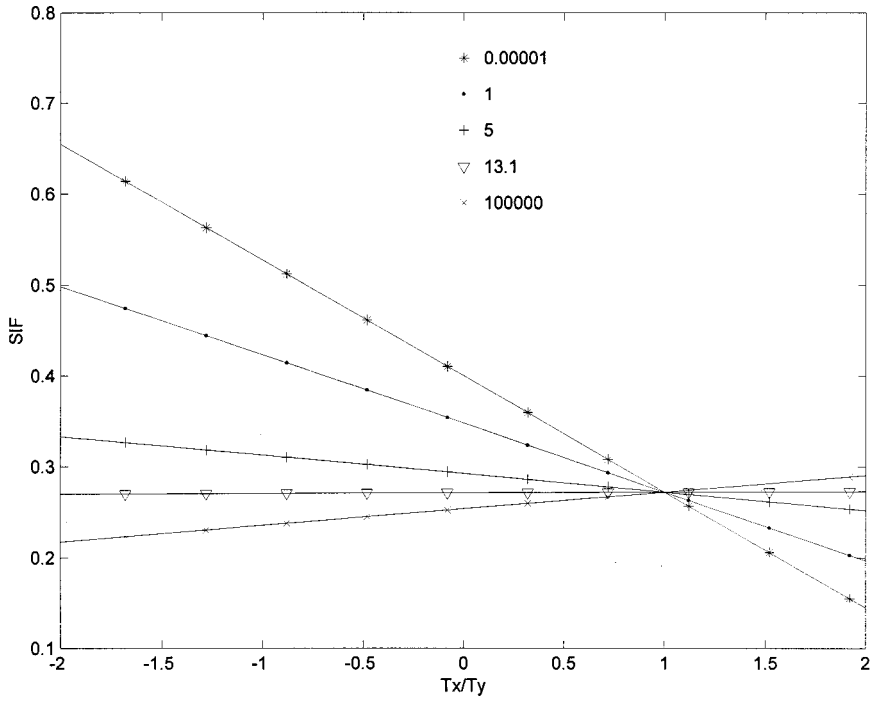


Fig. 5 Effects of $\sigma_{xx}^{\infty}/\sigma_{yy}^{\infty}$ and χ on stress intensity factor

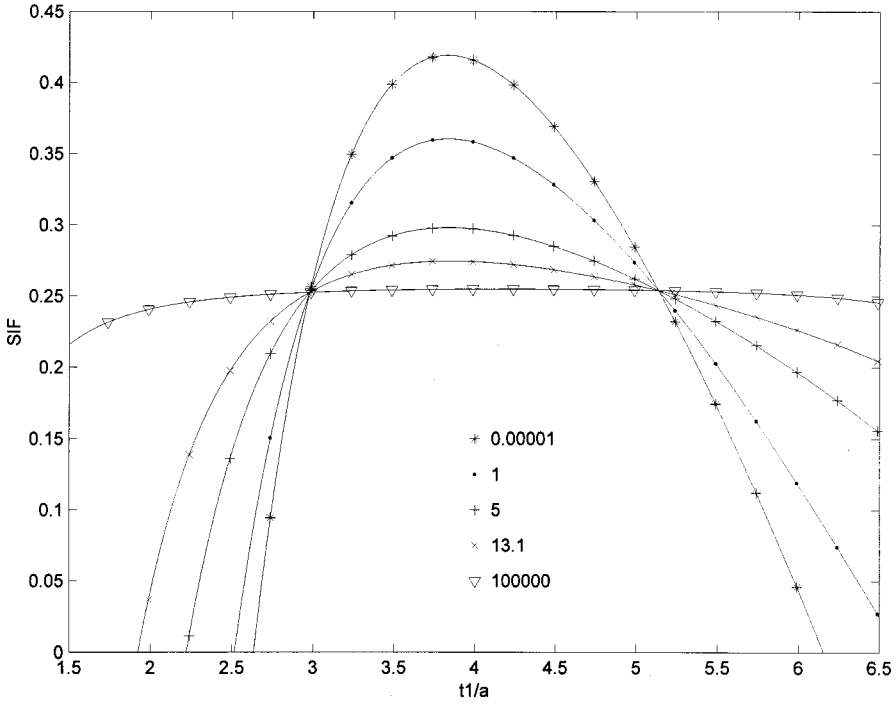


Fig. 6 Effects of t_1/a and χ on stress intensity factor

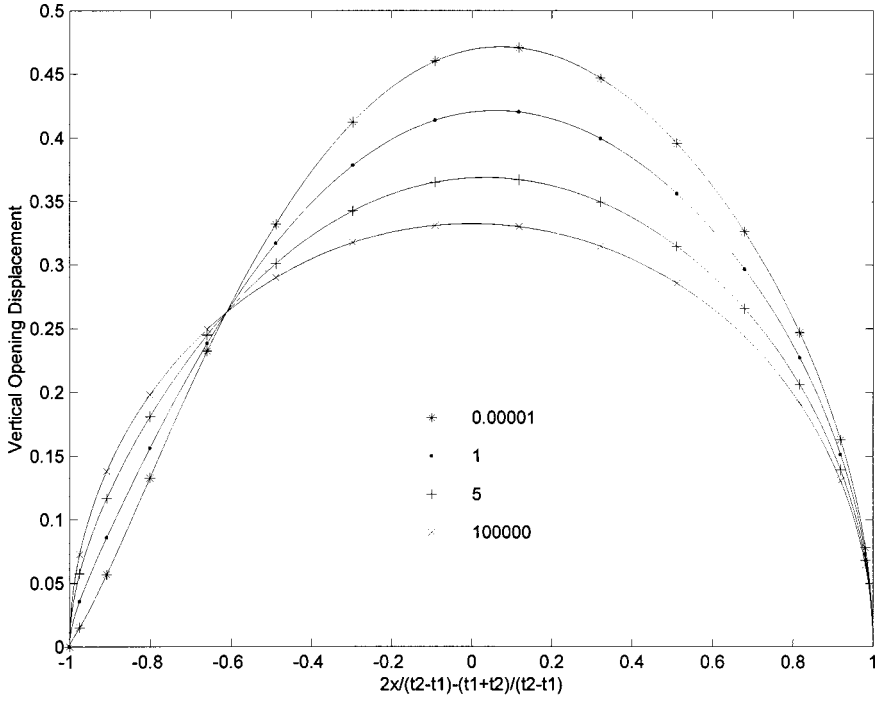


Fig. 7 Effects of χ on vertical opening displacement $\Delta u_y G_2/(a\sigma_{yy}^{\infty})$

5 Conclusion

The present research analytically investigates the interaction problem between an edge dislocation and a three-phase cylinder model with a sliding interface. In the three-phase model, the damage on the inclusion-matrix interface and interaction between inclusions can be simultaneously taken into account. Since the damage on the inclusion-matrix interface is assumed to be uniform, then the conventional power series method is still effective; since the three-phase composite system is considered, then only the power series method is appropriate to treat the problem. All of the coefficients in the Laurent series can be determined in a decoupled manner. This decoupling strategy will make one successfully solve the interaction problem between an edge dislocation and arbitrary number of concentric circular inclusions with sliding or perfect interfaces in an unbounded medium. In addition, the problem of a dislocation in the interfacial zone of the three-phase elliptical inclusion, which was not solved successfully by Qaissaunee and Santare [3], can also be solved by a similar approach

adopted in the present study. Based on the analytical solution derived in the present paper, we then investigate a crack in the matrix interacting with the sliding circular inclusion, which is an important mode of failure in composite materials, to illustrate the influence of the imperfect interface on the stress intensity factor.

Acknowledgment

This research was partially supported by the National Natural Science Foundation of China (10132010), and partially supported by the Doctorate Foundation of Xi'an Jiaotong University.

References

[1] Christensen, R. M., and Lo, K. H., 1979, ''Solution for Effective Shear Properties in Three Phase Sphere and Cylinder Models,'' J. Mech. Phys. Solids, **27**, pp. 315–330.
[2] Luo, H. A., and Chen, Y., 1991, ''An Edge Dislocation in a Three-Phase Composite Cylinder Model,'' ASME J. Appl. Mech., **58**, pp. 75–86.
[3] Qaissaunee, M. T., and Santare, M. H., 1995, ''Edge Dislocation Interacting With an Elliptical Inclusion Surrounded by an Interfacial Zone,'' Q. J. Mech. Appl. Math., **48**, pp. 465–482.
[4] Ru, C. Q., 1999, ''Three-Phase Elliptical Inclusions with Internal Uniform Hydrostatic Stresses,'' J. Mech. Phys. Solids, **47**, pp. 259–273.
[5] Xiao, Z. M., and Chen, B. J., 2000, ''A Screw Dislocation Interacting With a Coated Fiber,'' Mech. Mater., **32**, pp. 485–494.
[6] Xiao, Z. M., and Chen, B. J., 2001, ''On the Interaction Between an Edge Dislocation and a Coated Inclusion,'' Int. J. Solids Struct., **38**, pp. 2533–2548.
[7] Ru, C. Q., 1999, ''A New Method for an Inhomogeneity With Stepwise Graded Interphase Under Thermomechanical Loadings,'' J. Elast., **56**, pp. 107–127.
[8] Dundurs, J., and Gangadharan, A. C., 1969, ''Edge Dislocation Near an Inclusion With a Slipping Interface,'' J. Mech. Phys. Solids, **17**, pp. 459–471.
[9] Mura, T., and Furuhashi, R., 1984, ''The Elastic Inclusion With a Sliding Interface,'' ASME J. Appl. Mech., **51**, pp. 308–310.
[10] Mura, T., Jasiuk, I., and Tsuchida, E., 1985, ''The Stress Field of a Sliding Inclusion,'' Int. J. Solids Struct., **21**, pp. 1165–1179.
[11] Jasiuk, I., Tsuchida, E., and Mura, T., 1987, ''The Sliding Inclusion Under Shear,'' Int. J. Solids Struct., **23**, pp. 1373–1385.
[12] Lee, M., Jasiuk, I., and Tsuchida, E., 1992, ''The Sliding Circular Inclusion in an Elastic Half-Space,'' ASME J. Appl. Mech., **59**, pp. 57–64.
[13] Gharpuray, V. M., Dundurs, J., and Keer, L. M., 1991, ''A Crack Terminating at a Slipping Interface Between Two Materials,'' ASME J. Appl. Mech., **58**, pp. 960–963.
[14] Ru, C. Q., 1998, ''A Circular Inclusion With Circumferentially Inhomogeneous Sliding Interface in Plane Elastostatics,'' ASME J. Appl. Mech., **65**, pp. 30–38.
[15] Lubarda, V. A., and Markenscoff, X., 1999, ''Energies of Circular Inclusions: Sliding Versus Bonded Interfaces,'' Proc. R. Soc. London, Ser. A, **455**, pp. 961–974.
[16] Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff, Groningen.
[17] England, A. H., 1971, *Complex Variable Method in Elasticity*, John Wiley and Sons, New York.
[18] Honein, T., and Herrmann, G., 1990, ''On Bonded Inclusions With Circular or Straight Boundaries in Plane Elastostatics,'' ASME J. Appl. Mech., **57**, pp. 850–856.
[19] Worden, R. E., and Keer, L. M., 1991, ''Green's Functions for a Point Load and Dislocation in an Annular Region,'' ASME J. Appl. Mech., **58**, pp. 954–959.
[20] Chao, C. K., and Tan, C. J., 2000, ''On the General Solution for Annular Problems With a Point Heat Source,'' ASME J. Appl. Mech., **67**, pp. 511–518.
[21] Erdogan, F., and Gupta, G. D., 1972, ''On the Numerical Solution of Singular Integral Equations,'' Q. Appl. Math., **29**, pp. 525–534.
[22] Erdogan, F., Gupta, G. D., and Ratwani, M., 1974, ''Interaction Between a Circular Inclusion and an Arbitrarily Oriented Crack,'' ASME J. Appl. Mech., **41**, pp. 1007–1013.
[23] Luo, H. A., and Chen, Y., 1991, ''Matrix Cracking in Fiber-Reinforced Composite Materials,'' ASME J. Appl. Mech., **58**, pp. 846–848.

A Cracked Piezoelectric Material Strip Under Transient Thermal Loading

B. L. Wang

Y. W. Mai[1]

Department of Manufacturing Engineering
and Engineering Management,
City University of Hong Kong,
Tat Chee Avenue,
Kowloon Tong, Hong Kong

Considered in this paper is a piezoelectric material strip containing an embedded crack or an edge crack perpendicular to its boundaries. The problem is solved for a strip that is suddenly heated or cooled from the top surface. The bottom surface is assumed to be zero temperature or thermally insulated. First the transient temperature and the stress distributions in an uncracked strip are calculated. Then, these stresses are used as the crack surface traction with opposite sign to formulate the mixed boundary value problem. This leads to a singular integral equation of Cauchy-type, which is then solved numerically. The numerically results for stress intensity factor are computed as a function of the normalized time and the crack size. The temperature and the thermal stress distributions for the uncracked problem are also included. [DOI: 10.1115/1.1429935]

1 Introduction

Recently, piezoelectric materials and structures have received considerable attention due to the potential for designing adaptive structures that are both light in weight and possess adaptive control capabilities. Because of their brittleness, piezoelectric materials have a tendency to develop critical cracks during the manufacturing and poling process.

The problem of a strip with a crack perpendicular to its edges under thermal stress has been studied by many authors. In [1], an elastic strip with an embedded and edge crack under transient thermal stresses was investigated. The cracking in a plate of finite thickness due to sudden thermal transient stresses was considered for an edge crack ([2]). A cracked thin layer bonded to a very thick substrate under ''thermal shock'' conditions has been considered in [3]. In [4] an unconstrained elastic layer under statically self-equilibrium thermal or residual stresses is studied. Discussed in [5] is the thermal stress problem of a functionally graded plate as one of the advanced high-temperature materials capable of withstanding the extreme temperature environments, with and without an edge crack.

The mechanical, electrical, and thermal fields are coupled in most of the physical problems. Thermal effects could be important when piezoelectric materials have to undergo high or low temperature gradients. Accordingly, the analysis of the fracture process of piezoelectric materials could provide information to improve the design of piezoelectric devices operated in thermal environments. Isothermal behaviors of piezoelectric materials with cracks have attracted many research activities in recent years ([6–13]). Due to mathematical difficulties, thermal effects have not received high concern. The problem of a two-dimensional piezoelectric material with an elliptic cavity under a uniform heat flow was discussed in [14]. The limit situation was that the hole was reduced to a slit crack. In what follows, we are interested in solving the problem of a piezoelectric material strip under transient thermal stresses. The strip contains an internal crack or an edge crack perpendicular to its surface. The superposition technique is used to solve the governing equations. Once the temperature distribution is obtained, the thermal stresses of the uncracked problem are evaluated. The perturbation problem is formulated by using the thermal stresses from the uncracked problem with an opposite sign as the crack surface traction. The problem is reduced to a singular integral equation of Cauchy-type, which is solved numerically. The results are obtained for various parameters of the problem.

2 Temperature Distribution

The problem under consideration is a linear piezoelectroelastic medium containing a crack normal to the surfaces of the strip, as shown in Fig. 1. The crack problem may be solved by superposition. That is, one may first solve the thermal stress problem in the absence of any cracks and then solve the isothermal crack problem by using the equal and opposite of the thermal stresses as the crack surface traction. In the problem considered, the heat conduction is one-dimensional, a straight crack does not obstruct the heat flow in this arrangement, determination of the temperature distribution and the resulting thermal stresses would be quite straightforward and the related crack problem would be one of mode I.

Let us start with a one-dimensional thermal conduction problem. The temperature field T under consideration is one of one-dimensional and is governed by

$$k_3\frac{\partial^2 T(z,t)}{\partial z^2}=\rho c_v\frac{\partial T(z,t)}{\partial(t)}, \tag{1}$$

where k_3 is the coefficient of thermal conductivity, ρ the density, and c_v the specific heat. The corresponding thermal flows are

$$q_z=-k_3\partial T/\partial z. \tag{2}$$

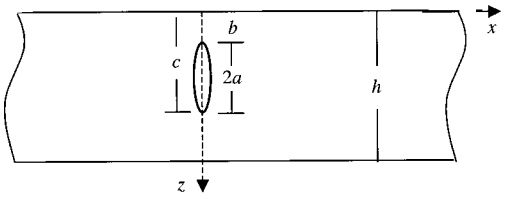


Fig. 1 A piezoelectric material strip with a crack ($a=(c-b)/2$), if b is larger than zero the crack is embedded in the strip, whereas b equals zero corresponds to an edge crack problem.

[1]Corresponding author.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 18, 2000; final revision, July 13, 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.

The initial conditions are specified to be

$$T(z,0)=0. \tag{3}$$

The following two kinds of boundary conditions are considered:

Boundary Condition a

$$T(0,t)=T_0H(t), \tag{4}$$

$$T(h,t)=0. \tag{5}$$

Boundary Condition b

$$T(0,t)=T_0H(t), \tag{6}$$

$$\partial T(h,t)/\partial z=0. \tag{7}$$

In Eqs. (4) and (5), $H(t)$ is the Heaviside function.

The thermal Eq. (1) can be solved by means of separate variables. Without going into details, the results are as follows:

Boundary Condition a

$$T(z)=-\sum_{m=1}^{\infty}\frac{2T_0}{m\pi}\sin\left(\frac{m\pi}{h}z\right)e^{-m^2\pi^2t/t_0}+\frac{T_0}{h}(h-z)$$

$$=T_0\sum_{m=1}^{\infty}\frac{2}{m\pi}(1-e^{-m^2\pi^2t/t_0})\sin\frac{m\pi}{h}z. \tag{8}$$

Boundary Condition b

$$T(z)=T_0\left(1-4\sum_{m=1,3,5}^{\infty}\frac{\sin(m\pi/2)}{m\pi}e^{-(m\pi/2)^2t/t_0}\cos\frac{m\pi}{2h}(h-z)\right), \tag{9}$$

where $t_0=\rho\ c_v\ h^2/k_3$.

If the plate is infinite (i.e., $h=\infty$) and the temperature and thermal flow vanish at infinity, the temperature field is given by the well-known solution

$$T(z)=T_0\,\mathrm{erfc}\left(\frac{z}{2\sqrt{(k_3/\rho c_v)t}}\right)$$

$$=\frac{2}{\pi}T_0\int_0^{\infty}(1-e^{-(k_3/\rho c_v)s^2t})\ \sin(sz)\frac{ds}{s}. \tag{10}$$

3 Thermal Stresses in the Uncracked Medium

Once $T(z)$ is known, the stresses and electric displacements can be found from

$$\begin{Bmatrix}\sigma_{xx}\\ \sigma_{zz}\\ \tau_{xz}\\ D_x\\ D_z\end{Bmatrix}=\begin{bmatrix}c_{11}&c_{13}&0&0&-e_{31}\\ c_{13}&c_{33}&0&0&-e_{33}\\ 0&0&c_{44}&-e_{15}&0\\ 0&0&e_{15}&\in_{11}&0\\ e_{31}&e_{33}&0&0&\in_{33}\end{bmatrix}\begin{Bmatrix}\varepsilon_{xx}\\ \varepsilon_{zz}\\ 2\varepsilon_{xz}\\ E_x\\ E_z\end{Bmatrix}-\begin{Bmatrix}\lambda_{11}\\ \lambda_{33}\\ 0\\ 0\\ -\beta_3\end{Bmatrix}T, \tag{11}$$

where σ_{ij} $(i,j=x,z)$ and D_i are stresses and electric displacements, respectively. λ_{ij}, β_i, c_{ij}, e_{ij}, and $\in_{ii}$ stand, respectively, for stress-temperature coefficients, pyroelectric constants, elastic constants, piezoelectric constants, and dielectric permittivities.

If we now assume that the plate is infinite along the x and z-axis, (i.e., $-\infty<(x,z)<\infty$), free of surface traction and electric charge at $z=0$ and $z=h$, it may be shown that

$$\sigma_{zz}=0,\quad \sigma_{xz}=0,\quad D_x=0,\quad D_z=0, \tag{12}$$

and all nonvanishing field quantities are independent of x and z. By solving ε_{zz} and E_z from the second and the fifth equations of (11), in terms of ε_{xx} and T, and then substituting into the first equation of (11), it is found that

$$\sigma_{xx}=\overline{c}_{11}\varepsilon_{xx}-\overline{\lambda}_{11}T, \tag{13}$$

where

$$\overline{c}_{11}=c_{11}-\frac{(c_{13}\in_{33}+e_{31}e_{33})c_{13}-(e_{31}c_{33}-c_{13}e_{33})e_{31}}{c_{33}\in_{33}+e_{33}^2}, \tag{14}$$

$$\overline{\lambda}_{11}=\lambda_{11}-\frac{(c_{13}\in_{33}+e_{31}e_{33})\lambda_{33}+(e_{31}c_{33}-c_{13}e_{33})\beta_3}{c_{33}\in_{33}+e_{33}^2}. \tag{15}$$

The compatibility condition that needs to be satisfied is

$$d^2\varepsilon_{xx}/dz^2=0, \tag{16}$$

giving

$$\varepsilon_{xx}=Az+B, \tag{17}$$

$$\sigma_{xx}=\overline{c}_{11}(Az+B)-\overline{\lambda}_{11}T, \tag{18}$$

where A and B are unknown constants to be obtained from the boundary conditions for the plate. For example,

- if the plate is unconstrained along its boundaries, we have
$$\int_0^h\sigma_{xx}(z)dz=0,\quad \int_0^h\sigma_{xx}(z)z\,dz=0; \tag{19}$$
- for a fully constrained plate $\varepsilon_{xx}=0$, giving $A=0$, $B=0$;
- and if the plate stretches uniformly but does not bend, then $A=0$, B is determined from the first one of Eqs. (19).

In the crack problem under considered the equal and opposite of the stress given by Eq. (18) will be used as the crack surface traction and the medium will be assume to be under plane-strain conditions.

4 The Crack Problem

The crack problem may be solved by treating the problem as isothermal and quasi-static, and by using the equal and opposite value of the thermal stress obtained from Eq. (18) as the crack surface traction. The plane piezoelasticity problem requires the solution of the following equilibrium equations:

$$\left.\begin{aligned}&c_{11}\frac{\partial^2u}{\partial x^2}+c_{44}\frac{\partial^2u}{\partial z^2}+(c_{13}+c_{44})\frac{\partial^2w}{\partial x\partial z}+(e_{31}+e_{15})\frac{\partial^2\phi}{\partial x\partial z}=0\\ &(c_{13}+c_{44})\frac{\partial^2u}{\partial x\partial z}+c_{44}\frac{\partial^2w}{\partial x^2}+c_{33}\frac{\partial^2w}{\partial z^2}+e_{15}\frac{\partial^2\phi}{\partial x^2}+e_{33}\frac{\partial^2\phi}{\partial z^2}=0\\ &(e_{31}+e_{15})\frac{\partial^2u}{\partial x\partial z}+e_{15}\frac{\partial^2w}{\partial x^2}+e_{33}\frac{\partial^2w}{\partial z^2}-\in_{11}\frac{\partial^2\phi}{\partial x^2}-\in_{33}\frac{\partial^2\phi}{\partial z^2}=0\end{aligned}\right\}, \tag{20}$$

subject to the following boundary conditions on the crack faces:

$$\sigma_{xx}(0,z)=\sigma_0(z), \quad c\geqslant z\leqslant b, \tag{21}$$

where $\sigma_0(y)$ is the equal and opposite value give in Eq. (18).

Out side of the crack, the displacement should be continuous. Therefore,

$$u(0,z)=0, \quad h\geqslant z\geqslant c, \quad b\geqslant z\geqslant 0. \tag{22}$$

Due to the symmetry of the problem with respect to the plane $x=0$, the following conditions should be satisfied:

$$\tau_{xz}(0,z)=0, \quad D_x(0,z)=0. \tag{23}$$

On the surfaces $z=0$ and $z=h$, the traction and electric displacement are considered to be

$$\tau_{xz}(x,0)=0, \quad \sigma_{zz}(x,0)=0, \quad D_z(x,0)=0, \tag{24}$$

$$\tau_{xz}(x,h)=0, \quad \sigma_{zz}(x,h)=0, \quad D_z(x,h)=0. \tag{25a}$$

Equations (24) and (25*a*) represent traction and electric charge free boundary conditions. If the strip stretches freely but without bending, the boundary condition (25a) on $z=h$ should be replaced by

$$\tau_{xz}(x,h)=0, \quad \partial w(x,h)/\partial x=0, \quad D_z(x,h)=0. \tag{25b}$$

Using the displacements and electric potential solutions shown in Appendix A, the stresses and electric displacements for the piezoelectric strip are given as

$$\tau_{xz}(x,z)=\frac{2}{\pi}\int_0^\infty s\sin sx\sum_{m=1}^{6} C_{1m}e^{s\lambda_m z}F_m ds + \sum_{n=1}^{3}\frac{1}{2\pi}\int_{-\infty}^{\infty}\xi e^{|\xi|\lambda_n x}D_{1n}G_n e^{-i\xi z}d\xi, \tag{26}$$

$$\sigma_{zz}(x,z)=\frac{2}{\pi}\int_0^\infty s\cos sx\sum_{m=1}^{6} C_{2m}e^{s\lambda_m z}F_m ds + \sum_{n=1}^{3}\frac{1}{2\pi}\int_{-\infty}^{\infty}\xi e^{|\xi|\lambda_n x}D_{2n}G_n e^{-i\xi z}d\xi, \tag{27}$$

$$D_z(x,z)=\frac{2}{\pi}\int_0^\infty s\cos sx\sum_{m=1}^{6} C_{3m}e^{s\lambda_m z}F_m ds + \sum_{n=1}^{3}\frac{1}{2\pi}\int_{-\infty}^{\infty}\xi e^{|\xi|\lambda_n x}D_{3n}G_n e^{-i\xi z}d\xi, \tag{28}$$

$$\sigma_{xx}(x,z)=\frac{2}{\pi}\int_0^\infty s\cos sx\sum_{m=1}^{6} C_{4m}e^{s\lambda_m z}F_m ds + \sum_{n=1}^{3}\frac{1}{2\pi}\int_{-\infty}^{\infty}\xi e^{|\xi|\lambda_n x}D_{4n}G_n e^{-i\xi z}d\xi, \tag{29}$$

$$D_x(x,z)=\frac{2}{\pi}\int_0^\infty s\sin sx\sum_{m=1}^{6} C_{5m}e^{s\lambda_m z}F_m ds + \sum_{n=1}^{3}\frac{1}{2\pi}\int_{-\infty}^{\infty}\xi e^{|\xi|\lambda_n x}D_{5n}G_n e^{-i\xi z}d\xi, \tag{30}$$

where the coefficients C_{jm} and D_{jn}, $(j=1,\ldots,5)$, are given in Appendix B.

The six unknowns for F_m and three unknowns for G_n will be determined from the boundary conditions (21)–(25). To do this, we introduce a displacement discontinuity function $g(z)$ along the cracked plane

$$\partial u(0,z)/\partial z=g(z). \tag{31}$$

Boundary condition (22) requires that

$$g(z)=0, \quad h\geqslant z>c, \quad b>z\geqslant 0. \tag{32}$$

In the case of an embedded crack, $g(z)$ satisfies the following single-valued condition

$$\int_b^c g(z)dz=0. \tag{33}$$

In the case of an edge crack, the single-valued condition is $g(y)=0$ at $(x=0,y=h)$.

After substituting from Eq. (*A*5) in Appendix A into Eq. (31), and by inverting the Fourier integrals we find

$$\sum_{n=1}^{3} B_{1n}G_n=i\xi^{-1}\int_b^c g(r)e^{i\xi r}dr. \tag{34}$$

Conditions (23) require that

$$\sum_{n=1}^{3} D_{1n}G_n=0 \tag{35}$$

$$\sum_{n=1}^{3} D_{5n}G_n=0. \tag{36}$$

From Eqs (34)–(36), G_n can be determined as

$$G_n(\xi)=i\xi^{-1}b_n(\xi)\int_b^c g(r)e^{i\xi r}dr, \quad n=1,2,3 \tag{37}$$

where the coefficients b_n $(n=1,2,3)$ are given in Appendix C.

By applying Eqs. (26)–(28) and (37) to the homogeneous boundary conditions (24) and (25*a*), and using the well-known integral [15]

$$\exp(s\lambda_n x)=\frac{2}{\pi}\int_0^\infty \frac{\xi\sin(\xi x)}{s^2\lambda_n^2+\xi^2}d\xi, \quad \mathrm{Re}(\lambda_n)<0,$$

we obtain

$$\sum_{m=1}^{6} C_{1m}e^{s\lambda_m h}F_m=\frac{i}{2\pi s}\int_b^c g(r)f_1(s,r)dr, \tag{38}$$

$$\sum_{m=1}^{6} C_{1m}F_m=\frac{i}{2\pi s}\int_b^c g(r)f_2(s,r)dr, \tag{39}$$

$$\sum_{m=1}^{6} C_{2m}e^{s\lambda_m h}F_m=\frac{i}{2\pi s}\int_b^c g(r)f_3(s,r)dr, \tag{40}$$

$$\sum_{m=1}^{6} C_{2m}F_m=\frac{i}{2\pi s}\int_b^c g(r)f_4(s,r)dr, \tag{41}$$

$$\sum_{m=1}^{6} C_{3m}e^{s\lambda_m h}F_m=\frac{i}{2\pi s}\int_b^c g(r)f_5(s,r)dr, \tag{42}$$

$$\sum_{m=1}^{6} C_{3m}F_m=\frac{i}{2\pi s}\int_b^c g(r)f_6(s,r)dr, \tag{43}$$

where the coefficient f_m, $(m=1,\ldots,6)$, are given in Appendix D. In the case of boundary condition (25*b*) considered, Eq. (40) has a different form as follows:

$$\sum_{m=1}^{6} A_{2m}e^{s\lambda_m h}F_m=\frac{i}{2\pi s}\int_b^c \overline{f_3}(s,r)g(r)dr, \tag{44}$$

where $\overline{f_3}$ is shown in Appendix D.

The linear algebraic Eqs. (38)–(43) can be used to determined all the remaining unknowns F_m, $(m=1,\ldots,6)$, in terms of g, the results are

$$F_m=\frac{i}{2\pi s}\int_b^c c_m(s,r)g(r)dr, \quad m=1,\ldots,6, \tag{45}$$

where the coefficients $c_m(s,r)$ are given in Appendix E.

The as-yet-unknown function g is then determined by applying (29) to crack surface boundary condition (21)

$$\int_b^c [h_1(z,r)+h_2(z,r)]g(r)dr=\sigma_0(z), \quad c\geqslant z\geqslant b, \tag{46}$$

where the kernels h_1 and h_2 are given by

$$h_1(z,r)=\frac{i}{\pi^2}\int_0^\infty \Lambda_1(s,z,r)ds, \tag{47}$$

$$h_2(z,r)=\frac{1}{2\pi i}\int_{-\infty}^\infty \Lambda_2(\xi)e^{i\xi(r-z)}d\xi, \tag{48}$$

with

$$\Lambda_1(s,z,r)=\sum_{m=1}^6 C_{4m}e^{s\lambda_m z}c_m(s,r), \tag{49}$$

$$\Lambda_2(\xi)=-\sum_{n=1}^3 D_{4n}b_n. \tag{50}$$

It can be shown that the value of $\Lambda_2(\xi)$ is $\Lambda_2(\xi)=\mathrm{sgn}(\xi)\Lambda_0$, where Λ_0 is a constant which depend only on the material properties. From Eq. (48) it is found that

$$h_2(z,r)=\frac{\Lambda_0}{\pi}\frac{1}{r-z}. \tag{51}$$

The singular integral equation for g can be derived by combining Eqs. (46)–(51), the result is

$$\frac{\Lambda_0}{\pi}\int_b^c \frac{1}{r-z}g(r)dr+\int_b^c h_1(z,r)g(r)dr=\sigma_0(z), \quad c\geqslant z\geqslant b, \tag{52}$$

which can be reduced to standard form by setting

$$(z,r)=\frac{c-b}{2}(\overline{z,r})+\frac{c+b}{2}. \tag{53}$$

The result is

$$\frac{\Lambda_0}{\pi}\int_{-1}^1 \frac{1}{\overline{r}-\overline{z}}g(r)d\overline{r}+a\int_{-1}^1 h_1(z,r)g(r)d\overline{r}=\sigma_0(z),$$
$$1\geqslant\overline{z}\geqslant -1. \tag{54}$$

Note that in this problem time t enters into the analysis through σ_0 only, and Eq. (54) must be solved for each value of t separately.

5 Thermal Stress Intensity Factor

The integral equation (54) has a Cauchy-type kernel. The crack-tip behavior can be characterized by a standard square-root singularity. Consequently, the solution of Eq. (54) may be expressed as

$$g(r)=\frac{1}{\sqrt{1-\overline{r}^2}}\sum_{n=1}^\infty a_nT_n(\overline{r}), \tag{55}$$

for the embedded crack, and

$$g(r)=\frac{1}{\sqrt{1-\overline{r}}}\sum_{n=0}^\infty a_nT_n(\overline{r}), \tag{56}$$

for the edge crack problem, where T_n is the Chebyshev polynomial of the first kind and $a_0,a_1,\ldots$ are unknown constants.

Applying Eqs. (55)–(56) to Eq. (54) and following the procedure outlined in [4], for the embedded crack problem the singular integral equation (54) is reduced to

$$\Lambda_0\sum_{n=1}^\infty a_nU_{n-1}(\overline{z})+a\sum_{n=1}^\infty a_n\int_{-1}^1 h_1(r,z)\frac{T_n(\overline{r})}{\sqrt{1-\overline{r}^2}}d\overline{r}=\sigma_0(z), \tag{57}$$

where U_n is the Chebyshev polynomial of the second kind.

Similarly, for the edge crack problem the result is

$$\sum_{n=0}^\infty a_n\left(\frac{\Lambda_0}{\pi}\frac{\log|B(\overline{z})|}{\sqrt{1-\overline{z}}}T_n(\overline{z})+\frac{\Lambda_0}{\pi}\int_{-1}^1\frac{T_n(\overline{r})-T_n(\overline{z})}{\sqrt{1-\overline{r}}(\overline{r}-\overline{z})}d\overline{r}\right.$$
$$\left.+\frac{c}{2}\int_{-1}^1 h_1(z,r)\frac{T_n(\overline{r})}{\sqrt{1-\overline{r}}}d\overline{r}\right)=\sigma_0(z), \tag{58}$$

where

$$B(\overline{z})=\frac{1+\sqrt{(1-\overline{z})/2}}{1-\sqrt{(1-\overline{z})/2}}. \tag{59}$$

The linear algebraic equations (57)–(58) can be solved by truncating the series and using a collocation technique ([4]). For the embedded crack problem, the stress intensity factor K ahead of the crack tips at $z=b$ and $z=c$ can be defined and calculated as

$$K(b)=(\sqrt{2[b-z]})_{z\to b^-}\sigma_{xx}(z)=\Lambda_0\sqrt{a}\sum_{n=1}^\infty (-1)^n a_n, \tag{60}$$

$$K(c)=(\sqrt{2[z-c]})_{z\to c^+}\sigma_{xx}(z)=-\Lambda_0\sqrt{a}\sum_{n=1}^\infty a_n. \tag{61}$$

For the edge crack problem the stress intensity factor is obtained as

$$K(c)=(\sqrt{2[z-c]})_{z\to c^+}\sigma_{xx}(z)=-\Lambda_0\sqrt{c}\sum_{n=0}^\infty a_n. \tag{62}$$

6 Numerical Results

Suppose the piezoelectric material strip undergoes a sudden temperature increase or decrease on the top surface, the bottom surface is assumed to be free of temperature (boundary condition a) or thermally insulated (boundary condition b). The numerical results of the temperature $T(y,t)$, the thermal stress $\sigma_{xx}(x,t)$, and the stress intensity factor K have been obtained for a centrally cracked strip, and for an edge cracked strip with various crack lengths. The following material properties is used:

$$c_{11}=7.41\times10^{10}\mathrm{N/m^2}, \quad c_{33}=8.36\times10^{10}\mathrm{N/m^2},$$
$$c_{12}=4.52\times10^{10}\mathrm{N/m^2},$$
$$c_{13}=3.93\times10^{10}\mathrm{N/m^2}, \quad c_{44}=1.32\times10^{10}\mathrm{N/m^2},$$
$$\lambda_{11}=0.621\times10^{6}\mathrm{N/m^2K}, \quad \lambda_{33}=0.551\times10^{6}\mathrm{N/m^2K},$$
$$e_{31}=-0.160\mathrm{C/m^2}, \quad e_{33}=0.347\mathrm{C/m^2}, \quad e_{15}=-0.138\mathrm{C/m^2},$$
$$\in_{11}=0.826\times10^{-10}\mathrm{C/Vm}, \quad \in_{33}=0.903\times10^{-10}\mathrm{C/Vm},$$
$$\beta_3=-2.94\times10^{-6}\mathrm{C/Km^2}.$$

Time is represented through the dimensionless Fourier number defined by

$$F_0=t/t_0=tk_3/\rho c_v h^2. \tag{63}$$

The Temperature and Thermal Stress Distributions. Assume the top surface of the strip is heating from initial temperature zero to T_0 suddenly, the calculated temperature distributions are shown in Fig. 2. The two kinds of boundary conditions on the other surface give the same temperature distributions at the early time. As a result, the stresses at the early time have the same values for the different boundary condition on the other surface. These stresses are plotted in Fig. 3. Note that in the heating problem the temperature T is positive. As time goes to infinity the temperature varies linearly from T_0 on the top surface to zero on the bottom surface for the case one. For the second kind of thermal boundary condition, as time goes to infinity the temperature approaches T_0 throughout the strip. In the problem under consideration, when F_0 is larger than 0.05, the two boundary conditions give different temperature and then thermal stress distribution that are plotted in Fig. 4.

Plotted in Fig. 5 are the variations of thermal stress with time at different position of the strip. The figure clearly shows that as t increases there is a general peak for each curve. The peak values of the thermal stress for various positions are reached at some time after the thermal shock. After that time the stress decreases with time. The thermal stress $\sigma_{xx}(x,t)$, approaches zero when t goes to infinity, means that the thermal stress occurs only in transient state. Since the thermal stress σ_{xx} is statically self-equilibrating, large compressive stress occurs near the surfaces, and the tensile stress appears inside the strip.

The Embedded Crack Problem. Suppose the crack is located in the center of the strip. As mentioned above, if the strip is heated suddenly on its surface, large tensile stress will occur inside the strip. Accordingly, the crack will open up and the stress intensity factor will be positive. Figures 6 and 7 show, respectively, the thermal stress intensity $K(c)$ for the upper crack tip and $K(b)$ for the lower crack tip. From the figures it may be observed that for a given crack length generally K first increases, goes through a maximum, and then decreases as t increases. As was expected, the stress intensity factor approaches zero when t goes to infinity.

The peak values of the thermal stress intensity factor for different crack length are depicted in Fig. 8. The K values are normalized such that they do not depend on crack length a. It is found that the stress intensity factors have lager values at the crack tip near the hotter surface than those at the tip near the cooler surface.

The Edge Crack Problem. If the top surface of the strip is cooled to T_0 suddenly, the strip will undergo tensile stresses near the surfaces. Some numerical results of the thermal stress intensity factor are shown in Figs. 9–11 and Table 1 for different boundary conditions on the other surface. Similar to the embedded crack problem, the stress intensity factor occurs at transient state. The steady thermal stress intensity factor is zero. Figure 11 plots the peak values of the thermal stress intensity factor with crack length. The results are also tabulated in Table 1. It can be shown that the peak value of the thermal stress intensity factor increases with crack length to a maximum value and then decreases. If we

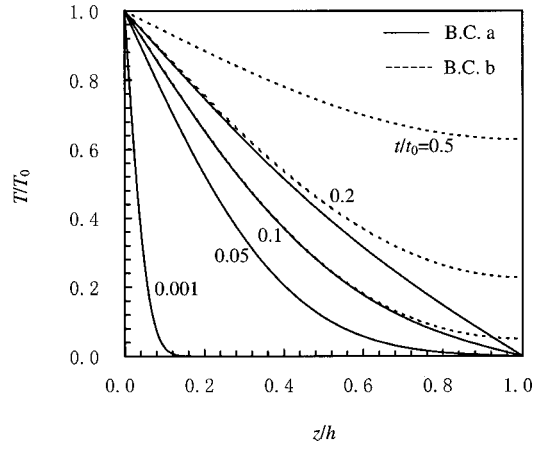


Fig. 2 Transient temperature distributions for different boundary conditions (B, C)

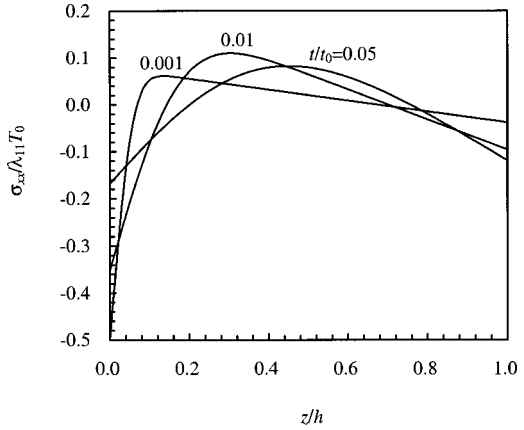


Fig. 3 Transient thermal stresses distributions in an uncracked strip

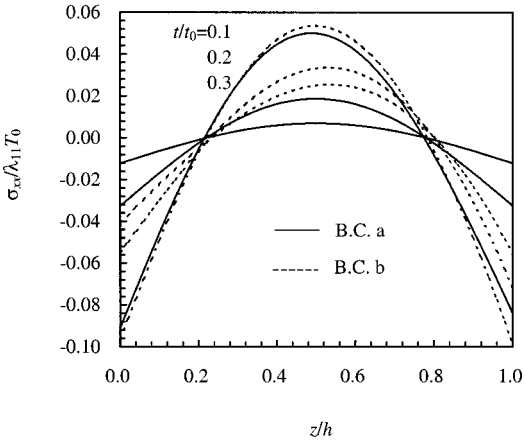


Fig. 4 Transient thermal stresses distribution in an uncracked strip (continued)

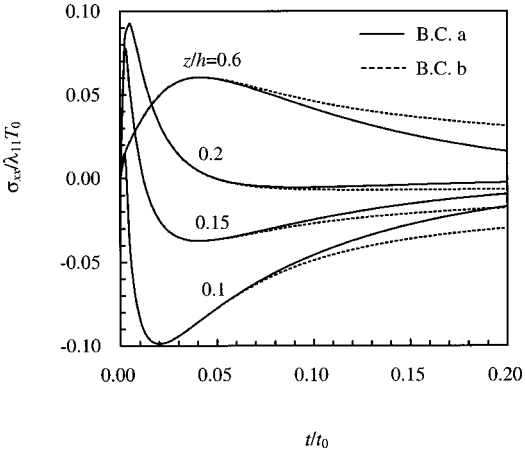


Fig. 5 Transient thermal stresses against time for different position

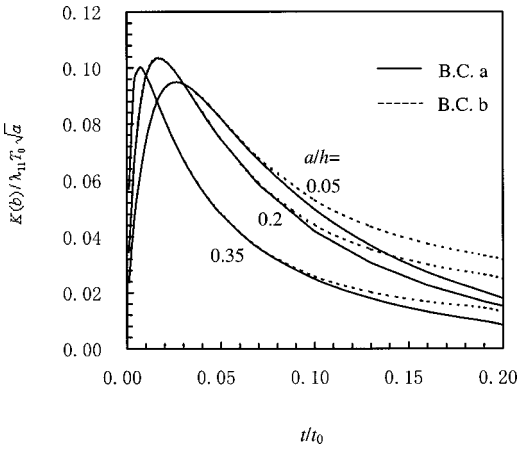


Fig. 6 Transient thermal stress intensity factors $K(b)$ at the embedded crack tip for different crack length ($b+c=h$)

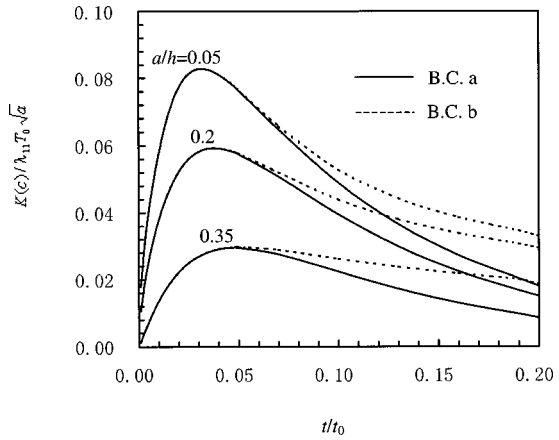


Fig. 7 Transient thermal stress intensity factors $K(c)$ at the embedded crack tip for different crack length ($b+c=h$)

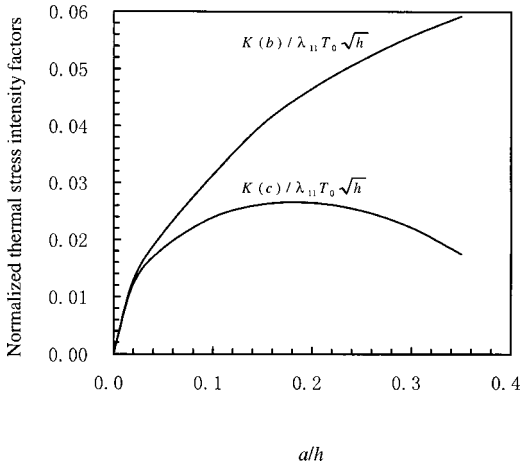


Fig. 8 Peak values of the thermal stress intensity factor at the embedded crack tips for different crack length ($b+c=h$)

assume a pre-existing crack, its length being set as c_0, it will start propagating when the thermal stress intensity $K(c_0)$ reaches the fracture toughness K_{Ic}. The crack will propagate continually at least to a new length c_1, where the thermal stress intensity factor $K(c_1)$ again reaches the fracture toughness K_{Ic}.

The results shown in Fig. 11 indicate that the thermal stress intensity factor with increasing crack length goes through a maximum with the region of crack instability bounded by two values of crack length. This behavior is in contrast with the Griffith criterion for constant load, where crack instability is bounded by one value of crack length, crack propagation being catastrophic whenever the crack length exceeds this value.

7 Conclusions

In conclusion, both an embedded crack and edge crack in a piezoelectric material strip under transient thermal loading have been investigated theoretically. The top surface of the strip undergoes a sudden heating (for the embedded crack problem) or cooling (for the edge crack problem). Two kinds of boundary conditions on the bottom surface of the strip are considered, i.e., zero temperature and zero thermal flow on the bottom surface. For the geometry and thermal conditions under consideration, the thermal stresses and hence the thermal stress intensity factors occur only at transient state. The steady values of the thermal stress intensity factor are always zeroes. At the early time of heating or cooling, the two boundary conditions on the other surface give the same results. A large difference is observed when heating or cooling time becomes large.

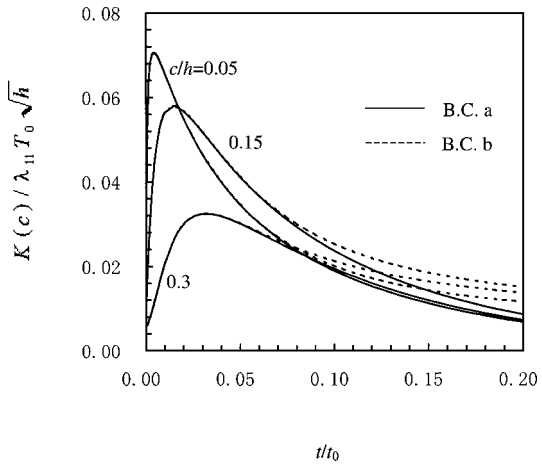


Fig. 9 Transient thermal stress intensity factors against time for the edge crack problem

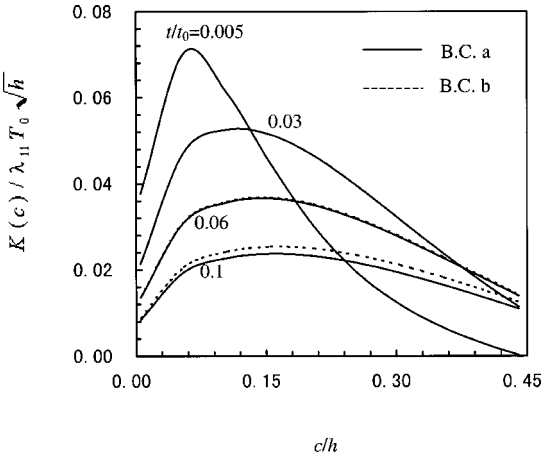


Fig. 10 Transient thermal stress intensity factors with crack length for the edge crack problem

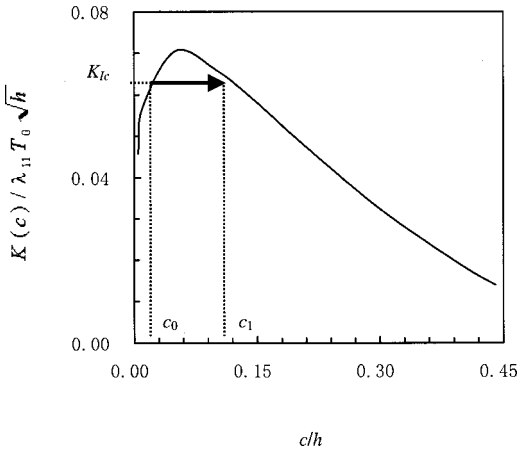


Fig. 11 Peak values of the thermal stress intensity factor for the edge crack problem

Table 1 The peak values of thermal stress intensity factor for an edge cracked strip with different crack length under sudden cooling

	$K(c)/\lambda_{11}T_0\sqrt{h}$	
c/h	Case 1	Case 2
0.005	0.0458	0.0458
0.01	0.0568	0.0568
0.05	0.0705	0.0705
0.1	0.0660	0.0660
0.15	0.0581	0.0581
0.2	0.0492	0.0492
0.3	0.0324	0.0324
0.4	0.0186	0.0186
0.44	0.0140	0.0142

Acknowledgments

Part of this work was finished during B. L. Wang's postdoctoral stay at Shizuoka University, Japan, supported by the Japan Society for the Promotion of Science (JSPS). B. L. Wang thanks Prof. Naotake Noda of Shizuoka University for many valuable discussions with him.

Appendix A

Due to the symmetry with respective to $x=0$ plane, the solutions to the governing Eqs. (20) for displacements and electric potential can be expressed in terms of unknowns $F_m(s)$ and $G_n(s)$ in the following form:

$$(u\ \ w\ \ \phi)^T=\frac{2}{\pi}\int_0^\infty (A_{1m}\sin sx\ \ A_{2m}\cos sx\ \ A_{3m}\cos sx)^T \times F_m e^{s\lambda_m z}ds+\frac{1}{2\pi} \times\int_{-\infty}^{\infty}(B_{1n}\ \ B_{2n}\ \ B_{3n})^T G_n e^{|\xi|\lambda_n x}e^{-i\xi z}d\xi, \quad (A1)$$

where λ_m, and λ_n are eigenvalues, while (A_{1m},A_{2m},A_{3m}) and (B_{1n},B_{2n},B_{3n}) are eigenvectors of the following characteristic equations:

$$\begin{bmatrix} c_{11}-c_{44}\lambda_m^2 & (c_{13}+c_{44})\lambda_m & (e_{31}+e_{15})\lambda_m \\ (c_{13}+c_{44})\lambda_m & c_{33}\lambda_m^2-c_{44} & e_{33}\lambda_m^2-e_{15} \\ (e_{31}+e_{15})\lambda_m & e_{33}\lambda_m^2-e_{15} & \in_{11}-\in_{33}\lambda_m^2 \end{bmatrix}\begin{Bmatrix} A_{1m} \\ A_{2m} \\ A_{3m} \end{Bmatrix}=0, \quad (A2)$$

$$\begin{bmatrix} c_{44}-c_{11}\lambda_n^2 & i\frac{|s|}{s}(c_{13}+c_{44})\lambda_n & i\frac{|s|}{s}(e_{31}+e_{15})\lambda_n \\ i\frac{|s|}{s}(c_{13}+c_{44})\lambda_n & c_{33}-c_{44}\lambda_n^2 & e_{33}-e_{15}\lambda_n^2 \\ i\frac{|s|}{s}(e_{31}+e_{15})\lambda_n & e_{33}-e_{15}\lambda_n^2 & -\in_{33}+\in_{11}\lambda_n^2 \end{bmatrix} \times\begin{Bmatrix} B_{1n} \\ B_{2n} \\ B_{3n} \end{Bmatrix}=0. \quad (A3)$$

Since the stresses and electric displacements must vanish as x approaches infinity, the eigenvalues λ_n, $(n=1,2,3)$, are selected such that

$$\mathrm{Re}(\lambda_n)<0. \quad (A4)$$

There are six roots for λ_m, and three roots for λ_n. In terms of these eigenvalues, the solution to Eq. (20) is

$$u(x,z)=\frac{2}{\pi}\sum_{m=1}^{6}\int_0^\infty \sin sx A_{1m}e^{s\lambda_m z}F_m ds + \frac{1}{2\pi}\sum_{n=1}^{3}\int_{-\infty}^{\infty} e^{|\xi|\lambda_n x}B_{1n}e^{-i\xi z}G_n d\xi, \quad (A5)$$

$$w(x,z)=\frac{2}{\pi}\sum_{m=1}^{6}\int_0^\infty \cos sx A_{2m}e^{s\lambda_m z}F_m ds + \frac{1}{2\pi}\sum_{n=1}^{3}\int_{-\infty}^{\infty} e^{|\xi|\lambda_n x}B_{2n}e^{-i\xi z}G_n d\xi, \quad (A6)$$

$$\phi(x,z)=\frac{2}{\pi}\sum_{m=1}^{6}\int_0^\infty \cos sx A_{3m}e^{s\lambda_m z}F_m ds + \frac{1}{2\pi}\sum_{n=1}^{3}\int_{-\infty}^{\infty} e^{|\xi|\lambda_n x}B_{3n}e^{-i\xi z}G_n d\xi, \quad (A7)$$

where F_m and G_n, $(m=1,\ldots,6,n=1,2,3)$, can be determined from the boundary conditions.

Appendix B

$$C_{1m}=c_{44}\lambda_m A_{1m}-c_{44}A_{2m}-e_{15}A_{3m}$$
$$D_{1n}=-ic_{44}B_{1n}+c_{44}\,\mathrm{sgn}(\xi)\lambda_n B_{2n}+e_{15}\,\mathrm{sgn}(\xi)\lambda_n B_{3n}$$
$$C_{2m}=c_{13}A_{1m}+c_{33}\lambda_m A_{2m}+e_{33}\lambda_m A_{3m}$$
$$D_{2n}=c_{13}\,\mathrm{sgn}(\xi)\lambda_n B_{1n}-ic_{33}B_{2n}-ie_{33}B_{3n}$$
$$C_{3m}=e_{31}A_{1m}+e_{33}\lambda_m A_{2m}-\in_{33}\lambda_m A_{3m}$$
$$D_{3n}=e_{31}\,\mathrm{sgn}(\xi)\lambda_n B_{1n}-ie_{33}B_{2n}+i\in_{33}B_{3n}$$
$$C_{4m}=c_1 A_{1m}+c_{13}\lambda_m A_{2m}+e_{31}\lambda_m A_{3m}$$
$$D_{4n}=c_{11}\,\mathrm{sgn}(\xi)\lambda_n B_{1n}-ic_{13}B_{2n}-ie_{31}B_{3n}$$
$$C_{5m}=e_{15}\lambda_m A_{1m}-e_{15}A_{2m}+\in_{11}A_{3m}$$
$$D_{5n}=-ie_{15}B_{1n}+e_{15}\,\mathrm{sgn}(\xi)\lambda_n B_{2n}-\in_{11}\,\mathrm{sgn}(\xi)\lambda_n B_{3n}$$

Appendix C

$$b_1=(D_{12}D_{53}-D_{13}D_{52})/\Delta$$

$$b_2=(D_{51}D_{13}-D_{11}D_{53})/\Delta$$

$$b_3=(D_{11}D_{52}-D_{51}D_{12})/\Delta$$

$$\Delta=B_{11}(D_{12}D_{53}-D_{13}D_{52})+B_{12}(D_{51}D_{13}-D_{11}D_{53})+B_{13}(D_{11}D_{52}-D_{51}D_{12})$$

Appendix D

$$f_1(s,r)=-\sum_{n=1}^{3}\int_{-\infty}^{\infty}\frac{s}{\xi^2\lambda_n^2+s^2}D_{1n}b_n e^{i\xi(-h+r)}d\xi$$

$$=\pi i\sum_{n=1}^{3}\mathrm{Im}(D_{1n}b_n e^{s(h-r)/\lambda_n}/\lambda_n)$$

$$f_2(s,r)=-\sum_{n=1}^{3}\int_{-\infty}^{\infty}\frac{s}{\xi^2\lambda_n^2+s^2}D_{1n}b_n e^{i\xi r}d\xi$$

$$=\pi i\sum_{n=1}^{3}\mathrm{Im}(D_{1n}b_n e^{sr/\lambda_n}/\lambda_n)$$

$$f_3(s,r)=\sum_{n=1}^{3}\int_{-\infty}^{\infty}\frac{|\xi|\lambda_n}{\xi^2\lambda_n^2+s^2}D_{2n}b_n e^{i\xi(-h+r)}d\xi$$

$$=-\pi i\sum_{n=1}^{3}\mathrm{Re}(D_{2n}b_n e^{s(h-r)/\lambda_n}/\lambda_n)$$

$$f_4(s,r)=\sum_{n=1}^{3}\int_{-\infty}^{\infty}\frac{|\xi|\lambda_n}{\xi^2\lambda_n^2+s^2}D_{2n}b_n e^{i\xi r}d\xi$$

$$=\pi i\sum_{n=1}^{3}\mathrm{Re}(D_{2n}b_n e^{sr/\lambda_n}/\lambda_n)$$

$$f_5(s,r)=\sum_{n=1}^{3}\int_{-\infty}^{\infty}\frac{|\xi|\lambda_n}{\xi^2\lambda_n^2+s^2}D_{3n}b_n e^{i\xi r(-h+r)}d\xi$$

$$=-\pi i\sum_{n=1}^{3}\mathrm{Re}(D_{3n}b_n e^{s(h-r)/\lambda_n}/\lambda_n)$$

$$f_6(s,r)=\sum_{n=1}^{3}\int_{-\infty}^{\infty}\frac{|\xi|\lambda_n}{\xi^2\lambda_n^2+s^2}D_{3n}b_n e^{i\xi r}d\xi$$

$$=\pi i\sum_{n=1}^{3}\mathrm{Re}(D_{3n}b_n e^{sr/\lambda_n}/\lambda_n)$$

$$\overline{f}_3=\sum_{n=1}^{3}\int_{-\infty}^{\infty}\frac{s\ \mathrm{sgn}(\lambda)\lambda_n}{\xi^2\lambda_n^2+s^2}B_{2n}b_n e^{i\xi(r-h)}d\xi$$

$$=-\pi i\sum_{n=1}^{3}\mathrm{Im}(B_{2n}b_n e^{s(h-r)/\lambda_n})$$

Obviously, f_i and $\overline{f}_{3u}$, $i=1,\ldots,6$, are pure imaginaries. When evaluating the above infinity integrals we have used the theory of residues.

Appendix E

The coefficients $c_m(s,r)$, $(m=1,2,\ldots,6)$, are

$$(c_1\ \ c_2\ \ c_3\ \ c_4\ \ c_5\ \ c_6)^T=\mathbf{G}^{-1}(f_1\ \ \ f_2\ \ f_3\ \ f_4\ \ f_5\ \ f_6)^T,$$

where

$$\mathbf{G}=\begin{bmatrix} C_{11}e^{s\lambda_1 h} & C_{12}e^{i\lambda_2 h} & C_{13}e^{s\lambda_3 h} & C_{14}e^{s\lambda_4 h} & C_{15}e^{s\lambda_5 j} & C_{16}e^{s\lambda_6 h} \\ C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21}e^{s\lambda_1 h} & C_{22}e^{s\lambda_2 h} & C_{23}e^{s\lambda_3 h} & C_{24}e^{s\lambda_4 h} & C_{25}e^{s\lambda_5 h} & C_{26}e^{s\lambda_6 h} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31}e^{s\lambda_1 h} & C_{32}e^{s\lambda_2 h} & C_{33}e^{s\lambda_3 h} & C_{34}e^{s\lambda_4 h} & C_{35}e^{s\lambda_s h} & C_{36}e^{s\lambda_6 h} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \end{bmatrix}.$$

If a bending free boundary condition (25*b*) is considered, f_3 should be replaced with $\overline{f}_3$, and C_{2j}, $j=1,\ldots,6$, in the third row of **G** should be replaced, respectively, by A_{2j}, $j=1,\ldots,6$.

References

[1] Rizk, A. A., and Radwan, S. F., 1992, "Transient Thermal Stress Problem for a Cracked Semi-infinite Medium," J. Therm. Stresses, **15**, pp. 451–468.

[2] Rizk, A. A., and Radwan, S. F., 1993, "Fracture of a Plate Under Transient Thermal Stresses," J. Therm. Stresses, **16**, pp. 79–102.

[3] Rizk, A. A., and Erdogan, F., 1989, "Cracking of Coated Materials Under Transient Thermal Stresses," J. Therm. Stresses, **12**, pp. 125–168.

[4] Erdogan, F., and Wu, B. H., 1996, "Crack Problem in FGM Layers Under Thermal Stresses," J. Therm. Stresses, **19**, pp. 237–265.

[5] Noda, N., 1997, "Thermal Stresses Intensity Factor for Functionally Gradient Plate With an Edge Crack," J. Therm. Stresses, **20**, pp. 373–387.

[6] Shindo, Y., Watanabe, K., and Narita, F., 2000, "Electroelastic Analysis of a Piezoelectric Ceramic Strip With a Central Crack," Int. J. Eng. Sci., **38**, pp. 1–19.

[7] Wang, T. C., 2000, "Analysis of Strip Electric Saturation Model of Crack Problem in Piezoelectric Materials," Int. J. Solids Struct., **37**, pp. 6031–6049.

[8] Chao, L.-P., and Huang, J. H., 2000, "On a Piezoelectric Material Containing a Permeable Elliptical Crack," Int. J. Solids Struct., **37**, pp. 5161–5176.

[9] Pak, Y. E., 1990, "Crack Extension Force in a Piezoelectric Material," J. Appl. Mech., **57**, pp. 647–653.

[10] Kwon, S. M., and Lee, K. Y., 2000, "Analysis of Stress and Electric Fields in a Rectangular Piezoelectric Body With a Center Crack Under Anti-plane Shear Loading," Int. J. Solids Struct., **37**, pp. 4859–4869.

[11] Lee, K. Y., Lee, W. G., and Pak, Y. E., 2000, "Interaction Between a Semi-infinite Crack and a Screw Dislocation in a Piezoelectric Material," ASME J. Appl. Mech., **67**, pp. 165–170.

[12] Calderon-Moreno, J. M., 2001, "Stress Induce Domain Switching of PZT in Compression Tests," Mater. Sci. Eng., Ser. A., **315**, pp. 227–230.

[13] Wang, B. L., and Noda, N., 2000, "Mixed Mode Crack Initiation in Piezoelectric Ceramic Strip," Theor. Appl. Fract. Mech., **34**, pp. 34–47.

[14] Lu, P., Tan, M. J., and Liew, K. M., 1998, "Piezothermoelastic Analysis of a Piezoelectric Material With an Elliptic Cavity Under Uniform Heat Flow," Arch. Appl. Mech., **68**, pp. 719–733.

[15] Gradshteyn, I. S., and Ryzhik, I. M., eds., 1980, *Tables of Integral, Series, and Products*, Academic Press, San Diego, CA, p. 406.

Journal of Applied Mechanics

Brief Notes

A Brief Note is a short paper that presents a specific solution of technical interest in mechanics but which does not necessarily contain new general methods or results. A Brief Note should not exceed 1500 words *or equivalent* (a typical one-column figure or table is equivalent to 250 words; a one line equation to 30 words). Brief Notes will be subject to the usual review procedures prior to publication. After approval such Notes will be published as soon as possible. The Notes should be submitted to the 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.

Comparison of the Two Formulations of *w-u-v* and *w-F* in Nonlinear Plate Analysis

J. Lee
Air Force Research Laboratory (VASS), Wright-Patterson AFB, OH 45433
e-mail: Jon.lee@wpafb.af.mil

In a moderately large deflection plate theory of von Karman and Chu-Herrmann, one may consider thin-plate equations of either the transverse and in-plane displacements, w-u-v formulation, or the transverse displacement and Airy function, w-F formulation. Under the Galerkin procedure, we examine if the modal equations of two plate formulations preserve the Hamiltonian property which demands energy conservation in the conservative limit of no damping and forcing. In the w-F formulation, we have shown that modal equations are Hamiltonian for the first four symmetric modes of a simply-supported plate. In contrast, the corresponding modal equations of w-u-v formulation do not exhibit the Hamiltonian property when a finite number of sine terms are included in the in-plane displacement expansions.
[DOI: 10.1115/1.1458556]

1 Problem Statement

In the plate theory of von Karman and Chu-Herrmann ([1]), one considers not only the inertia and plate bending due to transverse displacement w, but also a first-order effect of the in-plane displacements u and v, giving rise to membrane stretching. This is w-u-v formulation of a moderately large deflection plate theory ([2]). However, when the inertia of in-plane motion is neglected in a thin plate, one can replace the two in-plane equations for u and v by a compatibility relation for the Airy function F, hence the alternate w-F formulation. Although the two plate formulations are completely equivalent in theory, they are not in practice. This is because w, u, v, and F are all reduced to a finite degree-of-freedom representation by whatever the means used for computation, i.e., the finite difference, finite element, or Galerkin representation. Hence, it is a matter of uncertainty to claim equivalence of the two plate formulations in actual computations.

Under a Galerkin representation, we compare modal equations for the first four symmetric modes of a simply-supported plate obtained under the two plate formulations. This calls for further clarifications. First, why use the Galerkin representation? For simply-supported and clamped plates, we have previously carried out the Galerkin procedure by symbolic manipulation software, such as, MATHEMATICA™ ([3]), though others could have been used as well. This permits us to readily verify equality of mathematical expressions, aside from the bookkeeping aid in tracking myriad expansion terms arising from the geometric nonlinearity ([4]). Second, why a simply-supported plate? It is known for a rectangular simply-supported plate that plate modes can be constructed by a product of simply-supported-beam modes which are simple sine functions, along each plate coordinate. Here, by plate modes we mean the plate bending modes, which are the eigenfunctions of biharmonic operator. The last and crucial issue is, how are we to compare the w-F and w-u-v formulations? In a recent dissertation work, Geveci [5] has attempted such a comparison on a clamped plate by visually examining numerically generated trajectories of the transverse displacement under the two plate formulations. Unfortunately, his outcome is inconclusive for there is apparently no frame of reference for comparison. Since plate equations are embodied by the existence of Hamiltonian, the sum of kinetic and strain (potential) energies ([1]), modal equations must also preserve the Hamiltonian property at any level of modal truncation. Hence, a metric is established for comparison of the plate formulations.

2 Synopsis of Plate Equations for a Moderately Large Deflection

Following Chu and Herrmann [1], we begin with the force balance equations integrated across the plate thickness in the usual Cartesian coordinates (x,y,z) with displacements (u,v,w)

$$\frac{\partial N_x}{\partial x}+\frac{\partial N_{xy}}{\partial y}=\rho h\frac{\partial^2 u}{\partial t^2},\quad \frac{\partial N_{xy}}{\partial x}+\frac{\partial N_y}{\partial y}=\rho h\frac{\partial^2 v}{\partial t^2}, \tag{1}$$

$$\frac{\partial^2 M_x}{\partial x^2}+2\frac{\partial^2 M_{xy}}{\partial x\partial y}+\frac{\partial^2 M_y}{\partial y^2}+\frac{\partial}{\partial x}\left(N_x\frac{\partial w}{\partial x}\right)+\frac{\partial}{\partial y}\left(N_y\frac{\partial w}{\partial y}\right)+\frac{\partial}{\partial x}\left(N_{xy}\frac{\partial w}{\partial y}\right)+\frac{\partial}{\partial y}\left(N_{xy}\frac{\partial w}{\partial x}\right)=\rho h\frac{\partial^2 w}{\partial t^2}. \tag{2}$$

Only the inertial terms are retained in the right-hand sides, and excluded are the external forcing, rotatory effects, and transverse shear forces (see Eqs. (1)-(54)–(1)-(60) of Chia [2]). Although they are general statements about the forces exerted on a plate

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 2, 2000; final revision, Nov. 18, 2001. Associate Editor: R. C. Benson.

segment of thickness h and mass density ρ, the phenomenology of plate theory enters through the stress resultants (N_x, N_y, N_{xy}) and bending and twisting moments (M_x, M_y, M_{xy}). For the von Karman-Chu-Herrmann plate equation, we have

$$N_x = \frac{Eh}{1-\nu^2}(\varepsilon_x + \nu\varepsilon_y), \quad N_y = \frac{Eh}{1-\nu^2}(\nu\varepsilon_x + \varepsilon_y),$$

$$N_{xy} = \frac{Eh}{2(1+\nu)}\,\varepsilon_{xy},$$

where

$$\varepsilon_x = \frac{\partial u}{\partial x} + \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2, \quad \varepsilon_y = \frac{\partial v}{\partial y} + \frac{1}{2}\left(\frac{\partial w}{\partial y}\right)^2,$$

$$\varepsilon_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} + \frac{\partial w}{\partial x}\frac{\partial w}{\partial y},$$

E is Young's modulus of elasticity and ν Poisson's ratio. Further,

$$M_x = -D\left(\frac{\partial^2 w}{\partial x^2} + \nu\frac{\partial^2 w}{\partial y^2}\right), \quad M_y = -D\left(\nu\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right),$$

$$M_{xy} = -D(1-\nu)\frac{\partial^2 w}{\partial x \partial y},$$

where $D = Eh^3/12(1-\nu^2)$.

Chu and Herrmann [1] have also presented total energy of a plate with sides L_x and L_y as sum of the kinetic energy

$$U_k = \frac{1}{2}\rho h \int_0^{L_x}\int_0^{L_y}(\dot{u}^2 + \dot{v}^2 + \dot{w}^2)dxdy, \tag{3}$$

where the overhead dot denotes d/dt, and the strain energy which for convenience we split into bending

$$U_b = -\frac{1}{2}\int_0^{L_x}\int_0^{L_y}\left(M_x\frac{\partial^2 w}{\partial x^2} + 2M_{xy}\frac{\partial^2 w}{\partial x \partial y} + M_y\frac{\partial^2 w}{\partial y^2}\right)dxdy, \tag{4}$$

and membrane stretching

$$U_m = \frac{1}{2}\int_0^{L_x}\int_0^{L_y}(N_x\varepsilon_x + N_{xy}\varepsilon_{xy} + N_y\varepsilon_y)dxdy. \tag{5}$$

We now define the kinetic potential $\Pi = U_k - U_b - U_m$. By Hamilton's variational principle, it is standard to rederive (1) and (2) from stationary variation $\delta\Pi = 0$ with respect to $(\delta u, \delta v, \delta w)$. Hence, the energy conservation is fully embodied by plate Eqs. (1) and (2).

For a thin plate, i.e., $h/L_x \ll 1$, one may neglect the inertia in (1), hence

$$\frac{\partial N_x}{\partial x} + \frac{\partial N_{xy}}{\partial y} = 0, \quad \frac{\partial N_{xy}}{\partial x} + \frac{\partial N_y}{\partial y} = 0. \tag{6}$$

This brings about amazing simplifications. First, (6) is satisfied by a function $F(x,y)$ defined by

$$N_x = \frac{\partial^2 F}{\partial y^2}, \quad N_y = \frac{\partial^2 F}{\partial x^2}, \quad N_{xy} = -\frac{\partial^2 F}{\partial x \partial y}. \tag{7}$$

Second, by virtue of (6) a half the terms involving (N_x, N_y, N_{xy}) drops out of (2), hence

$$\rho h\frac{\partial^2 w}{\partial t^2} + D\left(\frac{\partial^4 w}{\partial x^4} + 2\frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^2 w}{\partial y^4}\right) = N_x\frac{\partial^2 w}{\partial x^2} + 2N_{xy}\frac{\partial^2 w}{\partial x \partial y} + N_y\frac{\partial^2 w}{\partial y^2}. \tag{8}$$

3 Two Formulations of the Plate Equations

To proceed further, it is more convenient to express equations of Section 2 in a dimensionless form by introducing new variables denoted by an overhead bar,

$$(\bar{x}, \bar{y}) = (x,y)\frac{1}{L_x}, \quad \bar{w} = \frac{w}{h}, \quad (\bar{u}, \bar{v}) = (u,v)\frac{L_x}{h^2}, \quad \bar{t} = t\frac{\pi^2}{L_x^2}\sqrt{\frac{D}{\rho h}},$$

$$(\bar{N}_x, \bar{N}_y, \bar{N}_{xy}) = (N_x, N_y, N_{xy})\frac{L_x^2}{12D},$$

$$(\bar{U}_k, \bar{U}_b, \bar{U}_m) = (U_k, U_b, U_m)\frac{rL_x^2}{\pi^4 h^2 D}, \tag{9}$$

$$(\bar{\varepsilon}_x, \bar{\varepsilon}_y, \bar{\varepsilon}_{xy}) = (\varepsilon_x, \varepsilon_y, \varepsilon_{xy})\frac{L_x^2}{h^2}, \quad \bar{F} = F\frac{1}{12D},$$

where $r = L_x/L_y$. We point out that $(\bar{u}, \bar{v})$ are also scaled by the factor (h/L_x).

The *w-u-v* Formulation. After introducing (9) into (8), we drop all the overhead bars in the resulting equation

$$\frac{\partial^2 w}{\partial t^2} + \frac{1}{\pi^4}\left(\frac{\partial^4 w}{\partial x^4} + 2r^2\frac{\partial^4 w}{\partial x^2 \partial y^2} + r^4\frac{\partial^2 w}{\partial y^4}\right) = \frac{12}{\pi^4}\left(N_x\frac{\partial^2 w}{\partial x^2} + 2rN_{xy}\frac{\partial^2 w}{\partial x \partial y} + r^2 N_y\frac{\partial^2 w}{\partial y^2}\right). \tag{10}$$

We must now remember that not only (10) but all the subsequent equations are also henceforth dimensionless. The dimensionless (6) becomes

$$\frac{\partial N_x}{\partial x} + r\frac{\partial N_{xy}}{\partial y} = 0, \quad \frac{\partial N_{xy}}{\partial x} + r\frac{\partial N_y}{\partial y} = 0, \tag{11}$$

where

$$N_x = \frac{\partial u}{\partial x} + \nu r\frac{\partial v}{\partial y} + \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2 + \frac{1}{2}\nu r^2\left(\frac{\partial w}{\partial y}\right)^2,$$

$$N_y = \nu\frac{\partial u}{\partial x} + r\frac{\partial v}{\partial y} + \frac{1}{2}\nu\left(\frac{\partial w}{\partial x}\right)^2 + \frac{1}{2}r^2\left(\frac{\partial w}{\partial y}\right)^2,$$

and

$$N_{xy} = \frac{1}{2}(1-\nu)\left(r\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} + r\frac{\partial w}{\partial x}\frac{\partial w}{\partial y}\right).$$

Writing out (11) in (u,v) yields the in-plane displacement equations

$$\frac{\partial^2 u}{\partial x^2} + d_1 r^2\frac{\partial^2 u}{\partial y^2} + d_2 r\frac{\partial^2 v}{\partial x \partial y} + \frac{\partial w}{\partial x}\frac{\partial^2 w}{\partial x^2} + d_2 r^2\frac{\partial w}{\partial y}\frac{\partial^2 w}{\partial x \partial y} + d_1 r^2\frac{\partial w}{\partial x}\frac{\partial^2 w}{\partial y^2} = 0,$$

$$d_1\frac{\partial^2 v}{\partial x^2} + r^2\frac{\partial^2 v}{\partial y^2} + d_2 r\frac{\partial^2 u}{\partial x \partial y} + d_1 r\frac{\partial w}{\partial y}\frac{\partial^2 w}{\partial x^2} + d_2 r\frac{\partial w}{\partial x}\frac{\partial^2 w}{\partial x \partial y} + r^3\frac{\partial w}{\partial y}\frac{\partial^2 w}{\partial y^2} = 0, \tag{12}$$

where $d_1 = (1-\nu)/2$ and $d_2 = (1+\nu)/2$. Compare (10) and (12) with Eq. (2.19c) and (2.15a,b) of Geveci [5].

Under $\dot{u} = \dot{v} = 0$, (3) reduces to the dimensionless kinetic energy

$$U_k = \frac{1}{2}\int_0^1\int_0^1 \dot{w}^2 dxdy. \tag{13}$$

Also, (4) and (5) in dimensionless form are

$$U_b=\frac{1}{2\pi^4}\int_0^1\int_0^1\left\{\left(\frac{\partial^2 w}{\partial x^2}\right)^2+2\nu r^2\frac{\partial^2 w}{\partial x^2}\frac{\partial^2 w}{\partial y^2}+r^4\left(\frac{\partial^2 w}{\partial y^2}\right)^2 +2(1-\nu)r^2\left(\frac{\partial^2 w}{\partial x\partial y}\right)^2\right\}dxdy, \tag{14}$$

$$U_m=\frac{6}{\pi^4}\int_0^1\int_0^1\left(\varepsilon_x^2+2\nu\varepsilon_x\varepsilon_y+\varepsilon_y^2+\frac{(1-\nu)}{2}\varepsilon_{xy}^2\right)dxdy, \tag{15}$$

where

$$\varepsilon_x=\frac{\partial u}{\partial x}+\frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2,\quad \varepsilon_y=r\frac{\partial v}{\partial y}+\frac{1}{2}r^2\left(\frac{\partial w}{\partial y}\right)^2,\quad \text{and}$$

$$\varepsilon_{xy}=r\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}+r\frac{\partial w}{\partial x}\frac{\partial w}{\partial y}.$$

The *w-F* Formulation. By virtue of dimensionless (7)

$$N_x=r^2\frac{\partial^2 F}{\partial y^2},\quad N_y=\frac{\partial^2 F}{\partial x^2},\quad N_{xy}=-r\frac{\partial^2 F}{\partial x\partial y}, \tag{16}$$

we express (8) in w and F instead,

$$\frac{\partial^2 w}{\partial t^2}+\frac{1}{\pi^4}\left(\frac{\partial^4 w}{\partial x^4}+2r^2\frac{\partial^4 w}{\partial x^2\partial y^2}+r^4\frac{\partial^2 w}{\partial y^4}\right) =\frac{12r^2}{\pi^4}\left(\frac{\partial^2 w}{\partial x^2}\frac{\partial^2 F}{\partial y^2}-2\frac{\partial^2 w}{\partial x\partial y}\frac{\partial^2 F}{\partial x\partial y}+\frac{\partial^2 w}{\partial y^2}\frac{\partial^2 F}{\partial x^2}\right), \tag{17}$$

together with the compatibility relation

$$r^{-2}\frac{\partial^4 F}{\partial x^4}+2\frac{\partial^4 F}{\partial x^2\partial y^2}+r^2\frac{\partial^2 F}{\partial y^4}=(1-\nu^2)\left(\left(\frac{\partial^2 w}{\partial x\partial y}\right)^2-\frac{\partial^2 w}{\partial x^2}\frac{\partial^2 w}{\partial y^2}\right). \tag{18}$$

Again, (17) and (18) are comparable with Eqs. (2.21) and (2.23) of Geveci [5]. In contrast to (15), the membrane strain energy is now expressed in F instead,

$$U_m=\frac{6}{\pi^4(1-\nu^2)}\int_0^1\int_0^1\left\{\left(\frac{\partial^2 F}{\partial x^2}\right)^2-2\nu r^2\frac{\partial^2 F}{\partial x^2}\frac{\partial^2 F}{\partial y^2}+r^4\left(\frac{\partial^2 F}{\partial y^2}\right)^2 +2(1+\nu)r^2\left(\frac{\partial^2 F}{\partial x\partial y}\right)^2\right\}dxdy. \tag{19}$$

4 Modal Equations Under the Galerkin Representation

In theory, the two plate formulations are completely equivalent in that by (16) one can freely interchange in-plane displacements with the Airy function. It is, however, not so in practice. This is because u, v, w, and F are approximated by a finite degree-of-freedom representation by whatever the means that one uses for computation, i.e., the finite difference, finite element, or Galerkin representation. To test the interchangeability argument under Galerkin representation, we consider the first four symmetric plate modes

$$w=\sum_{n=1,3}\sum_{m=1,3}a_{n,m}(t)\varphi_n(x)\varphi_m(y), \tag{20}$$

where $\varphi_i(\xi)=\sqrt{2}\sin(i\pi\xi)$ are the orthonormalized simply-supported beam modes. It is indeed fortuitous that $\varphi_n(x)\varphi_m(y)$ are the plate modes of a simply-supported plate. We point out that the clamped-beam modes cannot form clamped-plate modes *a la* (20). This does not, of course, mean that clamped plate is not amenable to the Galerkin procedure; it is that results are somewhat more complicated to present than the simply-supported plate case.

In this note, we restrict ourselves to the immovable plate edges; i.e., $u(x,y)=v(x,y)=0$ for $x=(0,1)$ at $y=0$ and 1, and $y=(0,1)$ at $x=0$ and 1. In the *w-u-v* formulation, we introduce the following in-plane displacement expansions:

$$u=\sum_{i=2,4,6}\sum_{j=1,3,\ldots,N}b_{i,j}\phi_i(x)\phi_j(y),$$

$$v=\sum_{i=1,3,\ldots,N}\sum_{j=2,4,6}c_{i,j}\phi_i(x)\phi_j(y). \tag{21}$$

With $\phi_i(\xi)=\sqrt{2}\sin(i\pi\xi)$ we see that the immovable plate edge conditions are fully guaranteed. Also, one summation of (21) is limited to the indices (2,4,6) consistent with (20), whereas the other summation is truncated at the upper limit N. On the other hand, for the *w-F* formulation it is customary to express the Airy function by

$$F=-\frac{C_y x^2}{2}-\frac{C_x y^2}{2r^2}+(1-\nu^2) \times\sum_{\substack{p=0,2,4,\ldots\\(p=q\neq 0)}}\sum_{q=0,2,4,\ldots}f_{p,q}\cos(p\pi x)\cos(q\pi y). \tag{22}$$

Under expansion (20), it turns out $f_{p,q}=0$ for $p,q>6$ so that (22) is a truncated double series. It is fair to say that the controversy of *w-F* formulation arises from imposing the boundary conditions for F in an average (integral) sense ([6,7]). First, the cosine expansion is justified in that $\int_0^1(\partial^2F/\partial x\partial y)_{x=0,1}dy=\int_0^1(\partial^2F/\partial x\partial y)_{y=0,1}dx=0$ imply zero shear stresses around the plate edges in an average sense. Second, zero in-plane displacements around the plate edge are expressed by the following integral constraints:

$$\int_0^1\int_0^1\left\{\frac{1}{(1-\nu^2)}\left(r^2\frac{\partial^2 F}{\partial y^2}-\nu\frac{\partial^2 F}{\partial x^2}\right)-\frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2\right\}dxdy=0,$$
$$\int_0^1\int_0^1\left\{\frac{1}{(1-\nu^2)}\left(\frac{\partial^2 F}{\partial x^2}-\nu r^2\frac{\partial^2 F}{\partial y^2}\right)-\frac{r^2}{2}\left(\frac{\partial w}{\partial y}\right)^2\right\}dxdy=0, \tag{23}$$

from which C_x and C_y are evaluated. Much has been argued to justify the integral boundary conditions imposed on F (see Refs. [6–9]), and hence we have nothing new to add here. We shall, however, adopt here a utilitarian view to judge the efficacy of (22) *a posteriori* by the consistency of modal equations that are derived from it.

5 Two Plate Equations for *w* and *F*

For the *w-F* formulation, we first evaluate from (23)

$$C_x=-\frac{\pi^2}{2}((1+\nu r^2)a_{1,1}^2+(1+9\nu r^2)a_{1,3}^2+(9+\nu r^2)a_{3,1}^2 +9(1+\nu r^2)a_{3,3}^2),$$

$$C_y=-\frac{\pi^2}{2}((r^2+\nu)a_{1,1}^2+(9r^2+\nu)a_{1,3}^2+(r^2+9\nu)a_{3,1}^2 +9(r^2+\nu)a_{3,3}^2).$$

Next, by inserting (20) and (22) into (18) we express $f_{p,q}$ in the quadratic $a_{n,m}$

$$f_{0,2}=(a_{1,1}^2-2a_{1,1}a_{1,3}+9a_{3,1}(a_{3,1}-2a_{3,3}))/8r^2,$$

$$f_{0,4}=(a_{1,1}a_{1,3}+9a_{3,1}a_{3,3})/16r^2,$$

$$f_{0,6}=(a_{1,3}^2+9a_{3,3}^2)/72r^2,$$

$$f_{2,0}=r^2(a_{1,1}^2-2a_{1,1}a_{3,1}+9a_{1,3}(a_{1,3}-2a_{3,3}))/8,\ldots,$$

$$f_{6,4}=-36r^2a_{3,1}a_{3,3}/(36+16r^2)^2,\quad f_{6,6}=0,\quad \text{and}$$

$$f_{p,q}=0\quad \text{for}\quad p,q\geqslant 8.$$

They may be compared with $b_{p,q}$ in Table 1 of Levy [8] by letting $a_{i,j}\rightarrow(w_{i,j}/h)$, $f_{p,q}\rightarrow 4b_{p,q}/Eh^2$ and $r=1$.

With C_x, C_y, and $f_{p,q}$, we can sort out from (17) modal equations for $a_{i,j}$ in an oscillator form by defining column vector $\mathbf{q}=(q_1,q_2,q_3,q_4)\triangleq(a_{1,1},a_{1,3},a_{3,1},a_{3,3})$

$$\ddot{\mathbf{q}}+\mathbf{C}\mathbf{q}+\mathbf{K}=0, \tag{24}$$

where

$$\mathbf{C}=\begin{pmatrix}(1+r^2)^2 & 0 & 0 & 0\\ 0 & (1+9r^2)^2 & 0 & 0\\ 0 & 0 & (9+r^2)^2 & 0\\ 0 & 0 & 0 & 81(1+r^2)^2\end{pmatrix}$$

is the stiffness matrix. The diagonal $\mathbf{C}$ testifies to $\varphi_n(x)\varphi_m(y)$ being the plate modes. And, cubic amplitude (column) vector $\mathbf{K}=(k_1,k_2,k_3,k_4)$ has the components

$$\begin{aligned}k_1=&4\kappa_1q_1^3+3\kappa_5q_1^2q_2+2\kappa_{14}q_1q_2^2+3\kappa_6q_1^2q_3+2\kappa_{11}q_1q_2q_3\\&+\kappa_{12}q_2^2q_3+2\kappa_{15}q_1q_3^2+\kappa_{13}q_2q_3^2+2\kappa_9q_1q_2q_4+2\kappa_{10}q_1q_3q_4\\&+\kappa_{18}q_2q_3q_4+2\kappa_{19}q_1q_4^2,\\k_2=&\kappa_5q_1^3+2\kappa_{14}q_1^2q_2+\kappa_{11}q_1^2q_3+2\kappa_{12}q_1q_2q_3+\kappa_{13}q_1q_3^2\\&+\kappa_9q_1^2q_4+\kappa_{18}q_1q_3q_4+4\kappa_2q_2^3+2\kappa_{20}q_2q_3^2+3\kappa_7q_2^2q_4\\&+2\kappa_{16}q_2q_4^2,\\k_3=&\kappa_6q_1^3+\kappa_{11}q_1^2q_2+\kappa_{12}q_1q_2^2+2\kappa_{15}q_1^2q_3+2\kappa_{13}q_1q_2q_3\\&+\kappa_{10}q_1^2q_4+\kappa_{18}q_1q_2q_4+2\kappa_{20}q_2^2q_3+4\kappa_3q_3^3+3\kappa_8q_3^2q_4\\&+2\kappa_{17}q_3q_4^2,\\k_4=&\kappa_9q_1^2q_2+\kappa_{10}q_1^2q_3+\kappa_{18}q_1q_2q_3+2\kappa_{19}q_1^2q_4+\kappa_7q_2^3\\&+2\kappa_{16}q_2^2q_4+\kappa_8q_3^3+2\kappa_{17}q_3^2q_4+4\kappa_4q_4^3,\end{aligned}$$

where κ_n $(n=1-20)$ are relegated to the Appendix.

Let us now identify the Hamiltonian for (24). With the conjugate $\mathbf{p}=\dot{\mathbf{q}}$ the inertial term gives rise to kinetic energy $H_k=\frac{1}{2}\mathbf{p}^T\mathbf{p}$, and the stiffness matrix generates bending strain energy $H_b=\frac{1}{2}\mathbf{q}^T\mathbf{C}\mathbf{q}$, where the superscript T denotes the transposed. We then deduce the following membrane strain energy:

$$\begin{aligned}H_m=&\kappa_1q_1^4+\kappa_2q_2^4+\kappa_3q_3^4+\kappa_4q_4^4+\kappa_5q_1^3q_2+\kappa_6q_1^3q_3+\kappa_7q_2^3q_4\\&+\kappa_8q_3^3q_4+\kappa_9q_1^2q_2q_4+\kappa_{10}q_1^2q_3q_4+\kappa_{11}q_1^2q_2q_3+\kappa_{12}q_1q_2^2q_3\\&+\kappa_{13}q_1q_2q_3^2+\kappa_{14}q_1^2q_2^2+\kappa_{15}q_1^2q_3^2+\kappa_{16}q_2^2q_4^2+\kappa_{17}q_3^2q_4^2\\&+\kappa_{18}q_1q_2q_3q_4+\kappa_{19}q_1^2q_4^2+\kappa_{20}q_2^2q_3^2,\end{aligned}$$

by integrating $\Sigma_{i=1}^4\int k_i dq_i$ and eliminating all redundant quartic terms ([10]). From total Hamiltonian $H(\mathbf{p},\mathbf{q})=H_k+H_b+H_m$, we can rederive (24) directly from the Hamilton equations of motion ([11])

$$\dot{q}_i=\frac{\partial H}{\partial p_i},\quad \dot{p}_i=-\frac{\partial H}{\partial q_i}, \tag{25}$$

hence (24) is Hamiltonian. Note that being Hamiltonian is a stronger dynamical property than energy-conserving. In fact, the existence of Hamiltonian is anticipated from the energy discussion in the beginning of this note. Though tedious, we can show $H(\mathbf{p},\mathbf{q})$ is really the total energy $U_k+U_b+U_m$. That is, (13) reduces to $U_k=\frac{1}{2}(\dot{a}_{1,1}^2+\dot{a}_{1,3}^2+\dot{a}_{3,1}^2+\dot{a}_{3,3}^2)$ which is H_k, (14) becomes $U_b=\frac{1}{2}((1+r^2)^2a_{1,1}^2+(1+9r^2)^2a_{1,3}^2+(9+r^2)^2a_{3,1}^2+81(1+r^2)^2a_{3,3}^2)$ which is nothing but H_b, and finally $U_m=H_m$ by (19).

6 Three Plate Equations for *w*, *u*, and *v*

Instead of $f_{p,q}$, we evaluate $b_{i,j}$ and $c_{i,j}$ by inserting (21) into (12). Afterward, the modal equations are derived from (10) just as we have done so from (17) in the w-F formulation. They are again put in oscillator form (24) with the same $\mathbf{C}$ as in Section 5, but now $\mathbf{K}=(\Re_1,\Re_2,\Re_3,\Re_4)$ has the following components:

$$\begin{aligned}\Re_1=&\{\alpha_1^1q_1^3+\alpha_2^1q_1^2q_2+\alpha_3^1q_1q_2^2+\alpha_4^1q_1^2q_3+\alpha_5^1q_1q_2q_3+\alpha_6^1q_2^2q_3\\&+\alpha_7^1q_1q_3^2+\alpha_8^1q_2q_3^2+\alpha_9^1q_1q_2q_4+\alpha_{10}^1q_1q_3q_4+\alpha_{11}^1q_2q_3q_4\\&+\alpha_{12}^1q_1q_4^2\}+\{\beta_1^1q_2^3+\beta_2^1q_3^3+\beta_3^1q_1^2q_4+\beta_4^1q_2^2q_4+\beta_5^1q_3^2q_4\\&+\beta_6^1q_2q_4^2+\beta_7^1q_3q_4^2+\beta_8^1q_4^3\},\\\Re_2=&\{\alpha_1^2q_1^3+\alpha_2^2q_1^2q_2+\alpha_3^2q_1^2q_3+\alpha_4^2q_1q_2q_3+\alpha_5^2q_1q_3^2+\alpha_6^2q_1^2q_4\\&+\alpha_7^2q_1q_3q_4+\alpha_8^2q_2^3+\alpha_9^2q_2q_3^2+\alpha_{10}^2q_2^2q_4+\alpha_{11}^2q_2q_4^2\}\\&+\{\beta_1^2q_1q_2^2+\beta_2^2q_2^2q_3+\beta_3^2q_3^3+\beta_4^2q_1q_2q_4+\beta_5^2q_2q_3q_4\\&+\beta_6^2q_3^2q_4+\beta_7^2q_1q_4^2+\beta_8^2q_3q_4^2+\beta_9^2q_4^3\},\\\Re_3=&\{\alpha_1^3q_1^3+\alpha_2^3q_1^2q_2+\alpha_3^3q_1q_2^2+\alpha_4^3q_1^2q_3+\alpha_5^3q_1q_2q_3+\alpha_6^3q_1^2q_4\\&+\alpha_7^3q_1q_2q_4+\alpha_8^3q_2^2q_3+\alpha_9^3q_3^3+\alpha_{10}^3q_3^2q_4+\alpha_{11}^3q_3q_4^2\}\\&+\{\beta_1^3q_2^3+\beta_2^3q_1q_3^2+\beta_3^3q_2q_3^2+\beta_4^3q_2^2q_4+\beta_5^3q_1q_3q_4\\&+\beta_6^3q_2q_3q_4+\beta_7^3q_1q_4^2+\beta_8^3q_2q_4^2+\beta_9^3q_4^3\},\\\Re_4=&\{\alpha_1^4q_1^2q_2+\alpha_2^4q_1^2q_3+\alpha_3^4q_1q_2q_3+\alpha_4^4q_1^2q_4+\alpha_5^4q_2^3+\alpha_6^4q_2^2q_4\\&+\alpha_7^4q_3^3+\alpha_8^4q_3^2q_4+\alpha_9^4q_4^3\}+\{\beta_1^4q_1^3+\beta_2^4q_1q_2^2+\beta_3^4q_2^2q_3\\&+\beta_4^4q_1q_3^2+\beta_5^4q_2q_3^2+\beta_6^4q_1q_2q_4+\beta_7^4q_1q_3q_4+\beta_8^4q_2q_3q_4\\&+\beta_9^4q_1q_4^2+\beta_{10}^4q_2q_4^2+\beta_{11}^4q_3q_4^2\}.\end{aligned}$$

Table 1 Numerical coefficients for $\mathbf{K}=\Re$ under $N=7$, $r=1$, and $\nu=\sqrt{0.1}$

$\alpha_n^1\,(n=1-12)$	21.31/−8.95/127/−8.95/23.56/−70/127/−70/−47.93/−47.93/257/163.47
$\alpha_n^2\,(n=1-11)$	−3/127/11.73/−139.8/−70.21/−23.86/259.45/751.53/585.46/−680.04/1679.93
$\alpha_n^3\,(n=1-11)$	−3/11.73/−70.21/127/−139.8/−23.86/259.45/585.46/751.53/−680.04/1679.93
$\alpha_n^4\,(n=1-9)$	−24.1/−24.1/259.89/163.7/−228.32/1689/−228.32/1689/1712.69
$\beta_n^1\,(n=1-8)$	−1.11/−1.11/−1.39/2.36/2.36/6.49/6.49/−0.28
$\beta_n^2\,(n=1-9)$	−3.13/−4.03/−1.8/4.01/9.41/7.34/8.52/3.86/−0.8
$\beta_n^3\,(n=1-9)$	−1.8/−3.13/−4.03/7.34/4.01/9.41/8.52/3.86/−0.8
$\beta_n^4\,(n=1-11)$	−0.52/1.37/7.29/1.37/7.29/18.26/18.26/8.42/−0.34/−1.94/−1.94

Table 2 Numerical coefficients for $\partial U_m/\partial\mathbf{q}=\Re$ under $N=7$, $r=1$, and $\nu=\sqrt{0.1}$

$\alpha_n^1\,(n=1-12)$	21.32/−8.94/127.1/−8.94/23.43/−69.97/127.1/−69.97/−47.89/−47.89/260.5/164.03
$\alpha_n^2\,(n=1-11)$	−2.98/127.1/11.72/−139.95/−69.97/−23.94/260.5/752.26/586.62/−682.17/1693.46
$\alpha_n^3\,(n=1-11)$	−2.98/11.72/−69.97/127.1/−139.95/−23.94/260.5/586.62/752.26/−682.17/1693.46
$\alpha_n^4\,(n=1-9)$	−23.94/−23.94/260.5/164.03/−227.39/1693.46/−227.39/1693.46/1732.87
$\beta_n^1\,(n=1-8)$	−0.94/−0.94/−1.46/1.67/1.67/9.57/9.57/−0.14
$\beta_n^2\,(n=1-9)$	−2.82/−4.83/−1.61/3.34/15.99/8/9.57/4.26/2.06
$\beta_n^3\,(n=1-9)$	−1.61/−2.82/−4.83/8/3.34/15.99/9.57/4.26/2.06
$\beta_n^4\,(n=1-11)$	−0.49/1.67/8/1.67/8/19.13/19.13/8.52/−0.43/6.18/6.18

We observe the following. First, each $\mathfrak{R}_i$ has 20 (=6!/3!3!) cubic terms, corresponding to the combinations with repetitions of forming $q_l q_m q_n$ out of $\mathbf{q}$. Second, the cubic terms in each $\mathfrak{R}_i$ are split into two curly brackets. In the first curly brackets, the terms with coefficient α_n^i have the same cubic pattern as in k_i, and we call them Type I cubic terms. The remaining terms in the second curly brackets with coefficient β_n^i are called Type II cubic terms.

For a given N, there are $\frac{3}{2}(N+1)$ components each for $b_{i,j}$ and $c_{i,j}$ so that solution of (12) requires inverting a matrix of $3(N+1)\times 3(N+1)$, which for instance is 6×6 for $N=1$. It is therefore not feasible to present the analytical expressions for α_n^i and β_n^i, hence we shall present only their numerical values evaluated for $r=1$ and $\nu=\sqrt{0.1}$. For $N=7$ the values of α_n^i and β_n^i are summarized in Table 1. Beyond the huge array of numerical data presented in the table, the point of contention is that one cannot deduce H_m directly from $\mathbf{K}=\mathfrak{R}$. To see this, for instance, $\beta_1^1 \neq \frac{1}{3}\beta_1^2$ in Table 1, and β_1^1 and $\frac{1}{3}\beta_1^2$ are the coefficients of $q_2^3 q_1$. Hence, the modal equations for $\mathbf{q}$ are not Hamiltonian in the w-u-v formulation or, put it more informally, cubic vector $\mathbf{K}=\mathfrak{R}$ is not energy-conserving when we truncate (21) at $N=7$. Now, to infer energy-conserving cubic vector we first evaluate U_m from (15),

$$
\begin{aligned}
U_m = {} & 5.33q_1^4 + 188.01q_2^4 + 188.01q_3^4 + 433.22q_4^4 - 2.98q_1^3q_2 - 2.98q_1^3q_3 - 227.39q_2^3q_4 - 227.39q_3^3q_4 - 23.94q_1^2q_2q_4 - 23.94\kappa_{10}q_1^2q_3q_4 \\
& + 11.72q_1^2q_2q_3 - 69.97q_1q_2^2q_3 - 69.97q_1q_2q_3^2 + 63.55q_1^2q_2^2 + 63.55q_1^2q_3^2 + 846.73q_2^2q_4^2 + 846.73q_3^2q_4^2 + 260.5q_1q_2q_3q_4 \\
& + 82.01q_1^2q_4^2 + 293.31q_2^2q_3^2 - 0.94q_1q_2^3 - 0.94q_1q_3^3 - 1.61q_2^3q_3 - 1.61q_2q_3^3 - 0.49q_1^3q_4 + 1.67q_1q_2^2q_4 + 1.67q_1q_3^2q_4 + 7.99q_2^2q_3q_4 \\
& + 7.99q_2q_3^2q_4 + 9.57q_1q_2q_4^2 + 9.57q_1q_3q_4^2 + 4.26q_2q_3q_4^2 - 0.14q_1q_4^3 + 2.06q_2q_4^3 + 2.06q_3q_4^3,
\end{aligned}
$$

Table 3 Percent deviation ranges in the numerical coefficients of cubic terms

N	Type I cubic terms	Type II cubic terms
7	(−1.37,0.7%)	(−140%,94%)
11	(−0.32%,0.19%)	(−78%,12%)
15	(−0.13%,0.077%)	(−42%,4.3%)
19	(−0.063%,0.039%)	(−23%,2%)
23	(−0.036%,0.022%)	(−14%,1.1%)
27	(−0.022%,0.014%)	(−9%,0.7%)
31	(−0.015%,0.009%)	(−6%,0.46%)
35	(−0.01%,0.007%)	(−4.3%,0.32%)

with (21) truncated also at $N=7$. Then, by (25) $\partial U_m/\partial q_i = \mathfrak{R}_i$ are components of the energy-conserving cubic vector and their values of α_n^i and β_n^i are summarized in Table 2. We can therefore quantify departure of $\mathbf{K}=\mathfrak{R}$ from the energy-conserving $\partial U_m/\partial U_m/\partial \mathbf{q}=\mathfrak{R}$ by comparing α_n^i and β_n^i in Tables 1 and 2. It is found that α_n^i differ only slightly in small percent range (−1.37 percent, 0.7 percent), whereas the values of β_n^i deviate greatly within a large percent range of (−140 percent, 94 percent). These deviation deviation ranges are entered into the row entries of Table 3 for $N=7$ under Type I and II cubic terms, respectively. As summarized in Table 3, $\mathbf{K}$ approaches the energy-conserving $\partial U_m/\partial \mathbf{q}$ as N is incremented by 4. Finally, at $N=35$ the α_n^i have almost attained the energy-conserving values, yet the percent deviation range of β_n^i is still more than ±1 percent.

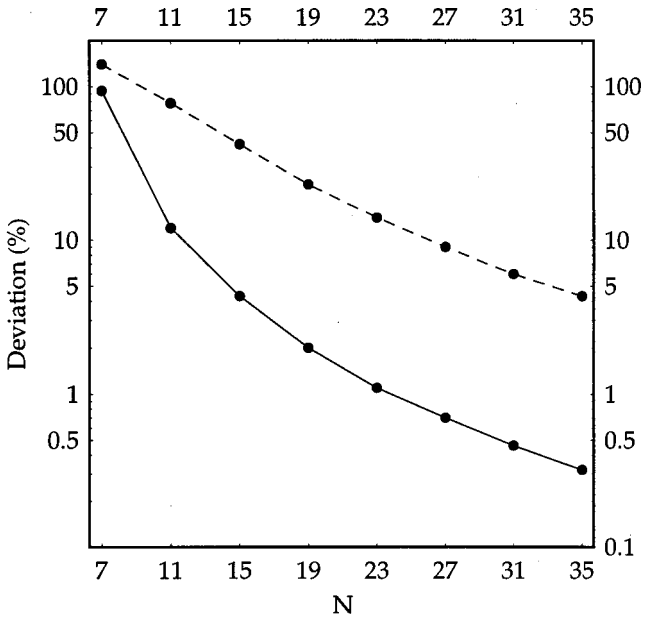


Fig. 1 Approach to the energy-conserving cubic vector. -●- maximum positive deviation; --●-- maximum absolute negative deviation.

7 Concluding Remarks

We began by presenting thin-plate equations of a moderately large deflection theory under the w-u-v and w-F formulations. The main issue here is to ask if Hamiltonian property survives through the Galerkin procedure by which modal equations are derived for the first four symmetric modes of a simply-supported plate. The modal equations are indeed Hamiltonian in the w-F formulation, whereas the corresponding modal equations of w-u-v formation are not and hence do not conserve the plate energy. In Table 3 the departure from energy conservation has been quantified by the percent deviations in Type I and II cubic amplitude terms. As the worst-case scenario, we plot in Fig. 1 the percent deviations of Type II cubic terms over $N=7-35$. By extrapolating the maximum absolute negative deviation out to 1 percent, we find $N\sim 55$ as a conservative estimate. In other words, the expansions for u and v must include at least 84 sine terms each to assure less than ±1 percent deviation in the coefficient β_n^i from the energy-conserving cubic vector.

Acknowledgments

I wish to thank Prof. David Walker for providing me with an earliest available copy of Geveci's thesis cited here in Ref. [5]. Also, correspondence with him and Dr. Geveci are sincerely appreciated.

Appendix

The Constants κ_n $(n=1-20)$.

$$\kappa_1=\frac{3}{4}(4r^2\nu+(r^4+1)(3-\nu^2)),$$

$$\kappa_2=\frac{3}{4}(36r^2\nu+(81r^4+1)(3-\nu^2)),$$

$$\kappa_3=\frac{3}{4}(36r^2\nu+(81+r^4)(3-\nu^2)),$$

$$\kappa_4=\frac{243}{4}(4r^2\nu+(r^4+1)(3-\nu^2)),$$

$$\kappa_5=-3(1-\nu^2),\quad \kappa_6=-3r^4(1-\nu^2),$$

$$\kappa_7=-243r^4(1-\nu^2),\quad \kappa_8=-243(1-\nu^2),$$

$$\kappa_9=-27r^4\left(1+\frac{1}{(1+4r^2)^2}\right)(1-\nu^2),$$

$$\kappa_{10}=-27\left(1+\frac{r^4}{(4+r^2)^2}\right)(1-\nu^2),\quad \kappa_{11}=\frac{48r^4(1-\nu^2)}{(1+r^2)^2},$$

$$\kappa_{12}=-24r^4\left(\frac{9}{8}+\frac{8}{(1+r^2)^2}+\frac{25}{8(1+4r^2)^2}\right)(1-\nu^2),$$

$$\kappa_{13}=-24r^4\left(\frac{9}{8r^4}+\frac{8}{(1+r^2)^2}+\frac{25}{8(4+r^2)^2}\right)(1-\nu^2),$$

$$\kappa_{14}=12r^4\left(\frac{9}{8}+\frac{1}{2r^4}+\frac{2}{(1+r^2)^2}+\frac{1}{8(1+4r^2)^2}\right)(1-\nu^2)+3((1+r^2\nu)+9r^2(r^2+\nu)),$$

$$\kappa_{15}=12r^4\left(\frac{1}{2}+\frac{9}{8r^4}+\frac{2}{(1+r^2)^2}+\frac{1}{8(4+r^2)^2}\right)(1-\nu^2)+3(9(1+r^2\nu)+r^2(r^2+\nu)),$$

$$\kappa_{16}=27\left(\frac{1}{2}+72r^4\left(\frac{1}{4}+\frac{1}{(1+9r^2)^2}+\frac{1}{16(4+9r^2)^2}\right)\right)(1-\nu^2)+27((1+r^2\nu)+9r^2(r^2+\nu)),$$

$$\kappa_{17}=27r^4\left(\frac{1}{2}+72\left(\frac{1}{4r^4}+\frac{1}{(9+r^2)^2}+\frac{1}{16(9+4r^2)^2}\right)\right)(1-\nu^2)+27(9(1+r^2\nu)+r^2(r^2+\nu)),$$

$$\kappa_{18}=12r^4\left(9\left(1+\frac{1}{r^4}\right)+\frac{225}{4}\left(\frac{1}{(4+r^2)^2}+\frac{1}{(1+4r^2)^2}\right)\right)(1-\nu^2),$$

$$\kappa_{19}=\frac{243}{2}r^4\left(\frac{1}{(4+r^2)^2}+\frac{1}{(1+4r^2)^2}\right)(1-\nu^2)+27((1+r^2\nu)+r^2(r^2+\nu)),$$

$$\kappa_{20}=24r^4\left(\frac{17}{(1+r^2)^2}+\frac{625}{16}\left(\frac{1}{(4+r^2)^2}+\frac{1}{(1+4r^2)^2}\right)\right)(1-\nu^2)+3(9+r^2\nu+9r^2(r^2+9\nu)).$$

References

[1] Chu, H.-N., and Herrmann, G., 1956, "Influence of Large Amplitudes on Free Flexural Vibrations of Rectangular Elastic Plates," ASME J. Appl. Mech., **23**, pp. 532–540.

[2] Chia, C.-Y., 1980, *Nonlinear Analysis of Plates*, McGraw-Hill, New York, Chapter 1.

[3] Wolfram, S., 1999, *The MATHEMATICA Book*, 4th Ed., Wolfram Media, Wolfram Research, IL.

[4] Lee, J., 1999, "Using Mathematica for the Galerkin Representation of Nonlinear Plate Equations," *Proc. 6th Pan Am. Cong. of Appl. Mech.*, Vol. 7, eds. P. B. Concalves, I. Jasiuk, D. Pamploma, C. Steele, H. I. Weber, L. Bevilacqua, Rio de Janeiro, pp. 821–824.

[5] Geveci, B., 1999, "Flow Induced Nonlinear Vibrations of Rectangular Plates," Ph.D. dissertation in Mechanical Engineering and Mechanics, Lehigh University, Dec 13. Lehigh, PA.

[6] Bolotin, V. V., 1963, *Nonconservative Problems of the Theory of Elastic Stability*, MacMillan, New York.

[7] Librescu, L., 1975, *The Elastostatic and Kinetics of Anisotropic and Heterogeneous Shell-Type Structures*, Noorhoff, Lyden, Chapters 1–7.

[8] Levy, S., 1942, "Bending of Rectangular Plates With Large Deflections," NACA Report No. 737.

[9] Dowell, E. H., 1966, "Nonlinear Oscillations of a Fluttering Plate," AIAA J., **4**, pp. 1267–1275.

[10] Lee, J., 1986, "Free Vibration of a Large-Amplitude Deflected Plate—Reexamination by the Dynamical Systems Theory," ASME J. Appl. Mech., **53**, pp. 633–640.

[11] Goldstein, H., 1980, *Classical Mechanics*, 2nd Ed., Addison-Wesley, Reading, MA, Chapter 8.

An Infinite Plate Weakened by Periodic Cracks

Y. Z. Chen
Division of Engineering Mechanics, Jiangsu University, Zhenjiang, Jiangsu 212013, P. R. China

K. Y. Lee
Department of Mechanical Engineering, Yonsei University, Seoul 120-749, South Korea

An infinite plate weakened by doubly distributing cracks is studied in this paper. Two loading cases, the remote tension and the remote shear stresses, are assumed. Analysis is performed for a cracked cell cut from the infinite plate. It is found that the eigenfunction expansion variational method is efficient to solve the problem. The stress intensity factor, the T-stress, and the elastic response are evaluated. The cracked plate can be equivalent to an orthotropic medium without cracks. The equivalent elastic constants are presented. [DOI: 10.1115/1.1458558]

1 Introduction

Multiple crack problems in plane elasticity were investigated by many investigators ([1–3]). A doubly periodic crack problem is a particular one in this field, as was studied in ([4,5]). In the foregoing studies, the obtained results are limited to evaluate the stress intensity factor and some discrepancy has been found between the sources. Also, the effective elastic properties of solids with many cracks were investigated ([6]). Recently, the doubly periodic cracks were modeled by a superposition of many rows of collinear cracks. The derivation was rather complicated and only the normal loading case was considered ([7,8]).

In this paper, an infinite plate weakened by doubly distributed cracks is studied. The plate is subjected to the remote tension or the remote shear stress. In both cases, the boundary value problems can be reduced into a complex mixed one for a cracked rectangular cell. It is found that the EEVM (eigenfunction expansion variational method) is efficient to solve the problem ([9]). The particular advantage of EEVM is that the relevant computa-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 12, 2001; final revision, Nov. 22, 2001. Associate Editor: J. R. Barber.

tion does not depend on any boundary collocation scheme which is necessary in the boundary collection method. Generally, if the boundary collocation method is used, the boundary collocation scheme will influence the final results. For both normal and shear loading cases, the stress intensity factors and the deformation responses are evaluated. From the obtained results it is shown that the cracked plate can be equivalent to an orthotropic medium without cracks.

2 Analysis

Based on the complex variable function method in plane elasticity ([10]), the stresses $(\sigma_x, \sigma_y, \sigma_y)$ and the displacements (u, v) are expressed in terms of two complex potentials $\phi(z)$ and $\omega(z)$ such that

$$\begin{aligned} \sigma_x + \sigma_y &= 4 \,\mathrm{Re}\, \phi'(z) \\ \sigma_y - i\sigma_{xy} &= \phi'(z) + (z-\bar{z})\overline{\phi''(z)} + \omega'(\bar{z}) \end{aligned} \tag{1}$$

$$2G(u+iv) = \kappa\phi(z) - (z-\bar{z})\overline{\phi'(z)} - \omega(\bar{z}) \tag{2}$$

where G is the shear modulus of elasticity, and $\kappa = (3-\nu)/(1+\nu)$ for the plane stress problem with ν being the Poisson's ratio.

In the following analysis, an infinite plate with doubly periodic cracks as shown in Fig. 1 is first considered under the remote tensions $\sigma_x = 0$, $\sigma_y = p$. The relevant elastic constants are denoted by ν_0, G_0, E_0 $(E_0 = 2G_0(1+\nu_0))$, and $\nu_0 = 0.3$ is used in this paper. In the analysis, it is convenient to cut a rectangular cracked cell from the infinite plate (Fig. 1(*a*)). Clearly, the boundary condition for the cracked rectangle will be

$$v = \bar{v} = \pm v_b, \quad \sigma_{xy} = 0 \quad (-b \leqslant x \leqslant b, y = \pm h) \tag{3a}$$

$$u = \bar{u} = \pm u_b, \quad \sigma_{xy} = 0 \quad (x = \pm b, -h \leqslant y \leqslant h) \tag{3b}$$

where v_b and u_b are two unknown values to be determined later. Clearly, the conditions in Eq. (3) are a complex mixed one.

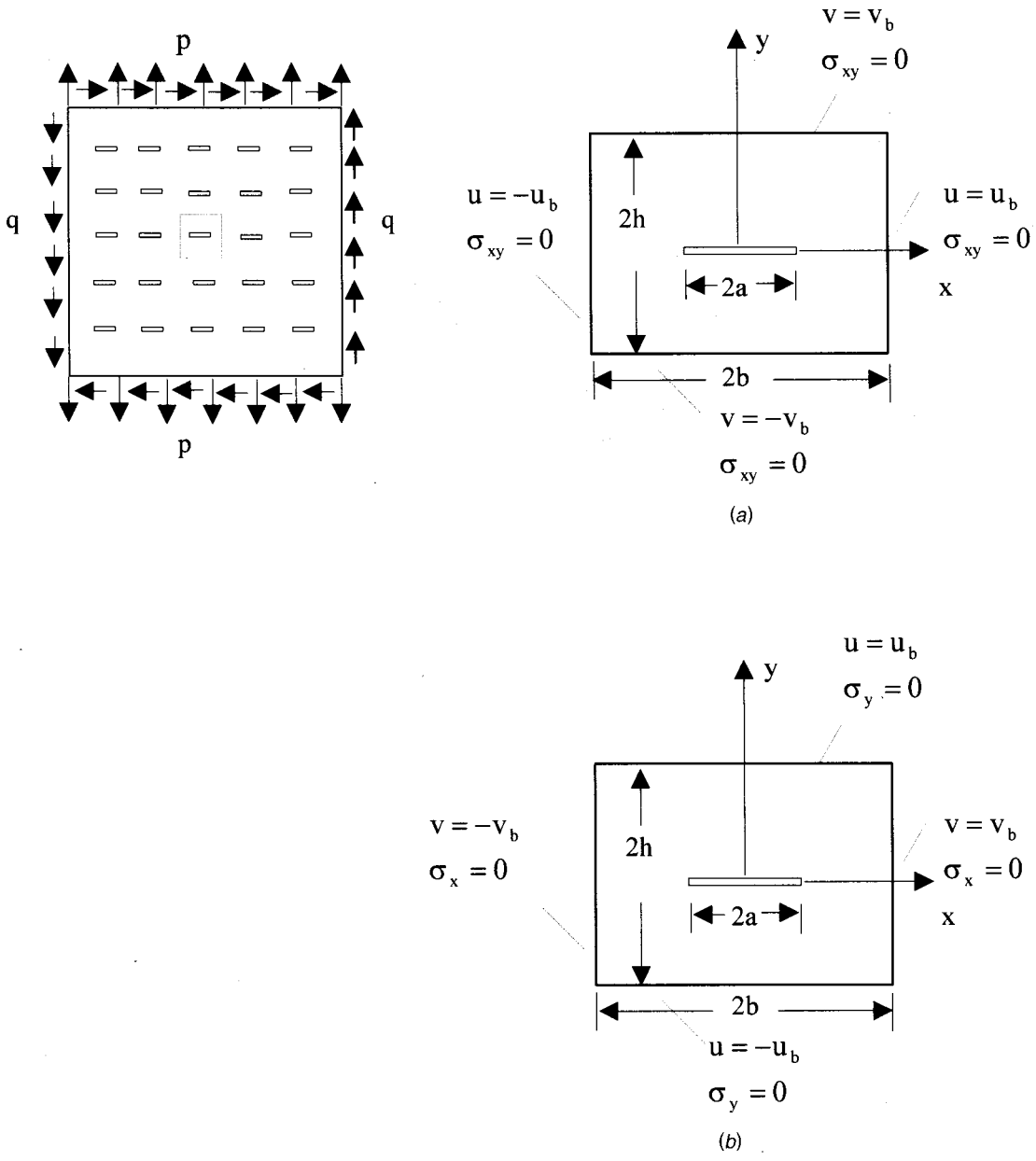


Fig. 1 An infinite plate with the doubly periodic cracks, (*a*) the cracked rectangle in tension loading, (*b*) the cracked rectangle in shear loading

To solve the problem, we introduce the following complex potentials ([9]):

$$\phi(z)=\sum_{k=1}^{2M} X_k\phi^{(k)}(z), \quad \omega(z)=\sum_{k=1}^{2M} X_k\omega^{(k)}(z) \tag{4}$$

where

$$\begin{aligned}\phi^{(k)}(z)&=\omega^{(k)}(z)=\sqrt{z^2-a^2}z^{2k-2}, \quad (1\leqslant k\leqslant M)\\ \phi^{(k)}(z)&=-\omega^{(k)}(z)=z^{2(M-k)-1}, \quad (M+1\leqslant k\leqslant 2M)\end{aligned} \tag{5}$$

and X_k are the unknown coefficients. As shown previously ([9]), the displacements, strains, and stresses derived from the complex potentials $\phi(z)$ and $\omega(z)$ in Eq. (4) satisfy all governing equations of plane elasticity and the traction-free condition on the crack faces.

It is proved that, for the following mixed boundary conditions for a cracked region ([9])

$$\sigma_{ij}n_j=\overline{p_i} \ \text{(on } C_p\text{)} \quad u_i=\overline{u_i} \ \text{(on } C_u\text{)}, \tag{6}$$

where C_p is the portion of boundary where the tractions are given, and C_u is the portion of boundary where the displacements are applied, and the actual solution can be obtained from the stationary condition of a functional Π defined as

$$\Pi=\int\int_{\Sigma} A(e_{ij})dF-\int_{C_p}\overline{p_i}u_i ds-\int_{C_u}\sigma_{ij}n_j(u_i-\overline{u_i})ds \tag{7}$$

where $A(e_{ij})$ is the strain energy and Σ is the region of integration. Simply using the following steps: (a) substituting all the quantities derived from the complex potentials in Eq. (4) and the boundary values in Eq. (6) into Eq. (7) and (b) letting the functional Π be a stationary value, we can construct a system of algebraic equations for the undetermined coefficients X_k ($k=1,2,\ldots 2M$) in Eq. (4). This method for evaluating the stress field for a cracked region is called EEVM [9].

The actual solution for the problem can be obtained in the following way. Under the conditions $v_b=1$ and $u_b=0$ in Eqs. (3*a*) and (3*b*), the obtained stresses are denoted by $\sigma_{x(1)}$, $\sigma_{y(1)}$, and $\sigma_{xy(1)}$. Similarly, under the conditions $v_b=0$ and $u_b=1$ in Eqs. (3*a*) and (3*b*), the obtained stresses are denoted by $\sigma_{x(2)}$, $\sigma_{y(2)}$, and $\sigma_{xy(2)}$. The loading conditions along the boundary in Fig. 1(*a*) lead to

$$\begin{aligned}v_b\int_0^b\sigma_{y(1)}(x,h)dx+u_b\int_0^b\sigma_{y(2)}(x,h)dx&=bp\\ v_b\int_0^h\sigma_{x(1)}(b,y)dy+u_b\int_0^h\sigma_{x(2)}(b,y)dy&=0.\end{aligned} \tag{8}$$

Thus, two unknown values v_b and u_b can be determined from the above equations, and the final solution is obtainable. In case of using $M=8$, the calculated results for the model I stress intensity factor and the T-stress can be, respectively, expressed as

$$K_1=A_1(h/b,a/b)p(\pi a)^{1/2}, \quad T=-B_1(h/b,a/b)\frac{1}{1-(a/b)}p. \tag{9}$$

The calculated $A_1(h/b,a/b)$ and $B_1(h/b,a/b)$ values are listed in Table 1. It is found that the obtained results coincide with the previous ones ([5,7]). For choosing the value of M, we performed computation for the cases of $M=3,4,\ldots$, and the value of M is determined when the stable numerical results are obtained.

Clearly, from the deformation response in the y-direction, the cracked rectangle can be modeled by an orthotropic medium without the cracks. It is known that the constitutive equation in the orthotropic medium takes the form ([11])

Table 1 The calculated nondimensional stress intensity factors and the *T*-stresses

$A_1(h/b,a/b)$ values (see Fig. 1 and Eq. (9))

a/b	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$h/b=0.4$	0.966	0.900	0.853	0.847	0.886	0.977	1.140	1.431
$h/b=0.6$	0.993	0.977	0.967	0.977	1.019	1.104	1.252	1.513
$h/b=0.8$	1.000	1.003	1.013	1.038	1.086	1.168	1.304	1.549
$h/b=1.0$	1.003	1.012	1.031	1.062	1.113	1.194	1.325	1.566
$h/b=1.5$	1.004	1.017	1.039	1.075	1.128	1.209	1.340	1.601
$h/b=2.0$	1.004	1.017	1.040	1.076	1.127	1.216	1.344	1.570

$B_1(h/b,a/b)$ values (see Fig. 1 and Eq. (9))

a/b	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$h/b=0.4$	0.846	0.636	0.436	0.278	0.167	0.090	0.045	0.019
$h/b=0.6$	0.883	0.743	0.597	0.461	0.342	0.242	0.161	0.097
$h/b=0.8$	0.894	0.779	0.662	0.546	0.437	0.335	0.241	0.156
$h/b=1.0$	0.898	0.793	0.687	0.581	0.477	0.376	0.278	0.184
$h/b=1.5$	0.900	0.800	0.699	0.599	0.499	0.399	0.299	0.200
$h/b=2.0$	0.900	0.800	0.700	0.600	0.500	0.401	0.301	0.200

$A_2(h/b,a/b)$ values (see Fig. 1 and Eq. (18))

a/b	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$h/b=0.4$	1.025	1.088	1.170	1.258	1.345	1.431	1.535	1.703
$h/b=0.6$	1.011	1.043	1.090	1.147	1.212	1.292	1.405	1.607
$h/b=0.8$	1.007	1.027	1.059	1.102	1.160	1.239	1.361	1.575
$h/b=1.0$	1.005	1.021	1.047	1.085	1.140	1.219	1.344	1.559
$h/b=1.5$	1.004	1.017	1.040	1.076	1.129	1.208	1.330	1.533
$h/b=2.0$	1.004	1.017	1.040	1.075	1.128	1.205	1.324	1.531

$$\varepsilon_x=\frac{1}{E_1}\sigma_x-\frac{\nu_{21}}{E_2}\sigma_y, \quad \varepsilon_y=-\frac{\nu_{12}}{E_1}\sigma_x+\frac{1}{E_2}\sigma_y, \quad \gamma_{xy}=\frac{1}{G_{12}}\sigma_{xy} \tag{10}$$

where there is a relation as follows:

$$(E_2\nu_{12})/(E_1\nu_{21})=1. \tag{11}$$

From the assumed loading condition and the numerical solution mentioned above, we have

$$\sigma_{x,av}=0, \quad \sigma_{y,av}=p, \quad \varepsilon_{x,av}=\frac{u_b}{b}, \quad \varepsilon_{y,av}=\frac{v_b}{h} \tag{12}$$

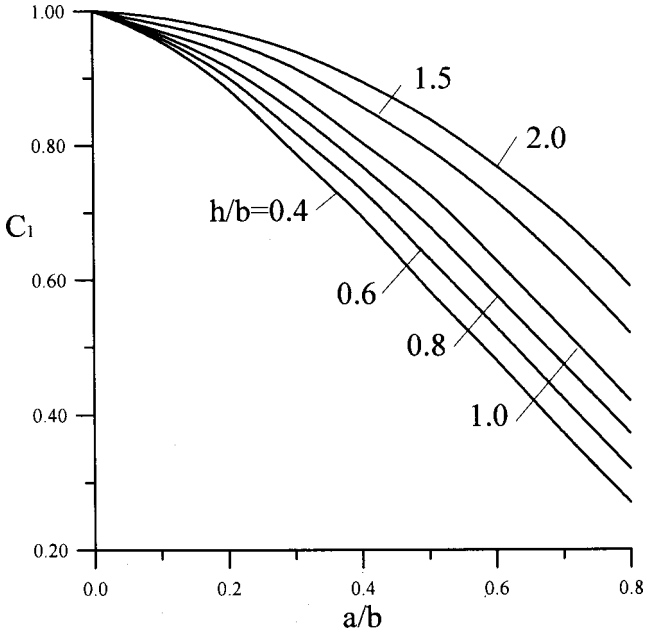


Fig. 2 Normalized elastic constant $C_1(h/b,a/b)(=E_2/E_0)$

where the subscript "av" means that the relevant quantity should be understood in the sense of average on some portion of the boundary.

Substituting Eq. (12) into Eq. (10) yields

$$E_2=\frac{hp}{v_b},\quad \nu_{21}=-\frac{hu_b}{bv_b}. \tag{13}$$

The calculated E_2 values are expressed by

$$E_2=C_1(h/b,a/b)E_0 \tag{14}$$

and the calculated C_1 values are plotted in Fig. 2.

Since the crack does not influence the relevant stress-strain relation in the x-direction, we have $E_1=E_0$ and $\nu_{12}=\nu_0$. Clearly, the results obtained for E_2, ν_{21} and $E_1(=E_0)$, $\nu_{12}(=\nu_0)$ may not satisfy the relation Eq. (11) exactly. In fact, in the range for h/b and a/b used in the numerical example, the ratios $(E_2\nu_{12})/(E_1\nu_{21})$ are varying within the range from 0.9990 to 1.0013. This is to say that the proposed assumption coincides with the physical situation very well.

Since the relation (11) is satisfied almost exactly, by using the relations $E_1=E_0$, $\nu_{12}=\nu_0$ and Eq. (14), it follows $\nu_{21}/\nu_0=C_1(h/b,a/b)$. Therefore, from the function $C_1(h/b,a/b)$ we can also obtain information for the reduction of Poisson's ratio.

The response of the periodic cracks to the shear loading can be investigated in a similar manner (Fig. 1(b)). In this case we assume that only the remote shear traction is nonzero such that $\sigma_{xy}=q$.

Similar to the tension case, the boundary conditions for the cracked cell are

$$u=\overline{u}=\pm u_b,\quad \sigma_y=0\quad (-b\leqslant x\leqslant b, y=\pm h) \tag{15a}$$

$$v=\overline{v}=\pm v_b,\quad \sigma_x=0\quad (x=\pm b, -h\leqslant y\leqslant h) \tag{15b}$$

where u_b and v_b are two unknown values to be determined by

$$\int_0^b \sigma_{xy}(x,h)dx=bq$$
$$\int_0^h \sigma_{xy}(b,y)dy=hq. \tag{16}$$

In this case, the eigenfunction expansion form Eq. (4) is still used with the following components:

$$\phi^{(k)}(z)=\omega^{(k)}(z)=i\sqrt{z^2-a^2}z^{2k-2},\quad (1\leqslant k\leqslant M)$$
$$\phi^{(k)}(z)=-\omega^{(k)}(z)=iz^{2(M-k)-1},\quad (M+1\leqslant k\leqslant 2M). \tag{17}$$

As before, letting the corresponding functional as in Eq. (7) be a stationary value yields the final solution. Finally, the calculated mode II stress intensity factor is expressed by

$$K_{\mathrm{II}}=A_2(h/b,a/b)q(\pi a)^{1/2}. \tag{18}$$

The calculated A_2 values are listed in Table 1.

From Table 1 we see that, in the case of $h/b=0.4$, $A_1(h/b,a/b)$ is generally lower than $A_2(h/b,a/b)$. However, in the case of $h/b\geqslant 1.0$, the two values $A_1(h/b,a/b)$ and $A_2(h/b,a/b)$ are approximately on the same level.

From the numerical solution we can obtain the average stress and strain for the cracked rectangle

$$\sigma_{xy,av}=q,\quad \gamma_{xy,av}=\frac{u_b}{h}+\frac{v_h}{b}. \tag{19}$$

As before, substituting Eq. (19) into Eq. (10) yields

$$G_{12}=q\left(\frac{u_b}{h}+\frac{v_b}{b}\right)^{-1}. \tag{20}$$

Similarly, the calculated G_{12} values can be expressed as

$$G_{12}=C_2(h/b,a/b)G_0. \tag{21}$$

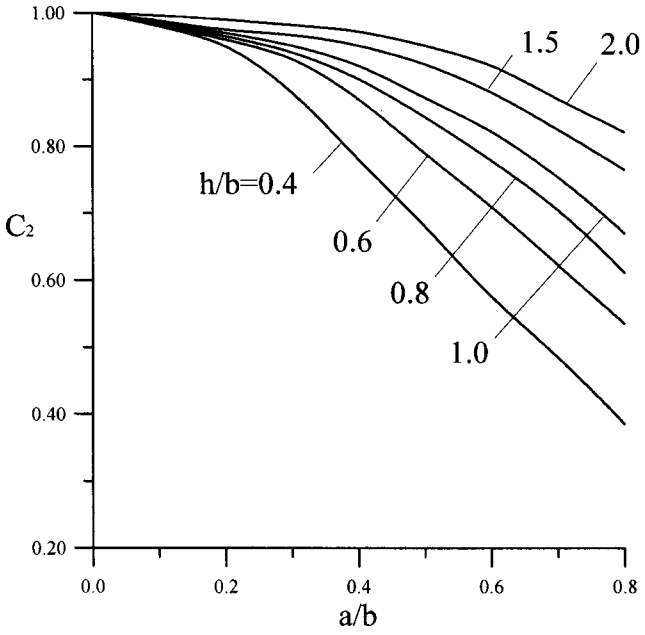


Fig. 3 Normalized elastic constant $C_2(h/b,a/b)(=G_{12}/G_0)$

The nondimensional values C_2 values are plotted in Fig. 3. From Figs. 2 and 3 we see that, for the same values of h/b and a/b, $C_1(h/b,a/b)$ is generally lower than $C_2(h/b,a/b)$. That is to say the reduction of rigidity in the normal loading case is higher than one in the shear loading case.

3 Conclusion

For an infinite plate with periodic cracks, the elastic response for any cracked cell to the external loading can be obtained. The eigenfunction expansion variational method is an effective way to solve the cracked cell problem with finite dimension. The suggested method can provide the accurate result for the stress intensity factors and the elastic response with relatively less effort.

References

[1] Savruk, M. P., 1981, *Two-dimensional Problems of Elasticity for Body with Cracks*, Naukova Dumka, Kiev (in Russian).

[2] Chen, Y. Z., 1995, "A Survey of New Integral Equations in Plane Elasticity Crack Problem," Eng. Fract. Mech., **51**, pp. 97–134.

[3] Kachanov, M., 1987, "Elastic Solids With Many Cracks; A Simple Method of Analysis," Int. J. Solids Struct., **23**, pp. 23–44.

[4] Delameter, W. R., Herrmann, G., and Barnett, D. M., 1975, "Weakening of Elastic Solid by a Rectangular Array of Cracks," ASME J. Appl. Mech., **42**, pp. 74–80.

[5] Isida, M., Usijima, N., and Kishine, N., 1981, "Rectangular Plate, Strips and Wide Plates Containing Internal Cracks under Various Boundary Conditions," Trans. Jpn. Soc. Mech. Eng., **47**, pp. 27–35.

[6] Kachanov, M., 1993, "Elastic Solids With Many Cracks and Related Problems," *Advances in Applied Mechanics*, Vol. 30, J. W. Hutchinson, and T. Wu, eds., Academic Press, San Diego, CA, pp. 259–445.

[7] Karihaloo, B. L., and Wang, J., 1997, "On the Solution of Doubly Array of Cracks," Mech. Mater., **26**, pp. 209–212.

[8] Wang, J., Fang, J., and Karihaloo, B. L., 2000, "Asymptotic of Multiple Crack Interactions and Prediction of Effective Modulus," Int. J. Solids Struct., **37**, pp. 4261–4273.

[9] Chen, Y. Z., 1983, "An Investigation of the Stress Intensity Factor for a Finite Internally Cracked Plate by Using Variational Method," Eng. Fract. Mech., **17**, pp. 387–394.

[10] Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff, Amsterdam.

[11] Lekhnitsy, S. G., 1963, *Theory of Elasticity of an Anisotropic Body*, Holden-Day, San Francisco.

Nonlinear Time-Dependent Thermoelastic Response in a Multilayered Anisotropic Medium

T.-C. Chen[1]
e-mail: ctcx831@mail.ncku.edu.tw
Mem. ASME

S.-J. Hwang

C.-Q. Chen

Department of Mechanical Engineering, National Cheng Kung University, Tainan 701, Taiwan, R.O.C.

A time-dependent nonlinear thermoelastic problem of a multilayered anisotropic medium with a certain specific form of temperature-dependent material properties in generalized plane deformation is analyzed by flexibility/stiffness matrix technique in the article. The closed-form general solutions of temperature, displacements, and stresses can then be obtained in the Fourier and the Laplace transform domains by using the technique of Kirchhoff transformation. The effects of temperature-dependent material properties on the distributions of temperature and thermal stresses are also calculated and discussed.
[DOI: 10.1115/1.1458555]

1 Introduction

Time-dependent thermal and mechanical analysis of anisotropic multilayered medium has attracted increasing attention over the past few decades due to its wide applications. This note is mainly concerned with nonlinear time-dependent thermal stress problems of a multilayered anisotropic medium with temperature-dependent material properties. Besides the nonlinear and transient characteristics, the inherent heterogeneous and anisotropic nature of layered composites in both thermal and elastic behaviors makes the analysis of such materials very difficult and complicated. Consequently, even for a nonlinear heat conduction problem with linearly temperature-dependent material properties, the analytical solutions are only confined to the cases in a homogeneous single layer (Halle [1], Peletier [2], Suzuki et al. [3], Ozisik [4], and Tao [5]) or two-layer slab (Chang and Payne [6]). The methods of flexibility/stiffness (Thangjitham and Choi [7], Chen and Jang [8], and Chen et al. [9]) and transfer matrix (Bahar and Hetnarski [10]) are very efficient analytical matrix approaches and have been successfully developed to solve the problems of multilayered medium by the investigators abovementioned in recent years. Existing literature reveals that the nonlinear transient thermoelastic problem of a multilayered anisotropic medium with temperature-dependent material properties, however, has seldom been solved by an analytical approach. Most of the investigations are devoted to the numerical study (Noor and Tenek [11]). In this note, the nonlinear time-dependent thermoelastic problem of a multilayered anisotropic medium subjected to thermal and/or mechanical loadings with temperature-dependent material properties under the state of generalized plane deformation is dealt with by the method of a flexibility/stiffness matrix. Exact solutions are obtained only when the temperature-dependent material properties are confined to some specific expressions. The original nonlinear time-dependent differential equation of heat conduction is first linearized by the Kirchhoff transformation. By utilizing the Fourier and Laplace transform techniques, the general solutions to this linearized heat conduction and thermoelastic problem for layers with monoclinic properties are then derived. The flexibility/stiffness matrix method is then adopted to obtain the complete solution of the entire layered medium by introducing the thermal and mechanical boundary and layer interface conditions. As a numerical illustration, the distributions of time-dependent temperature and thermal stresses in a laminated anisotropic slab subjected to a uniform surface temperature rise are presented for some stacking sequences of fiber-reinforced layers. Moreover, the effects of temperature-dependent material properties on the distributions of temperature and thermal stresses are also calculated and discussed.

2 Governing Equations and General Solutions

Suppose a layered slab, as shown in Fig. 1, composed of N fiber-reinforced layers is deformed under the state of generalized plane deformation such that no normal strain occurs in the y-direction and all field variables are functions of coordinates x, z and time t only.

$$T=T(x,z,t), \tag{1a}$$

$$u=u(x,z,t), \tag{1b}$$

$$v=v(x,z,t), \tag{1c}$$

$$w=w(x,z,t) \tag{1d}$$

where T is the temperature field; and u, v, and w represent the displacement components in the x, z, and z-directions, respectively.

The temperature field satisfies the transient-state nonlinear heat conduction equation

$$\frac{\partial}{\partial x}\left(\tilde{k}_{11}(T)\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial z}\left(\tilde{k}_{33}(T)\frac{\partial T}{\partial z}\right)=\rho(T)c(T)\frac{\partial T}{\partial t} \tag{2}$$

where $\tilde{k}_{ij}(T)$, $i,j=1,2,3$ are the coefficients of heat conductivity, which are temperature-dependent, in the structural coordinates of the medium as shown in Fig. 1.

The relations between heat fluxes and temperature gradients are

$$\begin{Bmatrix} q_x \\ q_y \\ q_z \end{Bmatrix}=\begin{bmatrix} \tilde{k}_{11}(T) & \tilde{k}_{12}(T) & 0 \\ \tilde{k}_{12}(T) & \tilde{k}_{22}(T) & 0 \\ 0 & 0 & \tilde{k}_{33}(T) \end{bmatrix}\begin{Bmatrix} -\dfrac{\partial T}{\partial x} \\ 0 \\ -\dfrac{\partial T}{\partial z} \end{Bmatrix} \tag{3}$$

where q_x, q_y, and q_z are the heat fluxes along x, y, and z-direction, respectively.

The Kirchhoff transformation is then adopted by introducing a function, U, which is called the equivalent or effective temperature defined as

$$U=\int_{T_0}^{T}\frac{\tilde{k}_{11}(T)}{\tilde{k}_{110}}dT \tag{4}$$

where $\tilde{k}_{ij0}$ is the material conductivity coefficients at reference temperature T_0.

After using this transformation, the nonlinear transient heat conduction Eq. (2) becomes a more convenient form as follows:

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 8, 2001; final revision, Oct. 5, 2001. Associate Editor: L. T. Wheeler.

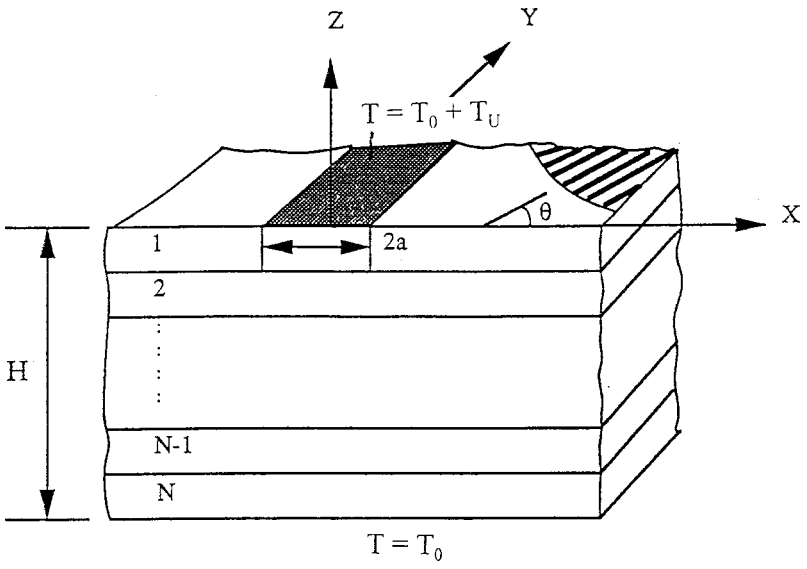


Fig. 1 Configuration of *N*-layer anisotropic medium

$$\frac{\partial^2 U}{\partial x^2}+\frac{\partial}{\partial z}\left[\left(\frac{\tilde{k}_{33}}{\tilde{k}_{11}}\right)\frac{\partial U}{\partial z}\right]=\left(\frac{\rho c}{\tilde{k}_{11}}\right)\frac{\partial U}{\partial t} \tag{5}$$

where ρ and c denote the density and the specific heat, respectively.

The heat conduction coefficients of material are assumed to be linear functions of the temperature as follows:

$$\tilde{k}_{ij}=\tilde{k}_{ij0}[1+m_1(T-T_0)]. \tag{6}$$

Therefore, $\tilde{k}_{33}/\tilde{k}_{11}=\tilde{k}_{330}/\tilde{k}_{110}$ and the relation between U and T can be obtained:

$$U(x,y,z)=T(x,y,z)-T_0+\frac{m_1}{2}[T(x,y,z)-T_0]^2. \tag{7}$$

Moreover, the equations of heat fluxes (3) can also be converted into the following equations in terms of U:

$$q_x=-\tilde{k}_{110}\frac{\partial U}{\partial x},\quad q_y=-\tilde{k}_{120}\frac{\partial U}{\partial x},\quad q_z=-\tilde{k}_{330}\frac{\partial U}{\partial z}. \tag{8}$$

Consequently, it is obvious that Eq. (5) becomes linear either for steady-state problems or for time-dependent problems in case the thermal diffusivity, $\tilde{k}_{11}/\rho c$, is independent of the temperature. In this note, the problem studied is only confined to the condition that the term of $\tilde{k}_{11}/\rho c$, is also temperature-independent. The thermal expansion coefficients of material, α_{ij}, are also assumed to be linear functions of the temperature as follows:

$$\alpha_{ij}=\alpha_{ij0}[1+m_2(T-T_0)] \tag{9}$$

where α_{ij0} is the thermal expansion coefficients at reference temperature T_0.

Under the condition of generalized plane thermoelasticity the equilibrium equations of a given layer expressed in terms of the displacement components and temperature can then be rearranged and reduced to the linear forms expressed as

$$\tilde{C}_{11}\frac{\partial^2 u}{\partial x^2}+\tilde{C}_{55}\frac{\partial^2 u}{\partial z^2}+\tilde{C}_{16}\frac{\partial^2 v}{\partial x^2}+\tilde{C}_{45}\frac{\partial^2 v}{\partial z^2}+(\tilde{C}_{13}+\tilde{C}_{55})\frac{\partial^2 w}{\partial x\partial z}=\beta_{10}\frac{\partial V}{\partial x} \tag{10a}$$

$$\tilde{C}_{16}\frac{\partial^2 u}{\partial x^2}+\tilde{C}_{45}\frac{\partial^2 u}{\partial z^2}+\tilde{C}_{66}\frac{\partial^2 v}{\partial x^2}+\tilde{C}_{44}\frac{\partial^2 v}{\partial z^2}+(\tilde{C}_{36}+\tilde{C}_{45})\frac{\partial^2 w}{\partial x\partial z}=\beta_{60}\frac{\partial V}{\partial x} \tag{10b}$$

$$(\tilde{C}_{13}+\tilde{C}_{55})\frac{\partial^2 u}{\partial x\partial z}+(\tilde{C}_{36}+\tilde{C}_{45})\frac{\partial^2 v}{\partial x\partial z}+\tilde{C}_{55}\frac{\partial^2 w}{\partial x^2}+\tilde{C}_{33}\frac{\partial^2 w}{\partial z^2}=\beta_{30}\frac{\partial V}{\partial z} \tag{10c}$$

where $\tilde{C}_{ij}(i,j=1\sim 3)$ denote the elastic stiffness constants; and

$$\beta_{10}=\tilde{C}_{11}\tilde{\alpha}_{110}+\tilde{C}_{12}\tilde{\alpha}_{220}+\tilde{C}_{13}\tilde{\alpha}_{330}+\tilde{C}_{16}\tilde{\alpha}_{120} \tag{11a}$$

$$\beta_{30}=\tilde{C}_{13}\tilde{\alpha}_{110}+\tilde{C}_{23}\tilde{\alpha}_{220}+\tilde{C}_{33}\tilde{\alpha}_{330}+\tilde{C}_{36}\tilde{\alpha}_{120} \tag{11b}$$

$$\beta_{60}=\tilde{C}_{16}\tilde{\alpha}_{110}+\tilde{C}_{26}\tilde{\alpha}_{220}+\tilde{C}_{36}\tilde{\alpha}_{330}+\tilde{C}_{66}\tilde{\alpha}_{120} \tag{11c}$$

$$V=(T-T_0)+\frac{m_2}{2}(T-T_0)^2. \tag{11d}$$

It can be seen that if $m_2=m_1$, then $U=V$; moreover, if $m_2=0$, then $V=T$. Since the heat conduction equation contains no elastic energy term, and the inertia is neglected in the momentum balance equations, these analyses may be, in general, referred to as weak coupled thermoelasticity in thermal stress literature (Boley and Weiner [12]).

To solve the linearized field Eqs. (5) and (10*a*) to (10*c*), the Fourier and Laplace transforms are applied over the variables x and t. The transform pairs for an arbitrary function $g(x,t)$ are defined as

$$\bar{g}(s,t)=\int_{-\infty}^{\infty}g(x,t)e^{isx}dx,\ g(x,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\bar{g}(s,t)e^{-isx}ds \tag{12a}$$

$$\bar{g}^*(s,t)=\int_0^{\infty}\bar{g}(s,t)e^{-pt}dt,\ \bar{g}(s,t)=\frac{1}{2\pi i}\int_{v-i\infty}^{v+i\infty}\bar{g}^*(s,p)e^{pt}dp \tag{12b}$$

where an overbar denotes the Fourier transformed quantity; a superscript * represents the Laplace transformed quantity; and s and p are the Fourier and Laplace transform parameters, respectively.

In the Fourier and Laplace transform domain, the governing equations for the thermoelasticity, Eqs. (10*a*)–(10*c*), under the condition of $m_1 = m_2$ can be arranged and rewritten by a system of linear nonhomogeneous ordinary differential equations expressed as follows:

$$\mathbf{M}_1 \frac{\mathbf{d}^2 \overline{\psi}^*}{\mathbf{dz}^2} + \mathbf{M}_2 \frac{\mathbf{d}\overline{\psi}^*}{\mathbf{dz}} + \mathbf{M}_3 \overline{\psi}^* = \overline{\psi}_0^* \tag{13a}$$

where $\overline{\psi}^*(s,z,p) = \{\overline{u}^*(s,z,p)\overline{v}^*(s,z,p)\overline{w}^*(s,z,p)\}^T$ is the displacement vector in the Fourier and Laplace transform domain; $\mathbf{M}_i(s,p)$, $i = 1,2,3$, are the 3×3 symmetric matrices which are functions of elastic stiffness constants and transform parameters s and p; and the vector $\overline{\psi}_0^*(s,z,p)$ is related to equivalent temperature field $\overline{U}^*(s,z,p)$ having the form

$$\overline{\psi}_0^*(s,z,p) = \left\{ -is\beta_{10}\overline{U}^*(s,z,p) - is\beta_{60}\overline{U}^*(s,z,p)\beta_{30}\frac{d}{dz}\overline{U}^*(s,z,p) \right\}. \tag{13b}$$

Consequently, it is easy to solve Eq. (13) to give the displacements and stresses in the transform domain ([7]).

Similarly, under the condition that the field variables and their first derivatives with respect to x vanish as $x \to \pm\infty$, and the initial value is zero, the heat conduction of Eq. (5) can be expressed as

$$\frac{\partial^2 \overline{U}^*}{\partial z^2} + \left(\frac{\rho c p + s^2 \tilde{k}_{11}}{\tilde{k}_{33}} \right) \overline{U}^* = 0. \tag{14}$$

Consequently, it is easy to solve Eq. (14) to give the equivalent temperature field in the transform domain as

$$\overline{U}^*(s,z,p) = H_1 \cosh(\kappa z) + H_2 \sinh(\kappa z) \tag{15}$$

where $H_j(s,p)$, $j = 1,2$, are the unknown constants to be evaluated from the proper boundary conditions and $\kappa = s\sqrt{(\rho c p/s^2 + \tilde{k}_{11})/\tilde{k}_{33}}$. It is noted that the variables s and p in the above equation are regarded as parameters. When the Laplace transform parameter $p = 0$, then $\kappa = s\sqrt{\tilde{k}_{11}/\tilde{k}_{33}}$, and this case reduces to the condition of steady-state. The transformed heat fluxes are then obtained by taking the Fourier and Laplace transforms of Eq. (8) and using Eq. (15).

3 Flexibility and Stiffness Matrix Formulation

The flexibility matrix equation of heat conduction can be formulated by establishing the relations between equivalent temperature and heat flux in the transform domains on the upper and lower surfaces of each layer. For an imperfectly bonded N-layer medium with interlayer thermal resistance subjected to arbitrary temperature variations on the bounding surfaces, the thermal boundary and interface conditions are prescribed as follows:

$$\overline{U}_1^{*+} = \overline{U}_U^* \tag{16a}$$

$$\overline{q}_k^{*-} = \overline{q}_{k+1}^{*+}, \tag{16b}$$

$$\overline{q}_k^{*-} = \frac{1}{R_k}(\overline{U}_{k+1}^{*+} - \overline{U}_k^{*-}), \quad k = 1,2,\ldots,(N-1) \tag{16c}$$

$$\overline{U}_N^{*-} = \overline{U}_L^* \tag{16d}$$

where $\overline{U}_k^{*-}$ and $\overline{U}_{k+1}^{*+}$ are the transformed equivalent temperature on the lower $(-)$ and upper $(+)$ surfaces of kth and $(k+1)$th layers, respectively; $\overline{q}_k^{*-}$ and $\overline{q}_{k+1}^{*+}$ are the transformed transverse heat fluxes on the lower and upper $(+)$ surfaces of kth and $(k+1)$th layers, respectively; R_k is the interlayer thermal contact resistance, which is defined as the reciprocal of interlayer thermal conductance h between the kth and $(k+1)$th layers, i.e., $R_k = 1/h$. For a perfect bonded interface, R_k is equal to null, and the temperature distributions become continuous between two adjacent surfaces, i.e., $\overline{U}_k^{*-} = \overline{U}_{k+1}^{*+}$; $\overline{U}_U^*$ and $\overline{U}_L^*$ represent the transformed equivalent temperature on the upper and lower surfaces of the multilayered medium, respectively.

By denoting $\overline{q}_1^* = \overline{q}_1^{*+}$ and $\overline{q}_{N+1}^* = \overline{q}_N^{*-}$ the transformed transverse heat fluxes on the upper and lower bounding surfaces of the multilayered medium, respectively, and $\overline{q}_{k+1}^* = \overline{q}_k^{*-} = \overline{q}_{k+1}^{*+}$, $k = 1,2,\ldots,(N-1)$, the common interfacial transverse heat fluxes between the kth and $(k+1)$th layers, the successive applications of the equations in Eqs. (16*a*)–(16*d*) lead to the global flexibility matrix equations for the N-layer medium written as

$$F_{11}^1 \overline{q}_1^* + F_{12}^1 \overline{q}_2^* = \overline{U}_U^* \tag{17a}$$

$$F_{12}^k \overline{q}_k^* + (F_{22}^k + F_{11}^{k+1} - R_k)\overline{q}_{k+1}^* + F_{12}^{k+1}\overline{q}_{k+2}^* = 0, \quad k = 1,2,\ldots,(N-1) \tag{17b}$$

$$F_{12}^N \overline{q}_N^* + F_{22}^N \overline{q}_{N+1}^* = -\overline{U}_L^*. \tag{17c}$$

The above equations can also be expressed in matrix form as

$$F\overline{q}^* = \overline{U}^* \tag{18}$$

where F is an $(N+1)\times(N+1)$ symmetric global flexibility matrix with half-bandwidth two and is the given function of thermal properties, the thickness of each layer and transformed parameters p and s; $\overline{q}^*$ is the global vector containing the unknown interfacial transverse heat fluxes in the transformed domain, i.e., $\{\overline{q}_1^* \overline{q}_2^* \ldots \overline{q}_{N+1}^*\}$, and $\overline{U}^*$ is the vector containing the transformed equivalent surface temperature and zero elements.

The stiffness matrix equations of moment balance can be formulated in the same way by establishing the relations between stresses and displacements in the transform domains on the upper and lower surfaces of each layer. By using the boundary conditions of applied tractions and/or displacements at the upper and the lower surfaces of the multilayered medium, the continuity conditions of transformed normal stress, $\overline{\sigma}_{zz}^*$, transformed shear stresses, $\overline{\tau}_{xz}^*$ and $\overline{\tau}_{yz}^*$, as well as the continuity conditions of transformed displacements, $\overline{u}^*$, $\overline{v}^*$ and $\overline{w}^*$, a system of algebraic equations can be formulated and expressed as

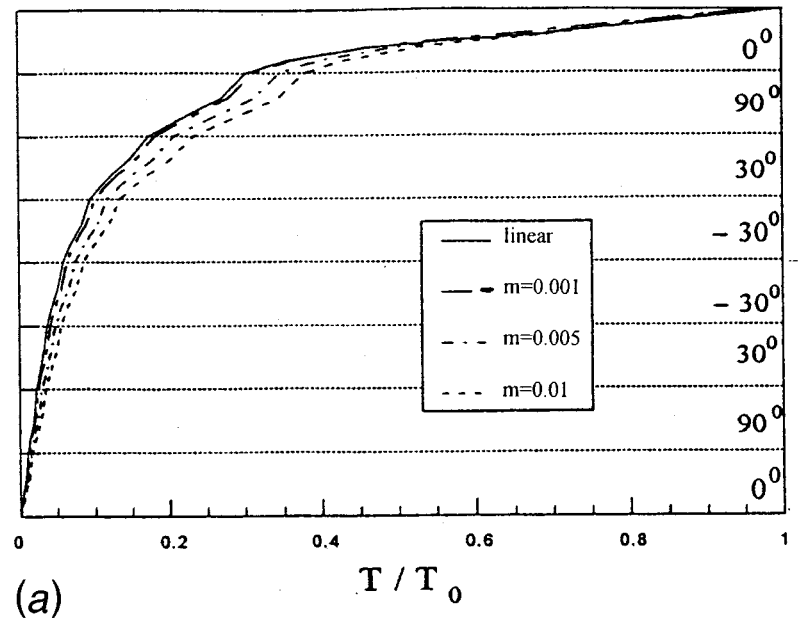
$$K\overline{\delta}^* = \overline{f}^* + \overline{r}^* \tag{19}$$

where K is a $3(N+1)\times3(N+1)$ symmetric global flexibility matrix with half-bandwidth six and is the given function of elastic stiffness constants, the thickness of each layer and transformed parameters p and s; $\overline{\delta}^*$ is a global vector containing the unknown displacements in the transform domain, i.e.,

$$\{i\overline{w}_1^* \overline{u}_1^* \overline{v}_1^* i\overline{w}_2^* \overline{u}_2^* \overline{v}_2^* \ldots i\overline{w}_{N+1}^* \overline{u}_{N+1}^* \overline{v}_{N+1}^*\},$$

$\overline{f}^*$ is a vector of transformed traction forces applied on the surface and zero elements, and $\overline{r}^*$ is a vector related to equivalent temperature.

The solution procedures are to solve the Eq. (18) to evaluate the heat fluxes and the equivalent temperature first and then to determine the displacements and the stresses by Eq. (19). Moreover, the inversion integral of Fourier transform with respect to Eq. (12*a*) is evaluated by numerical integration based upon trapezoidal rule, while the inversion integral of Laplace transform for Eq. (12*b*) is numerically carried out by Fourier series expansion developed by Durbin [13].



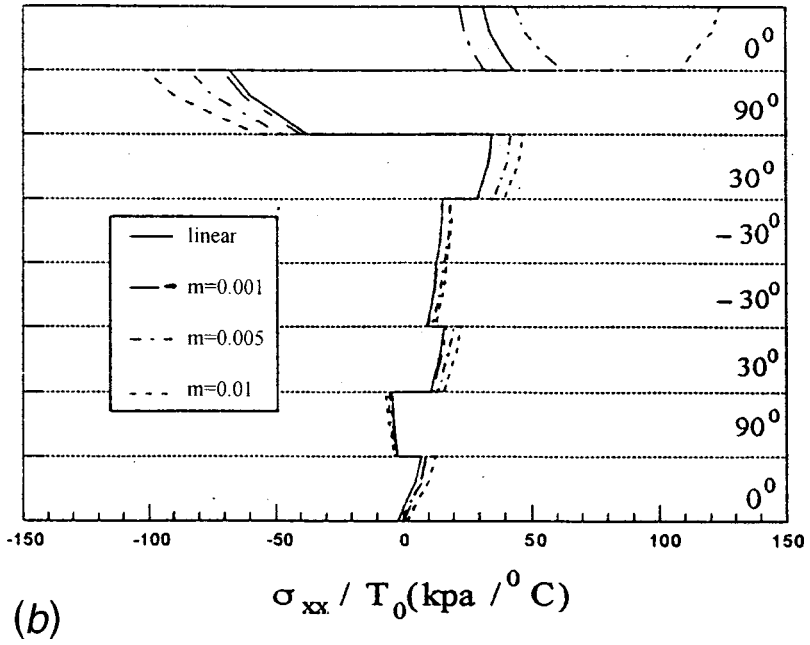


Fig. 2 Distributions of (*a*) temperature *T*, (*b*) distributions of normal stress σ_{xx}, (*c*) distributions of shear stress τ_{xy}, (*d*) normal stress σ_{yy} at time instant $t=2$ min. without interlayer thermal resistance ($\theta=30$ deg)

4 Numerical Examples and Discussions

Thermal and mechanical boundary conditions of a $[0\ \text{deg}/90\ \text{deg}/\theta/-\theta]_S$ balanced symmetrically laminated slab are prescribed as follows:

$$T_1^+(x)=\begin{Bmatrix} T_U+T_0; & |x|\leqslant a \\ T_0; & \text{otherwise} \end{Bmatrix} \tag{20a}$$

$$T_N^-(x)=T_0;\quad |x|<\infty \tag{20b}$$

$$\sigma_{ZZ}^{1+}(x)=\tau_{XZ}^{1+}(x)=\tau_{YZ}^{1+}(x)=0;\quad |x|<\infty \tag{20c}$$

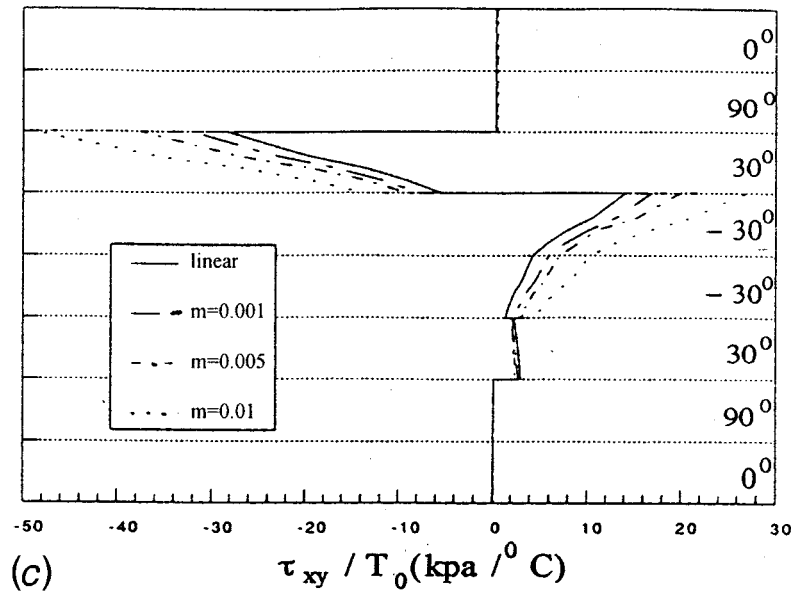
$$\sigma_{ZZ}^{N-}(x)=\tau_{XZ}^{N-}(x)=\tau_{YZ}^{N-}(x)=0;\quad |x|<\infty. \tag{20d}$$

These equations imply that a uniform temperature rise T_U is imposed in a region of $2a$ on the top surface of the laminated medium, while the surfaces otherwise remain the original temperature T_0. Moreover, all the surfaces are free of traction.

Therefore, the equivalent temperature distribution on the top surface of the slab becomes

$$U_1^+(x)=U_U(x)=\begin{Bmatrix} T_U+\dfrac{m_1}{2}T_U^2; & |x|\leqslant a \\ 0; & \text{otherwise} \end{Bmatrix}. \tag{21}$$

Taking Fourier and Laplace transforms over Eq. (21) yields



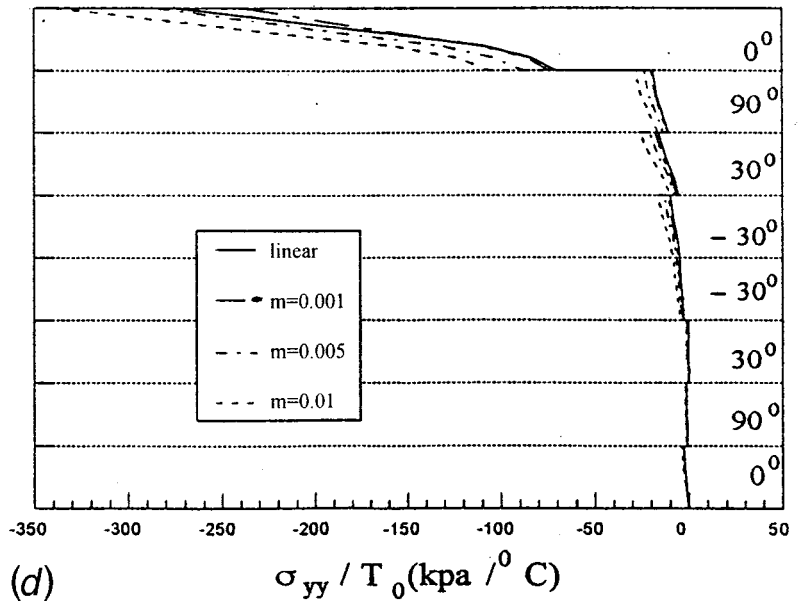


Fig. 2 (***continued***)

$$\tilde{U}_1^{+*}(s,z,p)=\frac{\sin sa}{ps}(2T_U+m_1T_U^2) \tag{22}$$

while the thermal and mechanical boundary conditions of Eqs. (20*b*)–(20*d*) in the Fourier and Laplace transform domains are still equal to zero. The same thermal and elastic properties as given in the literature ([8,9]) are adopted for our following numerical calculation in order to make comparison and investigate the effects of temperature-dependent material properties on time-dependent thermoelastic response. The unidirectional fiber-reinforced constituting layer is composed of fiber T300 graphite and matrix epoxy. For a specific fiber volume fraction $V_f=0.5$, the gross thermal and elastic properties at reference temperature are obtained as

$C_{110}=114$ Gpa, $C_{220}=C_{330}=8.7$ Gpa, $C_{120}=C_{130}=3.3$ Gpa,

$C_{230}=3.4$ Gpa, $C_{440}=2.7$ Gpa, $C_{550}=C_{660}=8.7$ Gpa,

$k_{110}=42.1$ W/m·K, $k_{220}=k_{330}=0.466$ W/m·K,

$\alpha_{110}=0.025\times10^{-6}$ m/m·K, $\alpha_{220}=\alpha_{330}=32.4\times10^{-6}$ m/m·K,

$\rho_0=1.44\times10^3$ kg/m^3, $c_0=1.76\times10^3$ J/kg·K.

Eight layered composite slabs with the same layer thickness 1.25 mm and fiber angles 30 and 90 deg. are considered, respectively, in the article. The parameters of temperature-dependent material properties for coefficients of thermal conductivity, m_1, and thermal expansion coefficients, m_2, are assumed to be identical (denoted thereafter by m) and equal to 0.001, 0.005, and 0.01, respectively. The dimension of a heated area is taken to be $2a=H=10$ mm, where H is the overall thickness of the slab. Figure 2 shows the distributions of temperature and stress components σ_{xx}, σ_{yy}, and τ_{xy}, through the thickness at $x=0$ with fiber angle $\theta=30$ deg for different nonlinear parameter m at time instant $t=2$ min without interlayer thermal contact conductance.

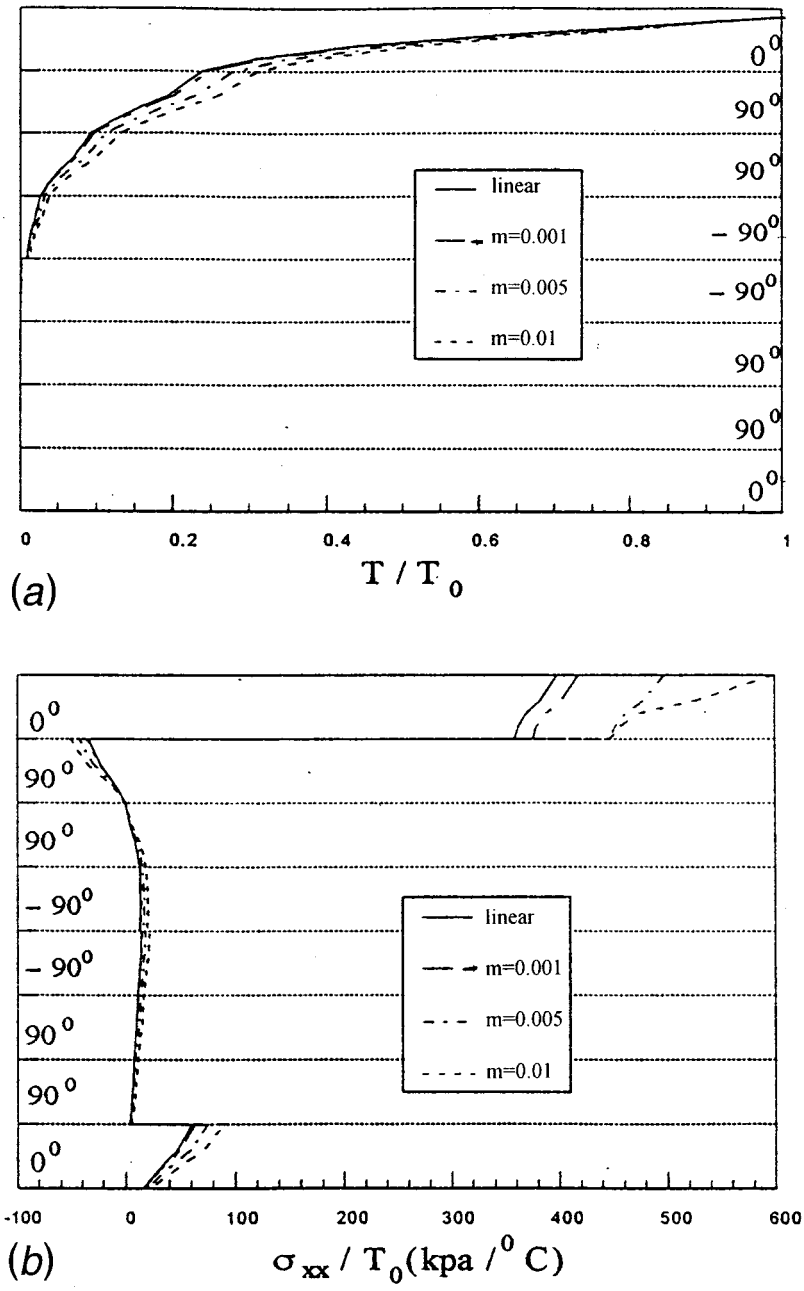
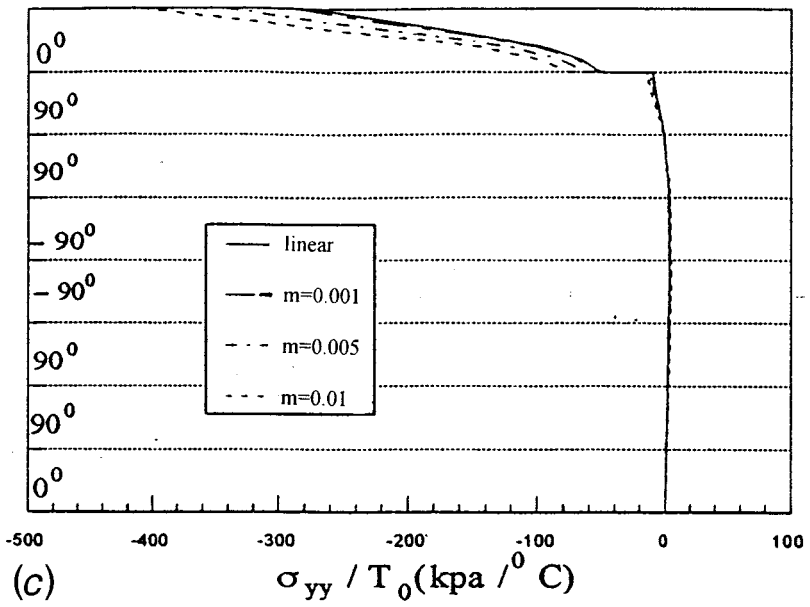


Fig. 3 Distributions of (*a*) temperature *T*, (*b*) normal stress σ_{xx}, (*c*) normal stress σ_{yy}, (*d*) normal stress σ_{zz} at time instant $t=10$ sec. with interlayer thermal resistance at interface between fourth and fifth layers ($h=100$ W/m^{2}K, $\theta=30$ deg)

Figure 3 shows the distributions of temperature and stress components σ_{xx}, σ_{yy}, and σ_{zz}, through the thickness at $x=0$ with fiber angle $\theta=90$ deg for different nonlinear parameter m at time instant $t=10$ sec with interlayer thermal contact conductance $h=100\ \mathrm{W/m^2\cdot K}$ at the interface between the fourth and fifth layers only. It can be seen that the temperature is continuously distributed through the thickness of the medium if there is no interlayer thermal resistance. However, the temperature distribution becomes discontinuous, as expected, at the interface between two layers having interlayer thermal resistance. Moreover, the distribution of stress component, σ_{zz}, is continuous through the thickness as expected; while the distributions of stress components, σ_{xx}, σ_{yy}, and τ_{xy}, are not continuous through the thickness, because the latter are not the continuity conditions of layered medium as described in the previous section. Since the value of thermal conductivity is assumed to increase with increasing temperature, the distributions of temperature predicted by nonlinear theory are, therefore, always higher than by linear theory due to a faster heat conduction rate. Consequently, the induced thermal stresses evaluated by nonlinear theory are also larger than by the linear theory. The maximum discrepancy in thermal stresses between two theories is significant and approximately more than 20 percent even when $m=0.001$.

5 Concluding Remark

The nonlinear time-dependent thermoelastic response of a multilayered anisotropic slab subjected to thermal loading with



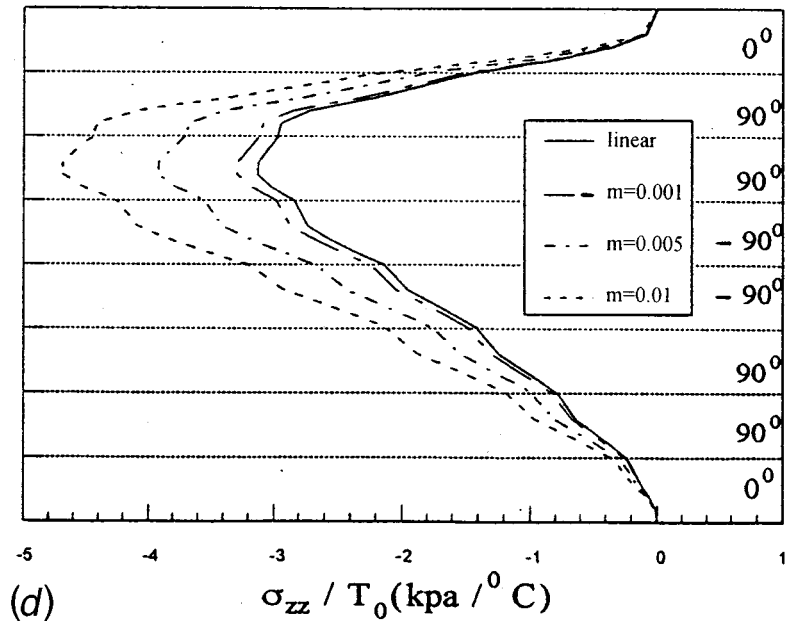


Fig. 3 (***continued***)

temperature-dependent material properties are analyzed by the technique of flexibility/stiffness matrix. Compared to the conventional solution procedure, it is well known that this method provides a higher computational efficiency by significantly reducing the number of equations that must be solved for the required unknown constants. It is seen that the closed-form general solutions are possible to obtain in the transform domains only in the case in which the material properties exhibit some specific linear functions of temperature. The analytical solutions and results under special condition given here may not only be used as a tool in design and analysis but also serve a useful purpose in providing a check upon other solution procedures. The results show that the effects of nonlinear parameter on the distributions of temperature and stress become appreciable if the nonlinear parameter m becomes larger. The discrepancy between linear and nonlinear solutions will be more than 20 percent when the nonlinear parameter m is larger than 0.001.

Acknowledgments

This work was supported by the National Science Council of Taiwan, R.O.C. (Grant No. NSC 84-2212-E-006-029).

References

[1] Halle, H., 1965, "Exact Solution of Elementary Transient Heat Conduction Problem Involving Temperature-Dependent Material Properties," ASME J. Heat Transfer, **87**, pp. 420–421.
[2] Peletier, L. A., 1970, "Asymptotic Behavior of Temperature Profiles of a Class of Nonlinear Heat Conduction Problems," Q. J. Mech. Appl. Math., **23**, pp. 441–447.
[3] Suzuki, M., Matsumoto, S., and Maeda, S., 1977, "New Analytical Method for Nonlinear Diffusion Problem," Int. J. Heat Mass Transf., **20**, pp. 883–889.
[4] Ozisik, M. N., 1980, *Heat Conduction*, John Wiley and Sons, New York, Chapter 11.
[5] Tao, L. N., 1989, "The Heat Conduction Problem With Temperature-Dependent Material Properties," Int. J. Heat Mass Transf., **32**, pp. 487–491.

[6] Chang, K. C., and Payne, U. J., 1991, ''Analytical Solution for Heat Conduction in a Two-Material-Layer Slab With Linearly Temperature Dependent Conductivity,'' ASME J. Heat Transf., **113**, pp. 237–239.
[7] Thangjitham, S., and Choi, H. J., 1991, ''Thermal Stresses in a Multilayered Anisotropic Medium,'' ASME J. Appl. Mech., **58**, pp. 1021–1027.
[8] Chen, T. C., and Jang, H. I., 1995, ''Thermal Stresses in a Multilayered Anisotropic Medium With Interface Thermal Resistance,'' ASME J. Appl. Mech., **62**, pp. 610–611.
[9] Chen, T. C., Jang, H. I., and Tseng, A. A., 1995, ''Transient Thermal Stresses in a Multilayered Anisotropic Medium,'' ASME J. Appl. Mech., **62**, pp. 1065–1067.
[10] Bahar, L. Y., and Hetnarski, R. B., 1980, ''Coupled Thermoelasticity of a Layered Medium,'' J. Therm. Stresses, **3**, pp. 141–152.
[11] Noor, A. K., and Tenek, L. H., 1992, ''Steady-State Nonlinear Heat Transfer in Multilayered Composite Panels,'' J. Eng. Mech., **118**, No. 8, pp. 1661–1678.
[12] Boley and Weiner, 1985, *Theory of Thermal Stresses*, Krieger, Melbourne, FL.
[13] Dubin, F., 1974, ''Numerical Inversion of Laplace Transform: An Efficient Improvement to Dubner and Abates's Method,'' Comput. J., **17**, pp. 371–376.