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. ded Piezoelectric Material Strip

Journal of Applied Mechanics

Published Bimonthly by ASME

VOLUME 71 • NUMBER 5 • SEPTEMBER 2004

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 71, Number 5** **SEPTEMBER 2004**

BRIEF NOTES

ANNOUNCEMENTS AND SPECIAL NOTES

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

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

X. Chen
Department of Civil Engineering and Engineering Mechanics,
Columbia University,
New York, NY 10027
Mem. ASME

John W. Hutchinson
Division of Engineering and Applied Sciences,
Harvard University,
Cambridge, MA 02138
Fellow ASME

Herringbone Buckling Patterns of Compressed Thin Films on Compliant Substrates

A thin metal film vapor deposited on thick elastomer substrate develops an equi-biaxial compressive stress state when the system is cooled due to the large thermal expansion mismatch between the elastomer and the metal. At a critical stress, the film undergoes buckling into a family of modes with short wavelengths characteristic of a thin plate on a compliant elastic foundation. As the system is further cooled, a highly ordered herringbone pattern has been observed to develop. Here it is shown that the herringbone mode constitutes a minimum energy configuration among a limited set of competing modes. [DOI: 10.1115/1.1756141]

1 Introduction

Recent studies (Bowden et al. [1], Huck et al. [2], and Yoo et al. [3]) have explored the feasibility of manipulating buckled films on compliant substrates to achieve highly ordered patterns with distinctive features. A metal film with a thickness measured in tens of nanometers can be vapor deposited at an elevated temperature on a thick elastomer substrate. When the film/substrate system is cooled, the large mismatch between the thermal expansion of the metal and elastomer produces a state of equi-biaxial compression in the film. At a critical temperature the film begins to buckle into modes with wavelengths typically measured in microns. As the temperature is further lowered the amplitude of the buckles grow and distinctive patterns emerge. Buckle patterns can be manipulated either by creating nonplanar substrate topography prior to deposition, [1,2], or by forcing a film on a smooth planar substrate to buckle into a nonplanar mold that is subsequently removed, [3]. If the surface of the substrate is nonplanar at the start of deposition, the buckle pattern is influenced by the underlying topography. On the right-hand side of Fig. 1, one sees the buckle pattern that arises when the substrate surface has a series of flat depressions running linearly in the vertical direction. Due to local deformation of the substrate, the in-plane film stress is relaxed in the direction perpendicular to the edges of the depressions resulting in the buckle alignment shown, [1,2]. This alignment persists well beyond the left-most depression edge in Fig. 1 where the substrate is perfectly smooth and flat prior to buckling. The parallel undulations in the left-center of Fig. 1 transition to the herringbone pattern. The herringbone buckle pattern appears to be the preferred mode whenever there exists a sufficiently large patch of smooth substrate and when the system has been cooled well below the onset of buckling. There are irregularities to the herringbone pattern, including local distortion most likely due to imperfections in either the film or substrate. Another example is shown in Fig. 2 where the substrate is pre-patterned with a single circular depression of several millimeters diameter at its center. The depression edge determines the orientation of the buckles in its vicinity, but away from the edge the herringbone pattern emerges. The example in Fig. 3, which is taken from [3], shows the highly ordered mode that forms when a mold with a square pattern is held against the film as the buckles form and is then removed when the buckles are fully developed. In this case the film is aluminum and the substrate has two layers, a relatively thin compliant polymer layer (polystyrene) bonded to a thick silicon layer.

All the films considered in this paper remain bonded to the substrate in the buckled state. They undergo little, or no, plastic deformation. Selected tests revealed that the buckles almost entirely disappeared when the temperature was reversed, [1,2]. The fact that the film locks into modes that are very different from what would be expected from a linear buckling analysis is due to the highly nonlinear character of buckling of an elastic film on a compliant substrate as the temperature drops well below critical. Nonlinearity also accounts for the unusual, highly ordered herringbone mode of buckling in Figs. 1 and 2 that is observed when the system is not manipulated in any way. In Figs. 1 and 2, the substrate is the elastomer, polydimethylsiloxane, (or PDMS, for short) and the film is 50 nm of gold. The crest-to-crest separation of the buckle undulations or ''waves'' is about 30 μm and the distance between jogs in the herringbone mode is about 100 μm. The change in direction of the waves at each jog is approximately 90 deg. The amplitude of the waves is on the order of a micron or smaller. Thus, although the amplitude is large compared to the film thickness, the mode is shallow in the sense that the slopes of the pattern are small. The strains associated with the buckling mode are also small, and both the film and the substrate materials are within their respective linear elastic ranges.

The herringbone pattern is very different from any mode one might suspect based on a linear stability analysis, as will be seen later. A clue to its existence is its ability to alleviate equally in all directions the biaxial in-plane stress driving buckling. A mode with undulations extending in only one direction such as that seen on the right in Fig. 1, which will be referred to hereafter as a *one-dimensional mode*, relieves in-plane stress only in the direction perpendicular to the undulations. The in-plane stress component parallel to the undulations is only slightly altered by buckling. On the other hand, the alternating directions of the local one-dimensional undulations in the herringbone mode reduce the overall in-plane stress in the film in all directions. The herringbone mode allows for an isotropic average in-plane expansion of the film, but otherwise has zero Gaussian curvature apart from the vicinity of the jogs. In addition to bending, some nonuniform stretching of the film necessarily occurs locally at the jogs. Nevertheless, the near-inextensionality (apart from the average uniform expansion) of the herringbone mode and its ability to alleviate the in-plane stress equally in all directions are the two features underlying its preferred existence. An origami pattern similar to it can be created from a series of folds of a piece of paper.

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 Applied Mechanics Division, May 14, 2003; final revision October 30, 2003. Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

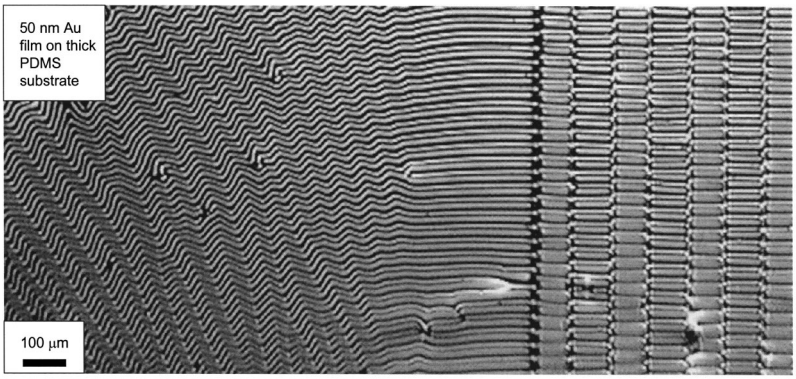


Fig. 1 Buckling of a 50 nm gold film on a thick elastomer (PDMS) substrate. On the right, the substrate has been patterned with alternating flat depressions, [2]. The substrate on the left two-thirds of the figure is flat and not patterned. The herringbone pattern is on the left. The wavelength of the pattern across the crests is approximately 30 μm while the distance between jogs of the herringbone mode is approximately 100 μm.

It will be shown that the herringbone mode of buckling is the minimum energy configuration among several competing modes. We confine attention to modes that are periodic, and we begin by presenting the result of the classical buckling analysis for the family of modes associated with the critical stress. Then, a closed-form analysis is presented of one-dimension undulations of finite amplitude at temperatures arbitrarily below the critical stress. Numerical analyses of the herringbone mode and a square checkerboard mode follow. The film is represented as an elastic thin plate satisfying the nonlinear von Karman plate equations. These equations are accurate for the shallow modes observed. In the analytical work, the elastomer substrate is represented by linear, small strain elasticity theory. This is an accurate representation because the strains in the substrate remain small. Moreover, a linear strain-displacement characterization of the substrate is justified because the substrate experiences no pre-stress—buckling is driven by the pre-stress in the film not the substrate. In the numerical analysis of the two other modes, nonlinear kinematics holds throughout the system (with linear stress-strain behavior), but nonlinearity in the substrate is negligible. It will be seen that the energy associated with the herringbone mode is distinctly below that of the other two modes. The energy minimum of the herringbone pattern is relatively flat in the sense that there is little change in the energy for a fairly wide range of the parameters characterizing the geometry of the pattern, especially the spacing between jogs. The paper ends with speculation on how the mode forms and a discussion on the limitations of approaches based on energy minimization.

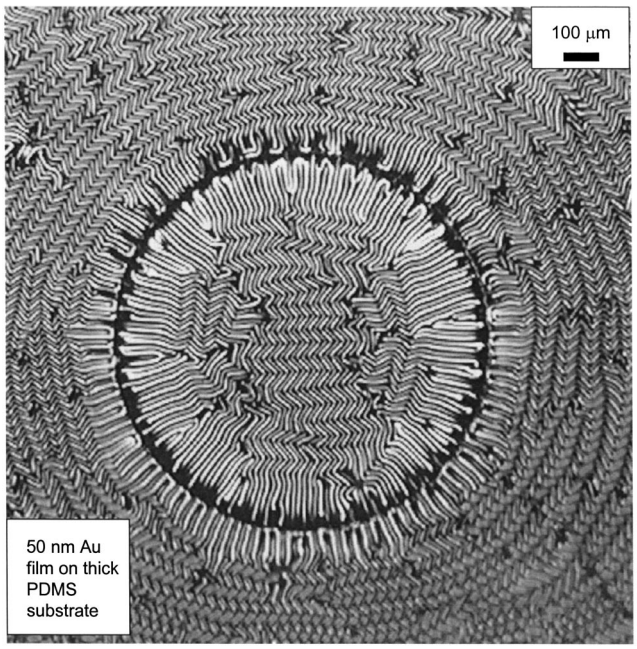


Fig. 2 Gold film on a substrate which has been patterned with a circular flat depression of several millimeters in diameter, [1]. The herringbone pattern emerges in the center of the spot and outside the edges of the spot.

The Young's modulus, Poisson's ratio and coefficient of thermal expansion of the film are denoted by E, ν, and α. The corresponding quantities for the substrate are denoted by E_s, ν_s, and α_s. The film thickness is t. The substrate is assumed to be infinitely thick and, thus, it imposes its in-plane strains on the film. Assuming the film is deposited on the substrate when both are at temperature T_D and the temperature of the system is then reduced by ΔT, and assuming the film is elastic and unbuckled, the compressive equi-biaxial pre-stress stress, σ_0, in the film is

$$\sigma_{11}=\sigma_{22}=-\sigma_0=-[E/(1-\nu)]\int_{T_D-\Delta T}^{T_D}\Delta\alpha dT \tag{1}$$

where $\Delta\alpha=\alpha_s-\alpha$. For the systems under consideration, $\Delta\alpha>0$ and $\sigma_0>0$.

The von Karman plate equations, [4], governing the deflection of the film are

$$D\nabla^4 w-(N_{11}w_{,11}+N_{22}w_{,22}+2N_{12}w_{,12})=-p \tag{2}$$

$$\frac{1}{Et}\nabla^4 F=w_{,12}^2-w_{,11}w_{,22}. \tag{3}$$

Here, ∇^4 is the bi-harmonic operator, $D=Et^3/[12(1-\nu^2)]$ is the bending stiffness of the plate, w is its displacement perpendicular to the plane, (x_1,x_2), p is the stress component acting perpendicular to the plate that is exerted by the substrate, $(\ )_{,\alpha}\equiv\partial(\ )/\partial x_\alpha$, $N_{\alpha\beta}=\int\sigma_{\alpha\beta}dx_3$ is the stress resultant in the plane of the plate and F is the Airy stress function with ($N_{11}=F_{,22}$, $N_{22}=F_{,11}$, $N_{12}=-F_{,12}$). Equation (2) is the moment equilibrium equation, and (3) is the compatibility equation ensuring the existence of in-plane displacement gradients, $u_{\alpha,\beta}$. Tangential components of the traction exerted by the substrate on the plate are ignored. This is a standard approximation in the analysis of wrinkling of a thin film on compliant substrate, [5], whose accuracy can be validated by a

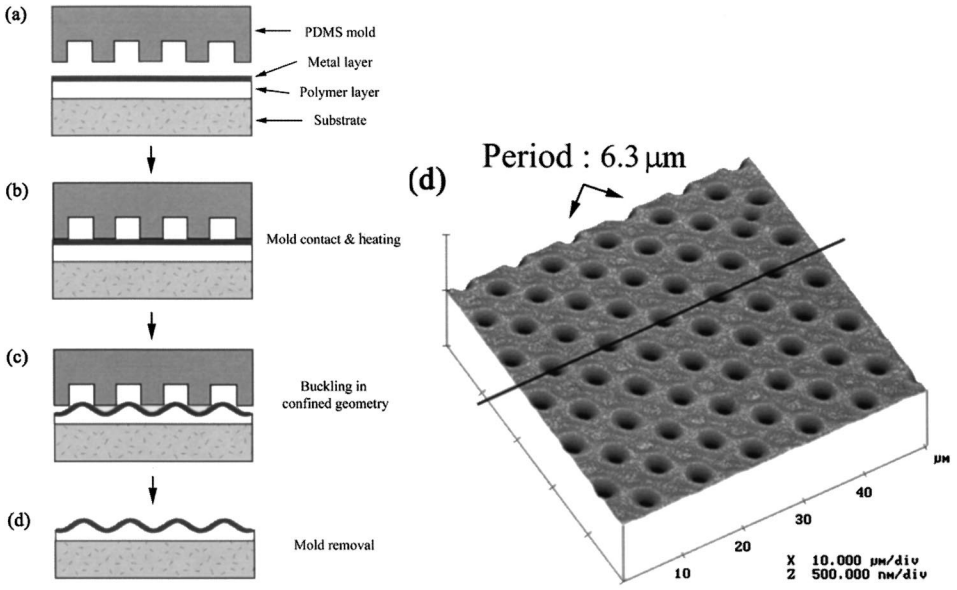


Fig. 3 An example from [3] of a highly organized buckling pattern for a film/substrate system. As depicted, the film is forced to buckle into a mold with a square pattern, after which the mold is removed.

more detailed analysis for one-dimensional modes (see below). Middle surface strains are related to displacements by $E_{\alpha\beta}=1/2(u_{\alpha,\beta}+u_{\beta,\alpha})+1/2w_{,\alpha}w_{,\beta}$; $N_{\alpha\beta}=[E/(1-\nu^2)]((1-\nu)E_{\alpha\beta}+\nu E_{\gamma\gamma}\delta_{\alpha\beta})$; and $M_{\alpha\beta}=D((1-\nu)w_{,\alpha\beta}+\nu w_{,\gamma\gamma}\delta_{\alpha\beta})$ are the constitutive relations with $M_{\alpha\beta}$ as the bending moment tensor.

2 Classical Buckling Based on Linearized Stability Analysis

The film is imagined to be infinite in extent. The unbuckled film has a uniform stress state given by $N_{11}=N_{22}=-\sigma_0 t$, $N_{12}=0$. The classical buckling analysis, based on linearization of (2) and (3) about the pre-buckling state, leads to

$$D\nabla^4 w+\sigma_0 t\nabla^2 w=-p \tag{4}$$

along with $\nabla^4\Delta F=0$ where $F=-1/2(x_1^2+x_2^2)\sigma_0 t+\Delta F$. The system of equations admits periodic solutions of the form

$$w=\hat{w}\cos(k_1x_1)\cos(k_2x_2),\quad p=\hat{p}\cos(k_1x_1)\cos(k_2x_2),$$
$$\Delta F=0 \tag{5}$$

with (4) giving $(Dk^4-\sigma_0 tk^2)\hat{w}=-\hat{p}$ where $k=\sqrt{k_1^2+k_2^2}$.

The exact solution for the normal deflection of the surface, δ, of the infinitely deep substrate under the normal loading p in (5) with zero tangential tractions at the surface is $\delta=\hat{\delta}\cos(k_1x_1)\cos(k_2x_2)$ where $\hat{\delta}=2\hat{p}/(\overline{E_s}k)$ with $\overline{E_s}=E_s/(1-\nu_s^2)$.[1] Combining $(Dk^4-\sigma_0 tk^2)\hat{w}=-\hat{p}$ and $\hat{\delta}=2\hat{p}/(\overline{E_s}k)$ subject to $\hat{w}=\hat{\delta}$ gives the eigenvalue equation $\sigma_0 t=Dk^2+\overline{E_s}/2k$. The critical buckling stress, σ_0^C, is the minimum eigenvalue with respect to k, which is attained for $k^Ct=(3\overline{E_s}/\overline{E})^{1/3}$ giving

$$\frac{\sigma_0^C}{E}=\frac{1}{4}\left(3\frac{\overline{E_s}}{E}\right)^{2/3} \tag{6}$$

with $\overline{E}=E/(1-\nu^2)$. This is the result for the one-dimensional, plane-strain wrinkling stress, which is widely known, [5]. Note, however, for an equi-biaxial pre-stress, the critical stress applies not only to the one-dimensional mode with $k_1=k^C$ and $k_2=0$ but to *any* mode whose wave numbers satisfy

$$\sqrt{k_1^2+k_2^2}t=k^Ct=(3\overline{E_s}/\overline{E})^{1/3}. \tag{7}$$

The compressed film in the equi-biaxial state has multiple modes associated the critical buckling stress. In what follows, both the one-dimensional mode and the square checkerboard mode with $k_1=k_2=k^C/\sqrt{2}$ will be investigated.

3 Nonlinear Analysis of the One-Dimensional Mode

An exact closed-form solution for the nonlinear von Karman plate coupled to the linearly elastic foundation is possible for the one-dimensional mode with nonzero k_1 and $k_2=0$. The eigenvalue (i.e., the stress at the onset of buckling) associated with arbitrary k_1 is now denoted by $\sigma_0^E t=Dk_1^2+\overline{E_s}/2k_1$ to distinguish it from the stress in the unbuckled state, σ_0. Results will be presented for various k_1 including the critical case with $k_1=k^C$ and $\sigma_0^E=\sigma_0^C$. The solution is produced for temperatures such that $\sigma_0>\sigma_0^E$.

The normal displacement in the finite amplitude state continues to be $w=\hat{w}\cos(k_1x_1)$, and (3) implies that the resultant stress in the buckled film, N_{11}, is independent of x_1. It follows, then, from (2) that the relation $N_{11}=-\sigma_0^E t$ remains in effect in the nonlinear regime. An additional constraint must be imposed to ensure that u_1 is consistent with the overall substrate deformation, i.e., $\int_0^{2\pi/k_1}\Delta u_1 dx_1=0$ where $\Delta u_1=u_1-u_1^0$ with u_1^0 as the displacement in the unbuckled film at σ_0. (Equation (3) ensures the existence of du_1/dx_1; this condition provides the underdetermined constant, ensuring the overall film displacement matches that of the substrate.) This constraint condition can be expressed using the strain-displacement relation and the stress-strain relation as

$$\frac{1}{Et}(\sigma_0 t+N_{11})=\frac{k_1}{4\pi}\int_0^{2\pi/k_1}\left(\frac{dw}{dx_1}\right)^2 dx_1=\frac{k_1^2}{4}\hat{w}^2. \tag{8}$$

Thus, the amplitude of the buckling mode is obtained by combining (8) with $N_{11}=-\sigma_0^E t$:

[1]The effect of the boundary conditions tangential to the surface of the substrate is minor. For example, if the tangential displacements are constrained to be zero, one finds $\hat{\delta}=2\hat{p}\lambda/\overline{E_s}k$ where $\lambda=(3-4\nu_s)/[4(1-\nu_s)^2]$. For $\nu_s=1/3$, $\lambda=15/16$; for $\nu_s=1/2$, $\lambda=1$.

$$\frac{\hat{w}}{t}=\frac{2}{k_1 t}\sqrt{\frac{(\sigma_0-\sigma_0^E)}{\bar{E}}}. \tag{9}$$

In the unbuckled state when the film stress is σ_0, the substrate is unstressed and the energy per unit area in the film/substrate system is

$$U_0=\frac{1-\nu}{E}\sigma_0^2 t. \tag{10}$$

The average energy per unit area in the buckled state can be expressed as

$$U=\frac{1}{2E}(N_{11}(N_{11}-\nu N_{22})+N_{22}(N_{22}-\nu N_{11}))$$
$$+\frac{k_1}{4\pi}\int_0^{2\pi/k_1} D\left(\frac{d^2w}{d^2x_1}\right)^2 dx_1+\frac{k_1}{4\pi}\int_0^{2\pi/k_1} pw\,dx_1. \tag{11}$$

In (11), the first contribution is from the uniform resultant in-plane stresses in the film, the second is the bending contribution from the film, and the third is the elastic energy in the substrate. Enforcing $\Delta E_{22}=0$, as measured from the unbuckled state at σ_0, one readily finds that $N_{22}=-(1-\nu)\sigma_0 t-\nu\sigma_0^E t$. Each of the contributions in (11) can be evaluated explicitly. In the same order as in (11), the ratios of the energy contributions in the buckled state to the energy in the unbuckled state are

$$\frac{U}{U_0}=\frac{1+\nu}{2}\left[\left(\frac{\sigma_0^E}{\sigma_0}\right)^2+\frac{(1-\nu)^2}{(1-\nu^2)}\right]+\frac{(1+\nu)(k_1 t)^4}{48}\left(\frac{\bar{E}}{\sigma_0}\right)^2\left(\frac{\hat{w}}{t}\right)^2$$
$$+\frac{(1+\nu)k_1 t}{8}\left(\frac{\bar{E}_s}{\bar{E}}\right)\left(\frac{\bar{E}}{\sigma_0}\right)^2\left(\frac{\hat{w}}{t}\right)^2. \tag{12}$$

With ν and $\bar{E}_s/\bar{E}$ specified and $\sigma_0/\bar{E}$ determined from (1), U/U_0 can be computed from (12) for any $k_1 t$ because $\sigma_0^E/\bar{E}=(k_1 t)^2/12+(\bar{E}_s/\bar{E})/(2k_1 t)$ and $\hat{w}/t$ is given by (9). For the critical mode with $k_1 t\equiv k^C t=(3\bar{E}_s/\bar{E})^{1/3}$ and $\sigma_0^E=\sigma_0^C$ given by (6), one finds $\hat{w}/t=\sqrt{\sigma_0/\sigma_0^C-1}$ and

$$\frac{U}{U_0}=\frac{1+\nu}{2}\left[\left(\frac{\sigma_0^C}{\sigma_0}\right)^2+\frac{(1-\nu)^2}{(1-\nu^2)}\right]+(1-\nu)\frac{\sigma_0^C}{\sigma_0}\left(1-\frac{\sigma_0^C}{\sigma_0}\right). \tag{13}$$

Plots of U/U_0 as a function of σ_0/σ_0^C are given in Fig. 4 for $\nu=1/3$, $\nu_s=0.48$, and $\bar{E}_s/\bar{E}=4100$, representative of a gold film on a PDMS substrate. Results for the normalized energy are presented for five values of the wavelength ratio, $L/L^C=k^C/k_1$, where $L=2\pi/k_1$ is the wavelength and $L^C\equiv 2\pi/k^C$ is the wavelength of the critical mode. The wavelength predicted for the gold film/PDMS substrate system is 3 μm which is significantly below the observed wavelength of roughly 20 to 30 μm seen in Figs. 1 and 2. The discrepancy, discussed in [1,2], is believed to be due to a layer of PDMS just below the film whose Young's modulus is much higher than the bulk elastomer due the high film deposition temperature. In effect, it is argued that there is a two-layer film whose thickness is substantially greater than the gold film. Other possibilities for the wavelength discrepancy include the possibility that the modulus used for PDMS in the range of very small strains applicable to this problem may not be correct. The experimental agreement with the theoretical wavelength prediction is considerably better in [3].

The results of Fig. 4 for U/U_0 for the one-dimensional mode show that the lowest energy state is associated with the critical mode ($L/L^C=1$) even at finite amplitude buckling deflections. The lowest energy state at values of σ_0/σ_0^C just above unity must be associated with the critical mode, but, in general, there is no a priori reason why lowest energy configuration should remain associated with the critical mode. Thus, the results of Fig. 4 emphasize the strong preference for the wavelength associated with the critical wavelength $L^C\equiv 2\pi/k^C$ when the mode is one-dimensional. The limit of the energy ratio as σ_0/σ_0^C becomes large is $U/U_0=(1-\nu)/2$. In this limit the in-plane compressive stress perpendicular to the buckles is completely relieved but the compressive stress parallel to the buckles is changed only by the Poisson effect.

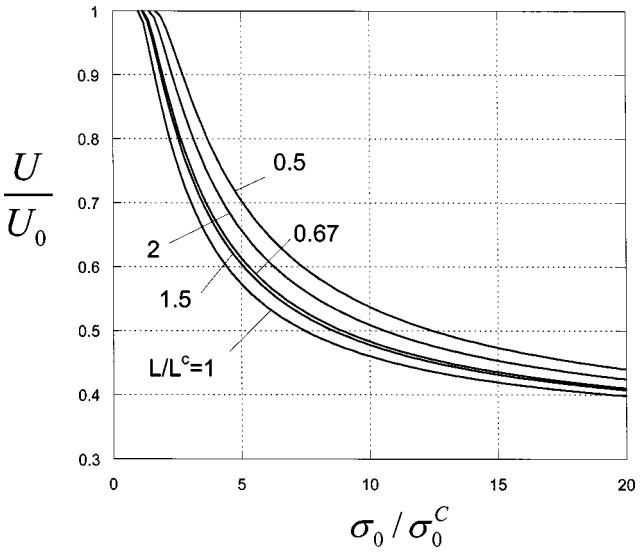


Fig. 4 Ratio of average elastic energy per unit area in the film/substrate system in the buckled state to that in the unbuckled state, U/U_0, as a function of σ_0/σ_0^C for the one-dimensional mode. Results are shown for several wavelengths. Note that the wavelength that is critical at the onset of buckling ($k_1=k_1^C$) produces the minimum energy in the buckled state even when σ_0 is well above σ_0^C.

4 Numerical Analysis of the Checkerboard and Herringbone Modes

An exact analytic solution such as that given for the one-dimensional mode cannot be obtained for either the checkerboard or the herringbone mode. The finite element code, ABAQUS, has been used to obtain a three-dimensional analysis of the periodic cell of these two modes. Within the cell, the plate is represented by 1000 three-dimensional eight-node, quadratic thin shell elements (with five degrees-of-freedom at each node and with reduced integration) that account for finite rotations of the middle surface. The stresses and strains within the plate are linearly related. The substrate is meshed with 20-node quadratic block elements with reduced integration. The constitutive relation of the substrate is also taken to be linear isotropic elasticity, but the geometry is updated. As mentioned earlier, nonlinear strain-displacement behavior of the substrate has essentially no influence on the results of interest. The substrate is taken to be very deep (depth d) compared to mode wavelength, and the boundary conditions along its bottom surface are zero normal displacement and zero tangential tractions. The film is assigned a temperature-independent thermal expansion mismatch, $\Delta\alpha$, and a temperature drop ΔT is imposed starting from the unstressed state. The biaxial compressive stress in the unbuckled film is therefore $\sigma_0=E\Delta\alpha\Delta T/(1-\nu)$ if the substrate is infinitely deep. For each mode, a unit periodic cell is identified and meshed with periodicity conditions imposed on the edges of cell, both for the film and the underlying substrate.

4.1 The Square Checkerboard Mode. Consider a square checkerboard mode such that the wavelength L in the x_1 and

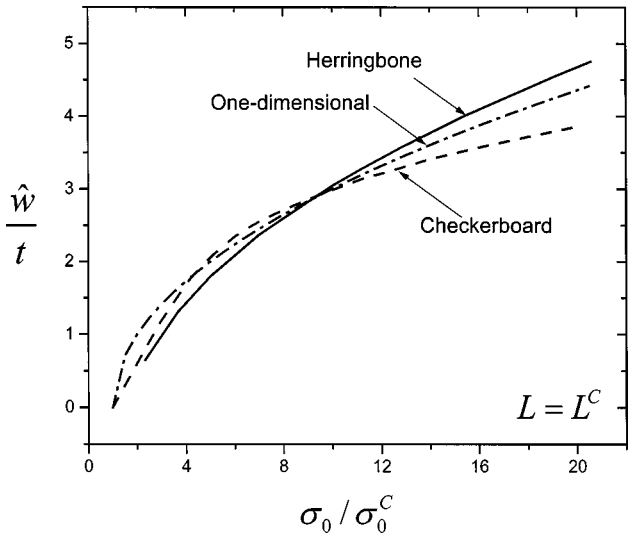


Fig. 5 Buckling amplitude of the film, $\hat{w}/t$, as a function of σ_0/σ_0^C for the three modes considered. The wavelengths (and inclination in the case of the herringbone mode) correspond to the critical at the onset of buckling.

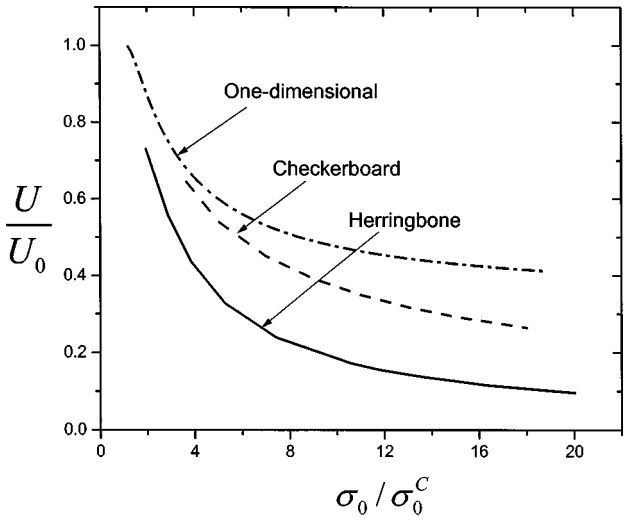


Fig. 6 Ratio of average elastic energy per unit area in the film/substrate system in the buckled state to that in the unbuckled state, U/U_0, as a function of σ_0/σ_0^C for the three modes considered. The wavelengths (and inclination in the case of the herringbone mode) correspond to the critical at the onset of buckling. At σ_0/σ_0^C well above unity, the herringbone mode lowers the energy more than the other two modes.

x_2-directions is set by the critical condition in (7), i.e., $2\pi/k_1 = 2\pi/k_2 = L = \sqrt{2}L^C$ where as defined before $L^C \equiv 2\pi/k^C$. The unit cell in this case is a rectangular parallelepiped of dimension $L \times L \times d$. A very small initial imperfection is prescribed such that the plate in the unstressed system at $\Delta T = 0$ has a slight middle surface deflection, $w = \hat{w}_I \cos(k_1 x_1)\cos(k_2 x_2)$, where $\hat{w}_I/t = 0.02$. Periodicity conditions are applied to the cell by enforcing all five nodal degrees-of-freedom to be the same on the two edges of the cell parallel to the x_1-coordinate, and similarly for the x_2-axis. In addition, at one of the corners of the cell the conditions $\partial w/\partial x_1 = 0$ and $\partial w/\partial x_2 = 0$ are enforced such that the mode crests and valleys align with the cell sides.

A square checkerboard mode is indeed determined with normal deflection that is roughly of the form $w = \hat{w}\cos(k_1 x_1)\cos(k_2 x_2)$. The computed relation between the mode amplitude, $\hat{w}/t$, defined as one half the difference between the maximum and minimum deflections, and σ_0/σ_0^C is plotted in Fig. 5. Included in this figure are the corresponding results for the one-dimensional mode and the herringbone mode, which is obtained in the next subsection. The results for the computed average strain energy per area in the film/substrate system are presented in Fig. 6 in normalized form as U/U_0 versus σ_0/σ_0^C. The results for the one-dimensional mode with $k_1 = k^C$ ($k_2 = 0$) are also plotted, as is the corresponding result for the critical herringbone mode obtained next. The energy per area of the critical checkerboard mode lies between that for the one-dimensional mode and the herringbone mode.

4.2 The Herringbone Mode. The unit top surface of the periodic cell for the herringbone mode is shown in Fig. 7. The parallelepiped is characterized by its width, a, breadth, L, and inclination angle, α. Periodicity conditions are applied to the top and bottom edges of the cell, and symmetry is imposed on the left and right edges. A small initial deflection ($\hat{w}_I/t = 0.02$) satisfying these edge conditions is introduced to initiate the mode.

Contours of the normal deflection of the film w within the cell are displayed in Fig. 8 for three values of L/L^C, at $\sigma_0/\sigma_0^C = 26$ (Figs. 8(a), 8(b)) or $\sigma_0/\sigma_0^C = 4$ (Fig. 8(c)) with $a/L = 2$ and $\alpha = 45°$. It will be seen below that the minimum energy configuration has $L/L^C \cong 1$, and for this value it can be seen that the deflection shape displays the features of the herringbone mode seen in Figs. 1 and 2. The mode has a curving ridge running along the center of the cell that aligns itself to merge smoothly at the jog with the ridge in the next cell. The excess breadth of the cell having $L/L^C = 2.4$ results in two ridges in the interior sector of the cell. The shape of the mode with narrow breadth, $L/L^C = 0.55$, is similar to that of the experimental herringbone pattern, but it will be seen that the energy for the narrow cell is well above that for $L/L^C = 1$.

The dependence of U/U_0 on the parameters of the cell geometry $(L/L^C, a/L, \alpha)$ is presented in Fig. 9. Figure 9(a), displays the clear trend whereby minimum energy is associated with $L/L^C \cong 1$. The energy of modes with $L/L^C = 1.6$ and $L/L^C = 0.7$ is distinctly above the minimum. In Fig. 9(b) it is seen that the energy in the buckled state is surprisingly insensitive to the normalized length of the cell, a/L. Only for very short cells, $a/L = 0.5$, is the energy noticeably above the minimum. Evidence for this insensitivity is reflected in the experimental herringbone patterns in Figs. 1 and 2, where it can be seen that the distance between jogs varies by at least a factor of two from one section of the film to another. Similarly, there is not a very strong dependence of the energy of the buckled system on the inclination of the cell, α, although the

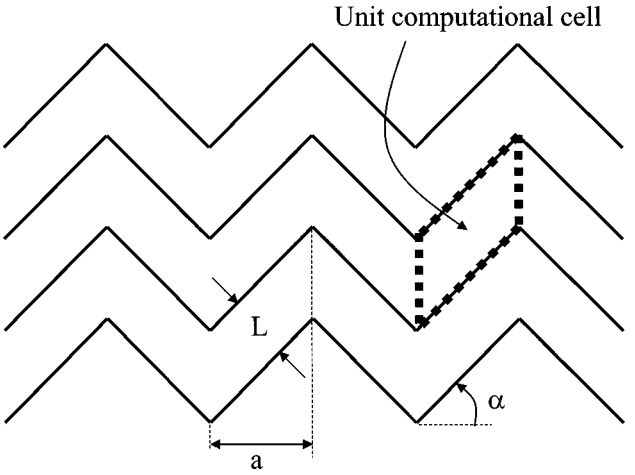
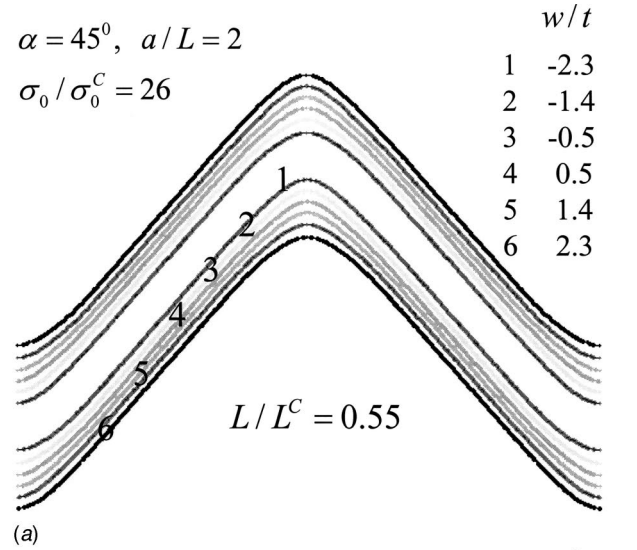
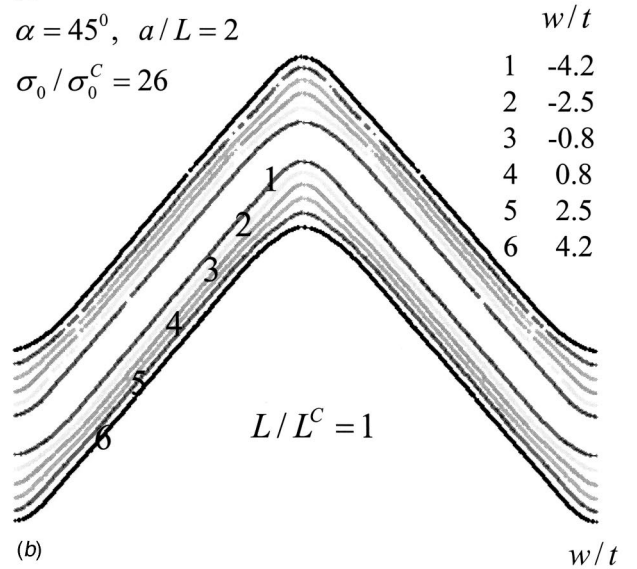


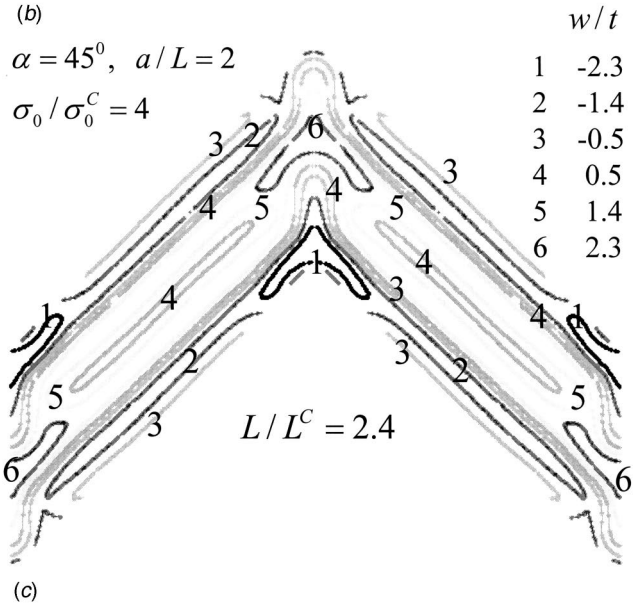
Fig. 7 Periodic cell of the herringbone mode



(*a*)

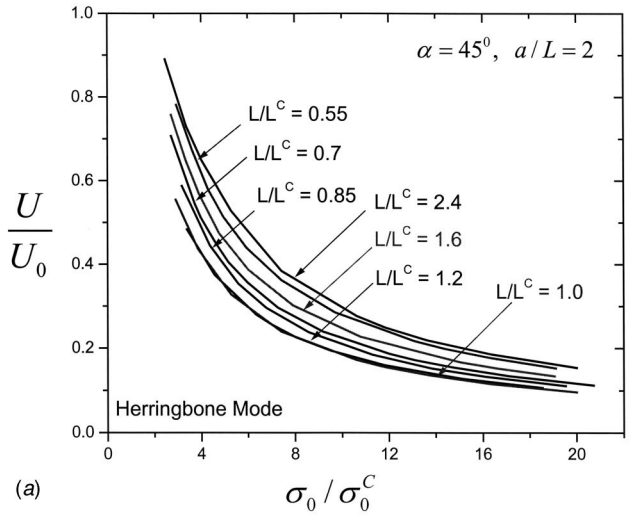


(*b*)

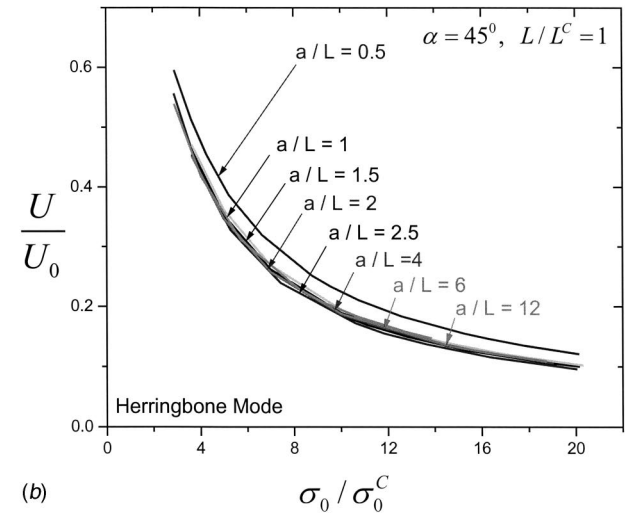


(*c*)

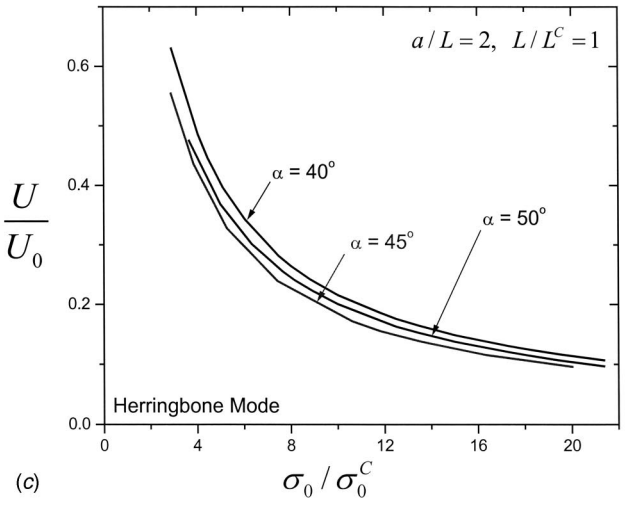
Fig. 8 Contour plots of the normal deflection of the film in the herringbone mode at $\sigma_0/\sigma_0^C=26$ (*a,b*) and $\sigma_0/\sigma_0^C=4$ (*c*) for several values of the breadth of the periodic cell, all with $a/L=2$ and $\alpha=45$ deg



(*a*)



(*b*)



(*c*)

Fig. 9 Variation of U/U_0 as a function of σ_0/σ_0^C for the herringbone mode. (*a*) Dependence on L/L^C with $a/L=2$ and $\alpha=45$ deg. (*b*) Dependence on a/L with $L/L^C=1$ and $\alpha=45$ deg. (*c*) Dependence on α with $L/L^C=1$ and $a/L=2$.

minimum is attained for $\alpha \cong 45$ deg (Fig. 9(*c*)). A feature seen in each of the plots in Fig. 9, as well as for the one-dimensional mode in Fig. 4, is the invariance of the ordering of the relative energy trends with respect to changes in σ_0/σ_0^C. Put another way, the parameters governing the geometry of the minimum energy mode do not change in a significant way as σ_0/σ_0^C increases. In both Figs. 5 and 9 the results for the herringbone mode are only plotted for values of σ_0/σ_0^C sufficiently large compared to unity. The herringbone mode is not a bifurcation mode, and it only becomes a preferred mode in the sense of having minimum energy at σ_0/σ_0^C somewhat above unity. At $\sigma_0/\sigma_0^C \approx 1$, the amplitude and normalized energy of the herringbone mode is dominated by the initial imperfection.

5 Conclusions

Among the three buckling modes considered, the herringbone mode produces the lowest average elastic energy of the film/substrate system for films stressed well above critical, as seen in Fig. 6. The herringbone mode is able to relax the biaxial pre-stress stress, σ_0, in the film in all directions while inducing relatively little concurrent stretch energy. The stretch energy associated with buckling that does occur is localized in the jog regions. By contrast, the one-dimensional mode requires essentially no stretch energy (it continues to exhibit zero Gaussian curvature), but it relaxes the biaxial pre-stress only in one direction. The checkerboard mode relaxes the pre-stress in all directions, but it develops non-zero Gaussian curvature and induces much more concurrent stretch energy than the herringbone mode. The minimum energy state of the herringbone mode has undulation width, L, which is very close to that of the one-dimensional mode, $L^C = 2\pi t(\bar{E}/3E_s)^{1/3}$, and jog angle $\alpha \cong 45$ deg. The minimum energy state is weakly dependent on the spacing between jogs, where stretch is localized. The experimental herringbone patterns in Figs. 1 and 2 show a spread in the jog spacing, and they also displays jog angles in reasonable agreement with the theoretical minimum energy state.

A question not addressed in this paper is how the herringbone mode emerges as σ_0 increases above critical. For a small range of σ_0 above critical, combinations of the classical modes of Section 2 necessarily have the lowest system energy; but as σ_0 increases, the herringbone mode emerges as the minimum energy mode. Does the herringbone mode spread across the film starting from some region of imperfection or from an edge? Or does it somehow emerge spontaneously over the entire film as a transition from a combination of classical modes? Bowden et al. [1,2] were not able to observe the evolution of the buckling patterns as their specimens were cooled from the film deposition temperature, and thus at this time it is not possible to give an experimental description of how the herringbone mode evolves.

A word of caution is in order about predicting mode patterns based on minimum energy states. The means by which deformations evolve to the minimum energy state is by no means obvious. Mechanics is replete with problems whose minimum energy states are not easily assessable. The pattern formed by forcing a film to buckle into a mold in Fig. 3 is just such an example. Once the finite amplitude mode has formed and the mold removed, it appears that the mode is locked in place and does not undergo changes towards a lower energy state unless further disturbed. To our knowledge, the nonlinear mechanics governing such behavior in buckled films has not been studied. In the case of the minimum energy herringbone mode, experimental observation confirms its existence, even though it has not been established how it evolves.

Acknowledgment

This work has been supported in part by Grant NSF DMR 0213805 and in part by the Division of Engineering and Applied Sciences, Harvard University.

References

[1] Bowden, N., Brittain, S., Evans, A. G., Hutchinson, J. W., and Whitesides, G. M., 1998, "Spontaneous Formation of Ordered Structures in Thin Films of Metals Supported on an Elastomeric Polymer," Nature (London), **393**, pp. 146–149.

[2] Huck, W. T. S., Bowden, N., Onck, P., Pardoen, T., Hutchinson, J. W., and Whitesides, G. M., 2000, "Ordering of Spontaneously Formed Buckles on Planar Surfaces," Langmuir, **16**, pp. 3497–3501.

[3] Yoo, P. J., Suh, K. Y., Park, S. Y., and Lee, H. H., 2002, "Physical Self-Assembly of Microstructures by Anisotropic Buckling," Adv. Mater. (Weinheim, Ger.), **14**, pp. 1383–1387.

[4] Timoshenko, S. P., and Gere, J. M., 1961, *Theory of Elastic Stability*, McGraw-Hill, New York.

[5] Allen, H. G., *Analysis and Design of Structural Sandwich Panels*, Pergamon, New York.

A Coupled Zig-Zag Third-Order Theory for Piezoelectric Hybrid Cross-Ply Plates

S. Kapuria
Associate Professor,
Department of Applied Mechanics,
Indian Institute of Technology, Delhi,
Hauz Khas, New Delhi 110016, India

A new zig-zag coupled theory is developed for hybrid cross-ply plates with some piezoelectric layers using third-order zig-zag approximation for the inplane displacements and sublayer wise piecewise linear approximation for the electric potential. The theory considers all electric field components and can model open and closed-circuit boundary conditions. The deflection field accounts for the transverse normal strain due to the piezoelectric d_{33} coefficient. The displacement field is expressed in terms of five displacement variables (which are the same as in FSDT) and electric potential variables by satisfying exactly the conditions of zero shear stresses at the top and bottom, and their continuity at layer interfaces. The governing equations are derived from the principle of virtual work. Comparison of the Navier solutions for the simply-supported plates with the analytical three-dimensional piezoelasticity solutions establishes that the present efficient zig-zag theory is quite accurate for moderately thick plates. [DOI: 10.1115/1.1767170]

1 Introduction

Smart composite laminates having piezoelectric sensors and actuators to achieve desired control, form part of a new generation of adaptive structures. The sensors and actuators can be in the form of distributed layers or patches, which are surface mounted or embedded. There have been many reviews (e.g., Chopra [1]) on the state of art of smart structures and integrated systems. Several reviews, [2–5], of three-dimensional continuum-based approaches and two-dimensional theories for laminated hybrid plates, have been presented. Analytical three-dimensional solutions are available only for some specific shapes and boundary conditions of plates, [6–8]. The three-dimensional finite element analysis, [9], of laminated plates results in large problem size which may become computationally costly for practical dynamics and control problems. Hence, efficient accurate electromechanical coupled two-dimensional plate models are required without too much loss of accuracy compared to the three-dimensional models. Some works have used various elastic laminated plate models, [10–13], with effective forces and moments due to induced strain of actuators. Classical laminate theory (CLT), [14–17], first-order shear deformation theory (FSDT), [18–19], and the refined third-order theory (TOT), [20], have been applied without electromechanical coupling to hybrid plates and shells. Coupled CLT, FSDT, [21–26], and TOT, [27,28], solutions for hybrid plates including the charge equation of electrostatics and electromechanical coupling, have been reported. These equivalent single layer theories are less computationally involved and are straight forward to implement. But these theories do not account for the zig-zag distribution of the inplane displacements and do not satisfy the interlaminar shear stress continuity conditions, yielding inaccurate results for moderately thick laminates and even thinner laminates with strong inhomogeneities across the thickness. A discrete layer theory (DLT) with layerwise approximation of displacements was developed by Robbins and Reddy [29] for elastic laminated beams with induced actuation strain in the piezoelectric layers. Heyliger et al. [30] have presented coupled DLT, using layerwise approximation for displacement and potential, which yields accurate results for thin and thick plates. But it is computationally expensive for practical dynamics and control problems since the number of displacement unknowns depends on the number of sublayers. Carrera [31] has presented a coupled DLT for plates with layerwise linear zig-zag approximation across the thickness for inplane displacements and quadratic one for transverse shear stresses and potential. But the axial electric field is neglected and the constitutive equation for shear stresses is only approximately satisfied. Bisegna et al. [32] have presented a layerwise coupled first-order shear deformation model for each layer of piezoelectric sandwich plate with linear variation of electric potential along the thickness of the actuator. Kim et al. [33] have developed a coupled DLT for hybrid shell laminates with first order global variation and layerwise hyperbolic variation across the thickness for the inplane displacements and across-the-thickness uniform variation for the transverse displacement, which satisfy the shear traction free conditions at the top and bottom surfaces and transverse shear stress continuity conditions at the layer interfaces for zero inplane electric fields. Except for the coupled DLT, [30], in which the transverse displacement is also taken as piecewise linear, no other two-dimensional theory discussed above considers the piezoelectric transverse normal strain induced due to piezoelectricity through d_{33} coefficient, which has been observed to have considerable effect on the response, especially for electrical load, [4]. The efficient zig-zag third order theory of laminated elastic plates, [34,35], has been extended by Kapuria et al. [36–38] for static and dynamic analysis of hybrid composite and sandwich beams. In this coupled efficient DLT, a third-order zig-zag approximation across the thickness for the axial displacement is used with a sublayerwise piecewise linear approximation for the potential ϕ. The conditions of zero transverse shear stress τ_{zx} at the top and bottom surfaces and the conditions of continuity of τ_{zx} at layer interfaces are enforced to formulate the theory in terms of three primary displacement variables, which are the same as in FSDT. The model considers both the axial and transverse electric fields and includes the transverse piezoelectric strain from d_{33} coefficient. Very accurate results have been reported by them for moderately thick beams using this theory.

This work presents an efficient coupled third-order zig-zag theory for cross-ply hybrid plates using combination of third-order variation across the thickness and a layerwise linear zig-zag approximation across the thickness for the inplane displacements.

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 Applied Mechanics Division, February 5, 2003; final revision, October 17, 2003. Associate Editor: R. C. Benson. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

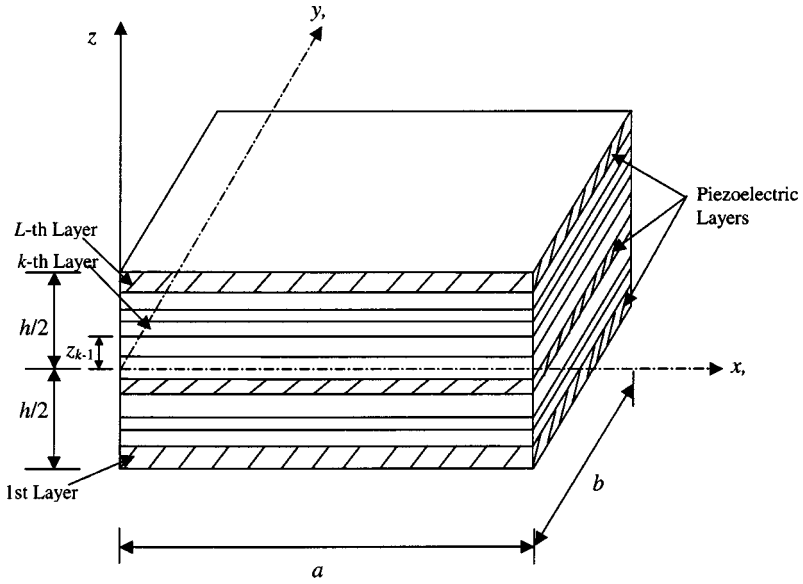


Fig. 1 Geometry of a hybrid plate

The electric potential ϕ is sublayerwise approximated as piecewise linear. The transverse displacement is approximated to take into account the piezoelectric strain in the thickness direction due to piezoelectric coefficient d_{33}. The shear traction free conditions at the top and bottom of the plate and shear continuity conditions at the layer interfaces are exactly satisfied to formulate the theory in terms of only five primary displacement variables and the potential variables. These primary displacement variables are the same as for the FSDT. The coupled stress and charge equilibrium equations, and boundary conditions are derived using the principle of virtual work. The theory includes the effect of in-plane electric field components which may be applied by actuation or induced by the piezoelectric effect. The theory can model open and closed-circuit boundary conditions. This theory is assessed by comparison of an analytical Navier solution for simply-supported rectangular plate, with the analytical three-dimensional piezoelastic solution and coupled FSDT solution. For this purpose highly inhomogeneous layups of a test case of six-layer hybrid plate, and a five-layer hybrid plate with composite elastic substrate are considered. The accuracy of the theory is checked for mechanical and electrical loads for different electrical conditions for thin and thick plates. The theory is computationally as efficient as the equivalent smeared plate coupled FSDT and yet yields quite accurate through-the-thickness variations of displacements, stresses and potential for moderately thick plates.

2 Approximation of Potential and Displacement Fields

Consider a hybrid cross-ply plate (Fig. 1) made of L orthotropic plys of total thickness h with the midplane chosen as the xy-plane $z=0$. The plate is loaded transversely on the bottom at $z=z_0=-h/2$ and on the top at $z=z_L=h/2$. Some of the layers can be orthorhombic piezoelectric materials of class mm^2 symmetry, with poling along z. The kth ply from the bottom has bottom surface at $z=z_{k-1}$. The reference plane $z=0$ either passes through or is the bottom surface of the k_0th layer. Let u_x, u_y; w be the in-plane and transverse displacements. Let ϕ be the electric potential and $E_x=-\phi_{,x}$, $E_y=-\phi_{,y}$, $E_z=-\phi_{,z}$, be the electric field. Denoting differentiation by a subscript comma, the strain-displacement relations are

$$\varepsilon_x=u_{x,x},\quad \varepsilon_y=u_{y,y},\quad \varepsilon_z=w_{,z},\quad \gamma_{xy}=u_{x,y}+u_{y,x},$$
$$\gamma_{yz}=u_{y,z}+w_{,y},\quad \gamma_{zx}=u_{x,z}+w_{,x}. \tag{1}$$

Unlike most other studies, E_x, E_y are not considered as zero, since these may be applied by actuation or may be induced by the piezoelectric coupling. The linear constitutive equations for the stresses σ, τ and electric displacements D_x, D_y, D_z are expressed, using the assumption of $\sigma_z \simeq 0$, as

$$\sigma=\bar{Q}\varepsilon-\bar{e}_3^{\mathrm{T}}E_z,\quad \tau=\hat{Q}\gamma-\hat{e}E,$$
$$D=\hat{e}^{\mathrm{T}}\gamma+\hat{\eta}E,\quad D_z=\bar{e}_3\varepsilon+\bar{\eta}_{33}E_z, \tag{2}$$

where, for cross-ply laminates,

$$\sigma=\begin{bmatrix}\sigma_x\\ \sigma_y\\ \tau_{xy}\end{bmatrix},\quad \tau=\begin{bmatrix}\tau_{zx}\\ \tau_{yz}\end{bmatrix},\quad D=\begin{bmatrix}D_x\\ D_y\end{bmatrix},\quad \varepsilon=\begin{bmatrix}\varepsilon_x\\ \varepsilon_y\\ \gamma_{xy}\end{bmatrix},$$
$$\gamma=\begin{bmatrix}\gamma_{zx}\\ \gamma_{yz}\end{bmatrix},\quad E=\begin{bmatrix}E_x\\ E_y\end{bmatrix},$$
$$\bar{Q}=\begin{bmatrix}\bar{Q}_{11} & \bar{Q}_{12} & 0\\ \bar{Q}_{12} & \bar{Q}_{22} & 0\\ 0 & 0 & \bar{Q}_{66}\end{bmatrix},\quad \hat{Q}=\begin{bmatrix}\bar{Q}_{55} & 0\\ 0 & \bar{Q}_{44}\end{bmatrix},$$
$$\hat{e}=\begin{bmatrix}\bar{e}_{15} & 0\\ 0 & \bar{e}_{24}\end{bmatrix},\quad \hat{\eta}=\begin{bmatrix}\bar{\eta}_{11} & 0\\ 0 & \bar{\eta}_{22}\end{bmatrix},\quad \bar{e}_3=[\bar{e}_{31}\ \bar{e}_{32}\ 0], \tag{3}$$

where $\bar{Q}_{ij}$, $\bar{e}_{ij}$, $\bar{\eta}_{ij}$ are the reduced elastic stiffnesses, piezoelectric stress constants and electric permittivities.

The potential ϕ is approximated as piecewise linear across the thickness, in terms of its values at n_ϕ points at z_ϕ^j, $j=1,2,\ldots,n_\phi$:

$$\phi(x,y,z)=\Psi_\phi^j(z)\phi^j(x,y) \tag{4}$$

where $\phi^j(x,y)=\phi(x,y,z_\phi^j)$. $\Psi_\phi^j(z)$ are linear interpolation functions and summation convention is used for indices j, j'. A piezoelectric layer is divided into sublayers for discretisation of ϕ whose number is determined by the required accuracy.

Three-dimensional solutions, [6], reveal that for moderately thick plates under electric potential load, w has significant variation across the thickness due to much greater electrical contribution to ε_z compared to that of σ_x, σ_y. Hence, herein w is ap-

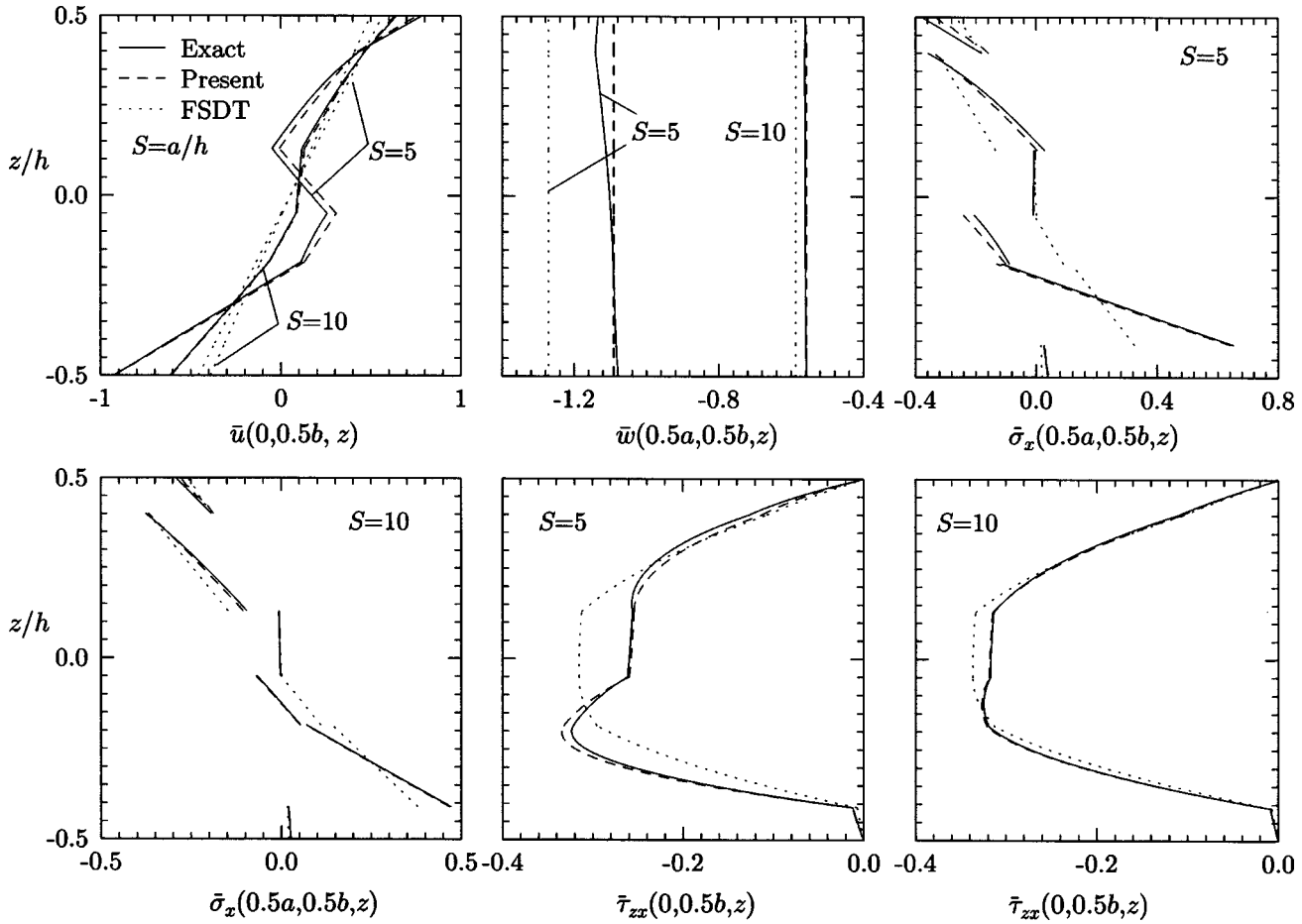


Fig. 2 The distributions of $\bar{u}$, $\bar{w}$, $\bar{\sigma}_x$, $\bar{\tau}_{zx}$ for test plate (a) under pressure load

proximated by integrating the constitutive equation for ε_z by neglecting elastic compliances S_{13}, S_{23}, i.e., $\varepsilon_z = w_{,z} = S_{13}\sigma_1 + S_{23}\sigma_2 + d_{33}E_z \simeq -d_{33}\phi_{,z} \Rightarrow$

$$w(x,y,z) = w_0(x,y) - \overline{\Psi}^j_\phi(z)\,\phi^j(x,y) \tag{5}$$

where $\overline{\Psi}^j_\phi(z) = \int_0^z d_{33}\Psi^j_{\phi,z}(z)dz$ is a piecewise linear function. For the kth layer, u_x, u_y are approximated as a combination of third-order variation in z across the thickness and layerwise piecewise linear variation:

$$u(x,y,z) = u_k(x,y) - z w_{0_d}(x,y) + z\psi_k(x,y) + z^2\xi(x,y) + z^3\eta(x,y), \tag{6}$$

where

$$u = \begin{bmatrix} u_x \\ u_y \end{bmatrix}, \quad w_{0_d} = \begin{bmatrix} w_{0,x} \\ w_{0,y} \end{bmatrix}, \quad u_k = \begin{bmatrix} u_{k_x} \\ u_{k_y} \end{bmatrix}, \quad \psi_k = \begin{bmatrix} \psi_{k_x} \\ \psi_{k_y} \end{bmatrix},$$
$$\xi = \begin{bmatrix} \xi_x \\ \xi_y \end{bmatrix}, \quad \eta = \begin{bmatrix} \eta_x \\ \eta_y \end{bmatrix}, \tag{7}$$

u_k is the translation and ψ_k is related to the shear rotation of the kth layer.

Substituting u_x, u_y, w from Eqs. (6) and (5), and ϕ from Eq. (4) into Eqs. (1) and using Eq. (2) yields τ as

$$\tau = \hat{Q}^k[\psi_k + 2z\xi + 3z^2\eta] + [\hat{e}^k\Psi^j_\phi(z) - \hat{Q}^k\overline{\Psi}^j_\phi(z)]\phi^j_d, \tag{8}$$

where $\phi^j_d = [\phi^j_{,x}\ \phi^j_{,y}]^{\mathrm{T}}$. For the k_0th layer, denote $u_0(x,y) = u_{k_0}(x,y) = u(x,y,0)$, $\psi_0(x,y) = \psi_{k_0}(x,y)$. The functions u_k, ψ_k, ξ, η are expressed in terms of u_0 and ψ_0 using the $(L-1)$ conditions each for the continuity of τ and u at the layer interfaces and the two-shear traction-free conditions $\tau = 0$ at $z = z_0$, z_L. The continuity condition of τ at interface $z = z_{i-1}$ between layers i and $i-1$ is expressed in the following recursive form so that the solution of ψ_i, ξ, η is easily tractable:

$$\begin{aligned}
&\hat{Q}^i[\psi_i + 2z_i\xi + 3z_i^2\eta] + [\hat{e}^i\Psi^j_\phi(z_i) - \hat{Q}^i\overline{\Psi}^j_\phi(z_i)]\phi^j_d \\
&\quad = \hat{Q}^{i-1}[\psi_{i-1} + 2z_{i-1}\xi + 3z_{i-1}^2\eta] + [\hat{e}^{i-1}\Psi^j_\phi(z_{i-1}) \\
&\qquad - \hat{Q}^{i-1}\overline{\Psi}^j_\phi(z_{i-1})]\phi^j_d + 2\hat{Q}^i(z_i - z_{i-1})\xi + 3\hat{Q}^i(z_i^2 - z_{i-1}^2)\eta \\
&\qquad + [\hat{e}^i\{\Psi^j_\phi(z_i) - \Psi^j_\phi(z_{i-1})\} - \hat{Q}^i\{\overline{\Psi}^j_\phi(z_i) - \overline{\Psi}^j_\phi(z_{i-1})\}]\phi^j_d .
\end{aligned} \tag{9}$$

Using Eq. (8), the shear traction free condition $\tau(x,y,z_0) = 0$, can also be written in the above pattern as

$$\begin{aligned}
&\hat{Q}^1[\psi_1 + 2z_1\xi + 3z_1^2\eta] + [\hat{e}^1\Psi^j_\phi(z_1) - \hat{Q}^1\overline{\Psi}^j_\phi(z_1)]\phi^j_d \\
&\quad = 2\hat{Q}^1(z_1 - z_0)\xi + 3\hat{Q}^1(z_1^2 - z_0^2)\eta + [\hat{e}^1\{\Psi^j_\phi(z_1) - \Psi^j_\phi(z_0)\} \\
&\qquad - \hat{Q}^1\{\overline{\Psi}^j_\phi(z_1) - \overline{\Psi}^j_\phi(z_0)\}]\phi^j_d .
\end{aligned} \tag{10}$$

Adding Eq. (10) and Eqs. (9) for $i = 2,3,\ldots,k$ yields

$$\begin{aligned}
&\hat{Q}^k(\psi_k + 2z_k\xi + 3z_k^2\eta) + [\hat{e}^k\Psi^j_\phi(z_k) - \hat{Q}^k\overline{\Psi}^j_\phi(z_k)]\phi^j_d \\
&\quad = 2C_1^k\xi + 6C_2^k\eta + C_{3j}^k\phi^j_d, \quad k = 2,\ldots,L,
\end{aligned} \tag{11}$$

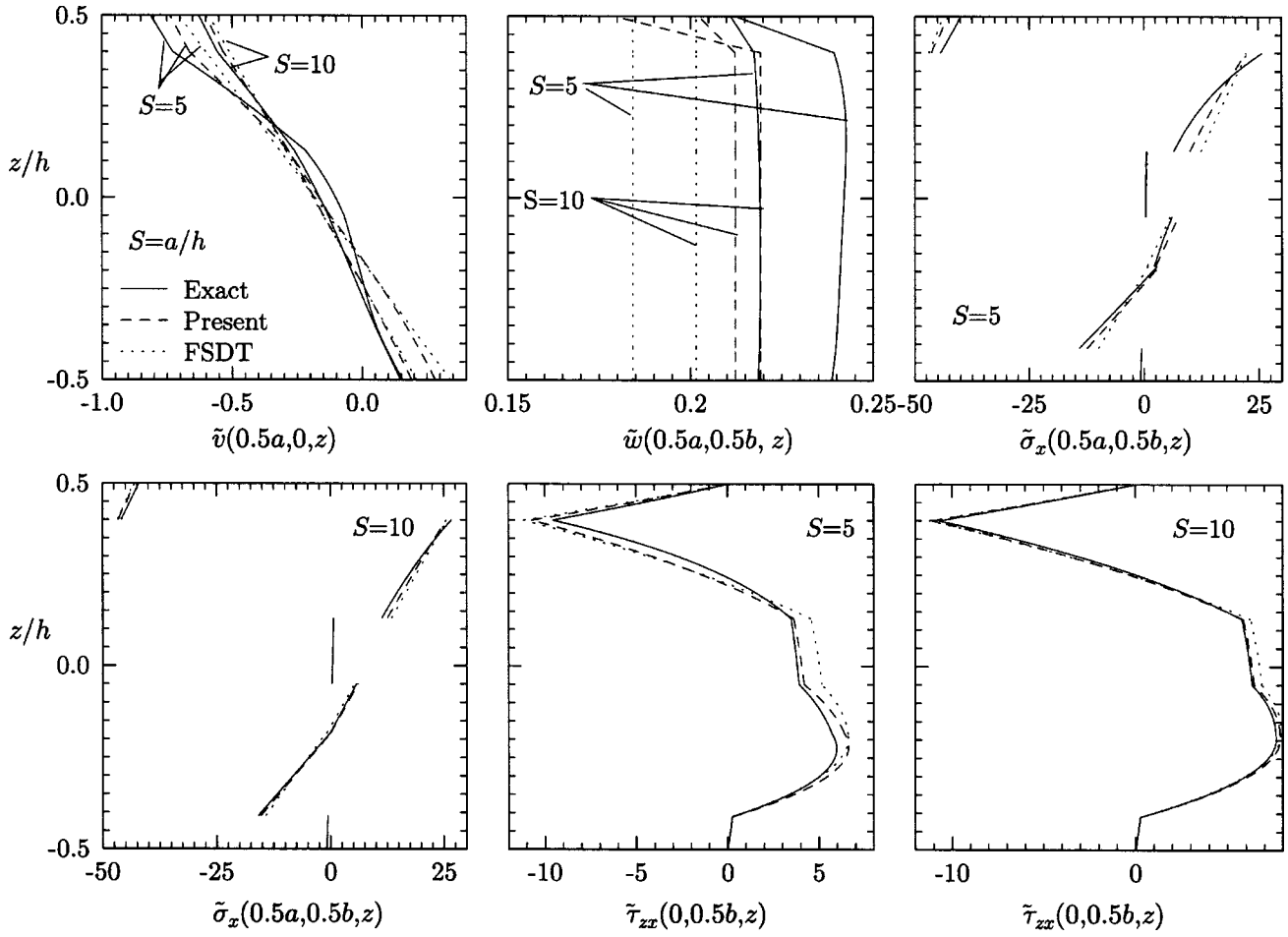


Fig. 3 The distributions of $\tilde{v}$, $\tilde{w}$, $\tilde{\sigma}_x$, $\tilde{\tau}_{zx}$ for test plate (a) under potential load

where $C_1^k=\Sigma_{i=1}^k\hat{Q}^i(z_i-z_{i-1})$, $C_{3j}^k=\Sigma_{i=1}^k[\hat{e}^i\{\Psi_\phi^j(z_i)-\Psi_\phi^j(z_{i-1})\}-\hat{Q}^i\{\Psi_\phi^j(z_i)-\Psi_\phi^j(z_{i-1})\}]$, $C_2^k=\Sigma_{i=1}^k\hat{Q}^i(z_i^2-z_{i-1}^2)/2$. Using Eq. (8), the condition $\tau_{zx}(x,z_L)=0$, can be written as

$$\hat{Q}^L[\psi_L+2z_L\xi+3z_L^2\eta]+[\hat{e}^L\Psi_\phi^j(z_L)-\hat{Q}^L\overline{\Psi}_\phi^j(z_L)]\phi_d^j=0. \tag{12}$$

Eliminating ψ_L from Eq. (12) and Eq. (11) for $k=L$, and rewriting Eq. (10) yields

$$2C_1^L\xi+6C_2^L\eta=-C_{3j}^L\phi_d^j,\quad 2z_0\xi+3z_0^2\eta=C_5^j\phi_d^j-\psi_1, \tag{13}$$

where $C_5^j=\overline{\Psi}_\phi^j(z_0)I_2-(\hat{Q}^1)^{-1}\hat{e}^1\Psi_\phi^j(z_0)$ and I_2 is a 2×2 identity matrix. The solution of Eq. (13) for ξ, η is

$$\xi=R_3\psi_1+R_5^j\phi_d^j,\quad \eta=R_4\psi_1+R_6^j\phi_d^j, \tag{14}$$

where $\Delta=4z_0^2C_1^L-8z_0C_2^L$, $R_3=4\Delta^{-1}C_2^L$, $R_4=-4\Delta^{-1}C_1^L/3$, $R_5^j=-\Delta^{-1}(2z_0^2C_{3j}^L+4C_2^LC_5^j)$, $R_6^j=\Delta^{-1}(4z_0C_{3j}^L+4C_1^LC_5^j)/3$. Substituting ξ, η from Eq. (14) into Eq. (11) yields

$$\psi_k=R_2^k\psi_1+R_{j1}^k\phi_d^j \tag{15}$$

where $R_2^k=a_1^kR_3+a_2^kR_4$, $R_{j1}^k=a_1^kR_5^j+a_2^kR_6^j+(\hat{Q}^k)^{-1}[C_{3j}^k-\hat{e}^k\Psi_\phi^j(z_k)]+\overline{\Psi}_\phi^j(z_k)I_2$, $a_1^k=2[(\hat{Q}^k)^{-1}C_1^k-z_kI_2]$, $a_2^k=3[2(\hat{Q}^k)^{-1}C_2^k-z_k^2I_2]$. Using Eq. (6), continuity of u between layers i and $i-1\Rightarrow u_i+z_{i-1}\psi_i=u_{i-1}+z_{i-1}\psi_{i-1}$ and using Eq. (15):

$$u_i=u_{i-1}+z_{i-1}[(R_2^{i-1}-R_2^i)\psi_1+(R_{j1}^{i-1}-R_{j1}^i)\phi_d^j],\quad i=2,\ldots,L. \tag{16}$$

Adding Eqs. (16) for $i=2$ to k yields u_k in terms of u_1:

$$u_k=u_1+\overline{R}_2^k\psi_1+\overline{R}_{j1}^k\phi_d^j \tag{17}$$

where $\overline{R}_2^k=\Sigma_{i=2}^kz_{i-1}(R_2^{i-1}-R_2^i)$, $\overline{R}_{j1}^k=\Sigma_{i=2}^kz_{i-1}(R_{j1}^{i-1}-R_{j1}^i)$. Equations (17) and (15) yield for the k_0th layer:

$$u_0(x,y)=u_{k_0}(x,y)=u_1+\overline{R}_2^{k_0}\psi_1+\overline{R}_{j1}^{k_0}\phi_d^j,$$
$$\psi_0(x,y)=\psi_{k_0}(x,y)=R_2^{k_0}\psi_1+R_{j1}^{k_0}\phi_d^j. \tag{18}$$

Substituting ξ, η from Eq. (14), u_k from Eq. (17) with u_1 from Eq. (18) and ψ_k from Eq. (15) in Eq. (6) yields

$$u(x,y,z)=u_0(x,y)-zw_{0_d}(x,y)+R_k(z)\psi_1(x,y)+R_{k\phi}^j(z)\phi_d^j(x,y), \tag{19}$$

where $R_k(z)=R_1^k+zR_2^k+z^2R_3+z^3R_4$, $R_{k\phi}^j(z)=R_1^{kj}+zR_{j1}^k+z^2R_5^j+z^3R_6^j$, with $R_1^k=\overline{R}_2^k-\overline{R}_2^{k_0}$, $R_1^{kj}=\overline{R}_{j1}^k-\overline{R}_{j1}^{k_0}$. Substituting ψ_1 in terms of ψ_0 from Eq. (18) into Eq. (19) yields the expression of u as

$$u(x,y,z)=u_0(x,y)-zw_{0_d}(x,y)+R^k(z)\psi_0(x,y)+R^{kj}(z)\phi_d^j(x,y), \tag{20}$$

where

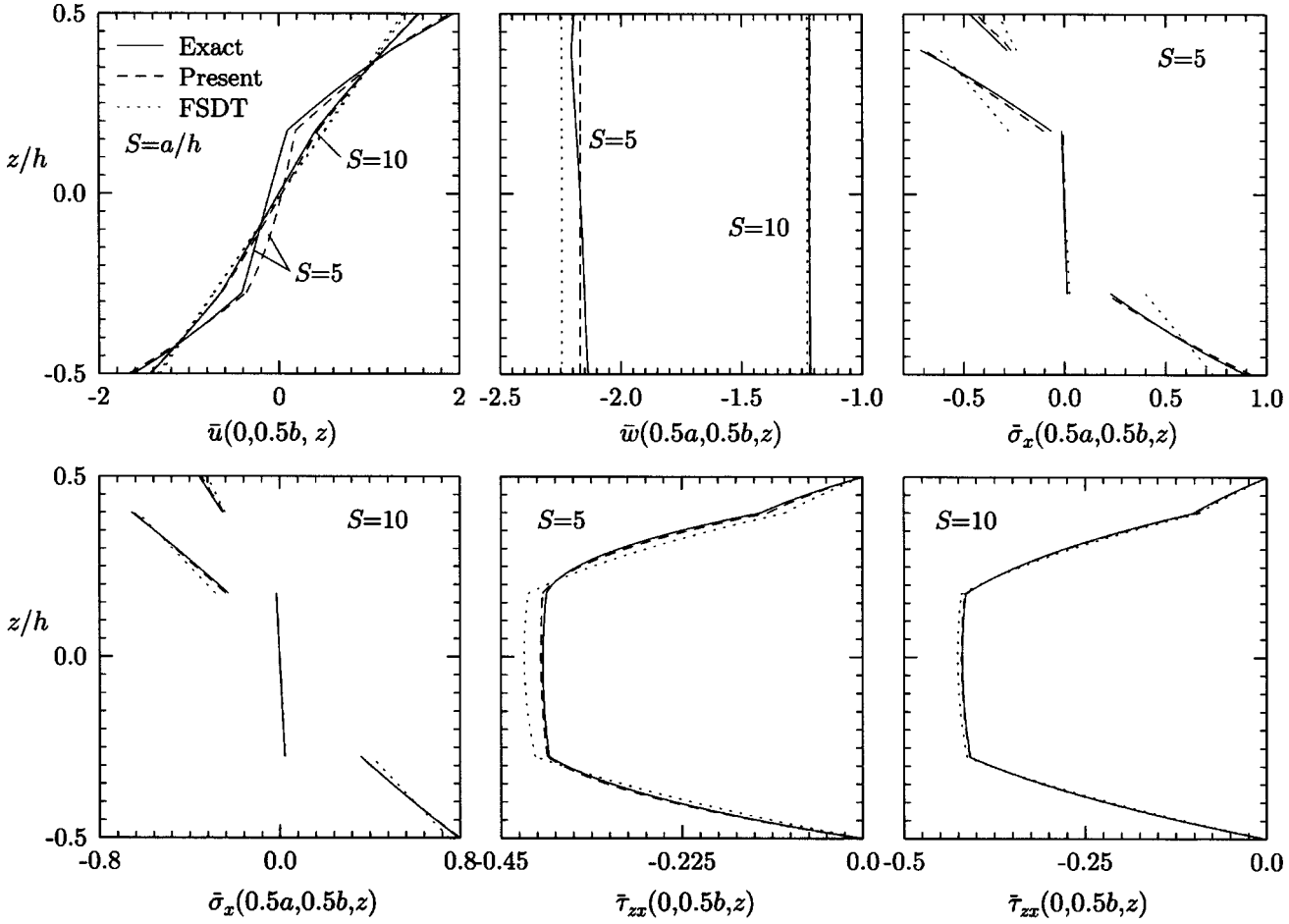


Fig. 4 The distributions of $\bar{u}$, $\bar{w}$, $\bar{\sigma}_x$, $\bar{\tau}_{zx}$ for hybrid composite plate (b) under pressure load

$$R^k(z)=R_k(z)(R_2^{k_0})^{-1}=\hat{R}_1^k+z\hat{R}_2^k+z^2\hat{R}_3+z^3\hat{R}_4,$$
$$R^{kj}(z)=R_{k\phi}^j(z)-R^k(z)R_{j1}^{k_0}=\hat{R}_1^{kj}+z\hat{R}_{j1}^k+z^2\hat{R}_5^j+z^3\hat{R}_6^j$$
$$(\hat{R}_1^k,\hat{R}_2^k,\hat{R}_3,\hat{R}_4)=(R_1^k,R_2^k,R_3,R_4)(R_2^{k_0})^{-1},$$
$$\hat{R}_1^{kj}=R_1^{kj}-\hat{R}_1^kR_{j1}^{k_0},\quad \hat{R}_{j1}^k=R_{j1}^k-\hat{R}_2^kR_{j1}^{k_0},$$
$$\hat{R}_5^j=R_5^j-\hat{R}_3R_{j1}^{k_0},\quad \hat{R}_6^j=R_6^j-\hat{R}_4R_{j1}^{k_0}. \tag{21}$$

R^k, R^{kj} are diagonal matrices. Thus ϕ, w, u are related to the primary variables u_0, w_0, ψ_0, ϕ^j by Eqs. (4), (5), (20).

3 Field Equations and Boundary Conditions

Let A be the surface area of the plate and at the interface $z=z_{\phi}^{j_i}$, ϕ^{j_i} be prescribed with q_{j_i} being the extraneous surface charge density on it. The total number of such prescribed potentials is $\bar{n}_{\phi}$. Let p_z^1, p_z^2 be the forces per unit area applied on the bottom and top surfaces of the plate in direction z. The principle of virtual work, [33], can be expressed, using the notation $\langle\ldots\rangle=\Sigma_{k=1}^{L}\int_{z_{k-1}^+}^{z_k^-}(\ldots)dz$, as

$$\int_A[\langle\sigma_x\delta\varepsilon_x+\sigma_y\delta\varepsilon_y+\tau_{xy}\delta\gamma_{xy}+\tau_{yz}\delta\gamma_{yz}+\tau_{zx}\delta\gamma_{zx}+D_x\delta\phi_{,x}$$
$$+D_y\delta\phi_{,y}+D_z\delta\phi_{,z}\rangle-p_z^1\delta w(x,y,z_0)-p_z^2\delta w(x,y,z_L)$$
$$+D_z(x,y,z_0)\delta\phi^1-D_z(x,y,z_L)\delta\phi^{n_\phi}-q_{j_i}\delta\phi^{j_i}]dA$$
$$-\int_{\Gamma_L}\langle\sigma_n\delta u_n+\tau_{ns}\delta u_s+\tau_{nz}\delta w+D_n\delta\phi\rangle ds$$
$$=0 \tag{22}$$

$\forall\,\delta u_0$, δw_0, $\delta\psi_0$, $\delta\phi^j$. Γ_L is the boundary curve of the midplane of the plate with normal n and tangent s. This variational equation is expressed in terms of δu_0, δw_0, $\delta\psi_0$, $\delta\phi^j$ and stress and electric displacement resultants to yield field equations and boundary conditions. The resultants $N=[N_x\,N_y\,N_{xy}]^{\mathrm{T}}$, $M=[M_x\,M_y\,M_{xy}]^{\mathrm{T}}$, $P=[P_x\,P_{yx}\,P_{xy}\,P_y]^{\mathrm{T}}$, $S=[S_x^j\,S_{yx}^j\,S_{xy}^j\,S_y^j]^{\mathrm{T}}$, $Q=[Q_x\,Q_y]^{\mathrm{T}}$, $\bar{Q}^j=[\bar{Q}_x^j\,\bar{Q}_y^j]^{\mathrm{T}}$, $V=[V_x\,V_y]^{\mathrm{T}}$, $V_\phi^j=[V_{\phi_x}^j\,V_{\phi_y}^j]^{\mathrm{T}}$, $H^j=[H_x^j\,H_y^j]^{\mathrm{T}}$, G^j are defined by

$$F_1=[N^{\mathrm{T}}\;\;M^{\mathrm{T}}\;\;P^{\mathrm{T}}\;\;S^{j^{\mathrm{T}}}]^{\mathrm{T}}=[\langle f_3^{\mathrm{T}}\sigma\rangle],$$
$$F_2=[Q_x\;\;Q_y\;\;\bar{Q}_x^j\;\;\bar{Q}_y^j]^{\mathrm{T}}=[\langle f_4^{\mathrm{T}}\tau\rangle] \tag{23a}$$

$$V=\langle\tau\rangle,\quad V_\phi^j=\langle\bar{\Psi}_\phi^j\tau\rangle,\quad H^j=\langle\Psi_\phi^j(z)D\rangle,\quad G^j=\langle\Psi_{\phi,z}^j(z)D_z\rangle. \tag{23b}$$

where $f_3=[I_3\,zI_3\,\Phi^k\,\Phi^{kj}]$, $f_4=[R_{,z}^k\,R_{,z}^{kj}-\bar{\Psi}_\phi^j(z)I_2]$, I_3 is a 3×3 identity matrix and

$$\Phi^k=\begin{bmatrix}R_{11}^k & 0 & 0 & 0\\ 0 & 0 & 0 & R_{22}^k\\ 0 & R_{11}^k & R_{22}^k & 0\end{bmatrix},$$

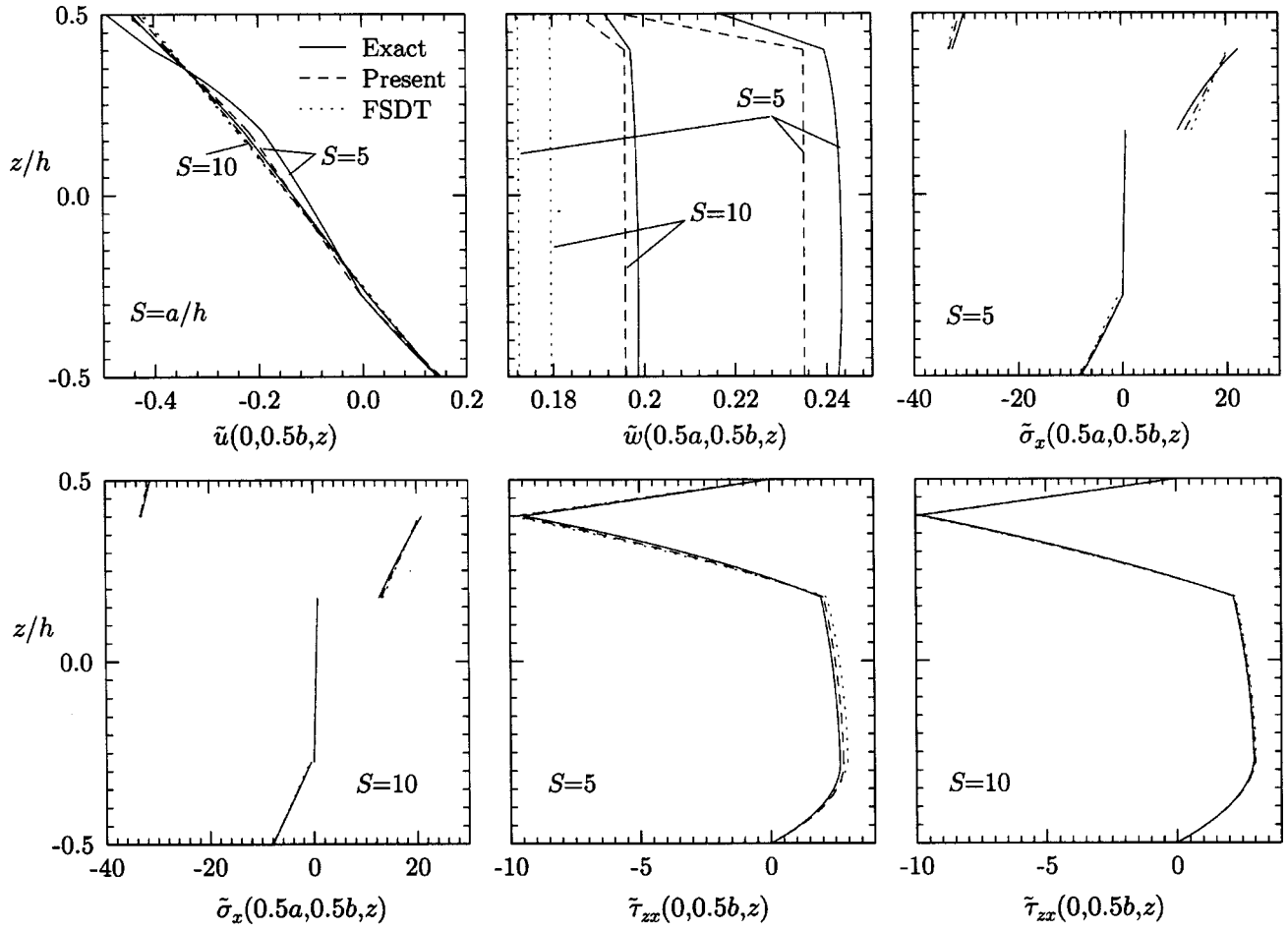


Fig. 5 The distributions of $\tilde{u}$, $\tilde{w}$, $\tilde{\sigma}_x$, $\tilde{\tau}_{zx}$ for hybrid composite plate (b) under potential load

$$\Phi^{kj}=\begin{bmatrix} R_{11}^{kj} & 0 & 0 & 0 \\ 0 & 0 & 0 & R_{22}^{kj} \\ 0 & R_{11}^{kj} & R_{22}^{kj} & 0 \end{bmatrix}. \tag{24}$$

It can be shown that elements of R^k, R^{kj}, N, M, P, S^j transform as second-order tensors and elements of V, V_ϕ^j, Q, $\overline{Q}^j$, H^j transform as vectors for the coplanar axes x, y and n, s.

Using expressions of ϕ, w, u from Eqs. (4), (5), (20) and using Eq. (23), the area integral in Eq. (22) becomes

$$\int_A [\,\delta\overline{\varepsilon}_1^{\mathrm{T}}F_1+\delta\overline{\varepsilon}_2^{\mathrm{T}}F_2+\delta\phi_d^{j^{\mathrm{T}}}H^j+\delta\phi^jG^j-F_3\delta w_0-F_6^j\delta\phi^j]dA \tag{25}$$

where

$$\overline{\varepsilon}_1=[u_{0_x,x}\;\; u_{0_y,y}\;\; u_{0_x,y}+u_{0_y,x}\;\; -w_{0,xx}\;\; -w_{0,yy}$$
$$-2w_{0,xy}\;\; \psi_{0_x,x}\;\; \psi_{0_x,y}\;\; \psi_{0_y,x}\;\; \psi_{0_y,y}\;\; \phi^j_{,xx}\;\; \phi^j_{,xy}\;\; \phi^j_{,yx}\;\; \phi^j_{,yy}]^{\mathrm{T}}$$
$$\overline{\varepsilon}_2=[\psi_{0_x}\;\; \psi_{0_y}\;\; \phi^j_{,x}\;\; \phi^j_{,y}]^{\mathrm{T}},$$
$$F_3=p_z^1+p_z^2,$$
$$F_6^j=-p_z^1\overline{\Psi}^j_\phi(z_0)-p_z^2\overline{\Psi}^j_\phi(z_L)+D_z(x,y,z_L)\,\delta_{jn_\phi}$$
$$-D_z(x,y,z_0)\,\delta_{j1}+q_{j_i}\delta_{jj_i}. \tag{26}$$

δ_{ij} is Kronecker's delta. Using Eq. (20), the relation for components n, s can be expressed as

$$\begin{bmatrix}\delta u_n\\ \delta u_s\end{bmatrix}=\begin{bmatrix}\delta u_{0_n}\\ \delta u_{0_s}\end{bmatrix}-z\begin{bmatrix}\delta w_{0,n}\\ \delta w_{0,s}\end{bmatrix}+\begin{bmatrix}R_{nn}^k & R_{ns}^k\\ R_{sn}^k & R_{ss}^k\end{bmatrix}\begin{bmatrix}\delta\psi_{0_n}\\ \delta\psi_{0_s}\end{bmatrix}+\begin{bmatrix}R_{nn}^{kj} & R_{ns}^{kj}\\ R_{sn}^{kj} & R_{ss}^{kj}\end{bmatrix}$$
$$\times\begin{bmatrix}\delta\phi^j_{,n}\\ \delta\phi^j_{,s}\end{bmatrix}. \tag{27}$$

Using expressions of ϕ, w, u from Eqs. (4), (5), (20) and resultant components for axes n, s defined analogous to Eq. (23), the line integral in Eq. (22) can be expressed using Eq. (27) as

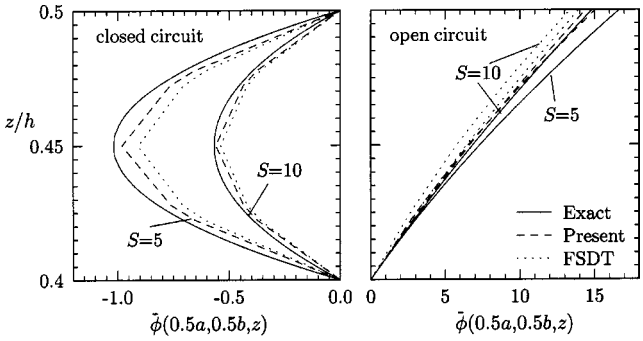


Fig. 6 Closed-circuit and open-circuit potential across the piezoelectric layer of plate (b) under pressure load

Table 1 Three-dimensional results and percentage error for present theory (Pres.) and FSDT for $n=1$

		plate (a)						plate (b)					
		load case 1			load case 2			load case 1			load case 2		
			% error in			% error in			% error in			% error in	
entity	S	3D	Pres.	FSDT	3D	Pres.	FSDT	3D	Pres.	FSDT	3D	Pres.	FSDT
w_c	5	-1.1026	-1.02	15.36	0.2420	-9.40	-23.88	-2.1693	-0.01	3.48	0.2430	-3.28	-29.06
	10	-0.5613	-0.55	4.67	0.2190	-3.08	-7.96	-1.2200	-0.24	0.7	0.1985	-1.33	-9.53
	100	-0.3558	-0.01	0.05	0.2030	-0.03	-0.09	-0.8896	0.00	0.01	0.1816	-0.01	-0.10
σ^e	5	0.6522	-1.46	-50.26	13.5130	16.50	29.32	0.9030	2.82	-18.44	22.3540	-6.96	-9.08
	10	0.4718	-0.86	-19.61	18.1530	4.23	8.00	0.7951	0.56	-5.66	21.0430	-2.01	-2.58
	100	0.4157	-0.01	-0.23	20.9670	0.04	0.09	0.7557	0.01	-0.05	20.5730	-0.02	-0.03
σ^p	5	-0.3786	-12.27	-31.03	-44.3840	4.44	5.72	-0.4696	-7.60	-30.99	-36.3650	0.56	-8.48
	10	-0.2916	-4.58	-11.61	-45.9040	1.28	1.65	-0.3571	-2.95	-10.92	-36.9900	0.09	-10.12
	100	-0.2602	-0.75	-0.84	-46.5590	-0.21	-0.21	-0.3181	-0.57	-0.66	-37.2290	-0.09	-10.58
τ_{zx}	5	-0.2597	-0.77	21.31	-9.5732	12.04	16.64	-0.3980	0.64	5.86	-9.4582	2.69	4.64
	10	-0.3167	0.36	6.29	-10.7010	3.56	4.97	-0.4195	0.22	1.51	-9.8089	0.71	1.21
	100	-0.3504	0.01	0.07	-11.2700	0.04	0.05	-0.4275	0.00	0.01	-9.9374	0.01	0.01
τ_{xy}	5	1.5811	-6.07	-16.78	-12.5470	-10.89	-15.48	0.9454	1.69	-5.22	-3.7263	-3.55	-12.08
	10	1.3169	-1.80	-6.19	-10.6190	-4.26	-6.18	0.6746	0.48	-2.50	-3.1249	-1.23	-4.07
	100	1.1842	-0.02	-0.08	-9.5554	-0.05	-0.08	0.5686	0.01	-0.03	-2.8946	-0.01	-0.04
ϕ (1) or D_z (2)	5	-0.9442	-5.50	-4.92	-2.0547	-1.09	-1.60	-1.0169	-3.58	-10.70	-1.3006	-0.34	-0.70
	10	-0.4692	-1.42	-3.78	-2.0162	-0.26	-0.45	-0.5669	-1.23	-5.03	-1.2911	-0.04	-0.13
	100	-0.2988	-0.02	-0.07	-1.9958	0.15	0.15	-0.4151	-0.01	-0.07	-1.2876	0.08	-0.08

$$\int_{\Gamma_L}[N_n\delta u_{0_n}+N_{ns}\delta u_{0_s}-M_n\delta w_{0,n}+(V_n+M_{ns,s})\delta w_0+P_n\delta\psi_{0_n}$$

$$+P_{ns}\delta\psi_{0_s}+S_n^j\delta\phi_{,n}^j+(H_n^j-V_{\phi_n}^j-S_{ns,s}^j)\delta\phi^j]ds$$

$$+\sum_i\,[\Delta M_{ns}(s_i)\delta w_0(s_i)-\Delta S_{ns}^j(s_i)\delta\phi^j(s_i)]=0, \qquad (28)$$

where the lateral surface has corners at $s=s_i$.

The area integral in Eq. (25) is expressed in terms of δu_{0_x}, δu_{0_y}, δw_0, $\delta\psi_{0_x}$, $\delta\psi_{0_y}$, $\delta\phi^j$, by using Green's theorem if required, and the terms involving δu_{0_x}, δu_{0_y}, $\delta\psi_{0_x}$, $\delta\psi_{0_y}$, $\delta w_{0,x}$, $\delta w_{0,y}$, $\delta\phi_{,x}^j$, $\delta\phi_{,y}^j$, in the integrand of Γ_L are expressed in terms of components n, s. The details are omitted. Thus Eq. (22) yields coupled field equations consisting of five equations of equilibrium and n_ϕ equations for electric potentials:

$$-N_{x,x}-N_{xy,y}=0, \quad -N_{xy,x}-N_{y,y}=0,$$
$$-M_{x,xx}-2M_{xy,xy}-M_{y,yy}-F_3=0,$$
$$-P_{x,x}-P_{yx,y}+Q_x=0, \quad -P_{xy,x}-P_{y,y}+Q_y=0,$$
$$-\overline{Q}_{x,x}^j-\overline{Q}_{y,y}^j+S_{x,xx}^j+S_{xy,xy}^j+S_{yx,xy}^j+S_{y,yy}^j-H_{x,x}^j-H_{y,y}^j+G^j$$
$$-F_6^j=0, \qquad (29)$$

with $j=1,2,\ldots,n_\phi$. The boundary conditions on Γ_L are the prescribed values of one of the factors of each of the following products:

$$u_{0_n}N_n, \quad u_{0_s}N_{ns}, \quad w_0(V_n+M_{ns,s}),$$

$$w_{0,n}M_n, \quad \psi_{0_n}P_n, \quad \psi_{0_s}P_{ns}, \quad \phi_{,n}^jS_n^j,$$

$$\phi^j[\overline{H}^j-\overline{V}_{\phi_n}^j-\overline{S}_{ns,s}^j-\{\overline{Q}_x^jn_x+\overline{Q}_y^jn_y+H_x^jn_x+H_y^jn_y$$
$$-(S_{x,x}^j+S_{yx,y}^j)n_x-(S_{y,y}^j+S_{xy,x}^j)n_y-S_{ns,s}^j\}] \qquad (30)$$

and at corners s_i:

$$w_0(s_i)\Delta M_{ns}(s_i), \quad \phi^j(s_i)\Delta S_{ns}^j(s_i).$$

The relations between the resultants F_1, F_2, H^j, G^j with $\overline{\varepsilon}_1$, $\overline{\varepsilon}_2$, ϕ_d^j, ϕ^j are obtained by substituting the expressions of σ, τ, D, D_z into Eqs. (23):

$$F_1=A\overline{\varepsilon}_1+\beta^{j'}\phi^{j'}, \quad F_2=\overline{A\varepsilon}_2+\overline{\beta}^{j'}\phi_d^{j'},$$

$$H^j=\overline{\beta}^{j^{\mathrm{T}}}\overline{\varepsilon}_2-\overline{E}^{jj'}\phi_d^{j'}, \quad G^j=\beta^{j^{\mathrm{T}}}\overline{\varepsilon}_1-E^{jj'}\phi^{j'} \qquad (31)$$

where

Table 2 Three-dimensional results and percentage error for present theory and FSDT for plate (b) for $n=3$

		load case 1			load case 2		
			% error in			% error in	
entity	S	3D	Pres.	FSDT	3D	Pres.	FSDT
w_c	5	0.13230	0.66	28.48	-0.05213	4.56	-103.9
	10	0.04762	0.39	7.59	-0.03216	-4.21	-55.33
	20	0.02107	-0.33	1.58	-0.02329	-2.66	-21.11
	100	0.01183	-0.03	0.04	-0.01985	-0.15	-1.03
σ^e	5	-0.17387	42.18	-50.52	-33.166	-24.65	-37.76
	10	-0.12604	7.09	-31.07	-25.186	-13.01	-17.79
	20	-0.09864	1.26	-11.56	-21.992	-4.32	-5.73
	100	-0.08785	0.04	-0.55	-20.796	-0.19	-0.25
σ^p	5	0.17387	-23.11	-75.98	36.688	2.89	-8.92
	10	0.07405	-12.10	-49.60	37.962	1.09	-11.40
	20	0.04597	-5.38	-21.31	38.477	0.25	-12.44
	100	0.03638	-0.76	-1.58	38.671	-0.06	-12.83
τ_{zx}	5	-0.10087	-6.92	43.69	-23.875	17.31	29.51
	10	-0.12966	0.55	12.37	-28.581	5.37	8.33
	20	-0.14146	0.35	3.22	-30.321	1.44	2.18
	100	-0.14593	0.02	0.13	-30.961	0.06	0.09
τ_{xy}	5	0.08273	4.62	7.14	-2.4917	-5.08	-22.94
	10	0.04545	6.61	-0.90	-1.7890	-2.90	-12.92
	20	0.02687	2.88	-2.18	-1.4755	-1.40	-4.95
	100	0.01912	0.16	-0.17	-1.3458	-0.07	-0.23
ϕ (1) or D_z (2)	5	0.63291	-23.78	-19.64	1.3451	-2.29	-4.03
	10	0.18876	-7.14	-14.81	1.2994	-0.79	-1.31
	20	0.08045	-2.57	-9.15	1.2841	-0.17	-0.31
	100	0.04524	-0.15	-0.67	1.2787	0.06	0.05

$$A=\langle f_3^{\mathrm{T}}(z)\overline{Q}f_3(z)\rangle,\quad \overline{A}=\langle f_4^{\mathrm{T}}(z)\hat{Q}f_4(z)\rangle,\quad \beta^{j'}=\langle f_3^{\mathrm{T}}(z)\overline{e}_3^{\mathrm{T}}\Psi_{\phi,z}^{j'}(z)\rangle,$$
$$\overline{\beta}^{j'}=\langle f_4^{\mathrm{T}}(z)\hat{e}\Psi_{\phi}^{j'}(z)\rangle,\quad E^{jj'}=\langle \overline{\eta}_{33}\Psi_{\phi,z}^{j}(z)\Psi_{\phi,z}^{j'}(z)\rangle,\quad \overline{E}^{jj'}=\langle \hat{\eta}\Psi_{\phi}^{j}(z)\Psi_{\phi}^{j'}(z)\rangle, \tag{32}$$

$$A=\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1,10} & A_{1,11}^{j'} & A_{1,12}^{j'} & A_{1,13}^{j'} & A_{1,14}^{j'} \\ A_{21} & A_{22} & \cdots & A_{2,10} & A_{2,11}^{j'} & A_{2,12}^{j'} & A_{2,13}^{j'} & A_{2,14}^{j'} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{10,1} & A_{10,2} & \cdots & A_{10,10} & A_{10,11}^{j'} & A_{10,12}^{j'} & A_{10,13}^{j'} & A_{10,14}^{j'} \\ A_{11,1}^{j} & A_{11,2}^{j} & \cdots & A_{11,10}^{j} & A_{11,11}^{jj'} & A_{11,12}^{jj'} & A_{11,13}^{jj'} & A_{11,14}^{jj'} \\ A_{12,1}^{j} & A_{12,2}^{j} & \cdots & A_{12,10}^{j} & A_{12,11}^{jj'} & A_{12,12}^{jj'} & A_{12,13}^{jj'} & A_{12,14}^{jj'} \\ A_{13,1}^{j} & A_{13,2}^{j} & \cdots & A_{13,10}^{j} & A_{13,11}^{jj'} & A_{13,12}^{jj'} & A_{13,13}^{jj'} & A_{13,14}^{jj'} \\ A_{14,1}^{j} & A_{14,2}^{j} & \cdots & A_{14,10}^{j} & A_{14,11}^{jj'} & A_{14,12}^{jj'} & A_{14,13}^{jj'} & A_{14,14}^{jj'} \end{bmatrix}=A^{\mathrm{T}},\quad \beta^{j'}=\begin{bmatrix} \beta_1^{j'} \\ \beta_2^{j'} \\ \vdots \\ \beta_{10}^{j'} \\ \beta_{11}^{jj'} \\ \beta_{12}^{jj'} \\ \beta_{13}^{jj'} \\ \beta_{14}^{jj'} \end{bmatrix},$$

$$\overline{A}=\begin{bmatrix} \overline{A}_{11} & \overline{A}_{12} & \overline{A}_{13}^{j'} & \overline{A}_{14}^{j'} \\ \overline{A}_{21} & \overline{A}_{22} & \overline{A}_{23}^{j'} & \overline{A}_{24}^{j'} \\ \overline{A}_{31}^{j} & \overline{A}_{32}^{j} & \overline{A}_{33}^{jj'} & \overline{A}_{34}^{jj'} \\ \overline{A}_{41}^{j} & \overline{A}_{42}^{j} & \overline{A}_{43}^{jj'} & \overline{A}_{44}^{jj'} \end{bmatrix}=\overline{A}^{\mathrm{T}},\quad \overline{\beta}^{j'}=\begin{bmatrix} \overline{\beta}_{11}^{j'} & \overline{\beta}_{12}^{j'} \\ \overline{\beta}_{21}^{j'} & \overline{\beta}_{22}^{j'} \\ \overline{\beta}_{31}^{jj'} & \overline{\beta}_{32}^{jj'} \\ \overline{\beta}_{41}^{jj'} & \overline{\beta}_{42}^{jj'} \end{bmatrix},\quad \overline{E}^{jj'}=\begin{bmatrix} \overline{E}_{11}^{jj'} & \overline{E}_{12}^{jj'} \\ \overline{E}_{21}^{jj'} & \overline{E}_{22}^{jj'} \end{bmatrix}. \tag{33}$$

Substitution of the expressions of the resultants from Eqs. (31) into Eqs. (29), yields following coupled equilibrium equations in terms of $\underline{U}=[u_{0_x}\, u_{0_y}\, w_0\, \psi_{0_x}\, \psi_{0_y}\, \psi_{0_x}\, \psi_{0_y}\, \phi^1\, \phi^2 \ldots \phi^{n_\phi}]^{\mathrm{T}}$, taking into account the zero elements of A, $\overline{A}$, β, $\overline{\beta}$, $\overline{E}$ for cross-ply laminates:

$$L\overline{U}=\overline{P} \tag{34}$$

where $\overline{P}=[P_1\, P_2\, P_3\, P_4\, P_5\, P_6^1\, P_6^2 \ldots P_6^{n_\phi}]^{\mathrm{T}}$. L is differential operator matrix with $L_{ip}=L_{pi}$ and

$$L_{11}=-A_{11}(\ )_{,xx}-A_{33}(\ )_{,yy},\quad L_{12}=-(A_{12}+A_{33})(\ )_{,xy},$$

$$L_{13}=A_{14}(\ )_{,xxx}+(A_{15}+2A_{36})(\ )_{,xyy},$$

$$L_{14}=-A_{17}(\ )_{,xx}-A_{38}(\ )_{,yy},\quad L_{15}=-(A_{1,10}+A_{39})(\ )_{,xy},$$

$$L_{22}=-A_{22}(\ )_{,yy}-A_{33}(\ )_{,xx},$$

$$L_{1,5+j'}=-A_{1,11}^{j'}(\ )_{,xxx}-(A_{3,12}^{j'}+A_{3,13}^{j'}+A_{1,14}^{j'})(\ )_{,xyy}-\beta_1^{j'}(\ )_{,x},$$

$$L_{23}=(A_{24}+2A_{36})(\ )_{,xxy}+A_{25}(\ )_{,yyy},$$

$$L_{24}=-(A_{27}+A_{38})(\ )_{,xy},$$

$$L_{2,5+j'}=-(A_{2,11}^{j'}+A_{3,12}^{j'}+A_{3,13}^{j'})(\ )_{,xxy}-A_{2,14}^{j'}(\ )_{,yyy}-\beta_2^{j'}(\ )_{,y}$$

$$L_{25}=-A_{39}(\ )_{,xx}-A_{2,10}(\ )_{,yy},\quad L_{33}=-A_{44}(\ )_{,xxxx}-(A_{45}+A_{54}+4A_{66})(\ )_{,xxyy}-A_{55}(\ )_{,yyyy},$$

$$L_{34}=A_{47}(\ )_{,xxx}+(A_{57}+2A_{68})(\ )_{,xyy},$$

$$L_{35}=(A_{4,10}+2A_{69})(\ )_{,xxy}+A_{5,10}(\ )_{,yyy},$$

$$L_{3,5+j'}=A_{4,11}^{j'}(\ )_{,xxxx}+(A_{4,14}^{j'}+2A_{6,12}^{j'}+2A_{6,13}^{j'}+A_{5,11}^{j'})(\ )_{,xxyy}+A_{5,14}^{j'}(\ )_{,yyyy}+\beta_4^{j'}(\ )_{,xx}+\beta_5^{j'}(\ )_{,yy},$$

$$L_{44}=\overline{A}_{11}-A_{77}(\ )_{,xx}-A_{88}(\ )_{,yy},\quad L_{45}=-(A_{7,10}+A_{89})(\ )_{,xy},$$

$$L_{4,5+j'}=-A_{7,11}^{j'}(\ )_{,xxx}-(A_{7,14}^{j'}+A_{8,12}^{j'}+A_{8,13}^{j'})(\ )_{,xyy}+(\overline{A}_{13}^{j'}+\overline{\beta}_{11}^{j'}-\beta_7^{j'})(\ )_{,x},$$

$$L_{55}=\overline{A}_{22}-A_{10,10}(\ )_{,yy},$$

$$L_{5,5+j'}=-(A_{9,12}^{j'}+A_{9,13}^{j'}+A_{10,11}^{j'})(\ )_{,xxy}-A_{10,14}^{j'}(\ )_{,yyy}+(\overline{A}_{24}^{j'}+\overline{\beta}_{22}^{j'}-\beta_{10}^{j'})(\ )_{,y}$$

$$L_{5+j,5+j'}=-A_{11,11}^{jj'}(\ )_{,xxxx}-(A_{11,14}^{jj'}+A_{12,12}^{jj'}+A_{12,13}^{jj'}+A_{13,12}^{jj'}+A_{13,13}^{jj'}+A_{14,11}^{jj'})(\ )_{,xxyy}-A_{14,14}^{jj'}(\ )_{,yyyy}+[\overline{A}_{33}^{jj'}-\beta_{11}^{jj'}-\beta_{11}^{j'j}+\overline{\beta}_{31}^{jj'}+\overline{\beta}_{31}^{j'j}-\overline{E}_{11}^{jj'}](\ )_{,xx}+[\overline{A}_{44}^{jj'}-\beta_{14}^{jj'}-\beta_{14}^{j'j}+\overline{\beta}_{42}^{jj'}+\overline{\beta}_{42}^{j'j}-\overline{E}_{22}^{jj'}](\ )_{,yy}+E^{jj'};$$

$$P_1=P_2=P_4=P_5=0,\quad P_3=-F_3,\quad P_6^j=-F_6^j \tag{35}$$

where $j,\ j'=1,\ldots,n_\phi$.

To assess the theory developed herein, by comparison with the analytical three-dimensional piezoelasticity solution, analytical Navier solution is obtained for simply-supported rectangular plates of sides a, b along the axes x, y for the boundary conditions

$$\text{at } x=0,a:\quad N_x, u_{0_y}, w_0, M_x, P_x, \Psi_{0_y}, \phi^j, S_x^j=0;$$

$$\text{at } y=0,b:\quad N_y, u_{0_x}, w_0, M_y, P_y, \Psi_{0_x}, \phi^j, S_y^j=0; \tag{36}$$

for $j=1,\ldots,n_\phi$. The solution is expanded as

$$\begin{bmatrix} w_0 & \phi^j \\ u_{0_x} & \psi_{0_x} \\ u_{0_y} & \psi_{0_y} \end{bmatrix} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \begin{bmatrix} [w_0 & \phi^j]_{nm}\sin(n\pi x/a)\sin(m\pi y/b) \\ [u_{0_x} & \psi_{0_x}]_{nm}\cos(n\pi x/a)\sin(m\pi y/b) \\ [u_{0_y} & \psi_{0_y}]_{nm}\sin(n\pi x/a)\cos(m\pi y/b) \end{bmatrix}.$$

Equations (34) yield algebraic equations for n, mth Fourier component. These are not listed for brevity. τ can be obtained using Eq. $(2)_2$ or more accurately by integrating the three-dimensional equations of equilibrium.

4 Numerical Results and Discussion

The accuracy of present theory is assessed against the analytical three-dimensional piezoelasticity solution, [6]. The three-dimensional solution is obtained by analytically solving the three-dimensional piezoelasticity constitutive relations, strain-displacement and electric field-potential relations, and the stress equilibrium and charge equilibrium equations subject to the exact satisfaction of the boundary and interface continuity conditions for simply supported plates with $u_y=0$, $w=0$, $\sigma_x=0$, $\phi=0$ at $x=0$, a and $u_x=0$, $w=0$, $\sigma_y=0$, $\phi=0$ at $y=0$, b. Since the number of displacement unknowns in the present theory is the same as in FSDT, results are also compared with the coupled FSDT, [25], with shear correction factors (SCFs) according to Whitney [39]. Two inhomogeneous hybrid simply-supported plates (a) and (b) consisting of an elastic substrate with a layer of PZT-5A, [40], of thickness $0.1h$ bonded to its top are analyzed. The top and the bottom of the substrate are grounded. The stacking order is mentioned from the bottom. The substrate of plate (a) has five plies of thickness $0.09h/0.225h/0.135h/0.18h/0.27h$ of materials 1/2/3/3/3 with orientation θ_k as [0°/0°/0°/90°/0°]. The plies have highly inhomogeneous stiffness in tension and shear and is a good test case, [41]. The substrate of plate (b) is a graphite-epoxy composite (material 4, [40]) laminate with 4 layers of equal thickness $.225h$ with symmetric layup [0°/90°/90°/0°]. The PZT-5A layer has poling in $+z$ direction. Plate (a) is a square plate and plate (b) is a rectangular plate with $b/a=3$. Convergence studies have revealed that converged results are obtained for plates (a) and (b) by dividing the PZT layer into four equal sublayers for discretizing ϕ. The material properties are:

$[(Y_1, Y_2, Y_3, G_{12}, G_{23}, G_{31}), \nu_{12}, \nu_{13}, \nu_{23}]=$

Material 1: [(6.9,6.9,6.9,1.38,1.38,1.38) GPa, 0.25, 0.25, 0.25]

Material 2: [(224.25,6.9,6.9,56.58,1.38,56.58) GPa, 0.25, 0.25, 0.25]

Material 3: [(172.5,6.9,6.9,3.45,1.38,3.45) GPa, 0.25, 0.25, 0.25]

Material 4: [(181,10.3,10.3,7.17,2.87,7.17) GPa, 0.25, 0.25, 0.33]

PZT-5A: [(61.0,61.0,53.2,22.6,21.1,21.1) GPa, 0.35, 0.38, 0.38], and

$[(d_{31}, d_{32}, d_{33}, d_{15}, d_{24}), (\eta_{11}, \eta_{22}, \eta_{33})]=[(-171, -171, 374, 584, 584)\times 10^{-12}\ \mathrm{m/V}, (1.53, 1.53, 1.5)\times 10^{-8}\ \mathrm{F/m}]$.

Two load cases considered are:

1. pressure $p_z^2=-p_0\sin(n\pi x/a)\sin(\pi y/b)$ on the top surface with closed circuit condition $\phi_{n_\phi}=0$ on it.
2. applied actuation potential $\phi^{n_\phi}=\phi_0\sin(n\pi x/a)\sin(\pi y/b)$ on the top surface with zero pressure ($p_z^1=p_z^2=0$).

The results for these cases are nondimensionalized with $S=a/h$, $d_T=374\times 10^{-12}\ \mathrm{CN}^{-1}$, $Y_T=6.9$ GPa for plate (a) and 10.3 GPa for plate (b):

1. $(\overline{u}, \overline{v}, \overline{w})=100(u, v, w/S)Y_T/hS^3p_0$, $(\overline{\sigma}_x, \overline{\sigma}_y, \overline{\tau}_{zx})=(\sigma_x, \sigma_y, S\tau_{zx})/S^2p_0$, $\phi=10^4\phi Y_T d_T/hS^2p_0$,

2. $(\tilde{u},\tilde{v},\tilde{w})=100(u,v,w/S)Sd_T\phi_0$, $(\tilde{\sigma}_x,\tilde{\sigma}_y,\tilde{\tau}_{zx})=(\sigma_x,\sigma_y, S\tau_{zx})h/Y_Td_T\phi_0$, $\tilde{\phi}=\phi/\phi_0$, $\tilde{D}_z=D_zh/100Y_Td_T^2\phi_0$.

The dimensionless entities are chosen in such a way that their values are almost independent of S for thin plates having large value of S. Hence the dimensionless parameter S appears in some of them. These dimensionless entities are meaningful since they have constant values in a coupled CLT.

The present results are compared with the three-dimensional results obtained as per Ref. [6] and the coupled FSDT results obtained as per Ref. [25] with Whitney's SCFs, [39]. The thickness distributions of u or v, w, σ_x, τ_{zx} are compared in Figs. 2 and 3 for plate (a) and in Figs. 4 and 5 for plate (b). The highly zig-zag in-plane displacements u and v are well predicted by the present theory except for the lower-half of test plate (a) and upper-half of hybrid composite plate (b) under potential load for the thick case $S=5$. The smeared laminate theories like FSDT cannot predict zig-zag variation. The thickness variation of w for load case 2 (Figs. 3 and 5) is very well captured by the present theory, even for thick plates in case of hybrid plate (b), since it includes the effect of strain ε_z induced by ϕ through d_{33}. The thickness distributions of σ_x, τ_{zx} in the present theory are in good agreement with the three-dimensional solution. The thickness distributions of electric potential ϕ for plate (b) with the top surface of the piezoelectric layer in closed circuit ($\phi_{n_\phi}=0$) and open circuit ($q_{n_\phi}=0$) conditions are compared in Fig. 6. The present theory yields good prediction of the potential field ϕ for both closed and open-circuit conditions, used in actuation and sensory modes, respectively.

The three-dimensional results and the error percentages with respect to the three-dimensional solution in the response predicted by the present theory and the coupled FSDT for plates (a) and (b) for the two load cases with the Fourier term $n=1$ are given in Table 1. w_c is the deflection w at the center. For plate (a), the stress σ^e is σ_x in the elastic substrate at $z=-.41h^+$ for case 1 and at $z=.265h$ for case 2. σ^p is the maximum stress σ_x in the piezoelectric layer occurring at its top and bottom faces for cases 1 and 2, respectively. For plate (b), σ^e is the maximum σ_x in the elastic substrate occurring at its bottom surface for case 1 and top surface ($z=.4h^-$) for case 2. σ^p is the maximum σ_x in the piezoelectric layer occurring at its top and bottom faces for cases 1 and 2, respectively. τ_{zx} is the maximum stress at $z=0$ for load case 1 and at the PZT interface for case 2. τ_{xy} is the maximum stress at the top surface. ϕ and D_z are the maximum values, respectively, at midsurface and top of PZT layer for load cases 1 and 2. The present central deflection agrees well with the three-dimensional solution for all cases, except for a relatively large error of 9.4% obtained in case of thick test plate (a) with $S=5$ under potential load.

For the thick test plate a with $S=5$, the error of the present theory with respect to the three-dimensional results for w, σ^e and τ_{zx} is within 1.5% for case 1 with $n=1$, whereas there is much larger error in FSDT of 15.4% for deflection w and upto 50.3% for the in-plane stresses. For a moderately thick plate (a) with $S=10$, the error in FSDT is 4.7% for w and upto 19.6% for the inplane stresses for case 1. For the moderately thick plate under load case 2 with $S=10$, the error in the present theory is 3.1% for w and within 4.2% all stresses, whereas the error in FSDT is larger being 8.0% for w and upto 8.0% for stresses. The transverse electric displacement D_z induced at the actuated PZT surface is accurately predicted by both the theories.

For the thick hybrid composite plate (b) with $S=5$, the error of the present theory is within 3.6% for all entities except σ^p(7.6%) for load 1 and σ^e(7.0%) for load 2, whereas there is very large error in FSDT of 29.1% for w for load case 2 and upto 31.0% and 12.1% for the stresses for load cases 1 and 2, respectively. For the moderately thick plate (a) with $S=10$, the error of the present theory is within 1.3% for all entities except σ^p(3.0%) and σ^e(2.0%), whereas the error in FSDT is large, being 9.5% for w for load case 2 and upto 10.9% and 10.1% for the stresses for loads 1 and 2, respectively.

The three-dimensional results and the error percentages in the present theory and FSDT for the response of plate (b) for the two load cases with $n=3$ are given in Table 2. The error in both the two-dimensional theories increases with higher n, since it effectively decreases the span-to-thickness ratio. The error of the present theory for w is 0.7% and 4.6% for loads 1 and 2, respectively, as against 28.5% and 103.9% in FSDT. In stresses too, the present theory shows remarkable improvement over FSDT in all cases.

The above results reveal that the present theory yields generally accurate results which are consistently superior to FSDT for both mechanical and electrical loads. Though FSDT yields relatively good prediction of the central deflection for the pressure load case, it gives highly erroneous results for the inplane stresses in the elastic as well as piezoelectric layers. This inferior performance of FSDT is essentially due to the globally linear approximation of the in-plane displacements across the laminate thickness in this model, which cannot account for the zigzag variation of these displacements across the thickness as obtained from the three-dimensional solution, shown in Figs. 2–5. Moreover, this theory does not satisfy the transverse shear stress continuity conditions at the layer interfaces. In contrast, the approximations for the displacement field across the thickness in the present theory are much closer to the three-dimensional solution. It also satisfies the shear stress continuity conditions at the layers interfaces. However, the transverse normal stress induced due to the piezoelectric stress constant e_{33} for the same deflection appears to have a greater value in the potential load case 2 than in the pressure load case 1. Therefore, the neglect of σ_z in the constitutive equation causes greater error in the deflection w in load case 2 than in load case 1 in the present model. It is important to note that the performance of FSDT strongly depends on the choice of the shear correction factors, which again depend on the boundary and loading conditions of the plate, [42]. The Whitney's SCFs used in the present study, which are calculated for cylindrical bending of plates, may not be valid for plates with more complex geometries, boundary and loading conditions of the plate. The much larger error in FSDT for deflection in load case 2 compared to load case 1 appears to be partly due to this reason. The present theory does not suffer from this drawback.

5 Conclusions

The zig-zag coupled theory presented herein, for hybrid plates with surface bonded or embedded piezo-electric layers, is the first coupled theory in which the shear stress continuity conditions and shear traction free conditions are satisfied exactly, even for the case of nonzero in-plane potential field and the effect of piezoelectric transverse normal strain is accounted for in the transverse displacement field. Its accuracy is established by comparison with the analytical three-dimensional piezoelasticity solution by considering two thick, moderately thick and thin plates with highly heterogeneous layups. The present accurate theory is also economical since the number of primary mechanical variables is the same as that of FSDT which yields poor results for moderately thick heterogeneous plates. The theory can accurately model closed and open-circuit electric boundary conditions in the sensor and actuator layers. The theory can readily be extended to any general anisotropic laminate configuration without any additional approximations. This work is in progress.

Nomenclature

a, b, h, L = sides along axes x, y, thickness, number of plys

$A, \bar{A}$ = beam stiffness matrices

$d_{ij}, \bar{e}_{ij}$ = piezoelectric strain and stress constants

$D_x, D_y, D_z; E_x, E_y, E_z$ = electric displacements, electric field

$F_6^j, G^j; H^j$ = electric load, resultants of D_z, D_x, D_y

$G_{ij}, Y_i, \nu_{ij}, \overline{Q}_{ij}$ = elastic modulii, Poisson's ratios, reduced stiffnesses

k_0, u_0, ψ_0 = midsurface layer and its displacement matrices

L = matrix of differential operators

N, M, P, S^j = stress resultants of inplane stress matrix σ

$\overline{P}, \overline{U}$ = load vector, vector of primary variables

q_{j_i} = extraneous charge density on the actuated surface at $z_\phi^{j_i}$

R^k, R^{kj} = matrices of cubic layerwise functions

S = thickness ratio a/h

$u; u_x, u_y, w; \phi$ = in-plane displacement matrix, displacements, potential

$\underline{u}_k, \psi_k$ = displacement matrices for layer k

V, Q, Q^j, V_ϕ^j = stress resultants of transverse shear stress matrix τ

z_{k-1} = z-coordinate of bottom of kth layer

$\beta^j, \overline{\beta^j}$ = beam electromechanical coupling matrices

$\sigma_x, \sigma_y, \tau_{xy}; \varepsilon_x, \varepsilon_y, \gamma_{xy}$ = in-plane stresses and strains

$\tau_{yz}, \tau_{zx}; \gamma_{yz}, \gamma_{zx}$ = transverse shear stresses and strains

$E^{jj'}, \overline{E^{jj'}}$ = plate dielectric matrices

$\overline{\eta_{ij}}$ = electric permittivities

$\Psi_\phi^j, \overline{\Psi_\phi^j}$ = interpolation function and related integral function

References

[1] Chopra, I., 2002, "Review of the State of Art of Smart Structures and Integrated Systems," AIAA J., **40**, pp. 2145–2187.

[2] Yu, Y. Y., 1995, "Some Recent Advances in Linear and Nonlinear Dynamical Modeling of Elastic and Piezoelectric Plates," J. Intell. Mater. Syst. Struct., **6**, pp. 237–254.

[3] Chee, C. Y. K., Tong, L., and Steven, G. P., 1998, "A Review on the Modeling of Piezoelectric Sensors and Actuators Incorporaed in Intelligent Structures," J. Intell. Mater. Syst. Struct., **9**, pp. 3–19.

[4] Saravanos, D. A., and Heyliger, P. R., 1999, "Mechanics and Computational Models for Laminated Piezoelectric Beams, Plates, and Shells," Appl. Mech. Rev., **52**, pp. 305–320.

[5] Gopinathan, S. V., Vardan, V. V., and Vardan, V. K., 2000, "A Review and Critique of Theories for Piezoelectric Laminates," Smart Mater. Struct., **9**, pp. 24–48.

[6] Heyliger, P., 1994, "Static Behavior of Laminated Elastic/Piezoelectric Plates," AIAA J., **32**, pp. 2481–2484.

[7] Lee, J. S., and Jiang, L. Z., 1996, "Exact Electroelastic Analysis of Piezoelectric Laminae via State Space Approach," Int. J. Solids Struct., **32**, pp. 977–990.

[8] Vel, S. S., and Batra, R. C., 2000, "Three-Dimensional Analytical Solution for Hybrid Multilayered Piezoelectric Plates," ASME J. Appl. Mech., **67**, pp. 558–567.

[9] Ha, S. K., Keilers, C., and Chang, F. K., 1992, "Finite Element Analysis of Composite Structures Containing Distributed Piezoceramic Sensors and Actuators," AIAA J., **30**, pp. 772–780.

[10] Crawley, E. F., and Lazarus, K. B., 1991, "Induced Strain Actuation of Isotropic and Anisotropic Plates," AIAA J., **29**, pp. 944–951.

[11] Wang, B. T., and Rogers, C. A., 1991, "Laminate Plate Theory for Spatially Distributed Induced Strain Actuators," J. Compos. Mater., **25**, pp. 433–452.

[12] Yang, J. S., 1997, "Equations for Elastic Plates With Partially Electroded Piezoelectric Actuators in Flexure With Shear Deformation and Rotory Inertia," J. Intell. Mater. Syst. Struct., **8**, pp. 444–451.

[13] Chattopadhya, A., and Seely, E., 1997, "A Higher Order Theory for Modeling Laminates With Induced Strain Actuators," Composites, Part B, **28B**, pp. 243–252.

[14] Lee, C. K., and Moon, F. C., 1989, "Laminated Piezopolymer Plates for Torsion and Bending Sensors and Actuators," J. Acoust. Soc. Am., **85**, pp. 2432–2439.

[15] Lee, C. K., 1990, "Theory of Laminated Piezoelectric Plates for the Design of Distributed Sensors/Actuators. Part 1: Governing Equations and Reciprocal Relationships," J. Acoust. Soc. Am., **87**, pp. 1144–1158.

[16] Tzou, H. S., 1989, "Distributed Sensing and Controls of Flexible Plates and Shells Using Distributed Piezoelectric Element," J. Wave Mat. Interact., **4**, pp. 11–29.

[17] Hwang, W. S., and Park, H. C., 1993, "Finite Element Modelling of Piezoelectric Sensors and Actuators," AIAA J., **31**, pp. 930–937.

[18] Chandrashekhra, K., and Agarwal, A. N., 1993, "Active Vibration Control of Laminated Composite Plates Using Piezoelectric Devices: A Finite Element Approach," J. Intell. Mater. Syst. Struct., **4**, pp. 496–508.

[19] Jonnalagadda, K. D., Blandford, G. E., and Tauchert, T. R., 1994, "Piezothermoelastic Composite Plate Analysis Using First Order Shear Deformation Theory," Comput. Struct., **51**, pp. 79–89.

[20] Chattopadhya, A., Gu, H., and Dragomir-Daescu, D., 1999, "Dynamics of Delaminated Composite Plates With Piezoelectric Actuators," AIAA J., **37**, pp. 248–254.

[21] Vel, S. S., and Batra, R. C., 2001, "Analysis of Piezoelectric Bimorphs and Plates With Segmented Actuators," Thin-Walled Struct., **39**, pp. 23–44.

[22] Detwiler, J. C., Shen, M. H., and Venkaya, V. B., 1995, "Finite Element Analysis of Laminated Composite Structures Containing Distributed Piezoelectric Actuators and Sensors," Finite Elem. Anal. Design, **20**, pp. 87–100.

[23] Yu, J. W., Kang, W. Y., and Kim, S. J., 1995, "Elastic tailoring of Laminated Composite Plate by Anisotropic Piezoelectric Polymers—Theory, Computation, and Experiment," J. Compos. Mater., **29**, pp. 1201–1221.

[24] Huang, J. H., and Wu, T. L., 1996, "Analysis of Hybrid Multilayered Piezoelectric Plates," Int. J. Eng. Sci., **34**, pp. 171–181.

[25] Saravanos, D. A., 1997, "Coupled Mixed-Field Laminate Theory and Finite Element for Smart Piezo-Electric Composite Shell Structures," AIAA J., **35**, pp. 1327–1333.

[26] Kapuria, S., and Dumir, P. C., 2000, "Coupled FSDT for Piezothermoelastic Hybrid Rectangular Plate," Int. J. Solids Struct., **37**, pp. 6131–6153.

[27] Mitchell, J. A., and Reddy, J. N., 1995, "A Refined Hybrid Plate Theory for Composite Laminates With Piezoelectric Laminae," Int. J. Solids Struct., **32**, pp. 2345–2367.

[28] Chattopadhyay, A., Li, J., and Gu, H., 1999, "Coupled Piezoelectric Mechanical Model for Smart Composite Laminates," AIAA J., **37**, pp. 1633–1638.

[29] Robbins, D. H., and Reddy, J. N., 1991, "Analysis of Piezoelectrically Actuated Beams Using a Layerwise Displacement Theory," Comput. Struct., **41**, pp. 265–279.

[30] Heyliger, P. R., Ramirez, G., and Saravanos, D. A., 1994, "Coupled Discrete-Layer Finite Elements for Laminated Piezoelectric Plates," Comm. Num. Meth. Eng., **10**, pp. 971–981.

[31] Carrera, E., 1997, "An Improved Reissner-Mindlin Type Model for the Electromechanical Analysis of Multilayered Plates Including Piezo Layers," J. Intell. Mater. Syst. Struct., **8**, pp. 232–248.

[32] Bisegna, P., Caruso, G., and Maceri, F., 2001, "A Layerwise Reissner-Mindlin-Type Model for the Vibration Analysis and Suppression of Piezoactuated Plates," Comput. Struct., **79**, pp. 2309–2319.

[33] Kim, H. S., Xu, Z., and Chattopadhya, A., 2002, "Interlaminar Stress Analysis of Shell Structures With Piezoelectric Patch Including Thermal Loading," AIAA J., **40**, pp. 2517–2525.

[34] Cho, M., and Parmerter, R. R., 1993, "Efficient Higher Order Composite Plate Theory for General Lamination Configurations," AIAA J., **31**, pp. 1299–1306.

[35] Shu, X., and Sun, L., 1994, "An Improved Simple Higher-Order Theory for Laminated Composite Plates," Comput. Struct., **50**, pp. 231–236.

[36] Kapuria, S., 2001, "An Efficient Coupled Theory for Multi-Layered Beams With Embedded Piezoelectric Sensory and Active Layers," Int. J. Solids Struct., **38**, pp. 9179–9199.

[37] Kapuria, S., Dumir, P. C., and Ahmed, A., 2003, "An Efficient Coupled Layerwise Theory for Dynamic Analysis of Piezoelectric Composite Beams," J. Sound Vib., **261**, pp. 927–944.

[38] Kapuria, S., Dumir, P. C., and Ahmed, A., 2003, "An Efficient Coupled Layerwise Theory for Static Analysis of Piezoelectric Sandwich Beams," Arch. Appl. Mech., **73**, pp. 147–159.

[39] Whitney, J. M., 1973, "Shear Correction Factors for Orthotropic Laminates Under Static Load," ASME J. Appl. Mech., **40**, pp. 302–304.

[40] Xu, K., Noor, A. K., and Tang, Y. Y., 1995, "Three-Dimensional Solutions for Coupled Thermoelectroelastic Response of Multilayered Plates," Comput. Methods Appl. Mech. Eng., **126**, pp. 355–371.

[41] Averril, R. C., and Yip, Y. C., 1996, "An Efficient Thick Beam Theory and Finite Element Model With Zig-Zag Sublaminate Approximation," AIAA J., **34**, pp. 1626–1632.

[42] Pai, P. F., 1995, "A New Look at Shear Correction Factors and Warping Functions of Anisotropic Laminates," Int. J. Solids Struct., **32**, pp. 2295–2313.

Firdaus E. Udwadia
Professor of Civil Engineering,
Aerospace and Mechanical Engineering,
Mathematics, and Information and Operations Management,
University of Southern California,
Los Angeles, CA 90089
e-mail: fudwadia@usc.edu

Robert E. Kalaba
Professor of Biomedical Engineering,
Electrical Engineering, and Economics,
University of Southern California,
Los Angeles, CA 90089

Phailaung Phohomsiri
Graduate Student,
Department of Aerospace and Mechanical Engineering,
University of Southern California,
Los Angeles, CA 90089

Mechanical Systems With Nonideal Constraints: Explicit Equations Without the Use of Generalized Inverses

In this paper we obtain the explicit equations of motion for mechanical systems under nonideal constraints without the use of generalized inverses. The new set of equations is shown to be equivalent to that obtained using generalized inverses. Examples demonstrating the use of the general equations are provided. [DOI: 10.1115/1.1767844]

1 Introduction

When constraints are applied to mechanical systems, additional forces of constraint are produced that guarantee their satisfaction. The development of the equations of motion for constrained mechanical systems has been pursued by numerous scientists and mathematicians, like Appell [1], Beghin [2], Chetaev [3], Dirac [4], Gauss [5], Gibbs [6], and Hamel [7]. All these investigators have used as their starting point the D'Alembert-Lagrange Principle. This principle, which was enunciated first by Lagrange in his *Mechanique Analytique*, [8], can be presumed as being, at the present time, at the core of classical analytical dynamics.

D'Alembert's principle makes an assumption regarding the nature of constraint forces in mechanical systems, and this assumption seems to work well in many practical situations. It states that the total work done by the forces of constraint under virtual displacements is always zero. In 1992 Udwadia and Kalaba [9] obtained a simple, *explicit* set of equations of motion, suited for general mechanical systems, with holonomic and/or nonholonomic constraints. Though their equations encompass time dependent constraints that are (1) not necessarily independent, and (2) nonlinear in the generalized velocities, their equations are valid only when D'Alembert's principle is observed by the constraint forces.

However, in many situations in nature, the forces of constraint in mechanical systems do not satisfy D'Alembert's principle. As stated in Pars's *A Treatise on Analytical Dynamics* [10], ''There are in fact systems for which the principle enunciated [D'Alembert's principle] . . . does not hold. But such systems will not be considered in this book.'' Such systems have been considered to lie beyond the scope of Lagrangian mechanics. Recently, Udwadia and Kalaba [11,12] have developed general, explicit equations of motion for constrained mechanical systems that may or may not satisfy D'Alembert's principle. The statement of their result involves the use of generalized inverses of various matrix quantities, and they derive their results by using the special properties of these generalized inverses.

In this paper we give a new, alternative set of *explicit* equations that describes the motion of constrained mechanical systems that may or may not satisfy D'Alembert's principle. Thus these equations are valid when the forces of constraint may do work under virtual displacements. We show here that there is no need to use any concepts related to generalized inverses in the development of these general equations. The explicit equations developed herein can handle time dependent constraints that are nonlinear in the generalized velocities, as do the equations obtained using generalized inverses. Instead of relying on the properties of generalized inverses, our explicit equations rely on a deeper understanding of virtual displacements as provided in Refs. [13,14].

After obtaining the new equations, we show that they are indeed equivalent to those given earlier by Udwadia and Kalaba [11,12] which make extensive use of generalized inverses. Three illustrative examples are provided showing the use of the new equations. The last example deals with sliding friction.

2 Explicit Equations of Motion for Mechanical Systems With Nonideal Constraints

For an unconstrained system of N particles, Lagrange's equation of motion for the system at time t can be written, using generalized coordinates, as

$$M(q,t)\ddot{q}=F(q,\dot{q},t); \quad q(0)=q_0,\dot{q}(0)=\dot{q}_0, \tag{1}$$

where, q is the generalized coordinate n-vector $q=[q_1,q_2,\ldots,q_n]^T$; M is an n by n symmetric positive definite matrix; and, $F(q,\dot{q},t)$ is the n-vector of the ''given'' force which is a known function of q, $\dot{q}$, and time, t. The number of degrees-of-freedom of the system is equal to the number of generalized coordinates, n, characterizing the configuration of the system at any time, t. The acceleration, $a(t)$, of the unconstrained system described by Eq. (1) is then given by $a(t)=M(q,t)^{-1}F(q,\dot{q},t)$.

Let the system described by Eq. (1) be now *further* constrained by the m constraint equations

$$\varphi_i(q,\dot{q},t)=0, \quad i=1,2,\ldots,m, \tag{2}$$

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 Applied Mechanics Division, Jan. 14, 2003; final revision, March 8, 2004. Associate Editor: O. O'Reilly. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

in which $k<m$ of these constraint equations are independent. We shall assume that the constraint equations satisfy the initial conditions given in Eq. (1). Equation set (2) includes both holonomic and nonholonomic constraints. Assuming sufficient smoothness, we can differentiate equation set (2) with respect to time t to obtain

$$A(q,\dot{q},t)\ddot{q}=b(q,\dot{q},t), \tag{3}$$

where the elements of A and b are known functions of q, $\dot{q}$, and t and the matrix A is an m by n matrix that has rank k.

The presence of the constraints causes additional constraint forces to arise at each instant of time to assure that the constraints are satisfied. The equation of motion for the constrained system can be then expressed as

$$M\ddot{q}=F(q,\dot{q},t)+F^c(q,\dot{q},t), \tag{4}$$

where, F is the given force, and F^c is the additional force engendered by the presence of the constraints.

Premultiplying Eq. (4) by $M^{-1/2}$, we have

$$M^{1/2}\ddot{q}=M^{-1/2}F(q,\dot{q},t)+M^{-1/2}F^c(q,\dot{q},t), \tag{5}$$

which can be written as

$$\ddot{q}_s(t)-F_s^c=a_s(t). \tag{6}$$

Here we have denoted the ''scaled'' acceleration of the constrained system, $M^{1/2}\ddot{q}$, by $\ddot{q}_s(t)$, the ''scaled'' force of constraint, $M^{-1/2}F^c$, by F_s^c, and the ''scaled'' acceleration of the unconstrained system, $M^{-1/2}F$, by $a_s(t)$. In the same manner, the constraint Eq. (3) at time t can be expressed as $(AM^{-1/2})(M^{1/2}\ddot{q})=b$. Denoting $AM^{-1/2}$ by the m by n matrix B, we obtain

$$B\ddot{q}_s=b. \tag{7}$$

A virtual displacement (see Refs. [13,14]) is any nonzero vector, w, that satisfies the equation

$$A(q,\dot{q},t)w=0. \tag{8}$$

When the constraints are nonideal, the work done, $W(t)$, by the constraint force, F^c, under virtual displacements, w, needs to be specified through knowledge of the n-vector C, so that, [11]

$$W(t)=w^TF^c(q,\dot{q},t)=w^TC(q,\dot{q},t), \tag{9}$$

where $C(q,\dot{q},t)$ is a known n-vector, and characterizes the *nature* of the nonideal constraint force, F^c. This is an extension of D'Alembert's principle.

Equation (8) can be rewritten as

$$(AM^{-1/2})(M^{1/2}w)=0. \tag{10}$$

Similarly, Eq. (9) can be rewritten as

$$(w^TM^{1/2})(M^{-1/2}F^c)=(w^TM^{1/2})(M^{-1/2}C). \tag{11}$$

Denoting $v=M^{1/2}w$, Eq. (10) becomes

$$Bv=0. \tag{12}$$

Since $M^{1/2}$ is nonsingular, v is then any nonzero vector such that relation (12) is satisfied. Furthermore, after denoting $C_s=M^{-1/2}C$, Eq. (11) can be written as

$$v^TF_s^c=v^TC_s. \tag{13}$$

Since B has rank k, there are $n-k$ linearly independent vectors, v, such that $Bv=0$. Assembling then such vectors $v_1 \ldots v_{n-k}$ in the matrix V, we obtain

$$V^TF_s^c=V^TC_s. \tag{14}$$

The matrix V can be constructed by a judicious use of the Gram-Schmidt procedure.

Consider the linear Eqs. (6), (7), and (14). These equations can be expressed as

$$Lr=\begin{bmatrix}[I]_{nxn} & -[I]_{nxn}\\ [B]_{mxn} & [0]_{mxn}\\ [0]_{(n-k)xn} & [V^T]_{(n-k)xn}\end{bmatrix}\begin{bmatrix}\ddot{q}_s\\ F_s^c\end{bmatrix}=\begin{bmatrix}a_s\\ b\\ V^TC_s\end{bmatrix}=s, \tag{15}$$

where L is a $(2n+m-k)$ by $2n$ matrix, r is a $2n$-vector, and s is a $(2n+m-k)$-vector.

The equation set (15) constitutes the fundamental linear set of equations that needs to be solved to obtain both the scaled acceleration, $\ddot{q}_s$, of the constrained system as well as the scaled constraint force, F_s^c. In what follows, we shall show that a solution to this linear system of equations exists and is unique.

We premultiply both sides of Eq. (15) by L^T to obtain the equation

$$L^TLr=\begin{bmatrix}I & B^T & 0\\ -I & 0 & V\end{bmatrix}\begin{bmatrix}I & -I\\ B & 0\\ 0 & V^T\end{bmatrix}r=\begin{bmatrix}I & B^T & 0\\ -I & 0 & V\end{bmatrix}\begin{bmatrix}a_s\\ b\\ V^TC_s\end{bmatrix}. \tag{16}$$

Let us denote

$$D=B^TB, \tag{17}$$

and

$$E=VV^T. \tag{18}$$

Equation (16) can be written as

$$Gr=\begin{bmatrix}[I+D]_{nxn} & -[I]_{nxn}\\ -[I]_{nxn} & [I+E]_{nxn}\end{bmatrix}r=\begin{bmatrix}a_s+B^Tb\\ -a_s+EC_s\end{bmatrix}, \tag{19}$$

where G is the $2n$ by $2n$ symmetric matrix L^TL. We next show that the inverse of the matrix G exists, and we determine it explicitly.

LEMMA 1.

Result 1: The inverse of the matrix G given in Eq. (19) exists and is

$$G^{-1}=\begin{bmatrix}P & J\\ J & S\end{bmatrix}, \tag{20}$$

where

$$J=(D+E)^{-1}=(B^TB+VV^T)^{-1}, \tag{21}$$

$$P=J(I+E), \tag{22}$$

and

$$S=J(I+D). \tag{23}$$

Result 2.

$$SE=I-JD, \tag{24}$$

which is a property that we shall use for the determination of the ''scaled'' force of constraint, F_s^c.

Proof.

Result 1. For simplicity, let us write G^{-1} as

$$G^{-1}=\begin{bmatrix}P & J\\ J & S\end{bmatrix}. \tag{25}$$

Beginning with the condition $G^{-1}G=I$, we obtain

$$\begin{bmatrix}P & J\\ J & S\end{bmatrix}\begin{bmatrix}I+D & -I\\ -I & I+E\end{bmatrix}=\begin{bmatrix}I & 0\\ 0 & I\end{bmatrix},$$

which can be written as

$$\begin{bmatrix}P(I+D)-J & -P+J(I+E)\\ J(I+D)-S & -J+S(I+E)\end{bmatrix}=\begin{bmatrix}I & 0\\ 0 & I\end{bmatrix}. \tag{26}$$

A comparison of the corresponding members on either side of the equality in Eq. (26) shows that

$$P(I+D)-J=I, \tag{27}$$

$$S=J(I+D), \tag{28}$$

$$P=J(I+E), \tag{29}$$

and

$$-J+S(I+E)=I. \tag{30}$$

Then replacing the matrix P obtained from Eq. (29) in Eq. (27) and simplifying that, we have

$$J(I+E)(I+D)-J=J(D+E+ED)=I. \tag{31}$$

Since $E=VV^T$, $D=B^TB$, and $BV=0$, we have $ED=VV^TB^TB=V(BV)^TB=0$.

Thus Eq. (31) can be simplified to

$$J(D+E)=I. \tag{32}$$

From Eqs. (17) and (18), it can be seen that the n by n matrix $D+E=B^TB+VV^T=[B^T\ V]\begin{bmatrix}B\\V^T\end{bmatrix}$. Since the matrix $[B^T\ V]$ has full rank, the rank of $[B^T\ V]\begin{bmatrix}B\\V^T\end{bmatrix}$ is n. Hence, $D+E$ has an inverse and from Eq. (32) the matrix J is given by

$$J=(D+E)^{-1}=(B^TB+VV^T)^{-1}. \tag{33}$$

By Eqs. (25), (28), and (29), the inverse of the matrix G can be then written as

$$G^{-1}=\begin{bmatrix}P & J\\ J & S\end{bmatrix}=\begin{bmatrix}J(I+E) & J\\ J & J(I+D)\end{bmatrix}, \tag{34}$$

where J is given by Eq. (33). □

Result 2: By substituting Eq. (28) in Eq. (30), we obtain $-J+S(I+E)=-J+S+SE=-J+J(I+D)+SE=I$, which can be simplified to

$$SE=I-JD. \qquad \square \tag{35}$$

From Eqs. (19) and (20), the vector $r=\begin{bmatrix}\ddot{q}_s\\F^c\end{bmatrix}$, can be uniquely found as

$$\begin{bmatrix}\ddot{q}_s\\F_s^c\end{bmatrix}=\begin{bmatrix}P & J\\ J & S\end{bmatrix}\begin{bmatrix}a_s+B^Tb\\-a_s+EC_s\end{bmatrix}. \tag{36}$$

Using Eq. (36), the "scaled" force of constraint can be expanded as

$$F_s^c=Ja_s+JB^Tb-Sa_s+SEC_s. \tag{37}$$

From Eqs. (23) and (24), Eq. (37) can be expressed as

$$F_s^c=Ja_s+JB^Tb-Ja_s-JDa_s+(I-JD)C_s$$
$$=JB^Tb-JDa_s+(I-JD)C_s.$$

Noting that $D=B^TB$, the last equation gives a simple form for the constraint force

$$F_s^c=JB^T(b-Ba_s)+(I-JB^TB)C_s. \tag{38}$$

Since the acceleration of the unconstrained system is defined as $a=M^{-1}F$, we have $Ba_s=(AM^{-1/2})(M^{-1/2}F)=A(M^{-1}F)=Aa$. Using this equality, and substituting C_s by $M^{-1/2}C$ in Eq. (38), we get

$$F^c=M^{1/2}F_s^c=M^{1/2}JB^T(b-Aa)+M^{1/2}(I-JB^TB)M^{-1/2}C$$
$$:=F_i^c+F_{ni}^c, \tag{39}$$

which gives the force of constraint F^c explicitly for the constrained system. The subscript i is used to describe the force of constraint were all the constraints to be ideal ($C\equiv 0$); the subscript ni is used to describe the contribution to the total constraint force because of the nonideal nature of the constraints. The explicit equation of motion with nonideal constraints can then be written as

$$M\ddot{q}=F+F^c=F+M^{1/2}JB^T(b-Aa)+M^{1/2}(I-JB^TB)M^{-1/2}C. \tag{40}$$

We emphasize that Eq. (40), which gives *explicitly* the motion of nonholonomic systems with nonideal constraint forces, does *not* involve any generalized inverses, or any Lagrange multipliers.

Previous investigators, so far as we know, have not obtained explicit equations of motion for non-ideal constraints. The only other general equation of motion for constrained mechanical systems with nonideal constraints available in the literature to date appears to be the one obtained in Refs. [11,12] and [15,16]. However, the results that have been obtained so far use the concept of the generalized inverse of a matrix, and the derivations are heavily dependent on the properties of generalized inverses. The equation obtained herein is: (1) explicit; (2) applicable to nonideal constraints; and (3) does not use generalized inverses. In the next section we shall compare our result with those obtained in Refs. [11,12].

There are, however, a number of formulations of the equations of motion for constrained mechanical systems under the assumption that the constraints are all ideal, i.e., when C in Eq. (40) in identically zero for all time. It is then perhaps worthwhile comparing Eq. (40) for $C\equiv 0$, thereby restricting it to only ideal constraints, with formulations that have been obtained by previous investigators. So, to elucidate our equation further, we compare the form of the equation obtained by us with those obtained previously. Though Eq. (40) is also valid for nonideal constraints, in the next paragraph we restrict ourselves, for purposes of comparison with other formulations of the equations of motion obtained by other researchers *only* to when all the constraints are ideal.

Unlike the results obtained in Beghin [2], Chataev [3], Hamel [7], and Lagrange [8], Eq. (40) *explicitly* gives the force of constraint; no Lagrange multipliers are involved. The use of Lagrange multipliers constitutes one approach to solving the problem of constrained motion. We use in this paper a different approach that is innocent of this notion. These multipliers, which were invented by Lagrange, are an intermediary *mathematical device* for solving the problem of constrained motion. As such, they are not intrinsic (essential) to either the description of the physical problem of constrained motion or to the final equation of motion that is obtained, as witnessed by the fact that we make no mention of Lagrange multipliers in our approach. Another important point of difference is that the constraint equations we use to obtain Eq. (40) are more general than those in Appell [1], Beghin [2], Chataev [3], Gibbs [6], Hamel [7], and Synge [17] because the elements of the matrix A are allowed to be not just functions of q and t, but also of $\dot{q}$. This greatly expands the scope of the type of constraints that we use. However, it entails a more delicate interpretation of the concept of virtual displacements (see, Ref. [14]). Furthermore, unlike the formulations of Gibbs [6] and Appell [1] the coordinates we use to describe the constrained motion are the *same* as those used to describe the unconstrained motion; no quasi-coordinates are used, and no coordinate transformations are needed. Dirac [4] developed a set of equations for the constrained motion of hamiltonian systems in which the constraints are not explicitly dependent on time. Our equation differs from his in that: (1) Eq. (40) (with $C(t)\equiv 0$) is also applicable to non-hamiltonian, and dissipative systems, and (2) it allows constraints that contain time explicitly in them. However, Eq. (40) assumes that M is positive definite, while Dirac's method can handle singular Lagrangians; such Lagrangians are more relevant to the field of quantum mechanics (for which Dirac developed his equation) and are seldom found in well-posed problems in classical mechanics.

One consequence of the fact that we use the same set of coordinates to describe the motion of the constrained system as we use to describe the unconstrained system is that our equation provides

the exceptional insight that the total force of constraint is the sum of two forces, as seen from the last two members on the right hand side of the last equality in Eq. (40). *The first corresponds to what would result were all the constraints ideal; the second corresponds to the force caused solely by the nonideal nature of the constraints*. Our ability to obtain the general equation of motion explicitly gives an additional insight when $C \equiv 0$. Nature appears to be acting like a ''control engineer,'' because the second term on the right-hand side of Eq. (40) may be viewed as a ''feedback control force'' proportional to the error, $(b-Aa)$, in the satisfaction of the constraint Eq. (3). We observe that the feedback ''control gain matrix,'' $M^{1/2}JB^T$, which nature uses turns out to be, in general, a highly nonlinear, time-dependent function of q, $\dot{q}$, and t. Such insights into the fundamental nature of constrained motion have been unavailable from previous formulations of the equations for constrained mechanical systems, such as those of Appell, Begin, Chataev, Hamel, Gibbs, Jacobi, Lagrange, and Synge.

3 Connection of Eq. (40) With Previous Results

In this section we show that the equation of motion obtained above is equivalent to the ones previously obtained in Refs. [11,12].

LEMMA 2.

$$JB^T = B^+, \tag{41}$$

where B^+ is the Moore-Penrose inverse of the matrix B.

Proof.

Let us consider a condition $GG^{-1} = G^{-1}G$,

$$\begin{bmatrix} I+D & -I \\ -I & I+E \end{bmatrix}\begin{bmatrix} P & J \\ J & S \end{bmatrix} = \begin{bmatrix} P & J \\ J & S \end{bmatrix}\begin{bmatrix} I+D & -I \\ -I & I+E \end{bmatrix},$$

which can be expanded to

$$\begin{bmatrix} (I+D)P-J & (I+D)J-S \\ -P+(I+E)J & -J+(I+E)S \end{bmatrix} = \begin{bmatrix} P(I+D)-J & -P+J(I+E) \\ J(I+D)-S & -J+S(I+E) \end{bmatrix}. \tag{42}$$

Equating the first element of the second column on either side of Eq. (42), we get

$$(I+D)J-S = -P+J(I+E). \tag{43}$$

After substituting Eqs. (28) and (29) in Eq. (43), we obtain

$$DJ = JD. \tag{44}$$

Similarly, equating the second element of the second column on either side of Eq. (42), we get $-J+(I+E)S = -J+S(I+E)$, which simplifies to

$$ES = SE. \tag{45}$$

As a result of Eqs. (44), (24), and (45), we have

$$DJ = JD = I-SE = I-ES. \tag{46}$$

To show that JB^T is the Moore-Penrose (MP) inverse of the matrix B, we need to prove the following conditions:

1. $B(JB^T)B = B$;
2. $(JB^T)B(JB^T) = JB^T$;
3. $(BJB^T)^T = BJB^T$;

and

4 $(JB^TB)^T = JB^TB$.

1. By using the relations obtained from Eqs. (17), (46), and (18), we have $B(JB^T)B = BJD = B(I-ES) = B-BES = B-(BV)V^TS$. Since $BV=0$, $B(JB^T)B = B$. Thus the first MP condition is satisfied.
2. Due to Eqs. (17) and (46), $(JB^T)B(JB^T) = J(B^TB)JB^T = JDJB^T = J(I-ES)B^T = JB^T - JESB^T$. Since $SE=ES$, $E = VV^T$, and $BV=0$, $JESB^T = JSEB^T$ and $EB^T = VV^TB^T = V(BV)^T = 0$; thus $(JB^T)B(JB^T) = JB^T$, and the second MP condition is satisfied.
3. Since the matrices D and E are symmetric, $J = (D+E)^{-1}$, is also symmetric. Hence $(BJB^T)^T = BJ^TB^T = BJB^T$; thus the third MP condition is satisfied
4. Using Eqs. (17) and (44) we get $(JB^TB)^T = (B^TB)J^T = DJ = JD = J(B^TB)$; thus the fourth MP condition is satisfied.

From the result of lemma 2, after substituting $B^+ = JB^T$ in Eq. (39), we obtain

$$F^c = M^{1/2}B^+(b-Aa) + M^{1/2}(I-B^+B)M^{-1/2}C. \tag{47}$$

The first member on the right of Eq. (47) is the force of constraint that would be generated were all the constraints ideal, the second member gives the contribution to total force of constraint because of its non-ideal nature. Since $B = AM^{-1/2}$, Eq. (47) can be rewritten as

$$F^c = M^{1/2}(AM^{-1/2})^+(b-Aa) + M^{1/2}(I-B^+B)M^{-1/2}C. \tag{48}$$

From Eq. (4), we have $\ddot{q} = M^{-1}F + M^{-1}F^c = a + M^{-1}F^c$. Hence, the explicit equation of motion of the constrained system can be expressed as

$$\ddot{q} = a + M^{-1/2}(AM^{-1/2})^+(b-Aa) + M^{-1/2}(I-B^+B)M^{-1/2}C, \tag{49}$$

which is identical to the equation given by Udwadia and Kalaba (Refs. [11,12]). When $C \equiv 0$, the constraint forces are ideal and D'Alembert's principle is satisfied. Equation (49) then reduces to the result given in Refs. [9] and [13].

4 Examples

In this section, we provide examples that demonstrate the use of the equations of motion (40) for systems with nonideal constraints. The last example deals with a problem of sliding friction.

(a) Consider a particle of unit mass traveling in a three-dimensional configuration space with ''given'' forces $f_x(x,y,z,t)$, $f_y(x,y,z,t)$ and $f_z(x,y,z,t)$ and satisfying the nonholonomic constraint $\dot{y} = z^2\dot{x} + \alpha g(x,t)$, where α is a constant and $g(x,t)$ is a given function of x and t. The initial conditions are taken to be compatible with the nonholonomic constraint.

Since the mass of particle is unity, the unconstrained acceleration is given by

$$a = \begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} f_x(x,y,z,t) \\ f_y(x,y,z,t) \\ f_z(x,y,z,t) \end{bmatrix}. \tag{50}$$

After differentiating the constraint equation with respect to time, we get

$$\begin{bmatrix} -z^2 & 1 & 0 \end{bmatrix}\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = 2z\dot{z}\dot{x} + \alpha g_x\dot{x} + \alpha g_t, \tag{51}$$

where g_x and g_t are partial derivatives of $g(x,t)$ with respect to x and t, respectively. A comparison with Eqs. (3) provides us

$$A=[-z^2 \;\; 1 \;\; 0] \tag{52}$$

and

$$b=2z\dot{z}\dot{x}+\alpha g_x\dot{x}+\alpha g_t. \tag{53}$$

Since $M=I_3$,

$$B=AM^{-1/2}=A. \tag{54}$$

In addition, the solution vectors v_1 and v_2 to Eq. (12) are

$$V=[v_1 \;\; v_2]=\begin{bmatrix} 1 & 1 \\ z^2 & z^2 \\ k_1 & k_2 \end{bmatrix}, \tag{55}$$

where, k_1 and k_2 are arbitrarily chosen, with $k_1 \neq k_2$, so that the column vectors v_1 and v_2 are linearly independent.

As previously shown in lemma 1, $J=(D+E)^{-1}=(B^TB+VV^T)^{-1}$. By Eqs. (54) and (55), we obtain (with $k_1 \neq k_2$)

$$J=\frac{1}{\Delta}\begin{bmatrix} k_1^2+k_2^2+z^4(k_1-k_2)^2 & 2k_1k_2z^2 & -(k_1+k_2)(z^4+1) \\ 2k_1k_2z^2 & z^4(k_1^2+k_2^2)+(k_1-k_2)^2 & -z^2(k_1+k_2)(z^4+1) \\ -(k_1+k_2)(z^4+1) & -z^2(k_1+k_2)(z^4+1) & 2(z^4+1)^2 \end{bmatrix},$$

where $\Delta=(k_1-k_2)^2(z^4+1)^2$.

This gives

$$JB^T=\frac{1}{(z^4+1)}\begin{bmatrix} -z^2 \\ 1 \\ 0 \end{bmatrix}. \tag{56}$$

We could have, of course, started by choosing, say, $k_1=1$ and $k_2=0$ in Eq. (55); we would then have arrived at relation (56) with much less algebra.

Suppose that the constraint force is nonideal and it does work under virtual displacements. Let us assume that the work done by the constraint force is given, for any virtual displacement, w, by

$$w^TF^c=-w^Ta_0(u^Tu)^{\beta}(u/|u|), \tag{57}$$

where $u=[\dot{x} \;\; \dot{y} \;\; \dot{z}]^T$ is the velocity of the particle, $|u|=\sqrt{u^Tu}$, and a_0 and β are constants. In this case, C is a known 3-vector, and can be written as

$$C=-a_0(u^Tu)^{\beta}(u/|u|)=-a_0(\dot{x}^2+\dot{y}^2+\dot{z}^2)^{\beta-1/2}\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix}. \tag{58}$$

After substituting Eqs. (50), and (52) through (58) in Eq. (39), we obtain

$$F^c=\left(\frac{2z\dot{z}\dot{x}+\alpha g_x\dot{x}+\alpha g_t+z^2f_x-f_y}{z^4+1}\right)\begin{bmatrix} -z^2 \\ 1 \\ 0 \end{bmatrix} -a_0\frac{(\dot{x}^2+\dot{y}^2+\dot{z}^2)^{\beta-1/2}}{z^4+1}\begin{bmatrix} \dot{x}+z^2\dot{y} \\ z^2\dot{x}+z^4\dot{y} \\ \dot{z}(1+z^4) \end{bmatrix}. \tag{59}$$

From Eq. (40), the equation of motion of the constrained system is then

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix}=\begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix}+\left(\frac{2z\dot{z}\dot{x}+\alpha g_x\dot{x}+\alpha g_t+z^2f_x-f_y}{z^4+1}\right)\begin{bmatrix} -z^2 \\ 1 \\ 0 \end{bmatrix} -a_0\frac{(\dot{x}^2+\dot{y}^2+\dot{z}^2)^{\beta-1/2}}{z^4+1}\begin{bmatrix} \dot{x}+z^2\dot{y} \\ z^2\dot{x}+z^4\dot{y} \\ \dot{z}(1+z^4) \end{bmatrix}. \tag{60}$$

The first member on the right-hand side of Eq. (60) is the impressed force. The second member is the constraint force that would be generated had the constraint been ideal, and the third member results from the nonideal nature of the constraint that is described by Eq. (57). When $\alpha=0$, and $\beta=1$, the equation of motion (60) becomes identical to that given by Udwadia and Kalaba [11]. We note that here the result is obtained without any reference to generalized inverses.

(b) Consider a bead having a mass m. Suppose that it moves on a circular ring of radius R as shown in Fig. 1. The motion can be described by the coordinates (x,y). The gravitational acceleration, g, is downwards. We assume that the initial conditions on the motion of the bead are compatible with the constraint that it lie on the ring.

Were the bead not constrained to lie on the ring, its unconstrained acceleration would be

$$a=\begin{bmatrix} 0 \\ -g \end{bmatrix}. \tag{61}$$

In this problem, the constraint equation is $x^2+y^2=R^2$. After differentiating the constraint equation twice, we obtain

$$[x \;\; y]\begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix}=-\dot{x}^2-\dot{y}^2, \tag{62}$$

so that

$$A=[x \;\; y], \tag{63}$$

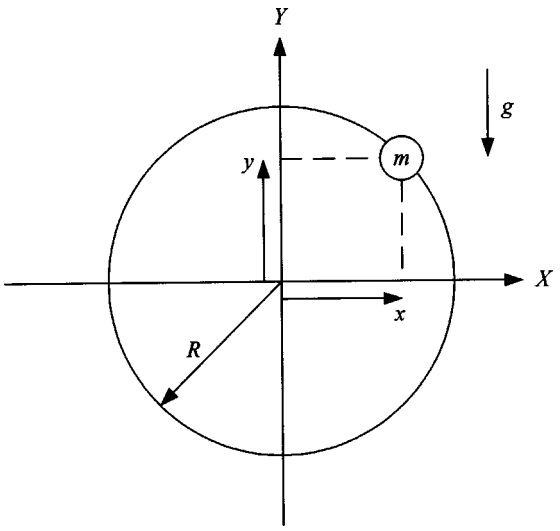


Fig. 1 A bead of mass, *m*, moving on a circular ring of radius, *R*

and

$$b=-\dot{x}^2-\dot{y}^2. \tag{64}$$

Since the mass matrix $M=mI_2$,

$$B=AM^{-1/2}=m^{-1/2}[x\ \ y]. \tag{65}$$

For any virtual displacement $w\neq 0$ such that $Aw=0$, we have $w=\begin{bmatrix} y \\ -x \end{bmatrix}$ so that

$$V=M^{1/2}w=m^{1/2}\begin{bmatrix} y \\ -x \end{bmatrix}. \tag{66}$$

Using Eq. (21), (65) and (66), we obtain

$$JB^T=(B^TB+VV^T)^{-1}B^T=\frac{1}{m^{3/2}R^4}\begin{bmatrix} x^2m^2+y^2 & xy(m^2-1) \\ xy(m^2-1) & y^2m^2+x^2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$=\frac{m^{1/2}}{R^2}\begin{bmatrix} x \\ y \end{bmatrix}. \tag{67}$$

Suppose that the nonideal constraint force, due to the rough surface of the ring, is given by

$$w^TF^c=-w^T\frac{h(x,\dot{x},y,\dot{y},t)}{\sqrt{\dot{x}^2+\dot{y}^2}}\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}, \tag{68}$$

for any virtual displacement w, where h is a known function of x, $\dot{x}$, y, $\dot{y}$, and t.

From the calculation in Eq. (39), the force of constraint on the bead can be expressed as

$$F^c=-\frac{m(\dot{x}^2+\dot{y}^2-yg)}{R^2}\begin{bmatrix} x \\ y \end{bmatrix}+\frac{h(x,\dot{x},y,\dot{y},t)}{\sqrt{\dot{x}^2+\dot{y}^2}}\cdot\frac{(x\dot{y}-y\dot{x})}{R^2}\begin{bmatrix} y \\ -x \end{bmatrix}. \tag{69}$$

Finally, by Eq. (40), the equation of motion of the constrained system is

$$\begin{bmatrix} m\ddot{x} \\ m\ddot{y} \end{bmatrix}=\begin{bmatrix} 0 \\ -mg \end{bmatrix}-\frac{m(\dot{x}^2+\dot{y}^2-yg)}{R^2}\begin{bmatrix} x \\ y \end{bmatrix}+\frac{h(x,\dot{x},y,\dot{y},t)}{\sqrt{\dot{x}^2+\dot{y}^2}}\cdot\frac{(x\dot{y}-y\dot{x})}{R^2}\begin{bmatrix} y \\ -x \end{bmatrix}. \tag{70}$$

The first member on the right-hand side of Eq. (70) is the given force acting on the unconstrained system; the second is the constraint force that would have been generated had the constraint been ideal; and, the last member accounts for the nonideal nature of the constraint.

(c) Consider a rigid block of mass m sliding on an inclined plane that oscillates in the vertical direction with amplitude β and frequency ω, the coefficient of Coulomb friction between the plane and the surface of the block being μ. See Fig. 2. We shall assume that the acceleration of the inclined plane is sufficiently small so that the block does not leave the surface of the plane as it moves under gravity.

In the absence of the inclined plane, the unconstrained equations of motion of the block of mass m and under gravity can be written as

$$\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix}\begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix}=\begin{bmatrix} 0 \\ mg \end{bmatrix}, \tag{71}$$

so that the acceleration, a, of the unconstrained system is given by $a=[0\quad g]^T$, and $M=mI_2$.

The unconstrained system is then subjected to the constraint, namely that the block must lie on the vibrating inclined plane. Hence, the constraint is given by the kinematic relation $y(t)=x(t)\tan\alpha-\beta\sin\omega t$, which can be expressed after differentiation with respect to time t as

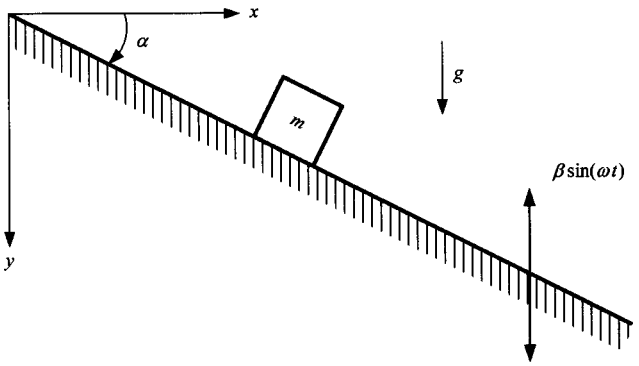


Fig. 2 A block sliding under gravity on an inclined plane ($0<\alpha<\pi/2$) that is vibrating vertically with amplitude β and frequency ω. The coefficient of Coulomb friction between the plane and the block is μ.

$$[-\tan\alpha\ \ 1]\begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix}=\beta\omega^2\sin\omega t. \tag{72}$$

Thus, $A=[-\tan\alpha\ \ 1]$, and b is the scalar $\beta\omega^2\sin\omega t$.

By Eq. (8), we have the virtual displacement

$$w=\delta\begin{bmatrix} 1 \\ \tan\alpha \end{bmatrix}, \tag{73}$$

where δ is any nonzero constant.

Hence, we get

$$V=v=M^{1/2}w=m^{1/2}\delta\begin{bmatrix} 1 \\ \tan\alpha \end{bmatrix}. \tag{74}$$

Since $B=AM^{-1/2}=m^{-1/2}A$, using Eq. (21), we have

$$JB^T=(B^TB+VV^T)^{-1}B^T=\left(m^{-1}\begin{bmatrix} -\tan\alpha \\ 1 \end{bmatrix}[-\tan\alpha\ \ 1]+m\delta^2\begin{bmatrix} 1 \\ \tan\alpha \end{bmatrix}[1\ \ \tan\alpha]\right)^{-1}\left(m^{-1/2}\begin{bmatrix} -\tan\alpha \\ 1 \end{bmatrix}\right),$$

which can be simplified to

$$JB^T=m^{1/2}\cos^2\alpha\begin{bmatrix} -\tan\alpha \\ 1 \end{bmatrix}. \tag{75}$$

Therefore, the force of constraint, were the constraint to be ideal, would then be given by

$$F_i^c=M^{1/2}JB^T(b-Aa)=m\cos^2\alpha\begin{bmatrix} \tan\alpha \\ -1 \end{bmatrix}(g-\beta\omega^2\sin\omega t). \tag{76}$$

In the presence of Coulomb friction, the magnitude of the frictional force is $\mu|F_i^c|$, where $|z|=+\sqrt{z^Tz}$. We note that Coulomb's law of friction is an approximate empirical relation (see Ref. [18]). The relative velocity of the block with respect to the inclined plane is given by $\dot{q}=[\dot{x}\quad \dot{x}\tan\alpha]^T$. The frictional force is in a direction opposite that of this relative velocity. The work done by Coulomb friction under a virtual displacement w is then

$$W=-w^T\left(\mu|F_i^c|\frac{\dot{q}}{|\dot{q}|}\right), \tag{77}$$

so that

$$C=-\mu|F_i^c|\frac{\dot{q}}{|\dot{q}|}=\frac{-\mu|F_i^c|}{|\dot{x}|\sec\alpha}\begin{bmatrix} \dot{x} \\ \dot{x}\tan\alpha \end{bmatrix}=-\mu|F_i^c|\cos\alpha\begin{bmatrix} 1 \\ \tan\alpha \end{bmatrix}s, \tag{78}$$

where, $s=\mathrm{sgn}(\dot{x})$.

Relation (76) yields

$$|F_i^c| = m\cos\alpha|(g-\beta\omega^2\sin\omega t)|. \tag{79}$$

The contribution to the total force of constraint generated by the non-ideal nature of the constraint is then

$$\begin{aligned} F_{ni}^c &= M^{1/2}(I-JB^TB)M^{-1/2}C = m^{1/2}\Bigg(I-m^{1/2}\cos^2\alpha\begin{bmatrix}-\tan\alpha\\1\end{bmatrix} \\ &\quad\times[-\tan\alpha\ \ 1]m^{-1/2}\Bigg)m^{-1/2}\Bigg(-\mu|F_i^c|\cos\alpha\begin{bmatrix}1\\\tan\alpha\end{bmatrix}s\Bigg) \\ &= -\mu|F_i^c|\cos\alpha\begin{bmatrix}1\\\tan\alpha\end{bmatrix}s \\ &= -\mu m\cos^2\alpha|(g-\beta\omega^2\sin\omega t)|\begin{bmatrix}1\\\tan\alpha\end{bmatrix}s. \end{aligned} \tag{80}$$

Note that if the block is to remain in contact with the plane we require $(g-\beta\omega^2\sin\omega t)\geqslant 0$. The equation of motion of the block sliding on the plane, by Eqs. (40), is then

$$\begin{aligned} \begin{bmatrix}m\ddot{x}\\m\ddot{y}\end{bmatrix} &= \begin{bmatrix}0\\mg\end{bmatrix} + m\cos^2\alpha\begin{bmatrix}\tan\alpha\\-1\end{bmatrix}(g-\beta\omega^2\sin\omega t) \\ &\quad -\mu m\cos^2\alpha\begin{bmatrix}1\\\tan\alpha\end{bmatrix}(g-\beta\omega^2\sin\omega t)s. \end{aligned} \tag{81}$$

We note that each of the three members on the right-hand side of Eq. (81) has a simple interpretation. And it is precisely to expose this essential simplicity with which nature seems to operate that we have desisted from simplifying the equation any further. For the first member is the ''given'' force; the second is the force of constraint were the constraint to be ideal; and, the third is the constraint force engendered by the nonideal nature of the constraint.

5 Conclusions

The explicit equations of motion for holonomic and nonholonomic mechanical systems with nonideal constraints have so far been obtained in terms of generalized inverses of matrices. These inverses were first proposed by Moore [19], and their properties were first extensively developed by Penrose [20]. Since the properties of generalized inverses have appeared to be essential in developing these explicit equations, it had been felt that it was because of their relatively recent introduction—in the 1950s—in the scientific literature that the explicit equations of motion for nonholonomic mechanical systems were unavailable until quite recently (see Refs. [11–13], and [15,16]).

In this paper we show that this line of reasoning does not appear to be correct. Rather than reliance on generalized inverses of matrices and their properties, what we may have needed to get the explicit equations of motion is a more refined understanding of the problem, and a further development of concepts that have long since been with us. Among these are: (1) a proper conceptualization of the problem of constrained motion in terms of an unconstrained system, which is then subjected to the imposed constraints, (2) the generalized concept of a virtual displacement vector, described in Ref. [14] and, (3) the use of linear algebra. It is somewhat surprising that though the equations of motion that govern even some of the simplest constrained mechanical systems are nonlinear, it is linear algebra that plays a central role in their development.

We point out that the explicit equations of motion obtained herein, like those obtained earlier (Refs. [11–16]), are completely innocent of the notion of Lagrange multipliers. Over the last 200 years, Lagrange multipliers have been so widely used in the development of the equations of motion of constrained mechanical systems that it is sometimes tempting to mistakenly believe that they possess an instrinsic presence in the description of constrained motion. This is not true. As shown in this paper, neither in the formulation of the physical problem of the motion of constrained mechanical systems nor in the equations governing their motion are any Lagrange multipliers involved. The use of Lagrange multipliers (a mathematical tool invented by Lagrange [8]) constitutes only *one* of the several *intermediary* mathematical devices invented for handling constraints. And, in fact, the direct use of this device appears inapplicable when the constraints are functionally dependent. Lagrange multipliers do not appear in the physical description of constrained motion, and therefore cannot, and do not, ultimately appear in the equations governing such motion.

The explicit equations of motion obtained in this paper apply to general, holonomic, and nonholonomic systems that may have nonideal constraint forces. These constraint forces may, in general, do positive, zero, or negative work under virtual displacements at any time during the motion of the system. The equations given here are the first of their kind that are explicit, and that do *not* require the use of any generalized inverses, nor use of any of their properties.

References

[1] Appell, P., 1899, ''Sur une Forme Generale des Equations de la Dynamique,'' C. R. Acad. Sci., Paris, **129**, pp. 459–460.
[2] Beghin, H., 1967, *Mecanique Theorique et Appliquee*, Gauthier-Villars, Paris.
[3] Chetaev, N. G., 1989, *Theoretical Mechanics*, Mir Publisher, Moscow.
[4] Dirac, P. A. M., 1964, *Lectures in Quantum Mechanics*, Yeshiva University, New York.
[5] Gauss, C. F., 1829, ''Uber Ein Neues Allgemeines Grundgsetz der Mechanik,'' J. Reine Angew. Math., **4**, pp. 232–235.
[6] Gibbs, J. W., 1879, ''On the Fundamental Formulas of Dynamics,'' Am. J. Math., **2**, pp. 49–64.
[7] Hamel, P. G., 1949, *Theoretische Mechanik: Eine Einheitliche Einfuhrung in die Gesamte Mechanik*, Springer-Verlag, Berlin.
[8] Lagrange, J. L., 1787, *Mechanique Analytique*, Mme Ve Courcier, Paris.
[9] Udwadia, F. E., and Kalaba, R. E., 1992, ''A New Perspective on Constrained Motion,'' Proc. R. Soc. London, Ser. A, **439**, pp. 407–410.
[10] Pars, L. A., 1979, *A Treatise on Analytical Dynamics*, Oxbow Press, Woodridge, CT, p. 14.
[11] Udwadia, F. E., and Kalaba, R. E., 2001, ''Explicit Equations of Motion for Mechanical Systems With Nonideal Constraints,'' ASME J. Appl. Mech., **68**, pp. 462–467.
[12] Udwadia, F. E., and Kalaba, R. E., 2002, ''On the Foundations of Analytical Dynamics,'' Int. J. Non-Linear Mech., **37**, pp. 1079–1090.
[13] Udwadia, F. E., and Kalaba, R. E., 1996, *Analytical Dynamics: A New Approach*, Cambridge University Press, Cambridge, UK.
[14] Udwadia, F. E., Kalaba, R. E., and Hee-Chang, E., 1997, ''Equation of Motion for Constrained Mechanical Systems and The Extended D'Alembert's Principle,'' Q. Appl. Math., **LV**(2), pp. 321–331.
[15] Udwadia, F. E., 2000, ''Fundamental Principles of Lagrangian Dynamics: Mechanical Systems With Non-Ideal, Holonomic, and Non-Holonomic Constraints,'' J. Math. Anal. Appl., **252**, pp. 341–355.
[16] Udwadia, F. E., and Kalaba, R. E., 2002, ''What is the General Form of the Explicit Equations of Motion for Constrained Mechanical Systems?'' ASME J. Appl. Mech., **69**, pp. 335–339.
[17] Synge, J. L., 1927, ''On the Geometry of Dynamics,'' Philos. Trans. R. Soc. London, Ser. A, **226**, pp. 31–106.
[18] Bowden, F. P., and Tabor, D., 1974, *Friction: An Introduction to Tribology*, Heinemann, London.
[19] Moore, E. H., 1920, ''On the Reciprocal of The General Algebraic Matrix,'' Bull. Am. Math. Soc., **26**, pp. 394–395.
[20] Penrose, R., 1955, ''A Generalized Inverse of Matrices,'' Proc. Cambridge Philos. Soc., **51**, pp. 406–413.

Applicability and Limitations of Simplified Elastic Shell Equations for Carbon Nanotubes

C. Y. Wang

C. Q. Ru[1]
e-mail: c.ru@ualberta.ca

A. Mioduchowski

Department of Mechanical Engineering,
University of Alberta,
Edmonton T6G 2G8, Canada

This paper examines applicability and limitations of simplified models of elastic cylindrical shells for carbon nanotubes. The simplified models examined here include Donnell equations and simplified Flugge equations characterized by an uncoupled single equation for radial deflection. These simplified elastic shell equations are used to study static buckling and free vibration of carbon nanotubes, with detailed comparison to exact Flugge equations of cylindrical shells. It is shown that all three elastic shell models are in excellent agreement (with relative errors less than 5%) with recent molecular dynamics simulations for radial breathing vibration modes of carbon nanotubes, while reasonable agreements for various buckling problems have been reported previously for Donnell equations. For general cases of buckling and vibration, the results show that the simplified Flugge model, which retains mathematical simplicity of Donnell model, is consistently in better agreement with exact Flugge equations than Donnell model, and has a significantly enlarged range of applicability for carbon nanotubes. In particular, the simplified Flugge model is applicable for carbon nanotubes (with relative errors around 10% or less) in almost all cases of physical interest, including some important cases in which Donnell model results in much larger errors. These results are significant for further application of elastic shell models to carbon nanotubes because simplified shell models, characterized by a single uncoupled equation for radial deflection, are particularly useful for multiwall carbon nanotubes of large number of layers. [DOI: 10.1115/1.1778415]

1 Introduction

Elastic shell models have been effectively used to study mechanical deformation of carbon nanotubes (CNTs), [1,2], especially buckling of CNTs under axial compression, [3–5], bending, [6,4], radial pressure, [7,8], or combined loadings, [9]. Previous work has shown that critical loading and the associated buckling mode predicted by simple isotropic elastic shell models are generally in reasonable agreement with available experiments and molecular dynamics simulations of singlewall nanotubes (SWNTs), [4,9], and multiwall nanotubes (MWNTs) of as many as 20 layers, [8]. In particular, because elastic shell models are relatively simple and cost-effective as compared to experiments and molecular dynamic simulations, they have the potential to offer simple general formulas in some important cases, identify major factors affecting mechanical behavior of CNTs, and explain or predict new physical phenomena. On the other hand, almost all of these studies are limited to reversible elastic deformation of nanotubes, and do not account for fracture and failure, [10,11].

So far, almost all existing elastic shell models for CNTs are based on the simplest Donnell equations of cylindrical shells. It is known that Donnell equations are based on several simplifications, [12–17]: (1) the contribution of two in-plane displacements to the changes in bending curvature is negligible; (2) the contribution of transverse shear forces to the equilibrium in the circumferential direction is negligible; (3) the in-plane inertia is negligible; and (4) the pre-stresses (for buckling problems only) are neglected in the two in-plane equations. The major advantage of Donnell equations over other more accurate shell equations (such as exact Flugge equations of cylindrical shells, [13], which do not reply on any of the above four assumptions) is the remarkably simple mathematical form of Donnell equations, especially due to an uncoupled single equation for the radial deflection. Since the radial deflection is the dominant displacement component for many problems, such as buckling and radial vibration of elastic shells, this advantage largely simplifies technical complexity of elastic shell analysis in many important cases. Here, it is emphasized that this advantage of Donnell equations is even more crucial when elastic shell models are applied to MWNTs of large number of layers, [5,8,9]. On the other hand, although Donnell equations are proved to be an excellent approximate model and almost indistinguishable from exact shell models in many important cases, they indeed led to substantial errors in some cases of practical significance, [18–22]. Hence, in view of unusual geometrical and material characteristics of CNTs, it is necessary to clarify the conditions under which Donnell equations or other simplified elastic shell models are applicable for CNTs.

The present paper gives a systematic study of applicability and limitations of simplified elastic shell equations for CNTs. Here, beside Donnell model, a simplified Flugge model is also examined, which is derived from the exact Flugge equations, [18], based on the last two assumptions of Donnell equations listed above and leads to an uncoupled single equation for the radial deflection, [12,15,17–21]. Various problems of static buckling and free vibration will be discussed with Donnell model (model I), simplified Flugge model (model II), and exact Flugge equations (model III). Since the applicability of Donnell model has been systematically demonstrated for static buckling of CNTs, [1–9], comparison between elastic shell models and available experiments or molecular dynamics simulations will focus on free vibration of CNTs. As will be seen below, the simplified Flugge model (model II), which retains mathematical simplicity of Donnell model (model I), enjoys improved accuracy and enlarged

[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 28, 2003; final revision, March 16, 2004. Associate Editor: H., Gao. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

range of applicability as compared to Donnell model, and thus is recommended for static and dynamic problems of CNTs.

2 Simplified Elastic Shell Models

Almost all existing elastic shell models for CNTs are based on the simplest Donnell equations of cylindrical shells. So, let us begin with a discussion of the assumptions of Donnell equations.

2.1 Donnell Equations (Model I). For almost all important problems of CNTs discussed so far in the literature, tangential external loadings along axial or circumferential direction of CNTs are absent. For example, for MWNTs, because interlayer friction is usually so small that adjacent concentric tubes can almost freely slide to each other, it is assumed, [5,8,9], that the interlayer friction is negligible between any two adjacent tubes. In the absence of any tangential external force, dynamics of the radial deflection $w(x,\theta,t)$ of an elastic cylindrical shell of radius R is uncoupled with other two in-plane displacements, u (axial displacement), and v (circumferential displacement), and is governed by the Donnell equation, [5,8,9]:

$$D\nabla^8 w = \nabla^4 p(x,\theta) + F_x^0 \frac{\partial^2}{\partial x^2}\nabla^4 w + 2\frac{F_{x\theta}^0}{R}\frac{\partial^2}{\partial x\,\partial\theta}\nabla^4 w + \frac{F_\theta^0}{R^2}\frac{\partial^2}{\partial\theta^2}\nabla^4 w - \frac{Eh}{R^2}\frac{\partial^4 w}{\partial x^4} - \rho h\frac{\partial^2}{\partial t^2}\nabla^4 w \tag{1}$$

where x and θ are axial coordinate and circumferential angular coordinate, respectively, w is the radial (inward) deflection, p is the net normal (inward) pressure, F_x^0, F_θ^0, and $F_{x\theta}^0$ are the known uniform membrane forces (called ''pre-stresses''), ρh is the mass density (per unit lateral area), D and h are the effective bending stiffness and thickness of the shell, and E is Young's modulus. Here, the effective bending stiffness D can be independent of the thickness h, and thus not necessarily proportional to h cube, [4]. Once the radial deflection w is determined from (1), other two in-plane displacements u (axial displacement) and v (circumferential displacement) can be determined from other two (axial and circumferential) equations, see e.g., Eqs. (6.33c) of [12], or Eqs. (6.13a,6.13b) of [15]. The Eq. (1) and the other two equations (the latter will not be used in the present paper) are called ''Donnell equations,'' while the single Eq. (1) is also often called ''Donnell equation.''

The single uncoupled Donnell Eq. (1) has been widely used in many problems of cylindrical shells, due to its mathematical simplicity compared to other more accurate shell equations, such as exact Flugge's coupled three equations for the displacement components (u,v,w) (see Eqs. (7a–c) of [13] or Eqs. (11)–(13) of [18] for static case). As mentioned before, the Donnell equations are based on four assumptions, [12–17]: (1) the contribution of two in-plane displacements to the bending curvature is negligible; (2) the contribution of transverse shear forces to the equilibrium in the circumferential direction is negligible; (3) the in-plane inertia is negligible; and (4) the pre-stresses (for buckling problems only) are neglected in two in-plane equations. The last two assumptions imply that the pre-stresses, (F_x^0, F_θ^0, and $F_{x\theta}^0$), and in-plane inertias ($\rho h\partial^2 u/\partial t^2$ and $\rho h\partial^2 v/\partial t^2$) are eliminated from two in-plane equations, and thus the two in-plane equations have the same form as their static counterparts without any pre-stresses. In particular, this means that Donnell Eq. (1) could give satisfactory results only when the in-plane inertia ($\rho h\partial^2 u/\partial t^2$ and $\rho h\partial^2 v/\partial t^2$) is less important than the radial inertia ($\rho h\partial^2 w/\partial t^2$). In addition, as long as the assumption (3) is concerned, it is noticed that the pre-stresses (F_x^0, F_θ^0, and $F_{x\theta}^0$) do occur in each of all three equilibrium equations of more accurate shell models (such as Flugge equations, see Eqs. (5a–c) of [13], or Eqs. (10)–(13) of [14]). Hence, the assumption (3) is indeed a simplifying approximation for buckling of elastic cylindrical shells and could lead to an error as compared to exact shell models such as Flugge equations.

2.2 Simplified Flugge Equations (Model II). It was shown by Kempner [15,18,20,21] that, in the absence of the in-plane inertias and pre-stresses in two in-plane equations, an uncoupled single equation for the radial deflection can be derived from exact Flugge equations of cylindrical shells by a procedure similar to the derivation of Donnell equation. In the presence of all pre-stresses and radial inertia, after neglecting powers of $h^2/12R^2$ (or $(1-\nu^2)D/Eh\cdot 1/R^2$ when D is a material parameter independent of the thickness h) compared to unity, this single equation can be written as (see Eqs. (6.34a–c) of [15], or [18–21]):

$$D\cdot\nabla^4\left(\nabla^2+\frac{1}{R^2}\right)^2 w + \frac{2D\cdot(1-\nu)}{R^6}\left(\frac{\partial^6 w}{\partial x^2\partial\theta^4}+\frac{\partial^4 w}{\partial x^2\partial\theta^2}-R^4\frac{\partial^6 w}{\partial x^6}\right) = \nabla^4 p(x,\theta) + F_x^0\frac{\partial^2}{\partial x^2}\nabla^4 w + 2\frac{F_{x\theta}^0}{R}\frac{\partial^2}{\partial x\,\partial\theta}\left(\nabla^4 w+\frac{\nabla^2 w}{R^2}\right) + \frac{F_\theta^0}{R^2}\nabla^4\left(\frac{\partial^2 w}{\partial\theta^2}+w\right) - \frac{Eh}{R^2}\frac{\partial^4 w}{\partial x^4} - \rho h\frac{\partial^2}{\partial t^2}\nabla^4 w \tag{2}$$

where the pre-stress terms are referred to Eq. (6.36) of [12].

In other words, the simplified Flugge single Eq. (2) is based on Donnell's assumptions (3) and (4) only (partially for the latter, because the pre-stress terms of (2) are approximate in nature), but not on the assumptions (1) and (2). On the other hand, it is seen from (1) and (2) that although the simplified Flugge Eq. (2) is slightly complicated than Donnell Eq. (1), it essentially retains mathematical simplicity of Donnell Eq. (1). Hence, it is of great interest to examine the range of applicability of the simplified Flugge Eq. (2) for CNTs, with a comparison to the Donnell Eq. (1). In connection with this, it should be stated that, in spite of known comparison between Donnell equations and exact Flugge equations, [15,20–22], no detailed comparison has been made between exact Flugge model and the simplified Flugge model (2) for buckling and dynamic problems of elastic shells. Here, we would mention that an interesting equation has been suggested by Morley [20] which is obtained by neglecting some terms in the simplified Flugge Eq. (2) and has an elegant form very close to Donnell Eq. (1). As commented by Donnell [12], ''the choice of coefficients in Morley's solution merely to give the desirable results in certain particular applications, rather than deriving them from basic principle as was done in Flugge's and our own (Donnell) solutions, makes their accuracy somewhat questionable in applications to unchecked problems.'' Hence, in the present paper, we shall focus on the simplified models (I) and (II).

In order to apply the model (I) or (II) to CNTs, it is sufficient to know the bending stiffness D, the in-plane stiffness Eh, the mass density per unit lateral area ρh, and Piosson's ratio ν of CNTs. In particular, these parameters are not dependent of the definition of the thickness h, [3,4]. In fact, the effective bending stiffness of a carbon SWNT is $D=0.85$ eV, its in-plane stiffness is $Eh=360$ J/m^2, the mass density per unite lateral area is $\rho=(2.27$ g/cm$^3)\times 0.34$ nm, and ν can be assumed to be equal to 0.2. For more recent studies on Young's modulus and Poisson ratio of carbon nanotubes and their sensitiveness to diameter smaller than 1 nm, see [23,24].

2.3 Exact Flugge Equations (Model III). Exact Flugge equations will be used in the paper as standard elastic shell model. For static buckling of CNTs, because the general formulas of exact Flugge equations for various buckling problems are available in Flugge's book, [13], the exact Flugge equations in the presence of the pre-stresses, which are somewhat lengthy, will not be cited here (see Eqs. (7a–c) of [13]). For free vibration of CNTs, on the other hand, the exact Flugge equations of cylindrical shells are relatively simple because of the absence of the pre-stresses, and are given by (see [16] and [18])

$$R^2\frac{\partial^2 u}{\partial x^2}+\frac{1}{2}(1-\nu)\frac{\partial^2 u}{\partial \theta^2}+\frac{R}{2}(1+\nu)\frac{\partial^2 v}{\partial x\partial \theta}-\nu R\frac{\partial w}{\partial x}$$
$$+(1-\nu^2)\frac{D}{EhR^2}\left[\begin{array}{c}\frac{1}{2}(1-\nu)\frac{\partial^2 u}{\partial \theta^2}+R^3\frac{\partial^3 w}{\partial x^3}\\ -\frac{R}{2}(1-\nu)\frac{\partial^3 w}{\partial x\partial \theta^2}\end{array}\right]$$
$$=\frac{\rho h}{Eh}(1-\nu^2)R^2\frac{\partial^2 u}{\partial t^2},$$

$$\frac{R}{2}(1+v)\frac{\partial^2 u}{\partial x\partial \theta}+\frac{R^2}{2}(1-v)\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial \theta^2}-\frac{\partial w}{\partial \theta}$$
$$+(1-\nu^2)\frac{D}{EhR^2}\left[\frac{3R^2}{2}(1-v)\frac{\partial^2 v}{\partial x^2}+\frac{R^2}{2}(3-v)\frac{\partial^3 w}{\partial x^2\partial \theta}\right]$$
$$=\frac{\rho h}{Eh}(1-v^2)R^2\frac{\partial^2 v}{\partial t^2}, \tag{3}$$

$$\nu R\frac{\partial u}{\partial x}+\frac{\partial v}{\partial \theta}-w-(1-\nu^2)\frac{D}{Eh}R^2\cdot\nabla^4 w$$
$$+(1-\nu^2)\frac{D}{EhR^2}\left[\begin{array}{c}-R^3\frac{\partial^3 u}{\partial x^3}+\frac{R}{2}(1-\nu)\frac{\partial^3 u}{\partial x\partial \theta^2}\\ -\frac{R^2}{2}(3-\nu)\frac{\partial^3 v}{\partial x^2\partial \theta}-w-2\frac{\partial^2 w}{\partial \theta^2}\end{array}\right]$$
$$=\frac{\rho h}{Eh}(1-\nu^2)R^2\frac{\partial^2 w}{\partial t^2}.$$

In order to apply Flugge Eqs. (3) to CNTs, it is sufficient to know the bending stiffness D, the in-plane stiffness Eh, the mass density per unit lateral area ρh, and Piosson's ratio ν of CNTs. In particular, these parameters are not dependent of the definition of the thickness h, [3,4]. Here, it is emphasized that, even for static problems without any pre-stresses (for which assumptions (3) and (4) are valid), Eq. (2) is not exactly the same as the exact Eqs. (3) because some powers of $h^2/12R^2$ (or $(1-\nu^2)D/Eh\cdot 1/R^2$ when D is a material parameter independent of the thickness h) have been neglected in (2) compared to unity. To our knowledge, no detailed comparison has been made between the simplified Flugge Eq. (2) (model II) and the exact Flugge equations (model III) for buckling and dynamic problems. In particular, in view of unusual geometric and materials characteristics of CNTs, it is relevant to compare the simplified shell models (1) and (2) with exact Flugge Eqs. (3) and available experiments or molecular dynamics simulations of CNTs.

3 Static Buckling of Carbon Nanotubes

First, let us examine static buckling of SWNTs. Since agreements between Donnell shell model (I) and available experiments or molecular dynamics simulations have been demonstrated previously for various static buckling problems of CNTs, [3,4,8,9], we shall focus on comparison between Donnell model (I), simplified Flugge model (II), and exact Flugge model (III). Throughout the paper, we shall consider simply supported boundary conditions for cylindrical shells given by

$$w=0, \quad v=0, \quad N_x=0 \quad \text{and} \quad M_x=0. \tag{4}$$

Thus, for all three models I, II, and III, the buckling modes are given by

$$w(x,\theta)=A\sin\frac{m\pi x}{L}\cos n\theta$$
$$u(x,\theta)=B\cos\frac{m\pi x}{L}\cos n\theta \tag{5}$$
$$v(x,\theta)=C\sin\frac{m\pi x}{L}\sin n\theta$$

where A, B, and C are some constants representing amplitudes of radial, axial and circumferential displacements, respectively, m is the axial half-wave number, and n is the circumferential wave number. In particular, $n=0$ represents axisymmetric mode. Substitution of (5) into Eqs. (1), (2), or (3) leads to homogeneous equations for the coefficient A or A, B, and C. Thus, eigenequation derived by existence condition of nonzero solution determines the buckling strain, as function of the wave numbers m and n. Finally, the critical strain for buckling is decided as the minimum buckling strain. Here, strictly speaking, m must be nonzero integer for simply supported end conditions. However, it is easily understood that buckling of axially uniform modes, which do not strictly meet simply supported end conditions, can be studied by the eigenequation based on the modes (5) by taking $m=0$.

3.1 Axial Compression. Let us first discuss buckling of SWNT (of typical diameter 1.3 nm) under axial compression. It is shown, [4], that the Donnell Eq. (1) (model I) gives good estimate of the critical stress, while the numbers m and n cannot be determined uniquely by Donnell model (I). On the other hand, the predicted axial wavelength of buckling mode, based on an empirical assumption, [4], that the axial wavelength is equal to the circumferential wavelength, is found to be in good agreement with available molecular dynamics simulations. It is well known, [12], that, because linear theories of shell buckling admit a large number of different buckling modes which correspond to almost the same buckling stresses, they usually cannot predict the actual buckling mode without aid of any empirical assumption like that mentioned above. Thus, one cannot expect that simplified or exact Flugge model could give an accurate theoretical prediction for the wavelengths of buckling mode of SWNTs without any empirical assumption. Therefore, our focus here is to compare three shell models and examine whether simplified Flugge Eq. (2) is in better agreement with exact Flugge equations than Donnell Eq. (1).

Because two key parameters for buckling modes are the circumferential wave number n and the dimensionless axial wavelength (normalized by the diameter $2R$) $L/(Rm)$, the dependency of the buckling strain on $(L/(Rm),n)$ is shown in Figs. 1–3 for the model I, II, and III, respectively. It is seen from Figs. 1–3 that although all three models give similar results for n larger than 4 or $L/(Rm)$ below unity, simplified Flugge model (II) is in much better agreement with exact Flugge model (III) than Donnell model (I) for $n=1$, 2 or 3 and $L/(R\cdot m)$ larger than unity. In fact, it is found from the data shown in Figs. 1–3 that the relative errors of the simplified Flugge model (II) is about less than 10% for n larger than 2 or $L/(Rm)$ below 2, while the relative error of Donnell model (I) is about less than 10% only for n larger than 4 or $L/(Rm)$ below 1. The critical strain given by the three models for several typical cases is shown in Table 1. In particular, because long SWNTs of larger aspect ratio exhibit beam-like buckling ($n=1$), the critical strain given by elastic beam-model (Euler formula) is also shown in Table 1 for larger values of $L/(Rm)$. It is seen from Table 1 that the simplified Flugge model (II) has much better accuracy than Donnell model (I) especially when $n=2$ or 3 which corresponds to the minimum buckling strain in many important cases. In addition, simplified Flugge model (II) is a better approximate model than Donnell model (I) for very long shells for which the buckling mode corresponds to $n=1$ (beam-mode). Indeed, in this case, Donnell Eq. (1) lead to an error in the order of magnitude, while simplified Flugge Eq. (2) differs than exact Flugge model (III) or the beam-model by a factor of two. This

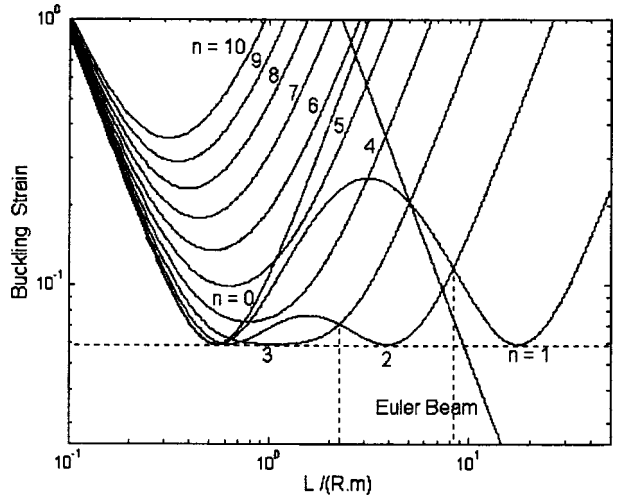


Fig. 1 The buckling strain given by Donnell model (I) for the SWNT of radius 0.65 nm under axial compression

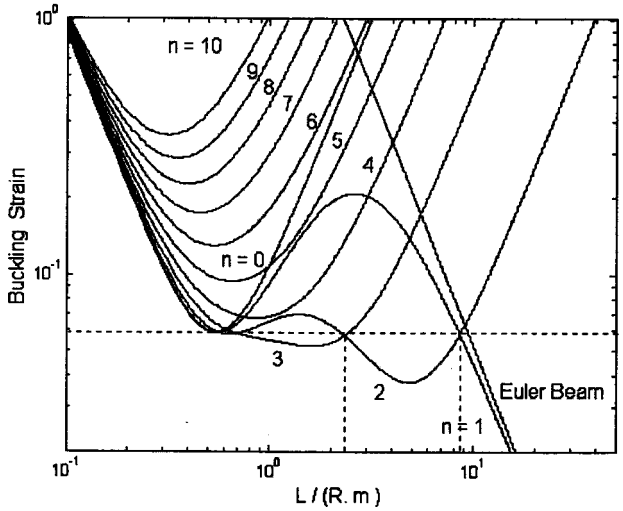


Fig. 3 The buckling strain given by exact Flugge model (III) for the SWNT of radius 0.65 nm under axial compression

conclusion for very long SWNTs shows that simplified Flugge model (II) is significantly better than Donnell model (I) even in extreme cases.

3.2 Radial Pressure. It has been shown, [8], that the Donnell model (I) gives satisfactory estimate of the critical pressure of CNTs. Here, our goal is to compare three shell models and examine whether simplified Flugge Eq. (2) is in better agreement with exact Flugge equations than Donnell Eq. (1). The dependency of buckling pressure on $(L/(Rm),n)$ (where, by the definition of buckling under radial pressure, n is not smaller than 2) is qualitatively similar for all three models I, II and III, and thus is shown only for the exact Flugge model (III) in Fig. 4. In addition, the relative errors of the model I and II compared to the exact model III are shown in Figs. 5 and 6, respectively. Particularly, the critical pressure given by three models (I, II and III) are shown in Table 2 for several typical cases. It is seen from Figs. 4–6 and Table 2 that simplified Flugge model gives much small relative errors than Donnell model for $n=2$ and $L/(Rm)$ larger than one, while relative errors of both models (I and II) are not much larger than 10% for n larger than 2. On one hand, for $L/(Rm)$ smaller than one, the minimum buckling pressure is decided by $n>2$ (Fig. 4) in which both model (I) and model (II) are applicable with relative errors not much larger than 10%. On other hand, for larger $L/(Rm)$ (it is the case when SWNTs are long or moderately long), the minimum buckling pressure is decided by $n=2$ and thus the simplified Flugge model (II) is much better than Donnell Model (I). Hence, it is concluded that, in any case, simplified Flugge model (II) is a better approximate model than Donnell model (I) for buckling of CNTs under radial pressure.

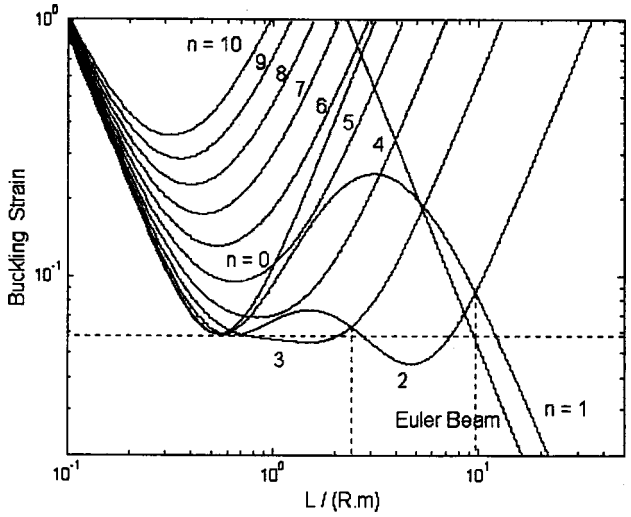


Fig. 2 The buckling strain given by simplified Flugge model (II) for the SWNT of radius 0.65 nm under axial compression

4 Free Vibration of Carbon Nanotubes

Free vibration of CNTs is a topic of major concern, largely because of the usefulness of frequency analysis to Raman spectra of CNTs, [25–28]. Here, we are particularly interested in comparison between resonant frequencies of radial breathing modes of CNTs predicted by elastic shell models and the available data obtained by experiments or molecular dynamics simulations, [25–28].

4.1 Radial Breathing Modes. Because of complicated vibration spectrum, existing data on vibration of CNTs have mainly focused on radial breathing modes of SWNTs and MWNTs. This is due to the fact that the radial breathing modes exhibit strong resonant characteristics in Raman spectra, and the frequency of radial breathing modes of SWNTs is simply proportional to the inversed radius. These features make radial breathing modes a very useful probe for structure and properties of CNTs, [25–28]. Here, our goal is to examine applicability of elastic shell models for vibration of CNTs with comparison to available data on radial breathing modes, and also to examine the accuracy of the van der Waals intertube interaction coefficient suggested in our previous work, [5,8,9].

For radial breathing modes, the radial deflection $w(t)$ is spatially uniform and thus independent of x and θ. In this special case, an uncoupled equation can be detained for the radial deflection $w(t)$ from all three models I, II, and III. In fact, the third equation of Donnell equations (see Eqs. (8.13c) of [15]), or the third equation of Flugge Eq. (3) gives almost the same uncoupled equation for $w(t)$ as

$$\frac{w}{R^2}+\frac{\rho h}{Eh}(1-\nu^2)\frac{\partial^2 w}{\partial t^2}=0 \tag{6}$$

where a small term $h^2/12R^2$ (or $(1-\nu^2)\cdot D/Eh\cdot 1/R^2$ when D is a material parameter independent of the thickness h) has been neglected compared to unity. For radial breathing mode of SWNTs, Eq. (6) gives the breathing frequency

Table 1 The comparison of minimum buckling strains, predicted by different models for the SWNT of radius 0.65 nm under axial compression

Model / $L/(Rm)$	Shell Model						Beam Model
	Model I		Model II		Model III		
0.7	$n=2$	0.0602		0.0592		0.0585	
1	$n=3$	0.0597	$n=3$	0.0565	$n=3$	0.0551	
2		0.0661		0.0574		0.0538	
3		0.0632		0.0552		0.0478	
4		0.0596		0.0470		0.0395	
5	$n=2$	0.0639	$n=2$	0.0455	$n=2$	0.0377	
6		0.0742		0.0488		0.0400	
7		0.0892		0.0556		0.0453	
8		0.1083		0.0652		0.0529	
9		0.1060		0.0969		0.0540	0.0651
10	$n=1$	0.0926	$n=1$	0.0818	$n=1$	0.0447	0.0527
20		0.0613		0.0235		0.0120	0.0132
50		0.2305		0.0039		0.0020	0.0021

$$f=230\ \ \mathrm{cm}^{-1}\,(\mathrm{nm}/2R) \tag{7}$$

which is in good agreement with experimental result $f=224\,\mathrm{cm}^{-1}\,(\mathrm{nm}/2R)$ [25,26], or Mahan's three-dimensional elasticity result $f=227\,\mathrm{cm}^{-1}\,(\mathrm{nm}/2R)$, [28]. In this case, it is noticed that bending stiffness of SWNTs does not appear because radial breathing vibration does not involve bending deformation.

Let us apply Eq. (6) to radial breathing modes of MWNTs and compare to Popov et al. results obtained by molecular dynamics simulations, [27]. Since all nested tubes of a MWNT are originally concentric and initial interlayer spacing is equal or very close to the equilibrium spacing, the van der Waals interaction pressure between any two adjacent tubes of a MWNTs is negligible prior to vibration. When radial vibration occurs, the interlayer spacing changes, and the van der Waals interaction pressure (per unit lateral area) at any point between any two adjacent tubes depends linearly on the difference of their radial deflections at that point (Δw). Thus, intertube pressure can be calculated as follows:

$$p=c\Delta w \tag{8}$$

where c is vdW interaction coefficient. In our previous work, [5,8,9], a value of c is suggested as follows:

$$c_0=\frac{320\times \mathrm{erg/cm}^2}{0.16d^2}\quad (d=1.42\times 10^{-8}\ \ \mathrm{cm}). \tag{9}$$

Here the curvature effect and the dependency on radius are neglected and thus c_0 given by (9) has been used for SWNTs of various radii.

Consider DWNTs and three-wall CNTs. The results based on the elastic model (6), (8), (9) are shown in Figs. 7 and 8, with comparison to Popov et al.'s results obtained by molecular dynamics simulations, [27]. It is seen that the elastic shell model (6), (8), (9) is in excellent agreement with molecular dynamics simulations of radial breathing modes for DWNTs and three-wall CNTs, with relative errors less than 5%. Further, radial breathing frequencies of MWNTs of innermost radii 1 nm are shown in Fig. 9 for various number of nested layers up to 10. Comparison between the present Fig. 9 and Popov et al.'s results based on a continuum model, [27], shows that the maximum of relative errors is less than 25%. Since no details are given for Popov et al. continuum model, [27], we cannot comment on the difference between their results and the present results.

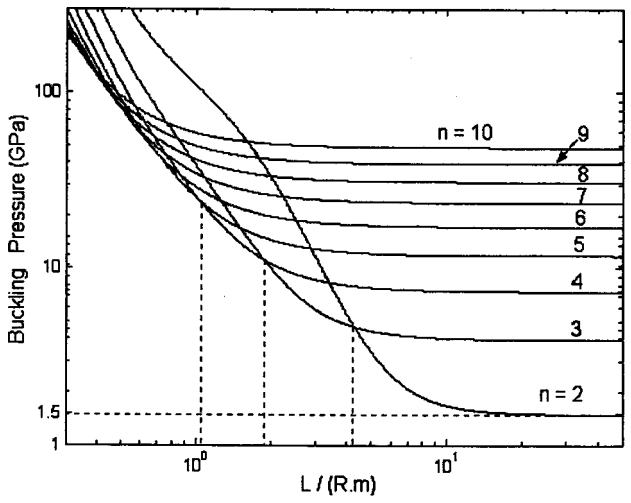


Fig. 4 The buckling pressure given by exact Flugge model (III) for the SWNT of radius 0.65 nm under radial pressure

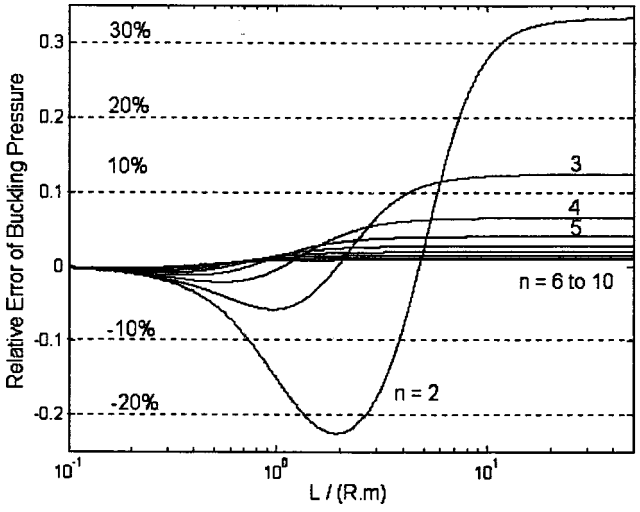


Fig. 5 The relative error of buckling pressure predicted by Donnell model (I) for the SWNT of radius 0.65 nm under radial pressure

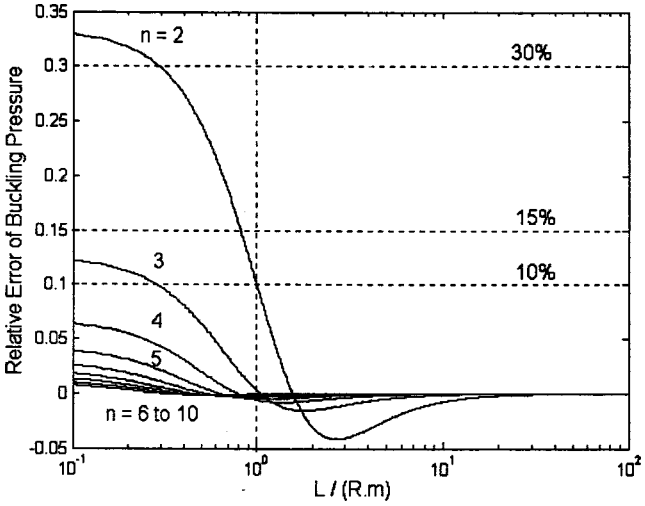


Fig. 6 The relative error of buckling pressure predicted by simplified Flugge model (II) for the SWNT of radius 0.65 nm under radial pressure

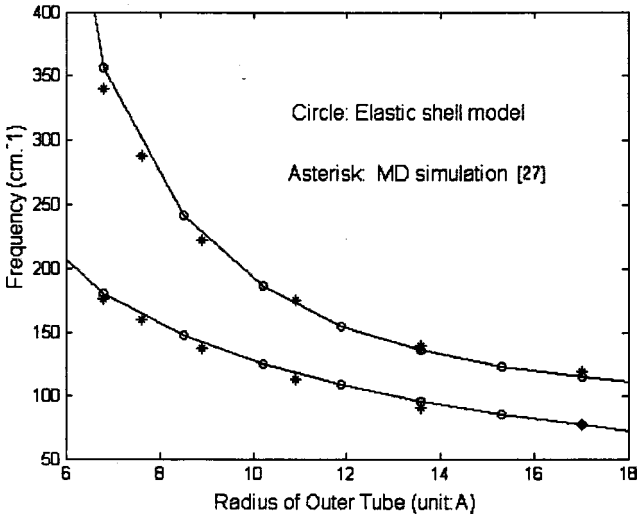


Fig. 7 Frequencies of radial breathing vibration predicted by elastic shell models and MD simulation for DWNTs of various outer radii

Finally, to examine the coefficient c_0 given by (9), we study the sensitiveness of radial breathing frequencies on the coefficient c (8). Since breathing frequencies are very high for small-radius SWNTs, the intertube interaction has little influence on the breathing frequencies of DWNTs or three-wall CNTs of small radii (such as $R_1 = 0.34$ nm). Thus, the frequencies of DWNTs or three-wall CNTs are insensitive to the coefficient c for small-radius DWNTs or three-wall CNTs, and comparison with molecular dynamics simulations of small-radius CNTs cannot identify the best value of the coefficient c. Here, some results are shown in Fig. 10 for a DWMT and a three-wall CNT of relatively large innermost radius 1.36 nm and 1.02 nm, respectively. Because the lowest frequency is insensitive to the value of c, only the dependency of higher frequencies is shown in Fig. 10. It is seen from Fig. 10 that a relative error of 20% occurs when the coefficient c_0 changes by a factor of 2. Therefore, the excellent agreement given in Figs. 7 and 8 between the model based on the coefficient c_0 (9) and molecular dynamics simulations suggests that the coefficient c_0 given by (9) can be regarded as a good value at least for CNTs of larger radii (not smaller than 1.02 nm). This offers an evidence for the accuracy of the coefficient (9) which was suggested in [5] and applied to MWNTs in [8,9].

Table 2 The comparison of critical buckling pressure, predicted by different models for the SWNTs of radius 0.65 nm or 5 nm under radial pressure

Radius (nm)	Vibration Modes m	Vibration Modes n	Aspect Ratio	Critical Buckling Pressure (GPa) Model I	Model II	Model III
0.65	1	4	0.7	15.58	15.16	15.28
		3	1.5	6.29	5.84	5.90
			5	2.15	1.67	1.68
		2	10	2.00	1.13	1.51
			20	1.97	1.48	1.48
	0	2		1.97	1.48	1.48
5.00	1	7	0.7	0.077	0.077	0.076
		5	1.5	0.035	0.033	0.034
		3	5	10.90×10^{-3}	9.90×10^{-3}	9.90×10^{-3}
		2	10	5.05×10^{-3}	4.19×10^{-3}	4.20×10^{-3}
			20	4.38×10^{-3}	3.31×10^{-3}	3.30×10^{-3}
	0	2		4.32×10^{-3}	3.31×10^{-3}	3.24×10^{-3}

4.2 Free Vibration of SWNTs. Let us now clarify the range of applicability of Donnell model (I) and simplified Flugge model (II) for free vibration of CNTs, in terms of the dimensionless axial wavelength $L/(Rm)$ and the circumferential wave number n. The vibration modes of a simply supported shell ($m>0$) are given by (ω is the circular frequency)

$$w = A \sin\frac{m\pi x}{L}\cdot\cos n\theta\cdot\exp(i\omega\cdot t)$$

$$u = B\cos\frac{m\pi x}{L}\cos n\theta\cdot\exp(i\omega\cdot t) \tag{10}$$

$$v = C\sin\frac{m\pi x}{L}\cos n\theta\cdot\exp(i\omega\cdot t)$$

where m should be nonzero integer for simply supported end conditions. Substitution of (10) into Eqs. (1), (2), or (3) leads to homogeneous equations, the existence condition for nonzero solution determine resonant frequencies as function of the numbers m and n.

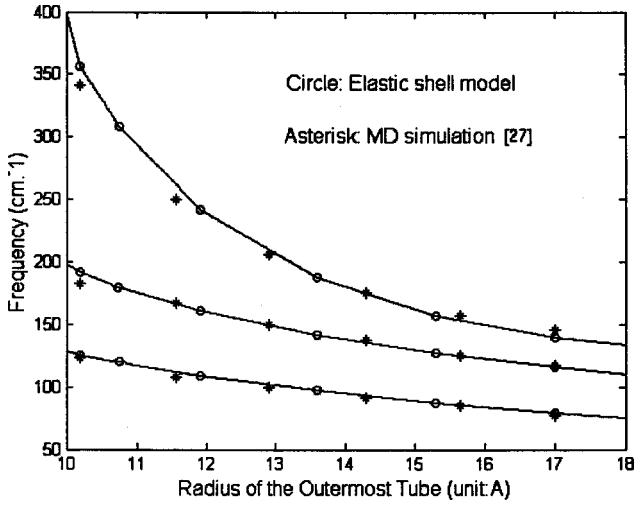


Fig. 8 Frequencies of radial breathing vibration predicted by elastic shell models and MD simulation for three-wall CNTs of various outermost radii

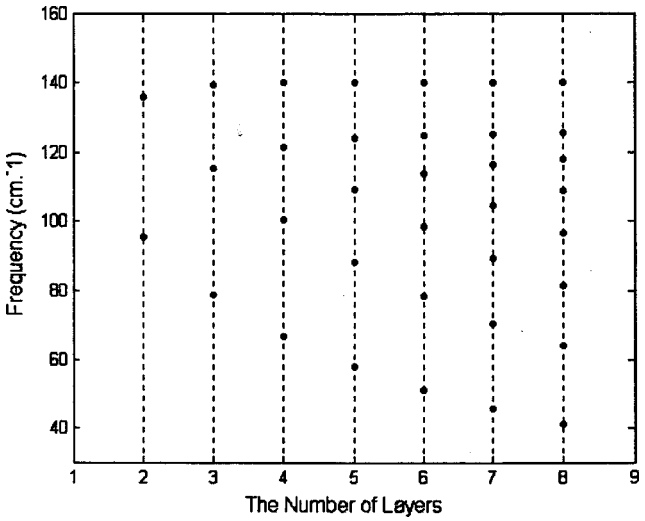


Fig. 9 Frequencies of radial breathing vibration predicted by the elastic shell models for MWNTs of the innermost radius 1 nm

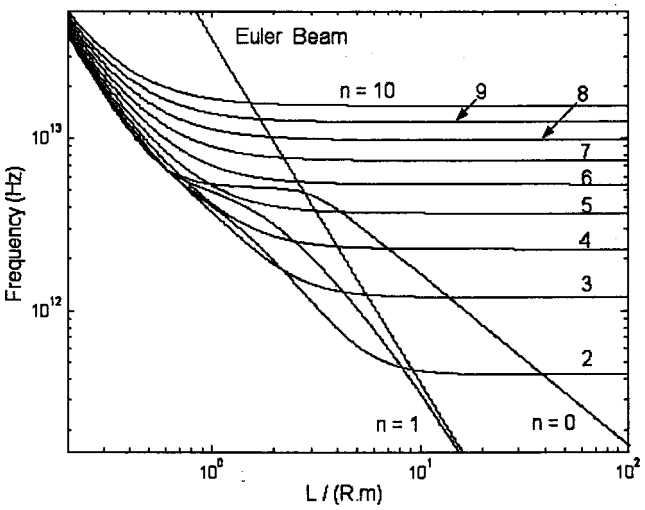


Fig. 11 The frequency of radial vibration given by exact Flugge model (III) for the SWNT of radius 0.65 nm

Here, the frequencies given by the exact model (III) are shown in Fig. 11 for the SWNT of radius $R=0.65$ nm, while the relative errors of the frequencies given by two simplified models (I) and (II) compared to the exact model (III), are shown in Figs. 12 and 13, respectively. Similar results for the radius $R=5$ nm, are given in Figs. 14–16. It is seen from Figs. 11–16 that the relative errors of the simplified Flugge model (II) are about 10% or less provided that n is larger than one or $L/(Rm)$ is smaller than two, while the relative errors of Donnell model (I) are about 10% or less provided that n is larger than 3 or $L/(Rm)$ is less than one. Hence, the simplified Flugge model (II) is a much better approximate shell model than Donnell model I. In particular, it is seen from Figs. 11 and 14 that the critical aspect ratio $(L/(2R))$ at which the vibrational mode corresponding to the minimum frequency transfers from $n=2$ to $n=1$ (beam-mode) is about 5 for SWNT of radius 0.65 nm, and is larger than 10 for SWNT of radius 5 nm. In addition, for long simply supported shells, the vibrational mode corresponding to the minimum frequency is characterized by $m=1$ and $n=1$ (beam-model). In this case, it is seen from Figs. 12–13 and Figs. 15–16 that Donnell model (I) leads to errors in the order of magnitude, while simplified Flugge model (II) is comparable to exact Flugge model (III) (with relative errors 40%). These conclusions are qualitatively consistent with those drawn for static buckling discussed in Section 3. It is emphasized that the improved accuracy of simplified Flugge model over Donnell model is significant for CNTs because the low-frequency vibrational modes often have circumferential wave number $n=2$ or 3. Therefore, these results are important for further application of simplified shell models to CNTs, especially to MWNTs of large number of layers.

Here, it should be stressed that one essential shortcoming of simplified shell models (I) and (II) is that they give only one frequency (radial vibration) for each combination (m,n), while the exact shell model (III) gives three frequencies for given m and n which represent radial, axial and circumferential vibrational modes, respectively. Therefore, the simplified shell models (I) and (II) cannot be used to discuss axial and circumferential vibration modes. In spite of this, because radial vibration is dominant in many important problems and the corresponding frequency is usually lower than the frequencies of axial and circumferential modes, the radial vibrational modes are of major concern. An

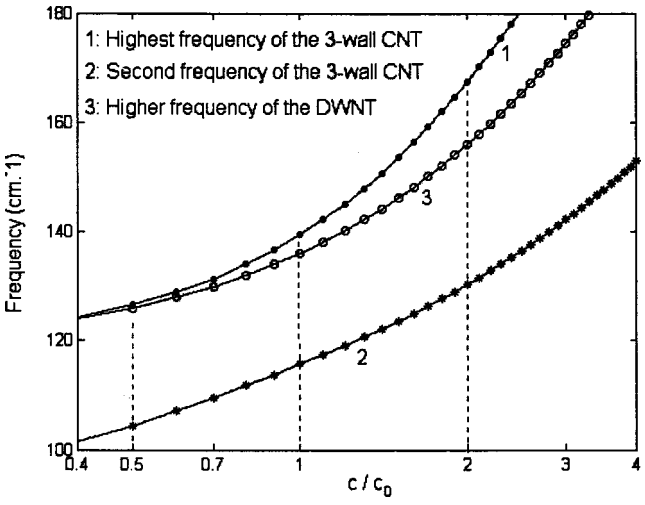


Fig. 10 Dependence of the breathing mode frequencies of a DWNT of inner radius 1.36 nm and a three-wall CNT of innermost radius 1.02 nm on the vdW interaction coefficient

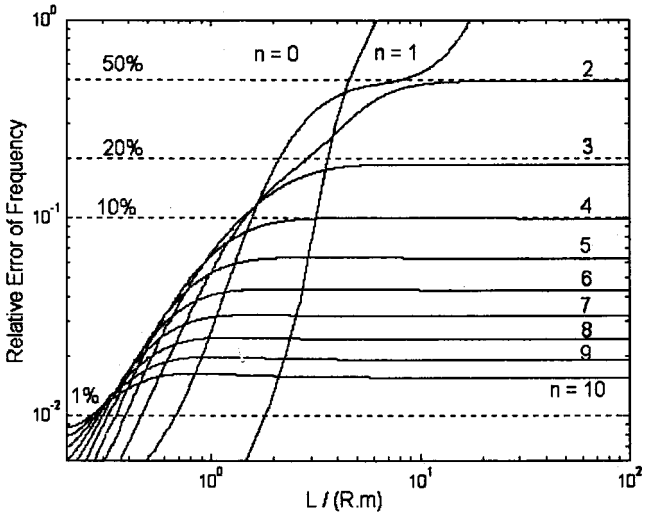


Fig. 12 The relative error of radial vibration frequency predicted by Donnell model (I) for the SWNT of radius 0.65 nm

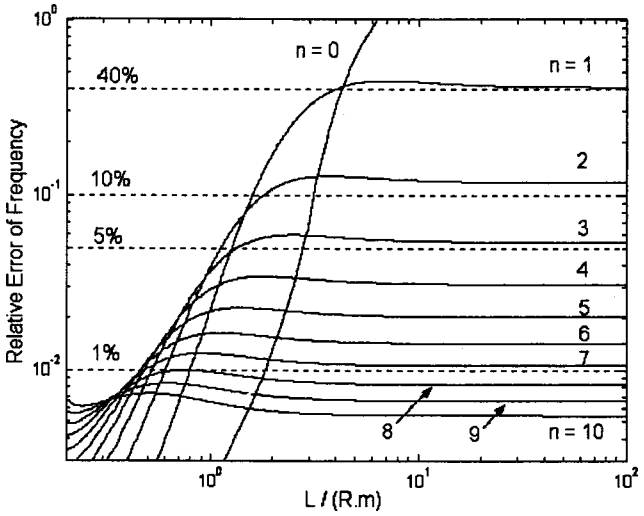


Fig. 13 The relative error of radial vibration frequency predicted by simplified Flugge model (II) for the SWNT of radius 0.65 nm

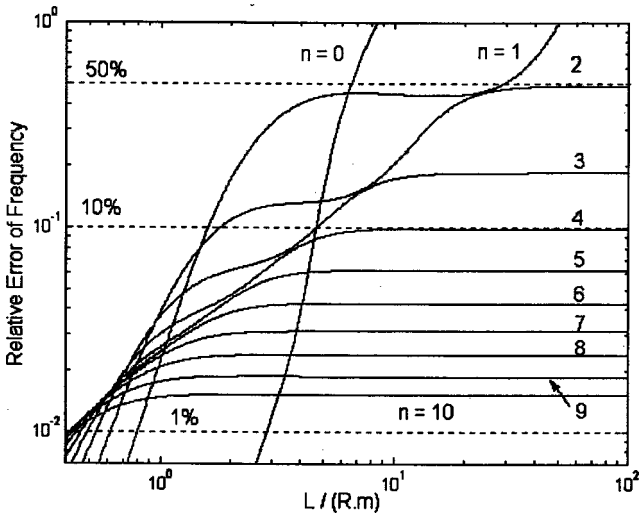


Fig. 15 The relative error of radial vibration frequency predicted by Donnell model (I) for the SWNT of radius 5 nm

example of SWNTs of radius 5 nm are given in Fig. 17 for three frequencies given by the exact model (III), and single frequency given by the simplified model (I) and (II), respectively, which indicates that the radial frequency is lower than other two (axial and circumferential) frequencies for almost all given $(L/(Rm),n)$ except $n=0$ which means axisymmetric vibrational modes.

4.3 Beam-Like Vibration. The above results indicate that large errors of simplified shell models (I) and (II) could occur when $n=1$ and $L/(Rm)$ is much larger than one or two. It is the case when low-frequency vibrational modes of long SWNTs are concerned. Here, to compare the three models in this special case, let us discuss beamlike vibration of simply supported SWNTs of larger aspect ratio whose lowest frequency corresponds to $n=1$ and $m=1$, [29–32]. All three shell models, together with the elastic beam-model, [10], are shown in Fig. 18 for SWNTs of radius 0.65 nm, as function of the aspect ratio $(L/(2R))$. In addition, detailed comparison is shown in Table 3 for several relevant cases. It is found that, for special case of beam-like vibration of CNTs, simplified Flugge model (II), in which the effect of in-plane inertia is neglected, [33,34], leads to as large as 40–50% relative errors as compared to exact Flugge model (III), while Donnell model (I) leads to errors in the order of magnitude.

4.4 Axially Uniform Vibrational Modes. The results shown above do not include the special case of axially uniform vibrational modes with $m=0$, which cannot strictly meet the simply supported end conditions. However, it is easily understood that axially uniform vibrational modes can be studied by the eigenequation based on the modes (10) by taking $m=0$. The frequencies given by the exact model (III) and the simplified model (I) and (II), are shown in Table 4 for axially uniform modes ($m=0$) of SWNTs of radius $R=0.65$ nm and 5 nm. Since the case of ($m=0$, $n=1$) represents a pure rigid-body motion, it is not included in Table 4. It is seen that the frequency for $n=0$ (radial breathing mode) is higher than all other frequencies given in Table 4. On the other hand, for $n\geqslant 2$, the frequencies shown in Table 4 monotonically increase with the circumferential wave number n.

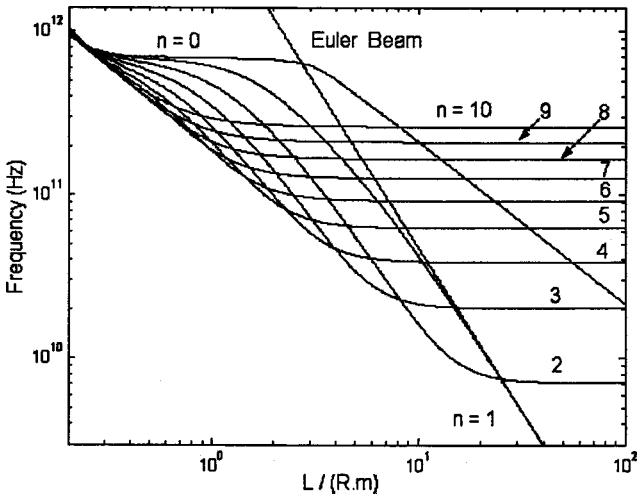


Fig. 14 The frequency of radial vibration given by exact Flugge model (III) for the SWNT of radius 5 nm

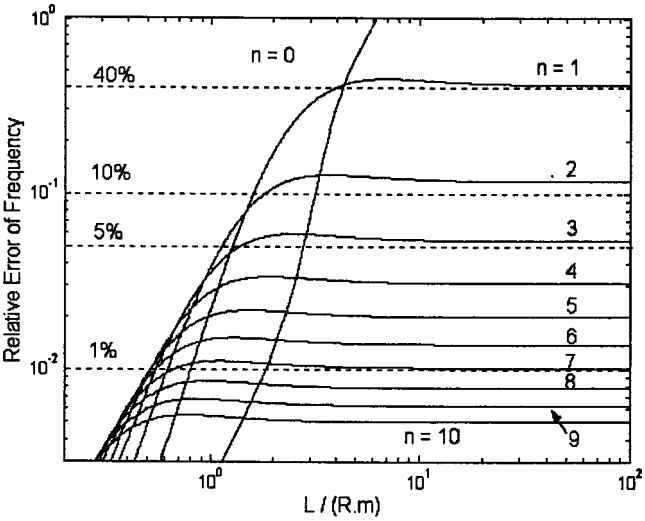


Fig. 16 The relative error of radial vibration frequency predicted by simplified Flugge model (II) for the SWNT of radius 5 nm

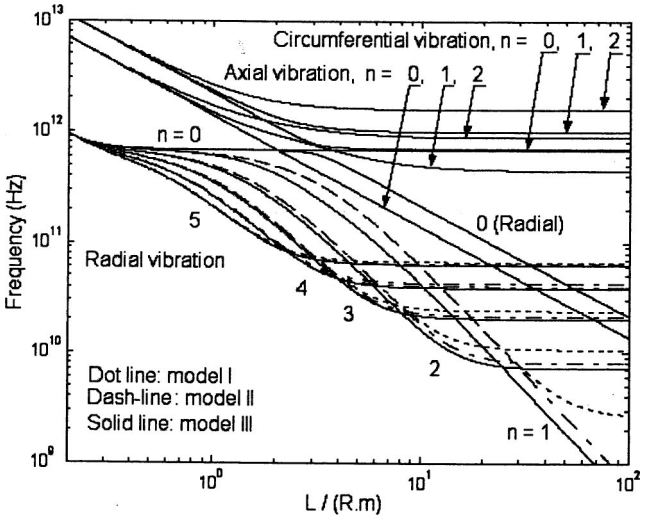


Fig. 17 The vibration frequencies given by three shell models for the SWNT of radius 5 nm

5 Conclusions

This paper examines applicability of the simplified Flugge and Donnell cylindrical shell equations for carbon nanotubes. Static buckling and free vibration of carbon nanotubes are studied using the two simplified shell models with comparison to the exact Flugge shell model. It is found that all three elastic shell models are in excellent agreement (with relative errors less than 5%) with recent molecular dynamics simulations for radial breathing modes of carbon nanotubes. For general cases of static buckling and free vibration, one major conclusion is that the relative errors of the simplified Flugge model (II) are generally less than 10% for n larger than 1 or 2, or for $L/(Rm)$ smaller than two, while the relative errors of Donnell model (I) are less than 10% only for n larger than 3 or 4, or for $L/(R.m)$ smaller than one. This conclusion is significant because the critical buckling mode and low-frequency vibrational modes often have circumferential wave number $n=2$ or 3. Hence, simplified Flugge model (II) has a significantly enlarged range of applicability compared to Donnell model (I), and covers almost all important cases of major concern. Almost the only missing significant case is the beam-like buckling or vibration of long CNTs. In this case, however, simple elastic beam model gives excellent approximate results as compared to the exact Flugge shell model and thus the latter is not necessarily needed. Here, we would emphasize that simplified shell models, characterized by a single uncoupled equation for radial deflection, will be particularly useful for MWNTs of large number of layers. On the other hand, such simplified shell models are applicable only for radial deformation of cylindrical shells with simpler end conditions (such as simply supported ends). In other words, when axial or circumferential deflection becomes essential or more complicated end conditions are involved, more accurate shell models (such as exact Flugge shell model) are usually needed.

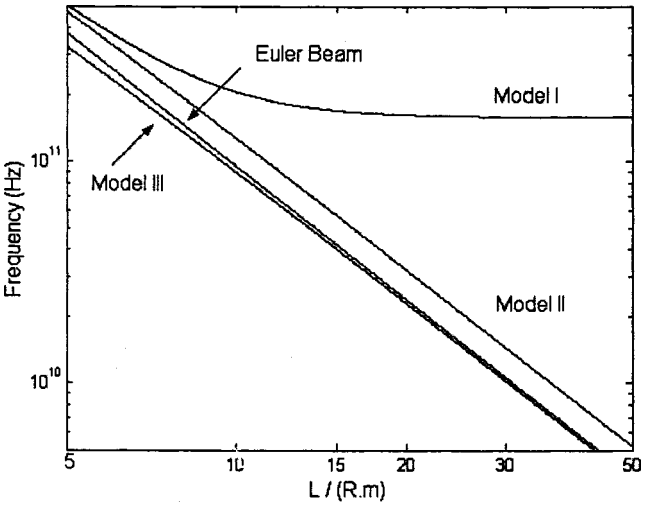


Fig. 18 Comparison of beam-like vibrational frequencies given by different models for SWNT of radius 0.65 nm

Table 3 The comparison of beam-like vibration frequencies ($n=1$), predicted by different models for SWNTs of radius 0.65 nm or 5 nm

Radius (nm)	$L/(Rm)$	Frequency (Hz) Model I	Model II	Model III	Beam model
0.65	0.5	8.21×10^{12}	8.18×10^{12}	8.16×10^{12}	
	1	5.09×10^{12}	5.08×10^{12}	4.97×10^{12}	
	5	1.51×10^{12}	1.50×10^{12}	1.04×10^{12}	1.53×10^{12}
	10	5.06×10^{11}	4.76×10^{11}	3.29×10^{11}	3.82×10^{11}
	50	1.59×10^{11}	2.08×10^{10}	1.47×10^{10}	1.53×10^{10}
	100	1.58×10^{11}	5.22×10^{9}	3.68×10^{9}	3.82×10^{9}
5	0.5	6.80×10^{11}	6.80×10^{11}	6.78×10^{11}	
	1	6.26×10^{11}	6.25×10^{11}	6.11×10^{11}	
	5	1.94×10^{11}	1.94×10^{11}	1.30×10^{11}	1.93×10^{11}
	10	6.19×10^{10}	6.18×10^{10}	4.28×10^{10}	4.81×10^{10}
	50	3.80×10^{9}	2.71×10^{9}	1.91×10^{9}	1.93×10^{9}
	100	2.80×10^{9}	6.78×10^{8}	4.79×10^{8}	4.81×10^{8}

Table 4 The comparison of vibration frequencies when $m=0$, predicted by different models for SWNTs of radius 0.65 nm or 5 nm

Radius (nm)	n	Frequency (Hz) Model I	Model II	Model III
0.65	0	5.31×10^{12}	5.31×10^{12}	5.30×10^{12}
	2	6.31×10^{11}	4.73×10^{11}	4.23×10^{11}
	3	1.42×10^{12}	1.26×10^{12}	1.20×10^{12}
	4	2.52×10^{12}	2.37×10^{12}	2.29×10^{12}
	5	3.94×10^{12}	3.78×10^{12}	3.71×10^{12}
5.00	0	6.90×10^{11}	6.90×10^{11}	6.90×10^{11}
	2	1.07×10^{10}	7.99×10^{9}	7.15×10^{9}
	3	2.40×10^{10}	2.13×10^{10}	2.02×10^{10}
	4	4.26×10^{10}	4.00×10^{10}	3.88×10^{10}
	5	6.66×10^{10}	6.39×10^{10}	6.27×10^{10}

Acknowledgment

The financial support of the Natural Science and Engineering Research Council of Canada (NSERC) is gratefully acknowledged.

References

[1] Qian, D., Wagner, G. J., Liu, W. K., Yu, M. F., and Ruoff, R. S., 2002, "Mechanics of Carbon Nanotubes," Appl. Mech. Rev., **55**, pp. 495–533.
[2] Huang, Y., and Wang, Z. L., 2003, "Mechanics of Nanotubes," *Comprehensive Structural Integrity Handbook*, Elsevier, New York, **8**, pp. 551–579.
[3] Yakobson, B. I., Brabec, C. J., and Bernholc, J., 1996, "Nanomechanics of Carbon Tubes: Instability Beyond Linear Response," Phys. Rev. Lett., **76**, pp. 2511–2514.
[4] Ru, C. Q., 2000, "Effective Bending Stiffness of Carbon Nanotubes," Phys. Rev. B, **62**, pp. 9973–9976.

[5] Ru, C. Q., 2001, ''Degraded Axial Buckling Strain of Multiwalled Carbon Nanotubes Due to Interlayer Slips,'' J. Appl. Phys., **89**, pp. 3426–3433.
[6] 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.
[7] Ru, C. Q., 2000, ''Elastic Buckling of Singlewalled Carbon Nanotube Ropes Under High Pressure,'' Phys. Rev. B, **62**, pp. 10405–10408.
[8] Wang, C. Y., Ru, C. Q., and Miduchowski, A., 2003, ''Elastic Buckling of Multiwall Carbon Nanotubes Under High Pressure,'' J. Nanosci. Nanotech. (special issue on Nanotubes), **3**, pp. 199–208.
[9] Wang, C. Y., Ru, C. Q., and Miduchowski, A., 2003, ''Axially Compressed Buckling of Pressured Multiwall Carbon Nanotubes,'' Int. J. Solids Struct., **40**, pp. 3893–3911.
[10] Zhang, P., Huang, Y., Gao, H., and Hwang, K. C., 2002, ''Fracture Nucleation in Single-Wall Carbon Nanotubes Under Tension,'' ASME J. Appl. Mech., **69**, pp. 454–458.
[11] Belytschko, T., Xiao, S. P., Schatz, G. C., and Ruoff, R. S., 2002, ''Atomistic Simulations of Nanotube Fracture,'' Phys. Rev. B, **65**, p. 235430.
[12] Donnell, L. H., 1976, *Beams, Plates and Shells*, McGraw-Hill, New York.
[13] Flugge, W., 1960, *Stresses in Shells*, Springer-Verlag, Berlin.
[14] Timoshenko, S. P., and Gere, J. M., 1961, *Theory of Elastic Stability*, McGraw-Hill, New York.
[15] Kraus, H., 1967, *Thin Elastic Shells: An Introduction to the Theoretical Foundations and the Analysis of Their Static and Dynamic Behavior*, John Wiley and Sons, New York.
[16] Markus, S., 1988, *The Mechanics of Vibration of Cylindrical Shells*, Elsevier, Amsterdam.
[17] Soedel, W., 1993, *Vibrations of Shells and Plates*, Marcel Dekker, New York.
[18] Kempner, J., 1955, ''Remarks on Donnell's Equation,'' ASME J. Appl. Mech., **77**, pp. 117–118.
[19] Hoff, N. J., 1955, ''The Accuracy of Donnell's Equation,'' ASME J. Appl. Mech., **22**, pp. 329–334.
[20] Morley, L. S. D., 1959, ''An Improvement on Donnell's Approximation for Thin-Walled Circular Cylinders,'' Q. J. Mech. Appl. Math., **XII**, pp. 88–99.
[21] Weaver, D. S., and Unny, T. E., 1973, ''On the Dynamic Stability of Fluid-Conveying Pipes,'' ASME J. Appl. Mech., **40**, pp. 49–52.
[22] Dym, C. L., 1973, ''On the Buckling of Cylinders in Axial Compression,'' ASME J. Appl. Mech., **40**, pp. 565–568.
[23] Li, C., and Chou, T. W., 2003, ''A Structural Mechanics Approach for the Analysis of Carbon Nanotubes,'' Int. J. Solids Struct., **40**, pp. 2487–2499.
[24] Chang, T., and Gao, H., 2003, ''Size-Dependence Elastic Properties of a Single-Wall Carbon Nanotube via a Molecular Mechanics Model,'' J. Mech. Phys. Solids, **51**, pp. 1059–1074.
[25] Rao, A. M., Richter, E., Bandow, S., Chase, B., Eklund, P. C., Williams, K. A., Fang, S., Subbaswamy, K. R., Menon, M., Thess, A., Smalley, R. E., Dresselhaus, G., and Dresselhaus, M. S., 1997, ''Diameter-Selective Raman Scattering From Vibrational Modes in Carbon Nanotubes,'' Science, **275**, pp. 187–191.
[26] Bandow, S., and Asaka, S., 1998, ''Effect of the Growth Temperature on the Diameter Distribution and Chirality of Single-Wall Carbon Nanotubes,'' Phys. Rev. Lett., **80**, pp. 3779–3782.
[27] Popov, V. N., and Henrard, L., 2002, ''Breathinglike Phonon Modes of Multiwalled Carbon Nanotubes,'' Phys. Rev. B, **65**, p. 235415.
[28] Mahan, G. D., 2002, ''Oscillations of a Thin Hollow Cylinder: Carbon Nanotube,'' Phys. Rev. B, **65**, p. 235402.
[29] Greenspon, J. E., 1960, ''Vibrations of a Thick-Walled Cylindrical Shell—Comparison of the Exact Theory With Approximate Theories,'' J. Acoust. Soc. Am., **32**, pp. 571–578.
[30] Forsberg, K., 1969, ''Axisymmetric and Beam-Type Vibrations of Thin Cylindrical Shell,'' AIAA J., **7**(2), pp. 221–227.
[31] Yoon, J., Ru, C. Q., and Miduchowski, A., 2002, ''Non-coaxial Resonance of an Isolated Multiwall Carbon Nanotube,'' Phys. Rev. B, **66**, p. 233402.
[32] Yoon, J., Ru, C. Q., and Miduchowski, A., 2003, ''Vibration of Embedded Multiwall Carbon Nanotubes,'' Compos. Sci. Technol., **63**, pp. 1533–1542.
[33] Shayo, L. K., and Ellen, C. H., 1978, ''Theoretical Studies of Internal Flow-Induced Insensitivities of Cantilever Pipes,'' J. Sound Vib., **56**, pp. 463–474.
[34] El-Raheb, M., and Babcock, C. D., 1981, ''Some Approximations in the Linear Dynamic Equations of Thin Cylinders,'' J. Sound Vib., **76**, pp. 543–559.

D. Q. Cao
e-mail: caojdq@yahoo.com.cn

Y. M. Ge[1]
e-mail: ymeige@yahoo.com.cn

Y. R. Yang
e-mail: yangyir@sohu.com

Department of Applied Mechanics and Engineering,
Southwest Jiaotong University,
Chengdu 610031, P.R. China

Stability Criteria for Nonclassically Damped Systems With Nonlinear Uncertainties

The asymptotic stability of nonclassically damped systems with nonlinear uncertainties is addressed using the Lyapunov approach. Bounds on nonlinear perturbations that maintain the stability of an asymptotically stable, linear multi-degree-of-freedom system with nonclassical damping are derived. The explicit nature of the construction permits us to directly express the algebraic criteria in terms of plant parameters. The results are then applied to the symmetric output feedback control of multi-degree-of-freedom systems with nonlinear uncertainties. Numerical examples are given to demonstrate the new stability criteria and to compare them with the previous results in the literature.
[DOI: 10.1115/1.1778719]

1 Introduction

Dynamic analysis of structures is commonly accomplished assuming deterministic behavior of model parameters and loads. Mathematical modeling of physical dynamic systems in many engineering problems, however, includes some degree of uncertainties due to structural parameter variations, unmodeled dynamics and control input constraints. Each of these constraints can lead to severe degradation in performance and even instability. Hence, the problem of maintaining the stability of a nominal stable system subjected to linear and/or nonlinear perturbations has been an active area of research for some time. To obtain stability measures of linear state-space systems with unstructured and/or structured uncertainties, the Lyapunov stability theory is utilized in the literature (see, for example, Yedavalli [1], Zhou and Khargonekar [2], Siljak [3], Bien and Kim [4], and the literature cited therein).

Stability measures of second-order systems have been relatively scarce compared to those in the first-order form, even for nominal cases. A good account of stability conditions for nominal second-order system is presented in Shieh et al. [5]. Stability robustness bounds on unstructured perturbations of second-order systems are presented using the Lyapunov function approach in Hsu and Wu [6]. Robust stability bounds on structured perturbations and dependent parametric perturbations are proposed in Cao and Shu [7]. A design procedure is carried out in Diwekar and Yedavalli [8] for stabilizing the matrix second-order systems with variations in inertia, damping and stiffness matrices. In 1997, Cox and Moro [9] studied the stability of a class of nonlinear dynamic systems whose linear part is almost classically damped and proposed a stability criterion that bounds the degree of the uncertain nonlinearity and deviation from classical damping. According to Cox and Moro [9], a linear second-order system is classically damped if the damping and stiffness matrices of the system are commuting ones. Although the Rayleigh damping models or other classical damping strategies (e.g., Cox and Moro [9]) are commonly used in the stability analysis due to their simplicity, they may not generally apply to real structures.

In this paper, we are concerned with the problem of the stability robustness of nonclassically damped systems with nonlinear uncertainties. By using a specific Lyapunov function, bounds on nonlinear perturbations that maintain the stability of an asymptotically stable, linear system with nonclassical damping are derived and directly expressed in terms of plant matrices. The main stability criterion obtained is shown to be less conservative than the criterion for the stability of almost classically damped system, proposed by Cox and Moro [9]. In order to stabilize the second-order system with nonlinear uncertainties a symmetric output feedback structure is introduced for the control law. As a consequence of the structure of the control law, the proposed approach can be employed to guarantee the global stability of the closed-loop system in the Lyapunov sense. Simple examples are studied for demonstrating the merit of the stability measures and to compare them with the previous results in the literature.

The following notation will be used throughout this paper.

R^n	=	n-dimensional Euclidean space
$R^{n\times m}$	=	$n\times m$ real matrix
$\lambda_j(A)$	=	j-th eigenvalue of square matrix A
$\lambda_M(P)(\lambda_m(P))$	=	maximum (minimum) eigenvalue of symmetric matrix P
I	=	$n\times n$ identity matrix
$\|x\|$	=	Euclidean norm of vector x
$\|A\|$	=	spectral norm of matrix A
$A>0$ $(A<0)$	=	positive (negative) definite matrix

2 Main Results

Consider a differential equation governing the motion of an n-degree-of-freedom system

$$M\ddot{x}(t)+D\dot{x}(t)+Kx(t)=f(x(t),\dot{x}(t),t) \tag{1}$$

where $x(t)\in R^n$ is configuration vector, $f\in R^n$ is a smooth function of x, $\dot{x}$ and t; M, D and $K\in R^{n\times n}$ are symmetric positive definite mass, damping and stiffness matrices, respectively. It is assumed that there exist positive constants ε_1 and ε_2 such that the nonlinear forcing function $f(x,\dot{x},t)$ satisfies

$$\|f(x,\dot{x},t)\|\leqslant\varepsilon_1\|x\|+\varepsilon_2\|\dot{x}\|, \quad \text{for all } t\geqslant t_0. \tag{2}$$

The state-space expression of system (1) can be written as

$$\dot{y}(t)=Ay(t)+F(y(t),t), \tag{3}$$

where

$$y(t)=\begin{bmatrix}x(t)\\ \dot{x}(t)\end{bmatrix}, \quad A=\begin{bmatrix}0 & I\\ -M^{-1}K & -M^{-1}D\end{bmatrix},$$

[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, July 7, 2003; final revision, February 2, 2004. Associate Editor: M. P. Mignolet. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

$$F(y(t),t)=\begin{bmatrix}0\\ M^{-1}f(x(t),\dot{x}(t),t)\end{bmatrix}. \tag{4}$$

The following lemma is well known (Skelton et al. [10]) and will be used in the proof of our main results.

Lemma 1. (Skelton et al. [10]): Let a symmetric matrix P be partitioned as

$$P=\begin{bmatrix}P_{11} & P_{12}\\ P_{12}^T & P_{22}\end{bmatrix}.$$

Then the following statements are equivalent:

(i) $P>0$;

(ii) $P_{11}>0$ and $P_{22}-P_{12}^TP_{11}^{-1}P_{12}>0$;

(iii) $P_{22}>0$ and $P_{11}-P_{12}P_{22}^{-1}P_{12}^T>0$.

THEOREM 1. Suppose that the nonlinear perturbation f satisfies condition (2); then system (1) is globally asymptotically stable if

$$\frac{1}{4}\left(\varepsilon_2-\frac{2\mu}{\gamma}\varepsilon_1\right)^2+\mu\varepsilon_1+\frac{2\mu\kappa}{\gamma}\varepsilon_2<\mu\kappa, \tag{5}$$

where $\mu=\lambda_M(M)$, $\kappa=\lambda_m(K)$, and $\gamma=\lambda_m(D)$.

Proof. Introduce a Lyapunov function candidate $V(y(t))=y^T(t)Py(t)$, where

$$P=\begin{bmatrix}D+\frac{2\mu}{\gamma}K & M\\ M & \frac{2\mu}{\gamma}M\end{bmatrix}. \tag{6}$$

Since M, D and $K\in R^{n\times n}$ are positive definite matrices, it follows from Lemma 1 that the symmetric matrix P is positive definite if and only if the symmetric matrix

$$D+\frac{2\mu}{\gamma}K-\frac{\gamma}{2\mu}M>0. \tag{7}$$

Taking notice of

$$\frac{\gamma}{2\mu}\lambda_M(M)=\frac{\gamma}{2}<\lambda_m(D)\leqslant\lambda_m\left(D+\frac{2\mu}{\gamma}K\right), \tag{8}$$

we see that (7) holds and therefore $P>0$. Taking the time derivative of $V(y(t))$ along the solution of system (3) yields

$$\begin{aligned}\dot{V}(y(t))&=\dot{y}^T(t)Py(t)+y^T(t)P\dot{y}(t)\\ &=y^T(t)(A^TP+PA)y(t)+2y^T(t)PF(y(t),t)\\ &=-2y^T(t)\begin{bmatrix}K & 0\\ 0 & \frac{2\mu}{\gamma}D-M\end{bmatrix}y(t)+2y^T(t)PF(y(t),t).\end{aligned} \tag{9}$$

Since

$$\begin{aligned}\dot{x}^T(t)\left(\frac{2\mu}{\gamma}D-M\right)\dot{x}(t)&=\frac{2\mu}{\gamma}\dot{x}^T(t)D\dot{x}(t)-\dot{x}^T(t)M\dot{x}(t)\\ &\geqslant 2\mu\|\dot{x}(t)\|^2-\lambda_M(M)\|\dot{x}(t)\|^2\\ &=\mu\|\dot{x}(t)\|^2,\end{aligned} \tag{10}$$

in terms of (2) and (4), we obtain

$$\begin{aligned}\dot{V}(y(t))&=-2x^T(t)Kx(t)-2\dot{x}^T(t)\left(\frac{2\mu}{\gamma}D-M\right)\dot{x}(t)\\ &\quad+2\left(x^T(t)+\frac{2\mu}{\gamma}\dot{x}^T(t)\right)f(x(t),\dot{x}(t),t)\\ &\leqslant-2(\kappa\|x(t)\|^2+\mu\|\dot{x}(t)\|^2)+2\left(\varepsilon_1\|x(t)\|^2\right.\\ &\quad\left.+\left(\varepsilon_2+\frac{2\mu}{\gamma}\varepsilon_1\right)\|x(t)\|\|\dot{x}(t)\|+\frac{2\mu}{\gamma}\varepsilon_2\|\dot{x}(t)\|^2\right)\\ &=-2[\|x(t)\|\|\dot{x}(t)\|]\begin{bmatrix}\kappa-\varepsilon_1 & -\frac{1}{2}\left(\varepsilon_2+\frac{2\mu}{\gamma}\varepsilon_1\right)\\ -\frac{1}{2}\left(\varepsilon_2+\frac{2\mu}{\gamma}\varepsilon_1\right) & \mu-\frac{2\mu}{\gamma}\varepsilon_2\end{bmatrix}\\ &\quad\times\begin{bmatrix}\|x(t)\|\\ \|\dot{x}(t)\|\end{bmatrix}.\end{aligned} \tag{11}$$

Then, $V(y(t))$ is a Lyapunov function if

$$\kappa-\varepsilon_1>0,\quad 1-\frac{2}{\gamma}\varepsilon_2>0,\quad \text{and}$$

$$\det\times\left(\begin{bmatrix}\kappa-\varepsilon_1 & -\frac{1}{2}\left(\varepsilon_2+\frac{2\mu}{\gamma}\varepsilon_1\right)\\ -\frac{1}{2}\left(\varepsilon_2+\frac{2\mu}{\gamma}\varepsilon_1\right) & \mu-\frac{2\mu}{\gamma}\varepsilon_2\end{bmatrix}\right)>0. \tag{12}$$

The inequalities in (12) hold if

$$\mu\kappa-\mu\varepsilon_1-\frac{2\mu\kappa}{\gamma}\varepsilon_2-\frac{1}{4}\left(\varepsilon_2-\frac{2\mu}{\gamma}\varepsilon_1\right)^2>0. \tag{13}$$

Therefore, if the inequality (5) holds, $V(y(t))$ is a Lyapunov function for system (3) and the nonlinear system (1) is globally asymptotically stable.

THEOREM 2. Suppose that the nonlinear perturbation f satisfies condition (2); then system (1) is globally asymptotically stable if

$$\varepsilon=\max\{\varepsilon_1,\varepsilon_2\}<\frac{2\gamma\kappa}{\sqrt{(2\kappa+\gamma)^2+\frac{\kappa}{\mu}(2\mu-\gamma)^2}+(2\kappa+\gamma)}. \tag{14}$$

Proof. Taking the same Lyapunov function candidate $V(y(t))=y^T(t)Py(t)$ and employing the very similar ways as the proof of Theorem 1, we have

$$\begin{aligned}\dot{V}(y(t))&\leqslant-2(\kappa\|x(t)\|^2+\mu\|\dot{x}(t)\|^2)+2\varepsilon\left(\|x(t)\|^2\right.\\ &\quad\left.+\left(1+\frac{2\mu}{\gamma}\right)\|x(t)\|\|\dot{x}(t)\|+\frac{2\mu}{\gamma}\|\dot{x}(t)\|^2\right)\\ &=-2[\|x(t)\|\|\dot{x}(t)\|]\\ &\quad\times\begin{bmatrix}\kappa-\varepsilon & -\frac{1}{2}\varepsilon\left(1+\frac{2\mu}{\gamma}\right)\\ -\frac{1}{2}\varepsilon\left(1+\frac{2\mu}{\gamma}\right) & \mu-\frac{2\mu}{\gamma}\varepsilon\end{bmatrix}\begin{bmatrix}\|x(t)\|\\ \|\dot{x}(t)\|\end{bmatrix}.\end{aligned} \tag{15}$$

Then, $V(y(t))$ is a Lyapunov function if

$$\kappa-\varepsilon>0,\quad \text{and}$$

$$\det\left(\begin{bmatrix}\kappa-\varepsilon & -\frac{1}{2}\varepsilon\left(1+\frac{2\mu}{\gamma}\right)\\ -\frac{1}{2}\varepsilon\left(1+\frac{2\mu}{\gamma}\right) & \mu-\frac{2\mu}{\gamma}\varepsilon\end{bmatrix}\right)>0. \tag{16}$$

The last inequality in (16) holds if

$$\mu\kappa-\left(\mu+\frac{2\mu\kappa}{\gamma}\right)\varepsilon-\frac{1}{4}\left(1-\frac{2\mu}{\gamma}\right)^2\varepsilon^2>0. \tag{17}$$

The inequality (17) is met if

$$\varepsilon<\begin{cases}\dfrac{\gamma\kappa}{\gamma+2\kappa}, & \text{for } \mu=\dfrac{\gamma}{2}\\ \dfrac{2\gamma}{(\gamma-2\mu)^2}[\sqrt{\mu^2(\gamma+2\kappa)^2+\mu\kappa(\gamma-2\mu)^2}-\mu(\gamma+2\kappa)], & \text{for } \mu\neq\dfrac{\gamma}{2}\end{cases}$$

$$=\frac{2\gamma\mu\kappa}{\sqrt{\mu^2(\gamma+2\kappa)^2+\mu\kappa(\gamma-2\mu)^2}+\mu(\gamma+2\kappa)}=\frac{2\gamma\kappa}{\sqrt{(\gamma+2\kappa)^2+\dfrac{\kappa}{\mu}(\gamma-2\mu)^2}+(\gamma+2\kappa)}.$$

Thus, taking notice of

$$\frac{2\gamma\kappa}{\sqrt{(\gamma+2\kappa)^2+\dfrac{\kappa}{\mu}(\gamma-2\mu)^2}+(\gamma+2\kappa)}\leqslant\frac{\gamma\kappa}{2\kappa+\gamma}\leqslant\min\left\{\kappa,\ \frac{\gamma}{2}\right\}, \tag{18}$$

it is easily to verify that the inequalities in (16) are met if the inequality (14) holds. This completes the proof of Theorem 2.

Consider a nonlinear system whose mass matrix has been normalized to the identity, which is a special case of system (1), described as

$$\ddot{x}(t)+D\dot{x}(t)+Kx(t)=f(x(t),\dot{x}(t),t), \tag{19}$$

where $x(t)\in R^n$, $f\in R^n$ satisfies condition (2); D and $K\in R^{n\times n}$ are symmetric positive definite damping and stiffness matrices, respectively. By Theorems 1 and 2, it is easy to obtain the following Corollary.

COROLLARY 1. Suppose that the nonlinear perturbation f satisfies condition (2); then system (19) is globally asymptotically stable if

$$\frac{1}{4}\left(\varepsilon_2-\frac{2}{\gamma}\varepsilon_1\right)^2+\varepsilon_1+\frac{2\kappa}{\gamma}\varepsilon_2<\kappa, \tag{20}$$

or

$$\varepsilon=\max\{\varepsilon_1,\varepsilon_2\}<\frac{2\gamma\kappa}{\sqrt{(2\kappa+\gamma)^2+\kappa(2-\gamma)^2}+(2\kappa+\gamma)} \tag{21}$$

where $\kappa=\lambda_m(K)$ and $\gamma=\lambda_m(D)$.

Remark 1. For system (19), Cox and Moro [9] worked out a stability criterion that bounds the degree of nonlinear perturbation and the deviation from classically damping. In Cox and Moro [9], the stiffness and damping matrices are assumed to be positive definite and can be written as

$$K=K_0+K_1 \quad \text{and} \quad D=D_0+D_1, \tag{22}$$

where $K_0D_0=D_0K_0$, i.e., K_0 and D_0 is a classically damped pair, and there exists a positive constant δ such that

$$\|K_1\|\leqslant\delta \quad \text{and} \quad \|D_1\|\leqslant\delta. \tag{23}$$

Based on a special Lyapunov function, Cox and Moro [9] derived a sufficient condition for asymptotic stability of system (19) that can be written as

$$\varepsilon+\delta<\frac{2\gamma_0\kappa_0}{\sqrt{(2\kappa_0+\gamma_0)^2+\kappa_0(2-\gamma_0)^2}+(2\kappa_0+\gamma_0)}, \tag{24}$$

where $\varepsilon=\max\{\varepsilon_1,\varepsilon_2\}$, $\kappa_0=\lambda_m(K_0)$ and $\gamma_0=\lambda_m(D_0)$. Obviously, the inequality (24) coincides with the inequality (21) if the linear part of the nonlinear system (19) is classically damped. If the stiffness and damping matrices K and D is not a classically damped pair, however, the criterion in Cox and Moro [9] is relatively conservative in comparing with the results in Corollary 1.

3 Robust Stabilization by Symmetric Output Feedback Control

In this section we extend the above approach to the study of robust controller design of second-order nonlinear systems. The symmetric output feedback is introduced for the control law. This idea builds directly upon the work reported in Junkins and Kim [11] and the literate cited therein.

Consider a class of nonlinear dynamical systems modeled by the equation of motion

$$M\ddot{x}(t)+D\dot{x}(t)+Kx(t)=f(x(t),\dot{x}(t),t)+Bu(t) \tag{25}$$

where $x(t)\in R^n$ and $u(t)\in R^m$ are configuration and control vectors, respectively; $f\in R^n$ is a smooth function of x, $\dot{x}$, and t; M, D, and $K\in R^{n\times n}$ are symmetric positive definite mass, damping and stiffness matrices, respectively; $B\in R^{n\times m}$ is the control influence matrix. Also, assume that there exist positive constants ε_1 and ε_2 such that the inequality (2) holds.

In order to maintain the symmetric property of the system, we introduce the following symmetric output feedback form of the control law:

$$u=-(G_1y+G_2\dot{y}), \quad y=B^Tx \tag{26}$$

where G_1 and G_2 are $m\times m$ positive definite symmetric gain matrices.

With the assumption that the sensors and actuators are perfect, i.e., linear and instantaneous, the closed-loop equations can be written as

$$M\ddot{x}(t)+(D+BG_2B^T)\dot{x}(t)+(K+BG_1B^T)x(t)=f(x(t),\dot{x}(t),t). \tag{27}$$

Since the gain matrices G_1 and G_2 are positive definite, the control-induced stiffness and damping terms BG_1B^T and BG_2B^T are positive semi-definite. Therefore, if the system is controllable, then the linear part of closed-loop system (27) is at least asymptotically stable so long as the gain matrices are chosen to be positive definite. The following theorem can be derived directly from Theorem 1.

THEOREM 3. Suppose that the nonlinear perturbation f satisfies condition (2); then the closed-loop system (27) is globally asymptotically stable if there exist positive definite gain matrices G_1 and G_2 such that

$$\kappa_c>\varepsilon_1+\frac{\varepsilon_2^2}{4\mu} \tag{28}$$

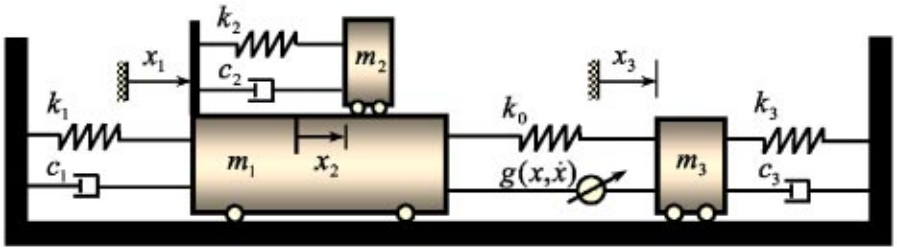


Fig. 1 Schematic of a three-degree-of-freedom vibrator

and

$$\gamma_c > \frac{2\mu(2\sqrt{(\varepsilon_2^2\kappa_c+\mu\varepsilon_1^2)(\kappa_c-\varepsilon_1)}+\varepsilon_2(2\kappa_c-\varepsilon_1))}{4\mu(\kappa_c-\varepsilon_1)-\varepsilon_2^2} \tag{29}$$

where $\mu=\lambda_M(M)$, $\kappa_c=\lambda_m(K+BG_1B^T)$, and $\gamma_c=\lambda_m(D+BG_2B^T)$.

For given bounds ε_1 and ε_2 on the nonlinear perturbations, Theorem 3 can be used to determine the output feedback control law (26) to stabilize the trivial solution of the nonlinear system (25).

4 Illustrative Examples

The following simple examples serve to illustrate the usefulness of the stability criteria presented here and to compare them with the existing criterion in the literature.

Example 1. Consider a three-degree-of-freedom nonlinear mechanical system depicted in Fig. 1. The system consists of three rigid bodies connected by springs and dampers.

The equations of motion in matrix form for the system in Fig. 1 can be written as

$$\begin{bmatrix} m_1+m_2 & m_2 & 0 \\ m_2 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{bmatrix}+\begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{bmatrix}\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix}$$
$$+\begin{bmatrix} k_1+k_0 & 0 & -k_0 \\ 0 & k_2 & 0 \\ -k_0 & 0 & k_3+k_0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} g(x,\dot{x}) \\ 0 \\ -g(x,\dot{x}) \end{bmatrix} \tag{30}$$

where $x=[x_1\ x_2\ x_3]^T$. We suppose that there exist positive constants ε_1 and ε_2 such that the nonlinear perturbation function $g(x,\dot{x})$ in (30) satisfy the following inequality:

$$|g(x,\dot{x})| \leqslant \frac{1}{2}\varepsilon_1|x_1-x_3|+\frac{1}{2}\varepsilon_2|\dot{x}_1-\dot{x}_3|. \tag{31}$$

Thus, in comparing with system (1), we have

$$\|f(x,\dot{x})\|=\left\|\begin{bmatrix} g(x,\dot{x}) \\ 0 \\ -g(x,\dot{x}) \end{bmatrix}\right\|$$
$$\leqslant\left\|\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}\left(\frac{\varepsilon_1}{2}\begin{bmatrix} |x_1| \\ |x_2| \\ |x_3| \end{bmatrix}+\frac{\varepsilon_2}{2}\begin{bmatrix} |\dot{x}_1| \\ |\dot{x}_2| \\ |\dot{x}_3| \end{bmatrix}\right)\right\|$$
$$\leqslant\varepsilon_1\|x\|+\varepsilon_2\|\dot{x}\|.$$

Obviously, system (30) is a nonclassically damped system if the viscous damping coefficient c_1 is not identical with c_3, i.e., $c_1 \neq c_3$. Consider the system (30) with

$$\begin{cases} m_1=2 \text{ kg}, \quad m_2=m_3=1 \text{ kg}; \\ c_1=0.9 \text{ Ns/m}, \quad c_2=0.6 \text{ Ns/m}, \quad c_3=0.6 \text{ Ns/m}; \\ k_1=4 \text{ N/m}, \quad k_2=5 \text{ N/m}, \quad k_3=4 \text{ N/m}, \quad k_0=2 \text{ N/m}. \end{cases}$$

It follows that

$$\mu=\lambda_M(M)=2+\sqrt{2}, \quad \gamma=\lambda_m(D)=0.6, \quad \kappa=\lambda_m(K)=4. \tag{32}$$

Application of the stability condition in Theorem 1 yields

$$(\varepsilon_2-5.690356\varepsilon_1)^2+13.656854\varepsilon_1+180.091390\varepsilon_2<54.627417. \tag{33}$$

From Theorem 2, we have

$$\varepsilon=\max\{\varepsilon_1,\varepsilon_2\}<0.245808. \tag{34}$$

The results are depicted in Fig. 2 where the stability region resulted from (33) is enclosed by the curve C and the axes, while the region given by (34) is enclosed by the lines L_1, L_2 and the axes. It can be seen from Fig. 2 that the stability bound given by condition (14) in Theorem 2 is relatively conservative than that in Theorem 1.

Example 2. Consider system (1) with

$$M=I, \quad D=\begin{bmatrix} 1 & -0.5 \\ -0.5 & 2 \end{bmatrix}, \quad K=\begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}. \tag{35}$$

Then, we have $\gamma=\lambda_m(D)=1/2(3-\sqrt{2})$, $\kappa=\lambda_m(K)=4$. Application of the stability condition (20) in Corollary 1 yields

$$(\varepsilon_2-2.522408\varepsilon_1)^2+4\varepsilon_1+40.358524\varepsilon_2<16. \tag{36}$$

Applying the stability condition (21) in Corollary 1, it can be easily computed that

$$\varepsilon=\max\{\varepsilon_1,\varepsilon_2\}<0.354144. \tag{37}$$

The corresponding stability region is depicted in Fig. 3. Obviously, the stability region given by condition (20) is much larger than that given by condition (21).

Since the matrices K and D do not commute, we cannot directly employ Theorem 3.1 of Cox and Moro [9]. Let

$$D_0=\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad D_1=\begin{bmatrix} 0 & -0.5 \\ -0.5 & 0 \end{bmatrix}.$$

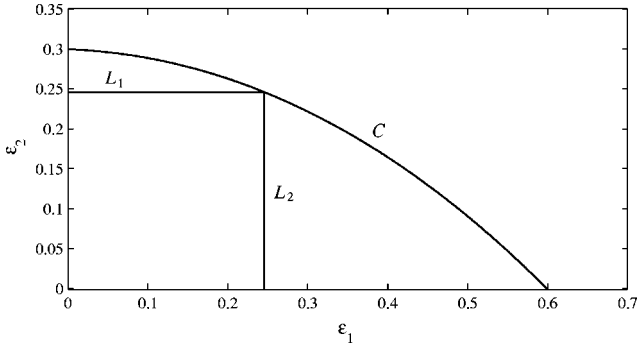


Fig. 2 Predicted stability regions for the system in Example 1

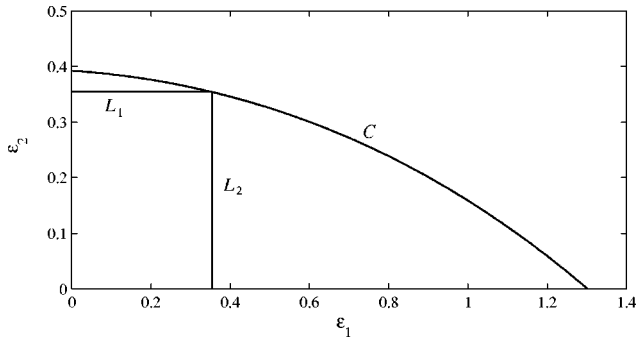


Fig. 3 Predicted stability regions for the system in Example 2

Then, we have $KD_0 = D_0K$ and

$$\gamma_0 = \lambda_m(D_0) = 1, \quad \kappa_0 = \lambda_m(K) = 4, \quad \delta = \|D_1\| = 0.5.$$

Therefore, according to Theorem 3.1 of Cox and Moro [9], the system is stable if $\varepsilon + \delta < 0.439089$. This inequality cannot be satisfied since $\delta = 0.5 > 0.439089$. Thus, for this example, the stability criterion of Cox and Moro [9] is not being able to draw any conclusion.

5 Conclusions

In conclusion, a specific Lyapunov function has been chosen to investigate the stability robustness of nonclassically damped systems with nonlinear perturbations. In terms of the positive definite mass, damping and stiffness matrices, algebraic criteria have been introduced to guarantee the asymptotic stability of the system. The checking of stability by our criteria can be carried out rather simply. This is believed to be a useful contribution to existing scarce literature on the problem of robust stability of multi-degree-of-freedom systems. Moreover, the symmetric output feedback control law is introduced to stabilize the second-order system with nonlinear uncertainties. Numerical examples have demonstrated that the new stability criteria are less conservative and more powerful comparing to those in the literature.

Acknowledgments

This work has been supported by a research grant from Southwest Jiaotong University. The first author would like to thank Prof. Ping He for beneficial discussions on this work.

References

[1] Yedavalli, R. K., 1985, ''Perturbation Bounds for Robust Stability in Linear State Space Models,'' Int. J. Control, **42**(6), pp. 1507–1517.
[2] Zhou, K. M., and Khargonekar, P. P., 1987, ''Stability Robustness Bounds for Linear State Space Models With Structured Uncertainty,'' IEEE Trans. Autom. Control, **3**2(7), pp. 621–623.
[3] Siljak, D. D., 1989, ''Parameter Space Methods for Robust Control Design: A Guided Tour,'' IEEE Trans. Autom. Control, **34**(7), pp. 674–688.
[4] Bien, Z. M., and Kim, J. H., 1992, ''A Robust Stability Bound of Linear Systems With Structured Uncertainty,'' IEEE Trans. Autom. Control, **37**(10), pp. 1549–1551.
[5] Shieh, L. S., Mehio, M. M., and Dib, H. M., 1987, ''Stability of the Second Order Matrix Polynomial,'' IEEE Trans. Autom. Control, **32**(3), pp. 231–233.
[6] Hsu, P., and Wu, J., 1991, ''Stability of Second-Order Multidimensional Linear Time-Varying Systems,'' J. Guid. Control Dyn., **14**(5), pp. 1040–1045.
[7] Cao, D. Q., and Shu, Z. Z., 1994, ''Robust Stability Bounds for Multi-Degree-of-Freedom Linear Systems With Structured Perturbations,'' Dyn. Stab. Syst., **9**(1), pp. 79–87.
[8] Diwekar, A. M., and Yedavalli, R. K., 1999, ''Robust Controller Design for Matrix Second-Order Systems With Structured Uncertainties,'' IEEE Trans. Autom. Control, **44**(2), pp. 401–405.
[9] Cox, S. J., and Moro, J., ''A Lyapunov Function for Systems Whose Linear Part is Almost Classically Damped,'' ASME J. Appl. Mech., **64**(4), pp. 965–968.
[10] Skelton, R. E., Iwasaki, T., and Geigoriadis, K., 1998, *A Unified Algebraic Approach to Linear Control Design*, Taylor & Francis, London.
[11] Junkins, J. L., and Kim, Y., 1993, *Introduction to Dynamics and Control of Flexible Structures*, AIAA, Washington, DC.

Dynamic Response of a Clamped Circular Sandwich Plate Subject to Shock Loading

X. Qiu

V. S. Deshpande

N. A. Fleck[1]
e-mail: naf1@eng.cam.ac.uk

Engineering Department,
Cambridge University,
Trumpington Street,
Cambridge CB1 1PZ, UK

An analytical model is developed for the deformation response of clamped circular sandwich plates subjected to shock loading in air and in water. The deformation history is divided into three sequential stages and analytical expressions are derived for the deflection, degree of core compression, and for the overall structural response time. An explicit finite element method is employed to assess the accuracy of the analytical formulas for the simplified case where the effects of fluid-structure interaction are neglected. The sandwich panel response has only a low sensitivity to the magnitude of the core compressive strength and to the degree of strain hardening in the face-sheets. The finite element results confirm the accuracy of the analytical predictions for the rigid ideally plastic sandwich plates. The analytical formulas are employed to determine optimal geometries of the sandwich plates that maximize the shock resistance of the plates for a given mass. The optimization reveals that sandwich plates have a superior shock resistance relative to monolithic plates of the same mass. [DOI: 10.1115/1.1778416]

1 Introduction

Clamped sandwich plates are representative of the structures used in the design of commercial and military vehicles. For example, the outermost structure on a ship comprises plates welded to an array of stiffeners. The superior performance of sandwich plates relative to monolithic solid plates is well known for applications requiring high quasi-static strength. However, the resistance of sandwich plates to dynamic loads remains to be fully investigated in order to quantify the advantages of sandwich design over monolithic design for application in shock resistant structures.

The response of monolithic beams and plates to shock type loading has been extensively investigated over the past 50 years or so. For example, Wang and Hopkins [1] and Symmonds [2] analyzed the response of clamped circular plates and beams, respectively, under impulsive loads. However, their analyses was restricted to small deflections and linear bending kinematics. By direct application of the principle of virtual work for an assumed deformation mode, Jones presented approximate solutions for simply supported and clamped beams, [3], and also simply supported circular plates, [4], undergoing finite deflections.

Recently, Xue and Hutchinson [5] carried out a preliminary finite element (FE) investigation of the resistance of clamped circular sandwich plates with a foamlike core to shock loading with the effects of fluid-structure interaction neglected. By employing a series of FE calculations they demonstrated that near-optimized circular sandwich plates offer a higher resistance to shock loading than monolithic plates of the same mass. In parallel studies, Fleck and Deshpande [6] proposed an analytical model for the response of clamped sandwich beams to shock loadings including the effects of fluid-structure interaction and showed that the analytical predictions are in close agreement with FE calculations, [7].

In this study we extend the analytical method of Fleck and Deshpande [6] to clamped circular sandwich plates. First, analytical formulas are presented for the response of clamped rigid-ideally plastic circular sandwich plates to a uniform shock wave, including the effects of fluid-structure interaction. Next, the analytical predictions of the response of sandwich plates are compared with FE predictions for the case where the effect of fluid-structure interaction is neglected: This loading represents shock loading in air. Finally, the analytical formulas are used to determine the optimal designs of sandwich plates that maximize the shock resistance in air for a given mass and the performance gain of these optimal sandwich plates over monolithic plates is quantified.

2 An Analytical Model for the Shock Resistance of Clamped Sandwich Plates

Fleck and Deshpande [6] have developed an analytical model for the response of clamped sandwich beams subject to air and underwater shock loading. This model is now extended to analyze the response of clamped axisymmetric sandwich plates to a spatially uniform air or underwater shock.

Consider a clamped circular sandwich plate of radius R with identical face-sheets of thickness h and a core of thickness c, as shown in Fig. 1. The face-sheets are made from a rigid ideally plastic solid of yield strength σ_{fY}, density ρ_f, and tensile failure strain ϵ_f. The core is taken to be a compressible isotropic solid of density ρ_f and deforms in uniaxial compression the core at a constant strength σ_c with no lateral expansion up to a densification strain ϵ_D; beyond densification the core is treated as rigid. Fleck and Deshpande [6] split the response of the sandwich structure into three sequential stages:

(i) Stage I—fluid-structure interaction phase,
(ii) Stage II—core compression phase, and
(iii) Stage III—plate bending and stretching phase.

Here, we assume a similar separation of time scales for the sandwich plate deformation history.

Stage I—The initial fluid-structure interaction phase

G. I. Taylor [8] developed the solution for a one-dimensional wave impinging a free-standing plate to compute the momentum transmitted to the plate by the shock pulse. Fleck and Deshpande [6] followed this approach and similarly computed the momentum transmitted to the sandwich beam by treating the outer face of the

[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 Applied Mechanics Division, September 30, 2003; final revision, January 20, 2004. Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

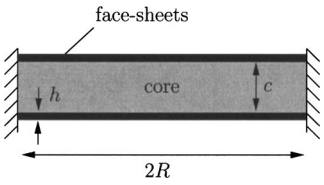


Fig. 1 Geometry of the clamped sandwich plate

sandwich beam as a free-standing plate. Their analysis also holds for the circular sandwich plate, and we briefly review the relevant equations.

The pressure p at any point in the fluid of density ρ_w engulfed by the pressure wave travelling at a velocity c_w is taken to be (starting at time $t=0$)

$$p = p_o e^{-t/\theta}, \tag{1}$$

where p_o is the peak pressure and θ the decay constant of the wave. When this pressure wave hits a stationary rigid plate at normal incidence it imparts an impulse I

$$I = 2\int_0^\infty p_o e^{-t/\theta} dt = 2p_o\theta, \tag{2}$$

to the plate. The factor of two arises in relation (2) due to full reflection of the pressure wave.

If instead, the pressure wave impacts a free-standing plate, the imparted impulse is less than I, and can be estimated as follows. When the pressure wave strikes a free-standing plate of thickness h made from a material of density ρ_f, it sets the plate in motion and is partly reflected. At the instant the plate achieves its maximum velocity, the pressure at the interface between the plate and the fluid is zero and cavitation sets in shortly thereafter. The momentum per unit area I_{trans} transmitted into the structure is then given by

$$I_{\text{trans}} = \zeta I, \tag{3a}$$

where

$$\zeta \equiv \psi^{\psi/(1-\psi)}, \tag{3b}$$

and $\psi \equiv \rho_w c_w \theta/(\rho_f h)$. It is assumed that this transmitted impulse imparts a uniform velocity $v_o = I_{\text{trans}}/(\rho_f h)$ to the outer front face of the sandwich plate.

In the present model, the effect of the fluid after the first cavitation event is neglected. This is consistent with the observation that the secondary shocks have a much smaller effect on the structure compared to the primary shock wave, see Cole [9].

Stage II—Core compression phase

At the start of this phase, the outer face has a velocity v_o while the core and inner face are stationary. The outer face compresses the core, while the core with compressive strength σ_c decelerates the outer face and simultaneously accelerates the inner face. The final common velocity of the faces and the core is dictated by momentum conservation and the ratio ϕ of the energy lost U_{lost} in this phase to the initial kinetic energy $I^2\zeta^2/2\rho_f h$ of the outer face is given by

$$\phi \equiv \frac{U_{\text{lost}}}{I^2\zeta^2/(2\rho_f h)} = \frac{1+\overline{m}}{2+\overline{m}}, \tag{4}$$

where $\overline{m} = \rho_c c/(\rho_f h)$ is the ratio of the mass of the core to the mass of a face-sheet. This energy lost is dissipated by plastic dissipation in compressing the core and thus the average through-thickness strain ϵ_c in the core is given by

$$\epsilon_c = \frac{\overline{I}^2\zeta^2}{2\overline{\sigma c}^2\overline{h}} \frac{\overline{h}+\overline{\rho}}{2\overline{h}+\overline{\rho}}, \tag{5}$$

where $\overline{h} \equiv h/c$, $\overline{c} \equiv c/R$, $\overline{\rho} \equiv \rho_c/\rho_f$, $\overline{I} \equiv I/(R\sqrt{\sigma_{fY}\rho_f})$ and $\overline{\sigma} \equiv \sigma_c/\sigma_{fY}$. However, if U_{lost} is too high such that ϵ_c as given by (5) exceeds the densification strain ϵ_D, then ϵ_c is set equal to ϵ_D and the model does not account explicitly for the additional dissipation mechanisms required to conserve energy. Rather it is assumed that inelastic impact of the outer face against the combined core and inner face leads to the additional dissipation. After the core has compressed by a strain of ϵ_c, the core height is reduced to $(1-\epsilon_c)c$. An approximate estimate of the time T_c for this second stage of motion (calculated by neglecting the mass of the core) is given by [6]

$$\overline{T}_c \equiv \frac{T_c}{R\sqrt{\rho_f/\sigma_{fY}}} = \begin{cases} \dfrac{\overline{I}\zeta}{2\overline{\sigma}}, & \text{if } \overline{I}^2\zeta^2 < 4\overline{\sigma}\overline{c}^2\overline{h}\epsilon_D \\ \dfrac{\overline{I}\zeta}{2\overline{\sigma}}\left[1-\sqrt{1-\dfrac{4\overline{\sigma c}^2\overline{h}\epsilon_D}{\overline{I}^2\zeta^2}}\right], & \text{otherwise.} \end{cases} \tag{6}$$

This time T_c is typically small compared to the structural response time and thus the transverse deflection of the inner face of the sandwich plate in this stage can be neglected.

Stage III—Plate bending and stretching phase

At the end of Stage II, the sandwich plate has a uniform velocity except for a boundary layer near the supports. The plate is brought to rest by plastic bending and stretching. The problem under consideration is a classical one: what is the dynamic response of a clamped plate of radius R with an initial uniform transverse velocity v? The structural response is broken down into two phases: (i) *small displacement analysis* as first considered by Wang and Hopkins [1] and (ii) *large displacement analysis.*

(i) Small displacement analysis

When the transverse displacement of the plate $w(t)$ is less than the total thickness $2h+c$, the dynamic response is governed by bending and transverse inertia of the plate. Wang and Hopkins [1] showed that the plate response comprises two sequential phases. Phase I comprises stationary plastic hinges at the supports and plastic hinges travelling inwards from each clamped support. After the moving hinges have coalesced at the center of the plate, continued rotation occurs about the central hinge until the plate is brought to rest in phase II.

We now introduce the appropriate nondimensional geometric parameters for the sandwich plate

$$\overline{c} \equiv \frac{c}{R}, \quad \hat{c} \equiv \overline{c}(1-\epsilon_c), \quad \overline{h} \equiv \frac{h}{c} \text{ and } \hat{h} \equiv \frac{\overline{h}}{1-\epsilon_c}, \tag{7}$$

and the nondimensional material properties of the core

$$\overline{\rho} \equiv \frac{\rho_c}{\rho_f}, \quad \text{and } \overline{\sigma} \equiv \frac{\sigma_c}{\sigma_{fY}}. \tag{8}$$

The nondimensional structural response time $\overline{T}$ and blast impulse $\overline{I}$ are

$$\overline{T} \equiv \frac{T}{R}\sqrt{\frac{\sigma_{fY}}{\rho_f}}, \quad \overline{I} \equiv \frac{I}{R\sqrt{\rho_f\sigma_{fY}}}. \tag{9}$$

In the small deflection regime, the maximum central deflection w of the inner face of the sandwich plate and the structural response time T are given by Eqs. (4.99) and (4.100), respectively, of Jones [4]. Noting that the plastic bending moment M_o of the circular sandwich plate is given by

$$M_o = \sigma_c \frac{(1-\epsilon_c)^2 c^2}{4} + \sigma_{fY} h[(1-\epsilon_c)c + h], \tag{10}$$

these equations reduce to

$$\bar{w} \equiv \frac{w}{R} = 0.28 \frac{\bar{I}^2 \zeta^2}{\bar{c}\hat{c}^2 \alpha_1 (2\bar{h} + \bar{\rho})}, \tag{11a}$$

and

$$\bar{T} \equiv \frac{T}{R} \sqrt{\frac{\sigma_{fY}}{\rho_f}} = 0.36 \frac{\bar{I}\zeta}{\hat{c}^2 \alpha_1}, \tag{11b}$$

where

$$\alpha_1 = (1 + 2\bar{h})^2 - 1 + \bar{\sigma}, \tag{12a}$$

and

$$\alpha_2 = \sqrt{\frac{2\bar{h} + \bar{\rho}}{2\hat{h} + \bar{\sigma}}}. \tag{12b}$$

(ii) Large displacement analysis

The above analysis ignores the buildup of membrane action associated with the lateral deflection of the clamped plates. Jones [4] has taken this into account by assuming that the plate deflects from the initial undeformed configuration with a velocity profile which decreases linearly from a maximum value at the center to zero at the supports. The analysis in Jones [4] is given for a simply supported circular plate and can be easily extended to clamped circular plates by assuming that stationary plastic hinges form at the center and at the clamped supports of the plate. Plastic dissipation is both by rotation about discrete plastic hinges and by uniform radial stretching of the plate due to its transverse displacement between the clamped supports.

The yield locus of an axisymmetric sandwich element subjected to a circumferential membrane force N_θ and a circumferential bending moment M_θ is well approximated by

$$\frac{M_\theta}{M_o} + \frac{N_\theta}{N_o} = 1, \tag{13}$$

where M_o is the plastic bending moment specified by (10) and N_o the circumferential plastic membrane force given by

$$N_o = 2h\sigma_{fY} + (1 - \epsilon_c) c \sigma_c, \tag{14}$$

where we have assumed the strength of the foam is unaffected by core compression. Analytical formulas for the deflection and structural response time of the circular plate can be obtained by approximating the above yield locus by either inscribing or circumscribing squares as sketched in Fig. 2(*a*). Employing a procedure similar to that detailed in Jones [4] the maximum central deflection w of the inner face and structural response time T of a clamped circular sandwich plate are given by

$$\bar{w} = \frac{\bar{c}\alpha_1}{2\hat{h} + \bar{\sigma}} \left(\sqrt{1 + \frac{2}{3} \frac{\bar{I}^2 \zeta^2}{\bar{c}\hat{c}^3 \alpha_1^2 \alpha_2^2}} - 1 \right), \tag{15a}$$

and

$$\bar{T} = \alpha_2 \sqrt{\frac{\bar{c}}{6\hat{c}}} \tan^{-1} \left(\sqrt{\frac{2}{3\bar{c}\hat{c}}} \frac{\bar{I}\zeta}{\hat{c}\alpha_1 \alpha_2} \right), \tag{15b}$$

respectively, for the choice of a circumscribing yield locus, and by

$$\bar{w} = \frac{\hat{c}\alpha_1}{2\hat{h} + \bar{\sigma}} \left(\sqrt{1 + \frac{4}{3} \frac{\bar{I}^2 \zeta^2}{\bar{c}\hat{c}^3 \alpha_1^2 \alpha_2^2}} - 1 \right), \tag{16a}$$

and

$$\bar{T} = \alpha_2 \sqrt{\frac{\bar{c}}{3\hat{c}}} \tan^{-1} \left(\frac{2}{\sqrt{3\bar{c}\hat{c}}} \frac{\bar{I}\zeta}{\hat{c}\alpha_1 \alpha_2} \right), \tag{16b}$$

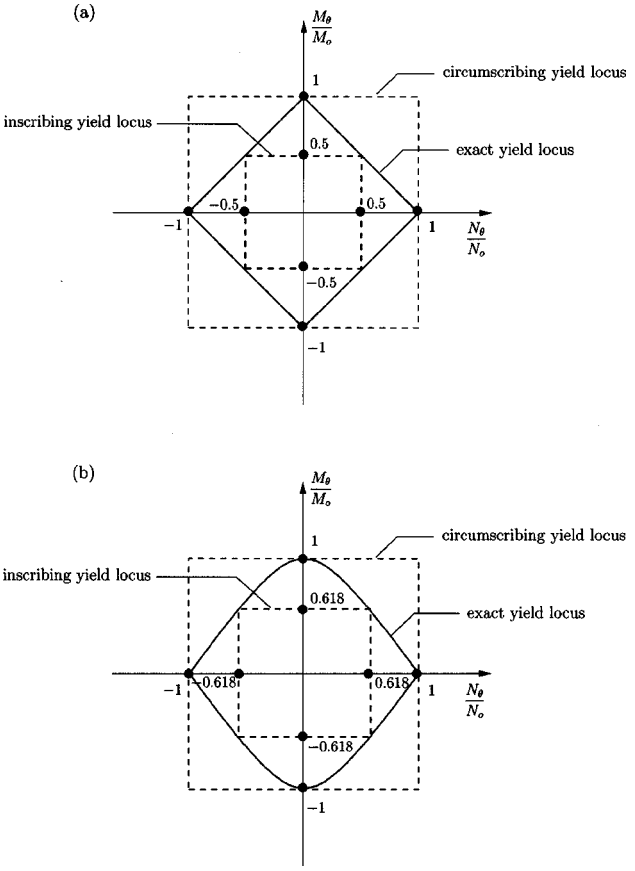


Fig. 2 Sketches of the exact, inscribing and circumscribing yield loci of (*a*) the sandwich plate and (*b*) the monolithic plate. Here, M_o and N_o are the fully plastic bending moments and axial loads, respectively, of the plates.

for an inscribing yield locus. A number of criteria can be devised for the transition from the small to the large deflection analysis. For example, the transition can be assumed to occur at an impulse level where both analyses predict equal displacements. It will be shown subsequently in the comparisons with the FE calculations that for most practical values of displacement or impulses, the large displacement solution suffices. Thus, we propose here to use the large displacement solution over the entire range of impulses.

The circumferential tensile strain ϵ_m in the face-sheets due to stretching is approximately equal to

$$\epsilon_m = \frac{1}{2} \bar{w}^2. \tag{17}$$

Neglecting the strains due to bending, an approximate failure criterion for the sandwich plates is given by setting the face-sheet tensile strain ϵ_m equal to the tensile ductility ϵ_f of the face-sheet material.

2.1 Response of a Monolithic Clamped Plate. Similar expressions exist for the deflection and structural response time of a monolithic clamped circular plate. For monolithic plates, no core compression phase exists and Stage II of the deformation history vanishes. Again the analysis is divided into the small and large displacement regimes. Consider a monolithic plate of thickness H and radius R made from a solid material with yield strength σ_{fY}. Then the analysis of Wang and Hopkins [1] implies that the maximum central deflection w and the time T to attain this deflection are given by

$$\bar{w} \equiv \frac{w}{R} = 0.28 \bar{I}^2 \zeta^2 \left(\frac{R}{H} \right)^3, \tag{18a}$$

and

$$\overline{T} \equiv \frac{T}{R}\sqrt{\frac{\sigma_{fY}}{\rho_f}} = 0.36\overline{I}\,\overline{\zeta}\left(\frac{R}{H}\right)^2. \tag{18b}$$

Next consider the large displacement regime. The yield locus for any plate element of the monolithic circular plate, subject to a circumferential membrane force N_θ and bending moment M_θ is given by

$$\left(\frac{M_\theta}{M_o}\right)^2 + \frac{N_\theta}{N_o} = 1, \tag{19}$$

where $N_o = H\sigma_{fY}$ and $M_o = \sigma_{fY}H^2/4$. The maximum central deflection w of the clamped circular plate and the structural response time in the large displacement regime can be calculated in the manner detailed in Jones [4] by approximating the yield locus by either inscribing or circumscribing squares as sketched in Fig. 2(*b*). Under the assumption of an inscribing yield locus, the nondimensional deflection $\overline{w}$ and structural response time $\overline{T}$ are given by

$$\overline{w} = \left(\frac{H}{R}\right)\left(\sqrt{1 + 1.079\overline{I}^2\overline{\zeta}^2\left(\frac{R}{H}\right)^4} - 1\right), \tag{20a}$$

and

$$\overline{T} = 0.519\tan^{-1}\left[1.039\overline{I}\,\overline{\zeta}\left(\frac{R}{H}\right)^2\right]. \tag{20b}$$

Similarly, the assumption of a circumscribing yield locus gives

$$\overline{w} = \left(\frac{H}{R}\right)\left(\sqrt{1 + \frac{2}{3}\overline{I}^2\overline{\zeta}^2\left(\frac{R}{H}\right)^4} - 1\right), \tag{21a}$$

and

$$\overline{T} = \frac{1}{\sqrt{6}}\tan^{-1}\left[\sqrt{\frac{2}{3}}\overline{I}\,\overline{\zeta}\left(\frac{R}{H}\right)^2\right]. \tag{21b}$$

Again, it will be shown via FE calculations that the large displacement solution is adequate over the entire range of deflections.

In the analytical formulas given above, we have ignored shear deflections of the plates. For the slender sandwich plate facesheets ($h/R \to 0$) and monolithic plates ($H/R \to 0$) under consideration here, Jones and Gomes de Oliveira [10] have shown that the shear deflections are negligible. Thus, it suffices to consider only the bending deflections of the plates as done above. Also, strain-rate effects in the parent material have been neglected in the current analysis. As a first-order approximation, Perrone and Bhadra [11] have shown that the effect of strain rate sensitivity can be captured by replacing σ_{fY} with the flow stress σ_o associated with the strain-rate in the beam at the representative transverse deflection $2w/3$.

3 Finite Element Study

In order to assess the accuracy of the above analytical model, a finite element (FE) study was conducted with the effects of fluid-structure interaction neglected. In the limit of no fluid-structure interaction ($\psi = 0$ and $\zeta = 1$) it is assumed that the entire shock impulse I is transferred uniformly to the outer face of the sandwich plate and to the full section of the monolithic plate. It is worth mentioning here that Xue and Hutchinson [5] demonstrated that impulsive loading of clamped circular sandwich plates suffices to capture the response of these plates subject to pressure versus time histories corresponding to most practical shock loadings.

In all the FE calculations presented here, loading corresponding to a nondimensional impulse $\overline{I}$ is specified by imparting an initial uniform velocity v_o

$$v_o = \frac{\overline{I}}{\overline{c}\,\overline{h}}\sqrt{\frac{\sigma_{fY}}{\rho_f}}, \tag{22}$$

to the outer face-sheet of the sandwich plate and by giving an initial velocity

$$v_o = \frac{\overline{I}R}{H}\sqrt{\frac{\sigma_{fY}}{\rho_f}} \tag{23}$$

uniformly to the monolithic plate.

3.1 Constitutive Description. Unless otherwise specified, the material properties of the sandwich plates are taken to be as follows. The face-sheets of the sandwich plate are assumed to be made from an elastic ideally plastic solid with yield strength σ_{fY}, a yield strain ϵ_{fY} and density ρ_f. The Young's modulus is specified by $E_f \equiv \sigma_{fY}/\epsilon_{fY}$. The solid is modeled as a J2 flow theory solid. The core is modeled as a compressible continuum using the foam constitutive model of Deshpande and Fleck [12]. This constitutive law employs an isotropic yield surface specified by

$$\hat{\sigma} - \sigma_c = 0, \tag{24a}$$

where the equivalent stress is defined by

$$\hat{\sigma}^2 \equiv \frac{1}{1 + (\alpha/3)^2}[\sigma_e^2 + \alpha^2\sigma_m^2]. \tag{24b}$$

Here, $\sigma_e \equiv \sqrt{3s_{ij}s_{ij}/2}$ is the von Mises effective stress with s_{ij} the deviatoric stress tensor and $\sigma_m \equiv \sigma_{kk}/3$ the mean stress. The yield strength σ_c is specified as a function of the equivalent plastic strain using uniaxial compression stress versus strain data. Normality of plastic flow is assumed, and this implies that the "plastic Poisson's ratio" $\nu_p = -\dot{\epsilon}_{22}^p/\dot{\epsilon}_{11}^p$ for uniaxial compression in the 1-direction is given by

$$\nu_p = \frac{1/2 - (\alpha/3)^2}{1 + (\alpha/3)^2}. \tag{25}$$

Numerical values for the reference material properties for the sandwich plate were taken to be as follows. The face-sheets are assumed to be made from a stainless steel of yield strength $\sigma_{fY} = 500$ MPa, yield strain $\epsilon_{fY} = 0.2\%$, elastic Poisson's ratio $\nu = 0.3$, and density $\rho_f = 8000\ \mathrm{kgm}^{-3}$. The strength of the core is taken to be representative of that for a lattice material such as the octet truss, [13], made from the same solid material as the face-sheets. Thus, the isotropic core yield strength is taken to be

$$\sigma_c = 0.5\overline{\rho}\sigma_{fY}, \tag{26}$$

where $\overline{\rho} \equiv \rho_c/\rho_f$ is the relative density of the core. As the reference case, we take $\overline{\rho} = 0.1$ (i.e., core density $\rho_c = 800\ \mathrm{kgm}^{-3}$) with $\alpha = 3/\sqrt{2}$ giving a plastic Poisson's ratio $\nu_p = 0$. The plastic crush strength σ_c of the foam core is taken to be independent of the effective plastic strain up to a densification strain $\epsilon_D = 0.5$: beyond densification, a linear hardening behavior is assumed with a very large tangent modulus $E_t = 0.2E_f$. Further, the core is taken to be elastically isotropic with a yield strain $\epsilon_{cY} = 0.2\%$ and an elastic Poisson's ratio $\nu_c = 0$.

3.2 Details on the Finite Element Method. All computations were performed using the explicit time integration version of the commercially available finite element code ABAQUS version 6.2. The plate was modeled using four-noded axisymmetric quadrilateral elements with reduced integration, (element type *CAX4R* in the ABAQUS notation). Numerical damping associated with volumetric straining in ABAQUS explicit was switched off by setting the bulk viscosity associated with this damping to zero; using the default viscosity in ABAQUS results in substantial artificial viscous dissipation due to the large volumetric compression of the core. For a typical plate of geometry $\overline{c} = 0.03$ and $\overline{h} = 0.1$, there were 2 and 8 elements through the thickness of the face-sheets and core, respectively, and 100 elements along the radius R.

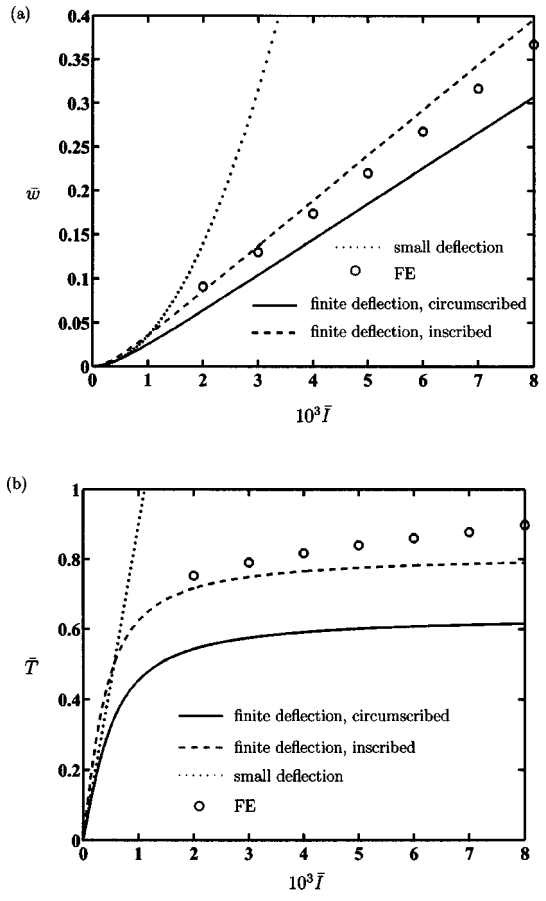


Fig. 3 Analytical and FE predictions of (*a*) maximum central deflection and (*b*) structural response time, of a monolithic plate with aspect ratio $R/H=50$ as a function of the applied impulse

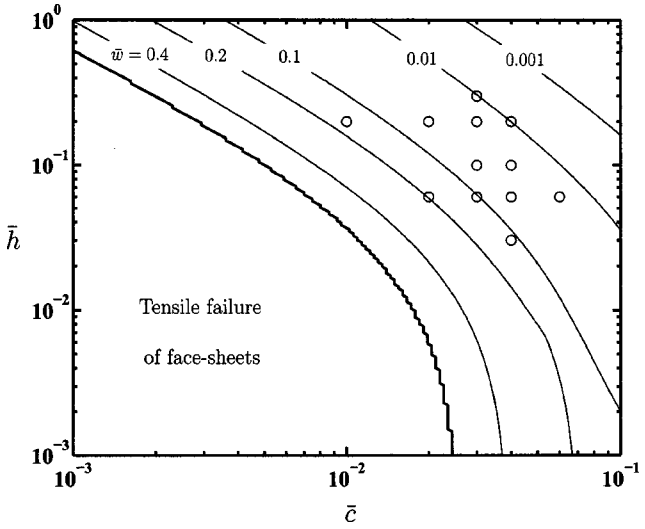


Fig. 4 Design chart for a clamped sandwich plate with core strength $\bar{\sigma}=0.05$ and densification strain $\epsilon_D=0.5$ and an assumed face-sheet material ductility $\epsilon_f=0.2$. Contours of the maximum normalized central deflection $\bar{w}$ of the inner face-sheet subject to a normalized impulse $\bar{I}=10^{-3}$ are included. The symbols denote the sandwich plate geometries selected for the FE calculations.

Mesh sensitivity studies revealed that further refinements did not improve the accuracy of the calculations appreciably.

4 Comparison of Finite Element and Analytical Predictions

4.1 Monolithic Plates. Comparisons between analytical and FE predictions of the dynamic response of monolithic plates made from the same material as the face-sheets of the reference sandwich plate (i.e., an elastic perfectly plastic solid with a yield strength $\sigma_{fY}=500$ MPa, yield strain $\epsilon_{fY}=0.2\%$, an elastic Poisson's ratio $\nu=0.3$ and a material density $\rho_f=8000\ \mathrm{kgm}^{-3}$) are presented in this section. The dependence of the normalized maximum central deflection $\bar{w}$ of the plate upon the uniformly applied normalized impulse $\bar{I}$ is shown in Fig. 3(*a*), for a plate with aspect ratio $R/H=50$. In the FE simulations, w is defined as the peak value of the central deflection versus time trace. Analytical predictions of this maximum deflection employing the small deflection analysis and the finite deflections analyses with the circumscribing and inscribing yield surfaces are included in Fig. 3(*a*). While the inscribing yield surface predictions are in good agreement with the FE results over the range of impulses investigated here, the circumscribing yield surface model underpredicts the deflections. Further, for realistic levels of the shock impulse, the FE results are captured accurately with the finite deflection analysis employing the inscribing yield locus.

The analytical and FE predictions of the normalized structural response time $\bar{T}$, as functions of the applied normalized impulse $\bar{I}$ are given in Fig. 3(*b*). In the FE simulations, T is defined as the time taken to reach the maximum deflection and written in nondimensional form via (9). For the range of impulses considered here, the response of the plate is governed by stretching and the structural response time is approximately independent of magnitude of the impulse. It is seen that the FE predictions of the structural response time T are also in good agreement with the analytical model employing the inscribing yield locus. Again the small deflection analysis is not relevant for realistic levels of shock impulses. Thus, in the subsequent discussion we only present comparisons with the finite deflection solution and neglect the small deflection analysis.

4.2 Sandwich Plates. Comparisons of dynamic finite element simulations and analytical predictions have been performed on sandwich plates made from the reference materials specified above. The comparisons between the analytical and FE predictions are carried out in two stages. First, for a fixed impulse, the response of the sandwich plate is investigated as function of the plate geometry and second, the response of a sandwich plate with a representative geometry is studied for varying levels of impulse.

For the purposes of selecting appropriate sandwich plate geometries for the FE calculations, we plot a design chart for sandwich plates subjected to a normalized impulse $\bar{I}=10^{-3}$, with an assumed face-sheet material ductility $\epsilon_f=0.2$. The design chart shown in Fig. 4 has been constructed using the analytical model with the circumscribing yield locus. Contours of the maximum normalized central deflection $\bar{w}$ of the inner face of the sandwich plates along with the regime of tensile failure of the face-sheet are shown on the chart. Twelve plate geometries in the range $0.03\leqslant\bar{h}\leqslant0.3$ and $0.01\leqslant\bar{c}\leqslant0.06$ (as indicated in Fig. 4) are selected for the FE calculations. This range of plate geometries is representative of practical plate geometries, and the analytic predictions for the central displacements of the inner face of the sandwich plate are in the range $0.01\leqslant\bar{w}\leqslant0.2$.

Comparisons of the FE and analytical predictions (inscribing yield locus) for the central deflection $\bar{w}$ of the inner face-sheet as functions of $\bar{h}$ are shown in Fig. 5(*a*) for $\bar{c}=0.03$ and $\bar{c}=0.04$, subject to $\bar{I}=10^{-3}$. Similar to the monolithic beam case, in the FE simulations w is defined as the peak value of the deflection versus time trace. Figure 5(*b*) shows comparisons of the analytical

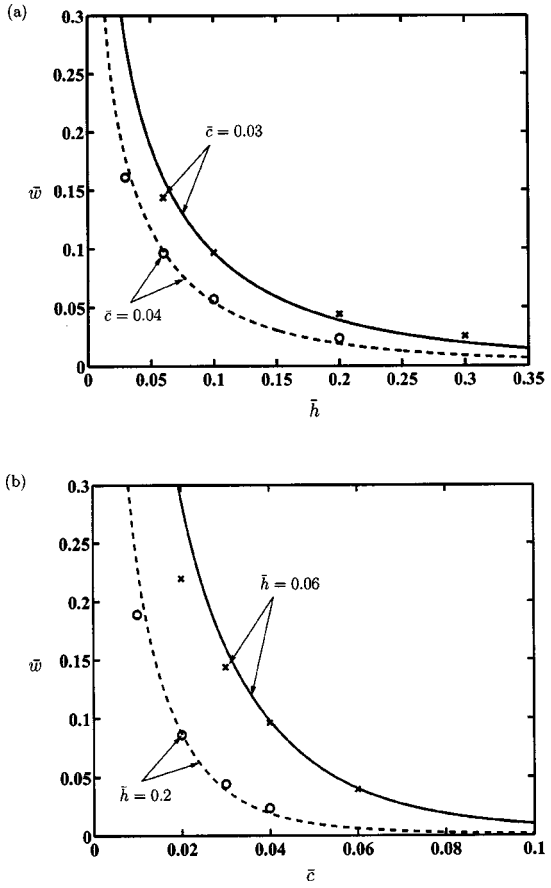


Fig. 5 Analytical and FE predictions of the maximum central deflection $\bar{w}$ of the inner face-sheet of sandwich plates with reference material properties subjected to a normalized impulse $\bar{I}=10^{-3}$. (*a*) $\bar{w}$ as a function of $\bar{h}$ for two values of $\bar{c}$. (*b*) $\bar{w}$ as a function of $\bar{c}$ for two values of $\bar{h}$.

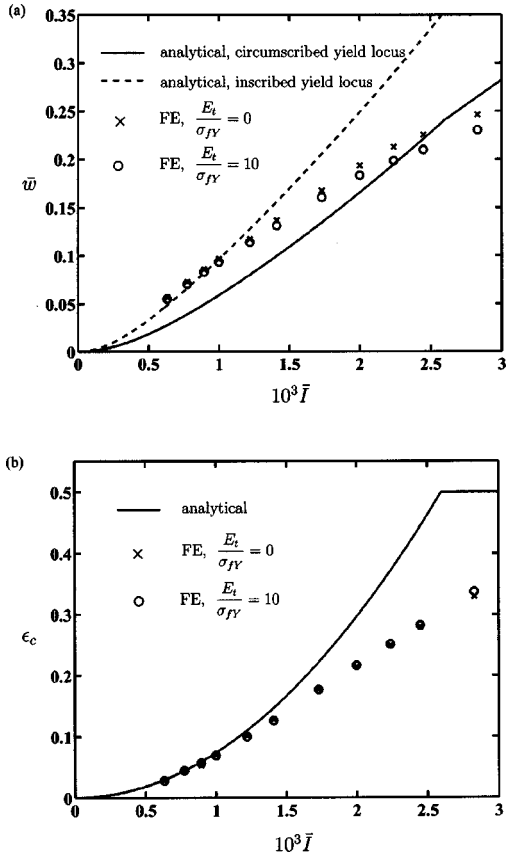


Fig. 6 Analytical and FE predictions of the (*a*) maximum central deflection $\bar{w}$ of the inner face-sheet and (*b*) core compression ϵ_c as a function of the applied impulse for sandwich plates. $\bar{c}=0.03$ and $\bar{h}=0.1$ and the sandwich plate is made from the reference core material, with both ideally plastic and strain hardening face-sheets.

and FE predictions of $\bar{w}$ versus $\bar{c}$ for plates with $\bar{h}=0.06$ and 0.2, again for the fixed impulse $\bar{I}=10^{-3}$. In all of the above cases good agreement is seen between the analytical and FE predictions, with the discrepancy in w between the analytical and FE predictions not exceeding 5%. As in the monolithic plate case, the analytical model employing the circumscribing yield locus underpredicts the deflections.

Next consider a representative sandwich plate of geometry, $\bar{c}=0.03$ and $\bar{h}=0.1$, subject to impulses in the range $7.5\times10^{-4} \leqslant \bar{I} \leqslant 3.2\times10^{-3}$. A comparison of the FE and analytical predictions of the maximum deflection $\bar{w}$ and core compression ϵ_c versus $\bar{I}$ are shown in Figs. 6(*a*) and 6(*b*), respectively. In the FE simulations, the strain ϵ_c is defined as the final through thickness nominal strain at the center of the plate. The choice of an inscribing yield surface for the analytical model leads to good agreement with the finite element predictions at low impulses, while at higher impulses the circumscribing yield surface appears to give better predictions. Figure 6(*b*) shows that the analytical calculation also substantially overpredicts the core compression in the high impulse domain. Similar to the sandwich beam case analyzed in Qiu et al. [7], these discrepancies can be rationalized by recalling that the analytical model neglects the reduction in momentum due to an impulse provided by the supports in the core compression phase. With increasing impulse this assumption is no longer valid as the higher core compression gives rise to significant stretching of the outer face-sheet at the supports and thus to a loss in momentum. This effect is not accounted for in the analytical model and consequently the analytical model overpredicts the deflections and core compression at high values of impulse.

In the FE simulations, the structural response time T is defined as the time taken to reach the maximum deflection and the core compression time T_c is defined as the time taken to first attain the final through thickness strain ϵ_c in the core. Comparisons of the analytical and FE predictions of the normalized structural response time $\bar{T}$ and the core compression time $\bar{T}_c$ as functions of $\bar{I}$ are shown in Fig. 7 for the sandwich plate with $\bar{c}=0.03$ and $\bar{h}=0.1$. Good agreement between the analytical and FE predictions is seen for the core compression time and, similar to the case of the monolithic plate, the inscribing yield locus model is in good agreement with the FE predictions of the structural response time. The normalized core compression time $\bar{T}_c$ is at least an order of magnitude smaller than the structural response time $\bar{T}$; this supports the assumption of a separation of time scales for the core compression phase and the plate bending and stretching phase in the analytical model.

4.2.1 Effect of Strain Hardening Upon the Dynamic Response of Sandwich Plates. The analytical models discussed in Section 2 and the FE calculations detailed above, both assume ideally plastic face-sheet materials. On the other hand, structural alloys, which are expected to be employed in shock resistant sandwich construction, can have a strong strain hardening response. The effect of strain hardening of the face-sheet material on the sandwich plate response is investigated here by suitable modifications of the FE model. The face-sheet material is assumed to be made

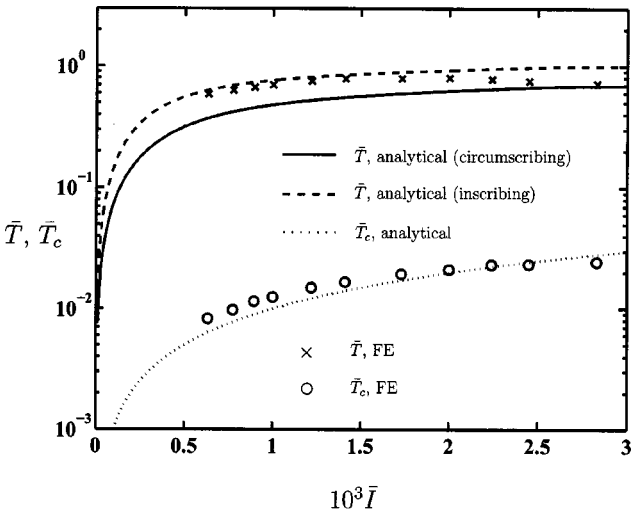


Fig. 7 A comparison between analytical and FE predictions of the structural response time $\bar{T}$ and core compression time $\bar{T}_c$, as a function of the applied impulse, for sandwich plates with $\bar{c}=0.03$ and $\bar{h}=0.1$ made from the reference materials

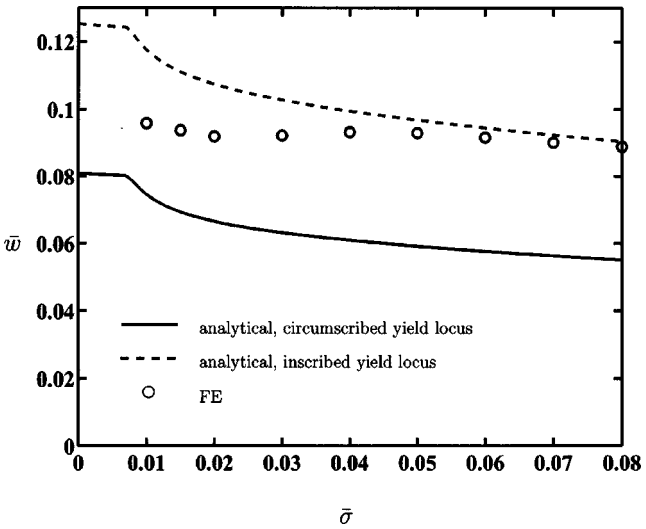


Fig. 8 A comparison between the analytical and FE predictions of the maximum central deflection $\bar{w}$ of the inner face-sheet of sandwich plates with $c=0.03$ and $\bar{h}=0:1$, subject to a normalized impulse $I=10^{-3}$ as a function of the normalized core strength $\bar{\sigma}$

from a elastic plastic material with yield stress and strain $\sigma_{fY}=500$ MPa and $\epsilon_{fY}=0.2\%$, respectively, and a density $\rho_f=8000\ \mathrm{kgm}^{-3}$. The strain hardening response is assumed to be linear with a tangent modulus $E_t/\sigma_{fY}=10$; this high rate of strain hardening is representative of the AL6XN stainless steel. The core properties are unchanged from the reference material case.

Consider a sandwich plate with core made from the reference material and elastic strain hardening plastic face-sheets, with geometry $\bar{c}=0.03$ and $\bar{h}=0.1$. The normalized maximum deflection $\bar{w}$ and core compression ϵ_c are plotted against $\bar{I}$ in Figs. 6(*a*) and 6(*b*), respectively, along with the deflections and core compressions of the sandwich plates made from the reference materials (with elastic-ideally plastic face-sheets). The strain hardening response of the face-sheets has only a small effect upon the deflection and core compression of the sandwich plate. This can be rationalised by recalling that the circumferential strain in the face-sheets is $\epsilon_m \approx 0.5\bar{w}^2 \approx 4.5\%$ for $\bar{w}\approx 0.3$. This level of straining does not increase the yield strength of the face-sheet material appreciably for the strain hardening considered here and hence the response is only mildly sensitive to the strain hardening behavior of the face-sheets. This conclusion should be moderated for the case of annealed face-sheets for which the flow strength at a uniaxial strain of 4.5% may be significantly above the yield strength.

4.2.2 Effect of Core Strength Upon the Dynamic Response of Sandwich Plates. In the calculations detailed above, the core strength was held constant. Here we investigate the effect of core strength on the sandwich plate response. Results are presented for sandwich plates of geometry $\bar{c}=0.03$ and $\bar{h}=0.1$, subjected to a normalized impulse $\bar{I}=10^{-3}$. Other than the core strength, the material properties of the sandwich plates were unchanged from the reference material properties. The normalized core strength $\bar{\sigma}$ was varied from 0.01 to 0.08, with a densification strain ϵ_D held fixed at 0.5; cores weaker than $\bar{\sigma}=0.01$ were not considered as numerical difficulties were encountered in such cases.

The maximum normalized deflection of the inner face of the sandwich plate $\bar{w}$ is plotted against the normalized core strength $\bar{\sigma}$ in Fig. 8. The FE results indicate that $\bar{w}$ is relatively insensitive to the core strength. Analytical predictions of $\bar{w}$ employing the inscribing and circumscribing yield surfaces are included in Fig. 8; the analytical model employing the inscribing yield surface agrees reasonably well with the FE predictions for plates with the high core strength but overpredicts the deflection of the plates with the weaker cores. Time histories of the plastic dissipation of the entire sandwich plate and of the core alone, each normalized by the initial kinetic energy of the outer face-sheet of the sandwich plate are shown in Fig. 9(*a*). These curves reveal two stages of deformation. In the first stage, plastic dissipation occurs primarily in the core, with the outer face-sheet approaching the inner face; at the end of this stage both face-sheets are moving at approximately the same velocity. Subsequently, plastic dissipation occurs primarily within the face-sheets, with the dissipation in the core increasing only gradually with time. It is worth noting that the plastic dissipation in the core at the end of the first stage is nearly independent of the core strength. Further, this stage lasts longer for the weaker cores. Consequently, the core compression phase overlaps with that for the face-sheet deformation for the choice of a sandwich plate with a weak core.

Finite element predictions of the plastic dissipation at the end of the first stage of the deformation (i.e., the plastic dissipation corresponding up to the ''knee'' in the plastic dissipation versus times curves of Fig. 9(*a*)) are shown in Fig. 9(*b*) as a function of the mass ratio $\bar{m}=\rho_c c/(\rho_f h)$, for the choices of core strength $\bar{\sigma}=0.04$ and $\bar{\sigma}=0.01$. These calculations were conducted on plates with the above geometry subject to a normalized impulse $\bar{I}=10^{-3}$. The ratio $\bar{m}$ was varied by changing the density of the core material from 80 kgm^{-3} to 1600 kgm^{-3}. The figure reveals that the plastic dissipation at the end of the core compression stage is independent of the core strength and increases with $\bar{m}$, in excellent agreement with the analytical predictions, Eq. (4).

4.2.3 Comparison of the Dynamic Response of Clamped Sandwich Plates and Beams. Comparisons between FE and analytical predictions of the impulsive response of clamped sandwich beams have already been presented by Qiu et al. [7]. Here we compare the analytical and FE predictions of the deflection w of sandwich plates with these existing results for sandwich beams.

Consider a clamped sandwich beam of span $2L$ comprising two identical face-sheets of thickness h and a core of thickness c made from the reference materials described above. Qiu et al. [7] presented FE results of the maximum normalized midspan defection $\bar{w}\equiv w/L$ of such sandwich beams with geometry $\bar{c}\equiv c/L=0.03$ and $\bar{h}\equiv h/c=0.1$ as a function of the applied normalized impulse $\bar{I}\equiv I/(L\sqrt{\rho_f\sigma_{fY}})$. These results are plotted in Fig. 10 along with the corresponding sandwich plate results. We note that the FE

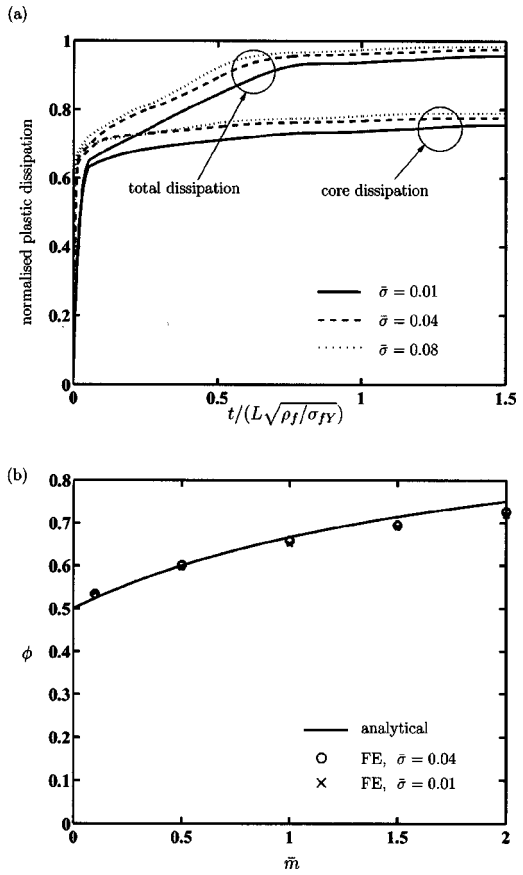


Fig. 9 (*a*) FE predictions of the time histories of the normalized plastic dissipation in sandwich plates for three selected core strengths. (*b*) Ratio ϕ of the plastic dissipation in the core compression stage to the initial kinetic energy of the outer face as a function of the mass ratio $\bar{m}$ for two selected core strengths. The sandwich plates in both cases have geometry $\bar{c}=0.03$ and $\bar{h}=0.1$ and are subjected to an impulse $\bar{I}=10^{-3}$.

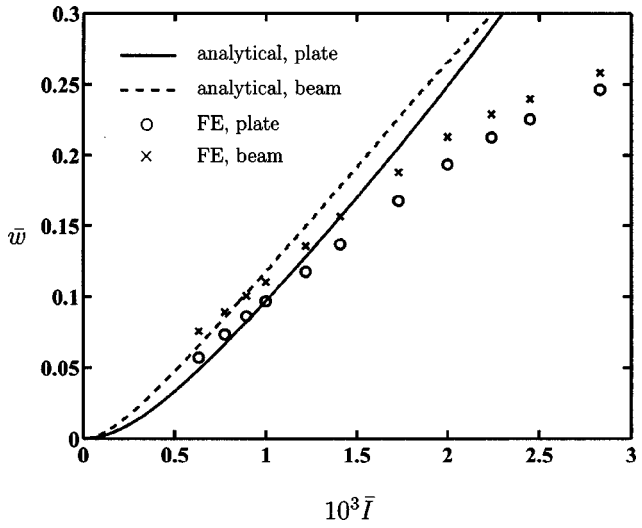


Fig. 10 Analytical and FE predictions of the maximum central deflections $\bar{w}$ of the inner face-sheet for clamped sandwich plates and beams, as a function of the applied impulse. Both the sandwich plates and beams have a geometry $\bar{c}=0.03$ and $\bar{h}=0.1$, and are made from the reference materials.

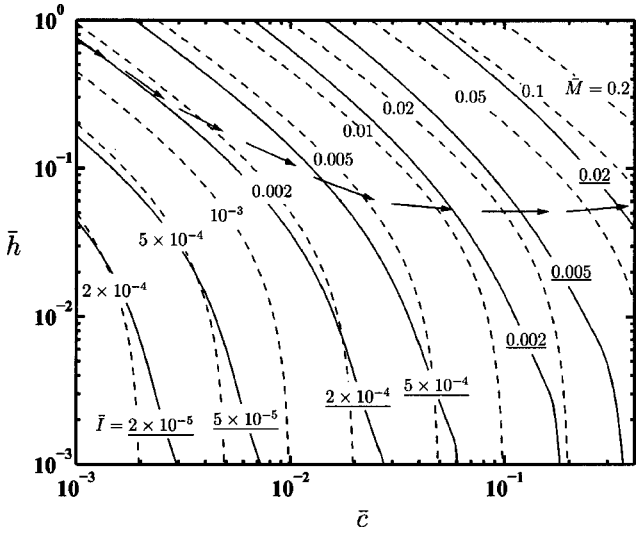


Fig. 11 Design chart for a clamped sandwich plate made from the reference materials for a fixed maximum central deflection of the inner face $\bar{w}=0.1$. Contours of the applied impulse $\bar{I}$ and nondimensional mass $\bar{M}$ are displayed. The underlined values denote the nondimensional impulse values while the arrows trace the path of the optimal designs with increasing $\bar{M}$.

predictions of the maximum deflections of the clamped sandwich beams and plates as functions of the applied normalized impulse are approximately equal when the half-span L of the sandwich beam is equated to the radius R of the sandwich plate. The analytical predictions (employing the inscribing yield locus) of the deflections of the beams [7] and plates are included in Fig. 10. In line with the FE predictions, the analytical predictions for the beams and plates are approximately equal.

5 Optimal Design of Sandwich Plates Subject to Shock Loading

In the preceding sections we have demonstrated that the finite deflection analytical formulas for the response of the clamped sandwich plates are in reasonable agreement with FE calculations. We now employ these analytical finite deflection formulas to determine the optimal designs of sandwich plates that maximise the resistance of a sandwich plate of given mass to shock loading subject to the constraint of a maximum allowable inner face deflection. The optimization is conducted by assuming that the entire shock impulse is transmitted to the sandwich plate ($\zeta=1$). This is representative of shock loading in air where the acoustic impedance of air is much less than that of the steel outer face-sheet of the sandwich plate as detailed in the Stage I analysis of Section 2.

To help with this optimization, it is instructive to construct a design chart relating the sandwich plate geometry to the shock impulse for a specified deflection. Such a design chart with axes $\bar{c}$ and $\bar{h}$ is plotted in Fig. 11 for a normalized deflection $\bar{w}=0.1$ of the inner face of the sandwich plate by employing the circumscribing yield locus analytical expressions. The chart is plotted for sandwich plates with reference materials properties, i.e., a core of relative density $\bar{\rho}=0.1$ and strength specified by (26). Contours of the nondimensional mass $\bar{M}$ of the sandwich plates have been added to Fig. 11, where

$$\bar{M} \equiv \frac{M}{\pi R^3 \rho_f} = 2\bar{h}\bar{c} + \bar{c}\bar{\rho}, \tag{27}$$

and M is the mass of the sandwich plate. The arrows in Fig. 11 trace the trajectory of $(\bar{c},\bar{h})$ which maximizes $\bar{I}$ for a given $\bar{M}$ with increasing $\bar{M}$.

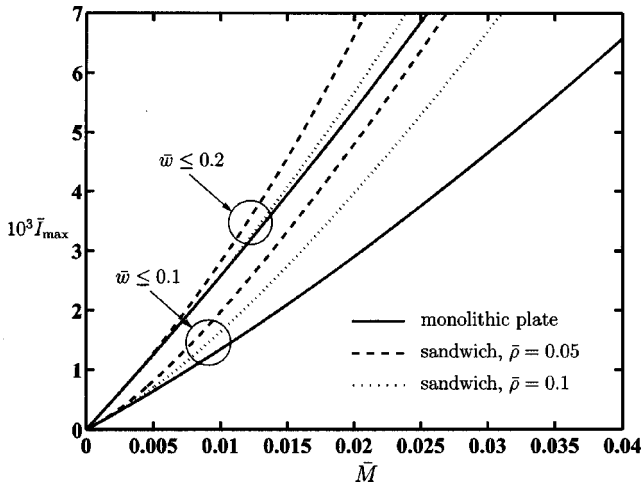


Fig. 12 A comparison of the maximum shock impulse sustained by monolithic plates and by optimal designs of the sandwich plates subject to the constraints $\bar{w}<0{:}1$ and $\bar{w}\leqslant 0{:}2$ for two relative densities $\bar{\rho}$ of the core material

The maximum normalized impulse $\bar{I}_{max}$ sustained by these optimal designs is plotted in Fig. 12 as a function of the nondimensional mass $\bar{M}$. Closed-form expressions for the optimal designs do not exist and hence only numerical results are presented here. Also included in Fig. 12 are the maximum impulses sustained by sandwich plates comprising a core with $\bar{\rho}=0.05$ and strength again specified by (26): Sandwich plates comprising the less dense core have a superior performance. The maximum impulses sustained by the sandwich plates with the constraint on the allowable deflection of the inner face relaxed to $\bar{w}\leqslant 0.2$ are also included in Fig. 12 for the sandwich plates with the $\bar{\rho}=0.05$ and 0.1 cores: the sandwich plates sustain about a 40% higher impulse with the constraint $\bar{w}\leqslant 0.2$ as compared to $\bar{w}\leqslant 0.1$.

For comparison purposes the impulses sustained by monolithic plates made from the same material as the sandwich plate facesheets and subject to the constraints $\bar{w}\leqslant 0.1$ and $\bar{w}\leqslant 0.2$ are included in Fig. 12 with the choice of the circumscribing yield locus. In the monolithic case the nondimensional mass of the monolithic plates is

$$\bar{M}\equiv \frac{M}{\pi R^3 \rho_f}=\frac{H}{R}, \tag{28}$$

in terms of the mass M of the monolithic plate. In all cases, the optimal sandwich plates out-perform monolithic plates. However, the performance gain obtained by employing sandwich construction reduces if the maximum allowable deflections $\bar{w}$ are larger: at large deflections the resistance of the plates is primarily due to the stretching action and the performance advantage of the sandwich plates in terms of their high bending resistance plays a much smaller role. It is worth mentioning here that the results will not qualitatively change if the optimisations were performed using the inscribing yield locus instead of the circumscribing yield locus.

6 Conclusions

An analytical model for the response of clamped circular plates subject to shock loading in air and underwater has been derived using the framework proposed by Fleck and Deshpande [6] for clamped sandwich beams. The predictions of the analytical model have been compared with dynamic FE results with the effect of fluid-structure interaction neglected and the analytical model used to determine optimal designs of the sandwich plates. The main findings include:

1. FE calculations demonstrate that the time scale for core compression separates from the time scale for plate bending/stretching of the sandwich plate, as assumed in the analytical model.
2. the analytical model employing the inscribing yield locus agrees well with the FE predictions at small deflections while the FE results are in better agreement with the analytical predictions employing the circumscribing yield locus at larger deflections.
3. for realistic levels of plates deflections, the presence of strain hardening representative that for most structural alloys has a negligible influence on the sandwich plate response.
4. both the FE calculations and the analytical model predict that the compressive strength of the core has only a limited influence on the sandwich plate response.
5. optimal designs of sandwich plates sustain larger shock impulses than monolithic plates of the same mass assuming that the face-sheets of the sandwich plate are made from the same solid as that of the monolithic plate.

Acknowledgments

The authors are grateful to ONR for their financial support through US-ONR IFO grant number N00014-03-1-0283 on the The Science and Design of Blast Resistant Sandwich Structures. We are pleased to acknowledge Profs. M. F. Ashby, T. Belytschko, A. G. Evans, and J. W. Hutchinson for many insightful discussions during the course of this work.

References

[1] Wang, A. J., and Hopkins, H. G., 1954, "On the Plastic Deformation of Built-in-Circular Plates Under Impulsive Load," J. Mech. Phys. Solids, **3**, pp. 22–37.
[2] Symmonds, P. S., 1954, "Large Plastic Deformations of Beams Under Blast Type Loading," *Proceedings of the Second US National Congress of Applied Mechanics*, pp. 505–515.
[3] Jones, N., 1971, "A Theoretical Study of the Dynamic Plastic Behavior of Beams and Plates With Finite Deflections," Int. J. Solids Struct., **7**, pp. 1007–1029.
[4] Jones, N., 1989, *Structural Impact*, Cambridge University Press, Cambridge, UK.
[5] Xue, Z., and Hutchinson, J. W., 2003, "A Preliminary Assessment of Sandwich Plates Subject to Blast Loads," Int. J. Mech. Sci., **45**, pp. 687–705.
[6] Fleck, N. A., and Deshpande, V. S., 2004, "The Resistance of Clamped Sandwich Beams to Shock Loading," ASME J. Appl. Mech., **71**, 386–401.
[7] Qiu, X., Deshpande, V. S., and Fleck, N. A., 2003, "Finite Element Analysis of the Dynamic Response of Clamped Sandwich Beams Subject to Shock Loading," Eur. J. Mech. A/Solids, **22**, 801–814.
[8] Taylor, G. I., 1963, *The Scientific Papers of G I Taylor, Vol III*, pages 287–303. Cambridge University Press, Cambridge, UK, pp. 287–303 (The Pressure and Impulse of Submarine Explosion Waves on Plates, 1941).
[9] Cole, R. H., 1948, *Underwater Explosions*, Princeton University Press, Princeton, NJ.
[10] Jones, N., and Gomes de Oliveira, J., 1980, "Dynamic Plastic Response of Circular Plates With Transverse Shear and Rotary Inertia," ASME J. Appl. Mech., **47**, pp. 27–34.
[11] Perrone, N., and Bhadra, P., 1984, "Simplified Large Deflection Mode Solutions for Impulsively Loaded, Viscoplastic, Circular Membranes," ASME J. Appl. Mech., **51**, pp. 505–509.
[12] Deshpande, V. S., and Fleck, N. A., 2000, "Isotropic Constitutive Models for Metallic Foams," J. Mech. Phys. Solids, **48**(6–7), pp. 1253–1283.
[13] Deshpande, V. S., Fleck, N. A., and Ashby, M. F., 2001, "Effective Properties of the Octet-Truss Lattice Material," J. Mech. Phys. Solids, **49**(8), pp. 1747–1769.

A Continuum Theory That Couples Creep and Self-Diffusion

Z. Suo
Division of Engineering and Applied Sciences,
Harvard University,
Cambridge, MA 02138
e-mail: suo@deas.harvard.edu

In a single-component material, a chemical potential gradient or a wind force drives self-diffusion. If the self-diffusion flux has a divergence, the material deforms. We formulate a continuum theory to be consistent with this kinematic constraint. When the diffusion flux is divergence-free, the theory decouples into Stokes's theory for creep and Herring's theory for self-diffusion. A length emerges from the coupled theory to characterize the relative rate of self-diffusion and creep. For a flow in a film driven by a stress gradient, creep dominates in thick films, and self-diffusion dominates in thin films. Depending on the film thickness, either stress-driven creep or stress-driven diffusion prevails to counterbalance electromigration. The transition occurs when the film thickness is comparable to the characteristic length of the material. [DOI: 10.1115/1.1781176]

1 Introduction

Self-diffusion can generate stress in a single-component material. During deposition, for example, a thin film sometimes develops a compressive stress. One possible mechanism has to do with the injection of atoms into the film, [1]. Impinging atoms may not have enough time to find equilibrium positions on the film surface, and may diffuse into the film. Oxidation leads to analogous phenomena. For some materials, during oxidation, atoms may emit from the materials, causing tension in the materials, [2]. For other materials, notably silicon, atoms may inject into the materials, causing compression in the materials, [3]. Electromigration provides yet another compelling example. The conduction electrons motivate atoms to diffuse, generating tension upstream and compression downstream, [4].

The stress generated in the material depends on the deformation mechanism of the material. Only elastic property enters the consideration if inelastic deformation (i.e., creep) is either so slow as to be negligible, or so fast as to relax the stress field locally to a hydrostatic state. For electromigration along a thin line, encapsulated in a stiff dielectric, it was thought that local stress relaxes to a hydrostatic state long before diffusion along the line reaches a steady state, [5,6]. Experiments, however, have shown large deviatoric stresses, [7]. Indeed, the initial discovery of electromigration-induced stress was made in a wide aluminum film, which could only sustain in-plane stresses, [4].

This paper formulates a theory to couple self-diffusion and creep in single-component materials. The new theory will contain Stokes's creep and Herring's diffusion as special cases. Stokes's creep, as formulated in fluid mechanics, describes a velocity field and a pressure field; it neglects self-diffusion. Herring's theory, [8], for self-diffusion is in terms of the chemical potential, a scalar; it makes no attempt to equilibrate stress tensor field.

Our theory parallels that of nonreciprocal diffusion in multicomponent solid solutions (i.e., the Kirkendall effect) due to Darken [9] and Stephenson [10], and extends our previous one-dimensional theory, [11]. The theory rests on a kinematic constraint: the divergence in the self-diffusion flux must be accommodated by deformation. The remainder of this section recalls a few historic highlights of the mechanistic picture of creep and self-diffusion. Sections 2–4 describe the kinematics, energetics, and kinetics of the theory. Section 5 gives the coupled partial differential equations for the velocity field and the chemical potential field, and identifies the characteristic length in the theory. Sections 6 discusses examples of flows driven by stress gradient, wind force, and atomic injection or emission. Stress gradient-driven channel flow is dominated by creep in thick channels, and by self-diffusion in thin channels. Section 7 discusses an anisotropic rule to place diffusion flux divergence as strain rates in various directions.

That creep and self-diffusion in some materials result from the same atomistic process has been known for a long time. In a liquid, self-diffusion and creep are different macroscopic manifestations of the same microscopic fact: Molecules change neighbors readily in the liquid. Einstein [12] related the Brownian movement of a macroscopic particle in a liquid to the viscosity of the liquid. The Stokes-Einstein formula, derived by Einstein using Stokes's continuum solution, has since been applied to diffusion of molecules in liquids, including self-diffusion.

Nabarro [13] and Herring [8] related creep in a polycrystal to self-diffusion mediated by the motion of vacancies. By itself, the motion of vacancies does not change the crystal shape, but the creation and annihilation of vacancies at the grain boundaries do. Consequently, creep in the polycrystal is fast when the self-diffusivity is high and the grains are small. When the grain size approaches the molecular dimension, the Nabarro-Herring formula for polycrystals reduces to the Stokes-Einstein formula for liquids, except for a numerical factor. Similar comments apply to the Coble creep, [14], mediated by atoms diffusing on grain boundaries.

Needleman and Rice [15] formulated a theory for polycrystals, where atoms diffuse on grain boundaries and creep in grains. Here creep can result from the motion of dislocations. The two processes, self-diffusion and creep, occur in different places, but couple through a kinematic constraint. For two grains meeting at a grain boundary, the creep in the two grains accommodates the divergence of the diffusion flux on the grain boundary.

Our theory neglects the microstructure. Regardless of the specific microstructure, when a wind force motivates atoms to diffuse, the material must deform to accommodate the divergence of the self-diffusion flux. The main advantage of the theory is that simple and enlightening solutions may be obtained for coupled problems. The main drawback is that the theory may lead to wrong predictions at the size scale approaching or smaller than the microstructural feature size. The new theory can be applied, with virtues and vices of a continuum theory, to complex materials, and to crystalline materials when atoms also diffuse in grains, among other situations for which the Needleman-Rice theory is not 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, February 4, 2004; final revision, March 17, 2004. Associate Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

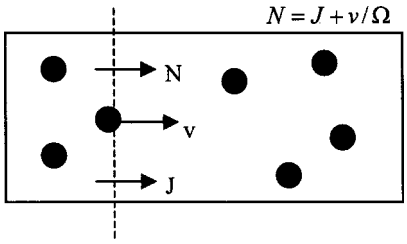


Fig. 1 The marker velocity v defines the convection flux, v/Ω. We can also independently measure the net atomic flux N. The atomic flux in excess of the convection flux defines the diffusion flux J.

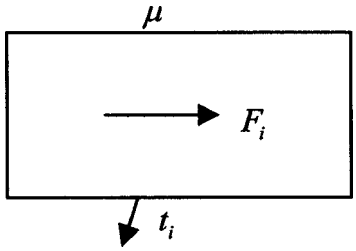


Fig. 2 The material is subject to three types of load: the wind force F_i in the volume, the traction t_i on the surface, and the chemical potential on the surface

tended. In particular, we will use the new theory to analyze electromigration-induced creep in Newtonian liquids.

2 Kinematics

When the two rate processes occur in separate places, creep in grains and diffusion on grain boundaries, there is no ambiguity about their distinct contributions to mass transport. When diffusion and creep occur in the same continuum space, how can their contributions be distinguished? We must give operational definitions of creep and self-diffusion without referring to the microstructure. Following Darken [9], we imagine that markers are dispersed throughout the material (Fig. 1). The markers in the material are analogous to leaves on a river. The flow of water carries the leaves, but is unaffected by their presence. The motion of the markers defines convection. The atomic flux in excess of convection defines diffusion. The markers should be small compared to the size scale in the flow of interest, but large compared to the atomic dimension so that the markers themselves diffuse negligibly.

We adopt the Eulerian approach. Let (x_1,x_2,x_3) be the coordinates of a fixed space. The field $v_i(x_1,x_2,x_3,t)$ is the velocity vector of the marker at position (x_1,x_2,x_3) at time t. Let Ω be the volume per atom in the body. Imagine a plane fixed in space and perpendicular to the axis x_i. The convection flux, v_i/Ω, is the number of atoms moving with the marker across the plane, per unit area per unit time. The net atomic flux, N_i, is the number of atoms across the plane, per unit area and per unit time. We can independently measure the marker velocity and the net atomic flux. The difference between the two fluxes defines the self-diffusion flux J_i, namely,

$$N_i = J_i + v_i/\Omega. \tag{1}$$

The net flux is the sum of the diffusion flux and the convection flux.

To demonstrate the new features of the theory with minimum complication, we neglect elasticity. Following Balluffi [16], we also neglect strains due to the space occupied by point defects such as vacancies or free volumes; enough dislocations climb or other defects move to maintain the point defects close to equilibrium concentrations, which are typically small. Consequently, a fixed volume contains a constant number of atoms at all time. The volume per atom, Ω, is constant. The net atomic flux is divergence-free, $N_{k,k}=0$, so that

$$v_{k,k} = -\Omega J_{k,k}. \tag{2}$$

A repeated subscript implies summation over 1, 2, and 3; a comma before a subscript indicates partial differentiation. Equation (2) has a clear interpretation. Imagine a volume fixed in space. When $J_{k,k}>0$, atoms diffuse out the volume; for the volume to maintain a constant number of atoms, convection must carry atoms into the volume, so that the markers converge. The opposite is true when $J_{k,k}<0$. In this theory, the material is incompressible, but the marker velocity has a divergence to compensate for the divergence in the diffusion flux. It is this kinematic constraint, Eq. (2), that couples creep and self-diffusion.

Markers at different locations may move at different velocities. When two markers move away from each other, atoms have to be inserted in the space between them. When two markers move toward each other, atoms have to be removed from the space between them. The gradient of the marker velocity field defines the strain-rate tensor:

$$d_{ij} = \tfrac{1}{2}(v_{i,j} + v_{j,i}). \tag{3}$$

If m_i is the unit vector pointing from one marker to the other, then $m_i d_{ij} m_j/\Omega$ is the number of atoms inserted or removed per unit time, per unit area normal to and per unit distance along the direction m_i. The strain rate is the sum of that due to diffusion, d_{ij}^D, and that due to creep, d_{ij}^C:

$$d_{ij} = d_{ij}^C + d_{ij}^D. \tag{4}$$

As suggested by Eq. (2), the divergence in the diffusion flux, $J_{k,k}$, causes the divergence in the marker velocity. We assume that the divergence in the diffusion flux causes an equal strain rate in all three directions:

$$d_{ij}^D = -\frac{\Omega}{3} J_{k,k}\delta_{ij}, \tag{5}$$

where $\delta_{ij}=1$ when $i=j$, and $\delta_{ij}=0$ when $i\neq j$.

A combination of Eqs. (2)–(5) gives the creep strain rates in terms of the marker velocity field:

$$d_{ij}^C = \tfrac{1}{2}(v_{i,j} + v_{j,i}) - \tfrac{1}{3} v_{k,k}\delta_{ij}. \tag{6}$$

The creep strain-rate tensor is symmetric and traceless.

3 Energetics

Figure 2 illustrates three types of load on the material. We identify them by the ways they supply power to the material. The wind force F_i supplies power $\int F_i J_i dV$, with the integral over the volume of the material. The traction t_i supplies power $\int t_i v_i dS$, with the integral over the material surface where the traction is prescribed. The chemical potential μ is the free energy difference between an atom on the material surface and an atom in a reference body (a bulk under no stress). The chemical potential is a scalar field defined on the material surface, in the same spirit as the traction is a vector field defined on the surface. We assume local equilibrium: The chemical potential of atoms in the material immediately beneath the surface equals that of atoms on the surface. Let the unit vector n_i be normal to the surface and point to the outside of the material, and $J_i n_i$ be the flux at which the atoms diffuse out the material. The chemical potential acts on atoms in the same way as the voltage acts on electrons. When atoms diffuse out of the material and join the reference body (i.e., when atoms move across the chemical potential), the chemical potential supplies power $-\int \mu J_i n_i dS$, with the integral over the material surface where the chemical potential is prescribed.

We next identify driving forces for the two rate processes, following an approach often used in constructing continuum theories of multiple rate processes and thermodynamic forces, e.g., [17]. We state the principle of virtual power in the form of the balance between power dissipation and power supply:

$$\int (s_{ij}d^C_{ij}+f_iJ_i)dV+\int \lambda(v_{k,k}+\Omega J_{k,k})dV = \int F_iJ_idV+\int t_iv_idS-\int \mu J_in_idS. \quad (7)$$

On the right-hand side of Eq. (7) are the three modes of power supply discussed above. On the left-hand side, the first integral contains two modes of power dissipation. Equation (7) defines the creep driving force, s_{ij}, as the power-conjugate of the creep strain rate. Because d^C_{ij} is a symmetric and traceless tensor, without loss of generality, we require that s_{ij} be a symmetric and traceless tensor. Eq. (7) defines the diffusion driving force, f_i, as the power-conjugate of the diffusion flux. The second integral enforces the kinematic constraint, Eq. (2), with λ as the Lagrange multiplier.

Because s_{ij} is a symmetric and traceless tensor, and d^C_{ij} relates to the marker velocity field by Eq. (6), we confirm that $s_{ij}d^C_{ij}=s_{ij}v_{i,j}$. Using the divergence theorem, we can express Eq. (7) as

$$-\int (s_{ij}+\lambda\delta_{ij})_{,j}v_idV+\int ((s_{ij}+\lambda\delta_{ij})n_j-t_i)v_idS+\int (f_i-F_i-\Omega\lambda_{,i})J_idV+\int (\Omega\lambda+\mu)J_in_idS=0. \quad (8)$$

This equation holds for arbitrary marker velocity and diffusion flux field, with no constraint. Consequently, the power balance requires that

$$(s_{ij}+\lambda\delta_{ij})_{,j}=0, \quad \text{in volume} \quad (9)$$

$$(s_{ij}+\lambda\delta_{ij})n_j=t_i, \quad \text{on surface} \quad (10)$$

$$f_i=F_i+\Omega\lambda_{,i}, \quad \text{in volume} \quad (11)$$

$$\mu=-\Omega\lambda, \quad \text{on surface.} \quad (12)$$

Equation (9) and (10) recover force balance equations. In familiar terms, the creep driving force s_{ij} is the deviatoric stress tensor, the Lagrange multiplier λ is the mean stress σ, and the combination $\sigma_{ij}=s_{ij}+\sigma\delta_{ij}$ is the Cauchy stress tensor. Equations (11) and (12) recover Herring's equations for the diffusion driving force, [4,8]. The quantity $-\Omega\sigma$ is the free energy change associated with transferring an atom in the stress-free reference body to a point inside the material under the mean stress σ. In short, $-\Omega\sigma$ is the chemical potential inside the material. Its gradient, together with the wind force F_i, drives diffusion. To maintain local equilibrium, the chemical potential in the material just beneath the surface matches the prescribed value on the surface.

Following the established usage in mechanics, we intend the phrase "virtual power" to mean that Eq. (7) holds true provided all the kinematic relations are satisfied, and that no constitutive relations are assumed between the kinematic quantities and the force-like quantities. We could have as well followed an equivalent approach by invoking stress potential and strain rate potential, [18]. This paper considers isothermal phenomena. Were we interested in phenomena with nonuniform temperature fields, we would follow the practice of the nonequilibrium thermodynamics, working with the entropy production, [19].

4 Kinetics

Familiar isotropic kinetic laws are prescribed for diffusion and creep. The diffusion flux is proportional to the diffusion driving force:

$$J_i=\frac{Df_i}{\Omega kT}, \quad (13)$$

where D is the self-diffusion coefficient, k Boltzmann's constant, and T the temperature. The creep strain rate relates to the deviatoric stresses as

$$d^C_{ij}=\frac{s_{ij}}{2\eta}, \quad (14)$$

where η is the viscosity. For linear creep, η is constant. For nonlinear creep, a standard approach is to assume that η is a function of either the effective stress $\sigma_e=(3s_{ij}s_{ij}/2)^{1/2}$, or the effective creep strain rate $d^C_e=(2d^C_{ij}d^C_{ij}/3)^{1/2}$. The function $\eta(\sigma_e)$ or $\eta(d^C_e)$ is determined by fitting to the relation between stress and strain rate measured under a simple stress state.

5 Governing Equations and Characteristic Length

Inserting the creep law (14) and creep strain-rate expression (6) into the force balance Eq. (9), we obtain that

$$[\eta(v_{i,j}+v_{j,i}-\tfrac{2}{3}v_{k,k}\delta_{ij})]_{,j}+\sigma_{,i}=0. \quad (15)$$

Inserting the diffusion law (13) and diffusion driving force expression (11) into the kinematic constraint (2), we obtain

$$v_{k,k}=-\left[\frac{D}{kT}(F_k+\Omega\sigma_{,k})\right]_{,k}. \quad (16)$$

When the diffusion flux divergence vanishes, the marker velocity divergence also vanishes, $v_{k,k}=0$; Eq. (15) reduces to Stokes's equation for creep, and Eq. (16) reduces to Herring's equation for self-diffusion. In general, $v_{k,k}\neq 0$, and Eqs. (15) and (16) are four coupled partial differential equations that govern the marker velocity v_i and the mean stress σ. Each point on the material surface requires four boundary conditions: three conditions of either velocities or tractions, one condition of either chemical potential or the diffusion flux component normal to the surface.

The theory has a characteristic length. When D and η are constant, Eqs. (15) and (16) are linear. Let σ_0 be a representative stress scale in a boundary problem, and Λ be the length scale to be determined. Scale the stresses by σ_0, the velocities by $\Lambda\sigma_0/\eta$, the wind forces by $\Omega\sigma_0/\Lambda$, and the spatial coordinates by Λ. Equations (15) and (16) become dimensionless and parameter-free provided

$$\Lambda=\sqrt{D\eta\Omega/kT}. \quad (17)$$

The length characterizes the relative rate of diffusion and creep, and is independent of the scale of the stress.

For polycrystals, when creep is facilitated by diffusion, either through grains or along grain boundaries, the viscosity scales with the grain size d_g as $\eta=kTd_g^2/42D\Omega$, [20], so that the characteristic length scales with the grain size, $\Lambda=d_g/\sqrt{42}$. For simple liquids, the self-diffusivity is estimated by the Stokes-Einstein formula, [12], $D=kT/6\pi a\eta$, where a is atomic radius, so that the characteristic length scales with the atomic size, $\Lambda=\sqrt{\Omega/6\pi a}$. It is important to determine this length for more complex materials, such as amorphous metals and polymer melts.

If the creep data under the uniaxial tensile stress state fit the power law, $d^C_{11}=B\sigma^n_{11}$, where B and n are constants, the function η is given by

$$\eta(\sigma_e)=\frac{1}{3B\sigma_e^{n-1}}, \quad \text{or} \quad \eta(d^C_e)=\frac{1}{3B^{1/n}(d^C_e)^{1-1/n}}. \quad (18)$$

For power-law creep, with a constant diffusivity D, the solution to Eqs. (15) and (16) has a remarkable scaling structure. Let σ_* be a representative stress scale in a boundary problem, $d^C_*=B\sigma^n_*$, $\eta_*=\eta(\sigma_*)$, and Λ_* be a length scale. Scale the stress field by σ_*, the strain-rate field by d^C_*, the velocity field by $d^C_*\Lambda_*$, the

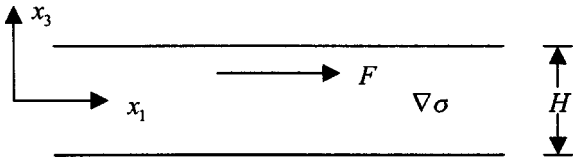


Fig. 3 A conductor film, sandwiched between dielectrics, is subject to an electron wind force and a stress gradient. The stress gradient can cause both a self-diffusion flux and a creep flow.

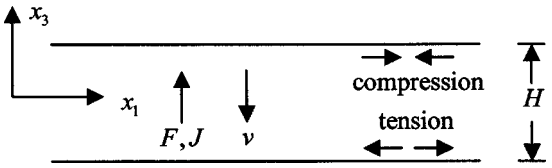


Fig. 4 A thin conductor subject to a through-thickness wind force. The diffision flux is in the same direction as the wind force, and the marker velocity is in the opposite direction. Tension is generated near one face of the film, and compression the other.

wind force field by $\Omega\sigma_*/\Lambda_*$, and the spatial coordinates by Λ_*. In terms of the dimensionless fields, the governing Eqs. (15)–(17) have only one parameter, the power index n, provided we identify the length Λ_* with Eq. (17), and replace η with η_* in the expression. For a power-law creep material, the length depends on the stress level σ_*. For polycrystals this length relates to a length L identified by Needleman and Rice [15], $\Lambda_*=\sqrt{L^3/3d_g}$. These authors also tabulated the experimental data for the length L for several metals.

6 Examples

6.1 Flow in a Film Driven by Stress Gradient and Electron Wind. In an interconnect line encapsulated in a dielectric, when the electron wind drives atoms to diffuse toward the anode, compression develops near the anode, and tension develops near the cathode. The stress gradient drives atoms to flow toward the cathode, in the direction opposite to the electron wind. It was discovered that the stress gradient could counter the electron wind, so that net mass flow vanished, [4]. This discovery has since become an effective means to avert electromigration failure; see recent reviews, [21,22]. In their original paper [4], Blech and Herring asserted that mass flow stops when the driving force for diffusion vanishes, namely, $F+\Omega\nabla\sigma=0$, where F is the electron wind force, and $\nabla\sigma$ the stress gradient. This assertion neglects a daily experience: The stress gradient also drives creep flow in a channel (e.g., in pumping water through a pipe). Given that the stress gradient can drive both a creep flow and a diffusion flow, will creep flow be also significant enough to counter electromigration?

To answer this question, consider a conductor film, thickness H, sandwiched between two dielectrics (Fig. 3). Let the axis x_3 be normal to the film, and the two faces of the film coincide with the planes $x_3=\pm H/2$. We will first analyze a steady flow subject to a constant electron wind force F and a constant stress gradient $\nabla\sigma$. For simplicity, we assume that both diffusivity and viscosity are constant.

In the steady flow, the only nonzero component of the marker velocity is in the flux direction, and varies in the thickness direction; that is, $v_2=v_3=0$ and $v_1=v_1(x_3)$. Consequently, the velocity field has no divergence, and convection and diffusion decouple. Of the pair of the governing equations, Eq. (16) is satisfied automatically, and Eq. (15) reduces to $\eta\partial^2 v_1/\partial x_3^2+\nabla\sigma=0$. This is an ordinary differential equation for the velocity profile $v_1(x_3)$. The gradient in the hydrostatic stress can induce a shear stress. Assuming the no-slip boundary condition at the conductor/dielectric interface, we obtain the familiar parabolic velocity profile: $v_1=((H/2)^2-x_3^2)\nabla\sigma/2\eta$.

The flow in the film has two contributions: the creep flow $Q^C=\int_{-H/2}^{+H/2}v_1 dx_3=H^3\nabla\sigma/12\eta$, and the diffusion flow $Q^D=HJ_1=H(D/\Omega kT)(F+\Omega\nabla\sigma)$. First consider flow under the stress gradient alone, in the absence of the wind force F. The ratio of creep flow to diffusion flow is

$$Q^C/Q^D=H^2/12\Lambda^2, \tag{19}$$

where Λ is the length defined by Eq. (17). For a flow driven by the stress gradient, creep dominates in thick films, and diffusion dominates in thin films. Recall that Λ scales with the grain size for polycrystals, and with the atomic size for liquids.

In the presence of both the wind force and the stress gradient, the combined flow due to creep and diffusion vanishes when

$$F+\left(1+\frac{H^2}{12\Lambda^2}\right)\Omega\nabla\sigma=0. \tag{20}$$

The contribution of creep in countering electromigration depends on the film thickness. This effect is small in the exiting technology. To enhance the creep effect, we have to accelerate creep relative to diffusion, so that the length Λ becomes much smaller than the film thickness. For example, Λ approaches the atomic dimension for a liquid metal. The effect of creep can probably be demonstrated in laboratories, but there is no clear way to implement the effect in the electronic industry. Such an implementation would call for a material with both a large creep rate and a high electric conductivity.

6.2 Stress Generated by Electromigration Through Film Thickness. Figure 4 illustrates a metal film sandwiched between two other conductors, with electric current through the film thickness. The two outside conductors do not suffer electromigration, but the film does. This setup idealizes a contact. The electron wind force, F, now in the x_3-direction, causes atoms of the film to diffuse from one side to the other. Consequently, a state of stress is generated, tensile on one side and compressive on the other. The stress state is biaxial, $\sigma_{11}=\sigma_{22}$; all other stress components vanish. The mean stress component is $\sigma=2\sigma_{11}/3$. The stress is the function of the depth, $\sigma_{11}(x_3)$, and is to be determined.

The diffusion flux is a long the x_3-direction, given by

$$J_3=\frac{D}{\Omega kT}\left(F+\frac{2\Omega}{3}\frac{\partial\sigma_{11}}{\partial x_3}\right). \tag{21}$$

Because the net atomic flux vanishes, the markers move in the direction opposite to the diffusion flux, $v_3=-\Omega J_3$. The diffusion flux induces a strain rate $d_{11}^D=-(\Omega/3)\partial J_3/\partial x_3$. The deviatoric stress component is $s_{11}=\sigma_{11}/3$. The creep strain rate is proportional to the deviatoric stress, $d_{11}^C=\sigma_{11}/6\eta$. The film is constrained by the refractory metals, so that the strain rates vanish in the two lateral directions: $d_{11}=d_{22}=0$. The strain rate is the sum of that due to diffusion, and that due to creep. The sum vanishes:

$$-\frac{2D\Omega}{9kT}\frac{\partial^2\sigma_{11}}{\partial x_3^2}+\frac{\sigma_{11}}{6\eta}=0. \tag{22}$$

This is a second-order differential equation for $\sigma_{11}(x_3)$.

Atoms do not diffuse in or out the refractory conductors, so that the diffusion flux vanishes at the two faces of the film. Subject to these the boundary conditions, the solution to the differential equation is

$$\sigma_{11}(x_3)=-\frac{3Fl\sinh(x_3/l)}{2\Omega\cosh(H/2l)}. \tag{23}$$

The characteristic length is $l=\sqrt{4D\eta\Omega/3kT}$, which differs from the length identified in Section 5 by a numerical factor. The stress

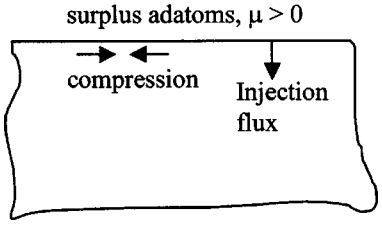


Fig. 5 An excess of the chemical potential on the surface drives atoms to inject into the material, leading to compressive stress

vanishes at the middle of the film, tensile on one side, and compressive on the other. When the diffusivity or viscosity is large, electromigration is rapid, and the material creeps slowly, so that the magnitude of the stress is large.

6.3 The Effect of Elasticity. When the electric current just starts, atoms have yet migrated much, and the stress in the film is negligible. This time-dependence is absent in the above solution. The problem arises because we have neglected elasticity. In this case, the geometric change is small, and we can include elasticity easily. The biaxial stress causes an elastic strain rate $d_{11}^E=[(1-\nu)/E]\partial\sigma_{11}/\partial t$. The combined strain rate due to diffusion, creep and elasticity vanishes:

$$-\frac{2D\Omega}{9kT}\frac{\partial^2\sigma_{11}}{\partial x_3^2}+\frac{\sigma_{11}}{6\eta}+\frac{1-\nu}{E}\frac{\partial\sigma_{11}}{\partial t}=0. \tag{24}$$

The problem now has a time scale $\tau=6(1-\nu)\eta/E$. The stress no longer changes instantaneously, but builds up gradually. The stress distribution given by (23) is the steady state, reached over the time scale τ.

6.4 Stress Generated by Atomic Injection or Emission The chemicalpotential μ on the surface can be varied in several ways, by applying a stress normal to the surface, by creating excess number of adatoms with an impinging flux, or by creating excess number of vacancies by oxidation. A surplus or deficit in the chemical potential will motivate atoms to diffuse in or out the material, leading to biaxial compression or tension (Fig. 5). Let the bulk of the material occupies the half-space $x_3<0$. The stress is prescribed by the chemical potential, $\sigma_{11}=\sigma_{22}=-3\mu/2\Omega$, on the surface, and vanishes as $x_3\to-\infty$. Equation (22) governs the stress as a function of the depth, giving

$$\sigma_{11}(x_3)=-\frac{3\mu}{2\Omega}\exp(x_3/l). \tag{25}$$

The chemical potential of the surface atoms sets the magnitude of the stress field. The stress decays exponentially over the length l.

6.5 Lateral Expansion or Contraction of a Free-Standing Film. Next consider a free-standing thin film. When the film thickness is on the order of the length l, the lateral constraint is partially relieved, and the film will expand or contract. The lateral strain rate d_{11} is independent of position x_3. The strain rate is the sum of that due to diffusion, and that due to creep:

$$d_{11}=-\frac{2D\Omega}{9kT}\frac{\partial^2\sigma_{11}}{\partial x_3^2}+\frac{\sigma_{11}}{6\eta}. \tag{26}$$

Force balance requires that the resultant force vanish: $\int_{-H}^{+H}\sigma_{11}(x_3)dx_3=0$. The solution to the ordinary differential equation is

$$\sigma_{11}(x_3)=\left(-\frac{3\mu}{2\Omega}\right)\frac{(H/2l)\cosh(x_3/l)-\sinh(H/2l)}{(H/2l)\cosh(H/2l)-\sinh(H/2l)}. \tag{27}$$

The lateral strain rate is

$$d_{11}=\left(\frac{\mu}{4\eta\Omega}\right)\frac{\sinh(H/2l)}{(H/2l)\cosh(H/2l)-\sinh(H/2l)}. \tag{28}$$

If atoms emit or inject preferentially at one of the film surfaces, the film will bend. The above analysis can be extended to calculate the rate at which the bending curvature increases.

6.6 Effect of Nonlinear Creep. Again consider the stress generated in a semi-infinite material by atomic injection or emission at the surface. Under the biaxial stress state, $\sigma_{11}=\sigma_{22}$, the equivalent stress is $\sigma_e=|\sigma_{11}|$. The lateral creep strain rate is $d_{11}^C=B/2\,|\sigma_{11}|^{n-1}\sigma_{11}$. The combined strain rate due to diffusion and creep vanishes:

$$-\frac{2D\Omega}{9kT}\frac{\partial^2\sigma_{11}}{\partial x_3^2}+\frac{B}{2}|\sigma_{11}|^{n-1}\sigma_{11}=0. \tag{29}$$

The solution of the boundary value problem is

$$\sigma_{11}(x_3)=-\frac{3\mu}{2\Omega}\left(1-\frac{x_3}{l_*}\right)^{-2/(n-1)}, \tag{30}$$

with

$$l_*=\frac{2}{n-1}\sqrt{\frac{2(n+1)\Omega D}{9|3\mu/2\Omega|^{n-1}BkT}}. \tag{31}$$

This length differs from the length identified in Section 5 by a numerical factor. Compared to the linear creep, the power-law creep modifies both the decay length and the decay function. For a large value of n the stress is substantial over a depth several times l_*.

7 Anisotropic Placement Rule

Equation (5) has been called the isotropic placement rule, [11]. We caution that this rule must be modified if atoms can be removed and inserted preferentially on some planes. For example, Fig. 6 illustrates a polycrystalline aluminum film, of columnar grain structure in the direction of the film thickness, and the native oxide covering the film surfaces. Aluminum diffusion is fast along the grain boundaries, and negligible on the film surfaces. Consequently the divergence in the self-diffusion flux will place atoms. We now extend the theory on the basis of an anisotropic placement rule.

Recall that, according to Eq. (2), the divergence in the diffusion flux causes the divergence of the marker velocity field. The issue is how to proportion this divergence in various directions. Our anisotropic placement rule stipulates that

$$d_{ij}^D=-\beta_{ij}\Omega J_{k,k}. \tag{32}$$

Here the coefficients β_{ij} weigh the placement in different directions. We require the tensor β_{ij} to be symmetric with a unit trace, $\beta_{ii}=1$. Consistent with this placement rule, the creep strain-rate tensor relates to the marker velocity field as

$$d_{ij}^C=\tfrac{1}{2}(v_{i,j}+v_{j,i})-\beta_{ij}v_{k,k}. \tag{33}$$

The creep strain-rate tensor is symmetric and traceless.

The statement of power balance still takes the form of Eq. (7). However, the creep dissipation rate now becomes $s_{ij}d_{ij}^C=s_{ij}v_{i,j}-\beta_{ij}s_{ij}v_{k,k}$. This will modify Eq. (8) by replacing $s_{ij}+\lambda\delta_{ij}$ in the two places by $s_{ij}+(\lambda-\beta_{pq}s_{pq})\delta_{ij}$; the rest of Eq. (8) remains unchanged. We now identify $(\lambda-\beta_{pq}s_{pq})$ as the mean stress σ, so that $\lambda=\sigma+\beta_{pq}s_{pq}=\beta_{pq}\sigma_{pq}$. It is the chemical potential $-\Omega\beta_{pq}\sigma_{pq}$ that enters Herring's equations of diffusion driving force, Eqs. (11) and (12). We may also wish to introduce anisotropy into the kinetic laws, Eqs. (13) and (14), which we will not pursue here.

As an example, when diffusion flux divergence is placed equally in the x_1 and x_2-directions, but not in x_3-direction, we let $\beta_{11}=\beta_{22}=1/2$, and all other components vanish. Consequently,

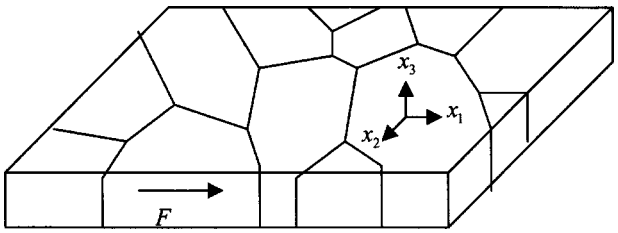


Fig. 6 An aluminum film has a columnar grain structure, with top and bottom surfaces covered by the native oxide. Under an electron wind force in the plane of the film, aluminum atoms diffuse fast on the grain boundaries, but negligibly on the film surfaces. The divergence of the diffusion flux will place atoms in the x_1 and x_2-directions, but not in the x_3-direction.

the chemical potential in the material becomes $-\Omega(\sigma_{11}+\sigma_{22})/2$. The work is done by the stresses acting in the direction where mass insertion or removal occurs.

8 Summary

Both convection and diffusion contribute to mass transport. Identify convection by the motion of markers dispersed in the material. Creep and self-diffusion couple because the markers must move to compensate for the diffusion flux divergence, Eq. (2). We stipulate rules to place the diffusion flux divergence to various planes; two versions are given: isotropic rule Eq. (5) and anisotropic rule Eq. (33). We define the driving force for creep and diffusion by a statement of power balance, Eq. (7), subject to the kinematic constraint. The theory leads to partial differential equations for the marker velocity field and the chemical potential field, Eqs. (15) and (16). The pair of equations generalizes Stokes's creep and Herring's diffusion. A length characterizes the relative rate of diffusion and creep, Eq. (17). Several boundary value problems illustrate the theory. In particular, a stress gradient can drive both a diffusion flow and a creep flow. Diffusion flow prevails in a thin channel, and creep flow prevails in a thick channel. The transition occurs when the channel thickness is comparable to the characteristic length of the material.

Acknowledgments

This work was supported by the ONR MURI grant entitled ''Prime Reliant Coatings,'' by NSF through grant CMS-9820713, and by the Division of Engineering and Applied Sciences at Harvard University. Discussions with D. V. Kubair, A. G. Evans, J. R. Rice, and F. Spaepen were helpful.

References

[1] Chason, E., Sheldon, B. W., Freund, L. B., , Freund, L. B., and Hearne, J. A., 2002, ''Origin of Compressive Residual Stress in Polycrystalline Thin Films,'' Phys. Rev. Lett., **88**, p. 156103.
[2] Vermilyea, D. A., 1957, ''On the Mechanism of Oxidation of Metals,'' Acta Metall., **5**, pp. 492–495.
[3] Prussin, S., 1972, ''Generation of Stacking Faults and Prismatic Dislocation Loops in Device-Processed Silicon Wafers,'' J. Appl. Phys., **43**, pp. 2850–2856.
[4] Blech, I. A., and Herring, C., 1976, ''Stress Generation by Electromigration,'' Appl. Phys. Lett., **29**, pp. 131–133.
[5] Korhonen, M. A., Borgesen, P., Brown, D. D., and Li, C. Y., 1993, ''Microstructure Based Statistical Model of Electromigration Damage in Confined Line Metallization in the Presence of Thermally Induced Stresses,'' J. Appl. Phys., **74**, pp. 4995–5004.
[6] Suo, Z., 1998, ''Stable State of Interconnect under Temperature Change and Electric Current,'' Acta Mater., **46**, pp. 3725–3732.
[7] Wang, P.-C., Cargill, G. S., Noyan, I. C., and Hu, C.-K., 1998, ''Electromigration-Induced Stress in Aluminum Conductor Lines Measured by X-ray Microdiffraction,'' Appl. Phys. Lett., **72**, pp. 1296–1298.
[8] Herring, C., 1950, ''Diffusional Viscosity of a Polycrystalline Solid,'' J. Appl. Phys., **21**, pp. 437–445.
[9] Darken, L. S., 1948, ''Diffusion, Mobility and Their Interrelation Through Free Energy in Binary Metallic Systems,'' Trans. AIME, **175**, pp. 184–201.
[10] Stephenson, G. B., 1988, ''Deformation during Interdiffusion,'' Acta Metall., **36**, pp. 2663–2683.
[11] Suo, Z., Kubair, D. V., Evans, A. G., Clarke, D. R., and Tolpygo, V. K., 2003, ''Stress Induced in Alloys by Selective Oxidation,'' Acta Mater., **51**, pp. 959–974.
[12] Einstein, A., 1926, *Investigations on the Theory of the Brownian Movement*, reprinted by Dover Publications, New York.
[13] Nabarro, F. R. N., 1948, ''Deformation of Crystals by the Motion of Single Ions,'' Report of a Conference on Strength of Solids (Bristol), pp. 75–90.
[14] Coble, R. L., 1963, ''A Model for Boundary Diffusion Controlled Creep in Polycrystalline Materials,'' J. Appl. Phys., **34**, pp. 1679–1682.
[15] Needleman, A., and Rice, J. R., 1980, ''Plastic Creep Flow Effects in the Diffusive Cavitation of Grain Boundaries,'' Acta Metall., **28**, pp. 1315–1332.
[16] Balluffi, R. W., 1953, ''The Supersaturation and Precipitation of Vacancies During Diffusion,'' Acta Metall., **2**, pp. 194–202.
[17] Suo, Z., 1997, ''Motions of Microscopic Surfaces in Materials,'' Adv. Appl. Mech., **33**, pp. 193–294.
[18] Cocks, A. C. F., Gill, S. P. A., and Pan, J. Z., 1999, ''Modeling Microstructure Evolution in Engineering Materials,'' Adv. Appl. Mech., **36**, pp. 81–162.
[19] de Groot, S. R., and Masur, P., 1984, *Nonequilibrium Thermodynamics*, Dover Publications, New York.
[20] Frost, H. J., and Ashby, M. F., 1982, *Deformation-Mechanism Maps*, Pergamon Press, Oxford, UK.
[21] Rosenberg, R., Edelstein, D. C., Hu, C.-K., and Rodbell, K. P., 2000, ''Copper Metallization for High Performance Silicon Technology,'' Annu. Rev. Mater. Sci., **30**, pp. 229–262.
[22] Suo, Z., 2003, ''Reliability of Interconnect Structures,'' pp. 265–324 in Volume 8: *Interfacial and Nanoscale Failure* W. Gerberich, and W. Yang, eds. (*Comprehensive Structural Integrity*, I. Milne, R. O. Ritchie, and B. Karihaloo, editors-in-chief), Elsevier, Amsterdam.

Sandwich Plates Actuated by a Kagome Planar Truss

N. Wicks

J. W. Hutchinson

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

Kagome truss plates have properties that suggest they should be uniquely effective as an actuation plane for sandwich plates: a Kagome truss plate has in-plane isotropy, optimal stiffness and strength, and its truss members can be actuated with minimal internal resistance. In this paper, sandwich plates are studied that are comprised of one solid face sheet and one actuated Kagome face sheet joined by a pyramidal truss core. Various aspects of the actuation behavior of these plates are investigated, including internal resistance and strains resulting from actuation and efficiency of actuation. Single and double curvature actuation modes are investigated. Contact is made with analytic results for actuation modes with long wavelength. [DOI: 10.1115/1.1778720]

1 Introduction

Recent studies of planar trusses based on the ancient Kagome basket weave pattern have shown that these truss plates have many properties that make them desirable for actuation planes of sandwich plates, [1,2]. In this study, we begin by analyzing the actuation characteristics of a single Kagome truss plate (Fig. 1) and follow with an actuation analysis of a sandwich plate comprised of one solid face sheet and one actuated Kagome face sheet joined by a pyramidal truss core (Fig. 2).

The single Kagome truss plate can be constructed from the unit cell shown in Fig. 1(*b*). The 120 deg symmetry of the structure ensures in-plane elastic isotropy assuming all the truss members are identical. Here, only solid circular members are considered, of length L and radius R. The Kagome-backed solid skin plate can be constructed from the unit cell depicted in Fig. 2(*c*). In the present study we limit consideration to plates with identical solid circular truss members of length L and radius R both for the Kagome face and the core. The solid skin thickness is denoted by t. In addition, to further limit the number of parameters in the system, we consider only plates in which both face sheets and the core members are constructed of the same material with Young's modulus E, Poisson's ratio ν and yield stress σ_Y. The Kagome-backed sandwich plate in Fig. 2 has isotropic bending and stretching stiffness.

The feature of the planar Kagome truss in Fig. 1 that makes it most advantageous for actuation is the ability to actuate members to achieve arbitrary in-plane nodal displacements with minimal internal resistance. Among infinite isotropic planar truss structures, a pin jointed planar Kagome truss is optimally stiff and strong to overall stressing. Although it has kinematic mechanisms, it is nevertheless able to carry arbitrary states of overall stress. Members of a pin jointed Kagome truss can be actuated (i.e., elongated or shortened) with no internal resistance, or equivalently, with no redundant stresses. When joints are welded, as will be assumed throughout this paper, actuation of a member does encounter internal resistance, but minimally so as will be shown. Welded joints also suppress the kinematic mechanisms and result in a structure with substantial in-plane buckling resistance. These and other aspects of the Kagome structure are explored elsewhere, [1–5].

In this study, we aim to explore the details of actuation of the planar Kagome truss and the Kagome-backed sandwich plate. Simulations of various periodic actuation modes of these infinite Kagome structures have been performed for a range of member aspect ratios. In its most general form, the approach is numerical. A calculation requires the formulation of a ''super element'' representing a unit cell followed by the assembly of the complete structure as a union of the super elements. The scaling of the energy required for actuating the unloaded structure with aspect ratio of the truss members is investigated. Bending and stretching strains in the members induced by actuation are also determined. For actuation modes with wavelengths long compared to the member length, the response computed with the numerical approach is compared to actuation predicted by an analytical long wavelength approximation outlined in a previous study, [1].

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, October 15, 2003; final revision, January 14, 2004. Associate Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 Planar Kagome Structure

2.1 The Planar Kagome Structure. Consider the infinite planar Kagome structure shown in Fig. 1(*a*) having members of length L and solid circular cross-sections of radius R and undergoing in-plane displacements. The unit cell of such a structure is shown in Fig. 1(*b*). The members are modeled as Euler-Bernoulli beams, with clamped conditions at each node representing welded joints (i.e., the displacement and rotational degrees-of-freedom are the same for all beams meeting at a given node). Actuation of any member comprises an elongation or contraction of the member by a strain e if the member were unconstrained. In other words, the actuation strain e is equivalent to a stress-free transformation strain.

2.2 Actuation Methodology. The actuation of beam members is modeled via the so-called cut-stress-reweld scheme employed by Eshelby in his study of the transformation strains in ellipsoidal inclusions, [6]. To actuate a given member, envision the following steps:

1. Remove the actuating member from the structure.
2. Allow the member to actuate (elongate or contract) freely by strain e.
3. Place equal and opposite forces $-E\pi R^2 e$ on the ends of the member to deform it back to its original configuration.
4. Place the member back into the structure and ''weld'' it in place.
5. Release the forces from the ends of the member by applying equal and opposite forces $E\pi R^2 e$ to the joints at the member ends.

Thus, the cut-stress-reweld procedure is equivalent to analyzing the complete truss subject to equal and opposite actuation forces of magnitude $E\pi R^2 e$ applied to the joints at the ends of the respective member.

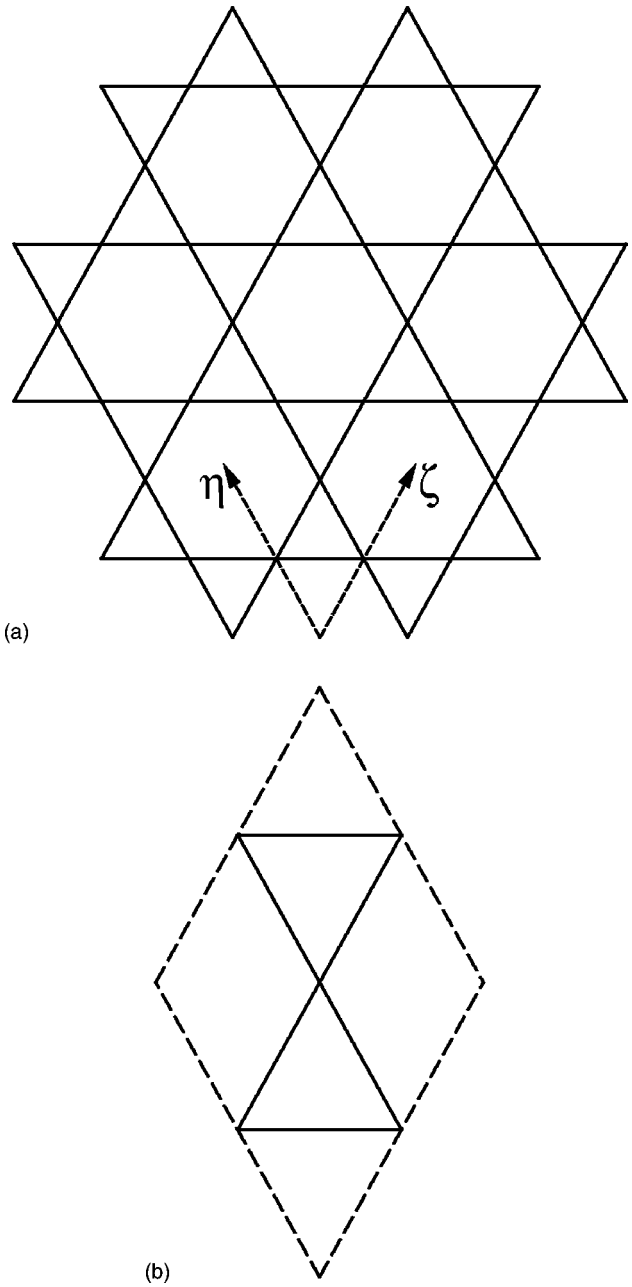


Fig. 1 (*a*) The Kagome planar truss. (*b*) The unit cell used for the Kagome planar truss analysis. The dashed lines are the outline of the cell. The solid lines are truss members of the unit cell.

2.3 Enforcement of Periodicity. Our goal is to simulate periodic actuations of an infinite structure by modeling the behavior of just one periodic cell subject to periodic boundary conditions. Consider the potential energy functional for a periodic cell:

$$\Phi = \frac{1}{2} K_{ij} u_i u_j - k_j^0 u_j \tag{1}$$

where $\boldsymbol{K}$ is the conventional stiffness matrix of the structure, $\mathbf{u}$ is the vector of displacements and rotations at the nodes of the structure, and $\mathbf{k}^0$ is the vector of applied nodal forces and moments (in this situation these correspond to the virtual actuation forces described earlier). If the structure were isolated and unconstrained

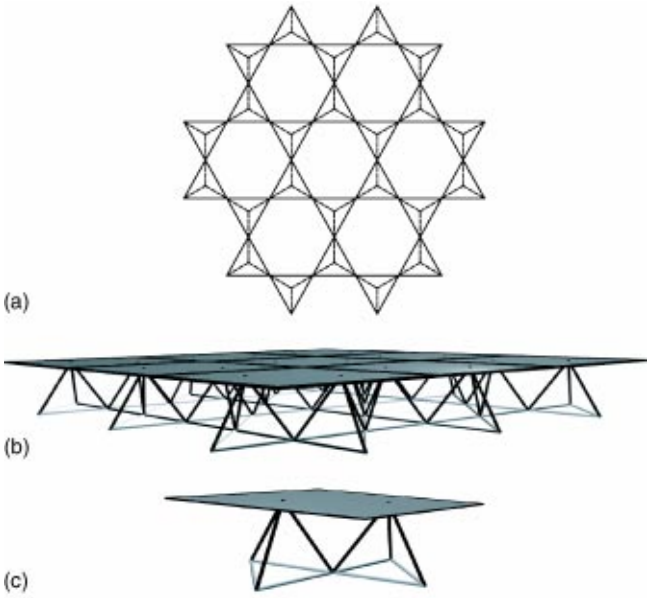


Fig. 2 (*a*) The Kagome plane and tetrahedral core of the Kagome plate structure. The solid members are the Kagome face members, and the dashed members the tetrahedral core members. (*b*) The Kagome plate structure. (*c*) The unit cell used for the Kagome plate analysis.

(apart from rigid-body motion), the problem would be solved in the usual manner by solving the system of equations represented by

$$K_{ij} u_j = k_i^0. \tag{2}$$

Now consider periodic boundary conditions represented by constraints of the form

$$a_{ij} u_j = 0. \tag{3}$$

The details of the actual periodic boundary conditions imposed for in-plane nodal displacements and rotations are included in the Appendix. It is important to note that once the appropriate periodic boundary conditions are imposed on the *displacements/rotations*, periodicity of *forces/moments* is satisfied by the solution. For the periodic cell in Fig. 3, the forces acting on a node along one edge of the periodic cell (corresponding to internal forces in the structure) will be equal and opposite to those acting on the equivalent node on the opposite edge, for a displacement field which satisfies the periodic conditions outlined in the Appendix. The displacements themselves are not, in general, periodic—we consider, for example, a displacement field corresponding to a constant strain. Thus, there are some cases for which the strains, forces and moments, will be periodic, but not the displacements.

To impose these additional conditions, Lagrangian multipliers are employed. The modified energy functional now takes the form:

$$\Phi = \frac{1}{2} K_{ij} u_i u_j - k_j^0 u_j - \lambda_i a_{ij} u_j \tag{4}$$

The systems of equations resulting from minimization of the energy functional with respect to displacements u_i and the Lagrangian multipliers λ_i are

$$K_{ij} u_j = k_i^0 + \lambda_j a_{ji} \tag{5}$$

$$a_{ij} u_j = 0. \tag{6}$$

Now consider periodic actuations of the infinite structure. The actuation forces exert no net force or moment on the structure.

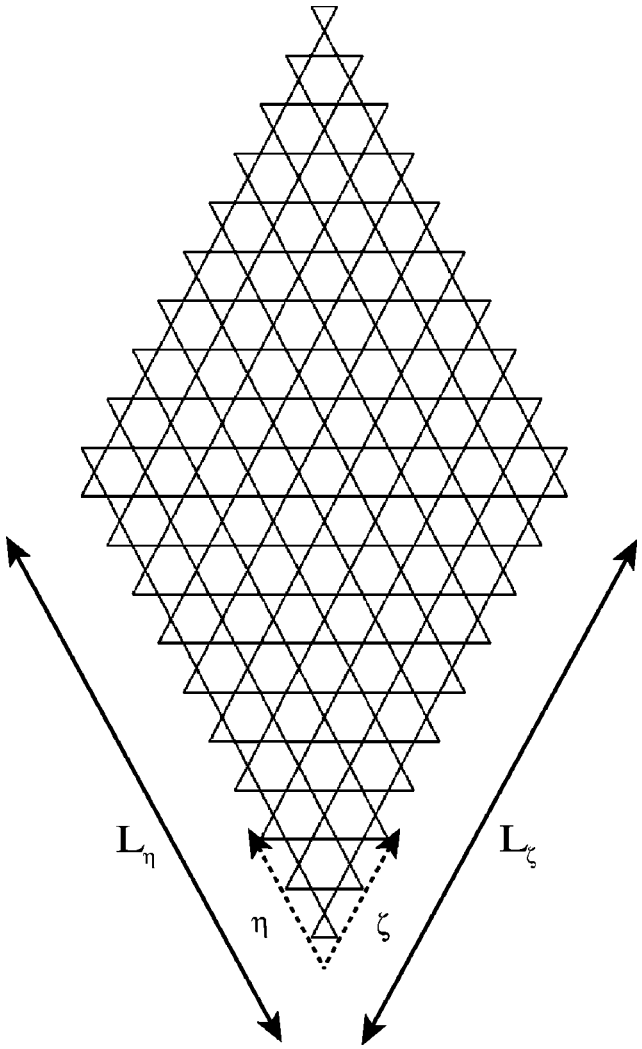


Fig. 3 The periodic cell used for the planar Kagome truss simulations

This means that if the energy functional given by (4) above is minimized over the periodic cell, the energy of the infinite structure is also minimized. Thus, it is possible to simulate the behavior of the infinite structure by modeling just the periodic cell with the appropriate boundary conditions.

2.4 Calculation of Actuations. The objective is to probe how effectively the planar Kagome structure can achieve arbitrary in-plane deformations through actuation of its members. To this end, periodic actuation of the structure is simulated as described in previous sections for actuation of each member in the periodic cell, tracking the nodal displacements in a matrix of influence coefficients, $\boldsymbol{A}$. The ijth component of this matrix is the ith displacement resulting from the actuation of the jth member. To generate this matrix, displacements are calculated at each of the nodes in the periodic cell resulting from a unit actuation of each member in that cell.

As this simulation is linear, once the matrix of influence coefficients, $\boldsymbol{A}$, has been constructed, the displacements of the nodes from actuation of any combination of members is easily computed as

$$u_i = A_{ij} e_j \tag{7}$$

where $\mathbf{e}$ is a vector of member actuations.

Now consider a target field of in-plane nodal displacements $\mathbf{u}^d$. The aim is to determine how well this field can be recreated by actuating members of the Kagome structure. Calculate elongations $\tilde{\mathbf{e}}$ and displacements $\tilde{\mathbf{u}}$ via the relations:

$$\tilde{e}_i = A^{\dagger}_{ij} u^d_j \tag{8}$$

$$\tilde{u}_i = A_{ij} \tilde{e}_j \tag{9}$$

where $\boldsymbol{A}^{\dagger}$ is the Moore-Penrose generalized inverse of $\boldsymbol{A}$, also called the pseudo-inverse of $\boldsymbol{A}$, [7]. Then, $\tilde{\mathbf{e}}$ is the vector of member actuations which minimize the squared error between $\tilde{\mathbf{u}}$ and $\mathbf{u}^d$. If there exist multiple vectors of actuations that minimize this squared error, $\tilde{\mathbf{e}}$ is such a vector of minimal length (i.e., $|\tilde{\mathbf{e}}|$ is minimized).

2.5 Example Target Displacement Fields. The simulations outlined above were run for several target displacement fields. The objective is to assess the ability of the structure to achieve specific actuations and to determine the associated energy required and stresses induced. The periodic cell used for these simulations is shown in Fig. 3. It contains a total of 100 unit cells. The axes used for describing the displacements fields described here are the ζ-axis and the η-axis shown in Fig. 3. It is important to note that the target displacement fields are all consistent with the periodic displacement boundary conditions.

The first target displacement field is described by

$$\mathbf{u}^d = A_0 \mathbf{e}_\zeta \sin\left(\frac{\pi\zeta}{L_\zeta}\right). \tag{10}$$

Here, A_0 is an amplitude factor, $\mathbf{e}_\zeta$ is a unit vector aligned with the ζ-axis, and L_ζ the length of the periodic cell in the ζ-direction. This target field is a displacement in the ζ-direction that is everywhere positive except along the edges $\zeta=0$ and $\zeta=L_\zeta$ of the periodic cell where it is zero and reaches a maximum of $A_0\mathbf{e}_\zeta$ along the line $\zeta=L_\zeta/2$ of each periodic cell. In this case, the displacement field is repeated in every periodic cell. This target field is depicted in Fig. 4. The Moore-Penrose best-fit actuations for this displacement field are calculated as described above. For this periodic cell, the number of nodal displacements in the target field is 640, while the number of members actuated is 600. One would expect some error between the target and achievable fields, however the actual displacements differ from the target displacements by less than 1% of the maximum target displacement. The actuations are quite close to the actuations predicted by the long wavelength approximation outlined in previous work, [1]. In this long wavelength theory, the actuation strain, ε^T, of a member connecting neighboring nodes I and J is $\varepsilon^T = (u^d_\alpha(\mathbf{x}^I) - u^d_\alpha(\mathbf{x}^J))t_\alpha/L$ where t_α is the unit vector parallel to the member and directed from J to I and u_α is the displacement derived from the target displacement field.

A second target displacement field is described by:

$$\mathbf{u}^d = A_0(\mathbf{e}_\zeta + \mathbf{e}_\eta)\sin\left(\frac{\pi\zeta}{L_\zeta}\right)\sin\left(\frac{\pi\eta}{L_\eta}\right). \tag{11}$$

Here, $\mathbf{e}_\eta$ is a unit vector aligned with the η-axis and L_η the length of the unit cell in the η direction. This displacement is zero along all of the edges of the periodic cell and reaches a maximum of $A_0(\mathbf{e}_\zeta+\mathbf{e}_\eta)$ at the center $(\eta=L_\eta/2, \zeta=L_\zeta/2)$ of each periodic cell. The target displacement field can be seen in Fig. 5. The Moore-Penrose best-fit actuations are predicted quite well by the long wavelength approximation. In this example, the maximum discrepancy between the actual displacements of the nodes and the target displacements is less than 2% of the maximum target displacement. For simulations run for a target displacement field in the same direction, but with a wavelength half of that of (11) (in both the ζ and η directions), this maximum discrepancy is less than 10^{-7}. The critical difference between the two sets of target displacement fields is that the field described by (11) with the full wavelength has a jump in slope across the boundaries between

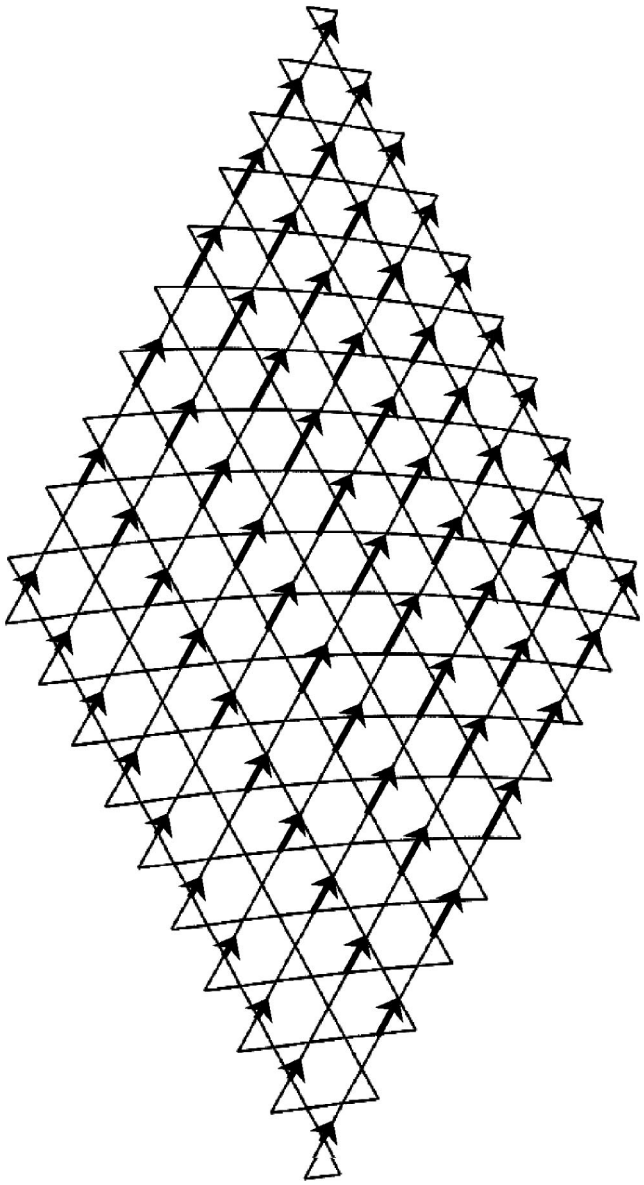


Fig. 4 The target field described by $\mathbf{u}^d = A_0 \mathbf{e}_\xi \sin(\pi\zeta/L_\zeta)$. The arrows show the displacement vectors of the nodes.

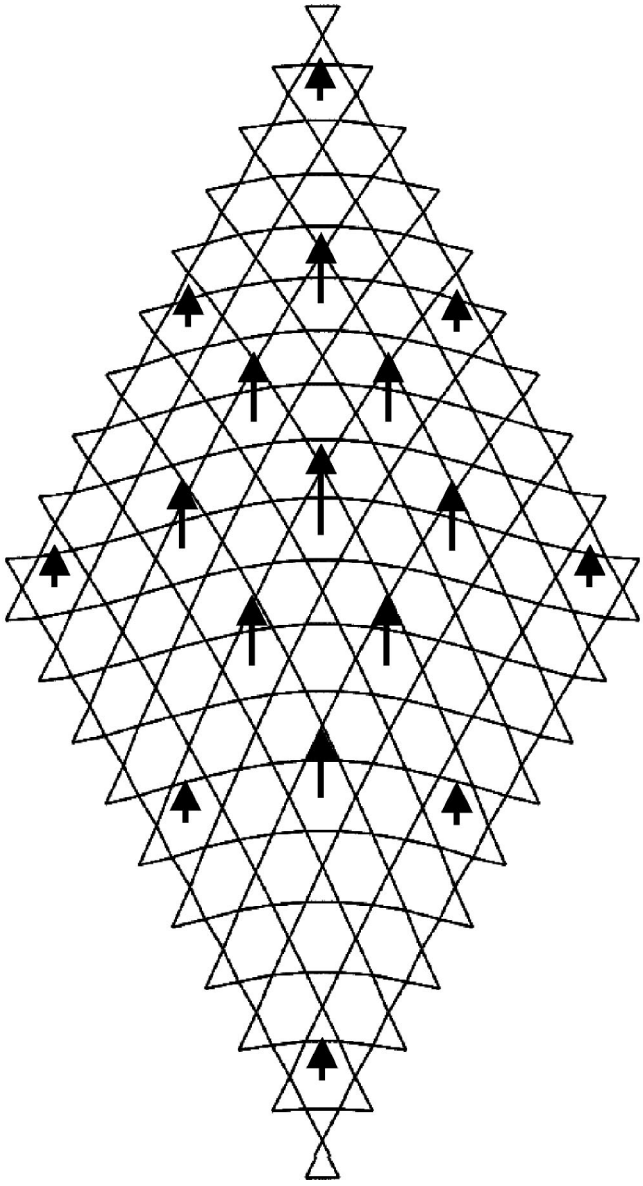


Fig. 5 The target displacement field described by $\mathbf{u}^d = A_0(\mathbf{e}_\xi + \mathbf{e}_\eta)\sin(\pi\zeta/L_\zeta)\sin(\pi\eta/L_\eta)$. The arrows show the displacement vectors of the nodes.

periodic cells (and, as a result, a jump in actuation). For the displacement field with the half-wavelength, the slope of displacement is continuous across periodic cell boundaries.

A final target displacement field of interest is described by

$$\mathbf{u}_d = A_0 \zeta \mathbf{e}_\zeta . \tag{12}$$

This corresponds to a constant strain of A_0 in the ζ direction. This field is shown in Fig. 6. In this example, the actual displacements of the nodes match the target displacements almost perfectly—the maximum discrepancy is less than 10^{-7} of the maximum target displacement. Note that for this displacement field, the slope of displacement is again continuous across periodic cell boundaries.

2.6 Actuation of Selected Kagome Members. In practical applications it will generally be desirable to manufacture structures in which only a small subset of the members will be actuated. The Moore-Penrose actuation scheme can be applied in a similar manner in such cases. This procedure is outlined in Section 4.4 for the Kagome plate structure. Given a restriction on the number of members to be actuated, systematic procedures for identifying the ''best'' subset of actuation members remain to be established. The behavior of planar Kagome structures with limited numbers of actuation members will be considered in subsequent work.

3 Energy of Actuation of Planar Kagome Structure

3.1 Energy of Actuation. One of the motivations behind the selection of the Kagome structure for actuation is the desire to find a structure that can be actuated with minimal internal resistance to actuation. Here, we present the total strain energy stored in the planar Kagome structure actuated to achieve the target displacement fields described above. In the limit of a pin-jointed Kagome structure, actuations can be achieved with no internal resistance, although mechanisms will also exist, [1]. For the Kagome structures considered here, simulated with Euler-Bernoulli beams welded together at their ends, the energy of actuation is expected to be due primarily to elastic bending of the beams.

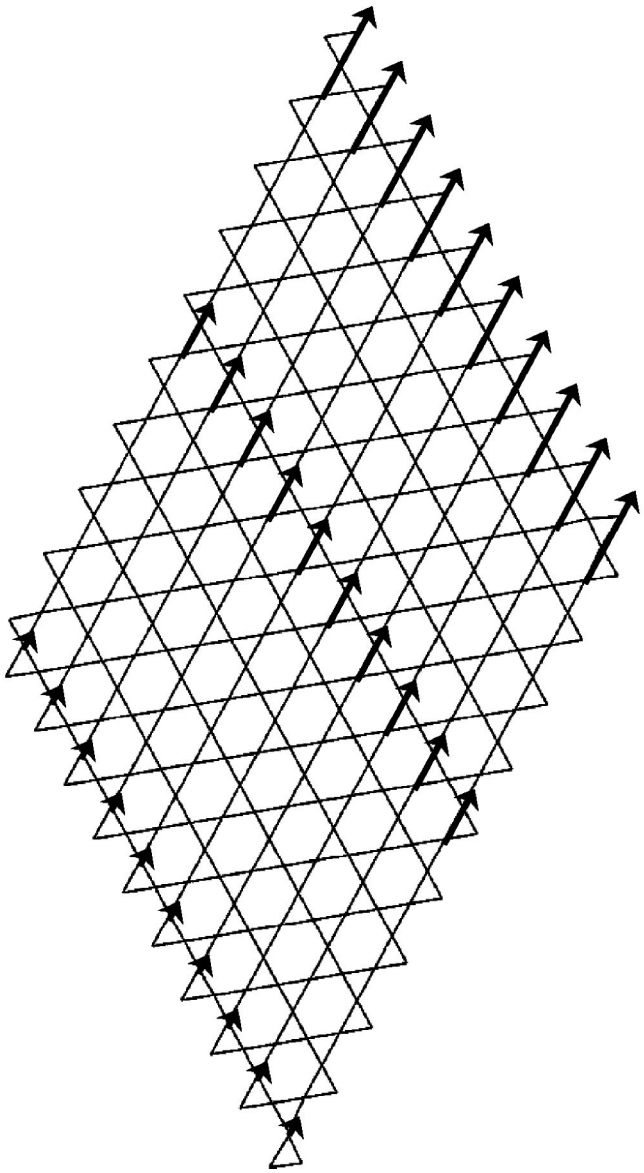


Fig. 6 The target displacement field described by $u_d = A_0 \zeta e_\zeta$

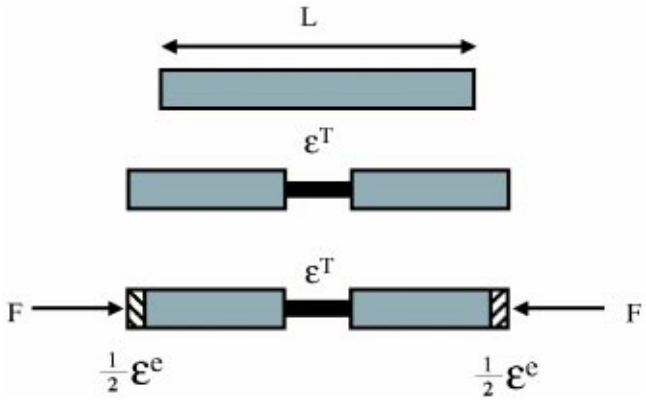


Fig. 7 The energy of actuation of a beam. The rest of the structure resists the imposed actuation strain ε^T, generating an elastic strain of ε^e and an internal force $F = \varepsilon^e EA$.

The strain energy of actuation is calculated in a straightforward manner depicted in Fig. 7. Consider an actuating member of length L, cross-sectional area A, and Young's modulus E. The imposed actuation strain of ε^T is the strain the member would undergo if the member were free to actuate (if the structure offered no resistance). As the structure will have some resistance to the actuation, the member will undergo some elastic strain that we denote ε^e. The axial force F experienced by the member is then easily calculated as $F = \varepsilon^e EA$. The work done by the actuator (acting on the structure) is the work of this axial force acting through the actuation strain ε^T, $-1/2\varepsilon^e \varepsilon^T EAL$, and it is stored as elastic strain energy in the structure.

3.2 Reference Energy. Now consider the work done by the same actuator, undergoing the same actuation strain ε^T, but in this case consider the structure to be rigid—that is, consider a structure that will completely resist the actuation. This corresponds to fixing both ends of the member as it is actuated. In this situation, the total strain of the member is zero, so $\varepsilon^e + \varepsilon^T = 0$. The energy of actuation in this case is therefore $1/2(\varepsilon^T)^2 EAL$, and this is chosen as the reference energy for each member in the energy calculations and comparisons below.

Actuation of multiple members does not pose any energy accounting difficulties—the total energy of actuation can be calculated by adding up the contributions from each individual member, with the elastic strain of each member calculated in response to all the actuations. In some circumstances, the work done by an individual member may be negative—the actuations of other members may result in the structure actually assisting, not resisting, an actuation. However, the overall energy of actuation will of course always be positive. Thus, the relevant energy ratio calculated for energy comparisons is

$$\hat{W} = \frac{\sum_{i=1}^{\#\text{members}} -\frac{1}{2}\varepsilon_i^e \varepsilon_i^T E_i A_i L_i}{\sum_{i=1}^{\#\text{members}} \frac{1}{2}(\varepsilon_i^T)^2 E_i A_i L_i}. \tag{13}$$

This energy ratio $\hat{W}$ is plotted for several slenderness ratios in Fig. 8 for the target displacement fields shown above. For a structure with energy storage dominated by bending energy, $\hat{W}$ can be

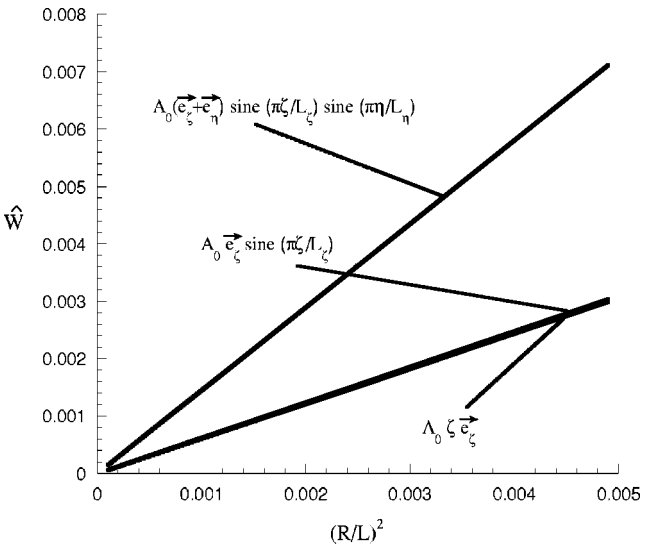


Fig. 8 Normalized actuation energy for the target displacement fields, as a function of $(R/L)^2$

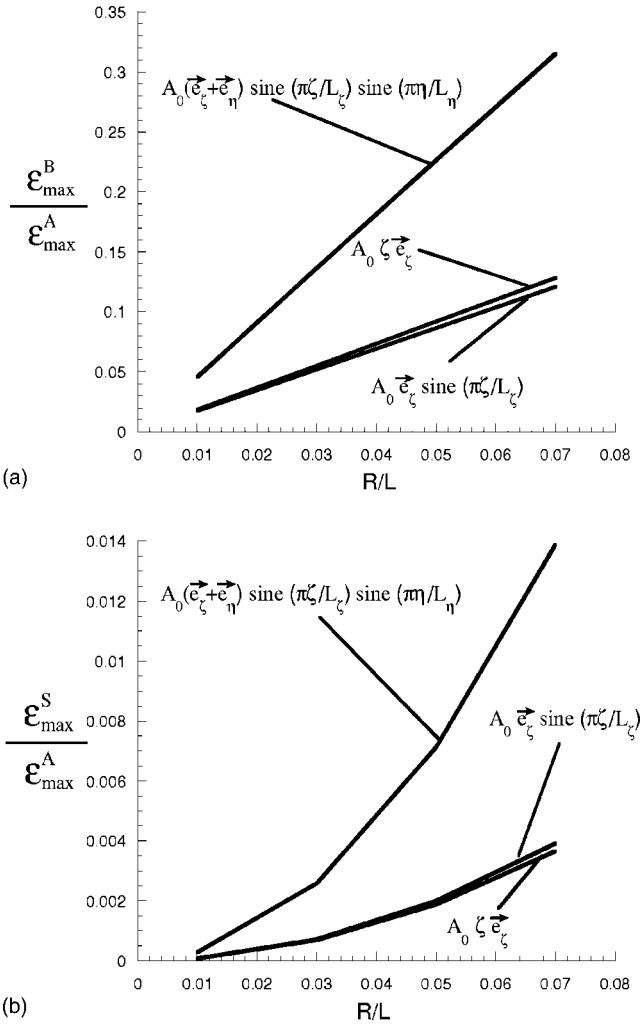


Fig. 9 (*a*) Maximum bending strain normalized by maximum actuation strain for the target displacement fields, as a function of *R*/*L*. (*b*) Maximum stretching strain normalized by maximum actuation strain for the target displacement fields, as a function of *R*/*L*.

shown to scale with the slenderness ratio squared, $(R/L)^2$, while for a statically overdetermined structure, such as a fully triangulated sheet, $\hat{W}$ is expected to be of order unity, independent of the slenderness ratio and whether pin jointed or weld jointed. The energy associated with actuation of the Kagome structure is clearly much smaller than that energy of a fully triangulated truss grid. The energy associated with actuation of an isolated beam in a Kagome structure is investigated in work by Wicks and Guest [5]. For such actuations, $\hat{W}$ scales linearly with slenderness, with energy equally partitioned between stretching and bending. However, for the target fields considered here, $\hat{W}$ clearly scales as the square of the slenderness.

3.3 Strain Levels Induced by Actuation. As the structures will generally undergo cyclic actuations, investigation of strain levels for fatigue design is also required. One quantity of interest is the maximum stretching strain occurring in the structure divided by the maximum actuation strain. For a bending controlled structure, this ratio is also expected to scale with slenderness ratio squared. Of more interest is the ratio of maximum bending strain to maximum actuation strain—which is expected to scale linearly with the slenderness ratio for bending controlled structures. These trends are seen to hold in Fig. 9 where the normalized maximum bending and stretching strains are plotted as a function of slenderness ratio for the target displacement fields described above. The maximum actuation strain, ε^A_{max}, is used to normalize the induced bending and stretching strains. The relatively larger bending strains will restrict the levels of actuation due to fatigue.

4 Kagome Plate Structure

4.1 The Kagome Plate Structure. Consider a solid face sheet of thickness t, backed by a planar Kagome structure. Core members attach to the face sheet and Kagome plane to form tetrahedral units, as shown in Fig. 2. The Kagome planar members and the core members have solid circular cross sections of radius R and are all of the same length L. As a result, the core thickness is $H_c=\sqrt{2/3}L$. While in some applications it may be desirable to use solid face sheets and Kagome planar sheets of different materials, for the purposes of this study we restrict the structure to be of one single isotropic material of Young's Modulus E and with a Poisson's ratio of $\nu=1/3$. The unit cell used in this analysis is shown in Fig. 2(*c*).

The stiffness matrix of the unit cell (Fig. 2(*c*)) of the Kagome plate structure is simulated through the use of a composite element comprised of beam and shell elements. We model the Kagome planar and core members as three-dimensional Euler-Bernoulli beams, with six degrees-of-freedom at each node. The behavior of the solid plate is simulated via linear shell elements in the commercial finite element package ABAQUS [8]. The in-plane behavior of these elements is plane stress, while the out-of-plane behavior corresponds with linear plate theory, [9], as these are flat shell elements.

Since the sandwich plate will generally be subject to applied transverse loads, the Kagome planar truss in its role as a face sheet must carry substantial in-plane loads in addition to undergoing actuation. For the sandwich plate to carry transverse loads efficiently, it is desirable for the solid and Kagome plane face sheets to have comparable in-plane stiffness. The in-plane stiffness of the Kagome planar truss is isotropic, with the relation between average in-plane strains and the overall stress resultants given by

$$\varepsilon_{11}=S^{-1}(N_{11}-\nu N_{22}),\quad \varepsilon_{22}=S^{-1}(N_{22}-\nu N_{11}),$$
$$\varepsilon_{12}=S^{-1}(1+\nu)N_{12} \tag{14}$$

with $S=EA/(\sqrt{3}L)$ and $\nu=1/3$. Equating the in-plane stiffness of the solid sheet and that of the Kagome sheet described by (14) leads to the following relation between the face sheet thickness and member radius:

$$\frac{t}{L}=\frac{\pi}{\sqrt{3}}\left(\frac{R}{L}\right)^2. \tag{15}$$

This relation will be used to specify the face sheet thickness for the various member radii used in the examples detailed below.

4.2 Actuation and Periodicity of the Infinite Kagome Plate.. First consider an infinite Kagome plate structure. The objective of this section is to probe how well the shape of the solid face sheet can be controlled by periodic actuations of the members of the planar Kagome face sheet. Actuation is simulated via the same cut-stress-reweld scheme outlined in Section 2.2. For the infinite plate, periodic target and actuation fields are considered and a corresponding periodic cell is adopted for performing the computations. The details of the periodic boundary conditions for out of plane behavior are given in the Appendix. The periodic boundary conditions, in addition to those suppressing rigid-body motions, are imposed via the use of Lagrangian multipliers, as shown in Section 2.3, resulting in solutions that show periodicity of the forces and moments exerted on the boundaries of the periodic cell.

As in the planar Kagome case, a simulation is run for a unit elongation of every member of the Kagome plane. Core members

Fig. 10 Periodic cell used for Kagome plate simulation

are not actuated. In this case, the vertical displacements of the nodes of the solid face sheet are assembled into a matrix of influence coefficients $\boldsymbol{B}$. The ijth component of $\boldsymbol{B}$ is the vertical displacement of the ith node of the solid face sheet resulting from a unit elongation of the jth member of the Kagome plane.

The linearity of the theory allows the displacements of the nodes from arbitrary actuation (elongation or contraction) of any combination of members to be calculated easily once $\boldsymbol{B}$ is assembled:

$$w_i = B_{ij} e_j \tag{16}$$

where $\mathbf{e}$ is the vector of member actuations and $\mathbf{w}$ the vector of vertical displacements of solid face sheet nodes.

Now consider a target field of the vertical displacements of the solid face sheet displacement field $\mathbf{w}^d$. We wish to determine how well this field can be recreated by actuating members of the Kagome plane. The Moore-Penrose generalized inverse is employed in a similar manner as before to calculate elongations $\tilde{\mathbf{e}}$ and displacements $\tilde{\mathbf{w}}$ via the relations

$$\tilde{e}_i = B^{\dagger}_{ij} w^d_j \tag{17}$$

$$\tilde{w}_i = B_{ij} \tilde{e}_j \tag{18}$$

where $\boldsymbol{B}^{\dagger}$ is the Moore-Penrose generalized inverse of $\boldsymbol{B}$. Here, $\tilde{\mathbf{e}}$ is the vector of member actuations which minimize the squared error between $\tilde{\mathbf{w}}$ and $\mathbf{w}^d$. As before, if there exist multiple vectors of actuations which minimize this squared error, $\tilde{\mathbf{e}}$ is such a vector of minimal length.

4.3 Example Target Displacement Fields. The simulations outlined above were run for several target displacement fields. The periodic cell used for these simulations is shown in Fig. 10. It contains a total of 64 unit cells. The axes used for describing the displacements fields described here are the same ζ-axis and η-axis used in the planar Kagome examples. It is important to note that the target displacement fields imposed are all consistent with the periodic displacement boundary conditions described in the Appendix. As in the planar Kagome examples, the displacements themselves are not, in general, periodic, while the stresses, strains and curvatures are periodic. Here we consider, for example, a displacement field corresponding to a constant curvature—the internal forces are periodic (with periodicity size of the periodic cell), while the displacements are clearly not periodic.

The first target displacement field is described by

$$\mathbf{w}^d = A_0 \zeta^2 \mathbf{e}_z . \tag{19}$$

Here, the displacement corresponds to the vertical displacement of the nodes of the solid face sheet corresponding to a state of constant curvature of $\kappa_{\zeta\zeta} = 2A_0$ with A_0 as the amplitude factor and $\mathbf{e}_z$ the unit vector perpendicular to the plate (aligned with the z-axis). The Moore-Penrose best-fit actuations for this displacement field are calculated as described above. The achievable displacement field is shown in Fig. 11—only the achievable field is shown, as it is indistinguishable from the target field. There are 384 members that are actuated in this simulation and only 209 target nodal displacements. However, the rank of $\boldsymbol{B}$ is only 194, so it is interesting that the achievable field is so close to the target field. Actuation energy and strains will be discussed in Section 5.

The second target displacement field is

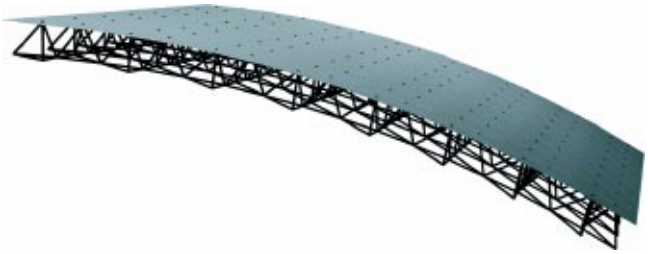


Fig. 11 Best-fit displacement field calculated using Moore-Penrose analysis for target field $\mathbf{w}^d = A_0 \zeta^2 \mathbf{e}_z$

$$\mathbf{w}^d = A_0 \mathbf{e}_z \sin\left(\frac{2\pi\zeta}{L_\zeta}\right) \tag{20}$$

where L_ζ is the length of the unit cell in the ζ-direction. This displacement field has zero displacement along two edges of the periodic cell, corresponding to $\zeta = 0$ and $\zeta = L_\zeta$. Along these same edges, the slope in the ζ-direction takes the value of $2\pi A_0 / L_\zeta$. This displacement field takes on a maximum value of $\pm A_0 \mathbf{e}_z$ at $\zeta = L_\zeta/4$ and $\zeta = 3L_\zeta/4$. The achievable displacement field shown in Fig. 12 again matches the target field.

One final target displacement field of interest is described by

$$\mathbf{w}^d = A_0 \mathbf{e}_z \sin\left(\frac{2\pi\zeta}{L_\zeta}\right) \sin\left(\frac{2\pi\eta}{L_\eta}\right). \tag{21}$$

Note that, unlike the previous target displacement fields, this field involves a nonzero Gaussian curvature of the solid face sheet. The achievable field shown in Fig. 13 also matches the target field.

4.4 Comparison With Long Wavelength Theory. The best-fit actuations for the displacement fields described above have been compared with the actuations predicted by the long wavelength approximation outlined in previous work, [1]. In this long wavelength theory, the extensional strain, ε^T, of a member is $\varepsilon^T = -H \kappa^d_{\alpha\beta} t_\alpha t_\beta$ where $\boldsymbol{\kappa}^d$ is the curvature tensor associated with $\mathbf{w}^d$ ($\kappa^d_{\alpha\beta} = w^d_{,\alpha\beta}$) and t_α is the unit vector specifying the orientation of the member.

Member actuations calculated via the Moore-Penrose analysis for the case of constant curvature agree well with actuations predicted by the long wavelength theory. As this is a displacement field with infinite wavelength, the agreement is not surprising. Actuations for the simulation corresponding to the sinusoidal dis-

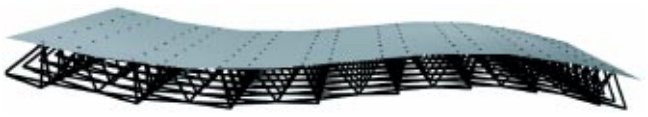


Fig. 12 Best-fit displacement field calculated using Moore-Penrose analysis for target field $\mathbf{w}^d = A_0 \mathbf{e}_z \sin(2\pi\zeta / L_\zeta)$

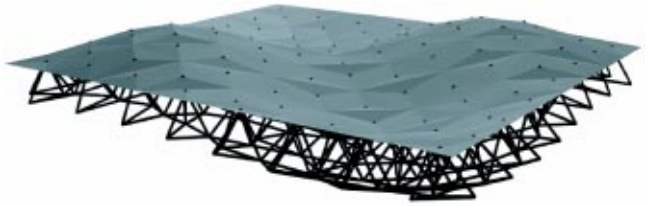


Fig. 13 Best-fit displacement field calculated using Moore-Penrose analysis for target field $\mathbf{w}^d = A_0 \mathbf{e}_z \sin(2\pi\zeta / L_\zeta) \times \sin(2\pi\eta / L_\eta)$. The faceting of the solid sheet is an artifact of the plotting—the actual shape of the solid sheet is smooth.

placement field (20) are also predicted well by the long wavelength theory. This displacement field has a wavelength of eight unit cells. Actuations were also calculated for a sinusoidal displacement field with a wavelength of only four unit cells:

$$\mathbf{w}^d = A_0 \mathbf{e}_z \sin\left(\frac{4\pi\zeta}{L_\zeta}\right). \tag{22}$$

The actuations predicted by the long wavelength theory show more discrepancy in this case, as would be expected, but the largest actuation strains are not in major disagreement.

4.5 Actuation of Selected Kagome Members. As a practical matter, it will usually be the case that only a small subset of the Kagome face members will be actuated. The Moore-Penrose actuation scheme described previously can still be used to probe how well such structures can achieve the target displacement fields.

Consider there are only M members which have been identified for actuation. The matrix $\boldsymbol{B}$ contains the displacements of the solid face sheet nodes corresponding to all member actuations. Thus, we can construct a new matrix $\boldsymbol{B}^*$ that contains only the M columns corresponding to the members to be actuated. Then the vertical displacements of the face sheet nodes are again easily calculated according to

$$w_i = B^*_{ij} e^*_j \tag{23}$$

where now $\mathbf{e}^*$ is the vector of actuations of those members that can be actuated.

Consider a target displacement field $\mathbf{w}^d$. The minimum length, minimum squared error set of actuations for the M members is

$$\tilde{e}^*_i = B^{*\dagger}_{ij} w^d_j . \tag{24}$$

Here, $\boldsymbol{B}^{*\dagger}$ is the Moore-Penrose generalized inverse of $\boldsymbol{B}^*$. The vertical displacements of the solid face nodes are also easily calculated:

$$\tilde{w}_i = B^*_{ij} \tilde{e}^*_j . \tag{25}$$

Reconsider the target displacement field that corresponds to constant curvature:

$$\mathbf{w}^d = A_0 \zeta^2 \mathbf{e}_z . \tag{26}$$

Now, however, assume that only members of the Kagome plane in the row corresponding to $\zeta \approx L_\zeta/2$ can actuate. These members are located in the middle of the periodic cell, as shown in Fig. 14(*a*). The Moore-Penrose best-fit displacement field is shown in Fig. 15. Note that the structure displays only local curvature along the line $\zeta = L_\zeta/2$.

Reconsider also the target sinusoidal displacement field:

$$\mathbf{w}^d = A_0 \mathbf{e}_z \sin\left(\frac{2\pi\zeta}{L_\zeta}\right). \tag{27}$$

Here, however, assume that only the selected members of the Kagome plane are actuated corresponding to those aligned in rows having $\zeta \approx L_\zeta/4$ and $\zeta \approx 3L_\zeta/4$, as shown in Fig. 14(*b*). These members lie within the zones of maximum curvature magnitude of the target displacement field. When the Moore-Penrose analysis is run under these conditions, the resulting displacement field is displayed in Fig. 16. While the shape looks very similar to that in Fig. 12 achieved by activating all the members of the Kagome face, the curvature in Fig. 16 is nevertheless limited to regions where members are actuating along $\zeta \approx L_\zeta/4$ and $\zeta \approx 3L_\zeta/4$.

5 Energy of Actuation of Kagome Plate Structure

5.1 Energy of Actuation. The strain energy of actuation is again calculated in the manner depicted in Fig. 7 and the work done by each actuator is $-1/2\varepsilon^e \varepsilon^T EAL$ where ε^T is its actuation strain and ε^e is the elastic strain it experiences as a consequence of all actuations. This work is stored as elastic strain energy in the structure. The work done by the same actuator, undergoing the same actuation strain ε^T, in a perfectly stiff structure is, as before, $1/2(\varepsilon^T)^2 EAL$. This energy is taken as the reference energy for the energy calculations presented below. As with the planar Kagome truss, actuation of multiple members does not pose any energy accounting difficulties—the total energy of actuation can be calculated by adding up the contributions from each individual member. The relevant energy ratio calculated for energy comparisons is again given by (13) where ε^e_i are computed for each specific set of actuations ε^T_i.

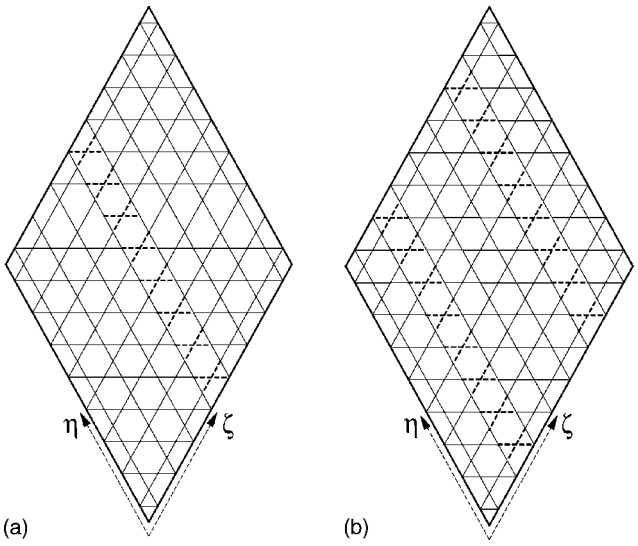


Fig. 14 The members which are allowed to actuate in the analysis of actuation of selected Kagome members. Only the Kagome plane is shown. The dashed members are the members selected to actuate.

As noted earlier, for a structure with energy storage dominated by bending energy in the beam members, the energy ratio $\hat{W}$ can be shown to scale with the slenderness ratio squared, $(R/L)^2$. It is important to note that the sandwich plate comprised of the Kagome face with a solid face sheet is indeterminate—not only are the joints welded, but a solid face is intrinsically indeterminate. It can be anticipated that this indeterminacy will result in

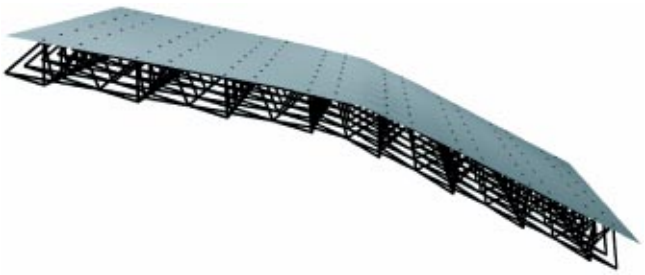


Fig. 15 The best-fit displacement field when limited members are allowed to actuate for the target displacement field $\mathbf{w}^d = A_0 \zeta^2 \mathbf{e}_z$

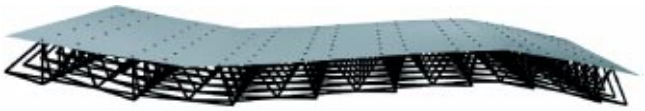


Fig. 16 The best-fit displacement field when limited members are allowed to actuate for the target displacement field $\mathbf{w}^d = A_0 \mathbf{e}_z \sin(2\pi\zeta/L_\zeta)$

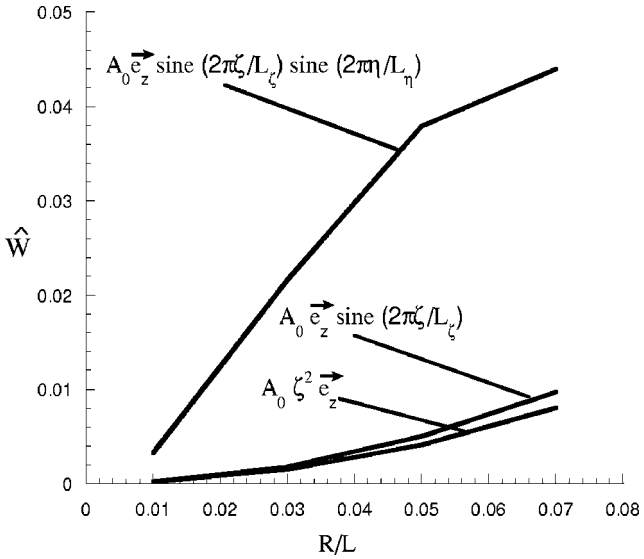


Fig. 17 Normalized actuation energy for the target displacement fields, as a function of *R*/*L*

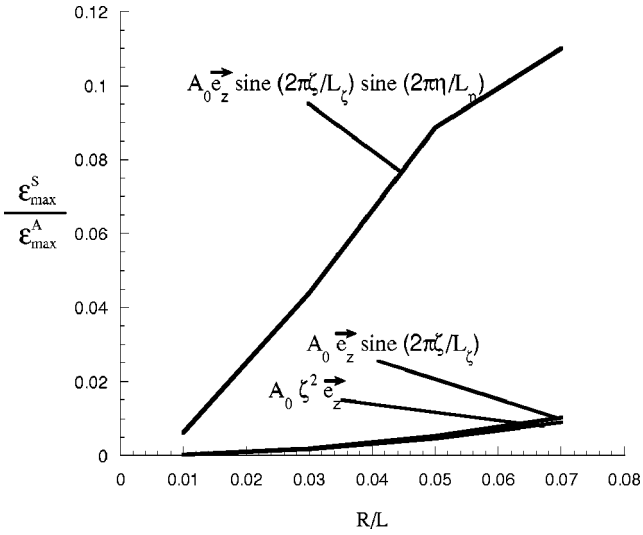


Fig. 19 Maximum stretching strain normalized by maximum actuation strain for the target displacement fields, as a function of *R*/*L*

somewhat larger actuation energies than those predicted for bending dominated structures. The energy ratio $\hat{W}$ is plotted against the slenderness ratio for the truss members in Fig. 17 for the three target periodic displacement fields described above. It is clear from this plot that the Kagome-backed sandwich plate offers considerably more resistance to actuation for modes with curvature in two directions than to actuations that bend the plate solely in one direction. Nevertheless, compared to the reference energy, the actuation energy is still small. A more meaningful interpretation of the actuation energy will be given in Section 5.2. Similar trends are seen in Figs. 18 and 19, where the maximum bending and stretching strain quantities for these displacement fields are plotted as a function of slenderness ratio.

Plates actuated to produce double curvature will necessarily be limited to smaller actuation strains (and thus displacements) to ensure they do not undergo plastic yield. It is important to note that the results here for the double curvature plate have been computed using linear theory. Large resistance will arise for sufficiently large deflections due to nonlinear coupling between bending and stretching in the solid face sheet whenever shapes with nonzero Gaussian curvature are actuated. This nonlinear effect is not considered in this paper. Thus, the results for doubly curved shapes presented here are restricted to small deflections.

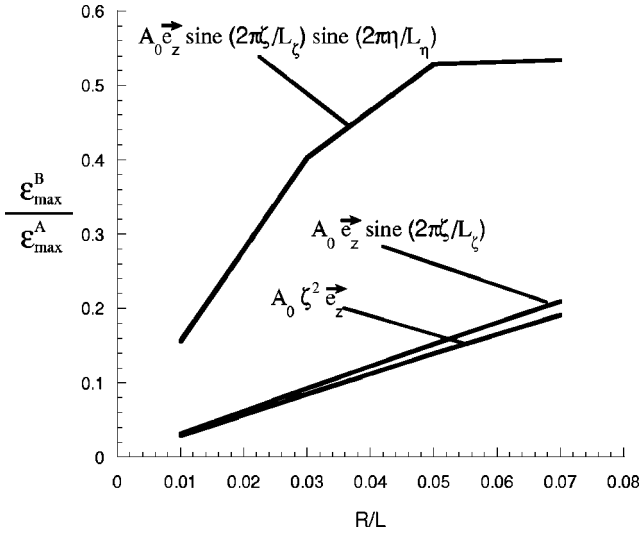


Fig. 18 Maximum bending strain normalized by maximum actuation strain for the target displacement fields, as a function of *R*/*L*

5.2 Comparison of Two Energy Quantities. In most applications, a multifunctional role of these actuating structures is anticipated where the structures will be required to both change shape and carry and lift significant loads. Here an approximate calculation is presented of the relative energies to perform these two different functions, highlighting the significance of structures with low resistance to actuation.

For specificity, consider a cantilevered Kagome plate structure of length ℓ subjected to a load per unit length P at free end. Now imagine that, via actuation of Kagome members, the cantilevered end is raised a distance δ_A. The work per unit length done to raise this load scales as

$$W_P \approx P\,\delta_A = P\,\varepsilon_A \ell \tag{28}$$

where ε_A is the typical actuation strain for a member near the clamped end. The energy per unit length stored as strain energy in the Kagome plate structure due to resistance to actuation scales as

$$W_A \approx kEA\varepsilon_A^2 \tag{29}$$

where A is the member cross-sectional area and k is a small fraction of unity for structures with low internal resistance to actuation, such as those reported above.

The structure must be designed to be able to carry the load per unit length P. For an optimally designed structure, we anticipate that face yielding or buckling will be an active constraint, [10]. With σ_c as the critical stress in a Kagome member (set by either buckling or yielding), the member must be sized to satisfy a relation that scales as $\sigma_c AL \approx PL\ell$. Thus, to carry the applied load, the member cross-sectional area will be sized according to the scaling law

$$A \approx P\ell/\sigma_c. \tag{30}$$

The maximum allowable actuation is also related to this critical stress σ_c according to

$$\varepsilon_A \approx f\sigma_c/E. \tag{31}$$

Here, f is the factor relating actuation strain and the maximum strain induced in the structure due to actuation as plotted for the

Kagome structures earlier in the paper. In most instances, the condition limiting actuation is likely to be yielding in bending, at least for the Kagome structures, but the possibility of buckling of members under compression must also be considered. Now calculate the ratio of the energy stored in the structure due to actuation to the amount of energy required to raise the load:

$$\frac{W_A}{W_P} \approx \frac{kEP\ell\varepsilon_A^2/\sigma_c}{P\varepsilon_A\ell} = \frac{k(E/\sigma_c)\varepsilon_A^2}{\varepsilon_A} = kf. \tag{32}$$

For the Kagome sandwich plates analyzed in Section 4, typical values of f are about 10, as the maximum induced bending strains in the Kagome plane are about 10% of the maximum actuation strain. Values of k depend upon the specifics of the actuation, but because of their low resistance to actuation they typically range from about 1/100 to 1/1000 for the Kagome structures. Thus, for such structures, the energy required to raise the applied loads will be 10–100 times the energy stored as strain energy in the structure. The sandwich plates subject to double curvature offer somewhat more internal resistance to actuation ($k \approx 1/40$), but values of f are also lower ($f \approx 2$) such that the energy stored in internal resistance is still relatively low, i.e., $W_A/W_P \approx 1/20$. In this fundamental sense, the Kagome plate structure offers minimal internal resistance to the actuation.

6 Concluding Remarks

Sandwich plates employing as the actuation plane a Kagome planar truss have been studied to assess their effectiveness and efficiency in the dual role of a load carrying structure capable of actuated shape changes. The advantage of the Kagome planar truss in this application is its in-plane stiffness and strength coupled with its low internal resistance to actuation. The sandwich plate offers more internal resistance than the isolated Kagome plane. Nevertheless, an actuated plate designed to carry specific loads can achieve a wide variety of shapes with relatively low expenditure of energy to overcome the internal resistance compared to the work expended in raising the loads. It remains to be seen from prototypes that are currently under construction, as well as from further theoretical work, just how large the actuated amplitudes can be and the range of modes shapes that can be produced. This is especially true for double curvature modes that require greater expenditure of energy to overcome internal resistance.

Acknowledgments

This work was supported by the grant Multifunctional Mechano-Electronic Materials (N00014-01-1-0523) and by the Division of Engineering and Applied Sciences, Harvard University.

Appendix

Periodic Displacement Boundary Conditions. Consider a periodic structure such as that shown in Fig. A1. With A as the reference periodic cell and $\mathbf{r}^{(i)}$ the vector from the origin in A to the ith node in A, we denote the displacements of the nodes in A as

$$\mathbf{u}_A^{(i)} = \mathbf{u}_0^{(i)}. \tag{A1}$$

Let B, C, and D be neighboring periodic cells, as shown in Fig. A1. We denote the displacements at the nodes in those cells as

$$\mathbf{u}_B^{(i)} = \mathbf{u}_0^{(i)} + \mathbf{u}_{BA} + \boldsymbol{\omega}_{BA} \times \mathbf{r}^{(i)} \tag{A2}$$

$$\mathbf{u}_C^{(i)} = \mathbf{u}_0^{(i)} + \mathbf{u}_{CA} + \boldsymbol{\omega}_{CA} \times \mathbf{r}^{(i)} \tag{A3}$$

$$\mathbf{u}_D^{(i)} = \mathbf{u}_C^{(i)} + \mathbf{u}_{BA} + \boldsymbol{\omega}_{BA} \times \mathbf{r}^{(i)} \tag{A4}$$

$$\mathbf{u}_D^{(i)} = \mathbf{u}_B^{(i)} + \mathbf{u}_{CA} + \boldsymbol{\omega}_{CA} \times \mathbf{r}^{(i)}. \tag{A5}$$

Combining (*A*4) with (*A*3) or (*A*5) with (*A*2) yields

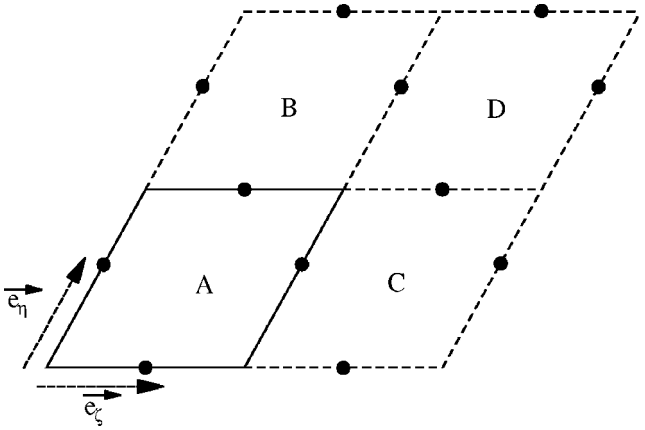


Fig. A1 A periodic structure. A, B, C, D are equivalent periodic cells. The dots correspond to nodes along the edges of the periodic cells. e_ζ and e_η are unit vectors sligned with the edges of the periodic cells.

$$\mathbf{u}_D^{(i)} = \mathbf{u}_0^{(i)} + \mathbf{u}_{BA} + \mathbf{u}_{CA} + (\boldsymbol{\omega}_{BA} + \boldsymbol{\omega}_{CA}) \times \mathbf{r}^{(i)}. \tag{A6}$$

Now consider the rotations at each node in A. Denote these as

$$\boldsymbol{\phi}_A^{(i)} = \boldsymbol{\phi}_0^{(i)}. \tag{A7}$$

Similarly for B, C, and D:

$$\boldsymbol{\phi}_B^{(i)} = \boldsymbol{\phi}_0^{(i)} + \boldsymbol{\omega}_{BA} \tag{A8}$$

$$\boldsymbol{\phi}_C^{(i)} = \boldsymbol{\phi}_0^{(i)} + \boldsymbol{\omega}_{CA} \tag{A9}$$

$$\boldsymbol{\phi}_D^{(i)} = \boldsymbol{\phi}_0^{(i)} + \boldsymbol{\omega}_{BA} + \boldsymbol{\omega}_{CA}. \tag{A10}$$

Now consider the edge joining cell A and cell B. With the displacements of the nodes along the top of A equal to the displacements of the nodes along the bottom of B:

$$\mathbf{u}_0^{(I)} = \mathbf{u}_0^{(i)} + \mathbf{u}_{BA} + \boldsymbol{\omega}_{BA} \times \mathbf{r}^{(i)}. \tag{A11}$$

Here, (I) is along the top of A and (i) along the bottom of B.

We can write $r^{(I)} = r^{(i)} + L_\eta \mathbf{e}_\eta$, where L_η is the length of the periodic cell in the $\mathbf{e}_\eta$ direction.

Now, equate the displacements along the edge joining C and D:

$$\mathbf{u}_0^{(I)} + \mathbf{u}_{CA} + \boldsymbol{\omega}_{CA} \times \mathbf{r}^{(I)} = \mathbf{u}_0^{(i)} + \mathbf{u}_{CA} + \mathbf{u}_{BA} + (\boldsymbol{\omega}_{BA} + \boldsymbol{\omega}_{CA}) \times \mathbf{r}^{(i)}. \tag{A12}$$

Rearranging and simplifying (*A*12) leads to:

$$\mathbf{u}_0^{(I)} - \mathbf{u}_0^{(i)} = \mathbf{u}_{BA} + \boldsymbol{\omega}_{BA} \times \mathbf{r}^{(i)} - \boldsymbol{\omega}_{CA} \times L_\eta \mathbf{e}_\eta \tag{A13}$$

Comparison of (*A*13) with (*A*11) above yields:

$$\boldsymbol{\omega}_{CA} \times L_\eta \mathbf{e}_\eta = 0 \Rightarrow \boldsymbol{\omega}_{CA} = \omega_{CA} \mathbf{e}_\eta \tag{A14}$$

Now equate the displacements of nodes along the edge between cells A and C:

$$\mathbf{u}_0^{(J)} = \mathbf{u}_0^{(j)} + \mathbf{u}_{CA} + \boldsymbol{\omega}_{CA} \times \mathbf{r}^{(j)} \tag{A15}$$

Here, (J) is along the right of A and (j) along the left of C.

Now we can write $r^{(J)} = r^{(j)} + L_\zeta \mathbf{e}_\zeta$, where L_ζ is the length of the periodic cell in the $\mathbf{e}_\zeta$ direction.

Equating displacements of nodes along the edge between cells B and D yields

$$\mathbf{u}_0^{(J)} + \mathbf{u}_{BA} + \boldsymbol{\omega}_{BA} \times (\mathbf{r}^{(j)} + L_\zeta \mathbf{e}_\zeta) = \mathbf{u}_0^{(j)} + \mathbf{u}_{CA} + \mathbf{u}_{BA} + (\boldsymbol{\omega}_{BA} + \boldsymbol{\omega}_{CA}) \times \mathbf{r}^{(j)}. \tag{A16}$$

Rearranging and simplifying (*A*16) yields:

$$\mathbf{u}_0^{(J)} - \mathbf{u}_0^{(j)} = \mathbf{u}_{CA} + \boldsymbol{\omega}_{CA} \times \mathbf{r}^{(j)} - \boldsymbol{\omega}_{BA} \times L_\zeta \mathbf{e}_\zeta. \tag{A17}$$

Comparison of (*A*17) with (*A*15) with above yields

$$\boldsymbol{\omega}_{BA}\times L_{\zeta}\mathbf{e}_{\zeta}=0\Rightarrow\boldsymbol{\omega}_{BA}=\omega_{BA}\mathbf{e}_{\zeta}. \quad (A18)$$

So, for edges parallel to $\mathbf{e}_{\zeta}$, displacements are related by

$$\mathbf{u}_0^{(I)}-\mathbf{u}_0^{(i)}=\mathbf{u}_{BA}+\omega_{BA}\mathbf{e}_{\zeta}\times\mathbf{r}^{(i)}. \quad (A19)$$

Now, write $\mathbf{r}^{(i)}=\zeta^{(i)}\mathbf{e}_{\zeta}+z^{(i)}\mathbf{k}$ and note that $\mathbf{k}$ is perpendicular to $\mathbf{e}_{\zeta}$ and $\mathbf{e}_{\eta}$.

Thus, we can write

$$\mathbf{u}_0^{(I)}-\mathbf{u}_0^{(i)}=\mathbf{u}_{BA}+\omega_{BA}z^{(i)}\mathbf{e}_c\times\mathbf{k}. \quad (A20)$$

Thus, for all pairs of points the same distance above the reference plane

$$\mathbf{u}_0^{(I)}-\mathbf{u}_0^{(i)}=\mathbf{u}_{BA}=\mathbf{u}_0^{(I+m)}-\mathbf{u}_0^{(i+m)}. \quad (A21)$$

Similar along the edges parallel to $\mathbf{e}_{\eta}$:

$$\mathbf{u}_0^{(J)}-\mathbf{u}_0^{(j)}=\mathbf{u}_{CA}=\mathbf{u}_0^{(J+m)}-\mathbf{u}_0^{(j+m)}. \quad (A22)$$

Now, compare the rotations. Setting equal the rotations at nodes along the edge joining cells A and B:

$$\boldsymbol{\phi}_0^{(I)}=\boldsymbol{\phi}_0^{(i)}+\boldsymbol{\omega}_{BA}=\boldsymbol{\phi}_0^{(i)}+\omega_{BA}\mathbf{e}_{\zeta}. \quad (A23)$$

Equating rotations at nodes along the edge joining cells C and D:

$$\boldsymbol{\phi}_0^{(I)}+\boldsymbol{\omega}_{CA}=\boldsymbol{\phi}_0^{(i)}+\boldsymbol{\omega}_{BA}+\boldsymbol{\omega}_{CA}\ \text{(same as (}A23\text{))}. \quad (A24)$$

Finally, equating rotations at nodes along the edge joining cells A and C:

$$\boldsymbol{\phi}_0^{(J)}=\boldsymbol{\phi}_0^{(j)}+\boldsymbol{\omega}_{CA}=\boldsymbol{\phi}_0^{(j)}+\omega_{CA}\mathbf{e}_{\eta}. \quad (A25)$$

References

[1] Hutchinson, R. G., Wicks, N., Evans, A. G., Fleck, N. A., and Hutchinson, J. W., 2003, ''Kagome Plate Structures for Actuation,'' Int. J. Solids Struct., **40**, pp. 6969–6980.

[2] Hutchinson, R. G., and Fleck, N. A., 2003, in preparation.

[3] Hyun, S., and Torquato, S., 2002, ''Optimal and Manufacturable Two-Dimensional Kagome-Like Cellular Solids,'' J. Mater. Res., **17**, pp. 137–144.

[4] Guest, S. D., and Hutchinson, J. W., 2003, ''On the Determinacy of Repetitive Structures,'' J. Mech. Phys. Solids, **51**, pp. 383–391.

[5] Wicks, N., and Guest, S. D., 2003, ''Single Member Actuation in Large Repetitive Truss Structures,'' Int. J. Solids Struct., **41**, pp. 965–978.

[6] Eshelby, J. D., 1957, ''The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems,'' Proc. R. Soc. London, Ser. A, **241**, pp. 376–396.

[7] Strang, G., 1988, *Linear Algebra and its Applications*, Harcourt Brace Jovanovich, Orlando, FL.

[8] Hibbit, Karlsson, and Sorensen, Inc., 2002, ABAQUS Version 6.3.

[9] Timoshenko, S. P., and Gere, J. M., 1961, *Theory of Elastic Stability*, McGraw-Hill, New York.

[10] Wicks, N., and Hutchinson, J. W., 2001, ''Optimal Truss Plates,'' Int. J. Solids Struct., **38**, pp. 5165–5183.

Size-Dependent Eshelby's Tensor for Embedded Nano-Inclusions Incorporating Surface/Interface Energies

P. Sharma
Department of Mechanical Engineering,
University of Houston,
Houston, TX 77204
e-mail: psharma@uh.edu

S. Ganti
General Electric Global Research Center,
Niskayuna, NY 12309

The classical formulation of Eshelby (Proc. Royal Society, **A241***, p. 376, 1957) for embedded inclusions is revisited and modified by incorporating the previously excluded surface/interface stresses, tension and energies. The latter effects come into prominence at inclusion sizes in the nanometer range. Unlike the classical result, our modified formulation renders the elastic state of an embedded inclusion size-dependent making possible the extension of Eshelby's original formalism to nano-inclusions. We present closed-form expressions of the modified Eshelby's tensor for spherical and cylindrical inclusions. Eshelby's original conjecture that only inclusions of the ellipsoid family admit uniform elastic state under uniform stress-free transformation strains must be modified in the context of coupled surface/interface-bulk elasticity. We reach an interesting conclusion in that only inclusions with a constant curvature admit a uniform elastic state, thus restricting this remarkable property only to spherical and cylindrical inclusions. As an immediate consequence of the derivation of modified size-dependent Eshelby tensor for nano-inclusions, we also formulate the overall size-dependent bulk modulus of a composite containing such inclusions. Further applications are illustrated for size-dependent stress concentrations on voids and opto-electronic properties of embedded quantum dots.* [DOI: 10.1115/1.1781177]

1 Introduction

Eshelby's linear elastic solution of an embedded inclusion, [1], has a distinguished place in the history of mechanics, materials science, and solid-state physics. Characterized by its insightful thought experiments, Eshelby's classic solution of the embedded inclusion has been fruitfully used in diverse areas and problems of physical sciences, e.g., localized thermal heating, residual strains, dislocation-induced plastic strains, phase transformations, overall or effective elastic, plastic and viscoplastic properties of composites, damage in heterogeneous materials, quantum dots, microstructural evolution; to name just a few. In this work, we seek to modify the classical elasticity original solution of an embedded inclusion to include surface/interface energies, tension and stresses. In the following we will simply use the word ''surface'' to signify both the free surface of a void in a material or the interface of a solid inclusion with that of the surrounding host matrix. As has been done tacitly in most elastic problems, the original elastic solution of the embedded inclusion ignored surface energies of the inclusion—for fairly good reasons. Surface energies only enter physics when surface to volume ratio becomes appreciable. For most technological problems (until recently) inclusions were of the order of microns and rarely were one concerned with nano-inclusions or related size effects. At the micron and higher length scales, the surface-to-volume ratios are negligible and indeed Eshelby's original assumptions hold true and so does his solution. In short, in the present work we seek to derive Eshelby's tensor in the context of coupled surface-bulk elasticity that then can be utilized to capture at least part of the length scale effects likely to be prominent for embedded nano-inclusions.

The literature on Eshelby's tensor and related problems is indeed rich and extensive. While we can hardly do justice in our review of all pertaining work, attempt is made in Section 2 to identify some pertinent literature. To make this article self-contained, a brief description of Eshelby's main conclusions in the classical elasticity context are also reviewed. In Section 3, we formulate the general problem of an embedded inclusion incorporating surface energy and related terms. Some simple closed-form expressions can be obtained for inclusions of constant curvature (i.e., spherical and cylindrical shapes). That is the object of Section 4. The inhomogeneity problem is briefly discussed in Section 5 after which several applications of this work are presented in Section 6 closing finally with summary and conclusions in Section 7.

2 Background

By way of introduction, consider a localized arbitrarily shaped region (Ω) in a material undergoing a stress-free inelastic deformation. Such strains are referred to as either transformation strains, [1], or eigenstrains, [2]. Various physical examples of such strains are thermal expansion, dislocation mediated inelastic strain, swelling strain, magnetomechanical strains, lattice mismatch, and so forth. If the inclusion is removed from the material and allowed to relax (thus enacting the eigenstrain), no stress is generated. However, due to the presence of the matrix or surrounding material, the final relaxed elastic state of the inclusion admits a state of stress. When the material properties of the inclusion and the matrix are the same, the problem of determining the elastic state is often referred to as Eshelby's first problem. The scenario where the inclusion elastic properties are different than those of the matrix is Eshelby's second problem (in which case the inclusion is referred to as an ''inhomogeneity''). This nomen-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, November 25, 2003; final revision, February 13, 2004. Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

clature was introduced by Mura [2]. Eshelby's interior or exterior tensor (**S** or **D**) relates the eigenstrain ($\boldsymbol{\varepsilon}^{\star}$) to the actual strain ($\boldsymbol{\varepsilon}$) in and out of the inclusion, [1,3,4]:

$$\boldsymbol{\varepsilon}(\mathbf{X}) = \mathbf{S}(\mathbf{X}) : \boldsymbol{\varepsilon}^{\star}(\mathbf{X}) \quad \mathbf{X} \in \Omega \tag{1a}$$

$$\boldsymbol{\varepsilon}(\mathbf{X}) = \mathbf{D}(\mathbf{X}) : \boldsymbol{\varepsilon}^{\star}(\mathbf{X}) \quad \mathbf{X} \notin \Omega \tag{1b}$$

Both boldfaced and index notation will be used as convenient. Eshelby's tensor, in classical elasticity context, depends solely on shape (i.e., aspect ratios) of the inclusions and is thus size-independent. Furthermore, for the family of ellipsoidal shaped inclusions (including spheres, cylinders, spheroids), this tensor is uniform within the interior of the inclusion. The latter fact greatly facilitates, for example, the calculation of effective properties of composites containing ellipsoidal inhomogeneities. For the sake of completeness, some additional details on classical Eshelby's tensor are recorded in Appendix A while an exhaustive account can be found in Refs. [1–4]. For nonuniform eigenstrains the Eshelby tensors are integral operators while for uniform eigenstrains numerical values can be established either analytically or numerically (depending upon the geometrical and material symmetry complications).

Since the original appearance of Eshelby's paper [1] several works have extended, modified and applied the concept of Eshelby's tensor to a diverse set of physical problems. There exist extensive reviews of this subject hence only selected representative papers are cited to establish appropriate context. We will, however, allude to some review articles; the references of which more or less contain an updated account of this topic:

(1) Anisotropicity: Several works have modified the classic (originally isotropic) formulation to incorporate anisotropic behavior. Progress has largely been made only in the plane case. An excellent, but somewhat dated, account of these aspects is given in the now classic monograph by Mura [2]. Some more recent works, which also contain extensive list of references on this subject, are: Ru [5] who discusses arbitrary shaped inclusions in anisotropic half and full plane, Li and Dunn [6] address coupled field anisotropic inclusion problems, Pan and Yang [7] who present a semi-analytical method for application to embedded quantum dots and Faux and Pearson [8] who have also applied an anisotropic formulation to quantum dots.

(2) Inclusion shapes: Chiu [9] has considered parallelepiped inclusion. Rodin [10] considers the general polyhedral inclusion. So do Nozaki and Taya [11].

(3) Bonding conditions of inclusion: The original assumption in Eshelby's work is that the inclusion is perfectly bonded to the matrix, i.e., the normal tractions are continuous and so are the displacements. Under certain conditions these conditions must be relaxed (e.g., grain boundary sliding, diffusive sliding, etc.). Various researchers have considered the imperfectly bonded inclusion, e.g., Furuhashi et al. [12], Ru and Schiavone [13], Zhong and Meguid [14], Qu [15,16] and Kouris et al. [17] to name a few.

(4) Coated inclusions: Frequently for technological reasons inclusions are embedded in a matrix with a coating (or which may be developed due chemical interaction with the matrix). A few representative works in this area are: Walpole [18], Luo and Weng [19], Cherkaoui et al. [20], among many others.

(5) Coupled problems: Due to possible applications in sensor and actuator technology, a large body of work has focused on coupled problems, e.g., magnetorestrictive inclusions, piezoelectric media, etc. See, for example, the works by Taya [21], Ru [22], Deng and Meguid [23], Mikata [24], Li and Dunn [6], and Pan [25,26].

(6) Nonuniform eigenstrains: Sendeckyj [27] and Moschovidis [28] considered general polynomial eigenstrains. Their work is useful for both nonuniform loadings as well as for taking into account interactions between inhomogeneities. Asaro and Barnett [29] and Mura and Kinoshita [30] addressed polynomial eigenstrains in an anisotropic media. Note also must be made of the recent work of Rahman [31] who presents simplified calculations of Eshelby type tensors for polynomial eigenstrains.

(7) Enriched elasticity: the classical theory of elasticity itself has been modified in several ways. Micromorphic elasticity takes into account additional microdegrees-of-freedom such as independent rotations, dilations, and shears. An extensive account of these theories can be found in Eringen [32]. As far as inclusion problem are concerned, it appears that the only two solutions that exist are due to Cheng and He [33,34] who, respectively, solve the spherical and cylindrical inclusion problem. Based on the latter work, Sharma and Dasgupta [35] have formulated the overall properties of micropolar composites.

In addition to the aforementioned group of papers, several other works exist in the context of nonlinear behavior and of course in application areas (such as effective medium theories, phase transformations, stability, among others.). A review of those works is beyond the scope of this paper. The following monographs, review articles, books and references therein are recommended for the interested reader: Mura [2], Nemat-Nasser and Hori [36], and Markov and Preziosi [37], Weng et al. [38], Bilby et al. [39], and Mura et al. [40].

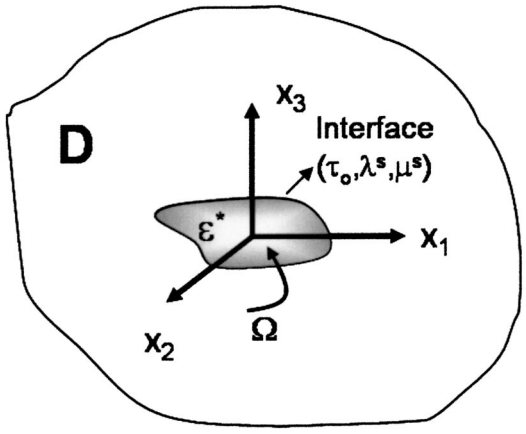


Fig. 1 Schematic of the problem

3 The General Size-Dependent Inclusion Problem in Coupled Bulk-Surface Elasticity

Consider, for now, an arbitrary shaped inclusion (Ω) embedded in an infinite amount of material. By definition of an inclusion, we suppose a prescribed stress-free transformation strain within the domain of the inclusion (Fig. 1). Consider the eigenstrain to be uniform. As a departure from the classical solution, we now require that the interface of the inclusion and the matrix be endowed with a deformation-dependent interfacial energy, Γ. The interfacial or surface energy is positive definite. This quantity is distinct from the bulk deformation-dependent energy due to the different coordination number of the surface/interface atoms, different bond lengths, angles, and a different charge distribution, [41]. Within the assumptions of infinitesimal deformations and a continuum field theory, the concept of surface stress and surface tension can be clarified by the following relation between interface/surface stress tensor, $\boldsymbol{\sigma}^{s}$, and the deformation-dependent surface energy, $\Gamma(\boldsymbol{\varepsilon}^{s})$ by

$$\boldsymbol{\sigma}^{s} = \tau_o \mathbf{I}^2 + \frac{\partial \Gamma}{\partial \boldsymbol{\varepsilon}^{s}}. \tag{2}$$

Where applicable, superscripts *B* and *S* indicate bulk and surface, respectively. Here, $\boldsymbol{\varepsilon}^{s}$ is the 2×2 strain tensor for surfaces or interfaces, $\mathbf{I}^2$ represents the identity tensor for surfaces while τ_o is

the deformation-independent surface/interfacial tension. It is worth pointing out that the concepts of surface tension, surface stress, and surface energy are often confused and used interchangeably. Only for liquids are all three the same. For solids, they are vastly different and must be carefully distinguished. See, for example, the excellent review article by Ibach [41]. A further source of confusion often is the sign of the surface stress. The latter can be negative but not the surface energy. Conceptual mistakes abound if one equates the surface stress to the surface energy. The determination of the surface tension and the surface elastic constants is often nontrivial and a discussion on this can be found elsewhere, e.g., Ibach [41], Miller and Shenoy [42], and Gurtin and Murdoch [43]. Some brief comments on this are also provided in Appendix C where the properties used in subsequent numerical calculations are listed.

Having introduced the essential concept of surface elasticity, the governing linearized isotropic equations can be written following Gurtin and co-workers, [44,45] who (along with previous works) can be credited for setting the theory of surface elasticity on a rational mechanics footing. The equilibrium and isotropic constitutive equations of bulk elasticity are written as usual:

$$\text{div}\,\boldsymbol{\sigma}^B = 0 \tag{3a}$$

$$\boldsymbol{\sigma}^B = \lambda \mathbf{I}^3 \text{Tr}(\boldsymbol{\varepsilon}) + 2\mu\boldsymbol{\varepsilon}. \tag{3b}$$

At the interface, the concept of surface or interface elasticity, [43–45], is introduced which is excluded in the classical elasticity formulation:

$$[\boldsymbol{\sigma}^B.\mathbf{n}] + \text{div}_s\,\boldsymbol{\sigma}^S = 0 \tag{4a}$$

$$\boldsymbol{\sigma}^S = \tau_o \mathbf{I}^2 + 2(\mu^s - \tau_o)\boldsymbol{\varepsilon}^S + (\lambda^s + \tau_o)\text{Tr}(\boldsymbol{\varepsilon}^S)\mathbf{I}^2. \tag{4b}$$

Here, λ and μ are the Lame' constants for the isotropic bulk material. Isotropic interfaces or surfaces can be characterized by surface Lame' constants λ^s, μ^s and surface tension, τ_o. $\mathbf{n}$ is the normal vector on the interface. It is to be noted that only certain strain components appear within the constitutive law for surfaces due to the 2×2 nature of the surface stress tensor (i.e., strains normal to the surface are excluded). Thus, $\mathbf{I}^2$ represents the 2×2 identity tensor while $\mathbf{I}^3$ represents the same for bulk 2nd rank tensor. Tr indicates the trace operation. The square brackets in Eq. (4*a*) indicate the jump of the field quantities across the interface. In absence of surface terms, Eq. (4) reduce to the usual normal traction continuity equations of classical elasticity. "div_s" represent the surface divergence. To define this further and well as the role of surface identity tensor $\mathbf{I}^2$ (i.e., the 2×2 nature of surface tensors), it would be convenient to first recall certain projection tensors ($\mathbf{P}^s$) employed by Gurtin et al. [44]:

$$\mathbf{P}^s = \mathbf{I} - \mathbf{n}\otimes\mathbf{n}. \tag{5}$$

Here $\mathbf{I}$ is the three-dimensional identity tensor and we have dropped the superscript "3." This surface projection tensor maps tensor fields from bulk to surface and vice versa. For example, the surface strain tensor projected into such a tangent space would be written as $\boldsymbol{\varepsilon}^s = \mathbf{P}^s\boldsymbol{\varepsilon}\mathbf{P}^s$. This notion of projection tensor (and related tensor machinery of superficial and tangential tensors) elegantly allows one to mix bulk and surface tensors in the same equations. To clarify the notion of surface divergence, consider a vector $\mathbf{v}$. The surface gradient and surface divergence, then, take the following form, [44]:

$$\begin{aligned} \boldsymbol{\nabla}_s\mathbf{v} &= \boldsymbol{\nabla}\mathbf{v}\mathbf{P}^s \\ \text{div}_s(\mathbf{v}) &= \text{Tr}(\boldsymbol{\nabla}_s\mathbf{v}). \end{aligned} \tag{6}$$

Noting that the transformation strain is only nonzero within the inclusion domain ($\mathbf{x}\in\Omega$), we can write the bulk-constitutive law for the inclusion-matrix as follows:

$$\boldsymbol{\sigma}^B = \mathbf{C}:\{\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^\star H(z(\mathbf{x}))\}. \tag{7}$$

Here "H" is the Heaviside function and $\mathbf{C}$ is the classic fourth-order stiffness tensor. We define $z(\mathbf{x})$ to be of the form

$$\begin{aligned} &\{z(\mathbf{x}) > 0 \,|\, \mathbf{x}\in\Omega\} \\ &\{z(\mathbf{x}) < 0 \,|\, \mathbf{x}\notin\Omega\}. \end{aligned} \tag{8}$$

Taking the divergence of Eq. (7) we obtain

$$\nabla.\boldsymbol{\sigma}^B = \nabla.(\mathbf{C}:\boldsymbol{\varepsilon}) - \nabla.\{\mathbf{C}:\boldsymbol{\varepsilon}^\star H(z(\mathbf{x}))\} - \underline{[\boldsymbol{\sigma}^B.\mathbf{n}]\delta(\bar{z}(\mathbf{x}))} = \mathbf{0}. \tag{9}$$

It can be readily seen that the eigenstrain and the underlined term appear as a body force. Note that in classical elasticity the last underlined expression in Eq. (9), i.e., $[\boldsymbol{\sigma}^B.\mathbf{n}]$ is typically omitted since the jump in the normal tractions is zero. $\delta(.)$ is the Dirac delta function while $\bar{z}(\mathbf{x}) = 0$ defines the interface. However, taking cognizance of Eq. (4*a*), i.e., coupling interface elasticity with bulk elasticity, we must rewrite Equation (9) as

$$\nabla.\boldsymbol{\sigma}^B = \nabla.(\mathbf{C}:\boldsymbol{\varepsilon}) - \underline{\nabla.\{\mathbf{C}:\boldsymbol{\varepsilon}^\star H(z(\mathbf{x}))\} + \delta(\bar{z}(\mathbf{x}))\text{div}_s\,\boldsymbol{\sigma}^s} = \mathbf{0}. \tag{10}$$

Using the underlined term as representing a body force in conjunction with the elastic Green's function, we can write the displacement field due to both the eigenstrain and the surface effect as

$$\mathbf{u} = -\int_V \mathbf{G}^T(\mathbf{y}-\mathbf{x}).(\nabla.\{\mathbf{C}:\boldsymbol{\varepsilon}^\star H(\mathbf{y})\})dV_y + \underline{\int_S \mathbf{G}^T(\mathbf{y}-\mathbf{x}).\text{div}_s\,\boldsymbol{\sigma}^s(\mathbf{y})dS_y}. \tag{11}$$

A more rigorous treatment of the interface conditions in Eq. (9)–(11) is provided in Appendix A. Here $\mathbf{G}$ is the Green's tensor for isotropic classical elasticity (Appendix B). The underlined term indicates the extra surface terms that we have incorporated in the present work. The first integral in Eq. (11) is simply the classical part. As customary, [1,36], we make use of Gauss theorem to cast Eq. (11) in a more attractive form:

$$\mathbf{u} = \int_V (\mathbf{C}:\boldsymbol{\varepsilon}^\star):(\nabla_x\otimes\mathbf{G}^T(\mathbf{y}-\mathbf{x}))dV_y + \underline{\int_S \mathbf{G}^T(\mathbf{y}-\mathbf{x}).\text{div}_s\,\boldsymbol{\sigma}^s(\mathbf{y})dS_y}. \tag{12}$$

Here we have also used the rule that, $\nabla_x\mathbf{G}(\mathbf{y}-\mathbf{x}) = -\nabla_y\mathbf{G}(\mathbf{y}-\mathbf{x})$. Invoking the linearized strain-displacement law: $\boldsymbol{\varepsilon} = \text{sym}\{\nabla\otimes\mathbf{u}\}$, we can then write

$$\boldsymbol{\varepsilon} = \mathbf{S}:\boldsymbol{\varepsilon}^\star + \text{sym}\left\{\nabla_x\otimes\underline{\int_S \mathbf{G}^T(\mathbf{y}-\mathbf{x}).\text{div}_s\,\boldsymbol{\sigma}^s(\mathbf{y})dS_y}\right\}. \tag{13}$$

Here we have invoked the definition of the classical size-independent Eshelby tensor, [1,2], based on which the first (non-underlined) integral in Eq. (12) reduces to the classical expression in Eq. (1). The notation, sym{.}, represents the symmetric part of a second-order tensor, $\mathbf{A}$, e.g.,

$$\text{sym}\{\mathbf{A}\} = \tfrac{1}{2}\{\mathbf{A} + \mathbf{A}^T\}.$$

Further simplification does not appear feasible without additional assumptions regarding inclusion shape. Note now that Eq. (13) implicitly gives the modified Eshelby's tensor for inclusions incorporating surface energies. This relation is implicit since the

surface stress depends on the surface strain, which in turn is the projection of the conventional strain ($\boldsymbol{\varepsilon}$) on the tangent plane of the inclusion-matrix interface. In the next section, using Eq. (13) we will derive explicit expressions for cylindrical and spherical inclusions. For now, however, it is worth noting some general features of the new Eshelby tensor.

In terms of the surface projection tensor the surface divergence of the surface stress tensor can be written as

$$\mathrm{div}_s\, \boldsymbol{\sigma}^s = \mathrm{div}_s\{\mathbf{C}^s\mathbf{P}^s\boldsymbol{\varepsilon}\mathbf{P}^s + \tau_o\mathbf{P}^s\}. \tag{14}$$

The surface divergence of surface stress tensor can only be uniform if the classical ''bulk'' strain as *well* as the projection tensor is uniform over the inclusion surface. Consider that, [44]:

$$\mathrm{div}_s\, \mathbf{P}^s = 2\kappa\mathbf{n}. \tag{15}$$

Here κ is the mean curvature of the inclusion. For a general ellipsoid the curvature is nonuniform and varies depending upon the location at the surface. Only for the special cases of spherical and cylindrical shape is the mean curvature uniform hence leading us to conclude the following:

PROPOSITION: *Eshelby's original conjecture that only inclusions of the ellipsoid family admit uniform elastic state under uniform eigenstrains must be modified in the context of coupled surface/interface-bulk elasticity. Only inclusions that are of a constant curvature admit a uniform elastic state, thus restricting this remarkable property to spherical and cylindrical inclusions.*

4 Inclusions With Constant Curvature (Spheres and Cylinders)

Spherical and cylindrical inclusions are endowed with a constant curvature and thus according to the previous section must admit a uniform elastic state in coupled bulk-surface elasticity. The new Eshelby's tensor will, of course, be size-dependent because of the presence of curvature terms.

Due to the constant curvature, Eq. (13) can be simplified considerably. The surface divergence of the surface stress can be simply taken out of the differential and integral operators. The surface integral is converted into a volume integral and we can then write:

$$\boldsymbol{\varepsilon} = \mathbf{S}:\boldsymbol{\varepsilon}^\star - \mathbf{C}^{-1}:\left(\mathrm{sym}\left\{\nabla_x\otimes C:\underbrace{\int_V \nabla_x\otimes\mathbf{G}(\mathbf{y}-\mathbf{x})dV_y}\right\}\right):2\kappa^s\mathbf{I} \tag{16}$$

where scalar ''s'' is defined from the relation:

$$\boldsymbol{\sigma}^s = s\mathbf{P}^s$$

$$\Rightarrow s = \tau_o + (\lambda^s+\mu^s)\mathrm{Tr}(\mathbf{P}^s\boldsymbol{\varepsilon}\mathbf{P}^s) \tag{17}$$

In the underlined integral term we have multiplied and divided by the elastic stiffness tensor to conveniently cast the term enclosed in the curly brackets in terms of the classical Eshelby tensor. Additionally we have used the surface constitutive law (Eq. 4(b)). We can rewrite Eq. (16) in the following simpler form:

$$\boldsymbol{\varepsilon} = \mathbf{S}:\boldsymbol{\varepsilon}^\star - (2\kappa s)\mathbf{C}^{-1}:(\mathbf{S}:\mathbf{I}). \tag{18}$$

Equation (18) can be made more explicit by noting that an isotropic fourth tensor, **A**, displaying the symmetries characteristic of the elastic stiffness tensor can be written in terms of two scalars a_1 and a_2 as: $A_{ijkl} = a_1\delta_{ij}\delta_{kl} + a_2(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})$. It is then straightforward to show that, $\mathbf{A}:\mathbf{I} = (3a_1+2a_2)\mathbf{I}$ which, after substituting $\mathbf{C}^{-1}:\mathbf{I} = (1/3K)\mathbf{I}$ in Eq. (18) directly leads to the following for spherical inclusions:

$$\boldsymbol{\varepsilon} = \mathbf{S}:\boldsymbol{\varepsilon}^\star - \frac{K^s}{3KR_o}(\mathbf{S}:\mathbf{I})\mathrm{Tr}(\mathbf{P}^s\boldsymbol{\varepsilon}\mathbf{P}^s) - \frac{2\tau_o}{3KR_o}(\mathbf{S}:\mathbf{I}). \tag{19}$$

Here we have used the fact that $\kappa = 1/R_o$ for spheres where R_o is the radius. K^s is defined by us to be the surface elastic modulus and is given as $2(\lambda^s+\mu^s)$ while K is the usual hydrostatic modulus, $\lambda + 2\mu/3$.

For an infinite circular cylindrical inclusion, in addition to the plane-strain conditions we have κ as $1/2R_o$ hence

$$\boldsymbol{\varepsilon} = \mathbf{S}:\boldsymbol{\varepsilon}^\star - \frac{K'^s}{3K'R_o}(\mathbf{S}:\mathbf{I})\mathrm{Tr}(\mathbf{P}^s\boldsymbol{\varepsilon}\mathbf{P}^s) - \frac{\tau_o}{3K'R_o}(\mathbf{S}:\mathbf{I}). \tag{20}$$

Here K'^s is the plane-strain surface modulus, $\lambda^s + 2\mu^s$ while K' is $2(\lambda+\mu)/3$. Note that for the interior solution, Eshelby's interior tensor (**S**) must be used while for exterior solution the corresponding exterior version (**D**) is required.

Substituting the well-known components of the classical Eshelby tensors for both spherical and cylindrical shapes, [2], we obtain the following simple expressions for spherical and cylindrical inclusions subjected to a dilatational eigenstrain, $\varepsilon^*_{11} = \varepsilon^*_{22} = \varepsilon^*_{33} = \varepsilon^*$.

Spherical inclusion (in spherical polar coordinates):

$$\varepsilon_{rr}(r) = \varepsilon_{\theta\theta}(r) = \varepsilon_{\phi\phi}(r) = \left.\frac{3K^M\varepsilon^* - 2\tau_o/R_o}{4\mu^M + 3K^M + 2K^s/R_o}\right| r < R_o \tag{21a}$$

$$\varepsilon_{rr}(r) = \left.\left[\frac{3K^M\varepsilon^* - 2\tau_o/R_o}{4\mu^M + 3K^M + 2K^s/R_o}\right]\frac{R_o^3}{r^3}\right| r > R_o \tag{21b}$$

$$\varepsilon_{\theta\theta}(r) = \varepsilon_{\phi\phi}(r) = -\left.\left[\frac{3K^M\varepsilon^* - 2\tau_o/R_o}{4\mu^M + 3K^M + 2K^s/R_o}\right]\frac{2R_o^3}{r^3}\right| r > R_o\,. \tag{21c}$$

Cylindrical inclusion (in cylindrical polar coordinates):

$$\varepsilon_{rr}(r) = \varepsilon_{\theta\theta}(r) = \left.\frac{3K'^M\varepsilon^* - \tau_o/R_o}{2\mu^M + 3K'^M + K'^s/R_o}\right| r < R_o \tag{22a}$$

$$\varepsilon_{rr}(r) = \left.\left[\frac{3K'^M\varepsilon^* - \tau_o/R_o}{2\mu^M + 3K'^M + K'^s/R_o}\right]\frac{R_o^2}{r^2}\right| r > R_o \tag{22b}$$

$$\varepsilon_{\theta\theta}(r) = -\left.\left[\frac{3K'^M\varepsilon^* - \tau_o/R_o}{2\mu^M + 3K'^M + K'^s/R_o}\right]\frac{R_o^2}{r^2}\right| r > R_o \tag{22c}$$

$$\varepsilon_{zz}(r) = 0. \tag{22d}$$

Wherever applicable, superscripts *H* and *M* will be used to represent inhomogeneity and matrix properties, respectively. The expressions (21)–(22) are exceptionally simple but clearly illustrate that elastic state is now size-dependent. The surface/interface tension is a residual strain-type term which, for example, should not impact the effective properties of composite. The effect of surface elasticity appears through K^s which (as shall be seen in Section 6(b)), leads to a size-dependent change in overall hydrostatic properties of a composite. By making the radius of the inclusion large we can trivially retrieve the known classical solution. Interestingly, although their treatment of a spherical precipitate was much more specialized, we can make contact with the results of Cahn and Larche [46]. Using an assumed displacement type method they (only taking into account surface tension) presented exactly the expression in Eq. (21*a*) with the surface elasticity effect (K^s) set to zero.

5 A Note on the Eshelby's 2nd Problem (Inhomogeneity)

Since the classical result for the strain within the inclusion is uniform for ellipsoids, Eshelby [1] was able to devise an elegant method to mimic an inhomogeneity by an inclusion containing a fictitious eigenstrain. The so-called equivalent inclusion method

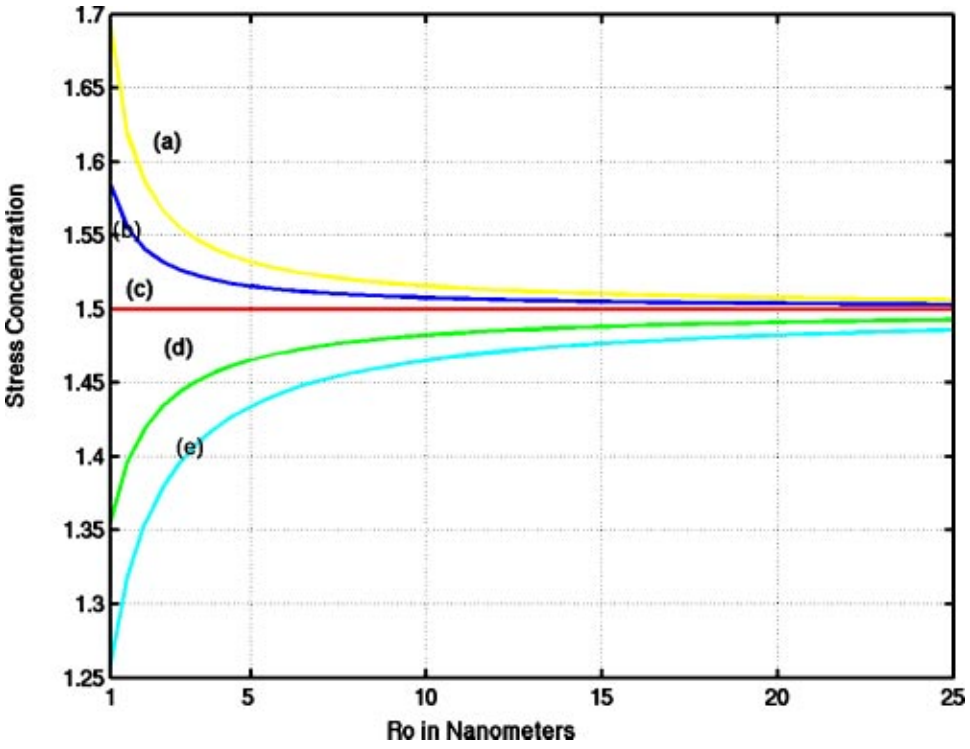


Fig. 2 Stress concentration as a function of surface properties and void radius. (*a*) Solution with surface modulus=2K^s, Al [1 0 0]. (*b*) Solution with surface modulus=nominal K^s for Al [1 0 0]. (*c*) Classical solution without surface effects, i.e., $K^s=0$. (*d*) Solution with surface modulus=2K^s Al [1 1 1]. (*e*) Solution with surface modulus=nominal K^s, Al [1 1 1].

simply entails equating the elastic state of an inhomogeneity to that of an inclusion albeit with the aforementioned fictitious eigenstrain, i.e.,

$$\mathbf{C}^M:\{\boldsymbol{\varepsilon}^\infty+\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^\star\}=\mathbf{C}^H:\{\boldsymbol{\varepsilon}^\infty+\boldsymbol{\varepsilon}\} \tag{23a}$$

$$\boldsymbol{\varepsilon}=\mathbf{S}:\boldsymbol{\varepsilon}^\star. \tag{23b}$$

$\boldsymbol{\varepsilon}^\infty$ is the externally applied strain. Since for spherical and cylindrical shapes the modified Eshelby's tensor with surface effects is also uniform, the equivalent inclusion method embodied in Eqs. 23(*a*,*b*) can be easily applied to study the size-dependent elastic state of inhomogeneities. Unlike the classical case, this fortuity, as was seen in Section 3, does not extend to ellipsoids. As an example, for a cylindrical *inhomogeneity* free of any external loading but containing a dilatational eigenstrain, the interior radial stress can be reduced to

$$\sigma_{rr}=2(\mu^H+\lambda^H)\frac{-2\mu^M\varepsilon^*-K'^s\varepsilon^*/R_o-\tau_o/R_o}{2(\mu^M+\mu^H+\lambda^H)+K'^s/R_o}. \tag{24}$$

6 Applications

The incorporation of surface size effects in the inclusion problem automatically reopens all the existing application areas of Eshelby tensor now extendable to the nanoscale. In the present paper we discuss three application areas: the first two are academic although classic in mechanics while the third is, currently, of immense technological importance.

(a) Size-Dependent Stress Concentration at a Spherical Void. Consider a spherical void under an applied hydrostatic tension. Based upon the preceding expressions (using Eshelby's exterior tensor and equivalent inclusion method), the stress concentration can then be derived to be

$$S.C.=\left.\frac{\sigma_{\theta\theta}}{\sigma_\infty}\right|_{r=R_o}=1+\frac{1}{2}\left(\frac{1-2K^s/3K^MR_o}{1+K^s/2\mu^MR_o}\right). \tag{25}$$

In this section we have set $\tau_o=0$, to study the effect of surface elastic constant, K^s. Results can then also be presented independent of loading conditions since the surface tension is a residual stress type of effect. Note that Eq. (25) trivially gives 1.5 as the stress concentration for the classical elasticity case when either surface modulus is small or void radius is relatively large (typically>25 nm). The numerical results are presented for Aluminum using free-surface properties computed by previous researchers ([42]—using molecular dynamics simulations). The surface properties are highly dependent upon crystallographic direction while ours is an isotropic formulation. The object of this section, however, is to simply use some realistic values to illustrate the physical effect. The stress concentration of the spherical cavity under hydrostatic tension is plotted as a function of the cavity radius R_o in Fig. 2 for two different set of surface properties (corresponding to [1 0 0] Al and [1 1 1] Al). To investigate a broader range of surface properties, curves of surface modulus twice that of Al [1 0 0] and Al [1 1 1] are also shown. The parameters as obtained from manipulation of data from Miller and Shenoy [42] are: [1 0 0], $\lambda^s=3.48912$ N/m, $\mu^s=-6.2178$ N/m$\Rightarrow K^s=-5.457$ N/m, while for [1 1 1], they are: $\lambda^s=6.842$ N/m, $\mu^s=-0.3755$ N/m$\Rightarrow K^s=12.932$ N/m. As depicted in Fig. 2, surface effects cause the stress concentration to reduce (increase) with decreasing pore size when $K^s>0$ ($K^s<0$). The classical case (without surface effects) corresponds to $K^s=0$ and is, as expected, independent of pore size. Below a critical void radius the void will sinter. This effect is closely related to the residual surface tension and is not investigated here.

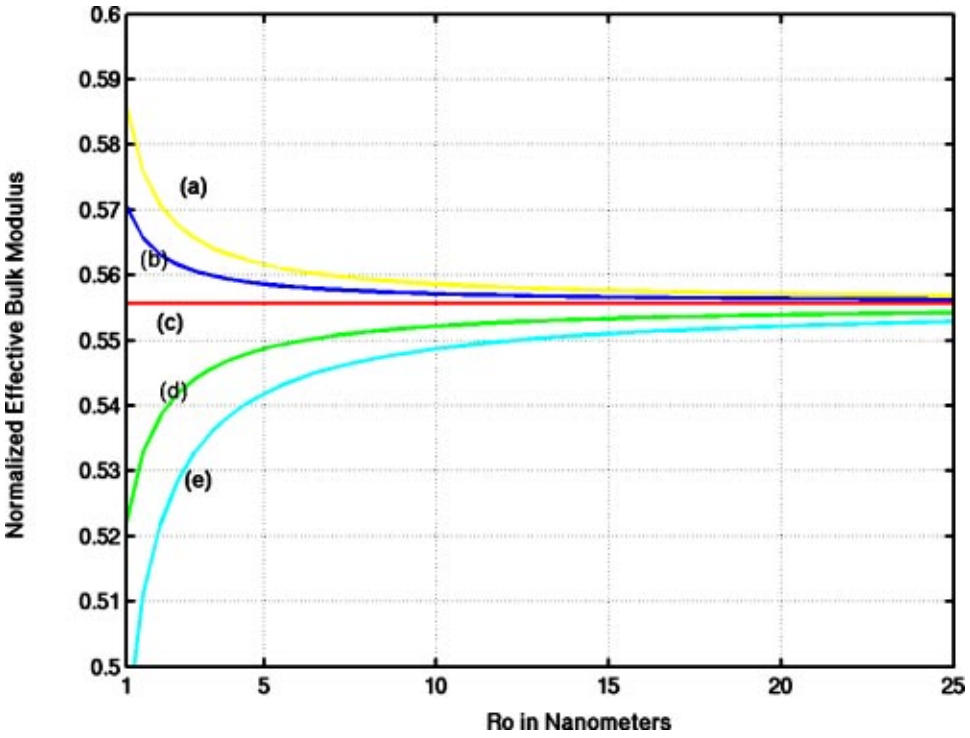


Fig. 3 Size-dependent effective hydrostatic modulus with surface effects versus void radius normalized with the matrix bulk modulus: (*a*) solution with surface modulus=2K^s, Al [1 0 0]; (*b*) solution with surface modulus=nominal K^s for Al [1 0 0]; (*c*) classical solution without surface effects, i.e., $K^s=0$; (*d*) solution with surface modulus=2K^s Al [1 1 1]; (*e*) solution with surface modulus=nominal K^s, Al [1 1 1]

(b) Size-Dependent Overall Properties of Composites One of the applications of this work is that surface effects can now be easily included in the determination of effective elastic properties of a composite. Here, as an example, a nanocomposite of Aluminum populated with a finite volume fraction of spherical voids is considered. Let the volume fraction of the inhomogeneities be denoted by "c." To take into account interactions between various inhomogeneities, we embed the single inhomogeneity (void) in a concentric spherical volume of matrix material with finite radius "$R_M > R_o$." Then, "c" is simply R_o^3/R_M^3. This is nothing but the classical spherical assemblage system, [36]. Using the interior and exterior Eshelby's tensor we can compute the displacement fields to be

$$u=\begin{cases} \Pr, & 0\leqslant r\leqslant R_o \\ Q+\dfrac{T}{r^2}, & R_o\leqslant r\leqslant R_M \end{cases}$$

$$Q=\frac{\sigma^{\infty}(4\mu^M+3K^H)}{3K^M(4\mu^M+3K^H)-4c\mu^M[3\Delta K+2K^s/R_o]} \tag{26}$$

$$\Delta K=K^M-K^H,\quad T=\frac{3\Delta K R_o^3}{4\mu^M+3K^H}Q,\quad P=Q+T/R_o^3.$$

Here we have used the kinematical relations ($\varepsilon_{rr}=\partial u/\partial r;\varepsilon_{\theta\theta}=\varepsilon_{\phi\phi}=u/r$). The overall applied stress is related to the total average strain via the effective bulk modulus as

$$\sigma^{\infty}=K^{\text{eff}}\langle\varepsilon\rangle. \tag{27}$$

As is well known, [36], the average strain can be completely determined through the surface integral of the displacement on the boundary of sphere R_M, i.e.,

$$\langle\boldsymbol{\varepsilon}\rangle=\frac{1}{V}\int_{S_M}u\mathbf{n}\otimes\mathbf{n}dS. \tag{28}$$

Here, $\mathbf{n}$ is the normal vector on the outer surface. The average strain field is obtained as

$$\langle\boldsymbol{\varepsilon}\rangle=3\left(Q+\frac{3K^M}{4\mu^M}Q-\frac{1}{4\mu^M}\right)\sigma^{\infty} \tag{29}$$

from which the effective hydrostatic modulus is deduced to be

$$K^{\text{eff}}=\frac{1}{3\left(Q+\dfrac{3K^M}{4\mu^M}Q-\dfrac{1}{4\mu^M}\right)}. \tag{30}$$

The overall hydrostatic modulus of the composite is size and surface-property-dependent (via "Q" which in turn depends on "K^s" weighted by the inhomogeneity curvature). The size effect is illustrated in Fig. 3, where the normalized effective hydrostatic modulus is plotted against the void radius for a constant volume fraction of $c=0.5$. The effective hydrostatic modulus with surface effects, shown in Fig. 3, is normalized by the hydrostatic modulus of the matrix material without voids. As can be observed, at small length scales, the size of inhomogeneities (at constant volume fraction) can cause a change in the macroscopic behavior of a composite. Asymptotically, as the inhomogeneity (void) size is increased, the surfaces effects begin to diminish and the normalized modulus approaches the classical solution.

(c) Size-Dependent Strain and Emission Wavelength in Quantum Dots. Quantum dots (QDs) have recently been the focus of several experimental and theoretical researchers due to the promise of improved and new opto-electronic properties, [47]. QDs are typically embedded in another semiconductor material

with differing elastic constants and lattice parameter. The ensuing elastic relaxation within the QD is well known to impact their opto-electronic properties. Several works, of varying sophistication (both analytical and numerical), have focused on the calculation of the strain state in buried quantum dots and the subsequent impact on opto-electronic properties (see, for example, the following works appearing in the mechanics literature: [48,49]). It would be of interest to see how much error (in strain and electronic properties) is incurred when surface effects are neglected considering that quantum dots are often "fabricated" in the sub 10-nm regime. An electronic property of interest in quantum dots is its bandgap, which in turn affects its emission wavelength. Using a simple effective mass theory, the deformation potential theory, [50], and the size effects from the present work, the shift in quantum dot emission wavelength due to surface effects can be written as

$$\Delta\lambda = \frac{hc(a_c+a_v)[\varepsilon^{cl}-\varepsilon^{s}(R_o)]}{\left(E_g^{\infty}+\dfrac{h^2(m_e^{\star}m_{lh}^{\star})}{8\pi^2R_o^2(m_e^{\star}+m_{lh}^{\star})}+(a_c+a_v)\varepsilon^{s}(R_o)\right)\left(E_g^{\infty}+\dfrac{h^2(m_e^{\star}m_{lh}^{\star})}{8\pi^2R_o^2(m_e^{\star}+m_{lh}^{\star})}+(a_c+a_v)\varepsilon^{cl}\right)}. \tag{31}$$

Here, h is Planck's constant while $m^{\star}$ is the effective mass of the carriers ("e" is electron and "lh" is light hole). E_g^{∞} is the bulk band gap of the material while (a_c+a_v) represents the dilatational deformation potential. ε^{cl} is the classical strain and $\varepsilon^{s}(R_o)$ is the size-dependent strain from the present work that includes surface effects. Note that for the purposes of band structure calculations, the eigenstrain must be subtracted from the compatible strain. The second term on the denominator of Eq. (31) is the usual quantum confinement effect, [51]. As an example, we have used an $In_{32}GaN$ quantum dot system embedded in a GaN matrix. The error in wavelength calculation is shown in Fig. 4. Numerical constants are listed in Appendix C. We have used a simple first-order approximation (Eq. (31)) to capture the wavelength shift and while a more sophisticated treatment is possible (following, say, Ref. [49]) the present expression in Eq. (31) suffices to provide a measure of the severity of surface effects on the opto-electronic properties.

As patent from Fig. 4 the error in wavelength calculation by neglecting the surface size effect is appreciable in certain size ranges. For large QD size, as expected, the classical and the new results are indistinguishable. For very small QD sizes, while surface effects are appreciable so are the quantum confinement effects (which scale as $1/R^2$) and hence dominate. In the "mid-regime" (still at the nanoscale), surface effects have the most impact. To be specific, in this particular material system, a maximum wavelength shift of ~40 nm (for a diameter of ~3 nm) is observed which is large enough to cause a shift of colors and indeed exceeds the strict optoelectronic design tolerances.

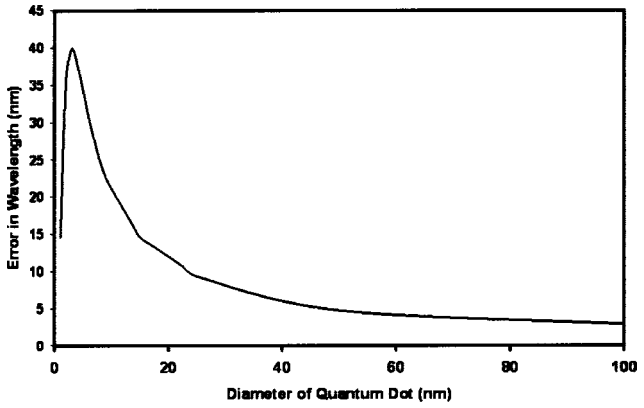


Fig. 4 Size-dependent wavelength shift due to surface elasticity effects

7 Summary and Conclusions

To summarize, we have modified Eshelby's classical approach towards inclusions and inhomogeneities to incorporate the effect of surface energies via the continuum field formulation of surface elasticity. As a consequence, the elastic state of inclusions is rendered size-dependent making possible the establishment of scaling laws that are valid at the nanoscale. Eshelby's original conjecture that only inclusions of the ellipsoid family admit uniform elastic state under uniform eigenstrains must be modified in the context of coupled surface/interface-bulk elasticity. Only inclusions that are of a constant curvature admit a uniform elastic state, thus restricting this remarkable property only to spherical and cylindrical inclusions. The modified size-dependent Eshelby tensor for the spherical and circular cylindrical shape is explicitly calculated in the present work.

Apart from the formal contribution, by way of illustration, the size-dependent stress concentration on a spherical void was demonstrated. Taking advantage of the fact that the modified Eshelby tensor is uniform for the spherical shape, we are also able to derive the exact size-dependent hydrostatic modulus of a heterogeneous solid. Perhaps the most technologically important application of the present work lies in arena of quantum dots and wires. While physicists routinely take into account the impact of strain on band structure and opto-electronic properties, the strain calculations are typically based upon classical elasticity and are size-independent. In the present work, a first-order calculation clearly shows that large errors in both the band structure and the emitted wavelength can be incurred if the surface size effects are neglected.

There are several limitations of the present work and a few are worth mentioning. They point naturally to future extensions:

(1) Isotropic behavior was assumed throughout. This is a rather dubious assumption when one is concerned with surfaces and interfaces. Unfortunately, matters are unlikely to be analytically tractable once the assumption of isotropy is abandoned. Numerical formulation of the coupled-surface bulk elasticity may be necessary to remove this restriction.

(2) Analytical formulas were restricted to the spherical and cylindrical shape. This limits our ability to study the effect of shape on the size-dependent elastic state of nano-inclusions. Derivation of the modified Eshelby tensor for the general ellipsoid (which surely must proceed numerically) would be a useful extension of the present work.

(3) It would be also of interest to see the behavior of nonsmooth inclusion shapes, e.g., parallelepipeds. Polyhedral inclusions with vertices essentially possess zero curvature everywhere except at the corners where singularities exist.

(4) Slip, twist, and wrinkling of surfaces/interfaces were ignored. One can expect some interesting physics to emerge from inclusion of such effects. Slip and twist of elastic interfaces were

recently included by Gurtin et al. [44] to supplement the original formulation, [45]. These notions are closely linked to the concept of coherency-incoherency and their discussion in relation to Eshelby's problems is relegated to a future work.

Acknowledgments

The present work greatly benefited from the comments by an anonymous reviewer. In particular, the rigorous treatment of Eqs. (9) and (16) and the note in Appendix A is a direct consequence of reviewer's suggestion.

Appendix A

Interfacial Conditions in Equations (9)–(11). The form of the underlined term in Eq. (9), i.e., the jump in traction across the inclusion-matrix interface, can be justified by considering the stress balance law in the following form:

$$\text{div}\,\boldsymbol{\sigma} = \int_S f(\mathbf{y})\,\delta(\mathbf{x}-\mathbf{y})\,dS_y = \mathbf{0} \tag{32}$$

Equation (32) defines a stress field perturbed by a force spread over the interface. An arbitrary trial function $\mathbf{w}(\mathbf{x})$ is introduced. Upon multiplication of this trial function with Eq. (32), integration over the volume and subsequent use of Gauss theorem yields

$$\int_S \{[\boldsymbol{\sigma}(\mathbf{x})].\mathbf{n}\}.\mathbf{w}(\mathbf{x})\,dS_x - \int_V \{\text{div}\,\boldsymbol{\sigma}(\mathbf{x})\}.\mathbf{w}(\mathbf{x})\,dS_x + \int_S \mathbf{f}(\mathbf{y}).\mathbf{w}(\mathbf{y})\,dS_y = 0. \tag{33}$$

Since, $\mathbf{w}(\mathbf{x})$ is completely arbitrary, Eq. (33) implies (a) the usual balance law within the bulk of the continuum, $\text{div}\,\boldsymbol{\sigma}=0$ and (b) the identification of the interface force with the jump in the normal tractions, i.e., $\mathbf{f}=-[\boldsymbol{\sigma}].\mathbf{n}$.

Table 1 Numerical values used in Fig. 4

Property	Value
E_g^{∞} (eV)	1.94 [52]
m_e^*	0.18 [53]*
m_{lh}^*	0.8 [54]*
a_c+a_v (eV)	8.3 Chin et al. [55]*
μ^M (Gpa)	67
K^M (Gpa)	102
K^H (Gpa)	168
μ^H (Gpa)	95
K^s (J/m^2)	161.73♣
τ_o (J/m^2)	1.33♠

*Linearly interpolated between InN and GaN using proportion of In concentration.
♣Estimated approximately using Gurtin and Murdoch's [43,45] analogy to membrane theory of Tiersten [56]. A transition from bulk constants to interface/surface properties can be made by the following transformation: $\{\mu^s,\lambda^s\}\rightarrow\{\mu h, 2\lambda\mu h/(\lambda+\mu)\}$. Here, h is the thickness over which surface/interface elasticity behavior differs from the bulk. Our molecular dynamics simulations indicate that such behavior is typically confined to about 1–2 lattice spacing. For the interface, we have assumed this value to be 5 Angstroms.
♠Interfacial tension for this material combination is not known. Conventional EAM potentials are not useful for Nitride structures (especially for surface property evaluation). Generally, however, it is well known that depending upon the degree of coherency the interfacial tension varies from 0.7 J/m^2 to 2 J/m^2. Our estimation proceeded as follows. For an indium concentration of zero, the interfacial tension is also zero while it should be the maximum for 100% In (which is completely incoherent with respect to GaN). Assuming tentatively a maximum interfacial tension of 2 J/m^2 for 100% In, we obtain 1.33 J/m^2 by simple proportion for the current composition of 32%.

Appendix B

Green's Function for Elasticity and Eshelby's Classical Tensor. The Green's function for elasticity $\mathbf{G}(\mathbf{y}-\mathbf{x})$ is the fundamental solution to the Kelvin's solution of a point load in an infinite solid. It is given by (for isotropic materials)

$$\mathbf{G}(\mathbf{y}-\mathbf{x}) = \frac{1}{16\pi\mu(1-\upsilon)|y-x|}\left\{(3-4\upsilon)\,\boldsymbol{\delta} + \frac{(\mathbf{y}-\mathbf{x})\otimes(\mathbf{y}-\mathbf{x})}{|y-x|^2}\right\}. \tag{34}$$

Substituting this expression in the first integral of Eq. (12) yields, [2]:

$$\varepsilon_{ij}(x) = \frac{1}{8\pi(1-\upsilon)}[\Psi_{kl,klij} - 2\upsilon\Phi_{kk,ij} - 2(1-\upsilon)(\Phi_{ik,kj} + \Phi_{jk,ki})] \tag{35}$$

where ψ and Φ are biharmonic and harmonic potentials of the inclusion shape (Ω). They are given as

$$\Psi_{ij}(\mathbf{x}) = \int_\Omega |\mathbf{x}-\mathbf{y}|\,\varepsilon_{ij}^{\star}(\mathbf{y})\,d^3\mathbf{y} \tag{36}$$

$$\Phi_{ij}(\mathbf{x}) = \int_\Omega \frac{1}{|\mathbf{x}-\mathbf{y}|}\,\varepsilon_{ij}^{\star}(\mathbf{y})\,d^3\mathbf{y} \tag{37}$$

Equation (35) can then be cast into the more familiar expression of Eq. (1*a*,*b*)

$$\begin{aligned}\boldsymbol{\varepsilon}(\mathbf{x}) &= \mathbf{S}(\mathbf{x}):\boldsymbol{\varepsilon}^{\star} \quad \mathbf{x}\in\Omega \\ \boldsymbol{\varepsilon}(\mathbf{x}) &= \mathbf{D}(\mathbf{x}):\boldsymbol{\varepsilon}^{\star} \quad \mathbf{x}\notin\Omega.\end{aligned} \tag{38}$$

Mura's book [2] contains detailed listing of **S** and **D** tensor for various inclusion shapes (spheres, cylinders, ellipsoids, and cuboids).

Appendix C

Numerical Constants for Wavelength Shift Calculation The numerical values used in the calculation of Fig. 4 are listed in Table 1.

References

[1] Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion and Related Problems," Proc. R. Soc. London, Ser. A, **A241**, pp. 376–396.
[2] Mura, T., 1987, *Micromechanics of Defects in Solids*, Martinus Nijhoff, Hague, Netherlands.
[3] Eshelby, J. D., 1959, "The Elastic Field Outside an Ellipsoidal Inclusion," Proc. R. Soc. London, Ser. A, **A252**, pp. 561–569.
[4] Eshelby, J. D., 1961, "Elastic Inclusions and Inhomogeneities," *Progress in Solid Mechanics 2*, I. N. Sneddon and R. Hill, eds., North Holland, Amsterdam, pp. 89–140.
[5] Ru, C. Q., 2003, "Eshelby Inclusion of Arbitrary Shape in an Anisotropic Plane or Half-Plane," Acta Mech., **160**(3–4), pp. 219–234.
[6] Li, J. Y., and Dunn, M. L., 1998, "Anisotropic Coupled-Field Inclusion and Inhomogeneity Problems," Philos. Mag. A, **77**(5), pp. 1341–1350.
[7] Pan, E., and Yang, B., 2001, "Elastostatic Fields in an Anisotropic Substrate due to a Buried Quantum Dot," J. Appl. Phys., **90**(12), pp. 6190–6196.
[8] Faux, D. A., and Pearson, G. S., 2000, "Green's Tensors for Anisotropic Elasticity: Application to Quantum Dots," Phys. Rev. B, **62**(8), pp. R4798–R4801.
[9] Chiu, Y. P., 1977, "On the Stress Field due to Initial Strains in Cuboid Surrounded by an Infinite Elastic Space," ASME J. Appl. Mech., **44**, pp. 587–590.
[10] Rodin, G. J., 1996, "Eshelby's Inclusion Problem for Polygons and Polyhedra," J. Mech. Phys. Solids, **44**(12), p. 1977.
[11] Nozaki, H., and Taya, M., 2001, "Elastic Fields in a Polyhedral Inclusion With Uniform Eigenstrains and Related Problems," ASME J. Appl. Mech., **68**, 441.
[12] Furuhashi, R., Huang, J. H., and Mura, T., 1992, "Sliding Inclusions and Inhomogeneities With Frictional Interfaces," ASME J. Appl. Mech., **59**, pp. 783–788.
[13] Ru, C. Q., and Schiavone, P., 1997, "A Circular Inclusion With Circumferentially Inhomogeneous Interface in Antiplane Shear," Proc. R. Soc. London, Ser. A, **A453**, pp. 2551–2572.
[14] Zhong, Z., and Meguid, S. A., 1996, "On the Eigenstrain Problem of a Spherical Inclusion With an Imperfectly Bonded Interface," ASME J. Appl. Mech., **63**, pp. 877–883.

[15] Qu, J., 1993, ''Effects of Slightly Weakened Interfaces on the Overall Elastic Properties of Composite Materials,'' Mech. Mater., **14**, pp. 269–281.
[16] Qu, J., 1993, ''Eshelby Tensor for an Elastic Inclusion With Slightly Weakened Interface,'' ASME J. Appl. Mech., **60**, pp. 1048–1050.
[17] Kouris, D. A., Tsuchida, E., and Mura, T., 1986, ''An Anomaly of Sliding Inclusions,'' ASME J. Appl. Mech., **53**, pp. 724–726.
[18] Walpole, L. J., 1978, ''Coated Inclusion in an Elastic Medium,'' Proc. Cambridge Philos. Soc., **83**, pp. 495–506.
[19] Luo, H. A., and Weng, G. J., 1989, ''On Eshelby's S-Tensor in Three Phase Cylindrically Concentric Solid,'' Mech. Mater., **8**(2–3), pp. 77–88.
[20] Cherkaoui, M., Sabar, H., and Berveiller, M., 1994, ''Micromechanical Approach of the Coated Inclusion Problem and Applications to Composite Problems,'' J. Eng. Mater. Technol., **116**(3), pp. 274–278.
[21] Taya, M., 1999, ''Micromechanics Modeling of Smart Composites,'' Composites, **30A**(4), pp. 531–536.
[22] Ru, C. Q., 2000, ''Eshelby's Problem for Two-Dimensional Piezoelectric Inclusions of Arbitrary Shape,'' Proc. R. Soc. London, Ser. A, **456**(1997), pp. 1051–1068.
[23] Deng, W., and Meguid, S. A., 1999, ''Closed Form Solutions for Partially Debonded Circular Inclusion in Piezoelectric Materials,'' Acta Mech., **137**(3–4), pp. 167–181.
[24] Mikata, Y., 2001, ''Explicit Determination of Piezoelectric Eshelby Tensors for a Spheroidal Inclusion,'' Int. J. Solids Struct., **38**(40–41), pp. 7045–7063.
[25] Pan, E., 2002, ''Elastic and Piezoelectric Fields Around a Quantum Dot: Fully Coupled or Semicoupled Model?'' J. Appl. Phys., **91**(6), pp. 3785–3796.
[26] Pan, E., 2002, ''Three-Dimensional Green's Functions in Anisotropic Magneto-Electro-Elastic Bimaterials,'' Z. Angew. Math. Phys., **53**(5), pp. 815–838.
[27] Sendeckyi, G. P., 1967, ''Ellipsoidal Inhomogeneity Problem,'' Ph.D. dissertation, Northwestern University, Evanston, IL.
[28] Moschovidis, Z. A., 1975, ''Two Ellipsoidal Inhomogeneities and Related Problems Treated by the Equivalent Inclusion Method,'' Ph.D thesis, Northwestern University, Evanston, IL.
[29] Asaro, R. J., and Barnett, D. M., 1975, ''The Non-uniform Transformation Strain Problem for an Anisotropic Ellipsoidal Inclusion,'' J. Mech. Phys. Solids, **23**, pp. 77–83.
[30] Mura, T., and Kinoshita, N., 1978, ''The Polynomial Eigenstrain Problem or an Anisotropic Ellipsoidal Inclusion,'' Phys. Status Solidi A, **48**, pp. 447–450.
[31] Rahman, M., 2002, ''The Isotropic Ellipsoidal Inclusion With a Polynomial Distribution of Eigenstrain,'' ASME J. Appl. Mech., **69**, pp. 593–601.
[32] Eringen, A. C., 1999, *Microcontinuum Field Theories I: Foundations and Solids*, Springer-Verlag, New York.
[33] Cheng, Z. Q., and He, L. H., 1995, ''Micropolar Elastic Fields due to a Spherical Inclusion,'' Int. J. Eng. Sci., **33**(3), pp. 389–397.
[34] Cheng, Z. Q., and He, L. H., 1997, ''Micropolar Elastic Fields due to a Circular Cylindrical Inclusion,'' Int. J. Eng. Sci., **35**(7), pp. 659–686.
[35] Sharma, P., and Dasgupta, A., 2002, ''Average Elastic Fields and Scale-Dependent Overall Properties of Heterogeneous Micropolar Materials Containing Spherical and Cylindrical Inhomogeneities,'' Phys. Rev. B, **66**, p. 224110.
[36] Nemat-Nasser, S., and Hori, M., 1999, *Micromechanics: Overall Properties of Heterogeneous Solids*, Elsevier, New York.
[37] Markov, K., and Preziosi, L., 2000, *Heterogeneous Media: Micromechanics Modeling Methods and Simulations*, Birkhauser Verlag, Switzerland.
[38] Weng, G. J., Taya, M., and Abe, H., eds, 1990, *Micromechanics and Inhomogeneity: The Toshio Mura Anniversary Volume*, Springer-Verlag, New York.
[39] Bilby, B. A., Miller, K. J., and Willis, J. R., 1984, IUTAM/IFC/ICM Symposium on Fundamentals of Deformation and Fracture, Sheffield, England, Apr. 2–5, Eshelby Memorial Symposium, Cambridge University Press, Cambridge, UK.
[40] Mura, T., Shodja, H. M., and Hirose, Y., 1996, ''Inclusion problems,'' Appl. Mech. Rev., **49**(10), Part 2, pp. S118–S127.
[41] Ibach, H., 1997, ''The Role of Surface Stress in Reconstruction, Epitaxial Growth and Stabilization of Mesoscopic Structures,'' Surf. Sci. Rep., **29**(5–6), pp. 193–263.
[42] Miller, R. E., and Shenoy, V. B., 2000, ''Size-Dependent Elastic Properties of Nanosized Structural Elements,'' Nanotechnology, **11**(3), pp. 139–147.
[43] Gurtin, M. E., and Murdoch A. I., 1978, ''Surface Stress in Solids,'' Int. J. Solids Struct., **14**(6), pp. 431–440.
[44] Gurtin, M. E., Weissmuller, J., and Larche, F., 1998, ''The General Theory of Curved Deformable Interfaces in Solids at Equilibrium,'' Philos. Mag. A, **78**, p. 1093.
[45] Gurtin, M. E., and Murdoch, A. I., 1975, ''A Continuum Theory of Elastic Material Surfaces,'' Arch. Ration. Mech. Anal., **59**, p. 389.
[46] Cahn, J. W., and Larche, F., 1982, ''Surface Stress and the Chemical Equilibrium of Small Crystals. II. Solid Particles Embedded in a Solid Matrix,'' Acta Metall., **30**(1), pp. 51–56.
[47] Bimberg, D., Grundmann, M., and Lendenstov, N. N., 1996, *Quantum Dot Heterostructures*, John Wiley and Sons, New York.
[48] Freund, L. B., and Johnson, H. T., 2001, ''Influence of Strain on Functional Characteristics of Nanoelectronic Devices,'' J. Mech. Phys. Solids, **49**, pp. 1925–1935.
[49] Johnson, H. T., and Freund, L. B., 2001, ''The Influence of Strain on Confined Electronic States in Semiconductor Quantum Structures,'' Int. J. Solids Struct., **38**, pp. 1045–1062.
[50] Singh, J., 1992, *Physics of Semiconductors & Their Heterostructures*, McGraw-Hill, New York.
[51] Brus, L. E., 1984, ''Electron-Electron and Electron-Hole Interactions in Small Semiconductor Crystallites: The Size-Dependence of the Lowest Excited Electronic State,'' J. Chem. Phys., **80**(9), p. 4403.
[52] LeBoeuf, S., 2003, private communication.
[53] Mohammad, S. N., and Morkoc, H., 1996, ''Progress and Prospects of Group-III Nitride Semiconductors,'' Prog. Quantum Electron., **20**, p. 361.
[54] Shur, M. S., and Khan, M. A., 1997, ''GaN/AlGaN Heterostructure Devices: Photodetectors and Field-Effect Transistors,'' Mater. Res. Bull., **22**(2), p. 44.
[55] Chin, V. W. L., Tansley, T. L., and Osotchan, T., 1994, ''Electron Mobilities in Gallium, Indium, and Aluminum nitrides,'' J. Appl. Phys., **75**, p. 7365.
[56] Tiersten, H. F., 1969, ''Elastic Surface Waves Guided by Thin Films,'' J. Appl. Phys., **40**, p. 770.

Defect Green's Function of Multiple Point-Like Inhomogeneities in a Multilayered Anisotropic Elastic Solid

B. Yang
Department of Mechanical and Aerospace Engineering,
Florida Institute of Technology,
Melbourne, FL 32901
e-mail: boyang@fit.edu

Defect Green's function (GF) of multiple point-like inhomogeneities in a multilayered solid has been derived within the theory of linear anisotropic elasticity. It is related to the (reference) GF of the multilayered matrix excluding the inhomogeneities through the continuum Dyson's equation. While the reference GF is available, the defect GF can be solved. The expressions are first analytically reduced by realizing the point-likeness of the inhomogeneities. The subsequent procedure involves the solution of the response of each individual inhomogeneity to a far-field straining in the multilayered matrix and a matrix inversion on the order of the number of inhomogeneities. Furthermore, the defect GF is applied to derive the field induced by inhomogeneous substitutions in a multilayered solid. Numerical results are reported for arrays of cubic and semispherical Ge inclusions in a Si/Ge superlattice. The numerical results have demonstrated the validity and efficiency of the present formulation. [DOI: 10.1115/1.1781179]

1 Introduction

Multilayered materials have been the focal research area in structural composites for decades, [1]. Recently, such a material setting has drawn refreshed attention in the field of semiconductor micro/nanodevices, [2]. In both cases, the composites may contain various ''defects,'' such as cavities and reinforcing particles in the structural case, and dislocations, vacancies and impurities in the micro/nanoscale case. Some of these defects may be modeled as inclusions with eigenstrain, introduced by, for instance, a thermal expansion mismatch or a lattice-constant mismatch. The others should be modeled as inhomogeneous inclusions, exhibiting both elastic and eigenstrain mismatches with matrices, [3]. Due to the small size and large quantity of such defects (in real cases), the Green's function (GF) method is often the first choice, sometimes the only one, to investigate the problems. However, the method is limited to the cases of certain materials property and geometry, where the GF is available and can be evaluated efficiently.

In the theory of (three-dimensional) linear elasticity, various GFs have been derived for infinite-space, half-space, bimaterials, trimaterials, and multilayers, [3–9]. The GFs have been directly applied to solve the eigenstrain problem of defects in these structures. The defects are modeled as inclusions, taking into account their eigenstrain mismatch, but ignoring their elastic mismatch, with these matrices [3,10,11]. Recently, Yang and Tewary [12] introduced the continuum Dyson's equation and defect GF in a generally heterogeneous anisotropic solid. The Dyson's equation links the point-force responses of two systems of identical geometry and (homogeneous) boundary condition but of different media. Given the GF of either system (called a reference), the GF of the other (called a defect system with ''defect'' change of materials properties relative to the reference) can be obtained by solving this equation. The defect GF can then be applied directly to solve the problem of inhomogeneous inclusions, taking into account the elastic mismatch as well as the eigenstrain mismatch between the defects and matrix. This defect GF method is ideal for treating localized defect spaces. In the present work, we derive the defect GF for multiple point-like inhomogeneities, and apply it to derive the induced field by point-like inhomogeneous inclusions in a multilayered solid.

In Section 2, the Green's function problem is formulated for multiple inhomogeneities in a multilayered anisotropic and linearly elastic solid subjected to a point force. It is solved by relating the defect GF to the reference GF in the absence of the inhomogeneities through the continuum Dyson's equation, [12]. The expressions are analytically reduced to the case of point-like inhomogeneities before the final numerical solution. In Section 3, the defect GF is used to derive the induced field by inhomogeneous inclusions in a multilayered solid. In Section 4, numerical results are reported for arrays of cubic and semi-spherical GE substitutions (namely, GE quantum dots (QD)) in a Si/Ge superlattice. The present solution of cubic QDs is compared with that of finite-size QDs by a boundary element (BE) method, [13]. The agreement has demonstrated the validity of the present formulation. The solution of more than 100 semisherical QDs has also demonstrated the efficiency of the present formulation. In Section 5, conclusions are drawn.

2 Defect Green's Function of Point-Like Inhomogeneities

Consider a multilayered substrate occupying domain Ω. It is filled with N particles occupying subdomains D_n $(n=1,2,\ldots,N)$, as schematically shown in Fig. 1(*a*). A Cartesian coordinate frame, (x_1,x_2,x_3) is attached to the system, with the x_3-axis being perpendicular to the substrate surface and pointing inward the substrate. The substrate and particles are modeled as anisotropic and linearly elastic materials. The particles and the parts of substrate that are substituted by the particles have generally different elastic stiffness. Thus, these particles are termed

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 Applied Mechanics Division, August 28, 2003; final revision, April 15, 2004. Associate Editor: Z. Suo. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

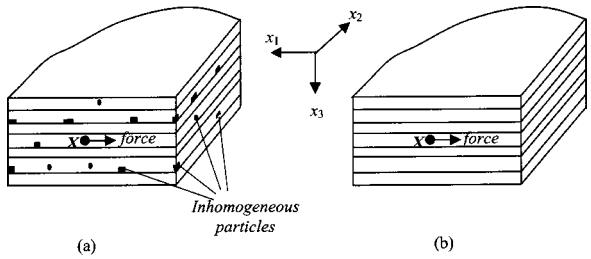


Fig. 1 (*a*) A multilayered solid embedded with multiple inhomogeneities; (*b*) a "clean" multilayered solid as reference to (*a*). Both structures are subjected to a unit point force along one of the axes at X. A global coordinate system is established with the x_3-axis being perpendicular to the top surface and pointing inward the substrate.

inhomogeneities in the substrate, [3]. The interfaces between the substrate and particles and between adjacent layers of the substrate are perfectly bonded. A unit point force is applied to the body along one of the axes, for example, along the *p*th direction, at **X**. In addition, a homogeneous boundary condition is imposed along the substrate surface $\partial\Omega$. The point force and all particles are remote to each other. Thus, the particles are treated as point-like (in size). Meanwhile, they may hold arbitrary shape, which can play a significant role in determining the induced elastic fields.

The constitutive law of the multilayered substrate (before filling particles) is given by

$$\sigma_{ij}(\mathbf{x}) = C_{ijkl}(\mathbf{x})\varepsilon_{kl}(\mathbf{x}), \tag{1}$$

where σ_{ij} is the stress component, $\varepsilon_{kl}(\equiv 1/2(u_{i,j}+u_{j,i}))$ is the infinitesimal strain component, C_{ijkl} is the elastic stiffness component, and the repeated subscripts imply the conventional summation over their range. In the definition of $\boldsymbol{\varepsilon}$, $\mathbf{u}$ is the displacement vector, and the comma in the subscript indicates the partial derivative with respect to the coordinate that follows. The constitutive law of the multilayered substrate after filling particles is given by

$$\sigma_{ij}(\mathbf{x}) = C^*_{ijkl}(\mathbf{x})\varepsilon_{kl}(\mathbf{x}), \tag{2}$$

with

$$C^*_{ijkl}(\mathbf{x}) = \begin{cases} C^{(n)}_{ijkl}(\mathbf{x}) & \mathbf{x}\in D_n \\ C_{ijkl}(\mathbf{x}) & \mathbf{x}\in \Omega - \sum_n D_n \end{cases}, \tag{3}$$

where the superscript *n* in the bracket indicates the attachment to the *n*th particle.

The equilibrium of the multilayered substrate filled with particles requires

$$[C^*_{ijkl}(\mathbf{x})G^*_{pk,l}(\mathbf{X},\mathbf{x})]_{,j} + \delta_{pi}\delta(\mathbf{x}-\mathbf{X}) = 0, \tag{4}$$

where $G^*_{pi}(\mathbf{X},\mathbf{x})$ is the *i*th displacement component at $\mathbf{x}$ due to the unit point force applied along the *p*th direction at $\mathbf{X}$, $\delta(\mathbf{x}-\mathbf{X})$ is the Dirac delta function, and δ_{pi} is the Kronecker delta function. The homogeneous boundary condition is given by

$$G^*_{pi}(\mathbf{X},\mathbf{x}) = 0 \quad \text{or} \quad \Sigma^*_{pij}(\mathbf{X},\mathbf{x})n_j(\mathbf{x}) = 0, \ \mathbf{x}\in\partial\Omega, \tag{5}$$

on each component *i*, where $\Sigma^*_{pij}(\mathbf{X},\mathbf{x}) \equiv C^*_{ijkl}(\mathbf{x})G^*_{pk,l}(\mathbf{X},\mathbf{x})$. Thus, $G^*_{pi}(\mathbf{X},\mathbf{x})$ and $\Sigma^*_{pij}(\mathbf{X},\mathbf{x})$ are, respectively, the Green's displacement and stress that satisfy Eqs. (4) and (5). Through the text, indices *p*, *q*, *s*, and *t* are used to indicate a component of the source point, $\mathbf{X}$ (i.e., the first variable of the GFs), and indices *g*, *h*, *i*, *j*, *k*, and *l* to indicate a component of the field point, $\mathbf{x}$ (i.e., the second variable of the GFs). Indices *m* and *n* are used to number the particles.

The above (defect) GF, $\mathbf{G}^*$ can be related to the (reference) GF, $\mathbf{G}$ of the multilayered substrate before filling the particles (Fig. 1(*b*)) by, [12]

$$G^*_{pi}(\mathbf{X},\mathbf{x}) = G_{pi}(\mathbf{X},\mathbf{x}) - \int_\Omega [G^*_{pj,k}(\mathbf{X},\mathbf{x}')\Delta C_{jksh}(\mathbf{x}')]_{,h} G_{si}(\mathbf{x}',\mathbf{x})dV(\mathbf{x}'), \tag{6}$$

where $\Delta C_{jksh}(\mathbf{x}) \equiv C_{jksh}(\mathbf{x}) - C^*_{jksh}(\mathbf{x})$. The GF, $\mathbf{G}$ is defined by

$$[C_{jikl}(\mathbf{x})G_{pk,l}(\mathbf{X},\mathbf{x})]_{,j} + \delta_{pi}\delta(\mathbf{x}-\mathbf{X}) = 0, \tag{7}$$

under the same homogeneous boundary condition as given in Eq. (5) for $\mathbf{G}^*$. Equation (6) is called the continuum Dyson's equation, the continuum counterpart of the Dyson's equation in the lattice-statics theory, [14]. It links the point-force response of the defect system of multilayered materials filled with inhomogeneous particles to that of the reference system of "clean" multilayered matrix under the same interfacial and boundary conditions. Given the reference GF, $\mathbf{G}$, of which an efficient evaluation scheme has been developed recently, [9], the defect GF, $\mathbf{G}^*$ can be obtained by solving this equation.

In the case of embedded particles, ΔC_{ijkl} is localized inside the subdomains D_n. Applying the Gauss divergence theorem, the above Dyson's Eq. (6) can be reduced to

$$G^*_{pi}(\mathbf{X},\mathbf{x}) = G_{pi}(\mathbf{X},\mathbf{x}) + \sum_n \int_{D_n} G^*_{pk,j}(\mathbf{X},\mathbf{x}') \Delta C^{(n)}_{jkst}(\mathbf{x}')G_{si,t}(\mathbf{x}',\mathbf{x})dV(\mathbf{x}'), \tag{8a}$$

Taking derivative of Eq. (8*a*) with respect to $\mathbf{x}$ yields

$$G^*_{pi,l}(\mathbf{X},\mathbf{x}) = G_{pi,l}(\mathbf{X},\mathbf{x}) + \sum_n \int_{D_n} G^*_{pk,j}(\mathbf{X},\mathbf{x}') \Delta C^{(n)}_{jkst}(\mathbf{x}')G_{si,lt}(\mathbf{x}',\mathbf{x})dV(\mathbf{x}'). \tag{8b}$$

In the above derivation from Eqs. (1) to (8), we have not used the property of point-likeness of the particles. The last Eqs. (8a,b) are applicable to embedded particles of any size, shape, and elastic property. In the following, we shall analytically reduce the integral by realizing the property of point-likeness of the particles. Although in Eqs. (8a,b) the distance between $\mathbf{X}$ and any $\mathbf{x}'$ is large, it is generally invalid to take $G^*_{pk,j}(\mathbf{X},\mathbf{x}')$ out of the integral by approximating it as constant within each particle. It may vary significantly inside a particle—it may be singular if the particle has an asterisk shape, for example. However, one may expect that the field at the location of a particle is regular if that particle is excluded and filled with the original matrix material. This leads us to write the defect GF $G^*_{pk,j}(\mathbf{X},\mathbf{x})$ as follows:

$$G^*_{pj,k}(\mathbf{X},\mathbf{x}) = G^{*(-n)}_{pg,h}(\mathbf{X},\mathbf{x})s^{(n)}_{ghjk}(\mathbf{x}), \tag{9}$$

where $G^{*(-n)}_{pg,h}(\mathbf{X},\mathbf{x})$ is the (less) defect GF in the presence of all but the *n*th particles, and $s^{(n)}_{ghjk}(\mathbf{x})$ is a forth-rank tensor that transforms $G^{*(-n)}_{pg,h}(\mathbf{X},\mathbf{x})$ into $G^*_{pk,j}(\mathbf{X},\mathbf{x})$ upon the inclusion of the *n*th particle. It is a function of location $\mathbf{x}$, different for particles of different shape and material property. Substituting Eq. (9) in Eqs. (8a,b) yields

$$G^*_{pi}(\mathbf{X},\mathbf{x}) = G_{pi}(\mathbf{X},\mathbf{x}) + \sum_n \int_{D_n} G^{*(-n)}_{pg,h}(\mathbf{X},\mathbf{x}')s^{(n)}_{ghjk}(\mathbf{x}') \Delta C^{(n)}_{jkst}(\mathbf{x}')G_{si,t}(\mathbf{x}',\mathbf{x})dV(\mathbf{x}'), \tag{10a}$$

$$G^*_{pi,l}(\mathbf{X},\mathbf{x}) = G_{pi,l}(\mathbf{X},\mathbf{x}) + \sum_n \int_{D_n} G^{*(-n)}_{pg,h}(\mathbf{X},\mathbf{x}')s^{(n)}_{ghjk}(\mathbf{x}') \Delta C^{(n)}_{jkst}(\mathbf{x}')G_{si,lt}(\mathbf{x}',\mathbf{x})dV(\mathbf{x}'). \tag{10b}$$

If both **X** and **x** are remote to all particles, Eqs. (10a,b) can be reduced to

$$G^*_{pi}(\mathbf{X},\mathbf{x}) = G_{pi}(\mathbf{X},\mathbf{x}) + \sum_n G^{*(-n)}_{pg,h}(\mathbf{X},\mathbf{x}^{(n)}) T^{(n)}_{ghst} G_{si,t}(\mathbf{x}^{(n)},\mathbf{x}), \tag{11a}$$

$$G^*_{pi,l}(\mathbf{X},\mathbf{x}) = G_{pi,l}(\mathbf{X},\mathbf{x}) + \sum_n G^{*(-n)}_{pg,h}(\mathbf{X},\mathbf{x}^{(n)}) T^{(n)}_{ghst} G_{si,lt}(\mathbf{x}^{(n)},\mathbf{x}), \tag{11b}$$

with

$$T^{(n)}_{ghst} = \int_{D_n} s^{(n)}_{ghkj}(\mathbf{x}) \Delta C^{(n)}_{jkst}(\mathbf{x}) dV(\mathbf{x}), \tag{12}$$

where $\mathbf{x}^{(n)}$ represents the location of the nth particle and may be chosen arbitrarily inside the particle with minimal effect on the final solution. Thus, given $G^{*(-n)}_{pg,h}(\mathbf{X},\mathbf{x}^{(n)})$ and $T^{(n)}_{ghst}$ for all particles, the defect GF, $G^*_{pi,l}(\mathbf{X},\mathbf{x})$ can be evaluated by using Eqs. (11a,b).

In order to find $T^{(n)}_{ghst}$, we consider the multilayered matrix with only the nth particle. Since the above derivation applies to the special case, we have

$$G^{*(n)}_{pi,l}(\mathbf{X},\mathbf{x}) = G_{pi,l}(\mathbf{X},\mathbf{x}) + \int_{D_n} G^{*(n)}_{pj,k}(\mathbf{X},\mathbf{x}') \Delta C^{(n)}_{jkst}(\mathbf{x}') G_{si,lt}(\mathbf{x}',\mathbf{x}) dV(\mathbf{x}'), \tag{13}$$

$$G^{*(n)}_{pj,k}(\mathbf{X},\mathbf{x}) = G_{pg,h}(\mathbf{X},\mathbf{x}) s^{(n)}_{ghjk}(\mathbf{x}). \tag{14}$$

Substituting Eq. (14) in Eq. (13) and approximating $G_{pg,h}(\mathbf{X},\mathbf{x}')$ as constant for $\mathbf{x}' \in D_n$ yield

$$G_{pg,h}(\mathbf{X},\mathbf{x}) s^{(n)}_{ghil}(\mathbf{x}) = G_{pi,l}(\mathbf{X},\mathbf{x}) + G_{pg,h}(\mathbf{X},\mathbf{x}^{(n)}) \int_{D_n} s^{(n)}_{ghjk}(\mathbf{x}') \Delta C^{(n)}_{jkst}(\mathbf{x}') G_{si,lt}(\mathbf{x}',\mathbf{x}) dV(\mathbf{x}'). \tag{15}$$

Setting $\mathbf{x} \in D_n$ and eliminating $G_{pg,h}(\mathbf{X},\mathbf{x})$ result in

$$s^{(n)}_{ghil}(\mathbf{x}) = \delta_{gi}\delta_{hl} + \int_{D_n} s^{(n)}_{ghjk}(\mathbf{x}') \Delta C^{(n)}_{jkst}(\mathbf{x}') G_{si,lt}(\mathbf{x}',\mathbf{x}) dV(\mathbf{x}'). \tag{16}$$

The above integral equation can be used to numerically solve for the **s** tensor (as a function of location **x**). To do so, the domain D_n is discretized into a number of elements. The field $s^{(n)}_{ghil}(\mathbf{x})$ is approximated by interpolating the nodal values of each element. Substituting the approximated field in Eq. (16), and assigning **x** to be the location of each node, a closed set of algebraic equations is obtained. It can be solved for the nodal values of $s^{(n)}_{ghil}$. Finally, the nodal values of $s^{(n)}_{ghil}$ are surrendered in Eq. (12) for approximate evaluation of $T^{(n)}_{ghst}$.

In order to find $G^{*(-n)}_{pg,h}(\mathbf{X},\mathbf{x}^{(n)})$, we set only **X** to be remote to all particles and **x** to be at $\mathbf{x}^{(m)}$, location of the mth particle in Eq. (10*b*). It results in

$$G^{*(-m)}_{pg,h}(\mathbf{X},\mathbf{x}^{(m)}) s^{(m)}_{ghil}(\mathbf{x}^{(m)}) = G_{pi,l}(\mathbf{X},\mathbf{x}^{(m)}) + \sum_{n\neq m} G^{*(-n)}_{pg,h}(\mathbf{X},\mathbf{x}^{(n)}) T^{(n)}_{ghst} G_{si,lt} \times (\mathbf{x}^{(n)},\mathbf{x}^{(m)}) + G^{*(-m)}_{pg,h}(\mathbf{X},\mathbf{x}^{(m)}) \times \int_{D_m} s^{(m)}_{ghjk}(\mathbf{x}') \Delta C^{(m)}_{jkst}(\mathbf{x}') G_{si,lt}(\mathbf{x}',\mathbf{x}^{(m)}) dV(\mathbf{x}'). \tag{17}$$

Applying Eq. (16) of the mth particle in the above equation and rearranging yield

$$G^{*(-m)}_{pi,l}(\mathbf{X},\mathbf{x}^{(m)}) = G_{pi,l}(\mathbf{X},\mathbf{x}^{(m)}) + \sum_{n\neq m} G^{*(-n)}_{pg,h} \times (\mathbf{X},\mathbf{x}^{(n)}) T^{(n)}_{ghst} G_{si,lt}(\mathbf{x}^{(n)},\mathbf{x}^{(m)}). \tag{18}$$

Equation (18) offers N equations with N unknowns $G^{*(-n)}_{pg,h}$. Thus, provided that **G** and $\mathbf{T}^{(n)}$ are known, the above system of equations can be solved for all $G^{*(-n)}_{pg,h}$.

In summary, given the reference GF, **G** (and its derivatives) of the multilayered matrix in Fig. 1(*b*), one may first solve Eq. (16) for $\mathbf{s}^{(n)}$ and consequently $\mathbf{T}^{(n)}$ of each individual particle. Then, Eq. (18) is solved for $G^{*(-n)}_{pg,h}$ of all particles. Finally, Eqs. (11a,b) are taken to compute $G^*_{pj,k}$ between any **X** and **x** remote to all particles. Often the derivatives of G^*_{pj} and $G^*_{pj,k}$ with respect to source coordinate **X**, i.e., $G^*_{pj,q}$ and $G^*_{pj,kq}$, are desired. They are the Green's fields due to a force dipole. Their solution procedure is very similar to the preceding one for G^*_{pj} and $G^*_{pj,k}$. For the sake of brevity, it is not repeated here.

3 Application to Eigenstrain/Eigenstress Problem

Often embedded particles (i.e., defects) exhibit thermal-expansion/lattice-constant mismatch as well as elastic mismatch with the matrix. The elastic mismatch has been accounted for above by developing the defect GF based on the continuum Dyson's equation. The thermal-expansion/lattice-constant mismatch may be modeled as eigenstrain within the particles. Such particles with both elastic and eigenstrain mismatches are termed inhomogeneous inclusions, [3]. The constitutive law of the layered solid embedded with inhomogeneous inclusions is given by

$$\sigma_{ij}(\mathbf{x}) = C^*_{ijkl}(\mathbf{x})(\varepsilon_{kl}(\mathbf{x}) - \varepsilon^0_{kl}(\mathbf{x})), \tag{19}$$

with $C^*_{ijkl}(\mathbf{x})$ given by Eq. (3) and eigenstrain field given by

$$\varepsilon^0_{kl}(\mathbf{x}) = \begin{cases} \varepsilon^{0(n)}_{kl}(\mathbf{x}) & \mathbf{x} \in D_n \\ 0 & \mathbf{x} \in \Omega - \sum_n D_n \end{cases}. \tag{20}$$

Within the defect GF formulation, we can derive the (equilibrium) displacement field induced by the inhomogeneous inclusions in the multilayered structure as, [12],

$$u_p(\mathbf{X}) = \sum_n \int_{D_n} G^*_{pi,j}(\mathbf{X},\mathbf{x}) \sigma^0_{ij}(\mathbf{x}) dV(\mathbf{x}), \tag{21}$$

where $G^*_{pi,j}(\mathbf{X},\mathbf{x})$ is the defect GF of the multilayered structure including the inhomogeneities as described in the previous section, and $\sigma^0_{ij} (\equiv C^*_{ijkl}\varepsilon^0_{kl})$ is the eigenstress. If the eigenstress is uniform within each particle, Eq. (21) is reduced to

$$u_p(\mathbf{X}) = \sum_n G^{*(-n)}_{pi,j}(\mathbf{X},\mathbf{x}^{(n)}) S^{(n)}_{ijkl} \sigma^{0(n)}_{kl}, \tag{22}$$

with

$$S^{(n)}_{ijkl} = \int_{D_n} s^{(n)}_{ijkl}(\mathbf{x}) dV(\mathbf{x}). \tag{23}$$

The **S** tensor is essentially the Eshelby's tensor, [3]. Taking derivative of Eq. (22) with respect to **X** and applying the definition of strain, the induced strain is obtained as

$$\varepsilon_{pq}(\mathbf{X}) = \sum_n \frac{1}{2}(G^{*(-n)}_{pi,jq}(\mathbf{X},\mathbf{x}^{(n)}) + G^{*(-n)}_{qi,jp}(\mathbf{X},\mathbf{x}^{(n)})) S^{(n)}_{ijkl} \sigma^{0(n)}_{kl}. \tag{24}$$

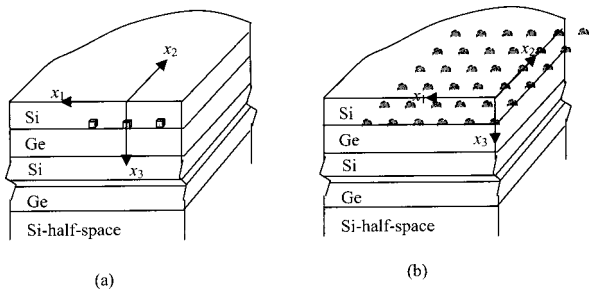


Fig. 2 A Si/Ge superlattice with (*a*) 3 cubic Ge-QDs and (*b*) 11×11 semispherical QDs

4 Numerical Examples

In this section, we apply the previous theory to solve the problem of Ge-QDs in a semiconductor superlattice with alternating Si/Ge layers on a half-space Si substrate. The problem holds strong technological interest owing to its potential application to novel nanodevices, [2]. The superlattice consists of five Si layers and five Ge layers. The layers have the same thickness, *t*. Both materials are cubic anisotropic. The elastic constants for Si are $C_{11}=165.8$, $C_{12}=63.9$ and $C_{44}=79.6$, and for Ge are $C_{11}=128.5$, $C_{12}=48.3$, $C_{44}=66.8$ (GPa), in their crystallographic base axes, [15]. The layers are aligned with their crystallographic orientations, respectively, parallel to the global axes (x_1,x_2,x_3). The Ge-QDs hold a hydrostatic eigenstrain of magnitude equal to 0.04, i.e., $\varepsilon^0_{kl}=\varepsilon^0\delta_{kl}$, and $\varepsilon^0=0.04$, relative to the Si matrix. Two examples are considered. In the first example, three cubic Ge-QDs are buried in the middle plane of the first Si layer, as shown in Fig. 2(*a*). They are aligned in the x_1-direction and separated in a distance of $0.5t$. In the second example, a square array of 11×11 semispherical Ge-QDs are embedded at the bottom of the first Si layer—right above the interface, as shown in Fig. 2(*b*). The array is aligned with the global x_1 and x_2-axes. The distance between the centers of nearest neighbors is *t*. The reference GF, **G** of the multilayered substrate is evaluated by using the efficient scheme recently developed by Yang and Pan [9].

In the first example (Fig. 2(*a*)), the induced strain is evaluated along $(x_1,0,0)$ on the top surface. The corresponding solution of finite-sized QDs is also obtained for comparison by using the BE method, [13]. The nonzero components of the normalized strain, $\varepsilon_{ij}t^3/\varepsilon^0 V$, where *V* is the volume of a QD, are plotted in Fig. 3. It can be seen that the finite-size solution converges toward the present solution of point-like QDs as the size decreases. The convergence has demonstrated the validity of the present formulation.

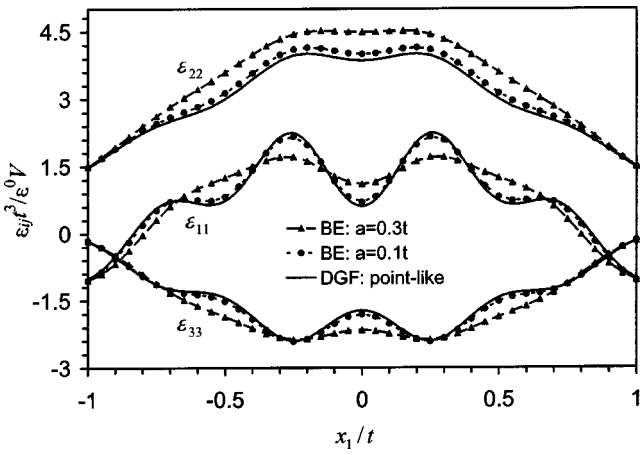


Fig. 3 Variation of (nonzero) normalized strain components along the line, $(x_1, x_2=0, x_3=0)$ on the top surface induced by the three buried cubic QDs as shown in Fig. 2(*a*). The solid lines indicate the present solution of point-like QDs, and the dashed lines with symbols indicate the BE solution of finite-size QDs with the size as indicated.

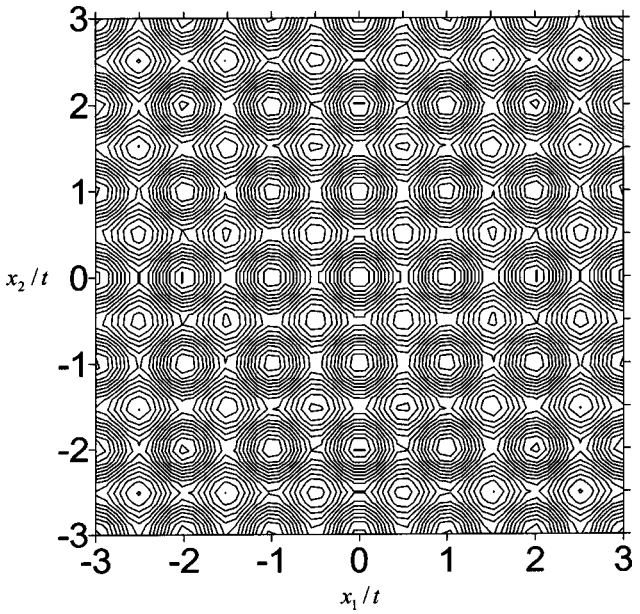


Fig. 4 Contour plot of the normalized hydrostatic strain ε_{kk} on the top surface in the case of 11×11 array of semispherical QDs as shown in Fig. 2(*b*). The magnitude of the contours can be read from the next Fig. 5.

In the second example (Fig. 2(*b*)), the induced field of hydrostatic strain $\varepsilon_{kk}(=\varepsilon_{11}+\varepsilon_{22}+\varepsilon_{33})$ is evaluated in a square area on the top surface. It is normalized by $\varepsilon_{kk}t^3/\varepsilon^0 V$, and plotted in contour in Fig. 4. The variation along the diagonal line from $(-3,-3)$ to $(3,3)$ is also plotted in Fig. 5. It can be seen that a minimum hydrostatic strain occurs above each QD. In between every four adjacent QDs, a maximum hydrostatic strain occurs. The average magnitude of ε_{kk} increases away from the central QD in this case. The solution of this example of 11×11 QDs in a ten-layer superlattice on a semi-infinite substrate has demonstrated the efficiency of the present formulation.

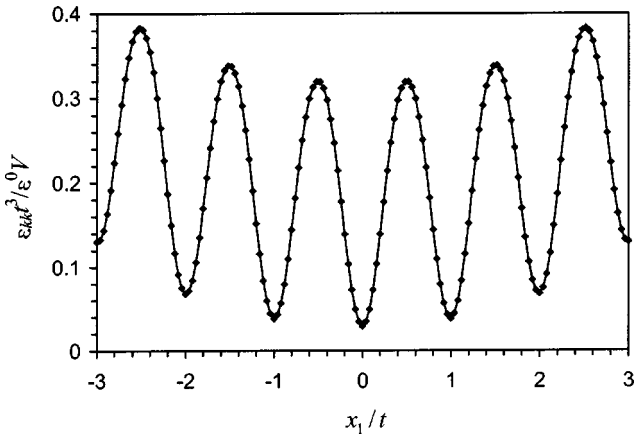


Fig. 5 Diagonal variation of the normalized hydrostatic strain ε_{kk} from $(-3,-3)$ to $(3,3)$ in Fig. 4, showing the magnitude and maximum and minimum locations

5 Conclusion

We have developed a defect GF for multiple point-like inhomogeneities in a multilayered solid within the theory of linear anisotropic elasticity. The defect GF is related to the reference GF of the multilayered matrix through the continuum Dyson's equation. Since the reference GF is available, the defect GF can be solved. In the present case of point-like inhomogeneities, the continuum Dyson's equation is analytically reduced, and the final numerical treatment involves the solution of a fourth-rank tensor of each individual particle and a matrix inversion on the order of the number of inhomogeneities. The fourth-rank tensor reserves the information of particle shape while the particles are treated as point-like in size—the size effect is neglected. Furthermore, the defect GF is applied to derive the field induced by multiple inhomogeneous inclusions that exhibit both elastic and eigenstrain mismatches with matrix. Numerical results are reported for the arrays of cubic and semispherical Ge-QDs in a Si/Ge superlattice. The numerical results have demonstrated the validity and efficiency of the present formulation.

References

[1] Lee, J. A., and Mykkanen, D. L., 1987, *Metal and Polymer Matrix Composites*, Noyes Data Corp, Park Ridge, NJ.

[2] Harrison, P., 2002, *Quantum Wells, Wires and Dots: Theoretical and Computational Physics*, John Wiley and Sons, New York.

[3] Mura, T., 1987, *Micromechanics of Defects in Solids*, Martinus Nijhoff, Boston.

[4] Mindlin, R. D., 1936, ''Force at a Point in the Interior of a Semi-Infinite Solid,'' Physica (Amsterdam), **7**, pp. 195–202.

[5] Love, A. E. H., 1944, *A Treatise on the Mathematical Theory of Elasticity*, Dover, New York.

[6] Ting, T. C. T., 1996, *Anisotropic Elasticity*, Oxford University Press, Oxford, UK.

[7] Pan, E., and Yuan, F. G., 2000, ''Three-Dimensional Green's Functions in Anisotropic Bimaterials,'' Int. J. Solids Struct., **37**, pp. 5329–5351.

[8] Yang, B., and Pan, E., 2002, ''Three-Dimensional Green's Functions in Anisotropic Trimaterials,'' Int. J. Solids Struct., **39**, pp. 2235–2255.

[9] Yang, B., and Pan, E., 2002, ''Efficient Evaluation of Three-Dimensional Green's Functions in Anisotropic Elastostatic Multilayered Composites,'' Eng. Anal. Boundary Elem., **26**, pp. 355–366.

[10] Pan, E., and Yang, B., 2001, ''Elastostatic Fields in an Anisotropic Substrate due to a Buried Quantum Dot,'' J. Appl. Phys., **90**, pp. 6190–6196.

[11] Yang, B., and Pan, E., 2003, ''Elastic Fields of Quantum Dots in Multilayered Semiconductors: A Novel Green's Function Method,'' ASME J. Appl. Mech., **70**, pp. 161–168.

[12] Yang, B., and Tewary, V. K., 2003, ''Continuum Dyson's Equation and Defect Green's Function in a Heterogeneous Anisotropic Solid,'' Mech. Res. Commun., **31**, pp. 405–414.

[13] Yang, B., and Pan, E., 2002, ''Elastic Analysis of an Inhomogeneous Quantum Dot in Multilayered Semiconductors Using a Boundary Element Method,'' J. Appl. Phys., **92**, pp. 3084–3088.

[14] Tewary, V. K., 1973, ''Green-Function Method for Lattice Statics,'' Adv. Phys., **22**, pp. 757–810.

[15] Schaffler, F., 2001, ''Silicon-Germanium ($Si_{1-x}Ge_x$),'' *Properties of Advanced Semiconductor Materials*, M. E. Levinshtein, S. L. Rumyantsev, and M. S. Shur, eds., John Wiley and Sons, New York.

The Maximal Lyapunov Exponent for a Three-Dimensional Stochastic System

K. M. Liew
Professor,
Fellow ASME,
e-mail: mkmliew@ntu.edu.sg

X. B. Liu

Center for Advanced Numerical Engineering Simulations,
School of Mechanical and Production Engineering,
Nanyang Technological University,
Nanyang Avenue,
Singapore 639798

This paper examines the almost-sure asymptotic stability condition of a linear multiplicative stochastic system, which is a linear part of a co-dimension two-bifurcation system that is on a three-dimensional central manifold and subjected to parametric excitation by an ergodic real noise. The excitation is assumed to be an integrable function of an n-dimensional Ornstein-Uhlenbeck vector process which is the output of a linear filter system, while both the detailed balance condition and the strong mixing condition are removed. Through a perturbation method and the spectrum representations of the Fokker Planck operator and its adjoint operator of the linear filter system, the explicit asymptotic expressions of the maximal Lyapunov exponent for three case studies, in which different forms of the coefficient matrix included in the noise excitation term are assumed, are obtained. [DOI: 10.1115/1.1782648]

1 Introduction

The investigation of the maximal Lyapunov exponent for dynamical systems which are excited by stochastic processes is the primary research focus in the fields of random dynamical systems and stochastic bifurcation. This is mainly attributed to the fact that for a linear stochastic system, the Lyapunov exponent is analogous to the real part of the eigenvalue, and this Lyapunov exponent characterizes the exponential rate of change of the response of a random system. Therefore, the sample or the almost-sure stability of the stationary solution of a random dynamical problem depends on the sign of the maximal Lyapunov exponent.

A general method for exact evaluation of the maximal Lyapunov exponent of a linear Ito stochastic differential equation was first presented by Khasminskii [1]. The main idea for this method is such that by projecting the system in space R^n onto the surface of an n-dimensional unit sphere, on which the stochastic differential equation for the variable $\rho=\log\|\mathbf{x}\|$ can be expressed explicitly in terms of the $n-1$ independent angle processes, which constitute themselves a $(n-1)$-dimensional diffusion process, the maximal Lyapunov exponent for this system can then be obtained. This method was then successfully employed by Mitchell and Kozin [2], Nishioka [3], and Ariaratnam and Xie [4] and Xie [5] to a two-dimensional Ito system. Among the works reported to date, limited results pertain to the cases of ergodic and nonwhite noise processes. In Arnold et al. [6], a perturbation method for asymptotic analysis was presented and employed to construct the asymptotic expansion of the maximal Lyapunov exponent of a two-dimensional system under a real noise excitation. To keep the solution tractable, the infinitesimal generator associated with the noise process was assumed to be a self-adjoint elliptic diffusion operator with an isolated simple zero eigenvalue.

Utilizing the method of stochastic averaging, the asymptotic expansions for the maximal Lyapunov exponents for two coupled oscillators with a real noise were obtained by Ariaratnam and Xie [4]. Instead of the stochastic averaging method, the same system and subsequently a more general four-dimensional linear stochastic system were also investigated by Namachchivaya and Roessel [7] and Doyle and Namachchivaya [8] by using the perturbation method.

For a linear multiplicative stochastic system, which was a linear part of a co-dimension two-bifurcation system that was driven by a parametric excitation of a real noise with a small intensity, Namachchivaya and Talwar [9] obtained an analytical expression of maximal Lyapunov exponent for the case where a diagonal matrix involved in the real noise excitation term. The co-dimension two-bifurcation system considered was on a three-dimensional central manifold and possesses one zero eigenvalue and a pair of pure imaginary eigenvalues, meanwhile the excitation was assumed to be a real noise that satisfied the strong mixing condition. The stochastic averaging method was employed to derive a set of approximated Ito equations for the (r,ϕ,z) process which was the first approximation of the original linear multiplicative stochastic system.

For the case that the stochastic excitation was given as the first component of an output of a linear filter system and conformed to the detailed balance condition [10], Liu and Liew [11,12] obtained the asymptotical expansions of the top Lyapunov exponents. In their works, a model of enhanced generality was considered, in which they removed the strong mixing condition which is the prerequisite of the stochastic averaging method. Instead the spectrum representation of the Fokker Planck operator of the linear filter system, [10,13], were employed in the construction of the asymptotic expansions of the stationary probability density functions and the top Lyapunov exponents for the relevant systems.

The present study attempts to obtain an asymptotic expansion of the maximal Lyapunov exponent for the linear multiplicative stochastic system, which is a further extension of the work in Liu and Liew [11,12]. For the excitation, both the detailed balance condition and the strong mixing condition are removed and the excitation itself is assumed to be an integrable function of the n-dimensional Ornstein-Uhlenbeck process. It is well known that the asymptotic expression of the top Lyapunov exponent depends on the form of matrix **B**, which is included in the noise excitation term. In this paper, a general form of matrix **B** is considered and furthermore, for the special cases of three different matrices **B** in which the complexity of the singular points of a one-dimensional phase diffusion process arises, we investigate the phenomena arising from these singular points and discuss in detail our findings.

This paper is organized as follows. Section 2 details the formulation of the problem. In Section 3, we recall the research results of Karlin and Taylor [14], Roy [13], and Liberzon and Brockett

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, October 2, 2003; final revision, April 13, 2004. Associate Editor: N. Sri Namachchivaya. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

[15] concerning the spectrum representation of the Fokker Planck operator and its adjoint of the linear filter system. In Sections 4 and 5, the asymptotical analysis is applied to obtain the expansion of the stationary probabilistic density function. The top Lyapunov exponents for three cases, in which the singularity of the diffusion coefficient arises, are evaluated in Section 6. The conclusions are drawn in Section 7.

2 Formulation

Consider a typical deterministic co-dimension two-bifurcation system which is on a three-dimensional central manifold and possesses one zero-eigenvalue and a pair of pure imaginary eigenvalues, [16],

$$\begin{aligned}\dot{r}&=\mu_1 r+a_1 rz+(a_2 r^3+a_3 r^2 z)+O(|r,z|^4)\\ \dot{z}&=\mu_2 z+c_1 r^2-z^2+(c_2 r^2 z+c_3 z^3)+O(|r,z|^4)\\ \dot{\Theta}&=\omega+O(|r,z|^2)\end{aligned} \tag{1}$$

where μ_1 and μ_2 are the unfolding parameters, and a_1, a_2, a_3, c_1, c_2, c_3, and ω are real constants. This normalized form arises in the classic fluid dynamic stability study of Couette flow, [16]. In the vicinity of equilibrium point $(r,z,\Theta)=(0,0,\omega t)$, via the transformation of $r=[x_1^2+x_2^2]^{1/2}$, $\Theta=\arctan[x_2/x_1]$, $z=x_3$, the model of the linearization of the original system (1), which is subjected to a stochastic parametric perturbation, is obtained as

$$\dot{\mathbf{x}}=\mathbf{A}_0\mathbf{x}-\varepsilon^2\mathbf{A}_1\mathbf{x}+\varepsilon f(\mathbf{u})\mathbf{B}\mathbf{x} \tag{2}$$

where

$$\mathbf{A}_0=\begin{bmatrix}0&\omega&0\\-\omega&0&0\\0&0&0\end{bmatrix},\quad \mathbf{A}_1=\begin{bmatrix}\delta_1&0&0\\0&\delta_1&0\\0&0&\delta_2\end{bmatrix},\quad \mathbf{B}=\begin{bmatrix}b_{11}&b_{12}&b_{13}\\b_{21}&b_{22}&b_{23}\\b_{31}&b_{32}&b_{33}\end{bmatrix} \tag{3}$$

and the parameters μ_1, μ_2 have been rescaled such that

$$\mu_1=-\varepsilon^2\delta_1,\quad \mu_2=-\varepsilon^2\delta_2$$

$f(\mathbf{u})$ is a scalar stochastic function of $\mathbf{u}(t)$, which is a real noise and defined as

$$\dot{\mathbf{u}}(t)=\mathbf{A}\mathbf{u}(t)+\dot{\mathbf{W}}(t) \tag{4}$$

where $\mathbf{A}=(a_{ij})_{n\times n}$; a_{ij} are real or complex numbers. $\dot{\mathbf{W}}(t)$ is an n-dimensional zero-mean Gaussian white noise with $E(\dot{\mathbf{W}}(t+\tau)\dot{\mathbf{W}}(t))=\mathbf{V}\delta(\tau)$, $\mathbf{V}=(v_{ij})_{n\times n}$ is a symmetric, non-negative defined constant matrix, and $\mathbf{u}=(u_1,u_2,\ldots,u_n)^{\mathrm{T}}$ is an Ornstein-Uhlenbeck vector process, which is in fact a zero-mean stationary Gaussian diffusion process. The matrix $\mathbf{A}$ is assumed to have a complete set of eigenvalues $\alpha_1,\ldots,\alpha_n$ along with the corresponding set of eigenvectors $\mathbf{e}_1,\ldots,\mathbf{e}_n$, which means that $\alpha_i\neq\alpha_j\cdot(i\neq j)$. Furthermore, as in Liberzon and Brockett [15], the following two conditions are assumed in the present paper, i.e.,

(a) each eigenvalue α_i is assumed to possess a negative real part, i.e., $R(\alpha_i)<0$ $(i=1,2,\ldots,n)$.
(b) $(\mathbf{A},\tilde{\mathbf{V}})$ is a controllable pair, i.e., $\mathrm{rank}(\tilde{\mathbf{V}},\mathbf{A}\tilde{\mathbf{V}},\ldots,\mathbf{A}^{n-1}\tilde{\mathbf{V}})=n$, where $\mathbf{V}=\tilde{\mathbf{V}}\tilde{\mathbf{V}}^{\mathrm{T}}$.

In fact the first condition ensures that the equilibrium solution, $\mathbf{u}=\mathbf{0}$, for the relevant deterministic system is Lyapunov asymptotically stable.

The following spherical polar transformation from (x_1,x_2,x_3) to (ρ,θ,ϕ)

$$x_1=R\cos\theta\sin\phi,\quad x_2=R\cos\theta\cos\phi,\quad x_3=R\sin\theta$$
$$\rho=\ln R,\quad \phi(t)=\omega t+\varphi(t)$$
$$\theta\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right],\quad \phi,\ \varphi\in[0,2\pi]$$

yields a set of equations of the arguments of ρ, θ, ϕ and the noise process $\mathbf{u}$, i.e.,

$$\dot{\rho}=\rho_\varepsilon,\quad \dot{\theta}=\theta_\varepsilon,\quad \dot{\phi}=\phi_\varepsilon \tag{5}$$

where

$$\begin{aligned}\rho_\varepsilon&=\varepsilon^2\rho_2+\varepsilon f(\mathbf{u})\rho_1\\ \theta_\varepsilon&=\varepsilon^2\theta_2+\varepsilon f(\mathbf{u})\theta_1\\ \phi_\varepsilon&=\omega+\varepsilon f(\mathbf{u})\phi_1\end{aligned} \tag{6}$$

and

$$\begin{aligned}\rho_2&=-\delta_1\cos^2\theta-\delta_2\sin^2\theta\\ \rho_1&=\tfrac{1}{2}(f_{r2}+f_{z1})\sin 2\theta+f_{r1}\cos^2\theta+f_{z2}\sin^2\theta\\ \theta_2&=\tfrac{1}{2}(\delta_1-\delta_2)\sin 2\theta\\ \theta_1&=\tfrac{1}{2}(f_{z2}-f_{r1})\sin 2\theta+(f_{z1}\cos^2\theta-f_{r2}\sin^2\theta)\\ \phi_1&=f_{\phi 1}+\tan\theta f_{\phi 2}\end{aligned} \tag{7}$$

$$f_{r1}=\tfrac{1}{2}[k_1+k_2\cos 2\phi+k_3\sin 2\phi],\quad f_{r2}=b_{13}\sin\phi+b_{23}\cos\phi$$
$$f_{\phi 1}=\tfrac{1}{2}[k_4+k_3\cos 2\phi-k_2\sin 2\phi],\quad f_{\phi 2}=b_{13}\cos\phi-b_{23}\sin\phi$$
$$f_{z1}=b_{31}\sin\phi+b_{32}\cos\phi,\quad f_{z2}=b_{33}$$
$$k_1=b_{22}+b_{11},\quad k_2=b_{22}-b_{11},\quad k_3=b_{12}+b_{21},\quad k_4=b_{12}-b_{21}.$$

Since the phase processes θ and ϕ are independent of the variable ρ, these together with the diffusion process, $\mathbf{u}(t)$, which is defined in Eq. (4), form a vector diffusion process $(\theta(t),\phi(t),\mathbf{u}(t))$ on $[-\pi/2,\pi/2]\times[0,2\pi]\times R^n$ of dimension $(n+2)$ with the following generator

$$\begin{aligned}L_\varepsilon^*&=L_0^*+\varepsilon L_1^*+\varepsilon^2 L_2^*\\ L_0^*&=L_{\mathrm{u}}^*+\omega\frac{\partial}{\partial\phi},\quad L_1^*=f(\mathbf{u})\phi_1\frac{\partial}{\partial\phi}+f(\mathbf{u})\theta_1\frac{\partial}{\partial\theta},\quad L_2^*=\theta_2\frac{\partial}{\partial\theta}\end{aligned} \tag{8}$$

and the adjoint operator.

$$\begin{aligned}L_\varepsilon&=L_0+\varepsilon L_1+\varepsilon^2 L_2\\ L_0&=-\omega\frac{\partial}{\partial\phi}+L_{\mathrm{u}},\quad L_1=-f(\mathbf{u})\frac{\partial}{\partial\theta}\theta_1-f(\mathbf{u})\frac{\partial}{\partial\phi}\phi_1,\quad L_2=-\frac{\partial}{\partial\theta}\theta_2\end{aligned} \tag{9}$$

where L_{u}^* and L_{u} are, respectively, the differential generator (backward Kolmogorov operator) and the Fokker-Planck operator for the diffusion process $\mathbf{u}(t)$, which are defined in Eq. (10), respectively.

3 The Spectral Analysis for a Linear Filter System

In this section, we mainly recall the existing results for the spectral analysis of an n-dimensional linear filter system.

For System (4), the stationary probability density function of $\mathbf{u}(t)$ is

$$p_s(\mathbf{u})=N\exp[-\tfrac{1}{2}\mathbf{u}^T\mathbf{K}_{\mathrm{u}}^{-1}\mathbf{u}],\quad N=(2\pi)^{-n/2}[\det\mathbf{K}_{\mathrm{u}}]^{1/2}$$

where N is the normalization constant, and $\mathbf{K}_{\mathrm{u}}=\langle\mathbf{u}(t)(\mathbf{u}(t))^T\rangle$ is the covariance matrix, which is the solution of the steady-state variance equation

$$\mathbf{A}\mathbf{K}_u+\mathbf{K}_u\mathbf{A}^T+\mathbf{V}=0.$$

$\mathbf{U}=(\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n)$ is assumed to be the relevant eigenmatrix of $\mathbf{A}$, which leads to $\mathbf{D}=\mathbf{U}^{-1}\mathbf{A}\mathbf{U}=\mathrm{diag}[\alpha_1,\alpha_2,\ldots,\alpha_n]$.

In the present formulation, for the *n*-dimensional Ornstein-Uhlenbeck process, the assumption of the detailed balance condition is removed.

For the diffusion process $\mathbf{u}(t)$, the differential generator (backward Kolmogorov operator) L_u^* and its adjoint, the Fokker-Planck operator, L_u are, respectively, given by

$$L_u^*=a_{ij}u_j\frac{\partial}{\partial u_i}+\frac{1}{2}v_{ij}\frac{\partial^2}{\partial u_i\partial u_j},\quad L_u=-\frac{\partial}{\partial u_i}[a_{ij}u_j]+\frac{1}{2}v_{ij}\frac{\partial^2}{\partial u_i\partial u_j} \tag{10}$$

where the repeated indices indicate the usual summation. Eigenvalue problems corresponding to the two operators arise in the form

$$L_u^*\psi_{\lambda'}^*(\mathbf{u})=\lambda'\psi_{\lambda'}^*(\mathbf{u}),\quad L_u\psi_\lambda(\mathbf{u})=\lambda\psi_\lambda(\mathbf{u}). \tag{11}$$

It can be verified that the spectrum of the operators L_u and L_u^* are discrete and they possess the same set of eigenvalues.

For the case of a one-dimensional Ornstein-Uhlenbeck process, the results of the spectral representation of the transition probability density function have been presented in [14] (p. 333). For the system considered in that reference, the drift and diffusion parameters were selected as $\mu(x)=-x$, $\sigma^2(x)=1$, i.e., $a_{11}=-1$, $v_{11}=1$. The differential generator and the relevant backward Kolmogonov equation can be written as

$$L^*=-x\frac{\partial}{\partial x}+\frac{1}{2}\frac{\partial^2}{\partial x^2},\quad \begin{cases}\dfrac{\partial}{\partial\tau}p(y,\tau|x)=\left[-x\dfrac{\partial}{\partial x}+\dfrac{1}{2}\dfrac{\partial^2}{\partial x^2}\right]p(y,\tau|x), & \tau\geqslant 0,\quad x,y\in(-\infty,+\infty)\\ p(y,0|x)=\delta(x-y).\end{cases} \tag{12}$$

Solving Eq. (12) by the direct integration, we obtain

$$p(y,\tau|x)=\frac{e^{-y^2}}{\sqrt{\pi}\sqrt{1-e^{-2\tau}}}\exp\left[-\frac{(x^2+y^2)e^{-2\tau}}{1-e^{-2\tau}}\right]\exp\left[\frac{2xye^{-\tau}}{1-e^{-2\tau}}\right]. \tag{13}$$

The associated eigenfunctions for the backward Kolmogonov operator satisfy

$$\frac{1}{2}\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2}-x\frac{\mathrm{d}\varphi}{\mathrm{d}x}=-\lambda\varphi(x),\quad x\in(-\infty,\infty)$$

and are subject to the integrability conditions $\int_{-\infty}^{\infty}e^{-x^2}[\varphi(x)]^2\mathrm{d}x<\infty$. It is easy for us to check that the solutions to Eq. (12) can be selected as the classical Hermite polynomials

$$H_n(x)=(-1)^n e^{x^2}\frac{\mathrm{d}}{\mathrm{d}x^n}(e^{-x^2})$$

with the relevant eigenvalues $\lambda_n=-n$, respectively. We can subsequently obtain the spectral representation for the transition probability density function as

$$p(y,\tau|x)=e^{-y^2}\sum_{n=0}^{\infty}e^{-n\tau}H_n(x)H_n(y)\pi_n \tag{14}$$

where $\pi_n^{-1}=\lfloor\sqrt{\pi}2^n n!\rfloor$. Based on that presented in [14] (p. 333), by virtue of a classical formula for the sum of the series, we find that Eq. (14) reduces to Eq. (13).

For the *n*-dimensional Ornstein-Uhlenbeck process, according to [13], the solutions to the associated eigenvalue problem of the backward Kolmogorov operator L_u^* possess two characteristics, namely

(i) Each of the eigenvalues can be expressed as $\lambda_{\mathbf{m}}=m_1\alpha_1+\cdots+m_n\alpha_n$, where $\mathbf{m}=(m_1,m_2,\ldots m_n)$; $m=m_1+m_2+\ldots+m_n$, $m_1\,(i=1,2,\ldots,n)$ are non-negative integers.

(ii) The corresponding eigenfunction is found to be an element of the set of multivariate Hermite polynomials, i.e.,

$$\psi_{\mathbf{m}}^*(\mathbf{u})=G_{\mathbf{m}}(\mathbf{v})=(-1)^m\exp\left[\frac{1}{2}\mathbf{v}^T\mathbf{C}\mathbf{v}\right]\frac{\partial^m}{\partial w_1^{m_1}\partial w_2^{m_2}\ldots\partial w_n^{m_n}}\times\exp\left[-\frac{1}{2}\mathbf{v}^T\mathbf{C}\mathbf{v}\right]$$

where

$$\mathbf{v}=\mathbf{U}^{-1}\mathbf{u},\quad \mathbf{C}=\mathbf{U}^T\mathbf{K}_u^{-1}\mathbf{U}=\mathbf{K}_v^{-1},\quad \mathbf{w}=\mathbf{U}^T\mathbf{K}_u^{-1}\mathbf{u}.$$

It has also been shown in [13] that if system (4) satisfies the condition of detailed balance, i.e.,

$$p(\mathbf{u}',\tau|\mathbf{u},0)p_s(\mathbf{u})=p(\boldsymbol{\varepsilon}\mathbf{u},\tau|\boldsymbol{\varepsilon}\mathbf{u}',0)p_s(\mathbf{u}'),\quad p_s(\mathbf{u})=p_s(\boldsymbol{\varepsilon}\mathbf{u})$$

where $\boldsymbol{\varepsilon}=\mathrm{diag}(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)$, $\varepsilon_i=1(\varepsilon_i=-1)$ for an even (odd) variable u_1, corresponding to the same eigenvalue $\lambda_{\mathbf{m}}$, $\psi_{\mathbf{m}}(\mathbf{u})$ is the eigenfunction of L_u, which can be expressed as

$$\psi_{\mathbf{m}}(\varepsilon\mathbf{u})=\psi_0(\mathbf{u})\psi_{\mathbf{m}}^*(\mathbf{u})=(-1)^m\frac{\partial^m}{\partial w_1^{m_1}\ldots\partial w_n^{m_n}}\psi_0(\mathbf{u}) \tag{15}$$

$$\psi_0(\mathbf{u})=\psi_0(\varepsilon\mathbf{u})=N\exp[-\tfrac{1}{2}\mathbf{u}^T\mathbf{K}_u^{-1}\mathbf{u}]=N\exp[-\tfrac{1}{2}\mathbf{v}^T\mathbf{K}_v^{-1}\mathbf{v}]$$

$$=N\exp[-\tfrac{1}{2}\mathbf{w}^T\mathbf{K}_v\mathbf{w}].$$

In fact, $\psi_{\mathbf{m}}(\mathbf{u})$ can also be expressed as

$$\psi_{\mathbf{m}}(\mathbf{u})=\psi_0(\mathbf{u})\prod_{i=1}^{n}(\mathbf{u}_i^{-1}\mathbf{u})^{m_i} \tag{16}$$

where $\mathbf{u}_i^{-1}$ is the *i*th row vector of $\mathbf{U}^{-1}$, which is the inverse of the matrix $\mathbf{U}$.

In this work, the condition of detailed balance is removed and Eq. (15) and Eq. (16) are thus not tenable. Therefore additional results for the solutions to the second equation in Eq. (13) are required here.

From [15], we know that under the conditions of (a) and (b) defined earlier, corresponding to each eigenvalue $\lambda_{\mathbf{m}}$, the associated eigenfunction $\psi_{\mathbf{m}}(\mathbf{u})$ of the Fokker Planck operator, L_u can be expressed as

$$\psi_{\mathbf{m}}(\mathbf{u})=\psi_0(\mathbf{u})\prod_{i=1}^{n}(\mathbf{u}_i^T\mathbf{u})^{m_i}$$

where $\mathbf{u}_i^T$ is the ith row vector of $\mathbf{U}^T$, which is the transpose of the matrix $\mathbf{U}$. With this result, it can easily be checked that if the system matrix $\mathbf{A}$ is a real and symmetrical, then without the condition of detailed balance, Eq. (16) is also tenable.

As in Caughey [17], we introduce a new operator $\tilde{L}_{\mathrm{u}}$

$$\tilde{L}_{\mathrm{u}}\tilde{\psi}_{\mathbf{m}}(\mathbf{u})=[\psi_0(\mathbf{u})]^{-1}L_{\mathrm{u}}(\psi_0(\mathbf{u})\psi_{\mathbf{m}}(\mathbf{u}))$$

$$\tilde{\psi}_{\mathbf{m}}(\mathbf{u})=[\psi_0(\mathbf{u})]^{-1}\psi_{\mathbf{m}}(\mathbf{u})$$

which is the adjoint operator of L_{u}^* with respect to the following inner product

$$\langle\tilde{\psi}_{\mathbf{m}}(\mathbf{u}),\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E=\int_{R^n}\psi_0(\mathbf{u})\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\psi_{\mathbf{m}}^*(\mathbf{u})\mathrm{d}\mathbf{u}=\mathrm{E}(\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\psi_{\mathbf{m}}^*(\mathbf{u}))$$

the expectation of $\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\psi_{\mathbf{m}}^*(\mathbf{u})$, from which, one can easily check that $\tilde{\psi}_{\mathbf{m}}(\mathbf{u})$ and $\psi_{\mathbf{m}}^*(\mathbf{u})$ are bi-orthogonal normal, i.e.,

$$\langle\tilde{\psi}_{\mathbf{m}1}(\mathbf{u}),\psi_{\mathbf{m}2}^*(\mathbf{u})\rangle_E=\delta_{\mathbf{m}1,\mathbf{m}2}=\begin{cases}1, & \mathbf{m}1=\mathbf{m}2\\ 0, & \mathbf{m}1\neq\mathbf{m}2.\end{cases}$$

Then the transition probability density of the process $\mathbf{u}(t)$ can be written as

$$p(\mathbf{u},\tau|\mathbf{u}')=\psi_0(\mathbf{u})\sum_{m_1=0,\cdots,m_n=0}^{\infty}\exp[\lambda_{\mathbf{m}}\tau]\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\psi_{\mathbf{m}}^*(\mathbf{u}'),\quad\tau\geqslant 0. \tag{17}$$

Via Eq. (17), we obtain the covariance matrix $\mathbf{R}_{\mathrm{u}}(\tau)$

$$\mathbf{R}_{\mathbf{u}}(\tau)=\int_{R^n}\mathrm{d}\mathbf{u}\int_{R^n}\mathrm{d}\mathbf{u}'[\mathbf{u}^T\mathbf{u}'p(\mathbf{u},\tau|\mathbf{u}')p_s(\mathbf{u}')]$$

$$=\sum_{m_1=0,\cdots,m_n=0}^{\infty}[\langle\mathbf{u},\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\rangle_E]^T\langle\mathbf{u},\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\exp[\lambda_{\mathbf{m}}\tau]$$

from which we obtain the spectral density function matrices, i.e.,

$$\mathbf{S}_{\mathrm{u}}(\omega)=2\int_0^{\infty}\mathbf{R}_{\mathrm{u}}(\tau)\cos(\omega\tau)\mathrm{d}\tau$$

$$=-\sum_{m_1=0,\cdots,m_n=0}^{\infty}[\langle\mathbf{u},\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\rangle_E]^T\langle\mathbf{u},\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\frac{2\lambda_{\mathbf{m}}}{\lambda_{\mathbf{m}}^2+\omega^2}$$

$$\mathbf{\Phi}_{\mathrm{u}}(\omega)=2\int_0^{\infty}\mathbf{R}_{\mathrm{u}}(\tau)\sin(\omega\tau)\mathrm{d}\tau$$

$$=-\sum_{m_1=0,\cdots,m_n=0}^{\infty}[\langle\mathbf{u},\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\rangle_E]^T\langle\mathbf{u},\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\frac{2\omega}{\lambda_{\mathbf{m}}^2+\omega^2}.$$

For a scalar stochastic function $f(\mathbf{u})$, which is integrable function of $\mathbf{u}$ in the sense that $\int_{R^n}[f(\mathbf{u})]^2\psi_0(\mathbf{u})\mathrm{d}\mathbf{u}<+\infty$ and

$$E[f(\mathbf{u})]=\int_{R^n}f(\mathbf{u})\psi_0(\mathbf{u})\mathrm{d}\mathbf{u}=0$$

the covariance and the spectral density function for $f(\mathbf{u})$ are obtained as

$$R_f(\tau)=\int_{R^n}\mathrm{d}\mathbf{u}\int_{R^n}\mathrm{d}\mathbf{u}'[f(\mathbf{u})f(\mathbf{u}')p(\mathbf{u},\tau|\mathbf{u}')p_s(\mathbf{u}')]$$

$$=\sum_{m_1=0,\cdots,m_n=0}^{\infty}\langle f(\mathbf{u}),\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\rangle_E\langle f(\mathbf{u}),\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\exp[\lambda_{\mathbf{m}}\tau]$$

$$S_f(\omega)=2\int_0^{\infty}R_f(\tau)\cos(\omega\tau)\mathrm{d}\tau=-\sum_{m_1=0,\cdots,m_n=0}^{\infty}\langle f(\mathbf{u}),\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\rangle_E$$

$$\times\langle f(\mathbf{u}),\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\frac{2\lambda_{\mathbf{m}}}{\lambda_{\mathbf{m}}^2+\omega^2} \tag{18}$$

$$\Phi_f(\omega)=2\int_0^{\infty}R_f(\tau)\sin(\omega\tau)\mathrm{d}\tau=-\sum_{m_1=0,\cdots,m_n=0}^{\infty}\langle f(\mathbf{u}),\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\rangle_E$$

$$\times\langle f(\mathbf{u}),\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\frac{2\omega}{\lambda_{\mathbf{m}}^2+\omega^2}.$$

4 Asymptotic Analysis

Corresponding to the Fokker Planck operator L_ε, defined in Eq. (9), the invariant probability density function $p_\varepsilon(\theta,\phi,\mathbf{u})$ satisfies the following FPK equation:

$$L_\varepsilon p_\varepsilon=(L_0+\varepsilon L_1+\varepsilon^2L_2)p_\varepsilon(\theta,\phi,\mathbf{u})=0. \tag{19}$$

In the present paper, $\mathbf{u}(t)$ is assumed to be an ergodic Markov process on R^n, and according to the multiplicative ergodic theorem of Oseledec, the top Lyapunov exponent for system (5) is

$$\lambda_\varepsilon=\langle\rho_\varepsilon,p_\varepsilon\rangle=\int_0^{2\pi}d\phi\int_{-\pi/2}^{\pi/2}d\theta\int_{R^n}\mathrm{d}\mathbf{u}[\rho_\varepsilon(\theta,\phi,\mathbf{u})p_\varepsilon(\theta,\phi,\mathbf{u})] \tag{20}$$

For the present work, the assumption $\varepsilon\ll 1$ holds and we do not need the exact solution $p_\varepsilon(\theta,\phi,\mathbf{u})$ of the FPK equation. A formal expansion of

$$p_\varepsilon(\theta,\phi,\mathbf{u})=p_0(\theta,\phi,\mathbf{u})+\varepsilon p_1(\theta,\phi,\mathbf{u})+\cdots+\varepsilon^N p_N(\theta,\phi,\mathbf{u})+\cdots \tag{21}$$

can be constructed such that

$$L_0p_0=0 \tag{22}$$

$$L_0p_1=-L_1p_0 \tag{23}$$

$$L_0p_2=-L_1p_1-L_2p_0,\cdots \tag{24}$$

and hence the top Lyapunov exponent for system (5) may possess an asymptotic expansion as follows:

$$\langle\rho_\varepsilon,p_\varepsilon\rangle=\langle\rho_0,p_0\rangle+\varepsilon[\langle\rho_1,p_0\rangle+\langle\rho_0,p_1\rangle]+\varepsilon^2[\langle\rho_2,p_0\rangle+\langle\rho_1,p_1\rangle+\langle\rho_0,p_2\rangle]+\cdots \tag{25}$$

of which the proof of the validity is required.

To show that Eq. (25) is correct, as in [6], we construct an adjoint expansion

$$L_\varepsilon^*F_\varepsilon=\rho_\varepsilon-(f_0+\varepsilon f_1+\cdots+\varepsilon^Nf_N)+\varepsilon^{N+1}(L_1^*F_N+L_2^*F_{N-1})+\varepsilon^{N+2}(L_2^*F_N) \tag{26}$$

with

$$F_\varepsilon(\theta,\phi,\mathbf{u})=F_0(\theta,\phi,\mathbf{u})+\varepsilon F_1(\theta,\phi,\mathbf{u})+\cdots+\varepsilon^NF_N(\theta,\phi,\mathbf{u})$$

where $f_0, f_1, \ldots, f_N$ are functions which do not depend on the variables θ and ϕ, but only on $\mathbf{u}(t)\in R^n$, which are chosen such that the sequence of the following problems obtained from Eq. (26)

$$L_0^*F_0=\rho_0-f_0,\quad L_0^*F_1=\rho_1-f_1-L_1^*F_0$$

$$L_0^*F_2=\rho_2-f_2-L_2^*F_0-L_1^*F_1$$

$$L_0^*F_3=-f_3-L_2^*F_1-L_1^*F_2$$

$$\cdots$$

$$L_0^*F_N+L_1^*F_{N-1}+L_2^*F_{N-2}=-f_N$$

is solvable.

For a fixed N, we introduce $p_\varepsilon^{(\mathrm{N})}=p_0+\varepsilon p_1+\varepsilon^2 p_2+\ldots+\varepsilon^{\mathrm{N}}p_{\mathrm{N}}$, as the truncated density function of p_ε, and $f_\varepsilon^{(\mathrm{N})}=(f_0+\varepsilon f_1+\ldots+\varepsilon^{\mathrm{N}}f_{\mathrm{N}})$. As $\varepsilon\to 0$, it is easy to verify that the difference between p_ε and $p_\varepsilon^{(\mathrm{N})}$ is of the order of $\varepsilon^{\mathrm{N}+1}$, and is expressed as $\varepsilon^{\mathrm{N}+1}(\delta_{\mathrm{P}})$.

With the foregoing preparation, one can arrive at the following equation:

$$\begin{aligned}\langle\rho_\varepsilon,p_\varepsilon\rangle-\langle\rho_\varepsilon,p_\varepsilon\rangle_N=&-\varepsilon^{N+1}\{\langle L_1^*F_N+L_2^*F_{N-1},p_\varepsilon\rangle-\langle L_1^*F_N\\&+L_2^*F_{N-1},p_\varepsilon^{(N)}\rangle+\langle F_\varepsilon,L_1p_N+L_2p_{N-1}\rangle\\&-\langle f_\varepsilon^{(N)},\delta_p\rangle-\langle\rho_1,p_N\rangle-\langle\rho_2,p_{N-1}\rangle\}\\&-\varepsilon^{N+2}\{\langle L_2^*F_N,p_\varepsilon\rangle+\langle F_\varepsilon,L_2p_N\rangle\\&-\langle L_2^*F_N,p_\varepsilon^{(N)}\rangle-\langle\rho_2,p_N\rangle\}\end{aligned}\tag{27}$$

where

$$\begin{aligned}\langle\rho_\varepsilon,p_\varepsilon\rangle_N=&\langle\rho_0,p_0\rangle+\varepsilon[\langle\rho_1,p_0\rangle+\langle\rho_0,p_1\rangle]+\varepsilon^2[\langle\rho_2,p_0\rangle\\&+\langle\rho_1,p_1\rangle+\langle\rho_0,p_2\rangle]+\cdots+\varepsilon^N[\langle\rho_2,p_{N-2}\rangle\\&+\langle\rho_1,p_{N-1}\rangle+\langle\rho_0,p_N\rangle].\end{aligned}$$

To furnish expression (27), the following relationship is employed

$$L_\varepsilon p_\varepsilon^{(N)}=\varepsilon^{N+1}[L_2p_{N-1}+L_1p_N]+\varepsilon^{N+2}L_2p_N.$$

In addition, in the present work, $\rho_0=0$. According to Theorem 3.1 in Section 3 of [6], suppose $N\geqslant 0$ is fixed, p_0, p_1, p_2, $\ldots$, p_{N} and F_1, F_2, $\ldots$, F_{N} are such that the inner products on the right side of Eq. (27) are well defined, and

$$\sup_{\phi,\mathrm{u}}|L_1^*F_N+L_2^*F_{N-1}|\leqslant C_1<\infty,\quad \sup_{\phi,\mathrm{u}}|L_2^*F_N|\leqslant C_2<\infty.$$

Then the asymptotic expansion (25) for the top Lyapunov exponent of system (5) is tenable.

In Eq. (21), all the functions $p_\varepsilon(\theta,\phi,\mathbf{u})$, $p_0(\theta,\phi,\mathbf{u})$, $\ldots$ are required to be 2π-periodic in variable ϕ, i.e.,

$$\begin{aligned}p_\varepsilon(\theta,\phi,\mathbf{u})&=p_\varepsilon(\theta,\phi+2\pi,\mathbf{u})\\p_0(\theta,\phi,\mathbf{u})&=p_0(\theta,\phi+2\pi,\mathbf{u})\\p_1(\theta,\phi,\mathbf{u})&=p_1(\theta,\phi+2\pi,\mathbf{u}),\cdots\end{aligned}\tag{28}$$

The normalization condition of the probabilistic density function $p_\varepsilon(\theta,\phi,\mathbf{u})$ then yields

$$\begin{aligned}&\int_0^{2\pi}\mathrm{d}\phi\int_{-\pi/2}^{\pi/2}\mathrm{d}\theta\int_{R^n}\mathrm{d}\mathbf{u}[p_0(\theta,\phi,\mathbf{u})]=1\\&\int_0^{2\pi}\mathrm{d}\phi\int_{-\pi/2}^{\pi/2}\mathrm{d}\theta\int_{R^n}\mathrm{d}\mathbf{u}[p_1(\theta,\phi,\mathbf{u})]\\&\quad=\int_0^{2\pi}\mathrm{d}\phi\int_{-\pi/2}^{\pi/2}\mathrm{d}\theta\int_{R^n}\mathrm{d}\mathbf{u}[p_2(\theta,\phi,\mathbf{u})]=0,\ldots.\end{aligned}\tag{29}$$

In general, each equation with the form $L_0p=q$ must satisfy the following solvability condition, i.e.,

$$\langle q,q^*\rangle=0,\quad \forall q^*\in\mathrm{Ker}(L_0^*)=\{q^*|L_0^*q^*=0\}\tag{30}$$

where $\langle\cdot,\cdot\rangle$ refers to the general scalar product, which is defined in Eq. (20), and L_0^* is the adjoint operator of L_0, i.e.,

$$L_0^*=\omega\frac{\partial}{\partial\phi}+L_{\mathrm{u}}^*.$$

Then via the definition of the scalar product $\langle\cdot,\cdot\rangle$, the solvability condition can be expressed as

$$\begin{aligned}\langle q,q^*\rangle&=\int_0^{2\pi}d\phi\int_{-\pi/2}^{\pi/2}d\theta\int_{R^n}d\mathbf{u}[q^*q]\\&=\int_0^{2\pi}d\phi\int_{-\pi/2}^{\pi/2}d\theta\int_{R^n}d\mathbf{u}[pL_0^*q^*]=0,\quad\forall q^*\in\mathrm{Ker}(L_0^*).\end{aligned}\tag{31}$$

To examine in detail each $q^*\in\mathrm{Ker}(L_0^*)$, we expand it as a series in terms of the eigenfunctions $\psi_{\mathbf{m}}^*(\mathbf{u})$ of L_{u}^*, i.e.,

$$q^*(\theta,\phi,\mathbf{u})=\sum_{m_1=0,\cdots,m_n=0}^{\infty}q_{\mathbf{m}}^*(\theta,\phi)\psi_{\mathbf{m}}^*(\mathbf{u}).\tag{32}$$

From Eq. (31), we know that each coefficient satisfies the following equation:

$$\left[\omega\frac{\partial}{\partial\phi}+\lambda_{\mathbf{m}}\right]q_{\mathbf{m}}^*(\theta,\phi)=0.$$

The condition $\omega>0$ along with the assumption that the real part of each eigenvalue $\lambda_{\mathbf{m}}$ except for $\lambda_0=0$, to which the associated eigenfunction $\psi_0^*(\mathbf{u})=1$, is less than zero leads to the fact that in Eq. (32) there exists only one non-zero periodic coefficient, i.e.,

$$q_{\mathbf{m}}^*(\theta,\phi)=\begin{cases}q_0^*(\theta), & m=0\\0, & m\neq 0.\end{cases}$$

From the above, we can conclude that each element in the space $\mathrm{Ker}(L_0^*)$ is an arbitrary integrable function of the variable θ. Hence, for the present problem, the solvability condition (30) reduces to

$$\int_0^{2\pi}d\phi\int_{R^n}d\mathbf{u}q(\theta,\phi,\mathbf{u})=0.\tag{33}$$

5 Expansion of Stationary Probability Density Function

To obtain the perturbation solution (21) to the FPK Eq. (19), a study on the recurrence Eqs. (22)–(24) will be conducted in the subsequent context.

5.1 FPK Equation of Order ε^0. Since the set of the eigenfunctions $\psi_{\mathbf{m}}(\mathbf{u})$ of L_{u} forms a complete set, then for Eq. (22), the solution $p_0(\theta,\phi,\mathbf{u})$ is sought in the form

$$p_0(\theta,\phi,\mathbf{u})=\sum_{m_1=0,\cdots,m_n=0}^{\infty}p_{\mathbf{m}}^{(0)}(\theta,\phi)\psi_{\mathbf{m}}(\mathbf{u})\tag{34}$$

which leads to the fact that each of the coefficients $p_{\mathbf{m}}^{(0)}(\theta,\phi)$ is, respectively, the solution to

$$\left[-\omega\frac{\partial}{\partial\phi}+\lambda_{\mathbf{m}}\right]p_{\mathbf{m}}^{(0)}(\theta,\phi)=0.\tag{35}$$

By solving Eq. (35), we know that there exists only one nonzero periodic solution, $p_0^{(0)}(\theta,\phi)=p_0^{(0)}(\theta)$, which corresponding to the eigenvalue $\lambda_0=0$. Furthermore, via the normalization condition (29), we finally obtain

$$p_0(\theta,\phi,\mathbf{u})=\frac{1}{2\pi}F(\theta)\psi_0(\mathbf{u})\tag{36}$$

where $F(\theta)$ is a function of θ yet to be determined by the solvability condition of Eq. (24).

5.2 FPK Equation of Order ε. Now we consider Eq. (23). Substitution of Eq. (36) into the right side of (23) yields

$$L_0 p_1(\theta,\phi,\mathbf{u}) = \frac{f(\mathbf{u})\psi_0(\mathbf{u})}{2\pi}\left[\theta_1\frac{\partial F(\theta)}{\partial\theta}+\left[\frac{\partial\theta_1}{\partial\theta}+\frac{\partial\phi_1}{\partial\phi}\right]F(\theta)\right]$$

$$= \frac{f(\mathbf{u})\psi_0(\mathbf{u})}{2\pi}[M_0+M_2^{(1)}\cos 2\phi+M_2^{(2)}\sin 2\phi + M_1^{(1)}\cos\phi+M_1^{(2)}\sin\phi] \quad (37)$$

where

$$M_0=\tfrac{1}{4}(2b_{33}-k_1)\Lambda_0, \quad \Lambda_0=\frac{d}{d\theta}[F(\theta)\sin 2\theta]$$

$$M_2^{(1)}=-k_2\Lambda_2, \quad M_2^{(2)}=-k_3\Lambda_2, \quad \Lambda_2=F(\theta)+\tfrac{1}{4}\Lambda_0$$

$$M_1^{(1)}=-b_{23}\Lambda_{12}+b_{32}\Lambda_{11}, \quad M_1^{(2)}=-b_{13}\Lambda_{12}+b_{31}\Lambda_{11}$$

$$\Lambda_{11}=\frac{d}{d\theta}[F(\theta)\cos^2\theta], \quad \Lambda_{12}=\frac{d}{d\theta}[F(\theta)\sin^2\theta]+F(\theta)\tan\theta.$$

In order for the problem to be tractable, the function $f(\mathbf{u})\psi_0(\mathbf{u})$, which arises in Eq. (37), is expressed as a series expansion along the eigenfunctions $\psi_{\mathrm{m}}(\mathbf{u})$, i.e.,

$$f(\mathbf{u})\psi_0(\mathbf{u}) = \sum_{m_1=0,m_2=0,\ldots,m_n=0}^{\infty}\langle f(\mathbf{u}),\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\psi_{\mathbf{m}}(\mathbf{u}) \quad (38)$$

Similarly for $p_1(\theta,\phi,\mathbf{u})$

$$p_1(\theta,\phi,\mathbf{u}) = \sum_{m_1=0,m_2=0,\ldots,m_n=0}^{\infty} p_{\mathbf{m}}^{(1)}(\theta,\phi)\psi_{\mathbf{m}}(\mathbf{u}). \quad (39)$$

Substitution of Eqs. (38) and (39) into Eq. (37) leads to the fact that each of the coefficients $p_{\mathbf{m}}^{(1)}(\theta,\phi)$ is, respectively, governed by

$$\left[-\omega\frac{\partial}{\partial\phi}+\lambda_{\mathbf{m}}\right]p_{\mathbf{m}}^{(1)}(\theta,\phi) = \frac{\langle f(\mathbf{u}),\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E}{2\pi}\{M_0+M_2^{(1)}\cos 2\phi + M_2^{(2)}\sin 2\phi+M_1^{(1)}\cos\phi + M_1^{(2)}\sin\phi\}$$

to which the solution is

$$p_{\mathbf{m}}^{(1)}(\theta,\phi) = \begin{cases} \dfrac{1}{2\pi}p_0^{(1)}(\theta), & m=0 \\ \dfrac{1}{2\pi}\langle f(\mathbf{u}),\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\{\Pi_{\mathbf{m}}^{(0)}\Lambda_0+\Pi_{\mathbf{m}}^{(2)}\Lambda_2+\Pi_{\mathbf{m}}^{(11)}\Lambda_{11}+\Pi_{\mathbf{m}}^{(12)}\Lambda_{12}\}, & m\neq 0 \end{cases} \quad (40)$$

where

$$\Pi_{\mathbf{m}}^{(0)}=\frac{1}{\alpha_{\mathbf{m}}}\frac{1}{4}(2b_{33}-k_1)$$

$$\Pi_{\mathbf{m}}^{(11)}=\frac{\{[b_{32}\lambda_{\mathbf{m}}+b_{31}\omega]\cos\phi+[b_{31}\lambda_{\mathbf{m}}-b_{32}\omega]\sin\phi\}}{\omega^2+\alpha_{\mathbf{m}}^2}$$

$$\Pi_{\mathbf{m}}^{(12)}=-\frac{\{[b_{23}\lambda_{\mathbf{m}}+b_{13}\omega]\cos\phi+[b_{13}\lambda_{\mathbf{m}}-b_{23}\omega]\sin\phi\}}{\omega^2+\alpha_{\mathbf{m}}^2}$$

$$\Pi_{\mathbf{m}}^{(2)}=-\frac{\{[k_2\lambda_{\mathbf{m}}+2k_3\omega]\cos 2\phi+[k_3\lambda_{\mathbf{m}}-2k_2\omega]\sin 2\phi\}}{4\omega^2+\alpha_{\mathbf{m}}^2}.$$

Finally, by synthesizing the foregoing results, we find that $p_1(\theta,\phi,\mathbf{u})$ takes on the following expression:

$$p_1(\theta,\phi,\mathbf{u}) = \frac{1}{2\pi}p_0^{(1)}(\theta)\psi_0(\mathbf{u}) + \sum_{\substack{m_1=0,m_2=0,\ldots,m_n=0\\ m\neq 1}}^{\infty} p_{\mathbf{m}}^{(1)} \times(\theta,\phi)\psi_{\mathbf{m}}(\mathbf{u}) \quad (41)$$

where $p_0^{(1)}(\theta)$ is a function to be determined by the solvability condition of the recurrence equation of order ε^3, which satisfies the normalization condition, i.e., $\int_0^{2\pi}p_0^{(1)}(\theta)d\theta=0$. Therefore by evaluating the asymptotic expansion for the maximal Lyapunov exponent, we find that $p_0^{(1)}(\theta)$ has no contribution to the maximal Lyapunov exponent. In addition, each $p_{\mathbf{m}}^{(1)}(\theta,\phi)(m\neq 0)$ contains the function $F(\theta)$, which are to be determined by the solvability condition of Eq. (24).

5.3 Solvability Condition and FPK Equation. To determine $F(\theta)$ in Eqs. (36) and (41), the solvability condition of Eq. (24) will be investigated. Since the solvability condition of Eq. (24) is

$$-\int_0^{2\pi}d\phi\int_{R^n}\mathrm{d}\mathbf{u}[L_1p_1+L_2p_0]=0. \quad (42)$$

Substitution of Eqs. (36) and (41) into Eq. (42) yields

$$-\int_{R^n}\mathrm{d}\mathbf{u}\int_0^{2\pi}L_1p_1\mathrm{d}\phi = \sum_{m_1,m_2,\ldots,m_n=0}^{\infty} I_3^{(\mathbf{m})}I_4^{(\mathbf{m})} = I_1I_2 + \sum_{\substack{m_1,m_2,\ldots,m_n=0\\ m\neq 0}}^{\infty} I_3^{(\mathbf{m})}I_4^{(\mathbf{m})}$$

where

$$I_1=\int_{R^n}\mathrm{d}\mathbf{u}[f(\mathbf{u})\psi_0(\mathbf{u})]=\int_{R^n}\mathrm{d}\mathbf{u}[f(\mathbf{u})p_s(\mathbf{u})]=0$$

$$I_3^{(\mathbf{m})}=\int_{R^n}\mathrm{d}\mathbf{u}[f(\mathbf{u})\psi_{\mathbf{m}}(\mathbf{u})]=\langle f(\mathbf{u}),\tilde{\psi}_{\mathbf{m}}(\mathbf{u})\rangle_E,$$

$$I_4^{(\mathbf{m})}=\frac{\partial}{\partial\theta}\left[\int_0^{2\pi}\theta_1p_{\mathbf{m}}^{(1)}\mathrm{d}\phi\right].$$

Then via computation, we know

$$
\begin{aligned}
&-\int_0^{2\pi} \mathrm{d}\phi \int_{R^n} L_1 p_1 \mathrm{d}\mathbf{u} \\
&\quad = -\frac{1}{8}\beta_1\left[\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}[\sin^2(2\theta)F(\theta)] - \frac{\mathrm{d}}{\mathrm{d}\theta}[\sin(4\theta)F(\theta)]\right] \\
&\qquad -\frac{1}{16}\beta_2\left[\frac{\mathrm{d}}{\mathrm{d}\theta}[\sin(2\theta)F(\theta)] + \frac{1}{4}\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}[\sin^2(2\theta)F(\theta)]\right. \\
&\qquad \left. -\frac{1}{4}\frac{\mathrm{d}}{\mathrm{d}\theta}[\sin(4\theta)F(\theta)]\right] - \frac{3}{8}\kappa_0\frac{\mathrm{d}}{\mathrm{d}\theta}[\sin(2\theta)F(\theta)] \\
&\qquad -\frac{1}{16}\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}[(\kappa_1\cos^2(2\theta) - 2\kappa_2\cos(2\theta) + \kappa_3)F(\theta)] \\
&\qquad -\frac{1}{16}\frac{\mathrm{d}}{\mathrm{d}\theta}[(\kappa_1\sin(4\theta) - 2\kappa_4\sin(2\theta) + 4\kappa_5\tan(\theta))F(\theta)]
\end{aligned} \tag{43}
$$

where

$$
\begin{gathered}
\beta_1 = S_f(0)\left[b_{33} - \frac{1}{2}k_1\right]^2 \\
\beta_2 = S_f(2\omega)[k_2^2 + k_3^2] \\
\kappa_0 = [-b_{13}b_{32} + b_{23}b_{31}]\Phi_f(\omega) \\
\kappa_1 = [(b_{13} + b_{31})^2 + (b_{23} + b_{32})^2]S_f(\omega) \\
\kappa_2 = [(b_{13}^2 - b_{31}^2) + (b_{23}^2 - b_{32}^2)]S_f(\omega) \\
\kappa_3 = [(b_{13} - b_{31})^2 + (b_{23} - b_{32})^2]S_f(\omega) \\
\kappa_4 = [(b_{13} + b_{31})(2b_{13} - b_{31}) + (b_{23} + b_{32})(2b_{23} - b_{32})]S_f(\omega) \\
\kappa_5 = [b_{13}^2 + b_{23}^2]S_f(\omega).
\end{gathered}
$$

Evaluation of the second term of the left side of Eq. (38) leads to

$$
\begin{aligned}
-\int_0^{2\pi} \mathrm{d}\phi \int_{R^n} \mathrm{d}\mathbf{u}[L_2 p_0] &= -\int_0^{2\pi} \mathrm{d}\phi \int_{R^n} \mathrm{d}\mathbf{u}\left[\frac{\psi_0(\mathbf{u})}{2\pi}\frac{\partial[\theta_2 F(\theta)]}{\partial\theta}\right] \\
&= \frac{\delta_1 - \delta_2}{2}\frac{\mathrm{d}[F(\theta)\sin 2\theta]}{d\theta}.
\end{aligned} \tag{44}
$$

Finally, by synthesizing the results of Eqs. (43) and (44), we find that the solvability condition (38) is equivalent to the following standard FPK equation:

$$
\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}[\sigma^2(\theta)F(\theta)] - \frac{\mathrm{d}}{\mathrm{d}\theta}[\mu(\theta)F(\theta)] = 0, \quad \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \tag{45}
$$

in which the relevant diffusion coefficient and drift coefficient are, respectively,

$$
\begin{aligned}
\sigma^2(\theta) &= \left[4\beta_1 + \frac{1}{2}\beta_2\right]\sin^2 2\theta + 2\kappa_1\cos^2 2\theta - 4\kappa_2\cos 2\theta + 2\kappa_3 \\
\mu(\theta) &= \left\{\left[2\beta_1 + \frac{1}{4}\beta_2\right] - \kappa_1\right\}\sin 4\theta + \{8(\delta_1 - \delta_2) + 2\kappa_4 - \beta_2 \\
&\quad - 6\kappa_0\}\sin 2\theta - 4\kappa_5\tan\theta.
\end{aligned} \tag{46}
$$

It should be noted that in Eq. (46), the parameter κ_5 is positive.

In view of the FPK Eq. (45), the process of $\theta(t)$ can be treated as a diffusion process on interval $[-\pi/2, \pi/2]$ with the relevant drift parameter $\mu(\theta)$ and the diffusion parameter $\sigma^2(\theta)$, respectively. In order to determine the solution to Eq. (45), the diffusion behaviors of such a process at the boundaries of $\theta = \pm\pi/2$ and other singular points within $[-\pi/2, \pi/2]$ should be investigated a priori.

The details of the definition and classification of singular points for one-dimensional diffusion processes can be found in Lin and Cai [18], from which we know that the first kind of singular points is the one at which $\sigma^2(\theta)$ vanishes, and the second kind is that at which $\mu(\theta)$ goes to infinity. With these definitions, we can conclude that on the interval $[-\pi/2, \pi/2]$, $\theta = \pm\pi/2$ are the singular points of the second kind.

For the singular boundary x_s (x_l and x_r represent the left and right boundaries, respectively) of the second kind, the diffusion exponent α_s, the drift exponent β_s and the character number c_s are introduced

$$
\begin{gathered}
\sigma^2(x) = O(|x - x_s|^{-\alpha_s}), \quad \alpha_s \geqslant 0, \quad x \to x_s \\
\mu(x) = O(|x - x_s|^{-\beta_s}), \quad \beta_s \geqslant 0, \quad x \to x_s
\end{gathered} \tag{47}
$$

$$
c_l = \lim_{x \to x_l^+} \frac{2\mu(x)[x - x_l]^{\beta_l - \alpha_l}}{\sigma^2(x)}, \quad c_r = -\lim_{x \to x_r^-} \frac{2\mu(x)[x_r - x]^{\beta_r - \alpha_r}}{\sigma^2(x)}.
$$

On $[-\pi/2, \pi/2]$, we consider only the situation of $\theta = \pi/2$, for the case of $\theta = -\pi/2$, the result is similar, while for $\theta \to \pi/2$, we obtain

$$
\tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{\sin\theta}{\sin\left[\frac{\pi}{2} - \theta\right]} \propto \frac{1}{\left[\frac{\pi}{2} - \theta\right]}.
$$

Equation (47) then leads to

$$
\alpha_r = 0, \quad \beta_r = 1, \quad c_r = 1 \tag{48}
$$

which are, respectively, the diffusion, drift exponents, and character value at $\theta = \pi/2$. After comparing these results with the terms in Table 4.53 in [18] (p. 135), which gives a detailed classification of the singular boundaries of the second kind, we know that $\pi/2$ is an entrance of $[-\pi/2, \pi/2]$ and the result is the same as that for $\theta = -\pi/2$, i.e., $\theta = -\pi/2$ is another entrance.

For the diffusion process θ, its scale and speed densities are defined, respectively, as ([14] (pp. 191–204))

$$
s(\theta) = \exp[-E(\theta)], \quad m(\theta) = \frac{1}{\sigma^2(\theta)s(\theta)}
$$

$$
E(\theta) = \int [2\mu(\theta)\sigma^{-2}(\theta)]\mathrm{d}\theta
$$

and in addition, the relevant scale and speed measures are

$$
S(\theta) = \int^{\theta} s(x)\mathrm{d}x, \quad M(\theta) = \int^{\theta} m(x)\mathrm{d}x. \tag{49}
$$

For Eq. (45), its solution can be represented as

$$
F(\theta) = m(\theta)[C_1 S(\theta) + C], \quad \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \tag{50}
$$

where C_1 and C are constants, which will be determined by the normality and boundary conditions at $\theta = \pm\pi/2$.

For Eq. (45), since the two boundaries $\theta = \pm\pi/2$ are both entrances, it is well known that $C_1 = 0$ ([14] (p. 241)).

Via direct integration

$$
E(\theta) = -\log\cos\theta + \tfrac{3}{4}\log F_2(\theta) + \tfrac{1}{2}c_1 \operatorname{arctanh}[F_1(\theta)]
$$

where

$$
F_1(\theta) = \frac{(\kappa_1 - \beta)\cos 2\theta - \kappa_2}{\sqrt{\beta^2 - \beta(\kappa_1 - \kappa_3) - 4\kappa_6^2}}
$$

$$
F_2(\theta) = (\kappa_1 - \beta)\cos^2 2\theta - 2\kappa_2\cos 2\theta + \beta + \kappa_3
$$

$$c_1=\frac{-8\Delta+(\beta+\beta_2)+(6\kappa_0-\kappa_1+\kappa_3-4\kappa_5)}{\sqrt{\beta^2-\beta(\kappa_1-\kappa_3)-4\kappa_6^2}}$$

$$\beta=2\beta_1+\tfrac{1}{4}\beta_2,\quad \kappa_6=[b_{13}b_{32}-b_{23}b_{31}]S_f(\omega),\quad \Delta=\delta_1-\delta_2.$$

Then the solution to Eq. (45) is

$$F(\theta)=\frac{C}{2\cos(\theta)}\left[\frac{1-F_1(\theta)}{1+F_1(\theta)}\right]^{1/4c_1}[F_2(\theta)]^{-1/4} \tag{51}$$

where C is a constant which can be determined by the normalization condition of $F(\theta)$, i.e., $\int_{-\pi/2}^{\pi/2}F(\theta)d\theta=1$.

Next, a rather simple case, $b_{31}=b_{13}$, $b_{23}=b_{32}$, which leads to

$$\begin{gathered}\kappa_0=\kappa_2=\kappa_3=0\\ \kappa_1=4\kappa_5,\quad \kappa_4=2\kappa_5\end{gathered} \tag{52}$$

is considered. Then Eq. (46) is reduced to

$$\begin{gathered}\sigma^2(\theta)=A_1\sin^2 2\theta+A_2\cos^2 2\theta\\ \mu(\theta)=B_1\sin 4\theta+B_2\sin 2\theta+B_3\tan\theta\end{gathered} \tag{53}$$

where

$$A_1=2\beta,\quad A_2=8\kappa_5$$

$$B_1=\beta-4\kappa_5,\quad B_2=8\Delta+4\kappa_5-\beta_2,\quad B_3=-4\kappa_5.$$

Since the two boundaries $\theta=\pm\pi/2$ are both singular points, via the same evaluation as in Eq. (48), we obtain, for the two boundaries, $\alpha_s=2$, $\beta_s=1$, $c_s=1$, from which we learn that $\theta=\pm\pi/2$ are both entrances.

To obtain the solution to Eq. (45), the two cases of $A_1>A_2$ and $A_1<A_2$, with the exception that $A_1=A_2$, which will be studied in the following section, should be investigated, respectively. For the first case $A_1>A_2$, we know that

$$\begin{gathered}E(\theta)=E^{(1)}+E^{(2)}+E^{(3)}\\ E^{(1)}=\tfrac{1}{2}\log[1-(1-\tau_1)\cos^2 2\theta]\\ E^{(2)}=-\frac{B_2}{A_1}\frac{1}{\sqrt{1-\tau_1}}\operatorname{arctanh}[\sqrt{1-\tau_1}\cos 2\theta]\\ E^{(3)}=\log\cos\theta-\tfrac{1}{4}\log[1-(1-\tau_1)\cos^2 2\theta]\\ -\tfrac{1}{2}\sqrt{1-\tau_1}\operatorname{arctanh}[\sqrt{1-\tau_1}\cos 2\theta]\\ 0<\tau_1=\frac{A_2}{A_1}=\frac{4\kappa_5}{\beta}<1.\end{gathered} \tag{54}$$

For the other case of $A_1<A_2$ $(\tau_1>1)$, the corresponding results are

$$\begin{gathered}E^{(1)}=\tfrac{1}{2}\log[1+(\tau_1-1)\cos^2 2\theta]\\ E^{(2)}=-\frac{B_2}{A_1}\frac{1}{\sqrt{\tau_1-1}}\arctan[\sqrt{\tau_1-1}\cos 2\theta]\\ E^{(3)}=\log\cos\theta-\tfrac{1}{4}\log[1+(\tau_1-1)\cos^2 2\theta]\\ +\tfrac{1}{2}\sqrt{\tau_1-1}\arctan[\sqrt{\tau_1-1}\cos 2\theta].\end{gathered} \tag{55}$$

Synthesizing Eq. (54) and Eq. (55) leads to

$$F(\theta)=C\begin{cases}\dfrac{1}{2\beta}\cos\theta\dfrac{[1-\sqrt{1-\tau_1}\cos 2\theta]^{1/2\eta_1-3/4}}{[1+\sqrt{1-\tau_1}\cos 2\theta]^{1/2\eta_1+3/4}}, & \tau_1<1\\ \dfrac{1}{2\beta}\cos\theta\dfrac{\exp[-\eta_1\arctan[\sqrt{\tau_1-1}\cos 2\theta]]}{[1+(\tau_1-1)\cos^2 2\theta]^{3/4}}, & \tau_1>1\end{cases} \tag{56}$$

where

$$\eta_1=\begin{cases}\dfrac{B_1+B_2}{A_1\sqrt{1-\tau_1}}=\dfrac{8\Delta+\beta-\beta_2}{2\beta\sqrt{1-\tau_1}}, & \tau_1<1\\ \dfrac{B_1+B_2}{A_1\sqrt{\tau_1-1}}=\dfrac{8\Delta+\beta-\beta_2}{2\beta\sqrt{\tau_1-1}}, & \tau_1>1\end{cases}$$

and C can be determined by the normalization condition.

5.4 Three Special Cases. From Eqs. (51) and (56), we know that it is impossible to obtain analytical results for the maximal Lyapunov exponent. Furthermore, the expressions of maximal Lyapunov exponents depend on the forms of the matrix **B**. Therefore in this work, three special cases for the coefficient matrix **B** in Eq. (3) are considered, i.e.,

Case I. $b_{13}=b_{23}=0$, which leads to

$$\mathbf{B}=\begin{bmatrix}b_{11}&b_{12}&0\\ b_{21}&b_{22}&0\\ b_{31}&b_{32}&b_{33}\end{bmatrix},\quad \kappa_0=\kappa_5=\kappa_6=0,\quad \kappa_1=\kappa_2=\kappa_3=-\kappa_4. \tag{57}$$

Case II. $b_{31}=b_{13}$, $b_{23}=b_{32}$ and in addition, $A_1=A_2=A$, which imply that

$$\begin{aligned}4S_f(\omega)[b_{13}^2+b_{23}^2]=2S_f(0)[2b_{33}-b_{11}-b_{22}]^2+S_f(2\omega)[(b_{22}\\ -b_{11})^2+(b_{12}+b_{21})^2].\end{aligned} \tag{58}$$

For the case of white noise excitation, with $b_{33}=b_{22}=b_{11}$, $b_{21}=b_{12}$, the condition expression (58) is equivalent to $b_{12}^2=4[b_{13}^2+b_{23}^2]$.

Case III. $b_{31}=b_{13}$, $b_{23}=b_{32}$, $b_{21}=-b_{12}$, $b_{33}=b_{22}=b_{11}$, from which we deduce that

$$\mathbf{B}=\begin{bmatrix}b_{11}&b_{12}&b_{13}\\ -b_{12}&b_{11}&b_{23}\\ b_{13}&b_{23}&b_{11}\end{bmatrix},\quad \beta_1=\beta_2=0. \tag{59}$$

In the subsequent procedure, for each case, we will investigate the stationary solution to the FPK Eq. (45).

For the first case, the associated diffusion and drift coefficients are, respectively,

$$\begin{gathered}\sigma^2(\theta)=8(\beta\sin^2\theta+\kappa_1\cos^2\theta)\cos^2\theta\\ \mu(\theta)=(\beta-\kappa_1)\sin 4\theta+[8\Delta-2\kappa_1-\beta_2]\sin 2\theta.\end{gathered} \tag{60}$$

In order to obtain the solution to Eq. (45) on the interval $[-\pi/2,\pi/2]$, the diffusion behaviors of the diffusion process θ at the boundaries and the singular points within the interval should be investigated a priori. On $[-\pi/2,\pi/2]$, since at $\theta=\pm\pi/2$ $\sigma^2(\theta)=\mu(\theta)=0$, it is easy for us to ascertain that there are no other singular points except these two boundaries, which are both the singular points of the first kind.

For the singular boundary x_s of the first kind, the diffusion exponent α_s, the drift exponent β_s and the character number c_s are defined as

$$\begin{gathered}\sigma^2(x)=O(|x-x_s|^{\alpha_s}),\quad \alpha_s\geqslant 0,\quad x\to x_s\\ \mu(x)=O(|x-x_s|^{\beta_s}),\quad \beta_s\geqslant 0,\quad x\to x_s\end{gathered} \tag{61}$$

$$c_l=\lim_{x\to x_l^+}\frac{2\mu(x)[x-x_l]^{\alpha_l-\beta_l}}{\sigma^2(x)},\quad c_r=-\lim_{x\to x_r}\frac{2\mu(x)[x_r-x]^{\alpha_r-\beta_r}}{\sigma^2(x)}.$$

Then at the left boundary $\theta=-\pi/2$, the evaluation in Eq. (61) leads to

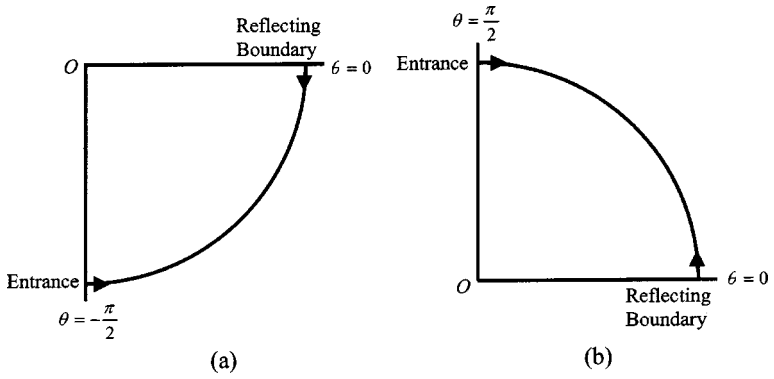


Fig. 1 Boundary diffusion behaviors of the intervals $[-\pi/2,0]$ and $[0,\pi/2]$

$$\alpha_l=2, \quad \beta_l=1, \quad c_l=1+\frac{\beta_2}{2\beta}-\frac{4}{\beta}\Delta$$

and at $\theta=\pi/2$, the result is the same.

According to [18] (p. 132), we know that a stationary solution to a FPK equation does not exist, if each of the two boundaries is either an exit, or attractively natural boundary (ANB), or strictly natural boundary (SNB). From this, we observe that under the condition of $c_s \leqslant \beta_s=1$ ($s=l$ or r), the two boundaries are both SNBs or ANBs, and therefore the invariant measure does not exist. In this work, the solution to Eq. (45) will be investigated only under the condition of $c_s>\beta_s=1$, i.e., $\delta_1-\delta_2<\beta_2/8$, in which case the two boundaries are both repulsively natural boundaries (RNBs) and the nontrivial stationary solution to Eq. (45) does exist. Since the two boundaries are both RNBs, an additional boundary condition is required to determine the solution. For the problem to be tractable, we assume that at the two boundaries, the probability current vanishes ([18] (pp. 169–173)), i.e.,

$$G|_{\theta=-\pi/2}=G|_{\theta=\pi/2}=0$$

$$G=\mu(\theta)F(\theta)-\frac{1}{2}\frac{\partial}{\partial\theta}[\sigma^2(\theta)F(\theta)].$$

Then via direct integration, the solution to Eq. (45) is

$$F(\theta)=\frac{C}{8}[\cos\theta]^{1/2[(-2\beta-8\Delta+\beta_2)/\beta]}[\beta-\beta\cos^2\theta + \kappa_1\cos^2\theta]^{1/4[(-2\beta-8\Delta+\beta_2)/\beta]} \tag{62}$$

where C can be determined by the normalization condition.

For the second case, which is under the condition that $\sigma^2(\theta)=A$, $B_1=0$, and $2B_3=-A$, we can obtain

$$\begin{aligned} E(\theta)&=\frac{2B_2}{A}\sin^2\theta+\log\cos\theta, \quad \theta\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right] \\ s(\theta)&=\sec\theta\exp[-\alpha\sin^2\theta] \\ m(\theta)&=\frac{1}{A}\cos\theta\exp[\alpha\sin^2\theta] \\ \alpha&=\frac{2B_2}{A}=1-\frac{\beta_2}{4\kappa_5}+\frac{2\Delta}{\kappa_5}. \end{aligned} \tag{63}$$

For the problem to be tractable, we divide the interval $[-\pi/2,\pi/2]$ into two subsets as $[-\pi/2,0]$ and $(0,\pi/2]$. Since the solution problem on $[-\pi/2,0]$ is the same as that on $(0,\pi/2]$, we will only investigate the solution problem on $(0,\pi/2]$. On both $[-\pi/2,0]$ and $(0,\pi/2]$, $\theta=0$ is not a singular point. To investigate the diffusion behavior at $\theta=0$, we employ the concepts of the scale and speed densities. As

$$s(\theta)=\sec\theta\exp[-\alpha\sin^2\theta], \quad \theta\in\left(0,\frac{\pi}{2}\right]$$

on $(0,\pi/2]$ and in the neighborhood of $\theta=0$, for the two cases of $\alpha>0$ and $\alpha\leqslant 0$, respectively,

$$e^{-\alpha}\sec\theta\leqslant s(\theta)\leqslant\sec\theta, \quad \alpha>0$$

$$e^{-\alpha}\sec\theta\geqslant s(\theta)\geqslant\sec\theta, \quad \alpha\leqslant 0.$$

Then for the scale measure $S(0,\theta]$, the following two inequalities are tenable:

$$\begin{aligned} -\infty<\int_0^\theta\{[e^{-\alpha}]\sec x\}dx\leqslant S(0,\theta]\leqslant\int_0^\theta\sec x dx<+\infty, \quad \alpha>0 \\ -\infty<\int_0^\theta\sec x dx\leqslant S(0,\theta]\leqslant\int_0^\theta\{[e^{-\alpha}]\sec x\}dx<+\infty, \quad \alpha\leqslant 0. \end{aligned} \tag{64}$$

For the speed measure $M(0,\theta]$, via the definition expression (49), we obtain

$$\begin{aligned} M(0,\theta]&=\int_0^{\sin\theta}\frac{1}{A}\{\exp[\alpha x^2]\}dx=\frac{1}{A}\mathrm{Erfi}[\sin\theta], \quad \alpha>0 \\ M(0,\theta]&=\int_0^{\sin\theta}\frac{1}{A}\{\exp[\alpha x^2]\}dx=\frac{1}{A}\mathrm{Erf}[\sin\theta], \quad \alpha\leqslant 0 \end{aligned} \tag{65}$$

where Erfi and Erf are the error functions.

Equations (64) and (65) tell us that the two measures are both finite. Thus according to the definition of a reflecting boundary ([14] (pp. 226–242)), we know that $\theta=0$ is a reflecting boundary for the two intervals, which is shown in Fig. 1. With this result, we can conclude that the diffusion process evolves on $[-\pi/2,0]$ and $(0,\pi/2]$ separately, and the solution to Eq. (45) will be analyzed on the two intervals, respectively.

For Eq. (45) which is restricted on $(0,\pi/2]$, the solution is

$$F(\theta)=\begin{cases} \dfrac{2\sqrt{\alpha}}{\sqrt{\pi}\mathrm{Erf}\,i[\sqrt{\alpha}]}\cos\theta\exp[\alpha\sin^2\theta], & \alpha>0, \\ \dfrac{2\sqrt{-\alpha}}{\sqrt{\pi}\mathrm{Erf}[\sqrt{-\alpha}]}\cos\theta\exp[\alpha\sin^2\theta], & \alpha\leqslant 0, \end{cases} \quad \theta\in\left(0,\frac{\pi}{2}\right]. \tag{66}$$

We can then easily verify that on $[-\pi/2,0]$, the stationary probability density is of the same expression.

For the third case, assumption (59) leads to

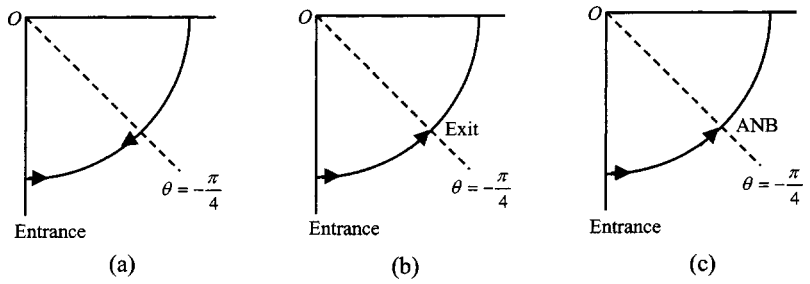


Fig. 2 Boundary diffusion behavior of interval $[-\pi/2,-\pi/4]$ for the cases of (*a*) $\delta_1>\delta_2$, (*b*) $\delta_1<\delta_2$, (*c*) $\delta_1=\delta_2$

$$\sigma^2(\theta)=8\kappa_5\cos^2 2\theta$$
$$\mu(\theta)=-4\kappa_5\sin 4\theta+\{8\Delta+4\kappa_5\}\sin 2\theta-4\kappa_5\tan\theta. \tag{67}$$

Since at $\theta=\pm\pi/4$, $\sigma^2(\theta)=0$, thus $\theta=\pm\pi/4$ are singular points of the first kind, and at points $\theta=\pm\pi/2$, $\mu(\theta)=-\infty$, so $\theta=\pm\pi/2$ are singular points of the second kind. Due to the different diffusion behaviors of the singular points, the interval $[-\pi/2,\pi/2]$ should be divided into three subintervals, i.e., $[-\pi/2,-\pi/4]$, $(-\pi/4,\pi/4]$ and $(\pi/4,\pi/2]$, of which the solutions to Eq. (45) should be investigated, respectively. First we investigate the diffusion behaviors at the singular points.

On $[-\pi/2,-\pi/4]$, according to the definition expressed in Eq. (47), we know that at the left boundary $\theta=-\pi/2$, the diffusion and the drift exponents, and the character value are

$$\alpha_l=0,\quad \beta_r=1,\quad c_r=1$$

which, in view of Table 4.5.3 in [18], leads to the deduction that $\theta=-\pi/2$ is an entrance for interval $[-\pi/2,-\pi/4]$.

Since at $-\pi/4$,

$$\sigma^2(\theta)|_{\theta=-\pi/4}=0,\quad \mu(\theta)|_{\theta=-\pi/4}=\begin{cases}-8(\delta_1-\delta_2)<0, & \delta_1>\delta_2\\ -8(\delta_1-\delta_2)>0, & \delta_1<\delta_2\\ 0, & \delta_1=\delta_2\end{cases}$$

$$\alpha_r=2,\quad \beta_r=\begin{cases}0, & \delta_1\neq\delta_2\\ 1, & \delta_1=\delta_2\end{cases}$$

the diffusion behaviors at such a boundary should be discussed for three cases. According to Table 4.5.2 in [18] (p. 134), which provides the classifications of singular boundaries of the first kind, we know that if $\delta_1>\delta_2$, $-\pi/4$ is an entrance, and if $\delta_1<\delta_2$, $-\pi/4$ is an exit instead. To determine the boundary type for the case of $\delta_1=\delta_2$, the character value, which is defined as

$$c_r=-\lim_{\theta\to-\pi^+/2}\left[-2\mu(\theta)\left[\theta+\frac{\pi}{2}\right]^{1-0}[\sigma^2(\theta)]^{-1}\right]=-\frac{1}{2} \tag{68}$$

is required. By contrasting Eq. (68) with the relevant terms in Table 4.5.2 of [18], we find that $-\pi/4$ is an ANB. These results are shown in Fig. 2.

Let us now consider the interval $(-\pi/4,\pi/4]$. It is easy for us to verify the following facts:

If $\delta_1>\delta_2$, $-\pi/4$ is an exit and $\pi/4$ is an entrance; if $\delta_1<\delta_2$, instead, $-\pi/4$ is an entrance and $\pi/4$ is an exit. For the case of $\delta_1=\delta_2$, $-\pi/4$ and $\pi/4$ are both ANBs. The results are depicted in Fig. 3.

On $(\pi/4,\pi/2]$, via the same procedure, we know: If $\delta_1>\delta_2$, $\pi/4$ is an entrance, if $\delta_1<\delta_2$, $\pi/4$ is an exit and if $\delta_1=\delta_2$, $\pi/4$ is an ANB. The other boundary $\pi/2$ is always an entrance on such an interval. The various situations are summarized in Fig. 4.

As was stated in [18] (p. 132), a stationary solution to a FPK equation does not exist, if each of the two boundaries is either an exit, attractively natural, or strictly natural, from which we find that under the conditions of $\delta_1>\delta_2$ and $\delta_1=\delta_2$, the invariant measures do not exist on $(-\pi/4,\pi/4]$ and $[-\pi/2,\pi/2]$, respectively. Thus in this study, the stationary solution to Eq. (45) will be able to be investigated only under the condition of $\delta_1<\delta_2$.

Next, upon each sub-interval, we determine the solution to Eq. (45). First we consider the interval $[-\pi/2,-\pi/4]$. Since $\theta=-\pi/2$ is an entrance while $\theta=-\pi/4$ is an exit, we know that on $[-\pi/2,-\pi/4]$, the solution to Eq. (45) is a Dirac Delta function of the following form:

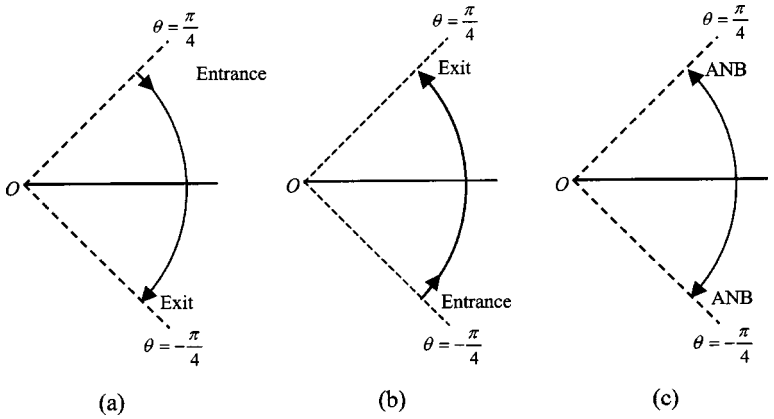


Fig. 3 Boundary diffusion behavior of interval $(-\pi/4,\pi/4]$ for the cases of (*a*) $\delta_1>\delta_2$, (*b*) $\delta_1<\delta_2$, (*c*) $\delta_1=\delta_2$

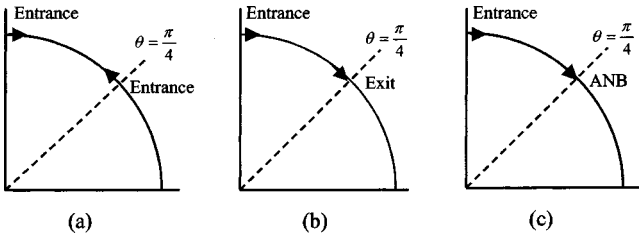


Fig. 4 Boundary diffusion behavior of interval $(\pi/4,\pi/2]$ for the cases of (*a*) $\delta_1>\delta_2$, (*b*) $\delta_1<\delta_2$, (*c*) $\delta_1=\delta_2$

$$F(\theta)=C\delta\left(\theta+\frac{\pi}{4}\right),\quad \theta\in\left[-\frac{\pi}{2},-\frac{\pi}{4}\right] \tag{69}$$

where C is an integral constant which can be determined by the normalization condition of $F(\theta)$ on the whole interval $[-\pi/2,\pi/2]$.

Similarly on interval $(\pi/4,\pi/2]$, the invariant measure is also a Dirac Delta function, i.e.,

$$F(\theta)=C\delta\left(\theta-\frac{\pi}{4}\right),\quad \theta\in\left(\frac{\pi}{4},\frac{\pi}{2}\right]. \tag{70}$$

On interval $(-\pi/4,\pi/4]$, since the two boundaries are both entrances, the invariant measure is

$$F(\theta)=Cm(\theta)=\frac{C}{8\kappa_5}[\sec 2\theta]^{3/2}\cos\theta\exp\left[\frac{\Delta}{\kappa_5}\sec 2\theta\right],\quad \theta\in\left(-\frac{\pi}{4},\frac{\pi}{4}\right]. \tag{71}$$

By synthesizing the results in Eqs. (69)–(71), we obtain the solution to Eq. (45) upon the whole interval $[-\pi/2,\pi/2]$, i.e.,

$$F(\theta)=\begin{cases} C\delta\left(\theta+\frac{\pi}{4}\right), & \theta\in\left[-\frac{\pi}{2},-\frac{\pi}{4}\right] \\ \frac{C}{8\kappa_5}[\sec 2\theta]^{3/2}\cos\theta\exp\left[\frac{\Delta}{\kappa_5}\sec 2\theta\right], & \theta\in\left(-\frac{\pi}{4},\frac{\pi}{4}\right] \\ C\delta\left(\theta-\frac{\pi}{4}\right), & \theta\in\left(\frac{\pi}{4},\frac{\pi}{2}\right]. \end{cases} \tag{72}$$

6 Asymptotic Expansion for Maximal Lyapunov Exponent

Under the assumption that the Fokker Planck operator L_ε defined by Eq. (9) is an ergodic operator, the maximal Lyapunov exponent on the domain $[0,2\pi]\times[-\pi/2,\pi/2]\times R^n$, for the stochastic bifurcation system (5) is therefore given as

$$\lambda_\varepsilon=\langle\rho_\varepsilon,p_\varepsilon\rangle=\int_0^{2\pi}d\phi\int_{-\pi/2}^{\pi/2}d\theta\int_{R^n}d\mathbf{u}[\rho_\varepsilon p_\varepsilon] \tag{73}$$

where $p_\varepsilon(\theta,\phi,\mathbf{u})$ is the stationary probability density which admits the asymptotic expansion (21), and ρ_ε has been defined in Eq. (6). According to the discussion in Section 3, it can be easily shown that the asymptotic expansion of the top Lyapunov exponent

$$\langle\rho_\varepsilon,p_0\rangle=\varepsilon\langle f(\mathbf{u})\rho_1,p_0\rangle+\varepsilon^2[\langle\rho_2,p_0\rangle+\langle\rho_1,p_1\rangle]+\cdots$$

is reasonable. In addition, the assumption on the stochastic function $f(\mathbf{u})$ leads to

$$\langle f(\mathbf{u})\rho_1(\theta,\phi),p_0\rangle=\frac{1}{2\pi}\int_{R^n}f(\mathbf{u})\psi_0(\mathbf{u})d\mathbf{u}\times\int_0^{2\pi}d\phi\int_{-\pi/2}^{\pi/2}\rho_1(\theta,\phi)F(\theta)d\theta=0.$$

As a result, the asymptotic expansion of the relevant maximal Lyapunov exponent for system (5) becomes

$$\lambda_\varepsilon=\varepsilon^2[\langle\rho_2,p_0\rangle+\langle f(\mathbf{u})\rho_1,p_1\rangle]+o(\varepsilon^2). \tag{74}$$

In calculating the asymptotic expansion of the top Lyapunov exponent, the computations of the solution functions $p_{\mathrm{m}}^{(1)}(\theta,\phi)$ are required.

6.1 Case I. For the first case, condition (57) leads to

$$M_0=\tfrac{1}{4}(2b_{33}-k_1)\Lambda_0,\quad \Lambda_0=\frac{d}{d\theta}[F(\theta)\sin 2\theta]$$

$$M_2^{(1)}=-k_2\Lambda_2,\quad M_2^{(2)}=-k_3\Lambda_2,\quad \Lambda_2=F(\theta)+\tfrac{1}{4}\Lambda_0$$

$$M_1^{(1)}=b_{32}\Lambda_{11},\quad M_1^{(2)}=b_{31}\Lambda_{11},\quad \Lambda_{11}=\frac{d}{d\theta}[F(\theta)\cos^2\theta] \tag{75}$$

$$p_{\mathbf{m}}^{(1)}(\theta,\phi)=\begin{cases}\frac{1}{2\pi}p_0^{(1)}(\theta), & m=0\\ \frac{1}{2\pi}\langle f(\mathbf{u}),\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\{\Pi_{\mathbf{m}}^{(0)}\Lambda_0+\Pi_{\mathbf{m}}^{(2)}\Lambda_2+\Pi_{\mathbf{m}}^{(11)}\Lambda_{11}\}, & m\neq 0.\end{cases}$$

Substitution of Eq. (75) into Eq. (74) yields

$$\begin{aligned}\langle f(\mathbf{u})\rho_1,p_1\rangle &= \frac{3}{16}\beta_2-\frac{3}{8}\frac{\beta_2}{2+\gamma_1}-\frac{\beta_1}{\gamma_2(2+\gamma_1)}\\ &+\frac{1}{4}\kappa_5\sqrt{\gamma_2}\frac{\gamma_1\gamma_2}{6+\gamma_1}F\left[\frac{3}{2},\frac{3}{2}+\frac{1}{4}\gamma_1;\frac{5}{2}+\frac{1}{4}\gamma_1;1-\gamma_2\right]\\ &-\frac{1}{16}\beta_2\gamma_1\sqrt{\gamma_2}\frac{\gamma_1\gamma_2-4\gamma_2+6}{(6+\gamma_1)(2+\gamma_1)}F\left[\frac{3}{2},\frac{3}{2}+\frac{1}{4}\gamma_1;\frac{5}{2}\right.\\ &\left.+\frac{1}{4}\gamma_1;1-\gamma_2\right]+3\beta_1\sqrt{\gamma_2}\frac{(\gamma_1\gamma_2+2)}{(6+\gamma_1)(2+\gamma_1)}\\ &\times F\left[\frac{5}{2},\frac{3}{2}+\frac{1}{4}\gamma_1;\frac{5}{2}+\frac{1}{4}\gamma_1;1-\gamma_2\right],\quad 0<\beta<2\kappa_1\end{aligned} \tag{76}$$

$$\langle\rho_2,p_0\rangle=-\delta_1-\frac{2\sqrt{\gamma_2}\Delta}{(2+\gamma_1)}F\left[\frac{3}{2},\frac{1}{2}+\frac{1}{4}\gamma_1;\frac{3}{2}+\frac{1}{4}\gamma_1;1-\gamma_2\right],\quad 0<\beta<2\kappa_1$$

where

$$\gamma_1=\frac{-8\Delta+\beta_2}{\beta}=\frac{-8(\delta_1-\delta_2)+\beta_2}{\beta},\quad \gamma_2=\frac{\beta}{\kappa_1}.$$

$F[a,b;c;z]$ is the confluent hypergeometric function ([19] (p. 41)) which is defined as

$$\begin{aligned}F[a,b;c;z] &= \sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{n!(c)_n}z^n\\ &=\sum_{n=0}^{\infty}\frac{[a(a+1)\cdots(a+n-1)][b(b+1)\cdots(b+n-1)]}{n![c(c+1)\cdots(c+n-1)]}z^n.\end{aligned}$$

The convergence condition for this series is $\|z\|<1$. Thus for the confluent hypergeometric functions in Eq. (76), the convergence condition is $0<\beta<2\kappa_1$.

With these results, the asymptotic expansion of the maximal Lyapunov exponent can be expressed as

$$\begin{aligned}\lambda_\varepsilon=\varepsilon^2\Bigg\{&-\delta_1+\frac{3}{16}\beta_2-\frac{3}{8}\frac{\beta_2}{2+\gamma_1}-\frac{\beta_1}{\gamma_2(2+\gamma_1)}\\ &+\frac{1}{4}\kappa_5[\gamma_2]^{3/2}\frac{\gamma_1}{6+\gamma_1}F\left[\frac{3}{2},\frac{3}{2}+\frac{1}{4}\gamma_1;\frac{5}{2}+\frac{1}{4}\gamma_1;1-\gamma_2\right]\\ &-\frac{1}{16}\beta_2\gamma_1\sqrt{\gamma_2}\frac{\gamma_1\gamma_2-4\gamma_2+6}{(6+\gamma_1)(2+\gamma_1)}F\left[\frac{3}{2},\frac{3}{2}+\frac{1}{4}\gamma_1;\frac{5}{2}+\frac{1}{4}\gamma_1;1\right.\\ &\left.-\gamma_2\right]+3\beta_1\sqrt{\gamma_2}\frac{(\gamma_1\gamma_2+2)}{(6+\gamma_1)(2+\gamma_1)}F\left[\frac{5}{2},\frac{3}{2}+\frac{1}{4}\gamma_1;\frac{5}{2}\right.\\ &\left.+\frac{1}{4}\gamma_1;1-\gamma_2\right]-\frac{2\sqrt{\gamma_2}\Delta}{(2+\gamma_1)}F\left[\frac{3}{2},\frac{1}{2}+\frac{1}{4}\gamma_1;\frac{3}{2}+\frac{1}{4}\gamma_1;1\right.\\ &\left.-\gamma_2\right]\Bigg\}+o(\varepsilon^2).\end{aligned} \tag{77}$$

6.2 Case II. For the second case, based on the results of Eq. (66), we obtain

$$M_0=\tfrac{1}{4}(2b_{33}-k_1)\Lambda_0,\quad \Lambda_0=\frac{d}{d\theta}[F(\theta)\sin 2\theta]$$

$$M_2^{(1)}=-k_2\Lambda_2,\quad M_2^{(2)}=-k_3\Lambda_2,\quad \Lambda_2=F(\theta)+\tfrac{1}{4}\Lambda_0$$

$$M_1^{(1)}=b_{23}[\Lambda_{11}-\Lambda_{12}]=b_{23}\Lambda_1,\quad M_1^{(2)}=b_{13}\Lambda_1$$

$$\Lambda_1=\frac{d}{d\theta}[F(\theta)\cos 2\theta]-F(\theta)\tan\theta \tag{78}$$

$$p_{\mathbf{m}}^{(1)}(\theta,\phi)=\begin{cases}\dfrac{1}{2\pi}p_0^{(1)}(\theta), & m=0\\ \dfrac{1}{2\pi}\langle f(\mathbf{u}),\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\{\Pi_{\mathbf{m}}^{(0)}\Lambda_0+\Pi_{\mathbf{m}}^{(2)}\Lambda_2+\Pi_{\mathbf{m}}^{(11)}\Lambda_1\}, & m\neq 0\end{cases}$$

where

$$\Lambda_0=\cos\theta[-1+3\cos 2\theta+\alpha\sin 2\theta]\exp[\alpha(\sin\theta)^2]$$

$$\Lambda_1=\tfrac{1}{2}[\cos\theta]^3[3+\alpha-\alpha\cos 2\theta]\exp[\alpha(\sin\theta)^2] \tag{79}$$

$$\Lambda_2=2\sin\theta[\cos\theta]^2[-3+\alpha\cos 2\theta]\exp[\alpha(\sin\theta)^2].$$

Substitution of Eqs. (78) and (79) into Eq. (74) yields

$$\langle f(\mathbf{u})\rho_1,p_1\rangle=\frac{1}{2}\left(\kappa_5+\frac{\beta_2}{4}\right)+\frac{J_2}{2J_0}\left(\kappa_5-\frac{\beta_2}{4}\right)$$

$$\langle\rho_2,p_0\rangle=-\delta_1+\frac{J_2}{J_0}\Delta=-\delta_1+\frac{J_2}{J_0}(\delta_1-\delta_2) \tag{80}$$

$$J_0=\int_0^1[\exp(\alpha x^2)]dx,\quad J_2=\int_0^1[x^2\exp(\alpha x^2)]dx.$$

The asymptotic expansion of the top Lyapunov exponent is then given as

$$\lambda_\varepsilon=\begin{cases}\varepsilon^2\left[-\delta_1+\dfrac{\kappa_5}{4}+\dfrac{\beta_2}{8}+\dfrac{\kappa_5}{4\sqrt{\pi}}\dfrac{i\exp[\alpha]\sqrt{\alpha}}{\mathrm{Erfi}[\sqrt{\alpha}]}\right]+o(\varepsilon^2), & \alpha>0\\ \varepsilon^2\left[-\delta_1+\dfrac{\kappa_5}{4}+\dfrac{\beta_2}{8}+\dfrac{\kappa_5}{4\sqrt{\pi}}\dfrac{\exp[\alpha]\sqrt{-\alpha}}{\mathrm{Erf}[\sqrt{-\alpha}]}\right]+o(\varepsilon^2), & \alpha\leqslant 0.\end{cases} \tag{81}$$

From Eq. (63), we find that condition $\alpha=0$ implies

$$\delta_1=\delta_2+[\tfrac{1}{8}S_f(2\omega)[k_2^2+k_3^2]-\tfrac{1}{2}S_f(\omega)[b_{13}^2+b_{23}^2]]$$

from which the inequalities in Eq. (81) can be easily obtained.

6.3 Case III. For the third case, since

$$M_0=M_2^{(1)}=M_2^{(2)}=0$$

$$M_1^{(1)}=b_{23}\Lambda_1,\quad M_1^{(2)}=b_{13}\Lambda_1$$

$$p_{\mathbf{m}}^{(1)}(\theta,\phi)=\begin{cases}\dfrac{1}{2\pi}p_0^{(1)}(\theta), & m=0\\ \dfrac{1}{2\pi}\langle f(\mathbf{u}),\psi_{\mathbf{m}}^*(\mathbf{u})\rangle_E\Pi_{\mathbf{m}}^{(11)}\Lambda_1, & m\neq 0\end{cases}$$

the terms in expression (74) can be evaluated as

$$\begin{aligned}\langle\rho_2,p_0\rangle&=-\frac{1}{2}(\delta_1+\delta_2)-\frac{\Delta}{2}\frac{I_1}{I_0}+C\Delta\frac{I_1}{I_0}\\&=-\frac{1}{2}(\delta_1+\delta_2)+\frac{1}{2}(\delta_2-\delta_1)\frac{I_1}{I_0}-(\delta_2-\delta_1)\frac{I_1}{I_0}C\end{aligned} \tag{82}$$

$$\langle f(\mathbf{u})\rho_1,p_1\rangle=\frac{1}{4}\kappa_5+2\kappa_5C+\frac{1}{4}\kappa_5\left[2\frac{I_2}{I_0}-\frac{I_1}{I_0}\right]-\frac{1}{2}\kappa_5C\left[2\frac{I_2}{I_0}-\frac{I_1}{I_0}\right]$$

where

$$I_0=\int_{-\infty}^{+\infty}\exp[-\kappa x^2]\mathrm{d}x=\frac{\sqrt{\pi}}{\sqrt{\kappa}}$$

$$I_1=\int_{-\infty}^{+\infty}\frac{1}{1+x^2}\exp[-\kappa x^2]\mathrm{d}x=\pi e^\kappa(1-\mathrm{erf}(\sqrt{\kappa}))$$

$$I_2=\int_{-\infty}^{+\infty}\left[\frac{1}{1+x^2}\right]^2\exp[-\kappa x^2]\mathrm{d}x=\sqrt{\pi\kappa}+\frac{1}{2}\pi e^\kappa(1-2\kappa)(1-\mathrm{erf}(\sqrt{\kappa}))$$

$$C^{-1}=\left[2+\frac{I_0}{8\sqrt{2}\kappa_5 e^\kappa}\right]$$

$$\kappa=-\frac{\Delta}{\kappa_5}=\frac{\delta_2-\delta_1}{\kappa_5}>0.$$

Then for the third case, the analytical expression of the maximal Lyapunov exponent of system (5) is obtained as

$$\lambda_\varepsilon=\varepsilon^2\left\{-\delta_1+\frac{\kappa_5}{4}+\lambda_2^{(1)}+\lambda_2^{(2)}\right\}+o(\varepsilon^2)$$

$$\lambda_2^{(1)}=\frac{1}{2}\left\{\left[(\delta_2-\delta_1)-\frac{1}{2}\kappa_5\right]R_1+\kappa_5R_2\right\} \tag{83}$$

$$\lambda_2^{(2)}=C\left\{(\delta_2-\delta_1)(R_1-1)2\kappa_5+\frac{1}{2}\kappa_5(R_1-2R_2)\right\}$$

where

$$R_1=\frac{I_1}{I_0}=\sqrt{\pi}\sqrt{\kappa}e^\kappa[1-\mathrm{erf}(\sqrt{\kappa})]$$

$$R_2=\frac{I_2}{I_0}-\kappa=\frac{1}{2}\sqrt{\pi}\sqrt{\kappa}e^\kappa[1-2\kappa][1-\mathrm{erf}(\sqrt{\kappa})].$$

7 Conclusions

In this paper, the explicit asymptotic expansions for the maximal Lyapunov exponent of a co-dimension two-bifurcation system driven by a small-intensity real noise process have been constructed. To account for a sufficiently general model, the real noise was assumed to be an integrable function of the output of a linear filter system, viz., a zero-mean stationary Gaussian diffusion process. The strong mixed condition and the detailed balance condition were removed in the present theoretical formulation. The method used in the present study involved the use of a perturbation method and the spectrum representations of the Fokker Planck operator and its adjoint operator of the linear filter system. Three special cases, one of which the singularity of the diffusion coefficient arose, were considered. The associated maximal Lyapunov exponents were evaluated accordingly.

References

[1] Khasminskii, R. Z., 1967, "Necessary and Sufficient Conditions for the Asymptotic Stability of Linear Stochastic Systems," Theor. Probab. Appl., **12**, pp. 144–147.

[2] Mitchell, P. R., and Kozin, F., 1974, "Sample Stability of Second Order Linear Differential Equations With Wide Band Noise Coefficients," SIAM (Soc. Ind. Appl. Math.) J. Appl. Math., **27**, pp. 571–605.

[3] Nishoka, K., 1976, "On the Stability of Two-Dimensional Linear Stochastic Systems," Kodai Math. Sem. Rep., **27**, pp. 211–230.

[4] Ariaratnam, S. T., and Xie, W. C., 1992, "Lyapunov Exponents and Stochastic Stability of Coupled Linear Systems Under Real Noise Excitation," ASME J. Appl. Mech., **59**, pp. 664–673.

[5] Xie, W. C., 2002, "Moment Lyapunov Exponents of Two-Dimensional Viscoelastic System Under Bounded Noise Excitation," ASME J. Appl. Mech., **69**, pp. 346–357.

[6] Arnold, L., Papanicolaou, G., and Wihstutz, V., 1986, "Asymptotic Analysis of the Lyapunov Exponents and Rotation Numbers of the Random Oscillator and Applications," SIAM (Soc. Ind. Appl. Math.) J. Appl. Math., **46**, pp. 427–450.

[7] Namachchivaya, N. S., and Van Roessel, H. J., 1993, "Maximal Lyapunov Exponent and Rotation Numbers for Two Coupled Oscillators Drivern by Real Noise," J. Stat. Phys., **71**, pp. 549–567.

[8] Doyle, M. M., and Namachchivaya, N. Sri., 1994, "Almost-Sure Asymptotic Stability of a General Four-Dimensional System Driven by Real Noise," J. Stat. Phys., **75**, pp. 525–552.

[9] Namachchivaya, N. S., and Taiwar, S., 1993, "Maximal Lyapunov Exponent and Rotation Numbers for Stochastically Perturbed Co-dimension Two Bifurcation," J. Sound Vib., **169**, pp. 349–372.

[10] Gardiner, C. W., 1997, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 2nd Ed., Springer, Berlin.

[11] Liu, X. B., and Liew, K. M., 2003, "The Lyapunov Exponent for a

Co-Dimension Two Bifurcation System Driven by a Real Noise,'' Int. J. Non-Linear Mech., **38**, pp. 1495–1511.
[12] Liu, X. B., and Liew, K. M., 2004, ''On the Almost-Sure Stability Condition for a Co-Dimensional Two Bifurcation System Under the Parametric Excitation of a Real Noise,'' J. Sound Vib., **272**, pp. 85–107.
[13] Roy, R. V., 1994, ''Stochastic Averaging of Oscillators Excited by Coloured Gaussian Processes,'' Int. J. Non-Linear Mech., **29**, pp. 461–475.
[14] Karlin, S., and Taylor, H. M., 1981, *A Second Course in Stochastic Processes*, Academic Press, San Diego, CA.
[15] Liberzon, D., and Brockett, R. W., 2000, ''Spectral Analysis of Fokker-Planck and Related Operators Arising From Linear Stochastic Differential Equations,'' SIAM J. Control Optim., **38**, pp. 1453–1467.
[16] Guckenheimer, G., and Holmes, P., 1983, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York.
[17] Caughey, T. K., 1971, ''Nonlinear Theory of Random Vibrations,'' *Advances in Applied Mechanics*, Academic Press, New York, **11**, pp. 209–253.
[18] Lin, Y. K., and Cai, G. Q., 1995, *Probabilistic Structural Dynamics, Advanced Theory and Applications*, McGraw-Hill, New York.
[19] Seaborn, J. B., 1991, *Hypergeometric Functions and Their Applications*, Springer-Verlag, New York.

A General Solution for Two-Dimensional Stress Distributions in Thin Films

R. Krishnamurthy[1]

D. J. Srolovitz

Princeton Institute for the Science and Technology of Materials, and Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08540

We present closed-form solutions for stresses in a thin film resulting from a purely dilatational stress-free strain that can vary arbitrarily within the film. The solutions are specific to a two-dimensional thin film on a thick substrate geometry and are presented for both a welded and a perfectly slipping film/substrate interface. Variation of the stress-free strain through the thickness of the film is considered to be either arbitrary or represented by a Fourier integral, and solutions are presented in the form of a Fourier series in the lateral dimension x. The Fourier coefficients can be calculated rapidly using Fast Fourier Transforms. The method is applied to determine the stresses in the film and substrate for three cases: (a) where the stress-free strain is a sinusoidal modulation in x, (b) where the stress-free strain varies only through the thickness, and (c) where a rectangular inclusion is embedded within the film, and the calculated stresses match accurately with the exact solutions for these cases. [DOI: 10.1115/1.1782649]

1 Introduction

A variety of processes occurring in thin films generate intrinsic strains. These include, for example, strains associated with heteroepitaxy, thermal expansion mismatch, defect incorporation, and compositional gradients. The residual stresses that develop as a result of the presence of these intrinsic strains can be obtained by solving a suitably formulated thermoelastic problem. Indeed, thermal and compositional stresses have been analyzed in this way, [1,2]. A general approach to solving the thermoelastic problem involves the use of the Goodier thermoelastic potential, [1]. However, this potential can be obtained analytically for only very simple geometries. A configuration in which a film rests upon a much thicker substrate is very common in thin film microelectronics, thermally grown oxide scales on bulk metals, and sputtered coatings. Typically, the film thicknesses are much smaller than any other dimension in the problem (i.e., much thinner than the substrate or the lateral extent of the film). If the intrinsic strain is independent of one of the lateral dimensions *y*, and there are no externally imposed strains, plane-strain conditions apply. Such plane strain problems can be addressed using Airy stress functions, [1]. An extremely rapid method of solving for mechanical equilibrium for simple geometries of this type involves the use of the Fourier series. Indeed, two-dimensional and three-dimensional stress distributions have been calculated using this approach, [1,3–6].

Glas considered a general modulation in the lattice parameter parallel to the substrate–film interface (i.e., intrinsic strains are a function of the lateral dimension *x* only), and analytically described the resultant stresses in the form of a Fourier series in *x*, [7]. He also applied this approach to determine the stresses in a capped film, in a film with steps at the film-substrate interface and in a film with embedded inclusions that are of the same thickness as the film, [8]. In all of these cases, the intrinsic strain does not vary through the thickness of the film. This limitation is unduly restrictive since many thin film growth situations necessarily involve the variation of composition or growth strains through the film thickness. This leads to the variation of stress in the *z*-direction. For example, a linear variation of stress along the growth direction was observed (using Raman spectroscopy) in diamond films grown on metal substrates by chemical vapor deposition, [9]. The observation that oxide films tend to curl on separation from the metal substrate upon which they were grown suggest the presence of such through thickness variations in growth strains, [10]. The presence of strain gradients within such films were confirmed by direct measurements, [11]. In the current work, we present the solution for stresses, resulting from a general intrinsic strain that varies both laterally and through the thickness of the film. We follow the Eshelby procedure for obtaining residual stresses to formulate the relaxation problem, [12]. Following Glas's approach, we represent the stress-free strain as a Fourier series in *x*, and solve for the Fourier coefficients that satisfy the equilibrium equations and boundary conditions. While the interface remains welded (i.e., continuity of tractions and displacements) for epitaxial deposition and growth, there are some instances such as oxide growth at elevated temperatures, where the interfacial diffusivity is high, and accordingly, the interface behaves like a liquid and cannot support shear. This interface may be better described as slipping (continuity of normal tractions and displacements, zero shear tractions) [13].

2 Problem Formulation

We consider a planar thin film of thickness H resting on a substrate that is semi-infinite and has the same elastic modulus E and poisson's ratio ν, as the film (see Fig. 1). The film and the substrate are infinite in the lateral directions, x and y, and the displacement in the y-direction is identically zero. We further assume that the stress-free strain in the film, ε^m, is purely dilatational and is only a function of x and z, i.e.,

$$\varepsilon^m = \varepsilon^m(x,z). \tag{1}$$

The stress-free strain could have different physical origins, such as strains induced due to thermal mismatch, composition strains, growth strains, etc., depending on the actual problem of interest. We also assume that there are no stress-free strains in the substrate. These assumptions imply that a state of plane strain applies for this system.

We follow Eshelby's procedure for obtaining residual stresses to develop the equilibrium equations and boundary conditions for a general thermoelastic problem with a stress-free strain of the

[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 Applied Mechanics Division, August 26, 2003; final revision, April 22, 2004. Associate Editor: K. Ravi-Chandar. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

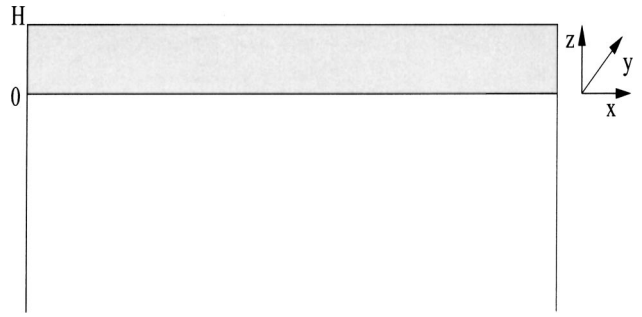


Fig. 1 A schematic illustration of a thin film on a thick rigid substrate. The coordinate axes in film thickness *H* are labeled.

form given in Eq. (1), [12]. A schematic description of Eshelby's procedure is shown in Fig. 2. The film is divided into infinitesimal cubes that are each removed from the film and allowed to freely transform (i.e., undergo the stress-free strain ε^m). Now, we impose external tractions on each of these infinitesimal elements to counteract the stress-free strain in order to deform them back to their original shape and size. These tractions, t_i^1, are given by

$$t_i^1 = -\frac{E}{1-2\nu}\varepsilon^m n_i, \tag{2}$$

where n_i is the surface normal. The stresses corresponding to these tractions are

$$\sigma_{ij}^1 = -\frac{E}{1-2\nu}\varepsilon^m \delta_{ij}, \tag{3}$$

where δ_{ij} is the Kronecker delta function. These elements are reassembled to form the original solid. This creates no additional stresses because the tractions on the cube surfaces are still present. More formally, however, it is appropriate to replace these tractions on the ''surfaces'' of the cubes that are now inside the homogeneous solid with equivalent body forces. These body forces, f_i^1, are

$$f_i^1 = \frac{E}{1-2\nu}\varepsilon_{,i}^m \quad i=x,z, \tag{4}$$

where $\varepsilon_{,i}^m$ represents the derivative of the stress-free strain with respect to the position coordinate *i*. At the film free surface, $z=H$, the surface traction, t_z^{s1}, is given by

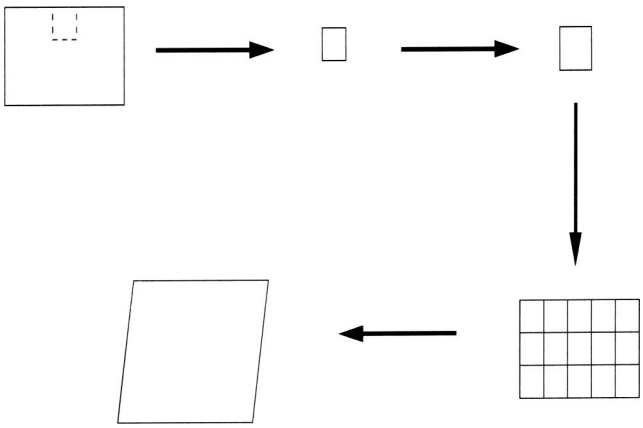


Fig. 2 A schematic illustration of the steps involved in the Eshelby procedure

$$t_z^{s1} = -\frac{E}{1-2\nu}\varepsilon^m(x,z=H). \tag{5}$$

At the interface between the substrate and the film ($z=0$), we have body forces, f_z^{i1}, given by

$$f_z^{i1} = \frac{E}{1-2\nu}\varepsilon^m(x,z=0)\,\delta(z). \tag{6}$$

Now, we remove the surface tractions and body forces that were externally imposed and let the system relax. This is achieved by imposing equal but opposite tractions and body forces (denoted with superscripts 2), to those in Eq. (2)–Eq. (6).

$$f_i^2 = -\frac{E}{1-2\nu}\varepsilon_{,i}^m, \tag{7}$$

$$t_z^{s2} = \frac{E}{1-2\nu}\varepsilon^m(x,z=H), \tag{8}$$

$$f_z^{i2} = -\frac{E}{1-2\nu}\varepsilon^m(x,z=0)\,\delta(z). \tag{9}$$

Mechanical equilibrium requires that

$$\begin{aligned} \delta\sigma_{ij,j} + f_i^2 = 0, \quad 0<z<H, \\ \delta\sigma_{ij,j} = 0, \quad z<0, \end{aligned} \tag{10}$$

where Einstein convention for repeated indices is adopted and $\delta\sigma_{ij}$ are the stresses induced by the relaxation and are related to the total stresses in the system, σ_{ij}, by

$$\sigma_{ij} = \delta\sigma_{ij} + \sigma_{ij}^1. \tag{11}$$

Tractions will vanish at the free surface of the film $z=H$, and therefore,

$$\begin{aligned} \sigma_{zz} = 0 \ \text{ or, equivalently, } \ \delta\sigma_{zz} = t_z^{s2}, \\ \sigma_{xz} = \delta\sigma_{xz} = 0. \end{aligned} \tag{12}$$

Continuity of tractions across the interface $z=0$ can be written as

$$\begin{aligned} \delta\sigma_{zz}^{\text{sub}} - \delta\sigma_{zz} = f_z^{i2}, \\ \delta\sigma_{xz}^{\text{sub}} - \delta\sigma_{xz} = 0, \end{aligned} \tag{13}$$

where the superscript sub is used to denote fields in the substrate. Deep into the substrate $z=-\infty$, displacement gradients $u_{x,i}$ and $u_{z,i}$, are expected to vanish, and accordingly

$$u_{x,i}(z=-\infty) = u_{z,i}(z=-\infty) = 0, \quad i=x,z. \tag{14}$$

Boundary conditions can be easily formulated for two special types of film-substrate interfaces, namely, when the interface is welded, and, when the interface is perfectly slipping. When the interface between the substrate and the film is welded, displacement fields are continuous across the interface, and hence

$$\begin{aligned} u_x^{\text{sub}} = u_x, \\ u_z^{\text{sub}} = u_z. \end{aligned} \tag{15}$$

These are supplemented by the traction boundary conditions presented in Eq. (13). When the interface between the substrate and the film is perfectly slipping, normal tractions are continuous across the interface while shear tractions at the interface are identically zero, and accordingly, we have

$$\begin{aligned} \delta\sigma_{zz}^{\text{sub}} - \delta\sigma_{zz} = f_z^{i2}, \\ \delta\sigma_{xz}^{\text{sub}} = \delta\sigma_{xz} = 0. \end{aligned} \tag{16}$$

We also require that normal displacements be continuous across the interface (as in Eq. 15(b)).

Equations (10), (12), (13), (14), (15), (16) along with the appropriate compatibility condition for displacements must be

solved to determine the relaxation displacement and stress fields. For the two-dimensional plane-strain problem considered here, the compatibility conditions can be combined with the mechanical equilibrium equations given in Eq. (10) to obtain the following relation for the stresses [14]:

$$\delta\sigma_{\gamma\gamma,\alpha\alpha}+\frac{1}{1-\nu}f^2_{\alpha,\alpha}=0, \quad \gamma,\alpha=x,z. \tag{17}$$

Equation (17) also makes use of the stress-strain constitutive relations valid for plane-strain elasticity:

$$\epsilon_{\alpha\beta}=\frac{1+\nu}{E}(\sigma_{ij}-\nu\sigma_{\gamma\gamma}\delta_{\alpha\beta}) \quad \alpha=\beta=\gamma=x,z, \tag{18}$$

where $\epsilon_{\alpha\beta}$ is the strain in the system. The stresses can be related to the Airy stress function, ϕ, by

$$\begin{aligned}\delta\sigma_{xx}&=\phi_{,zz}-\Omega,\\ \delta\sigma_{zz}&=\phi_{,xx}-\Omega,\\ \delta\sigma_{xz}&=-\phi_{,xz},\end{aligned} \tag{19}$$

where

$$\Omega=\frac{E}{1-2\nu}\varepsilon^m. \tag{20}$$

Accordingly, we can rewrite Eq. (17) as

$$\phi_{,\alpha\alpha\gamma\gamma}+\frac{E}{1-\nu}\varepsilon^m_{,\gamma\gamma}=0. \tag{21}$$

Equation (21) can be solved without the boundary conditions to obtain a particular solution. This solution can then be superposed with the general solution of the homogeneous equation,

$$\phi_{,\alpha\alpha\gamma\gamma}=0. \tag{22}$$

In the substrate, the stress-free strain is identically zero and the stress function satisfies the homogeneous equation. Equation (22) is solved for both the film and the substrate with the boundary conditions presented in Eqs. (12), (13), (14), (15), and (16).

3 Solution Method

We represent the stress-free strain ε^m in the form of a Fourier series in x, as

$$\varepsilon^m(x,z)=\sum_\alpha e^{i\alpha x}\overline{\varepsilon^m}(\alpha,z), \tag{23}$$

where $\overline{\varepsilon^m}$ are the Fourier coefficients, and $\alpha=2\pi/L$. Equation (23) represents a stress-free strain that is periodic in x with period L. While the summation in Eq. (23) can be written as a Fourier integral to represent any general function in x, Fourier integrals can be computed analytically for only a limited set of functions. Numerical computation of the Fourier integral uses the discrete Fourier transform which approximates the Fourier integral by a Fourier series. In accordance with Eq. (23), we look for solutions for the stress function ϕ that have the form

$$\phi=\sum_\alpha e^{i\alpha x}\gamma(\alpha,z), \tag{24}$$

where γ represent the Fourier coefficients of ϕ. Substituting, Eqs. (23) and (24) into (21), we find that γ satisfies the relation

$$\gamma''''-2\alpha^2\gamma''+\alpha^4\gamma+\frac{E}{1-\nu}(\overline{\varepsilon^m}''-\alpha^2\overline{\varepsilon^m})=0, \tag{25}$$

where the superscript $'$ denotes differentiation with respect to z. A particular solution of Eq. (25) is given by

$$\gamma=-\frac{E}{1-\nu}e^{\alpha z}I_2(\alpha,z), \tag{26}$$

where $I_2(\alpha,z)$ is defined by the equations

$$\begin{aligned}I_1(\alpha,z)&=\int e^{\alpha z}\overline{\varepsilon^m}(\alpha,z)dz,\\ I_2(\alpha,z)&=\int e^{-2\alpha z}I_1(\alpha,z)dz.\end{aligned} \tag{27}$$

The integrals I_1 and I_2 can be easily evaluated if we know the functional form of $\overline{\varepsilon^m}$. For example, if $\overline{\varepsilon^m}$, is given by a general polynomial of degree N with coefficients a_n,

$$\overline{\varepsilon^m}=\sum_{n=0}^{N}a_nz^n, \tag{28}$$

the integrals I_1 and I_2 are given by

$$\begin{aligned}I_1&=\sum_{n=0}^{N}\frac{a_nz^n}{\alpha}e^{\alpha z}\sum_{j=0}^{n}(-1)^j\frac{C^n_jj!}{(\alpha z)^j},\\ I_2&=\sum_{n=0}^{N}\frac{a_nz^n}{\alpha^2}e^{-\alpha z}\sum_{j=0}^{n}(-1)^{j+1}C^n_jj!\sum_{k=0}^{n-j}\frac{C^{n-j}_kk!}{(\alpha z)^{j+k}},\end{aligned} \tag{29}$$

where C^n_j represents the number of ways in which j objects can be distributed among n locations. When the functional form of ε^m is not known, we can replace the integrals I_1 and I_2, by definite integrals without any loss in generality and evaluate them numerically. We adopt two approaches, namely, writing $\overline{\varepsilon^m}$ as a Fourier integral in z and using Chebyshev polynomials to perform the numerical integration. In the first approach, we write

$$\overline{\varepsilon^m}(\alpha,z)=\int_{-\infty}^{\infty}e^{i\beta z}\tilde{\varepsilon}^m d\beta. \tag{30}$$

where $\tilde{\varepsilon}^m$ is the Fourier transform (in z) of $\overline{\varepsilon^m}$ and $\beta=2\pi/H$.

Particular solutions for stresses and strains can be obtained by combining Eq. (26) and Eq. (19). These are superposed with the general solution of the homogeneous equation, i.e., Eq. (22), to obtain the complete solutions for the relaxation stresses and displacements. We seek solutions of Eq. (22) that are of the form prescribed in Eq. (24). They must satisfy the relation

$$\gamma''''-2\alpha^2\gamma''+\alpha^4\gamma=0. \tag{31}$$

The general solution to Eq. (31) is

$$\begin{aligned}\gamma(\alpha,z)=&P_1(\alpha)\cosh\alpha z+P_2(\alpha)\sinh\alpha z+P_3(\alpha)z\cosh\alpha z\\ &+P_4(\alpha)z\sinh\alpha z.\end{aligned} \tag{32}$$

In the substrate, the general solution is of a similar form:

$$\begin{aligned}\gamma(\alpha,z)=&Q_1(\alpha)\cosh\alpha z+Q_2(\alpha)\sinh\alpha z+Q_3(\alpha)z\cosh\alpha z\\ &+Q_4(\alpha)z\sinh\alpha z.\end{aligned} \tag{33}$$

The coefficients P_1-P_4 and Q_1-Q_4 are obtained by substituting Eqs. (32) and (33) in Eqs. (12), (13), (14), and (15) when the film–substrate interface is welded and into Eqs. (12), (14), (15*b*), and (16) when the film–substrate interface is perfectly slipping and solving the ensuing sets of linear equations. For imposing displacement continuity across the welded interface (Eq. (15)), we follow the approach adopted by Glas [7]. Accordingly, continuity of u_x across the interface $z=0$ is ensured if the strain ϵ_{xx} is continuous across the interface, and continuity of u_z across the interface is ensured if du_z/dx is continuous across the interface. The resulting solutions for the coefficients for the welded film–substrate interface and $\alpha>0$ are

$$P_1=-HP_3+e^{-2\alpha H}\left(D\frac{G_2^i}{2\alpha}-G_1^He^{\alpha H}\right)-\frac{D}{2\alpha}G_2^i,$$

$$P_2=P_1+\frac{D}{\alpha}G_2^i,$$

$$P_3=-De^{-2\alpha H}(G_2^i-G_2^He^{\alpha H}),$$

$$P_4=Q_3=Q_4=P_3,\quad Q_1=Q_2=P_1+G_1^i, \tag{34}$$

and for $\alpha<0$ are

$$P_1=(e^{2\alpha H}G_1^i-G_1^He^{\alpha H})(1-2\alpha H)+DH(e^{2\alpha H}G_2^i-G_2^He^{\alpha H})-G_1^i+\frac{D}{2\alpha}G_2^i(1-e^{2\alpha H}),$$

$$P_2=-P_1+\left(\frac{D}{\alpha}G_2^i-2G_1^i\right),$$

$$P_3=-e^{\alpha H}(D(e^{\alpha H}G_2^i-G_2^H)-2\alpha(e^{\alpha H}G_1^i-G_1^H)), \tag{35}$$

$$P_4=-Q_3=Q_4=-P_3,\quad Q_1=-Q_2=P_1+G_1^i.$$

In Eqs. (34) and (35), $D=E/(1-\nu)$ and the parameters G_1^i, G_2^i, G_1^H and G_2^H are functions of α and are related to the values of the integrals I_1 and I_2 at $z=0$ and $z=H$ by

$$G_1^i=-DI_2(\alpha,z=0),\quad G_2^i=I_1(\alpha,z=0),$$
$$G_1^H=-De^{\alpha H}I_2(\alpha,z=H),\quad G_2^H=e^{-\alpha H}I_1(\alpha,z=H). \tag{36}$$

For a perfectly, slipping interface between the substrate and the film, the coefficients for $\alpha>0$ are

$$P_1=\Big((e^{3\alpha H}-e^{\alpha H})(G_1^H+DHG_2^H)+2\alpha He^{\alpha H}G_1^H+\frac{D}{2\alpha}G_2^i(1+e^{4\alpha H}-2e^{2\alpha H})-2\alpha^2H^2e^{2\alpha H}G_1^i\Big)\Big/N_p,$$

$$P_2=P_1-\frac{D}{\alpha}G_2^i(e^{4\alpha H}-e^{2\alpha H}+2\alpha He^{2\alpha H}(1-\alpha H))/N_p,$$

$$P_3=-\alpha(P_2+G_1^i)+DG_2^i, \tag{37}$$

$$P_4=\Big(\alpha(e^{\alpha H}+e^{3\alpha H})(G_1^H+DHG_2^H)+2\alpha^2H(e^{2\alpha H}G_1^i-e^{\alpha H}G_1^H)-2De^{3\alpha H}G_2^H+e^{4\alpha H}\left(\frac{D}{2}G_2^i-\alpha G_1^i\right)+(2D(1-\alpha H)G_2^i-\alpha G_1^i)e^{2\alpha H}-\frac{D}{2}G_2^i\Big)\Big/N_p,$$

$$Q_1=Q_2=P_1+G_1^i,\quad Q_3=Q_4=-\alpha Q_1,$$

$$N_p=e^{2\alpha H}(1-e^{2\alpha H}-2\alpha H+2\alpha^2H^2),$$

and for $\alpha<0$ are

$$P_1=\Big((e^{\alpha H}-e^{3\alpha H})(G_1^H+DHG_2^H)+(e^{4\alpha H}+1)\left(G_1^i-\frac{D}{2\alpha}G_2^i\right)+e^{2\alpha H}\left(\frac{D}{\alpha}G_2^i-2G_1^i\right)-2\alpha He^{\alpha H}(G_1^H+e^{\alpha H}\alpha HG_1^i)\Big)\Big/N_m,$$

$$P_2=(e^{3\alpha H}-e^{\alpha H})(G_1^H+DHG_2^H)+\left(\frac{D}{2\alpha}G_2^i-G_1^i\right)\Big(e^{2\alpha H}(2\alpha H-1)+e^{4\alpha H}\Big)+\alpha H(2G_1^He^{\alpha H}+DHe^{2\alpha H}G_2^i)/N_m+\frac{D}{\alpha}G_2^i-G_1^i,$$

$$P_3=-\alpha(P_2+G_1^i)+DG_2^i,$$

$$P_4=\Big(\alpha(e^{3\alpha H}-e^{\alpha H})(G_1^H+DHG_2^H-e^{\alpha H}G_1^i)+2e^{\alpha H}(1+\alpha H)\times(D(G_2^H-e^{\alpha H}G_2^i)-\alpha(G_1^H-e^{\alpha H}G_1^i))+\frac{D}{2}G_2^i(e^{4\alpha H}-1)\Big)\Big/N_m,$$

$$Q_1=-Q_2=P_1+G_1^i,\quad Q_3=-Q_4=\alpha Q_1,$$

$$N_m=e^{2\alpha H}(1+2\alpha H+2\alpha^2H^2)-1. \tag{38}$$

When we use Fourier integrals in z to represent the z-dependence of ε^m, the coefficients depend on only two parameters W_1 and W_2, defined by

$$W_1(\alpha)=\int_{-\infty}^{\infty}\frac{\beta}{\alpha^2+\beta^2}\tilde{\varepsilon}^m d\beta,$$
$$W_2(\alpha)=\int_{-\infty}^{\infty}\frac{\tilde{\varepsilon}^m}{\alpha^2+\beta^2}d\beta. \tag{39}$$

When the film–substrate interface is welded, the coefficients P_i and Q_i for $\alpha>0$ are

$$P_1=-\frac{De^{-2\alpha H}}{2\alpha}(W_2\alpha(-1-2\alpha H+2e^{\alpha H}(1+\alpha H)+e^{2\alpha H})+iW_1(1-e^{2\alpha H}+2\alpha H(1-\alpha H)),$$

$$Q_1=Q_2=P_1+DW_2,\quad P_2=Q_1-i\frac{DW_1}{\alpha},$$

$$P_3=-De^{-2\alpha H}(1-\alpha H)(W_2\alpha-iW_1), \tag{40}$$

$$Q_3=Q_4=P_4=P_3,$$

and for $\alpha<0$ are

$$P_1=-DW_2-\frac{D}{2}(e^{\alpha H}-1)\Big(W_2(1-e^{\alpha H}+2\alpha He^{\alpha H})+i\frac{W_1}{\alpha}(2\alpha He^{\alpha H}-1-e^{\alpha H})\Big),$$

$$Q_1=P_1+DW_2,\quad P_2=-Q_1-i\frac{D}{\alpha}W_1,$$

$$P_3=De^{\alpha H}(e^{\alpha H}-1)(\alpha W_2+iW_1), \tag{41}$$

$$Q_2=-Q_1,\quad P_4=Q_4=-Q_3=-P_3.$$

When the film–substrate interface is perfectly slipping, the coefficients for $\alpha>0$ are

$$P_1=R\Big(W_2(-e^{4\alpha H}-2e^{3\alpha H}(1+\alpha H)+2e^{2\alpha H}(1+2\alpha^2H^2)+2e^{\alpha H}(1-\alpha H)-1)+i\frac{W_1}{\alpha}(e^{4\alpha H}+2\alpha He^{3\alpha H}-2e^{2\alpha H}-2\alpha He^{\alpha H}+1)\Big),$$

$$P_2 = R\Big(W_2(e^{4\alpha H} - 2e^{3\alpha H}(1+\alpha H) + 4\alpha H e^{2\alpha H} + 2e^{\alpha H}(1-\alpha H) - 1) + i\frac{W_1}{\alpha}(-e^{4\alpha H} + 2\alpha H e^{3\alpha H} - 4\alpha H e^{2\alpha H}(1-\alpha H) - 2\alpha H e^{\alpha H} + 1)\Big),$$

$$P_3 = -\alpha P_2 - iDW_1,$$

$$P_4 = R(\alpha W_2(e^{2\alpha H} - 1)(e^{2\alpha H} - 1 + 2(1-\alpha H)e^{\alpha H}) + iW_1(e^{4\alpha H} - 2e^{3\alpha H}(2-\alpha H) + 4(1-\alpha H)e^{2\alpha H} + 2\alpha H e^{\alpha H} - 1)), \quad (42)$$

$$Q_2 = Q_1 = P_1 + DW_2, \quad Q_3 = Q_4 = -\alpha Q_1,$$

$$R = \frac{D}{2e^{2\alpha H}(e^{2\alpha H} - 2\alpha^2 H^2 + 2\alpha H - 1)},$$

and for, $\alpha < 0$ are

$$P_1 = Q_1 - DW_2, \quad P_2 = -\frac{P_3}{\alpha} - i\frac{D}{\alpha}W_1,$$

$$P_3 = R_1(\alpha W_2(-e^{4\alpha H} + 2e^{3\alpha H}(1+\alpha H) - 4\alpha H e^{2\alpha H} + 2e^{\alpha H}(\alpha H - 1) + 1) + iW_1(-e^{4\alpha H} - 2\alpha H e^{3\alpha H} + 2e^{2\alpha H} + 2\alpha H e^{\alpha H} - 1)),$$

$$P_4 = R_1(\alpha W_2(e^{4\alpha H} - 2e^{3\alpha H}(1+\alpha H) - 2e^{2\alpha H} + 2e^{\alpha H}(1+\alpha H) + 1) + iW_1(e^{4\alpha H} + 2\alpha H e^{3\alpha H} - 4e^{2\alpha H}(1+\alpha H) + 2e^{\alpha H}(\alpha H + 2) - 1)), \quad (43)$$

$$Q_1 = \frac{Q_3}{\alpha} = -\frac{Q_4}{\alpha} = P_3, \quad Q_2 = -Q_1,$$

$$R_1 = -\frac{D}{2e^{2\alpha H}(2\alpha H + 1 + 2\alpha^2 H^2 - e^{-2\alpha H})}.$$

The Fourier coefficients, $\gamma(\alpha,z)$, of the stress function ϕ can be obtained by substituting the coefficients given in Eqs. (34)–(41) into Eqs. (33) and (32), respectively. Equations (24) and (19) can then be employed to determine the relaxation stresses developed in the substrate and the film. The total stresses in the system can be obtained by adding the externally imposed stresses given in Eq. (3) to the relaxation stresses, [1].

4 Results

Stresses resulting from a general two-dimensional dilatational stress-free strain in a thin film have been described by closed-form expressions. These are in the form of a Fourier series in x, where the coefficients can be rapidly evaluated using fast fourier transforms (FFT). We now provide some simple examples to test the efficacy of our approach. Two extreme cases that lend themselves to easy comparison are the cases where the stress-free strain is only a function of one variable, i.e., film–substrate $\varepsilon^m = \varepsilon^m(x)$ and $\varepsilon^m = \varepsilon^m(z)$. The first case has been considered by Glas for a welded film–substrate interface and, not surprisingly, the solutions obtained using our approach for $\varepsilon^m = \epsilon_0 \cos 2\pi x$ and $\varepsilon^m = \epsilon_0 \sin 2\pi x$ match exactly with his results, [7]. Figure 3 shows both Glas's and our results for the stresses for $z = 0.005$ and $\varepsilon^m = 0.1 \cos 2\pi x$. In the second case, an exact solution for this geometry is $\sigma_{xx} = \sigma_{yy} = -E\varepsilon^m/1-\nu$, with the other stresses identically equal to zero. We have applied our procedure for the special case of $\varepsilon^m = (z/H)^3 - z/H$. Our results again show an excellent match with this exact solution, as seen in Fig. 4. In Figs. 4 and 5 the repeat distance along x is $L = 1$, the film is $H = 0.01$ thick, and $E = 100$.

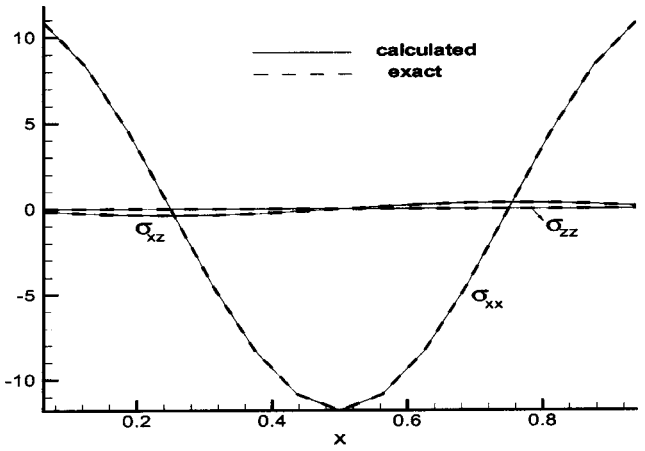


Fig. 3 Exact and calculated results for a stress-free strain that is sinusoidally modulated in x are shown at $z = 0.005$

We finally apply the method to a case where the stress-free strain ε^m varies with both x and z. We consider a periodic array of rectangular inclusions, that are infinite in extent in the y-direction, embedded in an otherwise homogeneous film. Exact solutions for this problem can be assembled using superposition from the solution given by Hu for a single inclusion embedded in a half-space, [15]. As an illustration of the results, contours of constant σ_{xx} within the film are shown in Fig. 5(a) for a mismatch of 0.0018 and an elastic modulus of 150 GPa. The film is $H = 0.1$ thick, the period in the x-direction is $L = 1$ and an inclusion of size (1/8) $\times$(1/80) is embedded in the center of the film. The values for the mismatch and the elastic modulus are from Hu's calculations for stresses in SiO_2 trenches in Si structures, [16]. Our solutions for the stresses match well with Hu's if we superpose the fields from several single inclusions (the calculations presented here are for an infinite periodic array of inclusions), as can be seen from Fig. 5(b).

5 Discussion and Conclusion

In summary, we have assumed a general form for the stress-free strain $\varepsilon^m(x,z)$ as represented by Eq. (23), and obtained solutions for the stress function ϕ that satisfy mechanical equilibrium and compatibility. Two different traction transfer modes, namely, perfectly slipping and welded, are assumed for the film–substrate interface. Solutions for the stress function are provided in terms of

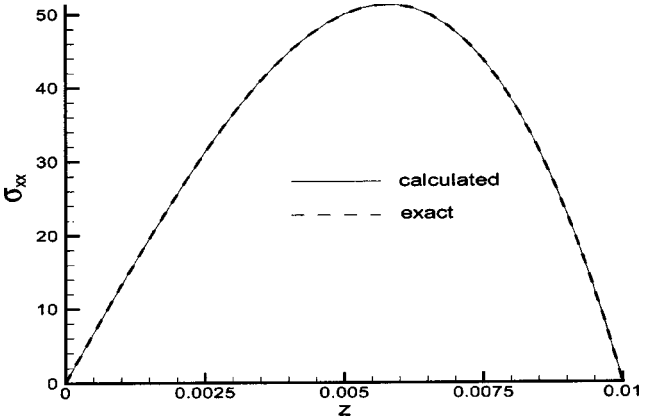


Fig. 4 Exact and calculated results for the in-plane stress for a stress-free strain that is a function of z alone, i.e., $\varepsilon^m = (z/H)^3 - z/H$

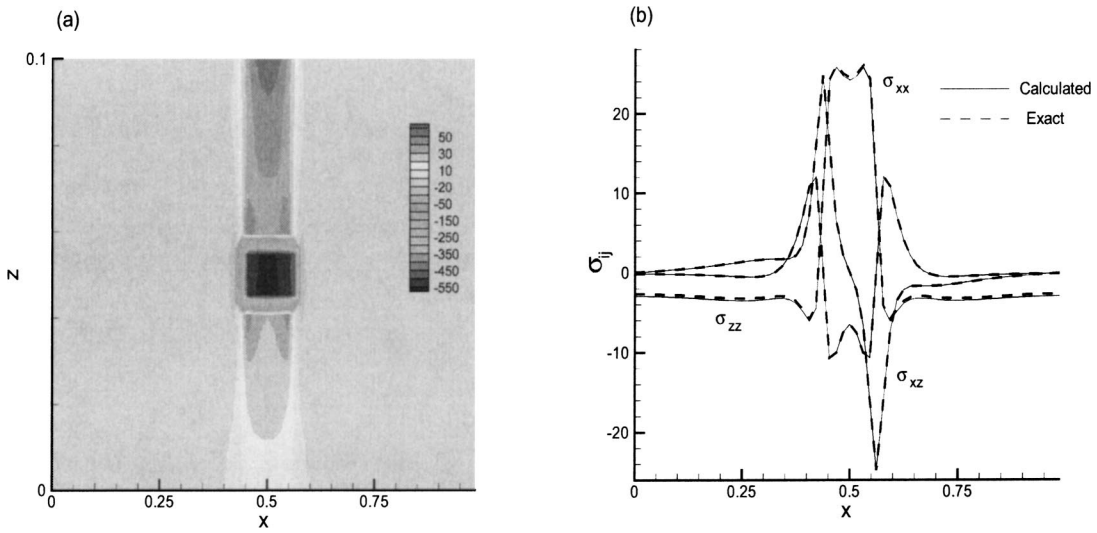


Fig. 5 Stresses due to a thermally mismatched inclusion; (*a*) shows contours of constant σ_{xx} in the film for a rectangular inclusion embedded in the center of the film and (*b*) is a comparison of the stresses obtained using Hu's formulas and those obtained from our calculations for $z=0.025$

Fourier coefficients, for the cases where the functional dependence of ε^m on z is arbitrary or where $\overline{\varepsilon^m}$ can be represented by a Fourier integral.

The Fourier transform approach is particularly attractive as it enables the use of FFT for calculating the particular stresses in addition to the parameters W_1 and W_2. However, the Fourier integral may converge poorly near the edges, i.e., at the film–substrate interface and the film free surface for certain functions $\overline{\varepsilon^m}(z)$, [17]. As we demonstrated for a general polynomial dependence for $\overline{\varepsilon^m}(z)$, the particular solution and the coefficients G_1^i, G_1^H, G_2^i, and G_2^H can be calculated exactly for a wide range of simple functions $\overline{\varepsilon^m}(z)$. When the functional form of $\overline{\varepsilon^m}(z)$ is not known, numerical integration can be employed to obtain the parameters G_1^i, G_1^H, G_2^i, and G_2^H. This method does not suffer from the poor convergence problems of the Fourier transform method described above, however, the computation is somewhat slower. However, both methods are much faster than the finite element techniques that are traditionally employed to solve the equilibrium equations. We envisage these methods to be particularly useful for materials evolution problems, where the stress-free strain evolves along with another function such as the composition or temperature field, and the stresses have to be calculated from the equilibrium equations at each and every time step in the coupled temperature/composition–stress field calculation.

Acknowledgments

This work was supported by NASA Glenn Research Center through agreement NRA-01-GRC-03.

References

[1] Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, McGraw-Hill, New York.

[2] Larche, F., and Cahn, J. W., 1985, "The Interactions of Composition and Stress in Crystalline Solids," Acta Metall., **33**, pp. 333–357.

[3] Downes, J. R., and Faux, D. A., 1997, "The Fourier-Series Method for Calculating Strain Distributions in Two Dimensions," J. Phys.: Condens. Matter, **9**, pp. 4509–4520.

[4] Pickett, G., 1944, "Application of the Fourier Method to the Solution of Certain Boundary Problems in the Theory of Elasticity," ASME J. Appl. Mech., **66**, pp. 176–182.

[5] Faux, D. A., 1994, "The Fourier-Series Method for the Calculation of Strain Relaxation in Strained-Layer Structures," J. Appl. Phys., **75**, pp. 186–192.

[6] Faux, D. A., and Haigh, J., 1990, "Calculation of Strain Distributions at the Edge of Strained-Layer Structures," J. Phys.: Condens. Matter, **2**, pp. 10,289–10,302.

[7] Glas, F., 1987, "Elastic State and Thermodynamical Properties of Inhomogeneous Epitaxial Layers: Application to Immiscible III–V Alloys," J. Appl. Phys., **62**, pp. 3201–3208.

[8] Glas, F., 1991, "Coherent Stress Relaxation in a Half Space: Modulated Layers, Inclusions, Steps and a General Solution," J. Appl. Phys., **70**, pp. 3556–3571.

[9] Fan, Q. H., Fernandes, A., and Periera, E., 1998, "Stress-Relief Behavior in Chemical-Vapor-Deposited Diamond Films," J. Appl. Phys., **84**, pp. 3155–3158.

[10] Dankov, P. D., and Churaev, P. V., 1950, Dokl. Akad. Nauk SSSR, **29**, pp. 529–582.

[11] Hou, P. Y., and Cannon, R. M., 1997, "The Stress State in Thermally Grown NiO Scales," Mater. Sci. Forum, **251–254**, pp. 325–332.

[12] Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," Proc. R. Soc. London, Ser. A, **241**, pp. 376–396.

[13] Head, A. K., 1953, "Edge Dislocations in Inhomogeneous Media," Proc. R. Soc. London, Ser. B, **66**, pp. 793–801.

[14] Malvern, L. E., 1969, *Introduction to the Mechanics of a Continuous Medium*, Prentice-Hall, Englewood Cliffs, NJ.

[15] Hu, S. M., 1989, "Stress From a Parallelepipedic Thermal Inclusion in a Semispace," J. Appl. Phys., **66**, pp. 2741–2743.

[16] Hu, S. M., 1989, "Stress From Isolation Trenches in Silicon Substrates," J. Appl. Phys., **67**, pp. 1092–1101.

[17] Churchill, R. V., 1963, *Fourier Series and Boundary Value Problems*, McGraw-Hill, New York.

Mechanical Response of a Metallic Aortic Stent—Part I: Pressure-Diameter Relationship

R. Wang

K. Ravi-Chandar[1]
Fellow ASME
e-mail: kravi@mail.utexas.edu

Center for Mechanics of Solids,
Structures and Materials,
The University of Texas at Austin,
Austin, TX 78712-1085

The mechanical response of a metallic stent is considered in this series of two papers. In Part I, the development of a test method for the characterization of the mechanical response of a metallic aortic stent subjected to internal or external pressure, and a model that captures the relationship between the pressure and diameter of the stent based on slender rod theory are described. The axial and radial deformation of a bare-metal stent were measured as the stent was subjected to loading ranging from an external pressure of about 80 mm of Hg to an internal pressure of about 160 mm of Hg. The pressure was applied using a polyethylene bag; the method of applying the pressure and measuring the strains was found to provide an accurate determination of the mechanical behavior of the stent. The stent was shown to exhibit two stiff limiting states corresponding to the fully collapsed and fully expanded diameters and an intermediate range between the two where the stiffness was an order of magnitude smaller than the typical stiffness of an aorta. A complete mathematical characterization of the pressure-diameter response of the wire stent was also developed; this model is a straightforward application of the theory of slender rods to the problem of the stent. Excellent agreement with the experimental measurements is indicated, opening the possibility for modeling of the coupled response of the stent and the vessel into which it is inserted. In Part II, we consider the effect of variations of pressure over the length of the stent that introduce changes in the diameter along the length of the stent which leads naturally to the formulation of the coupled problem of the stent within the blood vessel. [DOI: 10.1115/1.1782650]

1 Introduction

Metallic stents and stent grafts are commonly used to treat cardiovascular diseases such as occlusions and aneurysms. These stents are delivered to the desired location with the aid of catheters, and therefore collapsibility prior to deployment is an important factor in the design of the stent. On the other hand, the resistance to collapse under external forces must be high in service, especially in regions of hard occlusions in order for the stent to be effective. These considerations make the design of stents quite complicated, but also quite challenging. Two kinds of stents have been used in clinical practice: a stiff stent expanded into position using a balloon or through shape memory properties of the material, and a more compliant self-expanding stent made of braided wire. The focus in our work is on the self-expanding braided Wallstent®. Such bare metal stents have been inserted into aneurismal vessels, [1–3]; after deployment, a thin tissue layer develops that prevents the blood flow across the wire mesh and thereby starves the aneurysm of blood supply. The stent then experiences the full internal pressure of the blood and hence must possess appropriate stiffness in order to contain the blood pressure. In this case, the stiffness of the stent in relation to the vessel into which it is to be embedded also plays a key role in the stresses that develop in the vessel wall, stresses that can lead to the development of endoleaks or other long term problems with the treatment, [4,5].

Engineering analysis of the design of stents has been discussed in the literature before, some based on approximate models (see, for instance, Ref. [6]) and others based on empirical correlation with experiments, [7]. Rogers et al. [8] and Dumoulin and Cochelin [9] have recently analyzed the plastic deformation response of the stiff stents by finite element analysis. Here we show that a rigorous analysis of the self-expansion of the elastic stent is possible; the analysis is rigorous in the sense of the theory of mechanics of slender rods. A number of experimental studies of the mechanical response of stents can also be found in the literature, [10–14]. However, in most of these studies, the loads on the stents were not applied in a manner consistent with the pressure experienced in vivo. For example, Lossef et al. [11] applied a point load on the stent and measured the resulting change in the stent diameter to obtain a force vs diameter relationship. Flueckiger et al. [10] applied point loads and circular loads similar to that used by Fallone et al. [7]. While such measurements might be quite suitable in ranking different stents (see, for example, Dyet et al. [14]), they do not provide the appropriate pressure-diameter relationship that is needed in analysis of the coupled problem of stent-artery deformation for determination of arterial wall stresses, or in the analysis of the fluid flow through the stented artery, [15,16]. Schrader and Beyer [13] developed an interesting apparatus in which the stent was inserted into a rubber tube and then subjected to external pressure. The diameter was measured using an ultrasonic scheme. Reiu et al. [17] have also used a similar apparatus for applying external pressure on coronary stents. While both groups of investigators implemented the apparatus only for external pressure, it is possible to redesign the apparatus for internal pressure. However, a major drawback of the scheme for internal pressure is that the stiffness of the rubber tube must be measured and subtracted from each measurement of the stent. Since the wire stents exhibit a changing stiffness with expansion this method requires a rather complicated nonlinear inverse problem to be solved in order to determine the stiffness of the stent. The present

[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 Applied Mechanics Division, September 13, 2003; final revision, March 24, 2004. Associate Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

work is aimed at providing an accurate measurement of the pressure vs diameter relationship for stents without such complications.

We report on the development of an experimental scheme that can provide an accurate measurement of the mechanical behavior of the stent in a mechanical environment close to that seen in vivo. Although the results presented here are for one type of self-expanding stent, the method is quite general and applicable to all types of stents. We then interpret these measurements in terms of the mechanical response of the stent through a complete model of the behavior of the helical spring. This model is, of course, limited to the particular type of self-expanding helical stent. This model facilitates the formulation and solution of the problem of non-uniform pressure loading along the stent that is caused by the insertion of the stent into the blood vessel; these aspects as well as the inverse design problem—of determining the optimal stent geometry for patient specific data—are described in Part II (Wang and Ravi-Chandar [18]).

The paper is organized as follows: in Section 2, we present the experimental method developed to apply internal and external pressure on the stent; issues related to the application of pressure, as well as the measurement of the response are discussed. A description of the experimental measurements follows in Section 3. Comparison of the stiffness of the stent to the stiffness of arteries is also presented. Recognizing that the deformation of the stent is governed by the theory of slender rods, we develop a mathematically rigorous model that describes the response of the Wallstent in Section 4. Excellent comparison of the results of the model with the experimental measurements is demonstrated. Finally, we close with a discussion of how the experimental method and mathematical model could be used in addressing the problem of design of stents.

2 Experimental Methods

A bare metal Wallstent shown in Fig. 1 was used as a test sample in order to develop the apparatus for evaluating the pressure-diameter relationship under both internal and external pressure. It was made of a wire wrapped into a helix and interwoven into a simple-weave pattern; the wire-wire crossings are not bonded in any manner but held only by friction. This construction allows for easy analysis of the stent as discussed in Section 4. The stent consists of 36 wires each of diameter 170 μm, woven into a simple helical pattern with 18 right-handed and 18 left-handed helices. The nominal diameter of the stent in the zero-pressure state was 20 mm and the length of the cylindrical portion was 80 mm. The exact composition of the material is not known, but for the purpose of the mechanics analysis it was sufficient to know that it was a stainless steel with a modulus of elasticity $E = 200$ GPa. The defining parameters of the stent are listed in Table 1. Near the ends of the stent, the interweaving pattern became loose with handling and the diameter of the stent expanded slightly. In vascular applications, this enlarged segment gets embedded into the vessel wall and is considered to anchor the stent in the proper location. In our experiments, this segment of the stent was outside the tested region; external pressure was applied only over approximately 70 mm of the length of the stent and the length changes were measured only over the central region of 40 mm.

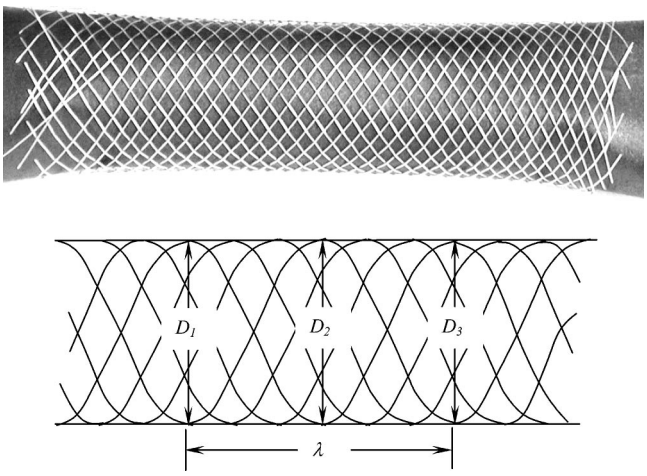


Fig. 1 A photograph and schematic diagram of the bare-metal Wallstent indicating the positions at which the diameter and the length of the stent was measured during the experiments under internal pressure. The stent is shown in the unloaded condition (zero pressure).

Table 1 Parameters of the Wallstent

n	Number of wires	36
E	Modulus of elasticity (assumed)	200 GPa
G	Shear modulus (assumed)	77 GPa
d	Diameter of the stent wire (measured)	170 μm
α_0	Pitch angle of the helix at zero pressure (measured)	34°
r_0	Radius of the stent at zero pressure (measured)	0.01 m
L	Length of the stent (measured)	0.08 m

The range of internal and external pressures considered in the experiments was determined by the fact that the stent has two limiting states: a fully collapsed state and a fully expanded state. Measurements were made over this entire range; the pressure experienced by the stent in vivo is contained within the range covered in these experiments. For the application of internal pressure on the stent, a polyethylene bag was squeezed inside the stent as shown in Fig. 2. In order to eliminate the resistance to inflation from the polyethylene film, many wrinkles were introduced into the film; upon inflation with internal pressure, the bag simply expanded by eliminating the wrinkles without stretching the film. In order to prevent friction between the polyethylene film and the stent as the stent expanded radially and contracted longitudinally, long, thin strips of an acetate sheet were lubricated with grease and inserted longitudinally between the stent and the polyethylene bag. Compressed air was introduced inside the bag through a tube attached with a control valve and a pressure gage. A laboratory compressed air supply line was used as the pressure source; the pressure gage had a sensitivity of ± 0.5 mm of Hg.

The end conditions must be controlled carefully on order to mimic the conditions in vivo. In the scheme shown in Fig. 2(a) the stent is free to expand/contract radially and to contract/expand axially. A variant of this end condition was also used: strings were attached to the wire braid at the two ends of the stent and these strings were collected together and fixed to a rigid post on one end and were taken over a pulley and attached to a weight on the other end. This end condition enabled the stent to expand freely in the radial direction, but its axial shortening was restrained by the weight; this arrangement is shown in Fig. 2(b). In practice, the radial expansion and axial shortening of the stent are constrained by the arterial wall. Our experiments provide an upper bound on the axial force and radial displacements that will be generated in the stent.

The apparatus for application of external pressure consists of a stiff cardboard tube inside which the stent is inserted, with the intervening gap filled by the polyethylene bag (see Fig. 3). Inflation of the bag results in the application of a uniform external

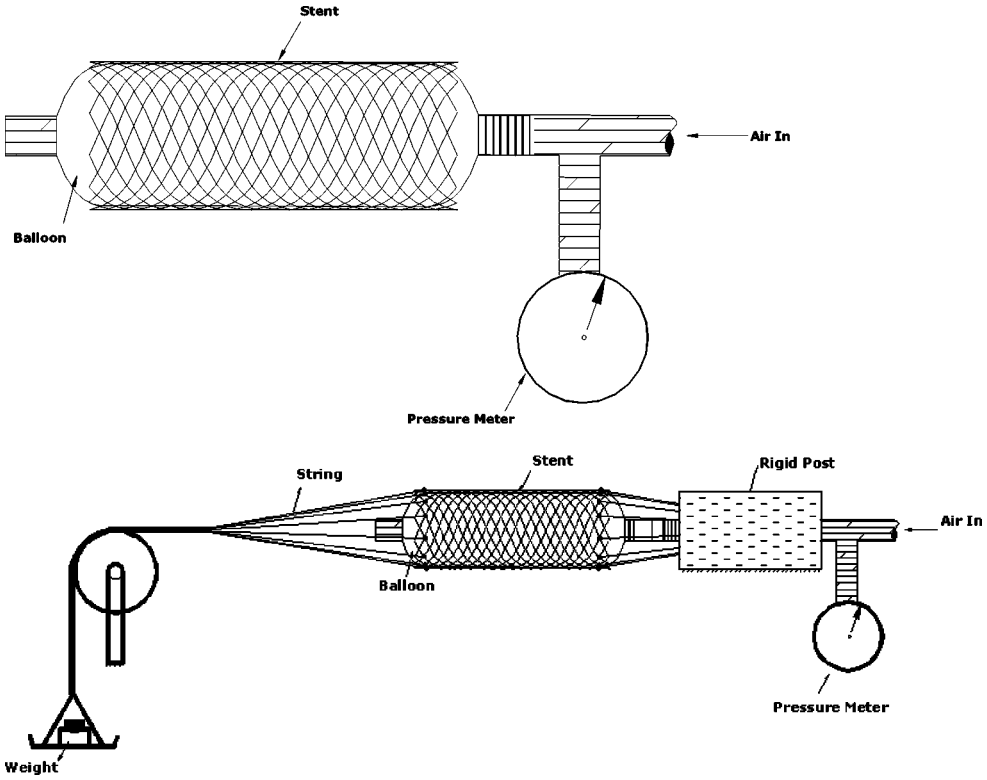


Fig. 2 (*a*) **Schematic representation of a stent under internal pressure produced by a compressed air-filled polyethylene bag.** (*b*) **Same as** (*a*)**, with the ends restrained from axial movement but not radial movement.**

pressure over the middle segment of the stent. Once again, the bag was wrinkled in order to eliminate its stiffness and the lubricated acetate sheets were placed between the polyethylene bag and the stent to minimize friction as in the case of the internal pressure experiment.

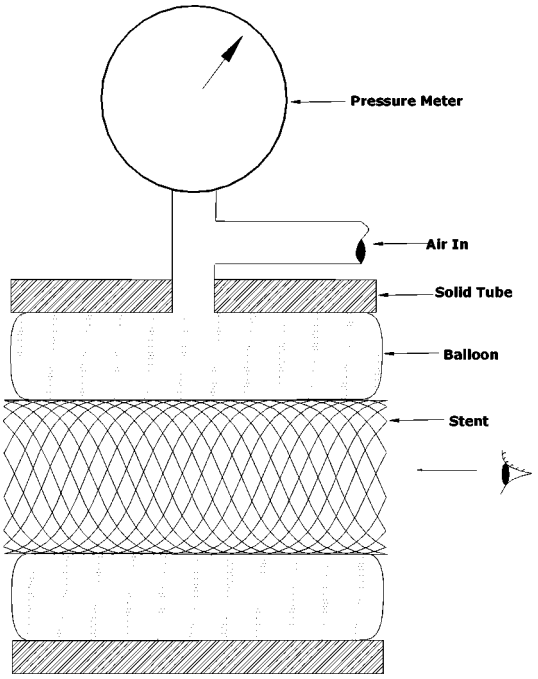


Fig. 3 Schematic representation of a stent under external pressure produced by a compressed air-filled polyethylene bag

In order to determine the pressure-diameter relationship, it is necessary in these experiments to measure the diameter at different pressure levels. This was accomplished in two different ways depending on whether the stent was under internal or external pressure. For the case of internal pressure, the procedure was quite straightforward: the values of the pressure were read from the pressure meter. After the pressure stabilized at each given value for about 3–5 minutes, a picture of the deformed stent was taken using a Nikon CoolPix 950 digital camera with a 1600 $\times$1200 pixel resolution. From measurements of the diameter at three points as indicated in Fig. 1, the average diameter at this pressure level was determined. For the wire stent, the axial contraction corresponding to the radial expansion can be significant. So, the length change was measured between 13 nodes in the middle of the stent, about 42 mm apart initially. Correlating the length and diameter changes with the applied pressure, the mechanical response was obtained. In contrast, under external pressure, the wire mesh grid points were not available for observation; therefore, the diameter changes were determined viewing axially and measuring the visible opening. Length changes could not be monitored, but due to the coupling between the radial and axial deformation as discussed below, the measurements of the radial deformation was used to calculate the axial deformation. The actual dimensional measurements were made in Adobe Photoshop, with the distance between two points obtained from the difference of their coordinates shown by the cursor. With the help of high magnification images, a spatial accuracy of 25 μm was obtained; the accuracy was determined by photographing a standard ruler in the apparatus and measuring the distance between two gage marks.

Coupling between the radial and axial deformation: The coupling between the radial and axial deformation can be determined by assuming that under the small loads imposed on the stent, the

stretching of the wire is negligible. Suppose that the radius of the stent is r and the axial length (pitch of the helix) is λ. Then the total length of the wire l over a single pitch of the helix is

$$l = \sqrt{\lambda^2 + 4\pi^2 r^2}. \quad (1)$$

During deformation of the stent, the total length of the wire remains unchanged and therefore changes in the radius of the stent, r, must be accommodated by changes in the pitch, λ, according to Eq. (1). Instead of using r and λ, one can also use radial and axial strain measures; for the large deformations encountered in our experiments, the true radial strain e_R, and the true axial strain, e_a, defined below were used as the appropriate measures

$$e_R = \ln \frac{r}{r_0}, \quad e_a = \ln \frac{\lambda}{\lambda_0}. \quad (2)$$

Since the length and radius are related as indicated in Eq. (1), the axial strain can be written as

$$e_a = \frac{1}{2} \ln \frac{l^2 - 4\pi^2 r^2}{l^2 - 4\pi^2 r_0^2}. \quad (3)$$

Therefore, from a measurement of the current radius, r, both the radial strain and axial strain can be calculated from Eqs. (2) and (3). In the results described below, we demonstrate that this procedure works well when compared with actual experiments under internal pressure and then use it to calculate the axial deformation under external pressure.

3 Experimental Results

First, we show that the method works well in evaluating the pressure-diameter relationship in stents. In Fig. 4, the measured variation of the true radial strain with the pressure from five repeated trials is shown. A trendline is shown to act merely as a guide for the nonlinear response of the stent; the actual form of the relationship will be derived later. The scatter in the data between the repeated trials is indicative of the errors in the experimental measurement. The scatter arises primarily from the control and measurement of pressure in the experiment; the data quality could be improved significantly through the use of an automated pressure regulator and monitor. Measurements were also taken during the decrease of the pressure from the maximum; the unloading behavior is also shown in Fig. 4 for one of the trials. A significant hysteresis was encountered with the diameter remaining almost constant until the pressure dropped to a very small value and then recovering its original dimensions with a very small change in pressure. We believe that this behavior is due to friction in the system; we have has been able to develop an analytical model of the stent and the loading system and to isolate the source of the friction conclusively to the loading system. Therefore, in the present paper we refrain from discussing the results of unloading, noting that this issue may be quite different in the coupled response of the stent and blood vessel.

In these trials with internal pressure, the axial deformation was also obtained from the photographs. From measurements of the length λ, the true axial strain was calculated as indicated in Eq. (2). In Fig. 5, we show the measured values of the axial strain as a function of the measured values of the radial strain. The axial

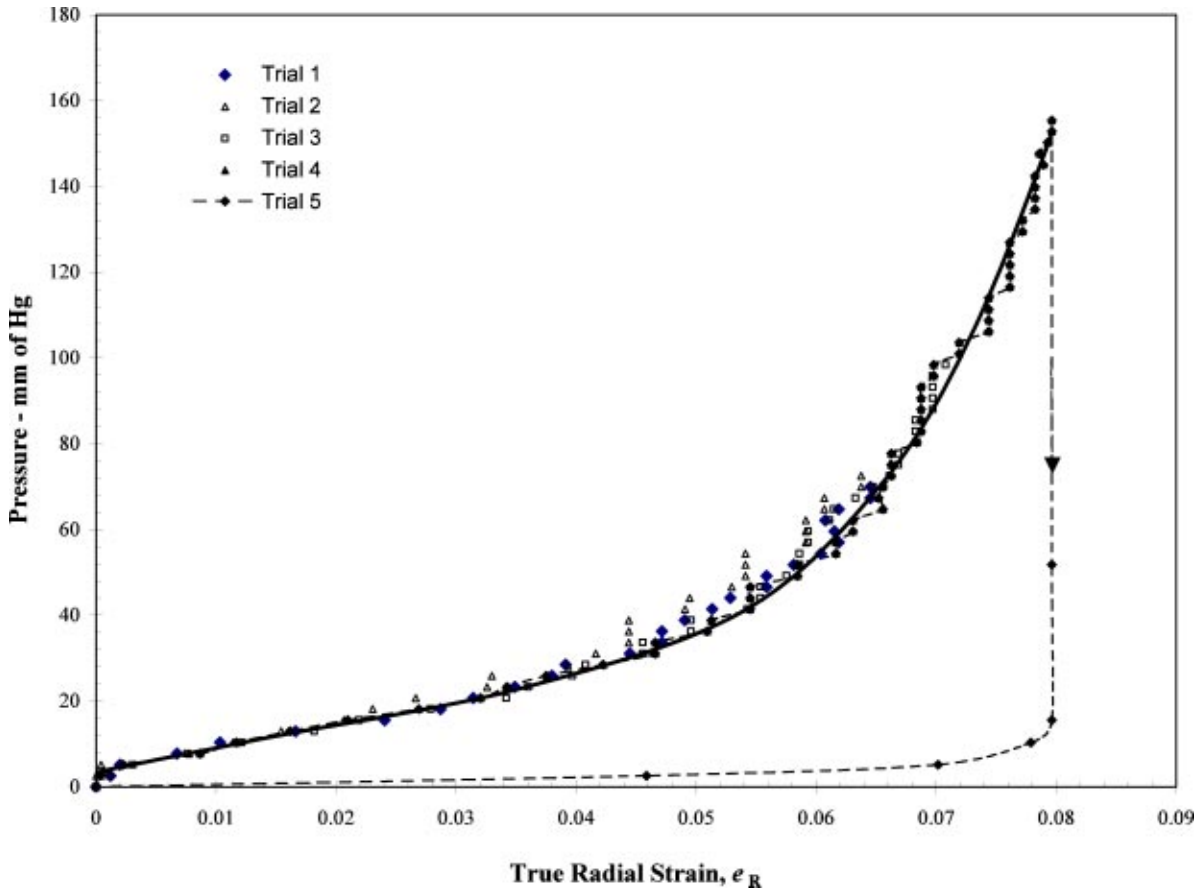


Fig. 4 Variation of the true radial strain, e_R with internal pressure, p. A polynomial curve fit to the experimental data is also shown simply to indicate the data trend. Data scatter is indicative of the errors encountered and is primarily due to the measurement and control of the pressure.

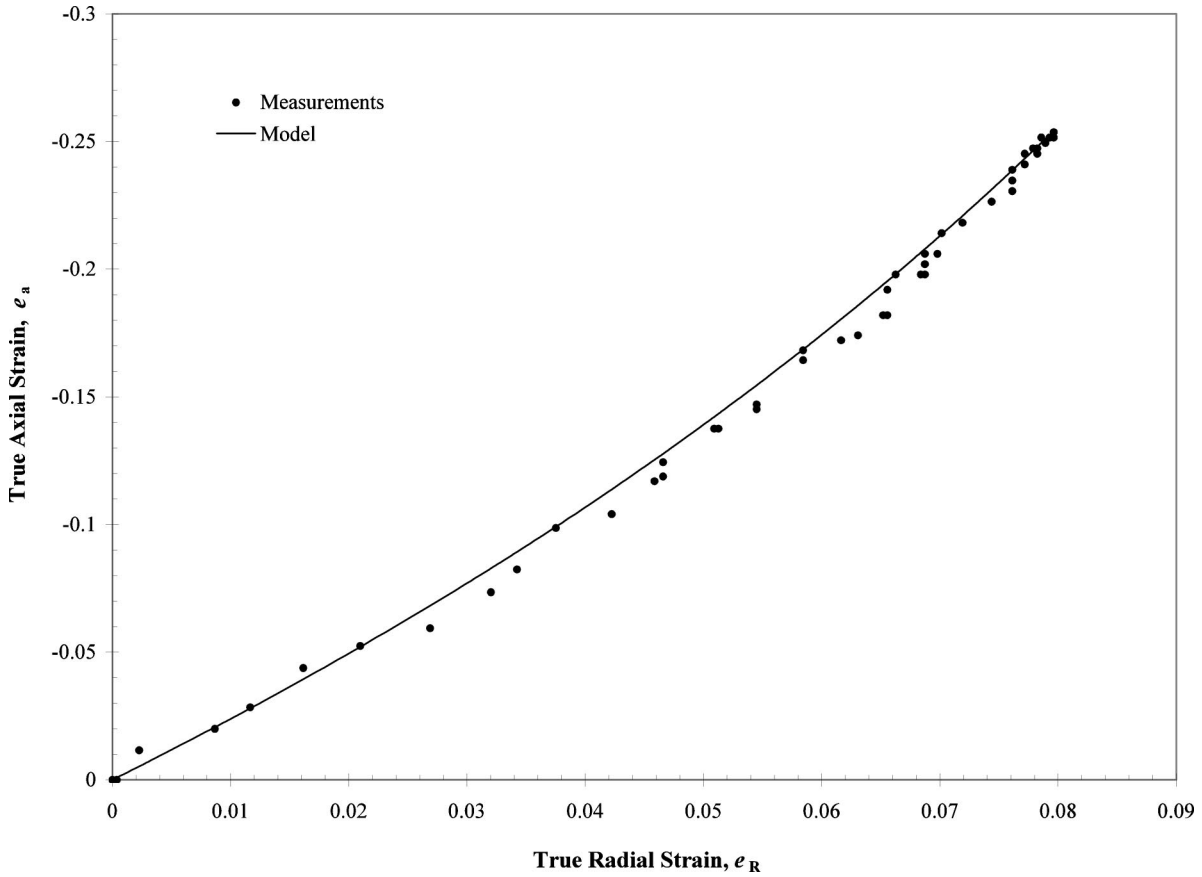


Fig. 5 Variation of the true axial strain with the true radial strain. The data points represent direct measurements of both the diameter and length. The line corresponds to a model of the axial strain from Eq. (4) with measured values of the diameter.

strain calculated from Eq. (3) is also plotted in this figure (identified as ''Model''). The agreement between the two is quite good and we conclude that the coupling between the radial and axial deformation can be calculated using the assumption of inextensibility of the wire under the loads imposed in these experiments. The main purpose of this demonstration is that in the experiments under external pressure, we can determine the axial deformation from Eq. (3) once the diameter has been measured. The nonlinearity of the relationship between the axial and radial strains also indicates that the small deformation approximation is not appropriate for characterizing the behavior of the stent.

The variation of the diameter D with the pressure p over the complete range of pressures used in our experiments—internal as well as external—is shown in Fig. 6. The variation of the length λ with pressure p is shown in Fig. 7. In both of these figures, the results of analytical models are also presented; we will discuss these in the next section. From the experimental results presented here, it is clear that the loading scheme and the measurement scheme described here are capable of evaluating the mechanical response of stents. While our demonstration was with a self-expanding Wallstent, this method will work with other designs of stents as long as the inflation pressures are not much larger than the range used here.

The response of the Wallstent shown in Figs. 6 and 7 suggests that there are three regimes in the response of the stent—two limiting states and an intermediate state connecting the limiting states. The first limiting state corresponds to a fully collapsed condition, similar to what exists when the stent is inserted into the catheter; here the diameter is nearly zero and the length is nearly double the unstressed length. The other limiting state corresponds to the fully expanded condition when the length is a minimum, at the maximum possible diameter. At both these limiting states, the stiffness of the stent is quite high, implying that a large change in pressure is required to cause a small to moderate change in the diameter. Connecting these two limiting states is an intermediate range where the siffness is quite small; here a very small change in the pressure could result in a significant change in the diameter. It is instructive to compare the response of the stent to the native properties of the vessel into which the stents are to be inserted. The stiffness of the vessel wall and the stent can be quantified by the ''pressure-strain modulus,'' E_p, defined as Peterson et al. [19]

$$E_p = \frac{r_1}{r_2 - r_1}\Delta p \tag{4}$$

where Δp is the change in pressure causing radius to change from r_1 to r_2. The smallest stiffness exhibited by the stent is about 8.7 kN/m^2. In comparison, the stiffness for a normal aorta lies in the range of 40 to 140 kN/m^2 and may be as large as 300 kN/m^2 in an aorta susceptible to aneurysm, [20]. The comparatively small stiffness of the stent must be a primary concern in sizing the stent for each insertion. In order to quantify this appropriately, it would be beneficial to have an analytical model of the stent. Therefore, we discuss below an analytical model of the stent covering the entire range of its response based on an analysis of the stent as a helical spring.

4 Analysis of the Response of the Stent

The Wallstent is a set of interwoven helical springs; the analysis of the mechanical response of a spring based on Kirchhoff's theory of slender rods is well established, [21], and thus can be drawn upon for this particular application. However, there are a

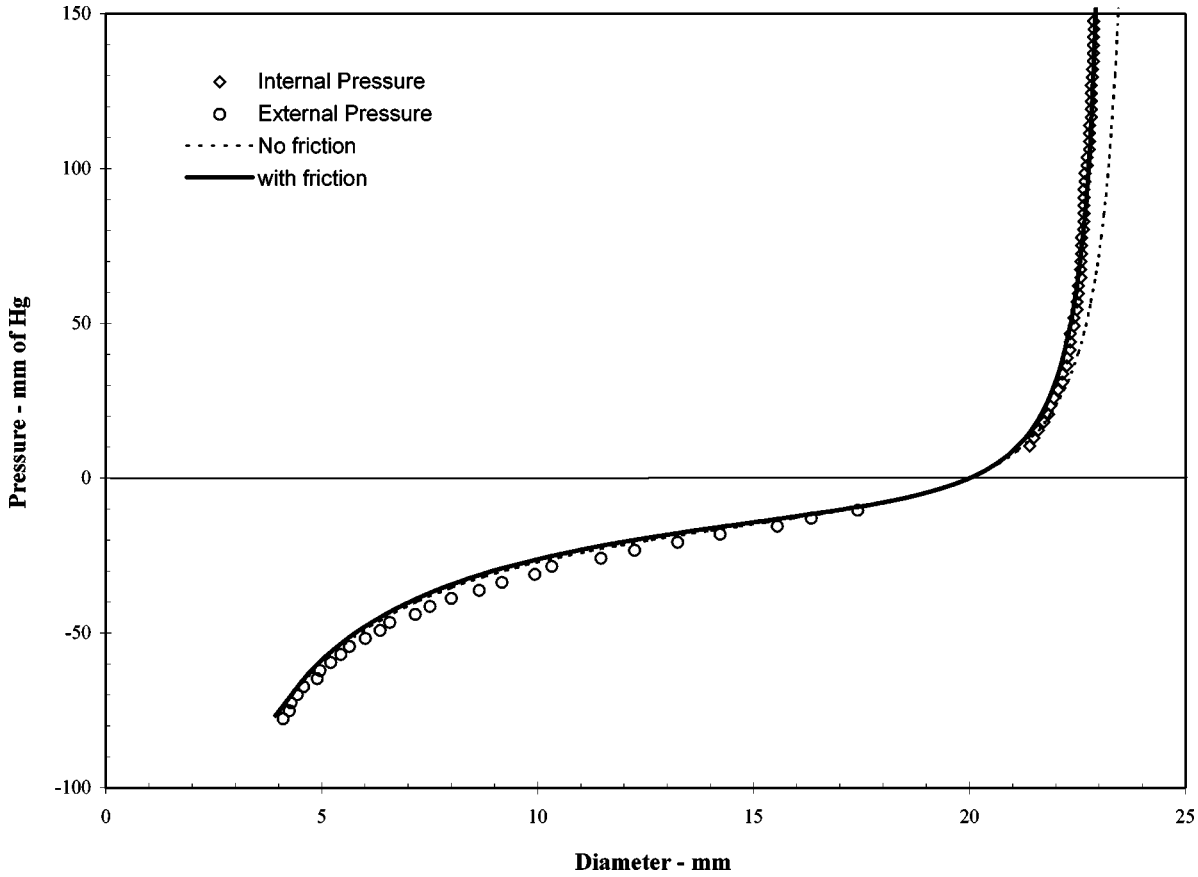


Fig. 6 Variation of the diameter of the stent with pressure. Internal pressure is indicated as positive and external pressure as negative. The lines (-- without friction, ——with friction) are calculated using a helical spring model for the deformation of the stent.

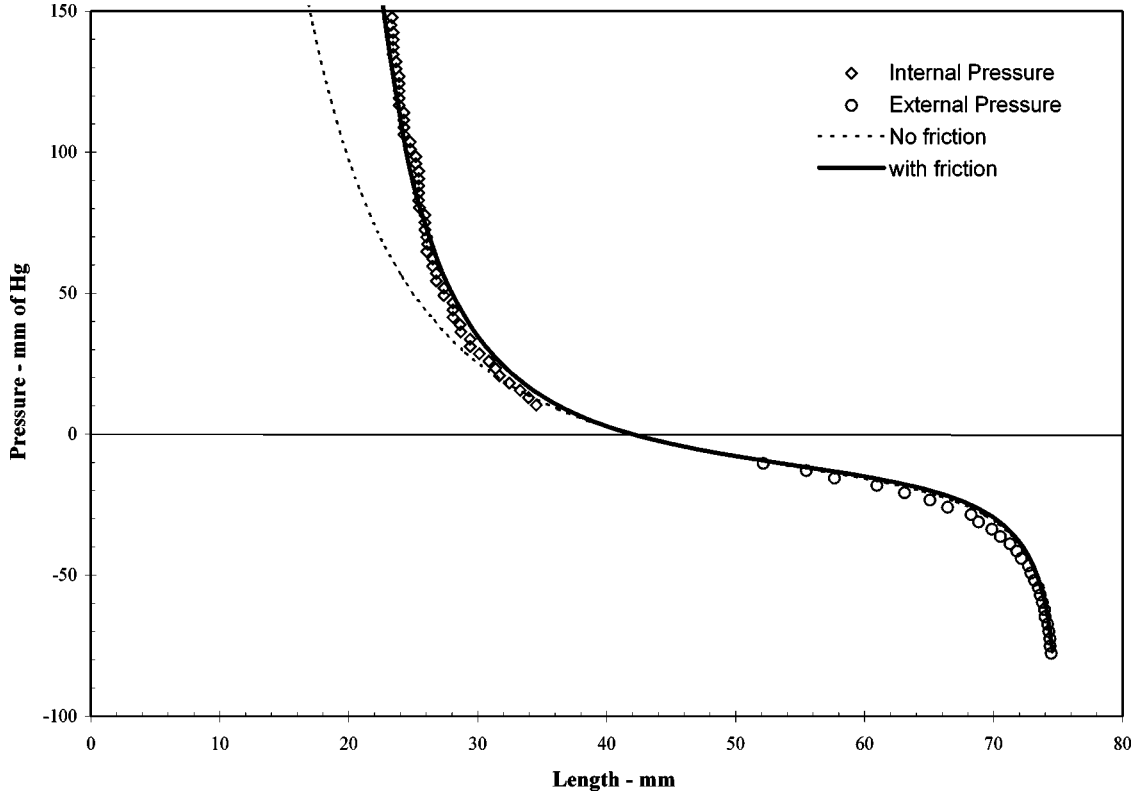


Fig. 7 Variation of the length of the stent with pressure. Internal pressure is indicated as positive and external pressure as negative. The lines (-- without friction, ——with friction) are calculated using a helical spring model for the deformation of the stent.

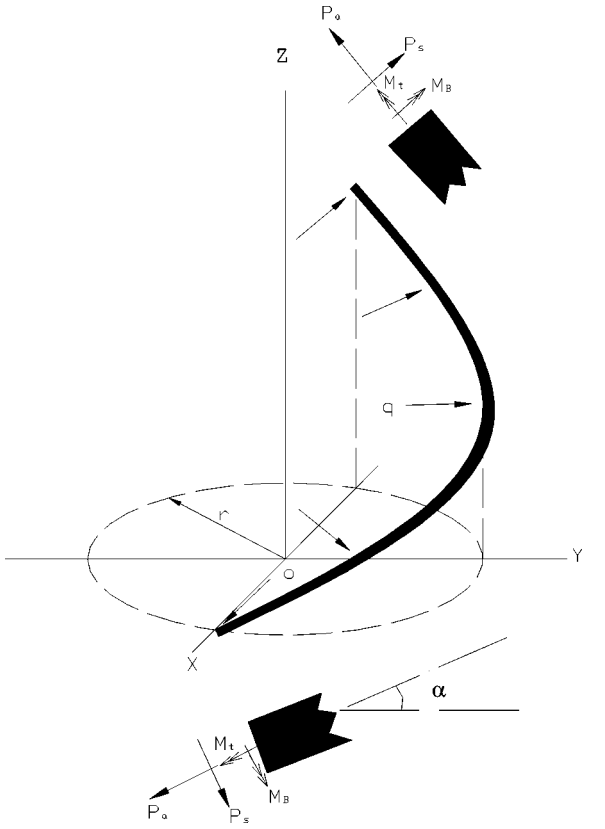


Fig. 8 Free-body diagram on one-half turn of one wire in the stent. Components of forces and moments in the direction of the tangent to the curve and normal to it are shown in the figure. *q* is the load per unit length along the wire that results from the pressure *p* in the stent. *r* is the radius of the helix, and α is the pitch angle.

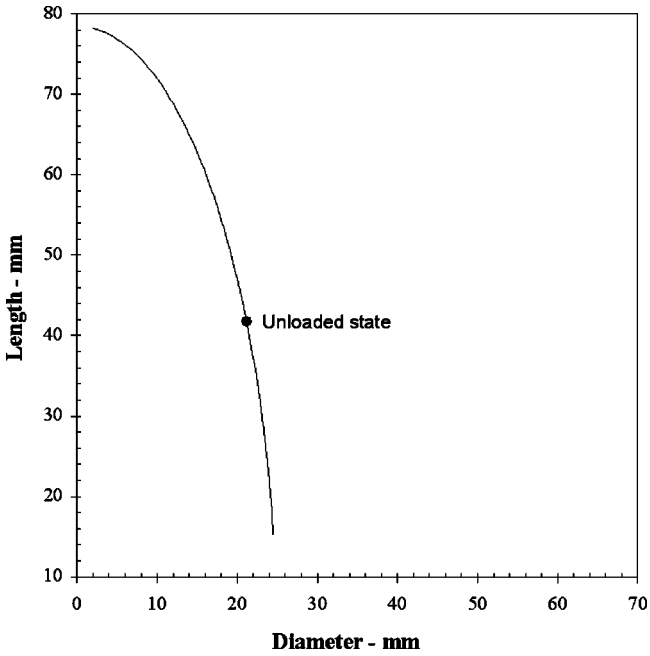


Fig. 9 Correlation of the length and diameter of the stent. This is a plot of Eq. (1); the unloaded state is indicated by the dot.

few significant departures that need attention. First, the loading on the stent is not along the axis as in a spring but on the lateral sides where it contacts the tissue or the polyethylene bag in the test setup. This difference is easily taken into account by considering the equilibrium of a spring under pressure on the lateral sides. Second, since there are many springs coiled together, there is likely to be frictional resistance—of as yet undetermined magnitude—to the sliding of the wires past each other. In fact, this frictional contact dictates that the length of the coil over each pitch must remain constant as the pitch and the pitch angle change due to deformation of the stent. Enforcing this condition results in a geometric constraint: $r/\cos\alpha = r_0/\cos\alpha_0$. Lastly, coiled springs are usually plastically deformed into the helical shape; the Wallstent is simply woven from straight filaments and the wires are not plastically deformed—upon removal of a filament from the stent, it straightens out by itself due to its elasticity, except at the braid crossover points. This has no impact on the deformation itself, but will affect stability of the deformation. In this section, we will first summarize the spring analysis, then show how this analysis can be adapted to the stent, and finally compare the predictions of the model to the measurements.

Let us denote the initial radius of the helix by r_0 and the initial pitch λ_0. The initial principal curvature and the initial twist of the spring are given by

$$\kappa_0 = \frac{\cos^2\alpha_0}{r_0}, \quad \tau_0 = \frac{\sin\alpha_0\cos\alpha_0}{r_0} \tag{5}$$

where $\alpha_0 = \arctan(\lambda_0/2\pi r_0)$ is the angle of the helix. Under the action of the axial force and the pressure p, the helix changes to a new radius, r, and pitch, λ. The corresponding curvature and twist, denoted by κ and τ, are given by

$$\kappa = \frac{\cos^2\alpha}{r}, \quad \tau = \frac{\sin\alpha\cos\alpha}{r} \tag{6}$$

where $\alpha = \arctan(\lambda/2\pi r)$ is the current angle of the helix. Consider the equilibrium of a one-half turn section of one wire in the stent as shown in Fig. 8. Let P_a, P_s, M_B, and M_t be the axial force, shear force, bending moment, and twisting moment, respectively. We note that these are components resolved in the directions of the tangent and normal to the helix. The pressure p (internal or external) in the stent is assigned to this segment of the wire as an equivalent load distributed per unit length and denoted by q. It is calculated as follows: the pressure acts on an effective area $2\pi r\lambda$; this force, $2\pi r\lambda p$, is carried in the n wires that make up the stent, each of length l. Therefore, the force per unit length that each wire experiences is then

$$q = \frac{2\pi r\lambda}{nl}p = \frac{2\pi r\sin\alpha}{n}p \tag{7}$$

where $\sin\alpha = \lambda/l$. In addition, let F_z denote the external force along the axis of the helix that is supplied by the end constraints; we note that this is zero if the stent is free to expand/contract axially. Introducing a nonzero F_z will allow us to calculate the axial force necessary if the stent is to be maintained at some final length due to end constraints. Equilibrium of forces and moments result in the following equations

$$P_a\cos\alpha - P_s\sin\alpha - \frac{qr}{\cos\alpha} = 0 \tag{8}$$

$$P_a\sin\alpha + P_s\cos\alpha = F_z \tag{9}$$

$$r(P_a\sin\alpha + P_s\cos\alpha) + M_B\sin\alpha - M_t\cos\alpha - \frac{qr^2\sin\alpha}{\cos^2\alpha} = 0. \tag{10}$$

The wires are considered to be slender and therefore their response to the bending and twisting moments are given by the Bernoulli-Euler beam theory and Coulomb torsion theory:

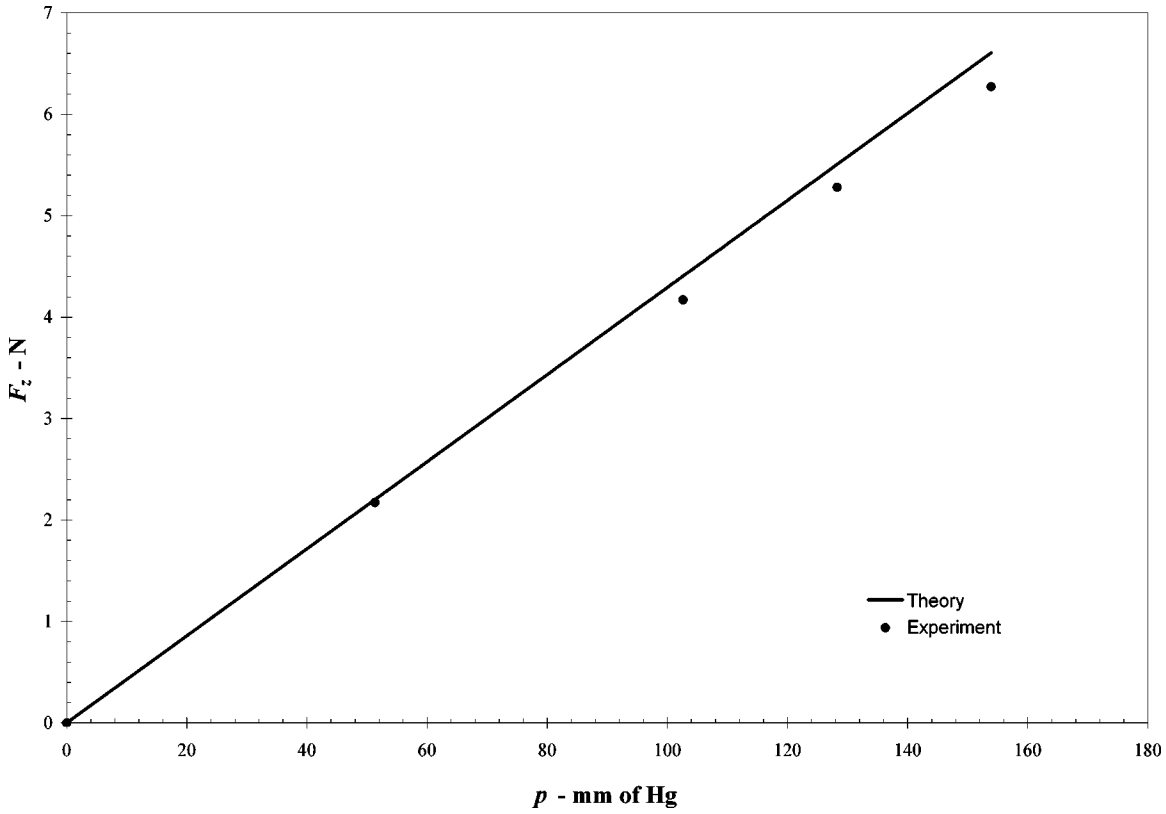


Fig. 10 Comparison of the calculated and measured axial force F_z as a function of the internal pressure, p, when the stent is free to expand radially, but is constrained to maintain its axial length

$$M_B = EI(\kappa - \kappa_0) = EI\left(\frac{\cos^2\alpha}{r} - \frac{\cos^2\alpha_0}{r_0}\right) \tag{11}$$

$$M_t = GI_p(\tau - \tau_0) = GI_p\left(\frac{\cos\alpha\sin\alpha}{r} - \frac{\cos\alpha_0\sin\alpha_0}{r_0}\right) \tag{12}$$

where E is the modulus of elasticity, G is the shear modulus, I is the second moment of the cross-sectional area and I_p is the second polar moment of the cross sectional area. Here, we consider wires of circular cross section and hence, $I_p = 2I = \pi d^4/32$ where d is the diameter of the wire. Eliminating P_a, P_s, M_B, and M_t between Eqs. (8)–(12) results in an expression for the axial force F_z

$$F_z = \frac{2\pi r^2\sin^2\alpha}{n\cos^2\alpha}p - \frac{EI\sin\alpha}{r}\left(\frac{\cos^2\alpha}{r} - \frac{\cos^2\alpha_0}{r_0}\right) + \frac{GI_p\cos\alpha}{r}\left(\frac{\cos\alpha\sin\alpha}{r} - \frac{\cos\alpha_0\sin\alpha_0}{r_0}\right). \tag{13}$$

If the applied axial force $F_z = 0$, then the relationship between the pressure and radius of the stent can be written as

$$p = \frac{n\cos^2\alpha}{2\pi r^2\sin^2\alpha}\left[\frac{EI\sin\alpha}{r}\left(\frac{\cos^2\alpha}{r} - \frac{\cos^2\alpha_0}{r_0}\right) - \frac{GI_p\cos\alpha}{r}\left(\frac{\cos\alpha\sin\alpha}{r} - \frac{\cos\alpha_0\sin\alpha_0}{r_0}\right)\right]. \tag{14}$$

As discussed above, we assume that the pitch and diameter change in such as way as to constrain the coil from unwinding upon deformation; this condition is expressed as: $r/\cos\alpha = r_0/\cos\alpha_0$.

The changes in the diameter as a function of pressure can be determined from Eq. (14); then the changes in the length of the stent (pitch) can be calculated using the assumption of inextensibility of the wire as in Eq. (1). Note that there is an explicit dependence on the radius in this equation as well as an implicit dependence through the dependence on r of α. The predictions of the model are shown in Figs. 6 and 7 by the dotted lines, identified with the label "no friction." In keeping with the experimental boundary condition, F_z was set equal to zero. Clearly, the spring model captures the essence of the behavior observed in the experiments. It must be borne in mind that there are *no adjustable parameters or empirical constants* that have been imposed on the model; the model is an exact implementation of spring theory. Deviations between the predictions and the experimental measurements are observed only under conditions of high internal pressure and can be attributed to friction effects that were neglected in this simple analysis. It is possible to obtain a very simple estimate of the effect of friction on the experimental measurements as described in the next paragraph.

Along the cylindrical surface, the wires of the stent are in contact with the polyethylene bag or the lubricated acetate sheet and must move in order to shorten or elongate as required by the applied pressure. If we assume that the frictional effects reduce the effective axial force that the stent experiences, we can modify Eq. (14) appropriately. Assume that the frictional force, F_f, opposing the axial elongation or contraction of the stent is proportional to the normal force; then $F_f = \mu Ap$, where μ is the (unknown) friction coefficient, $A = 2\pi r\lambda$ is the nominal area of contact, and p is the applied pressure. Decreasing the axial force by this amount results in the following relationship between the applied pressure and the radius of the stent

$$p=\frac{n\cos^3\alpha_0}{2\pi r_0(r_0^2\sin^2\alpha-\mu n\lambda r\cos^2\alpha_0)}\left[EI\frac{\sin\alpha}{r}(\cos\alpha-\cos\alpha_0)\right.$$
$$\left.-GI_p\frac{\cos\alpha}{r}(\sin\alpha-\sin\alpha_0)\right]. \quad (15)$$

Note that when $\mu=0$, this equation reduces to Eq. (14). The prediction of the pressure-diameter relationship from Eq. (15) is also shown in Figs. 6 and 7 and identified by the label ''with friction.'' A very small friction coefficient ($\mu=0.000001$) was sufficient to make the model predictions agree with the experimental measurements. We note that the small value of μ arises from the fact that the actual contact area is significantly smaller than what we have assumed. From these figures, it is evident that friction plays a significant role only when the system is under a large internal pressure, but more importantly that it can be taken into account by a proper model.

An important point to note in the behavior of the stent is the coupling between the axial and radial deformations as shown in Fig. 9. This figure is really a plot of Eq. (1) indicating the influence of the inextensibility of the helix. The unloaded state of the stent is identified in the figure by a dot. Clearly, a small change in the diameter must be accompanied by a large change in the axial length of the stent. As the diameter changes one unit due to the pulsatile pressure cycle, the length must alter by almost seven units. If the stent is simply riding inside the vessel, this axial deformation will result in the scraping of the inside of the vessel. On the other hand, if the stent is firmly anchored in the blood vessel, this axial deformation will be constrained and result in a large force at the points where the stent anchors into the tissue. The magnitude of this force can be determined from Eq. (13) by setting $r=r_0$ and evaluating F_z as a function of p. The axial force necessary to keep the overall length of the stent fixed as a function of internal pressure was measured using the variant of the end condition described earlier and is compared with the theoretical prediction in Fig. 10. If the stent is attached at the ends to a blood vessel, this force will be exerted by the stent on the vessel. Clearly, in the range of physiological pressures, the axial force can be quite large. We believe that the coupling of radial and axial deformations is an important consideration in the generation of endoleaks in the stents and for the conjectured remodeling of the stent in the vicinity of the anchoring site, [22]; the correspondence of this stress to the dilation and endoleaks needs to be examined further for this as well as other stent designs.

5 Conclusion

An experimental method for the evaluation of the mechanical response of stents has been developed; this method allows for both internal and external pressure to be applied on the stent and furthermore allows the measurements to be directly interpreted in terms of the pressure-diameter relationship of the stent. We note that the method was demonstrated with a Wallstent, but is applicable to all stent constructions. For the Wallstent, an analytical model based on the theory of slender rods was developed; the predictions of the model were compared with the experiments to show excellent agreement. The model is appropriate over the entire range of deformation of the stent—from fully collapsed to fully released. The observed response of the stent shows that the stiffness of the stent varies dramatically with its radius and therefore when confined inside an artery, the wall stress in the artery due to the mismatch or ''overexpansion'' could be quite large; calculation of this stress requires a model for the deformation of the artery and a model for the bending response of the stent. The latter is addressed in Part II with a beam-on-elastic foundation model.

Acknowledgment

We acknowledge Prof. Suncica Canic, Department of Mathematics, University of Houston for bringing this problem to our attention and for many discussions on the fluid-structure interaction aspects of the problem. We also acknowledge Dr. Krajcer of the Texas Heart Institute at St. Luke's Episcopal Hospital for providing us with the stents used in this investigation and for discussions on the clinical aspects of the problem.

References

[1] Ruiz, C. E., Zhang, H. P., Douglas, J. T., Zuppan, C. W., and Kean, C. J. C., 1995, ''A Novel Method for Treatment of Abdominal Aortic Aneurysms Using Percutaneous Implantation of a Newly Designed Endovascular Device,'' Circulation, **91**, pp. 2470–2477.

[2] Ruiz, C. E., Zhang, H. P., Butt, A. I., and Whittaker, P., 1997, ''Percutaneous Treatment of Abdominal Aortic Aneurysm in a Swine Model: Understanding the Behavior of Aortic Aneurysm Closure Through a Serial Histopathological Analysis,'' Circulation, **96**, pp. 2438–2448.

[3] Villareal, R. O., Howell, M. H., and Krajcer, Z., 2000, ''Regression of Inflammatory Abdominal Aortic Aneurysm After Endoluminal Treatment With Bare-Metal Wallstent® Endoprostheses,'' Tex Heart Inst. J., **27**, pp. 146–149.

[4] Chuter, T. A. et al., 2001, ''Endoleak After Endovascular Repair of Abdominal Aortic Aneurysm,'' J. Vasc. Surg., **34**, pp. 98–105.

[5] Bell, P. R. F., 2002, editorial, ''Endovascular Repair of Abdominal Aortic Aneurysms,'' Vasc. Med., **7**, pp. 253–255.

[6] Loshakove, A., and Azhari, H., 1997, ''Mathematical Formulation for Computing the Performance of Self-Expanding Helical Stents,'' Int. J. Med. Inf., **44**, pp. 127–133.

[7] Fallone, B. G., Wallace, S., and Gianturco, C., 1988, ''Elastic Characteristics of the Self-Expanding Metallic Stents,'' Invest. Radiol., **23**, pp. 370–376.

[8] Rogers, C., Tseng, D. Y., Squire, J. C., and Edelman, E. R., 1999, ''Baloon-Artery Interactions During Stent Placement,'' Circ. Res., **84**, pp. 37–383.

[9] Dumoulin, C., and Cochelin, B., 2000, ''Mechanical Behavior Modeling of Balloon-Expandable Stents,'' J. Biomech., **33**, pp. 1461–1470.

[10] Flueckiger, F., Sternthal, H., Klein, G. E., Aschauer, M., Szolar, D., and Kleinhappl, G., 1994, ''Strength, Elasticity, Plasticity of Expandable Metal Stents: In Vitro Studies With Three Types of Stress,'' J. Vasc. Interv Radiol., **5**, pp. 745–750.

[11] Lossef, S., Lutz, R., Mundorf, J., and Barth, K., 1994, ''Comparison of Mechanical Deformation Properties of Metallic Stents With Use of Stress-Strain Analysis,'' J. Vasc. Interv Radiol., **5**, pp. 341–349.

[12] Berry, J. L., Newman, V. D., Ferraria, C. M., Routh, W. D., and Dean, R. H., 1996, ''A Method to Evaluate the Elastic Behavior of Vascular Stents,'' J. Vasc. Interv Radiol., **7**, pp. 381–385.

[13] Schrader, S. C., and Bear, R., 1998, ''Evaluation of the Compressive Mechanical Properties of Endoluminal Metal Stents,'' Cathet Cardiovasc. Diagn., **44**, pp. 179–187.

[14] Dyet, J. F., Watts, W. G., Ettles, D. F., and Nicholson, A. A., 2000, ''Mechanical Properties of Metallic Stents: How do These Properties Influence the Choice of Stent for Specific Lesions?,'' Cardiovasc. Intervent Radiol., **23**, pp. 47–54.

[15] Taylor, C. A., Hughes, T. J. R., and Zarins, C. K., 1998, ''Finite Element Modeling of Blood Flow in Arteries,'' Comput. Methods Appl. Mech. Eng., **158**, pp. 155–196.

[16] Canic, S., 2002, ''Blood Flow Through Compliant Vessels After Endovascular Repair: Wall Deformations Induced by Discontinuous Wall Properties,'' Comput. Visual. Sci., **4**, pp. 147–155.

[17] Rieu, R., Barragan, P., Masson, C., Fuseri, J., Garitey, V., Silvestri, M., Roqubert, P., and Sainsous, J., 1999, ''Radial Force of Coronary Stents: A Comparative Analysis,'' Cath. Cardiovas. Interven., **46**, pp. 380–391.

[18] Wang, R., and Ravi-Chandar, K., 2004, ''Mechanical Response of a Metallic Aortic Stent: II. A Beam-on-Elastic Foundation Model,'' ASME J. Appl. Mech., **71**, pp. 706–712.

[19] Peterson, L. H., Jensen, R. E., and Parnel, J., 1960, ''Mechanical Properties of Arteries in Vivo,'' Circ. Res., **8**, pp. 622–639.

[20] MacSweeney, S. T. R., Young, G., Greenhalgh, R. M., and Powel, J. T., 1992, ''Mechanical Properties of the Aneurysmal Aorta,'' Br. J. Surg., **79**, pp. 1281–1284.

[21] Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, 4th Ed., Dover Publications, New York, p. 415.

[22] Wever, J. J. et al., 2000, ''Dilatation of the Proximal Neck of Infrarenal Aortic Aneurysms After AAA Repair,'' Eur. J. Vasc. Endovasc Surg., **19**, pp. 197–201.

Mechanical Response of a Metallic Aortic Stent—Part II: A Beam-on-Elastic Foundation Model

R. Wang

K. Ravi-Chandar
Fellow ASME
e-mail: kravi@mail.utexas.edu

Center for Mechanics of Solids, Structures and Materials,
The University of Texas at Austin,
Austin, TX 78712-1085

The main objective of the paper is to develop the mathematical analysis of the response of a metallic stent subject to axisymmetric loads over its length and to different boundary conditions. These situations introduce bending stresses in the stent and cannot be captured by a model of the stent that can be used to characterize the pressure-diameter relationship under axially uniform loading. The analysis presented here is based on an analogy between a thin-walled pressure vessel and a beam on elastic foundation; in the present application, we derive an equivalent beam model for the bending response of a stent. Using this model, we evaluate the shape of the stent exiting the catheter as well as the variation of the diameter along the length of the stent constrained by stiff end supports. This approach can be used to evaluate the coupled response of the stent and the blood vessel, if the mechanical properties of the blood vessel are known. The coupled problem and its implications in the design of stents are discussed.
[DOI: 10.1115/1.1782912]

1 Introduction

In Part I, [1], we examined the response of a woven metallic stent to internal and external pressure. A mathematical model of the pressure-diameter relationship was obtained based on the theory of slender rods and demonstrated to predict the response accurately. An experimental method was also developed for determining the pressure diameter relationship through appropriate tests. In the model as well as in the experiment, pressure was applied uniformly over the length of the stent. However, in vivo, the stent experiences an axially varying pressure due to interactions with the vessel wall and hard occlusions or aneurysms. For example, as shown schematically in Fig. 1, a stent inserted into a vessel with an aneurysm will experience external pressure in the intact regions of the vessel wall towards the ends of the aneurysm, but experience internal pressure in the aneurysm (after a sealing layer has formed). While the response determined from uniform pressure along the axis of the stent is essential in characterizing the stent, further analysis is required to determine the response of the stent under the axially varying pressures described above. In terms of the mechanics of the problem, uniform pressure along the axis corresponds to a ''membrane theory'' of deformation while the spatially varying pressure corresponds to a ''bending theory'' of deformation. While there are a large number of evaluations of the pressure-diameter relationship of stents (see Part I for a review of this work), to our knowledge the problem of bending due to axial variations in the loading has not been modeled or experimentally measured by any of the investigators. Ormiston et al. [2] used the three-point bend test to determine the overall longitudinal flexibility of different stents; the objective was simply to provide a quantitative comparison of the flexibility of stents. Others have used finite element analysis to determine the stresses generated in the stents as a result of deployment or flexing, [3,4]. In this paper, our goal is to consider the influence of variations in the loads along the length of the stent on the deformation of the stent; this will enable determination of the shape of the stent as it exits the catheter, as well as allow the formulation of the problem of coupled response of the stent and the artery.

This paper is organized as follows: in Section 2 we describe the mechanical basis of the analogy between the stent deformation and a beam-on elastic foundation; then we formulate the equation governing the deflection of the stent. The numerical solution procedure adopted to solve the differential equation is described in Section 3. Then, in Section 4, we describe how the model can be used to determine the shape of the stent (i) as it is pressurized with stiff end constraints and (ii) as it exits the catheter; in both cases good comparison to experiments are demonstrated. The purpose of these demonstrations is to really show that the model captures the mechanics of the stent response. In Section 5, the problem of evaluation of the coupled response of the stent and the artery is formulated and the design issues that arise are discussed.

2 Analysis of the Response of the Stent—Beam-on Elastic Foundation Model

The basic idea behind the analysis is the following: the stent can be considered to be a thin-walled pressure vessel, albeit with a nonlinear pressure-diameter relationship, $p=f(r)$ that is characteristic of the particular stent under consideration. The response of a thin-walled pressure vessel to axisymmetric loading that results in bending stresses may be determined by analogy of its response to that of an equivalent beam on an elastic foundation, [5]. This equivalence is demonstrated schematically in the meridional section in Fig. 2(*a*). Consider an axisymmetric loading on the stent at some point as indicated by the arrows in the figure; this loading changes the radius of the stent as a function of position along the axis, denoted by $r(x)$. Equivalently, this can be represented by the displacement $v(x)=r(x)-r_0$, where r_0 is the uniform initial radius of the stent without any load. Clearly, this loading introduces a bending component to the problem. A free-body diagram of a segment of the stent is shown in Fig. 2(*b*). At any point x along the stent, a circumferential force F_c is generated because of the variation in the radius. These forces balance out in the circumfer-

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 Applied Mechanics Division, September 13, 2003; final revision, March 24, 2004. Associate Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

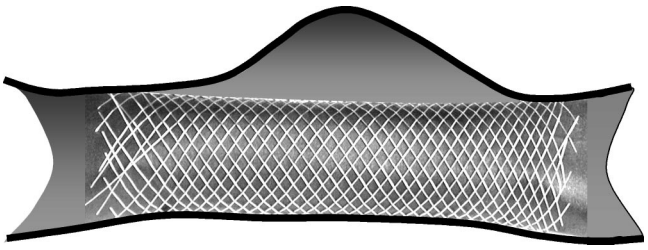


Fig. 1 Schematic diagram of a stent inside a blood vessel with an aneurysm

ential direction, but provide a component in the radial direction that opposes the radial displacement. This is how the stent is able to maintain its radius variation along the length. This force can be considered to be equivalent to the reaction from an elastic foundation. Since the relationship between the pressure and radius is in general nonlinear (see Part I for a particular model), we have a nonlinear elastic foundation. Therefore the differential equation for the bending of the equivalent beam on elastic foundation per unit width can be written as

$$EI_{\text{eff}}\frac{d^4r}{d^4x}+f(r)=p_a \tag{1}$$

where p_a is the applied pressure in the stent and $f(r)$ is the resistance of the equivalent elastic foundation. We note that for a typical thin walled pressure vessel, E is usually replaced by its plane strain equivalent due to the restraint provided against anticlastic curvature; however, in our implementation, since the wires are allowed to act independently, we should not use the plane strain equivalent. I_{eff} is the effective bending rigidity of the stent and is calculated easily. The equation for the calculation of the axisymmetric deformation is applicable for any stent design as as long as I_{eff} and $f(r)$ can be calculated or measured. We demonstrate this for one particular stent, the elastic, self-expanding Wallstent®.

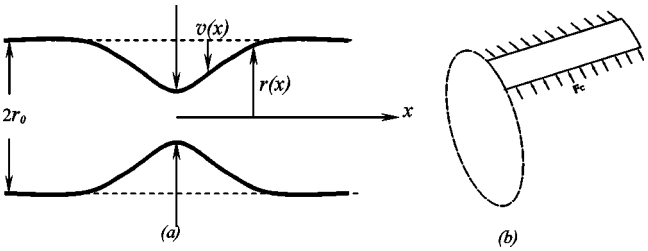


Fig. 2 Stent as an equivalent thin-walled pressure vessel and a beam on elastic foundation

The stent geometry of the woven metallic stent is shown in Fig. 1, inserted schematically into a blood vessel with an aneurysm. The stent consists of 36 wires each of diameter 170 μm, woven into a simple helical pattern with 18 right-handed and 18 left-handed helices; the geometric and material parameters of the stent are given in Table 1 of Part I. The length of the stent is about 80 mm and its diameter is 21 mm. The exact material composition is unknown, but it was sufficient to know that it was a stainless steel with a modulus of elasticity $E=200$ GPa. As can be seen from Fig. 1, the stent will experience support conditions and pressure differential that vary along its length, resulting in a bending of the stent.

The pressure-diameter relationship for the woven metallic Wallstent was derived in Part I (see Eq. (14)) and is reproduced below for convenience.

$$p=\frac{n\cos^2\alpha}{2\pi r^2\sin^2\alpha}\left[\frac{EI\sin\alpha}{r}\left(\frac{\cos^2\alpha}{r}-\frac{\cos^2\alpha_0}{r_0}\right)-\frac{GI_p\cos\alpha}{r}\left(\frac{\cos\alpha\sin\alpha}{r}-\frac{\cos\alpha_0\sin\alpha_0}{r_0}\right)\right]\equiv f(r) \tag{2}$$

where r_0 and λ_0 are the initial radius and pitch of the helix, r and λ are the radius and pitch at a pressure p, $\alpha_0=\arctan(\lambda_0/2\pi r_0)$ and $\alpha=\arctan(\lambda/2\pi r)$ are the initial and current pitch angle of the helix, n is the number of coils in the stent, E is the modulus of elasticity, G is the shear modulus, and I is the second moment, and I_p is the second polar moment of the cross-sectional area of the wire. Here, we consider wires of circular cross section and hence, $I_p=2I=\pi a^4/2$ where a is the radius of the wire.

In Fig. 3(a), the wire mesh of the stent is shown with the ends of the wires exaggerated by the dots; if we view the thin strip in Fig. 2(b) as a beam, per unit length there are $n/2\pi r$ wires. Each wire presents an elliptical cross section with semi-major and semi-minor axes of $a/\sin\alpha$ and a, respectively, dictated by the pitch angle of the helix. The second moment of the beam per unit width is then

$$I_{\text{eff}}=\frac{\pi a^4}{4\sin\alpha}\frac{n}{2\pi r}=\frac{na^4}{8r\sin\alpha}. \tag{3}$$

Substituting for I_{eff} from Eq. (3) and $f(r)$ from Eq. (2) into Eq. (1), the governing differential equation for the radius of the stent under applied pressure p_a and any other axisymmetric loading is obtained.

$$E\frac{na^4}{8r\sin\alpha}\frac{d^4r}{d^4x}+\frac{n\cos^2\alpha}{2\pi r^2\sin^2\alpha}\left[\frac{EI\sin\alpha}{r}\left(\frac{\cos^2\alpha}{r}-\frac{\cos^2\alpha_0}{r_0}\right)-\frac{GIp\cos\alpha}{r}\left(\frac{\sin\alpha\cos\alpha}{r}-\frac{\sin\alpha_0\cos\alpha_0}{r_0}\right)\right]=p_a \tag{4}$$

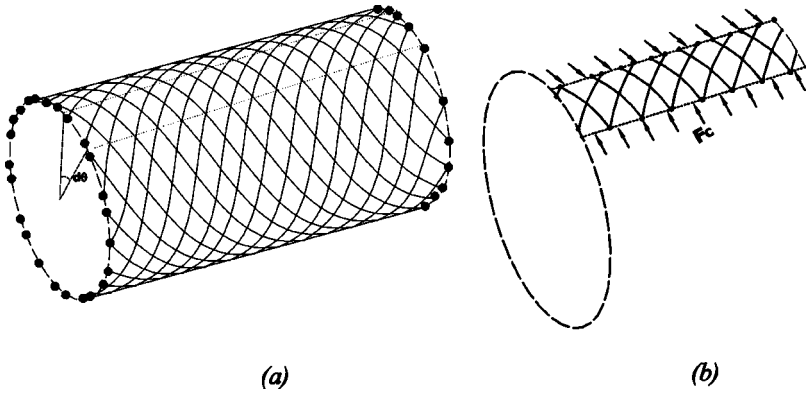


Fig. 3 Effective cross section of the stent that determines the bending rigidity

This equation must be supplemented by appropriate boundary conditions indicating how the stent is supported at the ends. The governing differential equation can be written succinctly by collecting terms

$$\frac{d^4r}{d^4x}=F[(r(x),p(x))] \tag{5}$$

where we have allowed for possible spatial variations of the applied pressure as well. In general, boundary conditions are specified at the support points of the stent. A number of possibilities exist, depending on the problem of interest. We will consider three possible boundary conditions—fixed, free, and compliant.

Case (i) At a *fixed* point on the beam, the appropriate boundary conditions are

$$r=R, \quad \frac{dr}{dx}=0 \tag{6}$$

where R is a prescribed value of the radius. Note that a fixed point need not be at an end of the beam, but can be anywhere along it as long as external constraints are applied to maintain the condition required by Eq. (6).

Case (ii) At a *free end* point on the beam, the appropriate boundary conditions are that the forces and moments are zero; therefore,

$$\frac{d^3r}{dx^3}=0, \quad \frac{d^2r}{dx^2}=0. \tag{7}$$

Case (iii) A stent supported by a blood vessel will experience a compliant boundary condition where the deflection and slope of the stent should match those of the blood vessel. The boundary conditions must be posed as a matching condition. The appropriate conditions are

$$r_{\text{stent}}=r_{\text{vessel}}, \quad \left.\frac{dr}{dx}\right|_{\text{stent}}=\left.\frac{dr}{dx}\right|_{\text{vessel}}. \tag{8}$$

Note that the deflection and slope at this point are not prescribed, but obtained as part of a coupled solution of the deformation of the stent and the vessel. Furthermore, the response of the blood vessel must be modeled though an equation equivalent to Eq. (4), but with modifications to take into account the appropriate properties of the blood vessel.

In any application, appropriate boundary conditions are selected from these choices. For example, as a completely collapsed stent is pushed out of the catheter, the boundary condition at one end is free and at the other end is fixed (the radius is fixed at the value of the inside radius of the catheter and the slope is set to zero assuming that the stiffness of the catheter is much larger than that of the stent). If the stent is supported between two rigid cylindrical blocks, as in the laboratory experiment described below, both ends of the stent experience fixed boundary conditions.

It should be noted that the above procedure for determination of the response under internal and external pressure is identical for this braided wire stent. We illustrate the procedure through two examples in the Section 4; in the first example, the shape of a stent exiting a catheter, and hence under nonuniform external pressure, is determined. The second example illustrates the application of internal pressure on a stent fixed rigidly at both ends. It is important to observe that under external pressure there is no buckling as there would be in a thin-walled pipe. This is because the external pressure is translated into a compressive force (P_a in Fig. 8 of Part I) of small magnitude along the axis of the wires. There are global bending and buckling modes for the entire stent, but these are beyond the scope of the present work.

3 Numerical Solution of the Differential Equation

The differential equation governing the bending of the stent under axisymmetric loading (Eq. (5)) is nonlinear and can only be solved numerically. We begin with the explicit Adams-Bashforth method, [6], to obtain the solution. The procedure is described here briefly since the boundary conditions have to be handled carefully. The whole length of the stent is partitioned into N equal segments of length h, with the $N+1$ nodes labeled by $0,1,2 \ldots k-1,\ k,\ k+1 \ldots N$. Thus the nodal positions are $x_k = kh$; nodal values of any quantity at the kth node is denoted with a subscript—for example, $F_k=F[r(x_k),p(x_k)]$. Applying the two-step explicit Adams-Bashforth method successively, we obtain

$$r_k'''=r_0'''+C_k \ \text{ with } \ C_k=\begin{cases} F_0h & k=1 \\ h\sum_{j=0}^{k-1} F_j+\frac{1}{2}(F_{k-1}-F_0)h & k=2\cdots N \end{cases} \tag{9a}$$

$$r_k''=r_0''+khr_0''+D_k, \quad \text{with}$$

$$D_k=\begin{cases} \frac{F_0h^2}{2} & k=1 \\ h\sum_{j=1}^{k-1} C_j+\frac{1}{2}C_{k-1}h & k=2\cdots N \end{cases} \tag{9b}$$

$$r_k'=r_0'+khr_0''+G_kr_0'''h^2+H_k, \ \text{with}$$

$$G_k=\begin{cases} \frac{1}{2} & k=1 \\ \frac{k^2-1}{2} & k=2\cdots N \end{cases} \ \text{and}$$

$$H_k=\begin{cases} \frac{F_0h^3}{6} & k=1 \\ h\sum_{j=1}^{k-1} d_j+\frac{1}{2}D_{k-1}h & k=2\cdots N \end{cases} \tag{9c}$$

$$r_k=\begin{cases} r_0+hr_0'+\frac{1}{2}h^2r_0''+\frac{1}{6}h^3r_0'''+\frac{1}{24}F_0h^4 & k=1 \\ r_0+khr_0'+\frac{k^2-1}{2}h^2r_0''+\left(\sum_{j=1}^{k-1} H_j+\frac{1}{2}H_{k-1}\right)h^3r_0'''+\left(\sum_{j=1}^{k-1} G_j+\frac{1}{2}G_{k-1}\right)h & k=2\cdots N. \end{cases} \tag{9d}$$

Equation (9*d*) is an explicit representation for the radius r_k in terms r_{k-1} for all k from 1 to N. Note that it involves r and the first three derivatives of r at $k=0$. However, as discussed in Section 2, boundary conditions—either r and r' or r'' and r'''—are prescribed at each end. Suppose both ends of the stent are fixed in the radial direction. Then

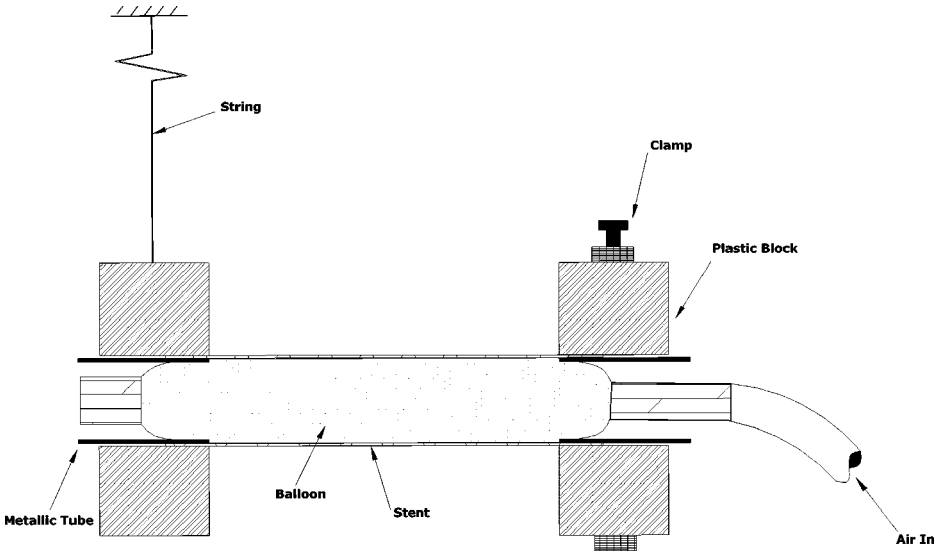


Fig. 4 Experimental scheme for the measurement of the bending of the stent under a fixed-fixed end condition

$$r_0 = R, \quad r_0' = 0 \quad r_N = R \quad r_N' = 0 \tag{10}$$

where R is a prescribed radius. Since two of the boundary conditions are given at $k=N$, these two equations can be used in ((9*c*) and (9*d*)) to eliminate two of the unknowns at $k=0$. This, however, destroys the explicit nature of Eq. (9*d*) by introducing nodal values at all nodes on the right hand side; therefore the solution cannot march from $k=0$ to $k=N$, but must be performed in an iterative manner. Assume the initial radius everywhere of the stent under internal pressure p_a is R; thus, $r_k^0 = R$ where the subscript k indicates the node number, and the superscript indicates the iteration number. All F_k are then calculated according to Eq. (2), followed by an estimation of C_k, D_k, G_k, and H_k in Eqs. (9). From the boundary conditions in Eq. (10), we have

$$\begin{aligned} r_N &= r_0 + Nhr_0' + \frac{N^2-1}{2}h^2 r_0'' + \left(\sum_{j=1}^{N-1} H_j + \frac{1}{2}H_{N-1}\right)h^3 r_0''' \\ &\quad + \left(\sum_{j=1}^{N-1} G_j + \frac{1}{2}G_{N-1}\right)h \\ r_N' &= r_0' + Nhr_0'' + G_N h^2 r_0''' + H_N . \end{aligned} \tag{11}$$

These are two linear equations for estimating r_0'' and r_0'''; after determining these and introducing in Eq. (9*d*) the update r_k^1 is obtained. This process is repeated until $|r_k^M - r_k^{M-1}| < \varepsilon$ for every node k from 0 to N. Then r_k^M are taken as the final solution of the problem; $\varepsilon \sim 10^{-5}$ was used in our simulations. In all the examples described in Section 4, the numerical procedure converged within a few seconds on a Pentium III class computer. This is perhaps not the most elegant solution procedure, but is adequate. A shooting method can also be used to solve this problem; Press et al. [7] suggest that for two-point boundary value problems of the type described here, the shooting method is preferable. Alternatively, a MATLAB script can also be generated to accomplish the solution.

4 Experimental Confirmation

In order to verify the beam-on-elastic foundation model of the stent as well as the numerical evaluation, an experiment was performed with a woven metallic stent shown in Fig. 1. The experimental arrangement is shown in Fig. 4. In this experiment, two rigid cylindrical blocks were used to constrain the radial expansion of the stent at the two ends. In each of the blocks, a hole with a given radius of R, slightly smaller than the free radius of the stent, was drilled. One block was held fixed with a clamp and the other was hung on a string so as to be at the same level; the string also allowed the block to move freely in the axial direction. The ends of the stent were inserted into the holes in the blocks, and two metallic tubes with proper outer diameter were inserted into the holes with their external surface touching the stent, strictly enforcing the fixed end support condition. Finally, a polyethylene bag was inserted into the stent, which when pumped with air provided internal pressure to the stent. A picture was taken when the pressure reached 153 mm of Hg to evaluate the deformed shape; a Nikon CoolPix 950 digital camera at a resolution of 1600×1200 pixels was used.

The image of the stent was imported into Adobe Photoshop software for quantitative measurements. On the lateral surface of the stent, the radius was measured at 23 selected points both before and after the pressure was applied. The measured final radius normalized by the initial radius (r/r_0) is plotted as a function of the position normalized by the length of the stent, (x/L) in Fig. 5. The numerical solution of Eq. (5) for this stent is also shown in the figure; the end conditions in Eq. (6) were used at both ends of the stent. The parameters for the stent are as follows: $r_0 = 10.6$ mm, $R = 9.9$ mm, $\alpha_0 = 34°$, $a = 85$ μm, $E = 200$ GPa, $G = 77$ GPa, and $n = 36$. From the comparison shown in Fig. 5, we see that the beam-on-elastic foundation model has captured the complete response of the stent quite well. The measured maximum radius and the calculated maximum radius differ by about 2%, which is acceptable considering that we have used a homogenized model for the bending stiffness of the wire mesh. When the stent is constrained at the ends from expanding radially, we observe that the maximum expansion in the radius is on the order of 15%; in contrast, for the free stent, from Eq. (1) we found that at the same pressure (153 mm of Hg), the maximum expansion of the radius is about 22% (we note that this depends critically on the initial radius of the stent). A somewhat counterintuitive aspect of the deformation of the stent under this condition might be the appearance of maximum radial expansion away from the midpoint if intuition is developed based on membrane deformation. The fundamental difference in the nature of the deflection is due to the incorporation of bending in the current model. This feature of the solution is rather well known for cylindrical pressure vessels with stiff end caps, [8]. We note that even if the length of the stent is long in comparison to the radius, the middle region will experi-

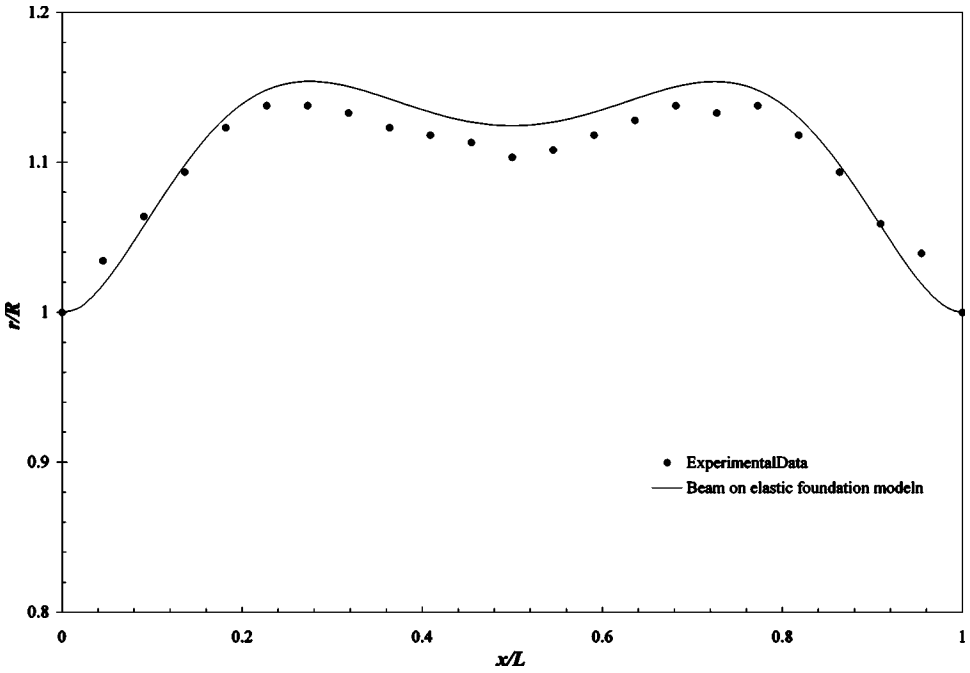


Fig. 5 Variation of the radius of the stent with position along the stent. The experimental measurements were obtained at a pressure of 153 mm of Hg; comparison to beam on elastic foundation model is also shown as "numerical solution."

ence a uniform expansion, but the maximum deflection will still occur close to the fixed supports. Another important point to note is that such variations in the stent diameter will influence the blood flow, and in particular the wall shear stress distribution; this is considered to play a key role in the biochemical response—formation of an endothelial layer, thrombus, etc.—and hence the coupled solid-fluid interaction problem should be of significant interest.

As a second example, consider the release of a stent from a catheter. A photograph of a stent partially pushed out from the catheter is shown in Fig. 6. The end of the stent outside the catheter experiences a free boundary condition while the end inside the catheter experiences a fixed boundary condition, with the radius equal to the internal radius of the catheter. In performing a numerical simulation with these boundary conditions, we began with an initial guess for the shape that varied parabolically from the radius of the catheter to the free radius of the stent; the iterative procedure is then turned on and a converged shape of the stent as it emerges from the catheter is predicted. It must be noted that we have not introduced any axial deformation into the beam model, but the stent experiences an unrestrained axial contraction as it comes out of the catheter. The axial contraction is incorporated in the model by simply mapping the nodal points x_k to the deformed position X_k given by the current radius r_k, and the inextensibility condition of the stent wire (Eq. (1) of Part I). Thus,

$$X_k = X_{k-1} - \frac{2r_k(r_k - r_{k-1})}{\sqrt{(r_0/\cos\alpha_0)^2 - r_k^2}} \quad k = 1 \ldots N \tag{12}$$

and $X_0 = x_0$. A comparison of the shape of the stent exiting the catheter as predicted by the model with the shape measured from Fig. 6 is shown in Fig. 7. Clearly, the beam-on-elastic foundation model provides a good representation of the deformability of the stent.

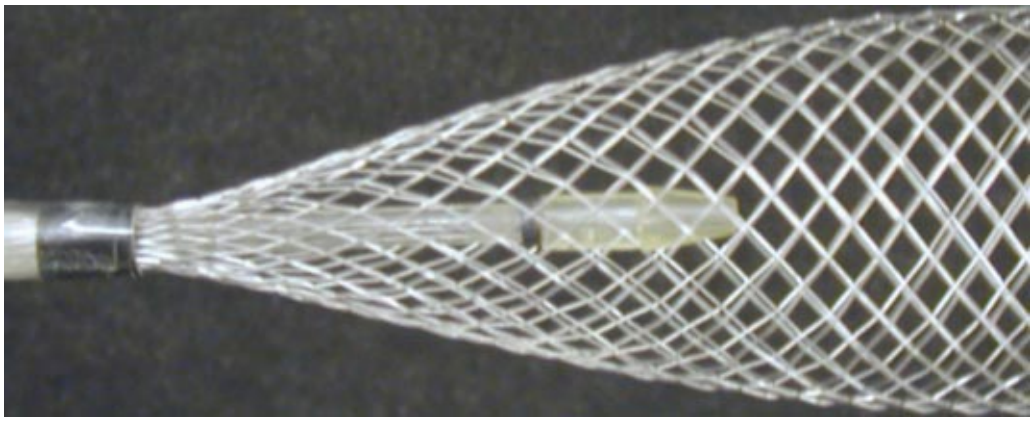


Fig. 6 Photograph showing the variation of the radius of the stent as it exits the catheter

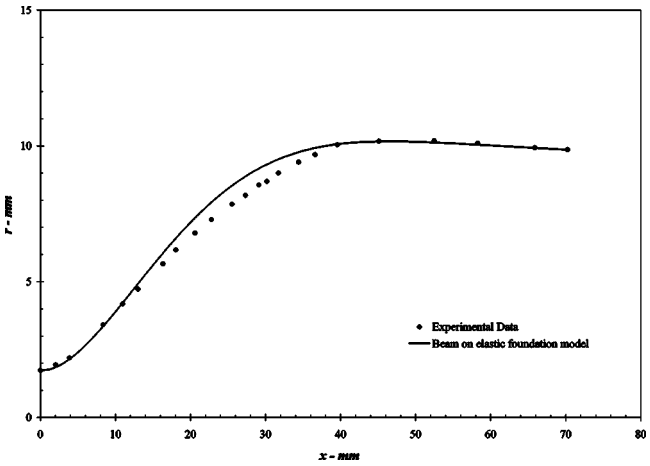


Fig. 7 Comparison of the measured variation of the radius of the stent with position along the stent as it exits the catheter with the predictions of the beam on elastic foundation model

5 The Design Problem

The two examples shown in Section 4 have demonstrated clearly that the mechanical response of the stent has been captured by our two mathematical models—the first model based on the deformation of a slender rod provides the pressure-diameter relationship of the stent and the second model based on the beam on elastic foundation analysis models the bending of the stent under axially varying, but axisymmetric loading. With these two models, we believe that we can now look into the coupled response of the stent and the artery in most applications of these metallic stents and evaluate design issues. Of course, before proceeding, an appropriate model describing the behavior of the artery into which the stent is to be inserted must be obtained. Several avenues are open to accomplish this; a common approach in the physiological literature is based on characterizing the pressure-diameter relationship through direct measurements, [9,10]. This is commonly used in coupled modeling of blood flow in distensible arteries, [11,12]. Constitutive equations based on the theory of finite deformations have also been presented in the literature, [13,14]; the strain energy density function is characterized in terms of the principal stretches, with model calibration provided from uniaxial, biaxial and tube-inflation experiments. While the framework based on finite deformation theory appears to be complete, definitive experimental characterization of arterial properties is not yet completely satisfactory (see Ref. [14] for a recent discussion). Nevertheless, from this analysis, one can obtain the nonlinear pressure-diameter relationship (see Eq. (7.50) in [14]):

$$p_A = f_A(r_i, r_A) \tag{13}$$

where r_i and r_A are the inner and outer radii of the artery. Note that this assumes axially homogeneous deformations. While general formulation of the axisymmetric problem is easily written down, there are very few attempts at solving these equations, since much of the focus has been on membrane problems. Finite element formulations of the problem are increasingly more accessible and may have to be relied upon for this problem. For example, Raghavan and Vorp [15] have recently evaluated the wall stresses in an abdominal aortic aneurysm with a finite deformation formulation and a generalized neo-Hookean constitutive model. To our knowledge, however, the coupled problem of a stent and an artery has not been fully examined. Rogers et al. [16] have performed finite element analysis to evaluate the development of stresses in the artery during placement of a balloon-expandable stent; however, they used a linearly elastic model for the artery. It would appear that treating the artery as a nonlinear membrane with an appropriate pressure-diameter relationship can provide a quick estimate of the magnitude of the stresses in the artery as well as the deformation of the stent itself.

Regardless of how the artery is modeled, interface conditions must be enforced at contact between the stent and the artery. Since the fully expanded stent is always larger in diameter than the artery, the net effect of the stent on the artery will always be an internal pressure; thus, the radius and slope of the stent must match that of the artery at every region of contact between the stent and the artery

$$r_{\text{stent}} = r_{\text{vessel}}, \quad \left.\frac{dr}{dx}\right|_{\text{stent}} = \left.\frac{dr}{dx}\right|_{\text{vessel}}. \tag{14}$$

Equation (13) must be solved simultaneously with the equation for the stent Eq. (5) and with interface conditions in Eq. (14) in order to determine the radius as a function of axial position. While this is a formal statement of the coupled problem, estimates of the arterial properties and simultaneous imposition of matching conditions over the overlapping regions still remain to be elucidated. Clearly a number of parameters are open for selection: the free diameter of the stent, the stiffness of the stent (the general construction of the stent), the lengths of the overlap between the stent and the artery in an aneurysm, etc; all of these parameters influence the radial force generated on the artery, the arterial opening presented by the stent, the wall stresses on the artery, etc. The coupled problem posed here can then be used to optimize the design of the stent to achieve particular objectives. The solution of the coupled problem of the stent and artery will be addressed in a future contribution.

6 Conclusion

In Part I of this series, we developed an equivalent pressure-diameter relationship for a helically wound wire metallic stent. In this paper, Part II, we took advantage of an analogy between a beam on elastic foundation and a thin-walled tube under pressure to determine the shape of the stent under spatially non-uniform loading. While the analogy has traditionally been used only for small deflections, within the linear range of material behavior, here we have extended it to account for large deformations and nonlinear pressure-diameter relationship that governs the response of the metallic stent. Predicted shapes of the stent for two very different sets of boundary conditions were compared favorably to experimental measurements. We believe that these two models are now well suited for evaluating the coupled response of the stent and the aorta; a formal statement of the coupled problem has been posed, pointing the need for the generation of an appropriate pressure-diameter relationship for the artery.

References

[1] Wang, R., and Ravi-Chandar, K., 2004, "Mechanical Response of a Metallic Aortic Stent: I. Pressure-Diameter Relationship," ASME J. Appl. Mech., **71**, pp. 697–705.
[2] Ormiston, J. A., Dixon, S. R., Webster, M. W. I., Ruygrok, P. N., Stewart, J. T., Minchington, I., and West, T., 2000, "Stent Longitudinal Flexibility: A Comparison of 13 Stent Designs Before and After Balloon Expansion," Cath. Cardiovas. Intervent., **50**, pp. 120–124.
[3] Dumoulin, C., and Cochelin, B., 2000, "Mechanical Behavior Modeling of Balloon-Expandable Stents," J. Biomech., **33**, pp. 1461–1470.
[4] Etave, F., Inet, G., Boivin, M., Boyer, J. C., Rioufol, G., and Thollet, G., 2001, "Mechanical Properties of Coronary Stents Determined by Using Finite Element Analysis," J. Biomech., **34**, pp. 1065–1075.
[5] Timoshenko, S. P., 1956, *Strength of Materials*, Part II, 3rd Ed., D. Van Nostrand Company, Princeton, NJ, pp. 124–128.
[6] Akai, T. J., 1994, *Applied Numerical Methods for Engineers*, John Wiley and Sons, New York, p. 242.
[7] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 2002, *Numerical Recipes in C++*, Cambridge University Press, New York, p. 753.
[8] Timoshenko, S. P., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, McGraw-Hill, New York, pp. 475–478.
[9] Peterson, L. H., Jensen, R. E., and Parnel, J., 1960, "Mechanical Properties of Arteries in Vivo," Circ. Res., **8**, pp. 622–639.
[10] Länne, T., Stale, H., Bengtsson, H., Gustafsson, D., Berqvist, D., Sonesson, B., Lecerof, H., and Dahl, P., 1992, "Noninvasive Measurement of Diameter

Changes in the Distal Abdominal Aorta in Man,'' Ultrasound Med. Biol., **18**, pp. 451–457.
[11] Dutta, A., Wang, D. M., and Tarbell, J. M., 1992, ''Numerical Analysis of Flow in an Elastic Artery Model,'' ASME J. Biomech. Eng., **114**, pp. 26–33.
[12] Quarteroni, A., Tuveri, M., and Veneziani, A., 2000, ''Computational Vascular Fluid Dynamics: Problems, Models and Methods,'' Comput. Visual. Sci., **2**, pp. 163–197.
[13] Fung, Y. C., 1993, *Biomechanics: Mechanical Properties of Living Tissues*, Springer-Verlag, New York.
[14] Humphrey, J. D., 2002, *Cardiovascular Solid Mechanics, Cells, Tissues and Organs*, Springer-Verlag, New York.
[15] Raghavan, M. L., and Vorp, D. A., 2000, ''Toward a Biomechanical Tool to Evaluate Rupture Potential of Abdominal Aortic Aneurysm: Identification of a Finite Strain Constitutive Model and Evaluation of Its Applicability,'' J. Biomech., **33**, pp. 475–482.
[16] Rogers, C., Tseng, D. Y., Squire, J. C., Edelman, E. R., 1999, ''Balloon-Artery Interactions During Stent Placement: A Finite Element Analysis Approach to Pressure, Compliance, and Stent Design as Contributors to Vascular Injury,'' Circ. Res., **84**, pp. 378–383.

Characterization of Plastic Deformation Induced by Microscale Laser Shock Peening

Hongqiang Chen

Jeffrey W. Kysar

Y. Lawrence Yao

Department of Mechanical Engineering,
Columbia University,
New York, NY 10027

Electron backscatter diffraction (EBSD) is used to investigate crystal lattice rotation caused by plastic deformation during high-strain rate laser shock peening in single crystal aluminum and copper sample on $(1\bar{1}0)$ and (001) surfaces. New experimental methodologies are employed which enable measurement of the in-plane lattice rotation under approximate plane-strain conditions. Crystal lattice rotation on and below the microscale laser shock peened sample surface was measured and compared with the simulation result obtained from FEM analysis, which account for single crystal plasticity. The lattice rotation measurements directly complement measurements of residual strain/stress with X-ray micro-diffraction using synchrotron light source and it also gives an indication of the extent of the plastic deformation induced by the microscale laser shock peening.
[DOI: 10.1115/1.1782914]

1 Introduction

Shot peening is a process involving multiple and repeated impacts by bombarding a surface with relatively hard particles with sufficient velocities to indent the surface, [1]. Shot peening is widely used to improve the fatigue behavior of mechanical components by introducing compressive stress on the peened surface.

Laser shock peening (LSP) has been studied since 1960s. As shown in Fig. 1, it is a surface treatment wherein, laser-induced shocks introduce compressive residual stresses relatively deep within the material resulting in an increased resistance of the material to various forms of failure, [2]. In particular, LSP can induce compressive residual stresses in the target surface and improve its fatigue life, which is important in applications such as turbine blades of aircraft engine. The potential benefits of laser peening over shot peening include a greater residual compressive stress depth and little change to either surface finish or shape. Also the process parameters such as laser intensity and laser pulse duration are much easier to control than shot peening. Finally, it is possible to apply LSP to only selected regions of a component, because of the ability to precisely dictate the position of the laser.

Recently, laser shock processing of polycrystalline aluminum and copper using a micron length scale laser beam has been studied, [3–5]. It has been shown that microscale laser shock peening (μLSP) can efficiently induce favorable residual stress distributions in bulk metal targets as measured by X-ray diffraction with micron-level spatial resolution, [6], and calculated through finite element analysis (FEM) simulations, [7]. Thus, microscale laser shock peening (μLSP) is a technique that can be used to manipulate the residual stress distributions in metal structures over regions as small as a few microns and thus improve the reliability of microdevices.

The mechanics of laser shock peening presents many exciting challenges, because it is a hybrid process involving many disciplines. Classical solid continuum mechanics has been very successful in describing shock-compression process for conventional shot peening. However, for microscale laser shock peening, high laser power intensity ($4\ \mathrm{GW/cm^2}$) and short shock peening times (laser pulse duration$=50$ ns) introduce high strain rate plastic deformation confined to the micron length scale. Micromechanical considerations of strain gradient plasticity, rate-dependent plasticity and its relationship to crystal structure, crystal lattice orientation, dislocation and cell structure formation under shock wave loading at the micron length scale require careful study. In addition, there is a solid-fluid interaction, because the specimen is submerged in water during the LSP process.

In this paper, electron backscatter diffraction (EBSD) is used to investigate crystal lattice rotation caused by high-strain rate microscale laser shock peening in single crystal aluminum and copper sample on $(1\bar{1}0)$ and (001) surfaces. For the first time, crystal lattice rotation on and below the microscale laser shock peened sample surface was measured; these are compared with FEM simulations based on single crystal plasticity. The experimental results provide useful insight into the high-strain rate shock peening process at the microscale. Also the experiments provide new methodologies for characterizing the microstructure formation and distribution of plastic deformation for microscale laser shock peening.

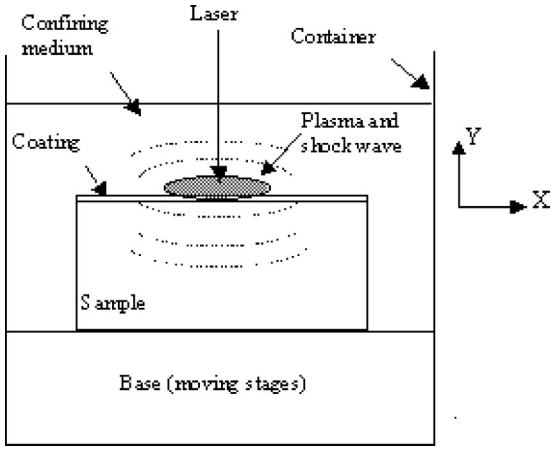


Fig. 1 Laser shock peening setup

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, September 4, 2003; final revision, April 29, 2004. Associate Editor: K. M. Liechti. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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.

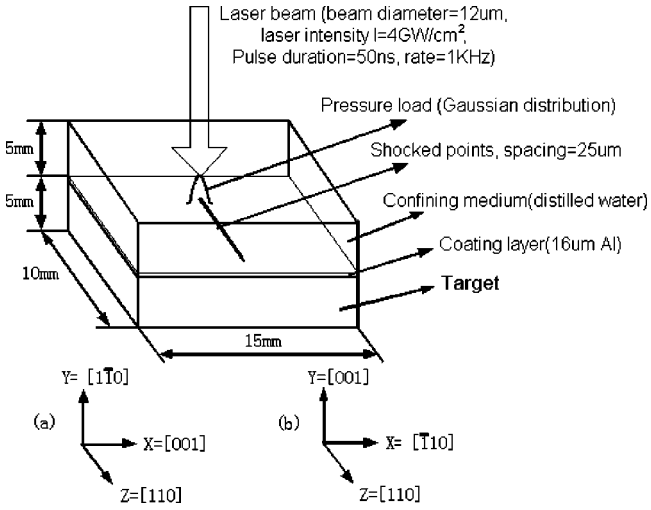


Fig. 2 Sample geometry and laser shock peening condition; (*a*) Al(1$\bar{1}$0) sample and Cu(1$\bar{1}$0) sample, (*b*) Al (001) sample

2 Material Preparation and Laser Shock Peening Conditions

Face-centered-cubic (FCC) metals such as copper and aluminum are routinely used in microdevices due to their good mechanical and electrical properties. Fully annealed single crystals of pure aluminum and copper (grown by the seeded Bridgman technique) were used for microscale laser shock peening herein. Laue X-ray diffraction was used to determine the crystal orientation within $\pm 1°$ and the sample was cut to shape using a wire electrical discharge machine (EDM). Regular mechanical polishing with diamond grit sizes 6 and 1μm was used to remove the heat affected zone of the cutting surface and electrochemical polishing was applied for all samples to eliminate any remaining deformed material prior to shock peening.

It is known that a line loading parallel to a $\langle 110 \rangle$ direction in an FCC crystal induces a state of plane deformation, [8]. Thus successive shock peens were applied to the material along a line parallel to [110] direction in an attempt to achieve a final deformation state that approximates a plane deformation state. In order to achieve a symmetric deformation field, the shock peens were applied to either the (1$\bar{1}$0) surfaces of Al and Cu or (001) surface of Al.

The samples, shown in Fig. 2, have the dimensions of 15 mm $\times$10 mm$\times$5 mm. The coordinate systems used throughout this paper are indicated in Fig. 2 and defined as follows: Z-axis is parallel to the shock line which has direction of [110], X-axis is parallel to [001] direction and Y-axis is parallel to [1$\bar{1}$0] normal to the shocked Al(1$\bar{1}$0) and Cu(1$\bar{1}$0) surfaces. For the Al (001) sample, the X-axis is parallel to [$\bar{1}$10] and Y is parallel to [001] normal of the shocked surface, with the Z-axis again parallel to the shocked line in direction of [110].

In the laser shock peening, a frequency tripled Q-switched neodymium: yttrium-aluminum-garnet (Nd:YAG) laser (wavelength 355 nm) in transverse electromagnetic modes 00 (TEM_{00}) mode was used. The pulse duration was 50 ns, spacing between consecutive pulses along a shock line was 25 μm, and pulse numbers were three on each shocked location at 1 KHz pulse repetition rate. Laser beam diameter was 12 μm and laser intensity was approximately 4 GW/cm^2. A thin layer of high vacuum grease (about 10 microns thick) was spread evenly on the polished sample surface, and a 16-μm thick polycrystalline aluminum foil, chosen for its relatively low threshold of vaporization, was tightly pressed onto the grease. The sample was placed in a shallow container filled with distilled water around 5 mm above the sample's top surface. After shock processing, the coating layer and the vacuum grease were manually removed. The induced deformation is due to shock pressure and not due to thermal effects since only the coating is vaporized by the laser shock. Further details of microscale LSP setup are given in [3–5].

3 Characterization of Laser Shock Peening

Several different experimental methods were employed to characterize the laser shock peening regions. Atomic force microscopy (AFM) was used to measure the deformation geometry on the shocked surfaces. Crystal lattice rotation was characterized by electron backscatter diffraction (EBSD) to measure crystallographic orientation as a function of position. Moreover, X-ray microdiffraction [6] was applied to measure crystal lattice rotation on the shocked surface, as well as to measure shock-induced residual stress. The result from each of these methods are discussed in detail in this section.

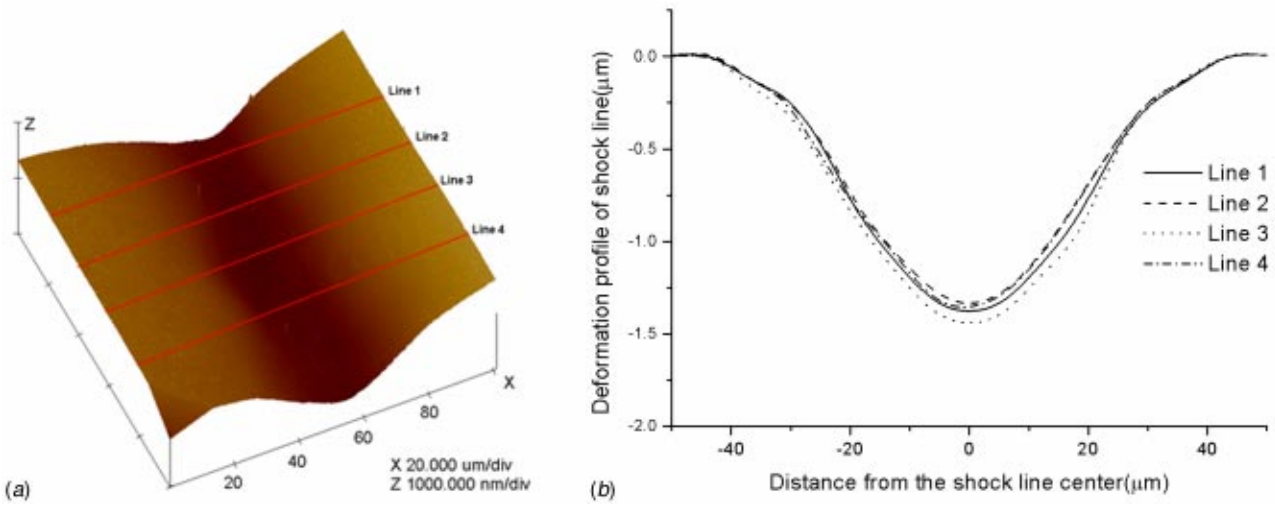


Fig. 3 Measurement of shocked line geometry using AFM for Al(1$\bar{1}$0) sample (scan area=100$\times$100 μm); (*a*) three-dimensional geometry, (*b*) cross section geometry at different positions (line spacing=20 μm)

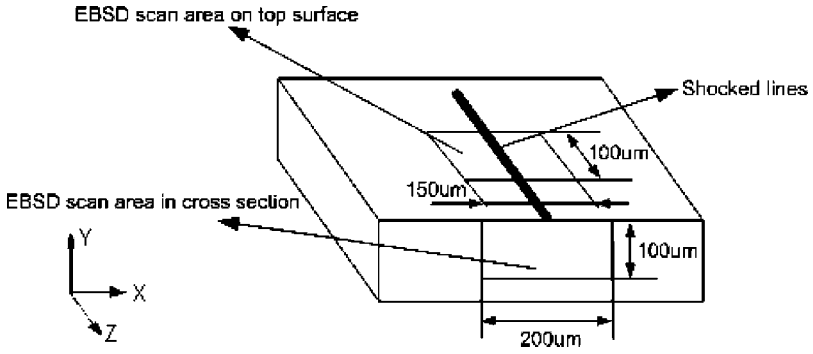


Fig. 4 EBSD automatic indexing map on top surface and on newly exposed cross section

3.1 Sample Deformation Measured by AFM. A typical three-dimensional geometry of the shocked region of Al (001) sample measured using AFM (Digital Instruments Nanoscope Inc.) is seen in Fig. 3(*a*); the scan area is 100×100 μm and 512 measurements were made along each direction. The scan direction is set to parallel to *Z*-direction (i.e., along the shock direction) to decrease the uncertainty of measurement. In order to check the deformation profiles at different location along the shocked line, detailed cross-section profiles at four different positions with spacing$=20$ μm (red lines 1–4 in Fig. 3(*a*)) are shown in Fig. 3(*b*). As is evident, the depth of the shock line is around 1.5 μm with width of 90 μm. It is clear that the shocked line is surprisingly uniform deformed along [110] direction in spite of the fact that the laser shocks were created sequentially. The lateral extent of the AFM measurements was not sufficient to show pileup on the edges of the shock line. However, additional measurement, to be discussed in Section 4.3, shows evidence of pileup.

3.2 Lattice Orientation Measurement With EBSD (Electron Backscatter Diffraction). From the work of Kysar and Briant [9], it is possible to measure the extent as well as character of the lattice rotation below the shocked surface by using electron backscatter diffraction (EBSD) to measure crystallographic orientation as a function of position. EBSD is a diffraction technique for obtaining crystallographic orientation with submicron spatial resolution from bulk samples or thin layers in a scanning electron microscope (SEM), [10].

The crystallographic orientation of the shock peened top surface was collected using EBSD, which provided information about the lattice rotation on the shocked surface. After that, in order to obtain the depth distribution and magnitude of lattice rotation below the shocked surface, the specimen was sectioned via wire EDM to expose a (110) plane in a region which experienced an approximate plane-strain deformation state due to μLSP. The newly exposed center surface was polished again after which the crystal orientation of the sectioned surface was mapped using EBSD, as indicated schematically in Fig. 4.

EBSD data was collected using a system supplied by HKL Technology, [11], and attached to a JEOL JSM 5600LV scanning electron microscope. All data were acquired in the automatic mode, using external beam scanning and employing a 1-μm step size. A typical scan area is 100 μm$\times$150 μm on the shocked surface and 200 μm$\times$100 μm on the cross section as in Fig. 4. The EBSD results from each individual scan comprise data containing the positioncoordinates and the three Euler angles which describe the orientation of the particular interaction volume of the crystal relative to the orientation of the specimen in the SEM. This information allows the in-plane and the out-of-plane lattice rotations to be calculated relative to the known undeformed crystallographic orientation, which serves as the reference state.

3.2.1 Image Pole Figure From Sample Top Surface. Pole figures or inverse pole figures are commonly used to analyze textures based on information of lattice orientation obtained from EBSD. The orientation of the crystal at each measurement position is represented by a point on the stereographic polar net, [12]. Figure 5 shows the inverse pole figure from the shocked surface for specimen Al (001) in which *Z* ([110]) is aligned with the shock line direction. The scan area covers a region ±50 μm in *X*-direction across the shocked line and 100 μm along the *Z*-direction with spatial resolution of 1 μm. It is clear from Fig. 5 that the [110] of the crystal remains closely aligned with the *Z*-axis after deformation. On the other hand, the inverse pole figures indicate a larger distribution of rotation of $[\bar{1}10]$ and [001] relative to the *X* and *Y*-axes, respectively. Thus, both the AFM and the EBSD results indicate that an approximate two-

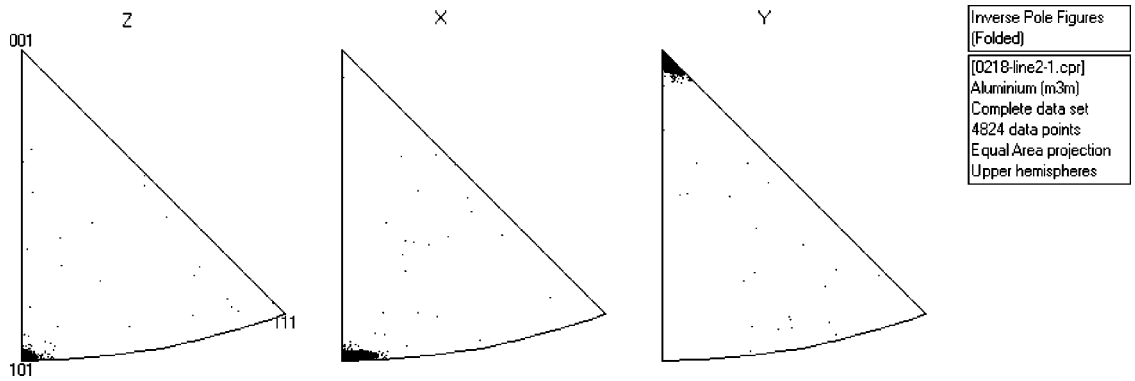


Fig. 5 Inverse pole figure of sample surface under shock peening (*Z*-direction is shock direction, *Y*-direction is the sample surface normal)

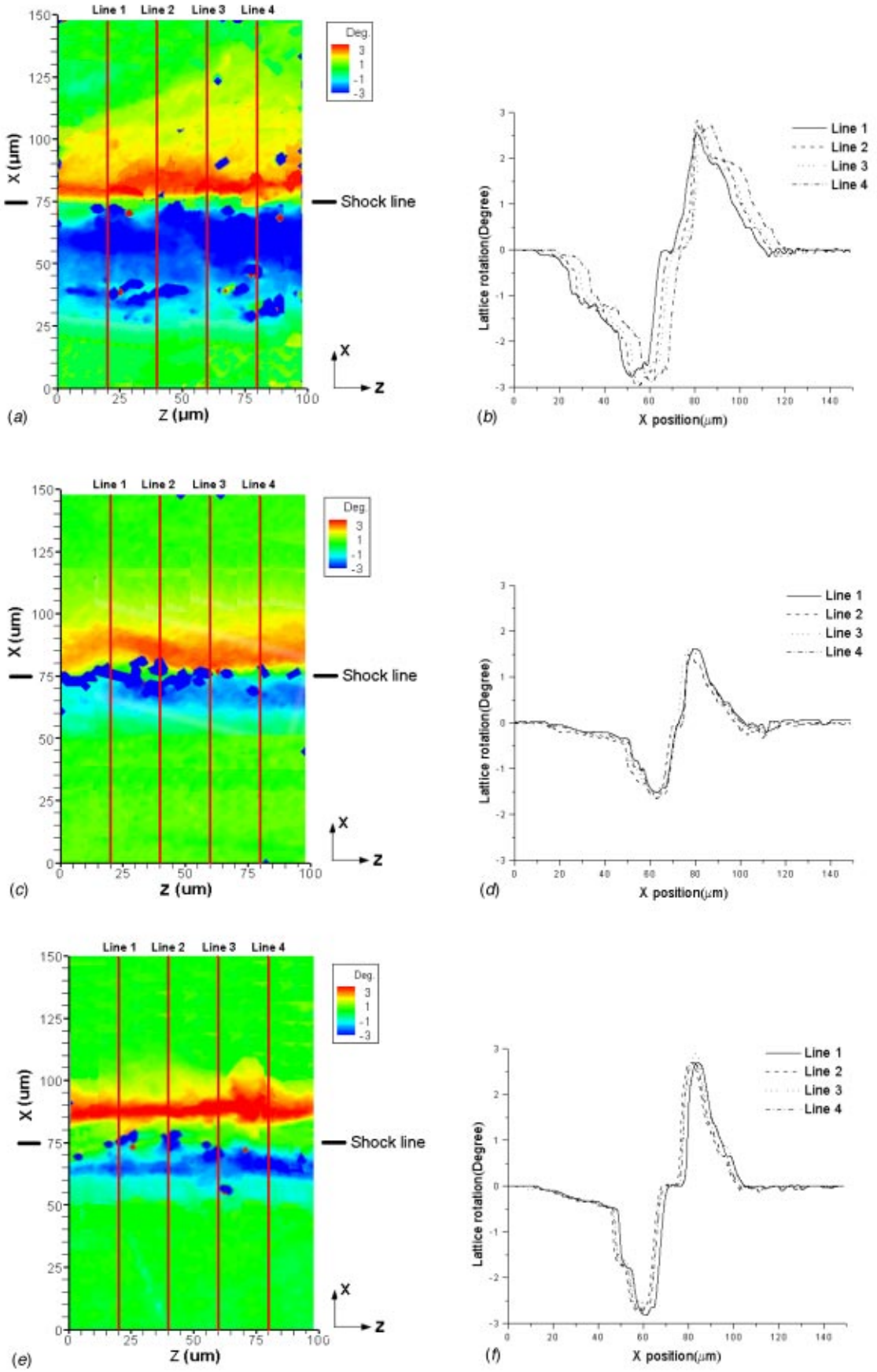


Fig. 6 Lattice rotation contour map on sample surface (line 1–4: four cross section with spacing=20 μm); (*a*) and (*b*): Al$(1\bar{1}0)$ sample; (*c*) and (*d*): Cu$(1\bar{1}0)$ sample; (*e*) and (*f*): Al (001) sample

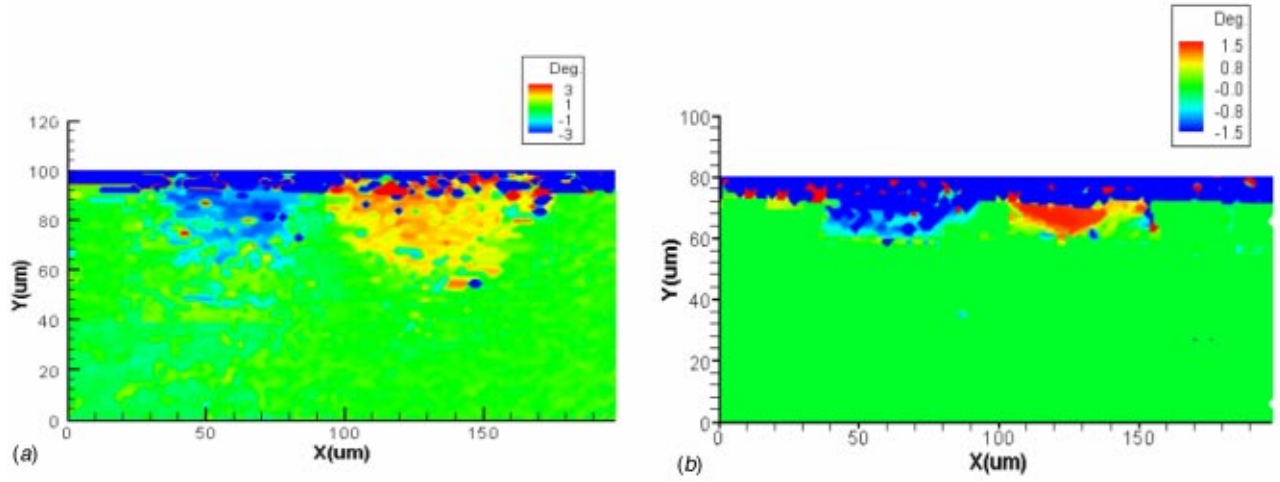


Fig. 7 Lattice rotation contour at the cross section (*a*) Al($1\bar{1}0$) sample; (*b*) Cu($1\bar{1}0$) sample

dimensional deformation state exists. We will appeal to this apparent two-dimensional deformation state when interpreting the other experimental data presented herein.

3.2.2 Lattice Rotation Measurement Results. Lattice rotation measured from top surface across the shocked line. The lattice rotation contour map on the shocked Al($1\bar{1}0$) sample's surface is shown in Fig. 6(*a*). Figure 6(*b*) shows the spatial distribution of lattice rotation along four lines across the shocked line with spacing$=20$ μm. The red region corresponds to counterclockwise rotation about the *Z*-axis which is positive and the blue region corresponds to clockwise rotation which is negative. It is clear to see that the lattice rotation is zero (green region) far away from the shocked line which corresponds to the shock-free region. Again, the lattice rotation distribution along the shocked line is quite uniform which further suggests the approximate two-dimensional deformation state mentioned before. The lattice rotation value is $\pm 3°$ between ± 35 μm from the center of shocked line and the rotation direction is anti-symmetric on both side of shocked line.

Figure 6(*c–d*) shows the lattice orientation change on the shocked Cu($1\bar{1}0$) sample surface. It is clear that both aluminum and copper shows the similar lattice rotation pattern. However, the region for lattice rotation in the copper sample is around ± 20 μm from the center of shocked line and the maximum value is about 1.5°, both of which are smaller than that of Al($1\bar{1}0$) sample.

In order to investigate the effect of crystal orientation on lattice rotation, an aluminum sample shocked on the (001) surface was also studied. Figure 6(*e–f*) shows the lattice rotation contour on the shocked surface. Compared with result of Al($1\bar{1}0$), the general trend of lattice rotation such as the rotation direction and magnitude is the same, but the shocked region is somewhat narrower.

Lattice rotation measured from cross section perpendicular to shocked line. From the measurement mentioned above, we obtained the lattice rotation result on the laser shock peened surface. In order to measure the lattice rotation below the sample surface and study the spatial distribution in the depth direction, the sample was sectioned via wire EDM and the crystallographic orientation of the newly exposed surface was mapped using EBSD. Then the in-plane lattice rotations beneath the laser shocked surface were measured via EBSD. Here we use the term ''in-plane'' because the experimental results indicate an approximate two-dimensional deformation state.

Figure 7(*a*) shows the lattice rotation in the cross section of the Al($1\bar{1}0$) sample. The lattice rotation varies between $\pm 3°$ in the region up to 40 μm below the sample surface. In the center of shocked line, the lattice rotation is nearly zero (green) and rotation direction reversed across the shocked line, which is consistent with the result from sample surface. The maximum lattice rotation occurs near the sample surface and the value decays as depth increases. Figure 7(*b*) shows the lattice rotation of on the cross section for the Cu($1\bar{1}0$) sample. The rotation distribution is similar in character to the Al sample, but the affected region in the depth direction is around 15 μm below the sample surface, smaller than that of Al sample. Also the total rotation angle varies between $\pm 1.5°$, rather than $\pm 3°$.

3.3 Lattice Rotation Measured by X-Ray Microdiffraction Spatially resolved residual stress/strain can be measured on the laser shock peened surface using X-ray microdiffraction from synchrotron radiation sources, [6]. It is also possible to determine lattice orientation on the shocked surface as a byproduct of the X-ray strain/stress measurement.

As discussed in [6] and illustrated in Fig. 8, two rotations, θ scan and χ scan were applied in the experiment by rotating the specimen until the maximum intensity is located in the detector in order to properly align the specimen in the X-ray apparatus. The θ scan ensures that the mean beam vector of incident X-ray, and not any other, is at the proper angle with respect to the surface and consequently, the proper diffraction angle is recorded by the detector arm. The χ scan ensures that the normal vector of the diffracting plane is contained in the same geometrical plane as the incoming and diffracted X-ray beams. These two scans applied

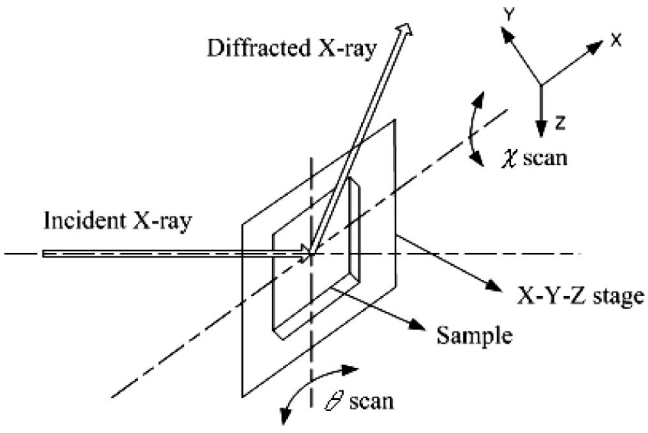


Fig. 8 Scan scheme of X-ray microdiffraction

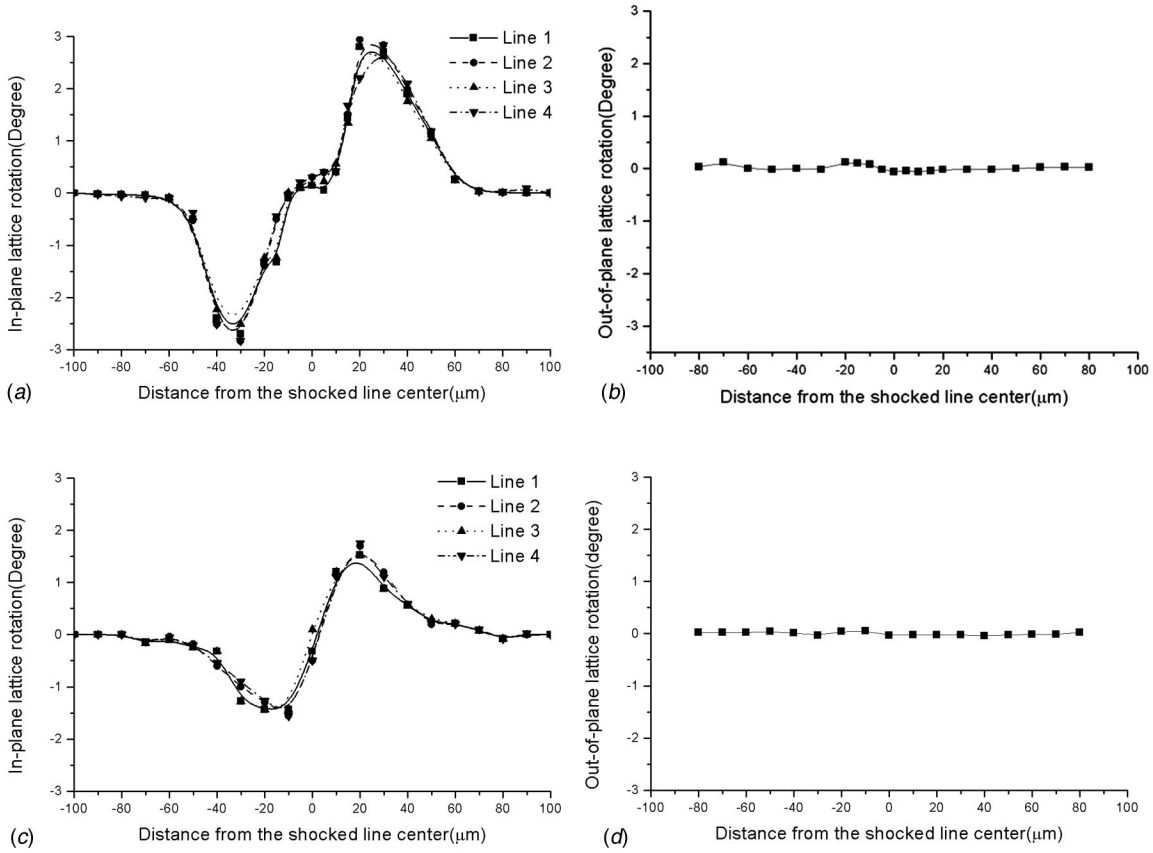


Fig. 9 (*a*) In-plane lattice rotation on shock peened surface of Al($1\bar{1}0$) sample. (*b*) Out-of-plane lattice rotation on shock peened surface of Al($1\bar{1}0$) sample. (*c*) In-plane lattice rotation on shock peened surface of Cu($1\bar{1}0$) sample. (*d*) Out-of-plane lattice rotation on shock peened surface of Cu($1\bar{1}0$) sample.

iteratively optimize the integrated intensity of the relevant reflection during alignment. In essence, the sample is rotated about its Z-axis to perform the θ scan, and about its X-axis to perform the χ scan. Therefore, the in-plane and out-plane lattice rotation can be obtained from the θ and χ scans as shown in Fig. 8.

Results of these measurements in Fig. 9(*a*) indicate that the spatial distribution of in-plane lattice rotation for the Al($1\bar{1}0$) sample is very similar to the EBSD results in Fig. 6(*a*) and 6(*b*). The maximum rotation angle is around $\pm 3°$ at position nearly ± 30 μm away from the center of shock line. While the variation of out-of-plane lattice rotation in Fig. 9(*b*) is only $\pm 0.1°$ which is quite small relative to in-plane lattice rotation. So this measured lattice rotation under shock peening is consistent with the uniform AFM profile along shock direction and two-dimensional in-plane lattice rotation assumption. Figures 9(*c*) and (*d*) shows similar results for Cu($1\bar{1}0$) sample. Thus, the lattice rotation measurements directly complement the material residual strain/stress measurements. Moreover, it also gives an indication of the extent of the plastic deformation induced by the microscale laser shock peening. Lattice rotation measured by X-ray microdiffraction is apparently more uniform than that from EBSD measurement, probably because X-rays penetrate deeper (20–30 μm), [13], than the electron beam used in EBSD (a few microns) and thus average the orientation over a large volume of material. The residual stress measurements that resulted from these experiments are discussed in detail in [6].

4 Theoretical Explanation and Simulations

In this section, we present results of elementary simulations of microscale laser-shock peening. Since the surface deformation and lattice rotation under laser shock peening indicate that an approximate two-dimensional deformation state exists, we will assume that the induced deformation state is strictly two-dimensional, which may be greatly oversimplified. However, it turns out that such an approach can shed significant insight into the mechanics of deformation which will be useful when the full three dimensional problem is addressed in future studies. Simulations of LSP pose many challenges because of the high transient pressures, fluid-solid interaction and high strain rates in a single crystal at the micrometer length scale which raises the possibility of the necessity to account for strain gradient effects. Given the

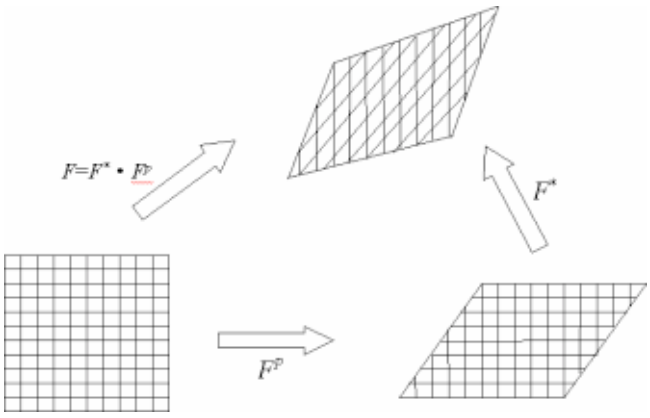


Fig. 10 Plastic deformation in single crystal plasticity

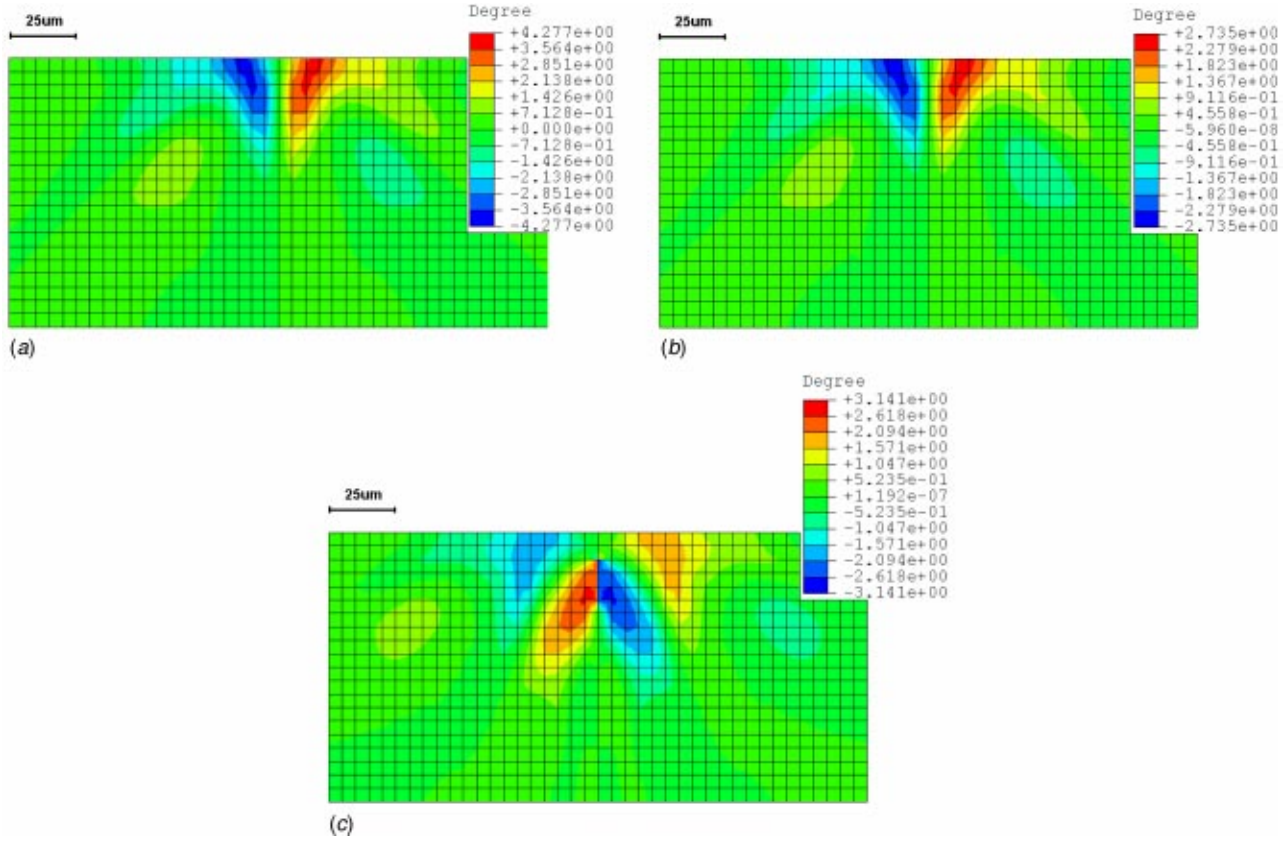


Fig. 11 Lattice rotation field on cross section of after laser shock peening; (*a*) Al$(1\bar{1}0)$, (*b*) Cu$(1\bar{1}0)$, (*c*) Al (001)

absence of constitutive data in this regime, it is impossible to incorporate realistically these effects into the model. Hence, we will make a grossly simplified assumption of ideal plastic behavior under quasi-static plane-strain conditions while implementing single crystal plasticity. Rate effects, hardening, strain gradient, and three-dimensional effects are neglected.

The goal of the simulation is then to attempt to understand the overall character of the deformation and lattice rotation fields and see how much can be predicted by such a simple simulation. In doing so, we can ascertain which of the dominate features of the fields are attributable to the anisotropic plastic behavior of the single crystal. Subsequent simulations, which account for more realistic material constitutive behavior, can then concentrate how the additional effects modify the baseline solution. Thus, this oversimplified approach can shed insight into the mechanics of deformation and lay the ground work for more realistic simulations in future studies which will include three-dimensional, dynamic, and strain-rate effects.

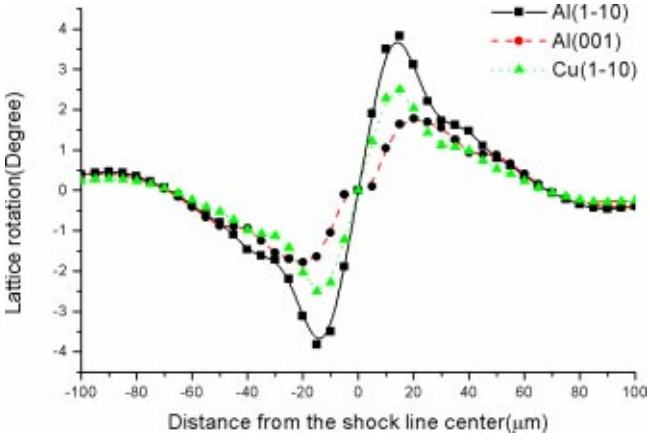


Fig. 12 Spatially distribution of latticed rotation on sample surface from simulation

4.1 Kinematical Theory of Single Crystal Mechanics

From single crystal plasticity theory, [14–16], there are two physically distinct mechanisms for deforming and reorienting materials—plastic slip and elastic lattice deformation. In general, the deformation gradient of a single crystal that undergoes plastic deformation can be written with reference to Fig. 10 as

$$F = F^* \cdot F^P \tag{1}$$

where F^P corresponds to the deformation caused by plastic shearing on crystallographic slip systems and F^* is caused by elastic stretching and rotation of the crystal lattice. The velocity gradient of material is given by a standard formula:

$$L = v\nabla = \dot{F} \cdot F^{-1} = D + \Omega. \tag{2}$$

The D and Ω terms are the symmetric rate of stretching tensor and the antisymmetric rate of spin tensor, respectively. They are then decomposed into parts due to plastic slip (D^P, Ω^P) and lattice deformation (D^*, Ω^*) as follows:

$$D = D^* + D^P, \quad \Omega = \Omega^* + \Omega^P. \tag{3}$$

The lattice rotation measured in the experiment is the Ω^* term integrated throughout the deformation history. The reader interested in the distinction between Ω^* and Ω^P-may refer to pg. 107 of Asaro's review paper, [16], for a full discussion of which rotation components leads to the measured lattice rotation field.

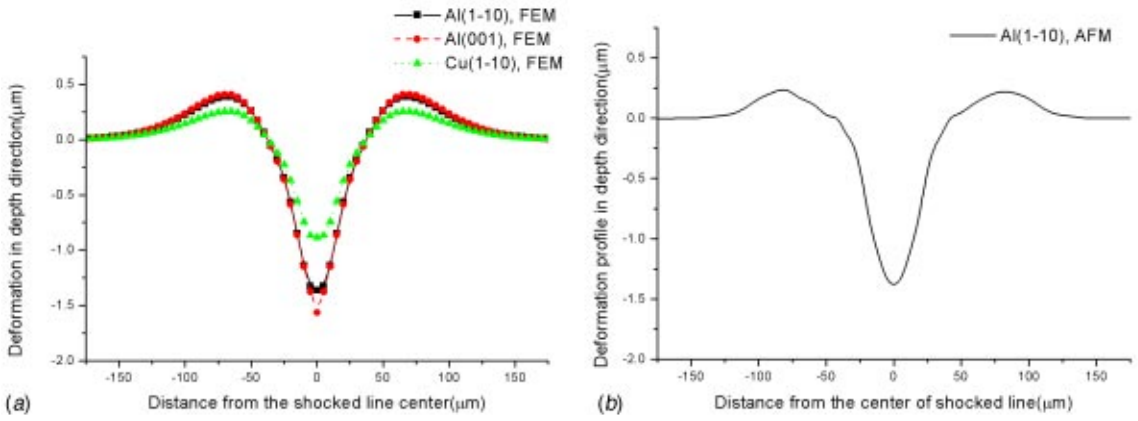


Fig. 13 Deformation profile in depth direction from simulation and AFM; (*a*) FEM result, (*b*) AFM result

4.2 FEM Analysis of Shock Peening With UMAT Incorporating Single Crystal Plasticity. A user-material subroutine (UMAT) for single crystal plasticity based on theory in [16] and written by Huang [17] and modified by Kysar [18] is incorporated into the finite element analysis using the general purpose finite element program ABAQUS/Standard, [19]. In the UMAT, the $\{111\}\langle 110\rangle$ slip systems in FCC metal are used for both single crystal Al and Cu. A critical shear strength $\tau_{\mathrm{CRSS}}=1$ MPa on each of the slip systems is assumed. The simulation is a two-step quasi-static loading and unloading process corresponding to the shock peening and relaxation processes. Following the work of Zhang and Yao [7], shock pressure obeys Gaussian spatial distribution, with its $1/e^2$ radius equals to $\sqrt{2}R$, where R is the radius of plasma. Letting x be the radial distance from the center of the laser beam, the spatially nonuniform shock pressure $P(x)$ is then given as

$$P(x)=P_0\exp\left(-\frac{x^2}{2R^2}\right) \tag{4}$$

on the shocked surface. P_0 is the peak value of shock pressure and the plasma radius $R=10\ \mu$m here. In order to make a dimensionless analysis, all simulation results are normalized as the function of two dimensionless parameters $(P_0/\tau_{\mathrm{CRSS}}, x/R)$. The boundary conditions of the plane strain model are as follows. At the top surface, surface traction equals the applied shock pressure, at the bottom surface, the vertical displacement is specified to be zero and the outer edges are traction-free. In the simulation, elastic-ideally plastic behavior is assumed so that hardening is neglected. In order to eliminate ''volume-locking'' which occurs in plastic deformation simulation, four-node linear elements with reduced

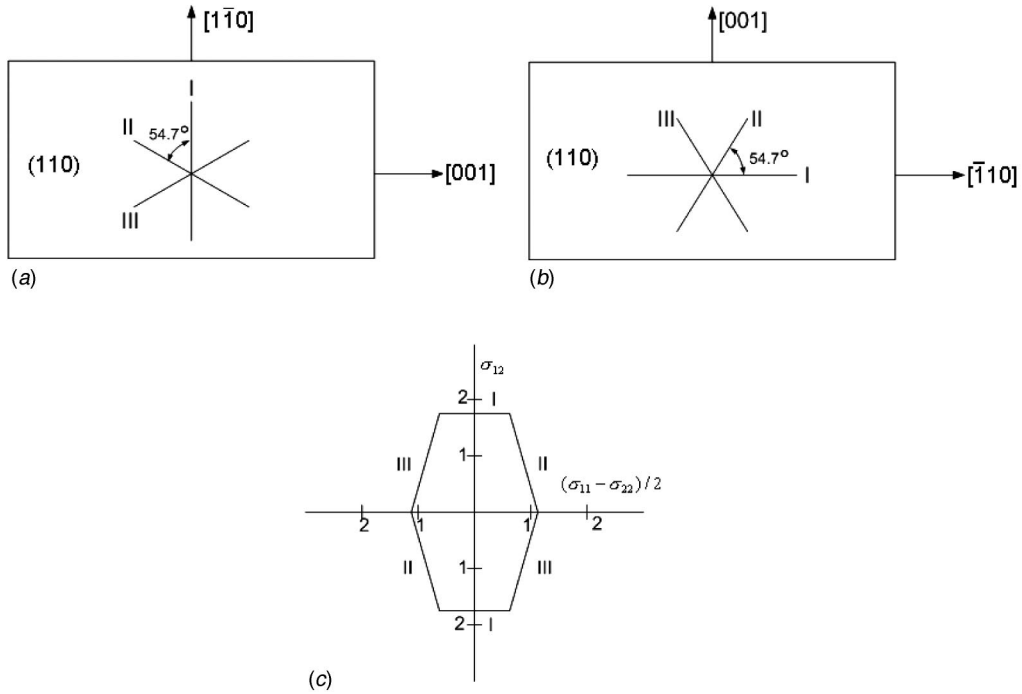


Fig. 14 Three plane-strain slip systems and yield surface in (110) plane; (*a*) Al(1$\bar{1}$0) and Cu(1$\bar{1}$0) sample, (*b*) Al (001) sample, (*c*) yield surface in (110) plane

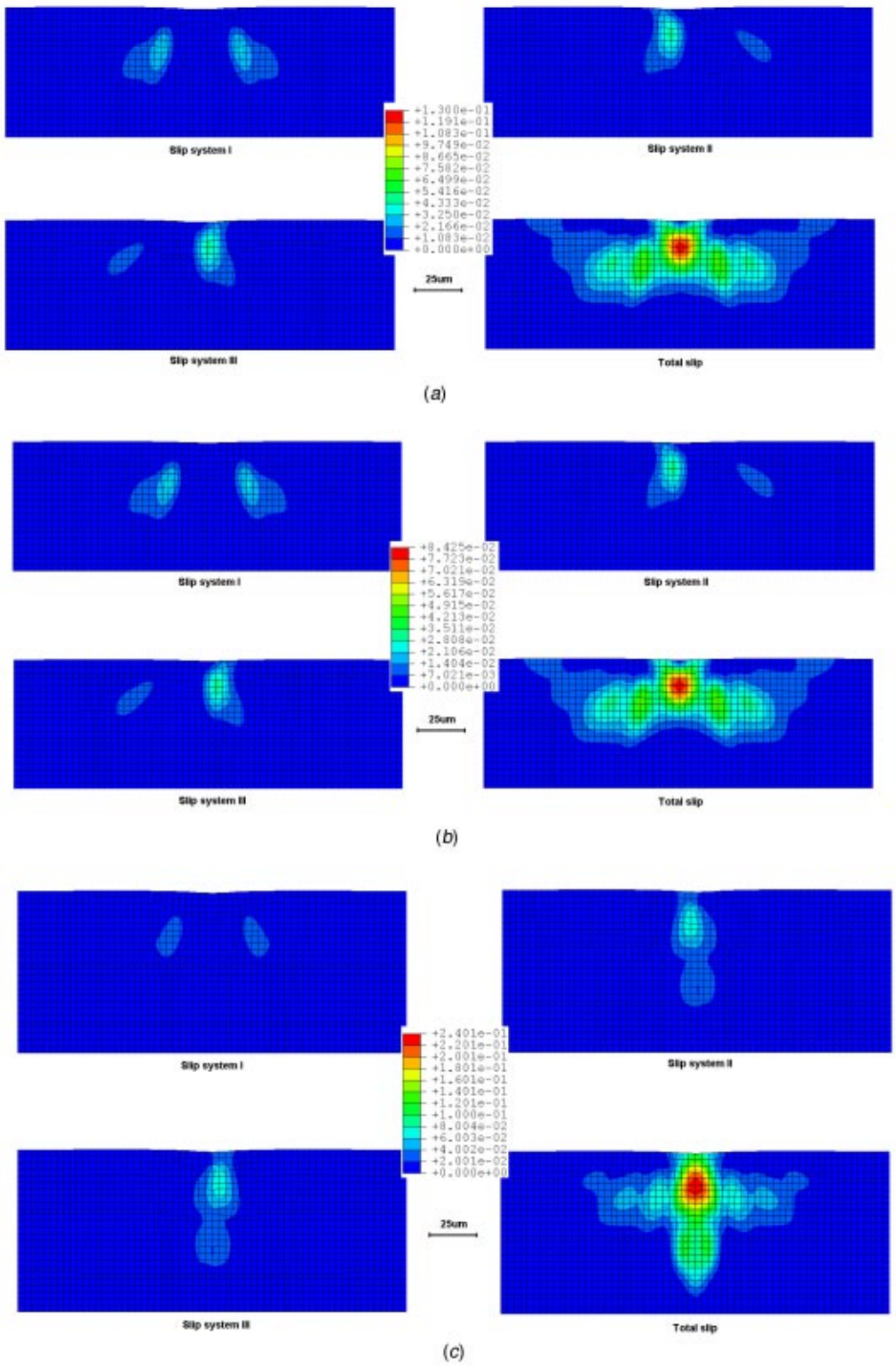


Fig. 15 Shear strain on active slips under laser shock peening from simulation; (*a*) shear strain of active slip systems for Al$(1\bar{1}0)$ sample, (*b*) shear strain of active slip systems for Cu$(1\bar{1}0)$ sample, (*c*) shear strain of active slip systems for Al (001) sample

integration and hourglass stiffness control are used. Two orientations and two materials are simulated, Al$(1\bar{1}0)$, Al (001), and Cu$(1\bar{1}0)$.

4.3 Finite Element Simulation Results. *Lattice rotation field for Al and Cu.* In Fig. 11(*a*), the calculated lattice rotation fields below the laser shocked surface of the Al$(1\bar{1}0)$ sample are shown. The shock pressure loading has been removed in the second step of simulation and the peak value of shock pressure is $P_0/\tau_{CRSS}=7$. The experimental results with the same material and same orientation is shown in Fig. 7(*a*). It can be seen that the antisymmetric pattern of lattice rotation field and the sign of ro-

tation (+, counterclockwise, −, clockwise) is consistent with the experiment result from EBSD (Fig. 7(*a*)). The magnitude of rotation is around $\pm 4°$ in simulation which is close to the $\pm 3°$ from experiment result. The lattice rotation field in the simulation extends over a region across the shock line center with width of $\pm 40\ \mu$m (0 corresponds to center) and depth of 35 μm.

Figure 11(*b*) shows the analogous lattice rotation fields for Cu$(1\bar{1}0)$ sample under laser shock peening and the peak value of shock pressure is $P_0/\tau_{CRSS}=7$ which is the same as that for Al. For Cu$(1\bar{1}0)$ sample, the lattice rotation field is similar with that of Al except the magnitude of rotation is only 2.7°, less than that of Al (4°). This is consistent with the experimental results from Fig. 7(*b*) and it is mainly due to the larger elastic modulus of Cu.

Figure 11(*c*) shows the lattice rotation field of Al (001) sample from simulation. On the shocked surface, the lattice rotation direction and magnitude is almost the same with that of Al$(1\bar{1}0)$ sample and consistent with the experiment result in Fig. 6(*e–f*). However, the predicted lattice rotation field in the depth direction is significantly different than that of Al$(1\bar{1}0)$ in that the affected region in depth direction is about two times deeper than that of Al$(1\bar{1}0)$ and the magnitude of rotation is $\pm 3°$, less than the $\pm 4°$ of Al$(1\bar{1}0)$. The change in sign of lattice rotation which occurs around 25 μm below the surface does not correspond to the transition from a compressive residual stress state to a tensile residual state, which occurs at approximately 80 μm below the surface.

Figure 12 shows the predicted spatial distribution of lattice rotation on the sample surface for Al$(1\bar{1}0)$, Cu$(1\bar{1}0)$, and Al(001) which can be compared with the experiment result from Fig. 6(*b*), (*d*), and (*f*). The lattice rotation distribution is quite similar to the experimental results. When the position changes from left of shock line to the right, the lattice rotation starts from zero degrees (beyond $\pm 40\ \mu$m) to maximum negative value ($-4°$ at $-15\ \mu$m) and after that, the magnitude of lattice rotation decreases to zero again close to the shocked line center. For the right side of shock line center, the distribution is antisymmetric with the left side.

According to the comparison above, it can be seen that the lattice rotation fields under shock peening depend mainly upon crystal orientation. For the same orientation of FCC material such as Al$(1\bar{1}0)$ and Cu$(1\bar{1}0)$, the lattice rotation fields are quite similar except the magnitude of rotation is less for Cu due to the larger elastic modulus and shear strength. If the orientation is different, even in same material such as Al(001) and Al$(1\bar{1}0)$, the lattice rotation fields on shock peened surface is still similar, however, they are quite different in depth direction below the sample surface.

Figure 13(*a*) and (*b*) compares the indentation profiles induced by laser shock peening between FEM simulation and AFM result for the Al$(1\bar{1}0)$ sample. Figure 13(*b*) shows "composite" surface profile from several AFM measurements across the shocked line. As is expected from the approximately incompressible material behavior, significant pile up around the indentation region (see Fig. 13(*b*)) is observed and agrees well with simulation result. Surface ablation is not observed by either SEM or AFM which is probably due to the protective coating layer on the top of samples.

Slip system and yield surface analysis. The plastic slip systems in a face-centered cubic crystal exhibit mirror symmetry about the (110) plane, so that sustained plastic flow under plane strain conditions in the (110) plane is possible as long as a slip system and its mirror image are activated in equivalent amounts. Thus, there are three pairs of effective slip systems that satisfy these conditions as shown in Fig. 14(*a*) and (*b*), [9]. The yield surface which defines graphically the criterion for plastic slip in a stress space with abscissa $(\sigma_{11}-\sigma_{22})/2$ and ordinate σ_{12}, [8], is shown in Fig. 14(*c*). Plastic slip occurs only when the stress state lies on the yield surface with stress increment directed out of the yield surface.

Figure 15(*a–c*) shows the predicted shear strain on each slip system, as well as the total accumulated slip summed over all slip systems for all three samples. It can be seen that the shear strain is close to zero on the top surface for slip system I in (*a*). Since on sample free surface, $\sigma_{12}=0$, the stress state must lie on the abscissa of the yield surface in Fig. 14(*c*). Therefore, plastic deformation near the surface is caused by slip systems II and III so that the shear strain for slip system I is zero near the free surface and only II and III slip systems are active. For the region around 20 μm below the surface, all three slip systems are active. Shear strain simulation on each slip system is shown in Fig. 15(*b*) for Cu$(1\bar{1}0)$ sample also and the result is quite similar with that of Al$(1\bar{1}0)$, except the magnitude of shear strain is 35% less than that of Al. Figure 15(*c*) shows the same simulation for Al (001) sample; it can be seen that the magnitude of shear strain is almost two times larger than that for Al$(1\bar{1}0)$ and the spatial distribution is different in that the affected region is two times deeper than Al$(1\bar{1}0)$. The shear strain in slip system I is also much smaller than other two slip systems.

5 Conclusions

In this study, new experimental methodologies using EBSD and X-ray microdiffraction are employed which enable measurement of the in-plane lattice rotation component of the deformation gradient under plane-strain conditions. The lattice rotation field under laser shock peening is found to be antisymmetric on and below the shock peened single crystal Al and Cu surface. For the Al$(1\bar{1}0)$ sample, the magnitude of rotation is $\pm 3°$ and covers a region around $\pm 35\ \mu$m across the shock line center on peened surface and reaches 40 μm below the surface. For Cu$(1\bar{1}0)$ sample, the magnitude of rotation is $\pm 1.5°$ and the affected region is $\pm 20\ \mu$m on surface and 15 μm below the sample surface. Single crystal plasticity FEM analysis shows an interesting correspondence between the experimental results and theoretical predictions. Lattice rotation fields are quite similar for Al and Cu with the same $(1\bar{1}0)$ orientation and different for Al with (001) orientation. FEM simulation shows only certain slip systems are active on shock peened surface with more active below the surface. Lattice rotation measurements made as a byproduct of residual strain/stress measurements by X-ray micro-diffraction using synchrotron light source also give an indication of the extent of the plastic deformation induced by the microscale laser shock peening.

The experimental methodology and results presented herein set the stage for further study of the microscale laser shock peening process both experimental and computational. It is now possible to systematically measure the extent and character of crystal lattice rotation fields, as well as to measure the induced residual stresses with micron spatial resolution. Thus it is possible to apply these techniques to determine the optimum laser-shock processing parameters (i.e., laser intensity, time of shock, shock spacing, etc.) which induce the maximum residual stress.

Acknowledgments

This work was supported by the National Science Foundation under grant DMI-02-00334. JWK would like acknowledge support by the National Science Foundation under the Faculty Early Career Development (CAREER) Program with grant CMS-0134226. Guidance in X-ray micro-diffraction provided by Dr. I. Cev Noyan and Dr. Jean Jordan-Sweet is appreciated. Assistance in technique details in sample preparation for EBSD provided by Dr. Yongxue Gang and Mr. J. B. Chou is also acknowledged.

References

[1] Johnson, K. L., 1968, "Deformation of A Plastic Wedge by a Rigid Flat Die Under the Action of a Tangential force," J. Mech. Phys. Solids, **16**, pp. 395–402.

[2] Clauer, A. H., and Holbrook, J. H., 1981, "Effects of Laser Induced Shock Waves on Metals," *Shock Waves and High Strain Phenomena in Metals-Concepts and Applications*, Plenum, New York, pp. 675–702.

[3] Zhang, W., and Yao, Y. L., 2000, ''Improvement of Laser Induced Residual Stress Distributions via Shock Waves,'' *Proc. ICALEO'00, Laser Materials Processing*, Laser Institute of America, Orlando, FL, **89**, pp. E183–192.
[4] Zhang, W., and Yao, Y. L., 2000, ''Microscale Laser Shock Processing of Metallic Components,'' ASME J. Manuf. Sci. Eng., **124**(2), pp. 369–378.
[5] Chen, H. Q., and Yao, Y. L., 2003, ''Modeling Schemes, Transiency, and Strain Measurement for Microscale Laser Shock Processing,'' ASME J. Manuf. Sci. Eng., submitted for publication.
[6] Chen, H. Q., Yao, Y. L., and Kysar, J. W., 2003, ''Spatially Resolved Characterization of Residual Stress Induced by Microscale Laser Shock Peening,'' ASME J. Manuf. Sci. Eng., to appear.
[7] Zhang, W., and Yao, Y. L., 2001, ''Feasibility Study of Inducing Desirable Residual Stress Distribution in Laser Micromachining,'' *Transactions of the North American Manufacturing Research Institution of SME (NAMRC XXIX) 2001*, Society of the Manufacturing Engineers, Dearborn, MI, pp. 413–420.
[8] Rice, J. R., 1987, ''Tensile Crack Tip Fields in Elastic-Ideally Plastic Crystals,'' Mech. Mater., **6**, pp. 317–315.
[9] Kysar, J. W., and Briant, C. L., 2002, ''Crack Tip Deformation Fields in Ductile Single Crystals,'' Acta Mater., **50**, pp. 2367–2380.
[10] Randle, V., 1992, *Microtexture Determination and its Applications*, The Institute of Materials, London.
[11] *HKL Channel 5™ User's Manual*, 2001, HKL Technology, Danbury, CT.
[12] Kocks, U. F., 1998, *Texture and Anisotropy*, Cambridge University Press, Cambridge, pp. 44–100.
[13] Cullity, B. D., 1978, *Elements of X-ray Diffraction*, 2nd Ed., Addison-Wesley, London, pp. 13–20.
[14] Hill, R., 1966, ''Generalized Constitutive Relations For Incremental Deformation of Metals By Multislip,'' J. Mech. Phys. Solids, **14**, pp. 95–102.
[15] Rice, R. J., 1977, ''The Localization of Plastic Deformation,'' *Proceedings of 14th International Congress of Theoretical and Applied Mechanics*, **I**., North-Holland, Amsterdam.
[16] Asaro, R. J., 1983, ''Micromechanics of Crystals and Polycrystals,'' Adv. Appl. Mech., **23**, pp. 1–115.
[17] Huang, Y., 1991, ''A User-Material Subroutine Incorporating Single Crystal Plasticity in the ABAQUS Finite Element Program, Mech Report 178,'' Division of Applied Sciences, Harvard University, Cambridge, MA.
[18] Kysar, J. W., 1997. Addendum to ''A User-Material Subroutine Incorporating Single Crystal Plasticity in the ABAQUS Finite Element Program, Mech Report 178,'' Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA.
[19] ABAQUS/Standard User's Manual, 2002. Version 6.2, Hibbit, Karlsson and Sorensen, Inc., Pawtucket, RI.

A Mechanical Model for Low-Gravity Sloshing in an Axisymmetric Tank

M. Utsumi
Machine Element Department,
Technical Research Laboratory,
Ishikawajima-Harima Heavy Industries
Company, Ltd.,
1 Shinnakaharacho,
Isogu-ku, Yokohama 235-8501, Japan

A mechanical model for low-gravity sloshing in an axisymmetric tank is developed using a newly developed slosh analysis method. In this method, spherical coordinates, whose origin is at the top of the cone that is tangent to the tank at the contact line of the meniscus with the tank wall, are used to analytically determine the characteristic functions for an arbitrary axisymmetric tank for which it is customary to resort to numerical methods. By this means, fast and cost-efficient computation can be conducted. Parameters of the mechanical model are determined such that the frequency responses of the resultant force and moment to lateral excitation coincide with those of the actual sloshing system. Influences of the Bond number and the liquid-filling level on the parameters of the mechanical model are examined. [DOI: 10.1115/1.1794700]

1 Introduction

Low-gravity propellant sloshing has received substantial attention in view of its relevance to operations of artificial satellites [1]. In low-gravity sloshing, the surface tension of the liquid plays an overwhelming role in comparison to the gravitational force. There have been a considerable number of studies made on low-gravity slosh dynamics. For a cylindrical tank, analytical investigations were presented for a rigid tank [2–5] and the hydroelastic problem of a flexible tank bottom [6,7]. Sloshing in an arbitrary axisymmetric tank was analyzed using a marker-and-cell method [8], a finite difference method [9,10], a finite element method [11], and a numerical approach based on the Ritz method [12]. Furthermore, experimental studies were conducted by producing low-gravity environments by means of small-scale models [2,13], drop-towers [14], or parabolic flight tests [5].

As mentioned above, the low-gravity sloshing problem for an arbitrary axisymmetric tank was generally solved by numerical methods. However, an ingenious application of curvilinear coordinates for which the Laplace equation is separable leads to a computationally efficient semi-analytical method that allows us to analytically determine the characteristic functions of the liquid motion. Based on this idea, the author developed a new method in previous papers [15,16]. This method uses spherical coordinates, whose origin is at the top of the cone that is tangent to the tank at the contact line of the meniscus with the tank wall, thereby expressing the characteristic functions in terms of the Gaussian hypergeometric function irrespective of the generatrix shape of the tank.

In the previous papers, the response of low-gravity sloshing was analyzed for the case where an arbitrary axisymmetric tank is exposed to lateral [15] and axial [16] excitations. It was confirmed that the present theoretical predictions for the eigenfrequency are in good agreement with the previous theoretical and experimental results [14]. Furthermore, the analytical method was extended to a case in which the shape of the static liquid domain is not axisymmetric as with the sloshing problem for teardrop tanks [17]. The purpose of the present paper is to develop an equivalent mechanical model for low-gravity sloshing in an arbitrary axisymmetric tank subject to lateral excitation. Representation of the slosh dynamics in terms of this model is useful for assessing the dynamic response of artificial satellite because the parameters of the equivalent mechanical model are determined in such a way as to satisfy a dynamic similarity condition that requires the resultant force and moment of the mechanical model be the same as those of the actual sloshing system.

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 Applied Mechanics Division, July 18, 2001; final revision, April 3, 2004. Associate Editor: W. S. Saric. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, 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 Method of Solution

2.1 Computational Model. The sloshing system to be considered is shown in Fig. 1. The tank is subjected to the lateral acceleration $\ddot{f}(t)$ in the x direction. The meanings of the symbols are given in the Nomenclature. Note that the static liquid surface M (meniscus) is curved strongly due to the surface tension effect. The meniscus M reduces to a plane surface under normal gravity. The analysis is performed under the following assumptions: 1) the liquid motion is inviscid, incompressible, and irrotational; 2) the tank is rigid; and 3) the oscillatory displacement of the liquid surface ζ from its equilibrium position M is small enough to be represented within the framework of the linear theory.

2.2 Spherical Coordinates. As shown in Fig. 1, we introduce spherical coordinates R, θ, and φ whose origin O is at the top of the cone that is tangent to the tank wall at the contact line of the meniscus. The origin O is above the tank for $z_C > b$ (Case 1) and below otherwise (Case 2). The liquid surface displacement ζ is considered in the R direction. In terms of the spherical coordinates, the undisturbed and disturbed liquid surfaces and the tank wall can be expressed as

$$M: \quad R = R_M(\theta) \tag{1}$$

$$F: \quad R = R_F(\theta,\varphi,t) = R_M(\theta) + \zeta(\theta,\varphi,t) \tag{2}$$

$$W: \quad R = R_W(\theta) \tag{3}$$

2.3 Variational Principle. The slosh analysis is performed based on a variational principle established as follows. In the absence of surface tension and gas pressure, the liquid pressure p_l gives the Lagrangian per unit volume [18]. For the low-g sloshing problem, we must take into account the potential energy due to the gas pressure p_g and the surface energy associated with the liquid-gas, liquid-solid, and gas-solid interfaces. Thus we obtain the following variational principle:

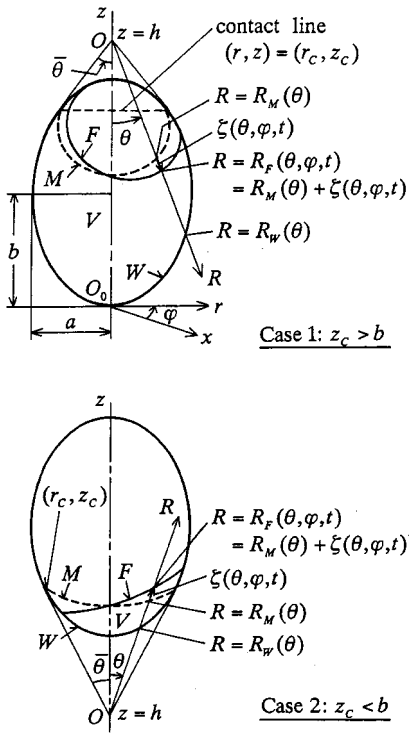


Fig. 1 Axisymmetric tank and coordinate systems

$$\delta\int_{t_1}^{t_2}\Bigg[\int\int\int_V (p_l-p_g)dV-\int\int_F \sigma dF-\int\int_{W_1}\sigma_1 dW_1 - \int\int_{W_2}\sigma_2 dW_2\Bigg]dt=0 \tag{4}$$

The liquid pressure p_l can be expressed in terms of the velocity potential ϕ, describing the liquid motion relative to the moving tank

$$p_l=p_C-\rho_f\Big\{\partial\phi/\partial t+g\varepsilon[R_M(\bar{\theta})\cos\bar{\theta}-R\cos\theta] + R\sin\theta\cos\varphi\,\ddot{f}(t)+\tfrac{1}{2}(\nabla\phi)^2+\dot{G}(t)\Big\} \tag{5}$$

where $\dot{G}(t)$ is an arbitrary time function.

An important work prior to the slosh analysis is to determine the meniscus shape $R_M(\theta)$ [Eq. (1)]. This static analysis can be conducted by reducing Eq. (4) to the principle of virtual work from which governing equations that are equivalent to the Young-Laplace equation and the contact angle condition can be derived. Because the detailed procedure is explained in Ref. [16], it is not presented here and $R_M(\theta)$ is considered to be a known function in this paper.

Substituting Eq. (5) into Eq. (4) and considering the variation with respect to ϕ, ζ, and G leads to [15,16]

$$\int_{t_1}^{t_2}\Bigg\{\rho_f\int\int\int_V\nabla^2\phi\,\delta\phi dV-\rho_f\int\int_W\nabla\phi\cdot\mathbf{N}_W\delta\phi dW - \rho_f\int\int_F\left[\frac{\partial\zeta}{\partial t}\cos(\mathbf{N}_F,R)-\nabla\phi\cdot\mathbf{N}_F\right]\delta\phi dF + \int\int_F(p_g-p_l-\sigma\,\mathrm{div}\,\mathbf{N}_F)\delta\zeta\cos(\mathbf{N}_F,R)dF + \varepsilon\int_C(\sigma\cos\theta_C'+\sigma_1-\sigma_2)\delta\zeta dC - \rho_f\delta G\int\int_F\frac{\partial\zeta}{\partial t}\cos(\mathbf{N}_F,R)dF\Bigg\}dt=0 \tag{6}$$

which yields the field equation (7), the boundary conditions (8)–(11), and the volume constant condition (12),

$$\nabla^2\phi=0 \quad \text{in } V \tag{7}$$

$$\nabla\phi\cdot\mathbf{N}_W=0 \quad \text{on } W \tag{8}$$

$$(\partial\zeta/\partial t)\cos(\mathbf{N}_F,R)-\nabla\phi\cdot\mathbf{N}_F=0 \quad \text{on } F \tag{9}$$

$$p_g-p_l-\sigma\,\mathrm{div}\,\mathbf{N}_F=0 \quad \text{on } F \tag{10}$$

$$\sigma\cos\theta_C'+\sigma_1-\sigma_2=0 \quad \text{along } C \tag{11}$$

$$\int\int_F\frac{\partial\zeta}{\partial t}\cos(\mathbf{N}_F,R)dF=0 \tag{12}$$

Since Eq. (12) can be derived from the other kinematic conditions (7)–(9), Eqs. (7)–(11) may be regarded as basic equations.

Expressing $\mathbf{N}_F$, $\mathbf{N}_W$, dF, dW, dC, and $\cos\theta_C'$ in Eq. (6) in terms of the spherical coordinates, making the linear approximation for the boundary conditions on F using Eq. (2), and using the static equations that are used to determine the meniscus shape, we transform Eq. (6) into

$$\begin{aligned}
&\rho_f\int_0^{2\pi}\int_0^{\bar{\theta}}\varepsilon\int_{R_M}^{R_W}\nabla^2\phi\,\delta\phi R^2\sin\theta dR d\theta d\varphi\\
&-\rho_f\int_0^{2\pi}\int_0^{\bar{\theta}}\varepsilon\Bigg(\left.\frac{\partial\phi}{\partial R}\right|_{R=R_W}-\frac{R_{W\theta}}{R_W^2}\left.\frac{\partial\phi}{\partial\theta}\right|_{R=R_W}\Bigg)\delta\phi|_{R=R_W}R_W^2\sin\theta d\theta d\varphi\\
&+\rho_f\int_0^{2\pi}\int_0^{\bar{\theta}}\varepsilon\Bigg(\left.\frac{\partial\phi}{\partial R}\right|_{R=R_M}-\frac{R_{M\theta}}{R_M^2}\left.\frac{\partial\phi}{\partial\theta}\right|_{R=R_M}-\frac{\partial\zeta}{\partial t}\Bigg)\delta\phi|_{R=R_M}R_M^2\sin\theta d\theta d\varphi\\
&+\int_0^{2\pi}\int_0^{\bar{\theta}}\Bigg\{\varepsilon\rho_f\left.\frac{\partial\phi}{\partial t}\right|_{R=R_M}+\varepsilon\rho_f R_M\sin\theta\cos\varphi\,\ddot{f}(t)\\
&-\rho_f g\zeta\cos\theta-\sigma\Bigg[S_{1M}(\theta)\zeta+S_{2M}(\theta)\frac{\partial\zeta}{\partial\theta}+S_{3M}(\theta)\frac{\partial^2\zeta}{\partial\theta^2}\\
&+S_{4M}(\theta)\frac{\partial^2\zeta}{\partial\varphi^2}\Bigg]\Bigg\}\delta\zeta R_M^2\sin\theta d\theta d\varphi-\int_0^{2\pi}\sigma\Bigg[R_M(R_M^2\\
&+R_{M\theta}^2)^{-3/2}\Bigg(R_M\frac{\partial\zeta}{\partial\theta}-R_{M\theta}\zeta\Bigg)\Bigg]_{\theta=\bar{\theta}}\delta\zeta|_{\theta=\bar{\theta}}R_M(\bar{\theta})\sin\bar{\theta}d\varphi=0,
\end{aligned} \tag{13}$$

where $S_{iM}(\theta)$ $(i=1-4)$ are functions depending on the meniscus shape $R_M(\theta)$. These are listed in Eq. (33) of Ref. [16].

2.4 Modal Equation. Because the Laplace equation (7) is separable for the spherical coordinates, ϕ and ζ can be analytically expressed as follows for the free vibration analysis:

$$\phi(R,\theta,\varphi,t)=i\omega\bar{\phi}(R,\theta,\varphi)e^{i\omega t} \tag{14}$$

$$\zeta(\theta,\varphi,t)=\bar{\zeta}(\theta,\varphi)e^{i\omega t} \tag{15}$$

with

$$\bar{\phi}(R,\theta,\varphi)=\sum_{k=1}^{\infty}\left\{a_k\left(\frac{R}{l_a}\right)^{\alpha_{1k}}+b_k\left(\frac{R}{l_b}\right)^{\alpha_{2k}}\right\}\Theta_k(\theta)\cos\varphi \tag{16}$$

$$\overline{\zeta}(\theta,\varphi)=\sum_{k=1}^{\infty} c_k\Theta_k(\theta)\cos\varphi \tag{17}$$

where a_k, b_k, and c_k are unknown constants; l_a and l_b are normalization parameters; α_{1k} and α_{2k} are characteristic exponents related to the separation variable λ by $\alpha(\alpha+1)=\lambda$, i.e.,

$$\alpha_{1k}=\tfrac{1}{2}[-1-(1+4\lambda_k)^{1/2}],\quad \alpha_{2k}=\tfrac{1}{2}[-1+(1+4\lambda_k)^{1/2}]; \tag{18}$$

and $\Theta_k(\theta)$ is the characteristic function expressed in terms of the Gaussian hypergeometric series F as

$$\Theta_k(\theta)=\sin\theta F[1-\alpha_{1k},\ \alpha_{1k}+2,\ 2,\ (1-\cos\theta)/2] \tag{19}$$

The characteristic value λ_k is determined such that

$$d\Theta_k/d\theta=0 \quad \text{at} \quad \theta=\overline{\theta} \tag{20}$$

Substituting Eqs. (14) and (15) into Eq. (13) and applying the Galerkin method leads to algebraic equations for a_k, b_k, and c_k. These equations can be reduced to an eigenvalue problem from which eigenfrequencies and mode shape functions can be determined. In terms of the fundamental mode shape functions, we express ϕ and ζ as

$$\phi(R,\theta,\varphi,t)=\dot{q}(t)\overline{\phi}(R,\theta,\varphi) \tag{21}$$

$$\zeta(\theta,\varphi,t)=q(t)\overline{\zeta}(\theta,\varphi) \tag{22}$$

where $q(t)$ is the modal coordinate. Substituting Eqs. (21) and (22) into Eq. (13) and considering the variation with respect to $q(t)$ leads to the modal equation of the form

$$\ddot{q}+\omega^2 q=\beta\ddot{f}(t) \tag{23}$$

2.5 Equivalent Mechanical Model. To develop the equivalent mechanical model, we first express the slosh force and moment exerted to the tank wall in terms of the modal coordinate $q(t)$. The slosh force due to the fluid pressures p_l and p_g can be calculated by taking the dynamical component of the force

$$\overline{F}_x=\int\int_{W_1} p_l(\mathbf{N}_W\cdot\mathbf{e}_x)dW_1+\int\int_{W_2} p_g(\mathbf{N}_W\cdot\mathbf{e}_x)dW_2 \tag{24}$$

That is

$$\overline{F}_{x,dy}=\int\int_{W_{1,st}} p_{l,dy}(\mathbf{N}_W\cdot\mathbf{e}_x)dW_{1,st}+\int\int_{W_{1,dy}} p_{l,st}\times(\mathbf{N}_W\cdot\mathbf{e}_x)dW_{1,dy}+\int\int_{W_{2,dy}} p_g(\mathbf{N}_W\cdot\mathbf{e}_x)dW_{2,dy} \tag{25}$$

where $W_{1,dy}$ and $W_{2,dy}$ are the variations in W_1 and W_2 respectively, due to the dynamical liquid surface displacement. Hence, the surface elements of $W_{1,dy}$ and $W_{2,dy}$ can be expressed as

$$dW_{1,dy}=-dW_{2,dy}=-\varepsilon\zeta dC_{st} \tag{26}$$

Substituting Eqs. (5) and (26) into Eq. (25) yields

$$\overline{F}_{x,dy}=\int\int_{W_{1,st}}(-\rho_f)\left[\frac{\partial\phi}{\partial t}+x\ddot{f}(t)\right](\mathbf{N}_W\cdot\mathbf{e}_x)dW_{1,st}+\varepsilon\int_{C_{st}}(p_g-p_C)\zeta(\mathbf{N}_W\cdot\mathbf{e}_x)dC_{st} \tag{27}$$

The slosh moment induced by the fluid pressures p_l and p_g can be calculated likewise. In addition to these fluid pressures, we should take into account the surface tension force vector applied along the moving contact line C. This force vector is perpendicular to the contact line C and parallel to the oscillating liquid surface F, having the magnitude σ. Thus, the slosh force and moment exerted on the tank wall are calculated by

$$F_x=F_{x1}+F_{x2}+F_{x3} \tag{28}$$

$$M_y=M_{y1}+M_{y2}+M_{y3} \tag{29}$$

where

$$F_{x1}=\int\int_{W_{1,st}}(-\rho_f)\left[\frac{\partial\phi}{\partial t}+x\ddot{f}(t)\right](\mathbf{N}_W\cdot\mathbf{e}_x)dW_{1,st} \tag{30}$$

$$M_{y1}=\int\int_{W_{1,st}}(-\rho_f)\left[\frac{\partial\phi}{\partial t}+x\ddot{f}(t)\right][(\mathbf{N}_W\cdot\mathbf{e}_x)z-(\mathbf{N}_W\cdot\mathbf{e}_z)x]dW_{1,st} \tag{31}$$

$$F_{x2}=\varepsilon\int_{C_{st}}(p_g-p_C)\zeta(\mathbf{N}_W\cdot\mathbf{e}_x)dC_{st} \tag{32}$$

$$M_{y2}=\varepsilon\int_{C_{st}}(p_g-p_C)\zeta[(\mathbf{N}_W\cdot\mathbf{e}_x)z-(\mathbf{N}_W\cdot\mathbf{e}_z)x]dC_{st} \tag{33}$$

$$F_{x3}=\int_C\sigma\frac{\mathbf{N}_F\times(\mathbf{N}_F\times\mathbf{N}_W)}{|\mathbf{N}_F\times(\mathbf{N}_F\times\mathbf{N}_W)|}\cdot\mathbf{e}_x dC \tag{34}$$

$$M_{y3}=\int_C\sigma\frac{\mathbf{N}_F\times(\mathbf{N}_F\times\mathbf{N}_W)}{|\mathbf{N}_F\times(\mathbf{N}_F\times\mathbf{N}_W)|}\cdot(\mathbf{e}_x z-\mathbf{e}_z x)dC \tag{35}$$

Using Eqs. (21) and (22) and expressing the normal vectors in terms of the spherical coordinates using Eqs. (2) and (3), we express each component of the slosh force and moment in terms of the modal coordinate $q(t)$ as

$$F_{x1}=A_1\ddot{q}(t)+B_1\ddot{f}(t) \tag{36}$$

$$M_{y1}=C_1\ddot{q}(t)+D_1\ddot{f}(t) \tag{37}$$

$$F_{x2}=A_2 q(t) \tag{38}$$

$$M_{y2}=C_2 q(t) \tag{39}$$

$$F_{x3}=A_3 q(t) \tag{40}$$

$$M_{y3}=C_3 q(t) \tag{41}$$

For brevity, constants A_1-A_3, B_1, C_1-C_3, and D_1 are not presented here. When surface tension is not present, we have

$$A_2=A_3=C_2=C_3=0 \tag{42}$$

because the components of the slosh force and moment F_{x2}, F_{x3}, M_{y2}, and M_{y3} do not arise.

Substituting the solution to the modal equation (23) for the sinusoidal excitation

$$\ddot{f}(t)=\sin\omega_f t \tag{43}$$

into Eqs. (36)–(41) and using Eqs. (28) and (29) leads to

$$F_x=\frac{A_1\beta\omega_f^2+B_1(\omega_f^2-\omega^2)-(A_2+A_3)\beta}{\omega_f^2-\omega^2}\sin\omega_f t \tag{44}$$

$$M_y=\frac{C_1\beta\omega_f^2+D_1(\omega_f^2-\omega^2)-(C_2+C_3)\beta}{\omega_f^2-\omega^2}\sin\omega_f t \tag{45}$$

On the other hand, for the mechanical model as shown in Fig. 2, the equation of motion of the slosh mass and the resultant force and moment are given by

$$m_1\ddot{u}+k_1 u=-m_1\ddot{f}(t) \tag{46}$$

$$F_{x,mech}=k_1 u-m_0\ddot{f}(t) \tag{47}$$

$$M_{y,mech}=k_1 l_1 u-m_0 l_0\ddot{f}(t)+m_1 g u \tag{48}$$

The responses of the force $F_{x,mech}$ and the moment $M_{y,mech}$ to the sinusoidal excitation (43) are

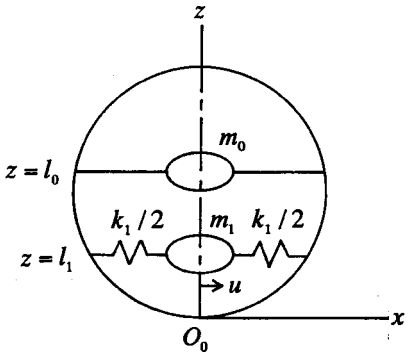


Fig. 2 Mechanical model

$$F_{x,mech}=\frac{k_1-m_0(\omega_f^2-\omega_{mech}^2)}{\omega_f^2-\omega_{mech}^2}\sin\omega_f t \tag{49}$$

$$M_{y,mech}=\frac{k_1 l_1-m_0 l_0(\omega_f^2-\omega_{mech}^2)+m_1 g}{\omega_f^2-\omega_{mech}^2}\sin\omega_f t \tag{50}$$

where

$$\omega_{mech}=(k_1/m_1)^{1/2} \tag{51}$$

The parameters of the mechanical model can be determined from the dynamic similarity condition that requires

$$F_{x,mech}=F_x \tag{52a}$$

$$M_{y,mech}=M_y \tag{52b}$$

for any frequency ω_f of the excitation. First, from Eq. (52*a*), the fixed and slosh masses and the spring constant are determined as

$$m_0=-A_1\beta-B_1 \tag{53}$$

$$m_1=A_1\beta-\frac{1}{\omega^2}(A_2+A_3)\beta \tag{54}$$

$$k_1=m_1\omega^2 \tag{55}$$

and then from Eq. (52*b*), the z coordinates of the slosh and fixed masses are determined by

$$l_0=\frac{C_1\beta+D_1}{-m_0} \tag{56}$$

$$l_1=\frac{-m_0 l_0\omega^2-D_1\omega^2-(C_2+C_3)\beta-m_1 g}{k_1} \tag{57}$$

From Eqs. (53) and (54), the sum of the fixed and slosh masses is

$$m_0+m_1=-B_1-\frac{1}{\omega^2}(A_2+A_3)\beta \tag{58}$$

Equations (30) and (36) indicate that the parameter B_1 is given by

$$B_1=-\rho_f\int\int_{W_{1,st}} x(\mathbf{N}_W\cdot\mathbf{e}_x)dW_{1,st} \tag{59}$$

By using the divergence theorem, Eq. (59) can be transformed into

$$B_1=-\rho_f\int\int\int_{V_0}\mathrm{div}(x\mathbf{e}_x)dV_0+\rho_f\int\int_S x(\mathbf{N}_S\cdot\mathbf{e}_x)dS$$

$$=-\rho_f V_0+\rho_f\int\int_S x(\mathbf{N}_S\cdot\mathbf{e}_x)dS \tag{60}$$

where S is the plane $z=z_C$, including the static contact line (see Fig. 1), $\mathbf{N}_S$ is the outer unit normal vector of the plane S, and V_0 is the domain bounded by the tank wall $W_{1,st}$ and the plane S. Since $\mathbf{N}_S\cdot\mathbf{e}_x=0$, Eq. (60) becomes

$$B_1=-\rho_f V_0 \tag{61}$$

which transforms Eq. (58) into

$$m_0+m_1=\rho_f V_0-\frac{1}{\omega^2}(A_2+A_3)\beta \tag{62}$$

For the case $\mathrm{Bo}\to\infty$, $\rho_f V_0$ is equal to the liquid mass $\rho_f V$ because the plane S coincides with the liquid surface. Furthermore, the second term on the right-hand side of Eq. (62) vanishes, as can be seen from Eq. (42). Hence, the sum of the slosh and fixed masses is equal to the liquid mass.

On the other hand, for finite Bond numbers, $\rho_f V_0$ is larger than the liquid mass and the following discussion can be made. First, when the frequency ω_f of the acceleration (43) of the tank is lower than the eigenfrequency ω of the sloshing, the slosh force component F_{x1} is out of phase with the acceleration (43) of the tank. Hence, it can be seen from Eq. (44) that

$$A_1\beta>0 \tag{63}$$

Second, the dynamic liquid pressure $-\rho_f\partial\phi/\partial t$ and the outward liquid surface displacement $-\varepsilon\zeta$ are in phase for an arbitrary φ, and p_g is larger than p_C. This fact renders the force components F_{x1} and F_{x2} given by Eqs. (30) and (32) out of phase (the phase difference is 180°) to one another. Hence, from Eqs. (44) and (63), we can see that

$$-A_2\beta<0 \tag{64}$$

Third, the dynamical component of F_{x3} is in phase with the φ derivative of the outward liquid surface displacement $-\varepsilon\partial\zeta/\partial\varphi$ at $\varphi=270°$, which is in phase with F_{x1}. Therefore, it is revealed from Eqs. (44) and (63) that

$$-A_3\beta>0 \tag{65}$$

It can be expected that the combined effect of F_{x2} and F_{x3} satisfies $-(A_2+A_3)\beta<0$ and consequently makes the sum of the fixed and slosh masses given by Eq. (62) close to the liquid mass. In fact, this conservation of mass is satisfied by the values of the fixed and slosh masses numerically computed based on the present theory.

3 Numerical Results

The numerical computation is carried out using dimensionless quantities normalized by the characteristic length b, mass $\rho_f b^3$, and frequency ω_{ch}. The Bond number defined by

$$\mathrm{Bo}=\frac{\rho_f g b^2}{\sigma} \tag{66}$$

is used as a dimensionless parameter relating the magnitude of gravity to surface tension. The characteristic frequency is defined by

$$\omega_{ch}=(g/b)^{1/2}\ \text{ for }\ \mathrm{Bo}\neq 0 \tag{67a}$$

$$\omega_{ch}=(\sigma/\rho_f b^3)^{1/2}\ \text{ for }\ \mathrm{Bo}=0 \tag{67b}$$

Figure 3 shows the dependence of the slosh mass on the Bond number and the liquid-filling level for the case of a spherical tank and a contact angle of 5° between meniscus and tank wall. It can be seen from Fig. 3 that when the Bond number is decreased for a given liquid-filling level, the slosh mass decreases. For high liquid-filling levels, the slosh mass is small because the liquid surface M is narrow and results in small kinematic energy of the liquid relative to the tank given by

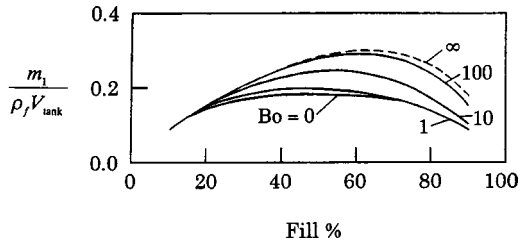


Fig. 3 Dimensionless slosh mass

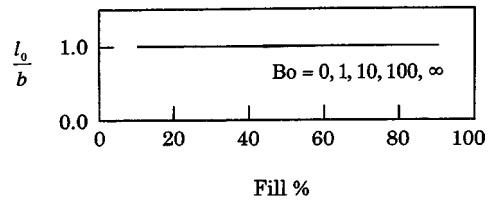


Fig. 4 Dimensionless position of fixed mass

$$\frac{1}{2}\int\int\int_V \cdot(\nabla\phi)^2 dV = -\frac{1}{2}\int\int_M \phi(\nabla\phi\cdot\mathbf{N}_M)dM + \frac{1}{2}\int\int_W \phi(\nabla\phi\cdot\mathbf{N}_W)dW \qquad (68)$$

where the integral over the tank wall W is zero in Eq. (8). The small values of the slosh mass for low liquid-filling levels are evident because the total liquid mass is small. Thus, the slosh mass is small for both high and low liquid-filling levels and consequently exhibits a maximum value at a certain intermediate liquid-filling level.

Another significant observation that can be made from Fig. 3 is that the liquid-filling level yielding the maximum slosh mass is lower for the case of finite Bond numbers than for the case Bo $\rightarrow\infty$, in which the surface tension effect is not taken into account. The reason for this can be discussed as follows. The magnitude of the dynamic liquid pressure applied along the generatrix $\varphi=0$ of the tank wall is large near the contact point $z=z_C$. Hence, the magnitude of the force F_{x1} given by Eq. (30) is greatly influenced by the direction cosine $\mathbf{N}_W\cdot\mathbf{e}_x$ of the unit normal vector $\mathbf{N}_W$ at the contact point $z=z_C$. This direction cosine is maximal at $z_C=1$, which corresponds to 50% filling level for Bo$\rightarrow\infty$. The force F_{x1} for Bo$\rightarrow\infty$ is maximum at somewhat higher 65% filling level, because an increase in the liquid-filling level extends the area of the tank wall subjected to the liquid pressure. When the Bond number is decreased while keeping the same filling level 65%, the contact point goes up reducing the value of $\mathbf{N}_W\cdot\mathbf{e}_x$ at the contact point and, therefore, the magnitude of the force F_{x1} decreases. Consequently, the liquid-filling level yielding the maximum magnitude of F_{x1} becomes lower. Also, the variations in the amplitudes of the force components F_{x2} and F_{x3} with the liquid-filling level exhibit a tendency similar to the dependence of F_{x1} on the liquid-filling level due to the following factors:

a. $\mathbf{N}_W\cdot\mathbf{e}_x$ that appears in Eq. (32) is maximum at $z_C=1$.

b. As can be seen from Eqs. (32) and (34), F_{x2} and F_{x3} are influenced by the length of the contact line and the amplitude of the liquid surface displacement ζ at the contact point; the former is maximum at $z_C=1$ while the latter increases as the liquid-filling level increases [15].

As a result, the amplitudes of the force components F_{x1}, F_{x2}, and F_{x3} are maximum at nearly the same liquid-filling level. Hence, although F_{x2} is out-of-phase with F_{x1} and F_{x3}, the total slosh force is maximum at a liquid-filling level that is lower than the liquid-filling level giving the maximum magnitude of F_{x1} for Bo$\rightarrow\infty$.

Figure 4 shows the z coordinate of the fixed mass. It can be seen from Fig. 4 that the z coordinate l_0 of the fixed mass is equal to the radius b of the tank irrespective of the Bond number and the liquid-filling level. This is due to the fact that for a spherical tank, the following relation holds.

$$(\mathbf{N}_W\cdot\mathbf{e}_x)z-(\mathbf{N}_W\cdot\mathbf{e}_z)x=b(\mathbf{N}_W\cdot\mathbf{e}_x) \qquad (69)$$

This relation results in

$$C_1=bA_1, \quad D_1=bB_1 \qquad (70)$$

as can be seen from Eqs. (30), (31), (36), and (37). Substituting Eqs. (53) and (70) into Eq. (56) leads to $l_0=b$.

Figure 5 shows the z coordinate l_1 of the slosh mass. For nonzero Bond numbers, the variation of l_1 with the liquid-filling level can be explained by the dependence of the dimensionless eigenfrequency $\omega/(g/b)^{1/2}$ on the liquid-filling level (Fig. 6) as follows. By using Eqs. (53)–(56), Eq. (57) can be transformed into

$$\frac{l_1}{b}=\frac{1}{b}\frac{\omega^2C_1-(C_2+C_3)}{\omega^2A_1-(A_2+A_3)}-\frac{1}{\omega^2/(g/b)} \qquad (71)$$

For the case Bo$\rightarrow\infty$, in which the surface tension effect is not taken into account at all, the relations (42) and (70) transform Eq. (71) into

$$\frac{l_1}{b}=1-\frac{1}{\omega^2/(g/b)} \qquad (72)$$

The dimensionless eigenfrequency $\omega/(g/b)^{1/2}$ increases with increasing liquid-filling level, as can be seen from Fig. 6. Therefore, l_1/b increases with the increase of the liquid-filling level. For finite Bond numbers, the variations of l_1/b with the liquid-filling level determined from Eqs. (71) and (72) exhibit a similar tendency because the relation $C_2=bA_2$ holds [see Eqs. (32), (33), (38), (39), and (69)] and C_3 does not largely differ from bA_3 for small contact angles for the following reason. The vector $\mathbf{N}_F\times(\mathbf{N}_F\times\mathbf{N}_W)$ appearing in Eqs. (34) and (35) can be expressed in terms of the linear combination of two vectors that are exactly and approximately proportional to $\mathbf{N}_W$ as

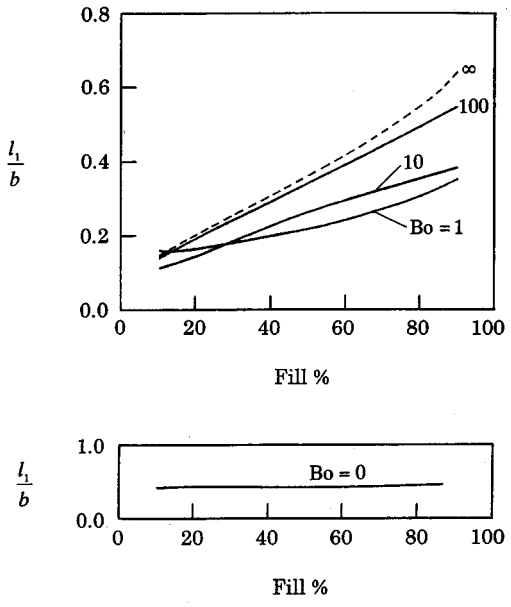


Fig. 5 Dimensionless position of slosh mass

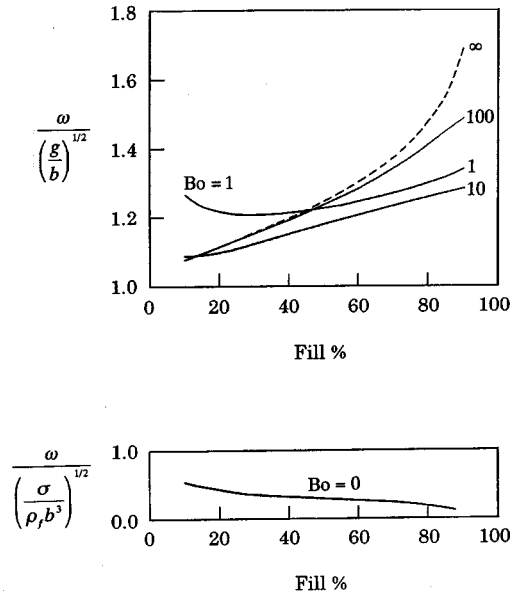


Fig. 6 Dimensionless eigenfrequency

$$\mathbf{N}_F\times(\mathbf{N}_F\times\mathbf{N}_W)=(\mathbf{N}_F\cdot\mathbf{N}_W)\mathbf{N}_F-(\mathbf{N}_F\cdot\mathbf{N}_F)\mathbf{N}_W=\mathbf{N}_F\cos\theta_C-\mathbf{N}_W \tag{73}$$

and the magnitude of this vector assumes a constant value sin θ_C, as can be readily seen from a geometrical consideration.

For the case of Bo$=0$, on the other hand, the second term on the right-hand side of Eq. (71) vanishes. This indicates that the variation of l_1/b with the liquid-filling level is not strongly influenced by the dependence of the dimensionless eigenfrequency on the liquid-filling level shown in Fig. 6. Therefore, l_1/b for Bo $=0$ is almost constant, as can be seen from Fig. 5.

4 Conclusions

A mechanical model for low-gravity sloshing in an axisymmetric tank has been developed in this paper. The parameters of the mechanical model were determined such that the resultant force and moment of the mechanical model are the same as the slosh force and moment. In evaluating the slosh force and moment, dynamical forces due to the pressure difference across the liquid-gas interface and the surface tension force applied along the moving contact line were taken into account in addition to the dynamical liquid pressure. Numerical results were presented for a spherical tank. The numerical results show that: 1) when the Bond number is decreased for a given liquid-filling level, the slosh mass decreases; 2) the liquid-filling level yielding the maximum slosh mass is lower for finite Bond numbers than for infinite Bond number; 3) the height of the fixed mass is equal to the radius of the tank; and 4) for nonzero Bond numbers, the variation of the height of the slosh mass with the liquid-filling level can be explained by the dependence of the eigenfrequency on the liquid-filling level, whereas for zero Bond number, the height of the slosh mass is almost constant since it is not strongly influenced by the dependence of the eigenfrequency on the liquid-filling level.

Nomenclature

A_1-A_3, B_1 = constants in slosh force [Eqs. (36), (38), and (40)]
Bo = Bond number [Eqs. (66)]
b = half height of tank (characteristic length, Fig. 1)
C = contact line
C_1-C_3, D_1 = constants in slosh moment [Eqs. (37), (39), and (41)]
$\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ = unit vectors in x, y, and z directions
F = disturbed liquid surface (Fig. 1)
$\ddot{f}(t)$ = acceleration of tank in x direction
g = gravitational acceleration
l_0, l_1 = z coordinates of fixed and slosh masses, respectively (Fig. 2)
M = meniscus (undisturbed liquid surface, Fig. 1)
m_0, m_1 = fixed and slosh masses respectively (Fig. 2)
$\mathbf{N}_F$ = unit normal vector of F pointing into liquid domain
$\mathbf{N}_M$ = unit normal vector of M pointing into liquid domain
$\mathbf{N}_W$ = unit normal vector of W pointing outwards from liquid domain
p_C = static liquid pressure at contact line
p_g = gas pressure
p_l = liquid pressure
$q(t)$ = modal coordinate
R, θ, φ = spherical coordinates (Fig. 1)
$R_F(\theta,\varphi,t)$ = function expressing shape of disturbed liquid surface F (Fig. 1)
$R_M(\theta)$ = function expressing shape of meniscus M (Fig. 1)
$R_{M\theta}$, $R_{M\theta\theta}$ = $dR_M/d\theta$, $d^2R_M/d\theta^2$
$R_W(\theta)$ = function expressing shape of tank wall (Fig. 1)
$R_{W\theta}$ = $dR_W/d\theta$
V = liquid domain (Fig. 1)
W = tank wall (Fig. 1)
W_1, W_2 = liquid-solid and gas-solid interfaces, respectively
β = coefficient in modal equation (23)
ε = 1 and -1, respectively, for Cases 1 and 2 (Fig. 1)
ζ = liquid surface displacement (Fig. 1)
$\bar{\theta}$ = maximum value of θ (Fig. 1)
θ_C = contact angle between meniscus and tank wall
θ_C' = contact angle between disturbed liquid surface and tank wall
ρ_f = liquid density
σ, σ_1, σ_2 = surface energy per unit area associated with liquid-gas, liquid-solid, and gas-solid interfaces, respectively
ϕ = velocity potential describing the liquid motion relative to the tank
ω = eigenfrequency of sloshing
ω_f = excitation frequency
ω_{mech} = eigenfrequency of mechanical model

References

[1] Abramson, H. N., ed, 1966, "The Dynamic Behavior of Liquids in Moving Containers," NASA SP-106.
[2] Dodge, F. T., and Garza, L. R., 1967, "Experimental and Theoretical Studies of Liquid Sloshing at Simulated Low Gravity," ASME J. Appl. Mech., **34**, pp. 555–562.
[3] Bauer, H. F., and Eidel, W., 1990, "Linear Liquid Oscillations in Cylindrical Container Under Zero-Gravity," Appl. Microgravity Technol., **2**, pp. 212–220.
[4] Satterlee, H. M., and Reynolds, W. C., 1964, "The Dynamics of the Free Liquid Surface in Cylindrical Containers Under Strong Capillary and Weak Gravity Conditions," Technical Report LG-2, Dept. of Mech. Eng., Stanford University, Stanford, CA.
[5] Peterson, L. D., Crawley, E. F., and Hansman, R. J., 1989, "Nonlinear Fluid Slosh Coupled to the Dynamics of a Spacecraft," AIAA J., **27**, pp. 1230–1240.
[6] Bauer, H. F., and Eidel, W., 1993, "Hydroelastic Vibrations in a Circular Cylindrical Container with a Flexible Bottom in Zero-Gravity," J. Fluids Struct., **7**, pp. 783–802.
[7] Tong, P., 1967, "Liquid Motion in a Circular Cylindrical Container With a Flexible Bottom," AIAA J., **5**, pp. 1842–1848.
[8] Hung, R. J., and Lee, C. C., 1992, "Similarity Rules in Gravity Jitter-Related Spacecraft Liquid Propellant Slosh Waves Excitation," J. Fluids Struct., **6**, pp. 493–522.
[9] Chu, W. H., 1970, "Low-Gravity Fuel Sloshing in an Arbitrary Axisymmetric Rigid Tank," ASME J. Appl. Mech., **37**, pp. 828–837.
[10] Concus, P., Crane, G. E., and Satterlee, H. M., 1969, "Small Amplitude Lateral

Sloshing in Spheroidal Containers Under Low Gravitational Conditions,'' NASA CR-72500.
[11] Dodge, F. T., Green, S. T., and Cruse, M. W., 1991, ''Analysis of Small-Amplitude Low Gravity Sloshing in Axisymmetric Tanks,'' Microgravity Sci. Technol., **4**, pp. 228–234.
[12] Yeh, C. K., 1967, ''Free and Forced Oscillations of a Liquid in an Axisymmetric Tank at Low-Gravity Environments,'' ASME J. Appl. Mech., **34**, pp. 23–28.
[13] Dodge, F. T., and Garza, L. R., 1970, ''Simulated Low-gravity Sloshing in Spherical, Ellipsoidal, and Cylindrical Tanks,'' J. Spacecr. Rockets, **7**, pp. 204–206.
[14] Coney, T. A., and Salzman, J. A., 1971, ''Lateral Sloshing in Oblate Spheroidal Tanks Under Reduced- and Normal-Gravity Conditions,'' NASA TN D-6250.
[15] Utsumi, M., 1998, ''Low-Gravity Propellant Slosh Analysis Using Spherical Coordinates,'' J. Fluids Struct., **12**, pp. 57–83.
[16] Utsumi, M., 2000, ''Low-gravity Sloshing in an Axisymmetrical Container Excited in the Axial Direction,'' ASME J. Appl. Mech., **67**, pp. 344–354.
[17] Utsumi, M., 2000, ''Development of Mechanical Models for Propellant Sloshing in Teardrop Tanks,'' J. Spacecr. Rockets, **37**, pp. 597–603.
[18] Seliger, R. L., and Whitham, G. B., 1968, ''Variational Principles in Continuum Mechanics,'' Proc. R. Soc. London, Ser. A, **305**, pp. 1–25.

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 2500 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.

Volumetric Constraint Models for Anisotropic Elastic Solids

Carlos A. Felippa
Department of Aerospace Engineering Sciences and Center for Aerospace Structures, University of Colorado, Boulder, CO 80309-0429 Mem. ASME

Eugenio Oñate
International Center for Numerical Methods in Engineering (CIMNE), Edificio C-1, c. Gran Capitán s/n, Universidad Politécnica de Cataluña, Campus Norte UPC, 08034 Barcelona, Spain

We study three "incompressibility flavors" of linearly elastic anisotropic solids that exhibit volumetric constraints: isochoric, hydroisochoric, and rigidtropic. An isochoric material deforms without volume change under any stress system. An hydroisochoric material does so under hydrostatic stress. A rigidtropic material undergoes zero deformations under a certain stress pattern. Whereas the three models coalesce for isotropic materials, important differences appear for anisotropic behavior. We find that isochoric and hydroisochoric models under certain conditions may be hampered by unstable physical behavior. Rigidtropic models can represent semistable physical materials of arbitrary anisotropy while including isochoric and hydroisochoric behavior as special cases. [DOI: 10.1115/1.1748318]

1 Introduction

An incompressible linearly elastic isotropic solid does not deform under hydrostatic stress. It does not change volume under pressure. Since deviatoric and volumetric deformations uncouple, no volume change occurs under any stress state. The three volumetric constraints just stated coalesce, and it is sufficient to qualify the material as incompressible.

A more careful study is necessary for anisotropic materials. In the present note we examine three volumetric constraint models of a linearly elastic anisotropic solid. The following definitions are used for that examination:

1. A material is called *rigidtropic* if it does not deform (i.e., experiences zero strains) under a specific stress pattern, which is a null eigenvector of the strain-stress (compliance) matrix. The term "rigidtropic" is used in the sense of "rigidity in a certain way" as defined by that eigenvector.
2. A material is called *isochoric* if it does not change volume under any applied stress system ([1], Sec. 77). Alternatively: the volumetric strain is zero under any stress state.
3. A material is called *hydroisochoric* if it is isochoric under hydrostatic stress. Isochoric materials are hydroisochoric but the converse is not necessarily true.

As noted the three models coalesce for an isotropic material. For an arbitrary anisotropic solid, however, it will be shown that imposing a isochoric or hydroisochoric constraint may produce a compliance matrix that has at least one negative eigenvalue. This means that under some stress system the material is able to create energy, contradicting the laws of thermodynamics. Such model cannot represent a physically stable material. On the other hand, for rigidtropic behavior it is easier to control material stability for any type of anisotropy because constraints are posed directly on the spectral form.

2 Compliance Relations

We consider a linearly elastic anisotropic solid in three dimensions referred to axes $\{x_i\}$. Stresses σ_{ij} and strains e_{ij} will be arranged as six-component column vectors constructed from the respective tensors through the usual conventions of structural mechanics:

$$\boldsymbol{\sigma}=[\sigma_{11}\ \sigma_{22}\ \sigma_{33}\ \sigma_{12}\ \sigma_{23}\ \sigma_{31}]^T,$$
$$\mathbf{e}=[e_{11}\ e_{22}\ e_{33}\ 2e_{12}\ 2e_{23}\ 2e_{31}]^T. \tag{1}$$

The strain-stress constitutive equations in matrix notation are

$$\mathbf{e}=\begin{bmatrix} e_{11}\\ e_{22}\\ e_{33}\\ 2e_{23}\\ 2e_{31}\\ 2e_{12}\end{bmatrix}=\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16}\\ & C_{22} & C_{23} & C_{24} & C_{25} & C_{26}\\ & & C_{33} & C_{34} & C_{35} & C_{36}\\ & & & C_{44} & C_{45} & C_{46}\\ & & & & C_{55} & C_{56}\\ \text{symm} & & & & & C_{66}\end{bmatrix}\begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{31}\\ \sigma_{12}\end{bmatrix}$$
$$=\mathbf{C}\boldsymbol{\sigma}. \tag{2}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, August 9, 2002, final revision, February 15, 2004. Associate Editor: K. R. Rajagopal.

Here C_{ij} are compliance coefficients arranged into the symmetric compliance matrix $\mathbf{C}$. All diagonal entries C_{ii} are assumed to be nonnegative with a positive sum. The compliance matrix is called *stable*, *semistable*, or *unstable* if $\mathbf{C}$ is positive definite, positive semidefinite, or indefinite, respectively. In the semistable case it will be assumed that $\mathbf{C}$ has a rank deficiency of at most one to simplify the analysis. The eigenvalues of $\mathbf{C}$ are γ_i for $i=1,2\ldots 6$, with $\mathbf{v}_i$ being the corresponding eigenvector normalized to length $\sqrt{3}$. (This nonstandard normalization simplifies linking up to the hydrostatic stress vector in Sections 4ff.) Accordingly the spectral decomposition is

$$\mathbf{C}=\frac{1}{3}\sum_{i=1}^{6}\gamma_i\mathbf{v}_i\mathbf{v}_i^T, \quad \mathbf{v}_i^T\mathbf{v}_j=3\delta_{ij}, \tag{3}$$

where δ_{ij} is the Kronecker delta. The eigenvalues will be arranged so that $\gamma_1=\gamma_{\min}$ is the algebraically smallest one and $\gamma_6=\gamma_{\max}$ the maximum. For stable or semistable models, $\gamma_1\geqslant 0$ and $\gamma_j>0$ for $j=2,\ldots 6$.

If $\gamma_1=0$ the material is rigidtropic according to the definition given in the Introduction, with $\mathbf{v}_1$ defining the corresponding stress pattern. The volumetric strain is $e_v=e_{11}+e_{22}+e_{33}$. Isochoric behavior is mathematically characterized by $e_v=0$ under any $\boldsymbol{\sigma}$. Hydroisochoric behavior means that $e_v=0$ under $\boldsymbol{\sigma}_p=p[1\ 1\ 1\ 0\ 0\ 0]^T$ for any p. These constraints are mathematically expressed in terms of $\mathbf{C}$ as follows:

$$\begin{aligned}&\text{rigidtropic:} && \gamma_1=0, \quad \gamma_i>0, \quad i=2,\ldots 6.\\ &\text{hydroisochoric:} && C_{11}+C_{22}+C_{33}+2C_{12}+2C_{13}+2C_{23}=0.\\ &\text{isochoric:} && C_{1j}+C_{2j}+C_{3j}=0, \quad j=1,2,3.\end{aligned} \tag{4}$$

Diagonal compliances are often known reliably from extensional and torsion tests. Off-diagonal entries are typically less amenable to accurate measurement. Volumetric constraints, for example on volume change, are checked with triaxial tests. In any case, such constraints may be satisfied only approximately. Reference [2] discusses projection and scaling techniques for finding a ''reference model'' that satisfies constraints accurately while removing spurious instabilities due to experimental noise.

3 Examples

The following examples of compliance matrices pertain to an orthotropic material with the $\{x_i\}$ aligned with the principal material axes. The diagonal entries are kept the same. The three nonzero off-diagonal entries are adjusted to meet the definitions (4).

Rigidtropic:

$$\mathbf{C}_{\text{rig}}=\begin{bmatrix} 1 & -3/8 & -3/16 & 0 & 0 & 0\\ -3/8 & 1/4 & -1/48 & 0 & 0 & 0\\ -3/16 & -1/48 & 1/9 & 0 & 0 & 0\\ 0 & 0 & 0 & 2 & 0 & 0\\ 0 & 0 & 0 & 0 & 5 & 0\\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix} = \frac{1}{144}\begin{bmatrix} 144 & -54 & -27 & 0 & 0 & 0\\ -54 & 36 & -3 & 0 & 0 & 0\\ -27 & -3 & 16 & 0 & 0 & 0\\ 0 & 0 & 0 & 288 & 0 & 0\\ 0 & 0 & 0 & 0 & 720 & 0\\ 0 & 0 & 0 & 0 & 0 & 432 \end{bmatrix}. \tag{5}$$

Eigenvalues: [5 3 2 1.181038 0.180074 0]. The compliance matrix is semistable. The null eigenvector defining the rigid mode is $\mathbf{v}_1=\sqrt{54/35}[1/2\ 5/6\ 1\ 0\ 0\ 0]^T$.

Hydroisochoric:

$$\mathbf{C}_{\text{hyd}}=\begin{bmatrix} 1 & -11/27 & -95/432 & 0 & 0 & 0\\ -11/27 & 1/4 & -23/432 & 0 & 0 & 0\\ -95/432 & -23/432 & 1/9 & 0 & 0 & 0\\ 0 & 0 & 0 & 2 & 0 & 0\\ 0 & 0 & 0 & 0 & 5 & 0\\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix} = \frac{1}{432}\begin{bmatrix} 432 & -176 & -95 & 0 & 0 & 0\\ -176 & 108 & -23 & 0 & 0 & 0\\ -95 & -23 & 48 & 0 & 0 & 0\\ 0 & 0 & 0 & 576 & 0 & 0\\ 0 & 0 & 0 & 0 & 1440 & 0\\ 0 & 0 & 0 & 0 & 0 & 864 \end{bmatrix}. \tag{6}$$

Eigenvalues: [5 3 2 1.208689 0.211580 −0.059158]. The compliance matrix is unstable. Isochoric:

$$\mathbf{C}_{\text{iso}}=\begin{bmatrix} 1 & -41/72 & -31/72 & 0 & 0 & 0\\ -41/72 & 1/4 & 23/72 & 0 & 0 & 0\\ -31/72 & 23/72 & 1/9 & 0 & 0 & 0\\ 0 & 0 & 0 & 2 & 0 & 0\\ 0 & 0 & 0 & 0 & 5 & 0\\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix} = \frac{1}{144}\begin{bmatrix} 144 & -82 & -62 & 0 & 0 & 0\\ -82 & 36 & 46 & 0 & 0 & 0\\ -62 & 46 & 16 & 0 & 0 & 0\\ 0 & 0 & 0 & 288 & 0 & 0\\ 0 & 0 & 0 & 0 & 720 & 0\\ 0 & 0 & 0 & 0 & 0 & 432 \end{bmatrix}. \tag{7}$$

Eigenvalues: [5 3 2 1.508781 0 −0.147669]. The compliance matrix is unstable.

4 Hydroisochoric Model

Assume that the material modeled by (2) is hydroisochoric. Consequently

$$\mathbf{C}\boldsymbol{\sigma}_p=\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16}\\ & C_{22} & C_{23} & C_{24} & C_{25} & C_{26}\\ & & C_{33} & C_{34} & C_{35} & C_{36}\\ & & & C_{44} & C_{45} & C_{46}\\ & & & & C_{55} & C_{56}\\ \text{symm} & & & & & C_{66} \end{bmatrix}\begin{bmatrix} p\\ p\\ p\\ 0\\ 0\\ 0 \end{bmatrix} = \begin{bmatrix} p(C_{11}+C_{12}+C_{13})\\ p(C_{12}+C_{22}+C_{23})\\ p(C_{13}+C_{23}+C_{33})\\ 2e_{12}\\ 2e_{23}\\ 2e_{31} \end{bmatrix} = \begin{bmatrix} e_{11}\\ e_{22}\\ e_{33}\\ 2e_{23}\\ 2e_{31}\\ 2e_{12} \end{bmatrix},$$

with $e_v = e_{11}+e_{22}+e_{33}$

$$= p(C_{11}+C_{22}+C_{33}+2C_{12}+2C_{13}+2C_{23})$$

$$= 0. \tag{8}$$

(The value of the shear strains is of no interest.) The complementary energy density produced by $\boldsymbol{\sigma}_p$ is

$$\mathcal{U}_p^* = \frac{1}{2}\boldsymbol{\sigma}_p^T \mathbf{C}\boldsymbol{\sigma}_p = \frac{1}{2}p(e_{11}+e_{22}+e_{33}) = \frac{1}{2}pe_v = 0. \tag{9}$$

But $\gamma_p = \mathcal{U}_p^*/(\boldsymbol{\sigma}_p^T\boldsymbol{\sigma}) = \mathcal{U}_p^*/(3p^2) = 0$ is the Rayleigh quotient of $\boldsymbol{\sigma}_p$ with $\mathbf{C}$. According to the Courant-Fisher theorem, [2], γ_p must lie in the closed interval $[\gamma_{\min}, \gamma_{\max}]$:

$$\gamma_1 \leqslant \gamma_p = 0 \leqslant \gamma_6. \tag{10}$$

If $\boldsymbol{\sigma}_p$ is not an eigenvector of $\mathbf{C}$: $\mathbf{C}\boldsymbol{\sigma}_p \neq \mathbf{0}$, the leftmost equality in (10) is not possible. Consequently

$$\gamma_1 < 0, \tag{11}$$

and the model is unstable.

If $\mathbf{C}\boldsymbol{\sigma}_p = \mathbf{0}$ the sum of the first three columns (or rows) of $\mathbf{C}$ must vanish. The hydroisochoric model then coalesces with the isochoric one, which is analyzed next.

5 Isochoric Model

The model is isochoric if the sum of the first three rows (or columns) of $\mathbf{C}$ is the null six-vector. Equivalently $\boldsymbol{\sigma}_p$ is a null eigenvector of $\mathbf{C}$. The Rayleigh quotient test (10) does not offer sufficient information on stability and a deeper look at $\mathbf{C}$ is required. Nonetheless a *sufficient* criterion for instability can be derived by considering the upper 3×3 principal minor $\tilde{\mathbf{C}}$. From the last of (4), $\tilde{\mathbf{C}}$ must have the form

$$\tilde{\mathbf{C}} = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ & C_{22} & C_{23} \\ & \text{symm} & C_{33} \end{bmatrix}$$

$$= \begin{bmatrix} C_{11} & \frac{1}{2}(C_{33}-C_{11}-C_{22}) & \frac{1}{2}(C_{22}-C_{11}-C_{33}) \\ & C_{22} & \frac{1}{2}(C_{11}-C_{22}-C_{33}) \\ & \text{symm} & C_{33} \end{bmatrix}. \tag{12}$$

This matrix is singular. Taking $\alpha = C_{11}/C_{22}$ and $\beta = C_{11}/C_{33}$ for convenience, an eigenvalue analysis shows that $\tilde{\mathbf{C}}$ is indefinite if

$$2\left(\frac{1}{\alpha}+\frac{1}{\beta}\right) < 1+\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)^2, \tag{13}$$

and is positive semidefinite if the inequality is reversed. If $\tilde{\mathbf{C}}$ is indefinite, so is $\mathbf{C}$ and the model is unstable. If $\tilde{\mathbf{C}}$ is semidefinite, an eigenvalue analysis of the complete $\mathbf{C}$ is required to decide on stability. The stability regions of $\tilde{\mathbf{C}}$ are shown in Fig. 1, where "potentially semistable" indicates that confirmation by a analysis of the full $\mathbf{C}$ is required. An exception is an orthotropic material referred to principal material axes, in which case no further tests are necessary if C_{44}, C_{55}, and C_{66} are positive.

Figure 1 illustrates that a wide range of diagonal compliances in $\tilde{\mathbf{C}}$ is detrimental to stability. For example if $\alpha = \beta$, instability is guaranteed to happen for $\alpha > 4$.

6 Rigidtropic Model

If $\mathbf{C}$ is nonnegative with $\gamma_1 = 0$ and $\mathbf{w} = \mathbf{v}_1$ is the only null eigenvector the material is rigidtropic under that stress mode. For an isotropic material $\mathbf{w} = [1\ 1\ 1\ 0\ 0\ 0]^T = \boldsymbol{\sigma}_p$, the hydrostatic stress mode. For an anisotropic material mode $\mathbf{w}$ generally will contain shear stresses. Introducing effective pressure as $p = 1/3\mathbf{w}^T\boldsymbol{\sigma}$ and effective volumetric strain as $e_v = \mathbf{w}^T\boldsymbol{\sigma}$, the volumetric and deviatoric energies can be uncoupled, [3].

If the rigid stress mode is $\boldsymbol{\sigma}_p$, rigidtropic reduces to isochoric. This inclusion is pictured in Fig. 2.

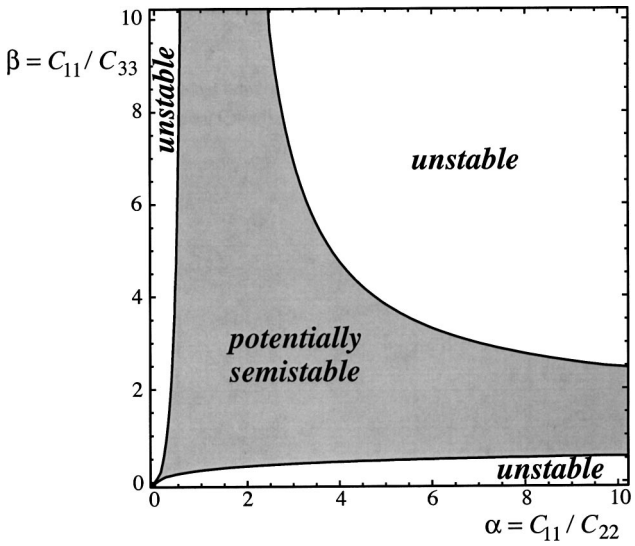


Fig. 1 Stability chart for the principal minor (12) of an isochoric material as function of the ratios C_{11}/C_{22} and C_{11}/C_{33}

7 Isotropic Material

If the solid is isotropic with elastic modulus $E > 0$ and Poisson's ratio ν,

$$\mathbf{C} = \frac{1}{E}\begin{bmatrix} 1 & -\nu & -\nu & 0 & 0 & 0 \\ & 1 & -\nu & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 \\ & & & 2(1+\nu) & 0 & 0 \\ & & & & 2(1+\nu) & 0 \\ \text{symm} & & & & & 2(1+\nu) \end{bmatrix}. \tag{14}$$

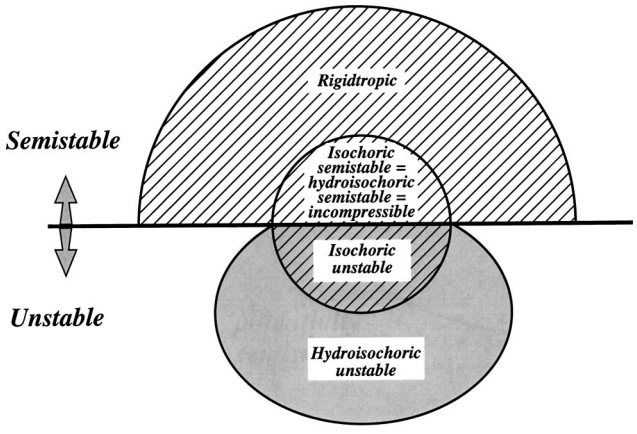


Fig. 2 Schematic of inclusions between rigidtropic, isochoric and hydroisochoric models. The crosshatched area marks a singular C matrix.

Under hydrostatic stress $\boldsymbol{\sigma}_p$, $e_v = 3(1-2\nu)p/E$, which vanishes for $\nu = 1/2$. It is easy to verify that if $\nu = 1/2$, $e_v = 0$ for any $\boldsymbol{\sigma}$ and the material is isochoric. Furthermore $\boldsymbol{\sigma}_p$ is the only null eigenvector of $\mathbf{C}$. Consequently $\gamma_p = \gamma_1 = 0$ and $\mathbf{C}$ has no negative eigenvalues. The definitions of rigidtropic, incompressible and isochoric behavior coalesce for this model.

8 Conclusions

It remains to pin down the label ''incompressible.'' In continuum mechanics this term means that the stress is determined by the deformation history only up to a hydrostatic pressure or ''extra stress'' p ([4], Sec. 30). This is equivalent to what we call here the hydroisochoric model, which as previously shown for semistable materials merges with the isochoric model. Restricting attention to the semistable case, the model nesting is:

$$\text{Isotropic semistable} = \text{Hydroisochoric semistable} \equiv \text{Incompressible} \in \text{Rigidtropic}. \quad (15)$$

These and related model inclusions are sketched in Fig. 2. From a mathematical standpoint, the splitting techniques used for the rigidtropic model by Felippa and Oñate [3] apply equally to isochoric behavior, and no special distinction for the incompressible case needs to be made.

We do not consider here the comparatively rare case of a compliance matrix possessing two or more zero eigenvalues. For those the analysis is complicated by the appearance of a multidimensional null space. Such ''multi-rigidtropic'' models require separate treatment.

Acknowledgments

Preparation of this paper has been supported by the National Science Foundation under Grant CMS-0219422, and by a faculty fellowship from the Spanish Ministerio of Educación y Cultura while visiting CIMNE over the period April through June 2002.

Nomenclature

- $\mathbf{C}$ = compliance matrix
- $\mathbf{C}$ = upper 3×3 minor of compliance matrix
- C_{ij} = entries of compliance matrix
- E = elastic modulus of isotropic model
- $\mathbf{e}$ = strain 6-vector arranged as per (1)
- e_{ij} = strain tensor components
- e_v = volumetric strain
- p = amplitude of hydrostatic stress state
- $\mathbf{v}_i$ = eigenvectors of compliance matrix
- $\mathbf{w}$ = null eigenvector of singular compliance matrix
- α = compliance ratio C_{11}/C_{22}
- β = compliance ratio C_{11}/C_{33}
- γ_i = eigenvalues of compliance matrix
- ν = Poisson's ratio of isotropic model
- $\boldsymbol{\sigma}$ = stress 6-vector, arranged as per (1)
- $\boldsymbol{\sigma}_p$ = hydrostatic stress 6-vector
- σ_{ij} = stress tensor components
- $\mathcal{U}$ = strain energy density
- $\mathcal{U}^*$ = stress (complementary) energy density

References

[1] Truesdell, C. A., and Toupin, R., 1960, ''The Classical Field Theories,'' *Handbook der Physik*, S. Flugge, ed., **III/1**, Springer-Verlag, Berlin.
[2] Courant, R., and Hilbert, D., 1952, *Methods of Mathematical Physics*, **I**, Interscience, New York.
[3] Felippa, C. A., and Oñate, E., 2003, ''Stress, Strain and Energy Splittings for Anisotropic Elastic Solids Under Volumetric Constraints,'' Comput. Struct., **81**, pp. 1343–1357.
[4] Truesdell, C. A., and Noll, W., 1965, ''The Nonlinear Field Theories of Mechanics,'' *Handbook der Physik*, S. Flugge, ed., **III/3**, Springer-Verlag, Berlin.

A Basic Power Decomposition in Lagrangian Mechanics

J. Casey
Mem. ASME,
Departments of Mechanical Engineering and Bioengineering, University of California,
6125 Etcheverry Hall,
Berkeley, CA 94720-1740.
e-mail: jcasey@me.berkeley.edu.

[DOI: 10.1115/1.1778413]

1 Introduction

In Lagrangian mechanics, under certain conditions, the Jacobi energy integral exists and plays a fundamental role (see [1–6]). More generally, when Jacobi's integral does not exist, it is still possible to gain useful engineering information from a consideration of power versus rate-of-energy relations. In the present note, we are concerned with a system of $N\,(\geqslant 1)$ particles subject to general holonomic and non-holonomic constraints. The unconstrained physical system may be represented by an abstract particle P in a $3N$-dimensional Euclidean configuration space. In the presence of holonomic constraints, the motion of P is confined to a submanifold M whose dimension is equal to the number of generalized coordinates needed to describe the system. In general, M moves through configuration space and may also change its shape with time.[1] Now, the velocity $\mathbf{v}$ of P can always be expressed as the vector sum of two components $\mathbf{v}'$ and $\mathbf{v}''$ such that $\mathbf{v}''$ is the velocity of the point A (say) of M that P occupies at time t, and $\mathbf{v}'$ is the velocity of P relative to A. It will be shown that when this decomposition is employed, the corresponding portions P' and P'' of the total power P of the forces acting on the particles, can be expressed as time derivatives (partial and total) of portions of the kinetic energy.[2] These expressions furnish a convenient means for calculating the power expended in moving the manifold M, and in moving P relative to M. This is particularly useful in the former case, because the constraint forces that move M would have been eliminated from the Lagrangian analysis.

The discussion is presented both in terms of physical quantities and abstract variables in configuration space. A few remarks regarding the desirability of the latter geometrical representation are in order:

Several different approaches to Lagrange's equations can be found in the literature. These vary both in generality and in the degree of physical insight that they provide. Some are based on d'Alembert's principle and the principle of virtual work, while some others use variational principles. With the advent of Riemannian geometry and tensor calculus in the 19th century, a new, abstract approach to dynamical theory arose, represented most cogently, perhaps, by Hertz [7]. In 1927, Synge [8] argued passionately for an approach to dynamics that is phrased in geometrical terms using the analytical apparatus of the tensor calculus (see also Section 186 of [1], as well as Synge's address, [9], to the American Mathematical Society in 1935). For constrained particle systems, a derivation of Lagrange's equations based on such an approach was first published by Synge and Schild [10], who proved that Lagrange's equations are just the covariant components, in the configuration manifold M of the constrained system, of Newton's second law. In this kind of derivation, no appeal to the concepts of virtual displacements, virtual velocities, nor virtual work is necessary (although the majority of authors of monographs on tensor calculus still prefer to employ them (see, e.g., [11,12])). Moreover, the exact physical content of Lagrange's equations is revealed clearly. Recently, Casey [13] showed that, in Synge and Schild's type of derivation, it is actually possible to bypass the cumbersome manipulations of Christoffel symbols, while maintaining the logical thrust of the original argument.[3] An additional major advantage of the geometrical approach to dynamics is that it places the subject in the rich mathematical environment of the global theory of differential equations on manifolds, a theory which is undoubtedly one of the most beautiful and powerful in all of mathematics.[4]

2 Representation of Particle Systems in Configuration Space

Consider a system consisting of particles $P_i\,(i=1,\ldots,N)$ moving relative to a Newtonian frame of reference under the influence of forces and possibly subject to time-dependent holonomic and non-holonomic constraints. Let $\mathbf{r}_i$, with rectangular Cartesian components (x_i^1,x_i^2,x_i^3), be the position vector of the particle P_i relative to the origin o of the Newtonian frame, let $M_i\,(>0)$ be the mass of P_i, and let the total mass of the system be m. Let the resultant force vector acting on P_i be $\mathbf{F}_i$, and denote its rectangular components by (F_i^1,F_i^2,F_i^3). We represent the physical system by an abstract particle P of mass m moving in a fixed $3N$-dimensional Euclidean vector space, *configuration space*, E^{3N}, as follows: The coordinates $u^i\,(i=1,2,\ldots,3N)$ of P, taken along mutually orthogonal axes through an arbitrarily chosen origin O, are identified as $(u^{3i-2},u^{3i-1},u^{3i})=(x_i^1,x_i^2,x_i^3)$, $(i=1,2,\ldots,N)$. Correspondingly, the position of P may be represented by its position vector $\mathbf{r}$. A metric on E^{3N} may be defined by

$$m d^2(P,O)=\sum_{i=1}^{N} M_i\mathbf{r}_i\cdot\mathbf{r}_i=\sum_{i=1}^{3N} m_i(u^i)^2, \tag{1}$$

where $m_{3i-2}=m_{3i-1}=m_{3i}=M_i\,(i=1,2,\ldots,N)$. Thus, the distance d of P from O is defined to be the radius of gyration of the particle system about the origin o in physical space.[5] The corresponding inner product is

$$[\mathbf{r},\mathbf{r}^*]=\frac{1}{m}\sum_{i=1}^{N} M_i\mathbf{r}_i\cdot\mathbf{r}_i^*, \tag{2}$$

where the asterisk denotes a second set of position vectors for the system. The position vector of P can be expressed as $\mathbf{r}=\Sigma_{i=1}^{3N}u^i\mathbf{e}_i$, where $\mathbf{e}_i\,(1=1,2,\ldots,3N)$ are pairwise orthogonal basis vectors, whose magnitudes are determined by (2).[6] A reciprocal basis $\mathbf{e}^i\,(i=1,2,\ldots,3N)$ may be defined by the conditions $[\mathbf{e}^i,\mathbf{e}_j]=\delta_j^i$, $(i=1,2,\ldots,3N;j=1,2,\ldots,3N)$, where δ_j^i is the Kronecker delta, having the value unity if $i=j$, and zero otherwise. We introduce an abstract force vector $\mathbf{\Phi}$ in E^{3N} by Φ

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 25, 2003, final revision, February 6, 2004. Associate Editor: I. Mezic.

[1] A simple example is provided by a heavy bead sliding on a spinning wire whose motion is prescribed as a function of time (a special case of this is analyzed in Example 3 at the end of the note). In practice, almost all mechanical engineering devices involve parts that move on other moving parts.

[2] See Eqs. (15a,b) and (6).

[3] The corresponding derivation of Lagrange's equations for a single rigid body and a system of rigid bodies may be found in [14,15]. A similar derivation of Lagrange's equations for a pseudo-rigid body is given in [16].

[4] The qualitative study of ordinary differential equations was initiated by Henri Poincaré (1854–1912) well over a century ago, and has truly blossomed during the twentieth century. See the monographs by Arnold [17], Abraham and Marsden [18], Guckenheimer and Holmes [19], Marsden [20], and Marsden and Ratiu [21] for an account of modern developments. See also Hirsch [22] and Smale [23].

[5] This *inertia metric*, or more precisely, the *moment of inertia metric*, was utilized by Hertz [7] (see [13,14] for further references).

[6] Thus, the magnitude of each of the first three of these vectors is $\sqrt{m_1/m}$.

$=\Sigma_{i=1}^{3N}\Phi_i\mathbf{e}^i$, with $(\Phi_{3i-2},\Phi_{3i-1},\Phi_{3i})=(F_i^1,F_i^2,F_i^3),(i=1,\ldots,N)$. Newton's second law, written for each particle P_i, is equivalent to the vector equation

$$\mathbf{\Phi}=m\ddot{\mathbf{r}} \tag{3}$$

in E^{3N} (see [13]). Thus, the dynamics of the unconstrained physical system is now represented by the dynamics of a single abstract particle P of mass m moving through configuration space.

In the presence of L holonomic constraints ($L<3N$), which are allowed to be time-dependent, P will be confined to a moving manifold M, called the *configuration space of the constrained system*, or simply *the constraint manifold*, of dimension $n=3N-L$.[7] Let $q^\alpha(\alpha=1,2,\ldots,n)$ be convected coordinates on M; these are our "generalized coordinates." The position vector of P can now be written as a function $\mathbf{r}=\mathbf{r}(q^\alpha,t)$. The covariant basis vectors in M are defined by

$$\mathbf{a}_\alpha=\frac{\partial\mathbf{r}}{\partial q^\alpha}\quad(\alpha=1,2,\ldots,n). \tag{4}$$

The inner products $a_{\alpha\beta}=[\mathbf{a}_\alpha,\mathbf{a}_\beta]$, $(\alpha=1,2,\ldots,n;\ \beta=1,2,\ldots,n)$ furnish a Riemannian metric on M. The velocity $\mathbf{v}$ of the particle P has the decomposition

$$\mathbf{v}=\mathbf{v}'+\mathbf{v}'', \tag{5a}$$

$$\mathbf{v}'=\sum_{\alpha=1}^{n}\dot{q}^\alpha\mathbf{a}_\alpha, \tag{5b}$$

$$\mathbf{v}''=\frac{\partial\mathbf{r}}{\partial t}. \tag{5c}$$

The component $\mathbf{v}''=\mathbf{v}''(q^\alpha,t)$ is the velocity of the point A of M that P instantaneously occupies at time t; the component $\mathbf{v}'=\mathbf{v}'(q^\alpha,\dot{q}^\alpha,t)$ is the velocity of P relative to A, and it lies in the tangent space to M at A. In general, these two components are not orthogonal to one another. The kinetic energy of the system can be expressed as

$$T=\tfrac{1}{2}m[\mathbf{v},\mathbf{v}]=T_2+T_1+T_0, \tag{6}$$

where[8]

$$T_2=\frac{1}{2}m[\mathbf{v}',\mathbf{v}']=\frac{1}{2}m\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}a_{\alpha\beta}\dot{q}^\alpha\dot{q}^\beta, \tag{7a}$$

$$T_1=m[\mathbf{v}',\mathbf{v}'']=m\sum_{\alpha=1}^{n}b_\alpha\dot{q}^\alpha,\quad b_\alpha=[\mathbf{v}'',\mathbf{a}_\alpha],\quad(\alpha=1,2,\ldots,n) \tag{7b}$$

$$T_0=\tfrac{1}{2}m[\mathbf{v}'',\mathbf{v}'']. \tag{7c}$$

The Lagrange's equations for the system can be written in the general form

$$\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{q}^\gamma}\right)-\frac{\partial T}{\partial q^\gamma}=Q_\gamma,\quad(\gamma=1,2,\ldots,n) \tag{8}$$

where

$$Q_\gamma=[\mathbf{\Phi},\mathbf{a}_\gamma]\quad(\gamma=1,2,\ldots,n) \tag{9}$$

are the covariant components of Φ in the manifold M.[9] These "generalized forces" need not be derivable from a potential, and no assumption whatsoever is being made regarding the nature of the constraint forces (which may, for example, be dissipative). Recalling the relationship that the abstract vectors $\mathbf{r}$, $\mathbf{\Phi}$ bear to physical position and force vectors, and making use of (4), we obtain

$$Q_\gamma=\sum_{i=1}^{N}\mathbf{F}_i\cdot\frac{\partial\mathbf{r}_i}{\partial q^\gamma}\quad(\gamma=1,2,\ldots,n). \tag{10}$$

Further, we note that the velocity of $P_i\,(i=1,2,\ldots,N)$ can be expressed as

$$\mathbf{v}_i=\dot{\mathbf{r}}_i=\mathbf{v}_i'+\mathbf{v}_i'',\quad(i=1,2,\ldots,N) \tag{11a}$$

with

$$\mathbf{v}_i'=\sum_{\alpha=1}^{n}\frac{\partial\mathbf{r}_i}{\partial q^\alpha}\dot{q}^\alpha,\ \ \mathbf{v}_i''=\frac{\partial\mathbf{r}_i}{\partial t}\quad(i=1,2,\ldots,N). \tag{11b}$$

3 The Power Decomposition

Let P be the power of all of the forces acting on the system, i.e.,

$$\mathsf{P}=\sum_{i=1}^{N}\mathbf{F}_i\cdot\dot{\mathbf{r}}_i=[\mathbf{\Phi},\mathbf{v}]. \tag{12}$$

It is obvious from (12), (3), and (6) that

$$\mathsf{P}=m[\dot{\mathbf{v}},\mathbf{v}]=\dot{T}. \tag{13}$$

Further, it is evident from (12), (5), and (11*a*,*b*) that P can always be decomposed as

$$\mathsf{P}=\mathsf{P}'+\mathsf{P}'', \tag{14a}$$

$$\mathsf{P}'=[\mathbf{\Phi},\mathbf{v}']=\sum_{\gamma=1}^{n}Q_\gamma\dot{q}^\gamma=\sum_{i=1}^{N}\mathbf{F}_i\cdot\mathbf{v}_i',\quad \mathsf{P}''=[\mathbf{\Phi},\mathbf{v}'']=\sum_{i=1}^{N}\mathbf{F}_i\cdot\mathbf{v}_i''. \tag{14b}$$

One can now establish the following results:[10]

$$\mathsf{P}'=\frac{d}{dt}(T_2-T_0)+\frac{\partial T}{\partial t}, \tag{15a}$$

$$\mathsf{P}''=\frac{d}{dt}(T_1+2T_0)-\frac{\partial T}{\partial t}. \tag{15b}$$

To prove (15*a*), note that by virtue of $(14b)_1$ and (8),

$$\begin{aligned}\mathsf{P}'&=\sum_{\gamma=1}^{n}\left(\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{q}^\gamma}\right)-\frac{\partial T}{\partial q^\gamma}\right)\dot{q}^\gamma\\&=\sum_{\gamma=1}^{n}\left(\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{q}^\gamma}\dot{q}^\gamma\right)-\frac{\partial T}{\partial\dot{q}^\gamma}\ddot{q}^\gamma-\frac{\partial T}{\partial q^\gamma}\dot{q}^\gamma\right).\end{aligned} \tag{16}$$

But, in view of (6) and (7*a*,*b*,*c*)

$$\frac{\partial T}{\partial\dot{q}^\gamma}=\frac{1}{2}m\sum_{\beta=1}^{n}a_{\gamma\beta}\dot{q}^\beta+\frac{1}{2}m\sum_{\beta=1}^{n}a_{\beta\gamma}\dot{q}^\beta+mb_\gamma,\quad(\gamma=1,2,\ldots,n). \tag{17}$$

Hence,

[7] Recall that non-holonomic constraints will not alter the dimension of the constraint manifold.

[8] Note that the length ds of the line element in M is related to T_2 by $ds^2=(2T_2/m)/mdt^2$. The kinetic energy T, depending on both generalized coordinates and generalized velocities, is a function defined on the $2n$-dimensional tangent bundle of M at time t.

[9] The left-hand side of (8) is equal to the n covariant components of mass times acceleration, i.e., $[m\dot{\mathbf{v}},\mathbf{a}_\gamma]$ (see [13] for details).

[10] The formula (15*a*) appears in published lecture notes ([24], p. 89) of Paul Painlevé (1863–1933). His course, given at the University of Lille in 1891, and at the Faculty of Sciences in Paris in 1895, was devoted to the integration of the equations of mechanics. I am not aware of any earlier reference. See also Painlevé [25], Appell [26], and Brillouin [11]. In addition to his contributions to mathematics and mechanics, Painlevé was active in politics. He became French minister for war in 1917, and subsequently was minister for aviation.

$$\sum_{\gamma=1}^{n} \frac{\partial T}{\partial \dot{q}^{\gamma}} \dot{q}^{\gamma} = 2T_2 + T_1. \tag{18}$$

Also, the total time derivative of T is given by

$$\dot{T} = \frac{\partial T}{\partial t} + \sum_{\gamma=1}^{n} \left(\frac{\partial T}{\partial q^{\gamma}} \dot{q}^{\gamma} + \frac{\partial T}{\partial \dot{q}^{\gamma}} \ddot{q}^{\gamma} \right). \tag{19}$$

From (16) and (19), it is clear that

$$\mathsf{P}' = \frac{d}{dt} \left(\sum_{\gamma=1}^{n} \frac{\partial T}{\partial \dot{q}^{\gamma}} \dot{q}^{\gamma} \right) + \frac{\partial T}{\partial t} - \dot{T}. \tag{20}$$

Equation (15*a*) follows immediately from (20), (18), and (6). Equation (15*b*) may be readily deduced from (13), (14*a*), (6), and (20).

We mention two important consequences of (15*a*):

(I) Suppose that: (a) the constraint forces do not contribute to the power P'; and (b) the remaining forces are derivable from a potential function V. Let $L = T - V$ be the Lagrangian function. Then,

$$\frac{d}{dt}(T_2 - T_0 + V) + \frac{\partial L}{\partial t} = 0. \tag{21}$$

To prove this, note that in view of assumption (b), the covariant components of the nonconstraint forces will be[11]

$$Q^*_{\gamma} = -\frac{\partial V}{\partial q^{\gamma}} \quad (\gamma = 1,2,\ldots,n). \tag{22}$$

The contribution of the components Q^*_{γ} to the power P' is

$$\sum_{\gamma=1}^{n} Q^*_{\gamma} \dot{q}^{\gamma} = -\sum_{\gamma=1}^{n} = 1 \frac{\partial V}{\partial q^{\gamma}} \dot{q}^{\gamma} = \frac{\partial V}{\partial t} - \dot{V}. \tag{23}$$

By assumption (a), the constraint forces contribute nothing to P'. With the help of $(14b)_1$ and (15*a*), it then follows that

$$\frac{\partial V}{\partial t} - \dot{V} = \frac{d}{dt}(T_2 - T_0) + \frac{\partial (L+V)}{\partial t}, \tag{24}$$

from which (21) can be concluded at once.

(II) (*Jacobi Integral*). In addition to the conditions (a) and (b) assumed in (I), suppose that the Lagrangian does not depend explicitly on t. Then, (21) immediately yields the integral:

$$T_2 - T_0 + V = \text{const.} = E' \quad (\text{say}). \tag{25}$$

4 Examples

Let us take some illustrative examples involving a single particle P of mass m. The configuration space now coincides with the physical three-dimensional space having the ordinary Euclidean metric.

Example 1. Suppose that P moves on a horizontal plane (elevator floor) that is being driven vertically upwards in a Newtonian frame. The floor is the constraint manifold. Using a fixed rectangular Cartesian coordinate system, we may write the velocity components in (5*b*,*c*) as $\mathbf{v}' = \dot{x}\mathbf{i} + \dot{y}\mathbf{j}$, $\mathbf{v}'' = \dot{h}\mathbf{k}$, where $h = h(t)$ is a prescribed function. The force acting on P is $\mathbf{F} = F_x\mathbf{i} + F_y\mathbf{j} + (N - mg)\mathbf{k}$, where N is the force supplied by the floor. The kinetic energy of P comprises

$$T_2 = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2), \quad T_1 = 0, \quad T_0 = \tfrac{1}{2}m\dot{h}^2. \tag{26}$$

Lagrange's equations yield $m\ddot{x} = F_x$, $m\ddot{y} = F_y$. The two portions of the power that appear in (14*b*) are $\mathsf{P}' = F_x\dot{x} + F_y\dot{y}$, $\mathsf{P}'' = (N - mg)\dot{h}$. Equations (15a,b) reduce to

$$\mathsf{P}' = \frac{dT_2}{dt}, \quad \mathsf{P}'' = \frac{dT_0}{dt} = m\ddot{h}\dot{h}. \tag{27}$$

Example 2. Suppose that P is constrained to move on a frictionless horizontal circle fixed at the origin of a Newtonian frame, and having a prescribed radius $l(t)$. Suppose that no other forces are applied to P. Here, the constraint manifold is a time-dependent circle centered at the origin and we may take the polar coordinate θ as our generalized coordinate. Let $(\mathbf{e}_r, \mathbf{e}_\theta)$ be the usual orthonormal basis of polar coordinates. The covariant basis vector is $\mathbf{a}_1 = \partial\mathbf{r}/\partial\theta = l\mathbf{e}_\theta$. The velocity components in (5*b*,*c*) are $\mathbf{v}' = \dot{\theta}(l\mathbf{e}_\theta)$, $\mathbf{v}'' = \dot{l}\mathbf{e}_r$. The kinetic energy of P comprises

$$T_2 = \tfrac{1}{2}ml^2\dot{\theta}^2, \quad T_1 = 0, \quad T_0 = \tfrac{1}{2}m\dot{l}^2. \tag{28}$$

The only force acting on P is that of the constraint and it points in the radial direction. The Lagrange's equation is $d/dt(\partial T/\partial\dot{\theta}) = 0$, which yields the angular momentum integral $ml^2\dot{\theta} = \text{const}$. The power decomposes as

$$\mathsf{P}' = 0 = \frac{d}{dt}(T_2 - T_0) + \frac{\partial T}{\partial t},$$

$$\mathsf{P} = \mathsf{P}'' = 2\frac{dT_0}{dt} - \frac{\partial T}{\partial t} = m(\ddot{l} - l\dot{\theta}^2)\dot{l}. \tag{29}$$

Example 3. Suppose that P is constrained to move on a frictionless rigid circle that is rotating with constant angular velocity Ω about a fixed vertical axis under the influence of gravity. Let us use spherical coordinates r, θ, φ, the angle φ being measured from the positive z-axis. Taking φ as our generalized coordinate, the position vector of P is $\mathbf{r} = \mathbf{r}(\varphi, t)$, and the velocity components in (5*b*,*c*) are $\mathbf{v}' = \dot{\varphi}(l\mathbf{e}_\varphi)$, $\mathbf{v}'' = l\Omega \sin\varphi\ \mathbf{e}_\theta$. The kinetic and potential energies are

$$T_2 = \tfrac{1}{2}ml^2\dot{\varphi}^2, \quad T_1 = 0, \quad T_0 = \tfrac{1}{2}ml^2\Omega^2\sin^2\varphi,$$
$$V = mgl\cos\varphi. \tag{30}$$

The constraint force has the form $\mathbf{N} = N_r\mathbf{e}_r + N_\theta\mathbf{e}_\theta$. The Lagrange's equation yields $\ddot{\varphi} - (g/l + \Omega^2\cos\varphi)\sin\varphi = 0$. The Lagrangian does not depend explicitly on time. We therefore have a Jacobi integral (25). The portion P'' of the power is

$$\mathsf{P}'' = 2\frac{dT_0}{dt} = ml^2\Omega^2\dot{\varphi}\sin 2\varphi, \tag{31}$$

i.e., the power supplied by N_θ.

Example 4. Suppose that P is confined to move on a fixed plane $z = 0$ under the action of potential forces. If fixed Cartesian coordinates are used, obviously $\mathsf{P}'' = 0$ and $T + V = \text{const}$. Instead, take another frame of reference, also with $z = 0$, but which rotates with constant angular velocity Ω about the z-axis. Let $(\mathbf{b}_1, \mathbf{b}_2)$ be an orthonormal basis fixed to the rotating frame and let the position vector of P be written as $\mathbf{r} = \rho_1\mathbf{b}_1 + \rho_2\mathbf{b}_2$. The constraint manifold M coincides with the rotating frame. Choosing ρ_1 and ρ_2 as generalized coordinates, we see that the covariant basis on M is $(\mathbf{b}_1, \mathbf{b}_2)$. The velocity components in (5*b*,*c*) are

$$\mathbf{v}' = \dot{\rho}_1\mathbf{b}_1 + \dot{\rho}_2\mathbf{b}_2, \quad \mathbf{v}'' = \Omega(-\rho_2\mathbf{b}_1 + \rho_1\mathbf{b}_2). \tag{32}$$

The kinetic energy of P has the three portions

$$T_2 = \tfrac{1}{2}m(\dot{\rho}_1^2 + \dot{\rho}_2^2), \quad T_1 = m\Omega(\rho_1\dot{\rho}_2 - \dot{\rho}_1\rho_2),$$
$$T_0 = \tfrac{1}{2}m\Omega^2(\rho_1^2 + \rho_2^2), \tag{33}$$

and $V = V(\rho_1, \rho_2, t)$. The Lagrange's equations yield

$$-\frac{\partial V}{\partial \rho_1} = m(\ddot{\rho}_1 - 2\Omega\dot{\rho}_2 - \Omega^2\rho_1),$$
$$-\frac{\partial V}{\partial \rho_2} = m(\ddot{\rho}_2 + 2\Omega\dot{\rho}_1 - \Omega^2\rho_2). \tag{34}$$

In view of the definitions (14*b*,*c*),

[11] The potential function in (22) may depend explicitly on t.

$$\mathsf{P}' = -\,grad V\cdot(\dot\rho_1\mathbf{b}_1+\dot\rho_2\mathbf{b}_2) = -\frac{\partial V}{\partial\rho_1}\dot\rho_1-\frac{\partial V}{\partial\rho_2}\dot\rho_2=\frac{\partial V}{\partial t}-\dot V, \tag{35a}$$

$$\mathsf{P}'' = -\,grad V\cdot\Omega(-\rho_2\mathbf{b}_1+\rho_1\mathbf{b}_2)=\Omega\left(\rho_2\frac{\partial V}{\partial\rho_1}-\rho_1\frac{\partial V}{\partial\rho_2}\right). \tag{35b}$$

Applying the relations (15*a*,*b*) to (33), we find that

$$\mathsf{P}' = \frac{d}{dt}(T_2-T_0)=m(\dot\rho_1\ddot\rho_1+\dot\rho_2\ddot\rho_2)-m\Omega^2(\rho_1\dot\rho_1+\rho_2\dot\rho_2), \tag{36a}$$

$$\mathsf{P}'' = \frac{d}{dt}(T_1+2T_0)=m\Omega(\rho_1\ddot\rho_2-\rho_2\ddot\rho_1)+2m\Omega^2(\rho_1\dot\rho_1+\rho_2\dot\rho_2). \tag{36b}$$

With the help of (34), Eqs. (35a,b) are seen to be equivalent to (36*a*,*b*).

Example 5. Suppose that P is confined to move on a fixed horizontal plane, with rectangular Cartesian coordinates x and y, under the action of potential forces and subject to the nonholonomic constraint $(-\sin\omega t)\dot x+(\cos\omega t)\dot y=0$, where $\omega=$const. Assume that the constraint force is parallel to the vector having the components $(-\sin\omega t,\ \cos\omega t)$. The kinetic energy is

$$T=T_2=\tfrac{1}{2}m(\dot x^2+\dot y^2). \tag{37}$$

The Lagrange's equations are

$$m\ddot x+\frac{\partial V}{\partial x}=-\lambda\sin\omega t,\qquad m\ddot y+\frac{\partial V}{\partial y}=\lambda\cos\omega t, \tag{38}$$

where the multiplier λ is plus or minus times the magnitude of the constraint force. The power P'' is zero and the constraint force makes no contribution to P'; (II) yields the energy integral $T+V=$const.

References

[1] Whittaker, E. T., 1944, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Second Ed., Cambridge University Press, Cambridge, UK.

[2] Pars, L. A., 1965, *A Treatise on Analytical Dynamics*, Heinemann, London.

[3] Goldstein, H., 1980, *Classical Mechanics*, Second Ed., Addison-Wesley, Reading, MA.

[4] Greenwood, D. T., 1977, *Classical Dynamics*, Prentice-Hall, Englewood Cliffs, NJ.

[5] Rosenberg, R. M., 1977, *Analytical Dynamics of Discrete Systems*, Plenum Press, New York.

[6] Papastavridis, J. G., 2002, *Analytical Mechanics: A Comprehensive Treatise on the Dynamics of Constrained Systems; for Engineers, Physicists, and Mathematicians*, Oxford University Press, Oxford, UK.

[7] Hertz, H., 1900, *The Principles of Mechanics Presented in a New Form*, Macmillan, New York (Dover Publications, 1956).

[8] Synge, J. L., 1927, ''On the Geometry of Dynamics,'' Philos. Trans. R. Soc. London, Ser. A, **226**, pp. 31–106.

[9] Synge, J. L., 1936, *Tensorial Methods in Dynamics* (University of Toronto Studies, Applied Mathematics Series No. 2), University of Toronto Press, Toronto, Canada.

[10] Synge, J. L., and Schild, A., 1949, *Tensor Calculus*, University of Toronto Press, Toronto, Canada (Dover, 1978).

[11] Brillouin, L., 1964, *Tensors in Mechanics and Elasticity*, Academic Press, San Diego, CA.

[12] Papastavridis, J. G., 1999, *Tensor Calculus and Analytical Dynamics*, CRC Press, Boca Raton, FL.

[13] Casey, J., 1994, ''Geometrical Derivation of Lagrange's Equations,'' Am. J. Phys., **62**, pp. 836–847.

[14] Casey, J., 1995, ''On the Advantages of a Geometrical Viewpoint in the Derivation of Lagrange's Equations for a Rigid Continuum,'' Z. Angew. Math. Phys., **46** (Special Issue), pp. S805–S847.

[15] Casey, J., and O'Reilly, O. M., 2004, ''Geometrical Derivation of Lagrange's Equations for a System of Rigid Bodies,'' Math. Mech. Solids, in press.

[16] Casey, J., 2004, ''Pseudo-Rigid Continua: Basic Theory and a Geometrical Derivation of Lagrange's Equations,'' Proc. R. Soc., London, Ser. A, **460**, pp. 2021–2049.

[17] Arnold, V. I., 1989, *Mathematical Methods of Classical Mechanics*, 2nd Ed., Springer-Verlag, New York.

[18] Abraham, R., and Marsden, J. E., 1978, *Foundations of Mechanics*, 2nd Ed., Addison-Wesley, Reading, MA.

[19] Guckenheimer, J., and Holmes, P., 1983, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York.

[20] Marsden, J. E., 1992, *Lectures in Mechanics* (London Mathematical Society Lecture Note Series, 174), Cambridge University Press, Cambridge, UK.

[21] Marsden, J. E., and Ratiu, T. S., 1999, *Introduction to Mechanics and Symmetry*, 2nd Ed., Springer-Verlag, New York.

[22] Hirsch, M. W., 1984, ''The Dynamical Systems Approach to Differential Equations,'' Bull. Am. Math. Soc., **11**, pp. 1–64.

[23] Smale, S., 1980, *The Mathematics of Time*, Springer-Verlag, New York.

[24] Painlevé, P., 1895, *Leçons sur l'Intégration des Equations Différentielles de la Mécanique et Applications*, A. Hermann, Paris.

[25] Painlevé, P., 1930, *Cours de Mécanique*, **I**, Gauthier-Villars et Cie., Paris.

[26] Appell, P., 1911, *Traité de Mécanique Rationnelle*, **2**, Gauthier-Villars, Paris.

System Identification Including the Load Environment

Z. R. Lu

S. S. Law

Civil and Structural Engineering Department,
The Hong Kong Polytechnic University,
Hunghom, Kowloon, Hong Kong

[DOI: 10.1115/1.1778412]

1 Introduction

A number of methods for the identification of the axle loads of vehicles using a bridge's static and pseudo-dynamic responses have been developed, [1–5]. Identification of moving loads has been studied separately by the authors, [6–8]. However existing methods to identify the system parameters or the moving loads do not consider the influence from the environment or the system itself, respectively. This paper attempts to include the external force system in the identification of system parameters of a simply supported beam. The dynamic response of a prestressed beam under moving load is studied based on modal superposition in a forward problem. An inverse problem to identify both the prestress force and the moving load is then formulated. It is further extended to include the flexural rigidity of the beam as variable in the identification. Results from the studies indicate that the identification of both the system parameters and the moving loads with normal modal testing technique is feasible even with noisy data.

2 Equation of Motion

The bridge deck is modeled as a single-span simply supported unbonded prestressed uniform Euler-Bernoulli beam subjected to a set of moving loads P_l $(l=1,2,\ldots,N_p)$ as shown in Fig. 1. These forces are assumed to be moving as a group at a prescribed velocity $v(t)$ along the axial direction of the beam from left to right. The equation of motion of the beam can be written as

$$\rho A\frac{\partial^2 y(x,t)}{\partial t^2}+c\frac{\partial y(x,t)}{\partial t}+T\frac{\partial^2 y(x,t)}{\partial x^2}+\frac{\partial^2}{\partial x^2}EI_0\frac{\partial^2 y(x,t)}{\partial x^2}=\sum_{l=1}^{N_p}P_l(t)\,\delta(x-x_l(t)) \tag{1}$$

where ρ is the mass density, A the cross-sectional area, c the damping of the beam, E the Young's modulus of material, I_0 the moment of inertia of the beam cross section, T the externally applied compressive axial force (note that compression is positive and tension is negative), $y(x,t)$ the transverse displacement function of the beam, $x_l(t)$ the location of the moving load $P_l(t)$ at the time t, N_p the number of the moving load, $\delta(t)$ the Dirac delta function and b is the width of the beam.

Expressing (1) using generalized coordinates,

$$\ddot{q}_i(t)+2\xi_i\overline{\omega}_i\dot{q}(t)+\overline{\omega}_i^2 q_i(t)=\frac{1}{m_i}f_i(t), \tag{2}$$

where $\overline{\omega}_i=\sqrt{EI_0/\rho A(i\pi/L)^4-T/\rho A(i\pi/L)^2}$, ξ_i and m_i are the reduced modal frequency, the damping ratio and the modal mass of the ith mode; $f_i(t)=\Sigma_{l=1}^{N_p}P_l(t)Y_i(x_l(t))$ is the modal force. Since the modal shape functions of the prestressed beam resemble those of a beam without prestress force, [9], and it can be written in the normalized form as $Y_i(x)=\sqrt{2/(\rho AL)}\sin i\pi/Lx$ for a simply supported beam. Writing (2) in matrix form

$$[I]\{\ddot{Q}(t)\}+[C]\{\dot{Q}(t)\}+([K]-[K'])\{Q(t)\}=\{F(t)\}, \tag{3}$$

the modal response of the system under load is computed in the time domain numerically by using the Newmark's integration scheme, [10].

3 Identification of the Prestress and Moving Forces

Expressing the measured displacements $\tilde{y}(x_m,t)$ at a point x_m from the left support in modal coordinates $\tilde{y}(x_m,t)=\Sigma_{i=1}^{N_m}Y_i(x)q_i(t)$, $(m=1,2,\ldots,N_m)$, or in vector

$$\{\tilde{y}\}_{N_m\times 1}=[Y]_{N_m\times N}\{q\}_{N\times 1} \tag{4}$$

where $\{\tilde{y}\}_{N_m\times 1}$ is the vector of displacements at N_m measurement locations. The vector of generalized coordinates can be obtained using the least-squares pseudo-inverse. A generalized orthogonal polynomial, [7], is used to model the measured displacement so as to reduce the computation error for the modal velocity and acceleration. The velocity and acceleration are then approximated by the first and second derivatives of the orthogonal polynomial. After some transformations, we have

$$[K']\{Q(t)\}+\{F(t)\}=[I]\{\ddot{Q}(t)\}+[C]\{\dot{Q}(t)\}+[K]\{Q(t)\}. \tag{5}$$

Matrix $[K']$ contains the prestress force T which is assumed constant throughout the length of the beam. The vector of generalized force $\{F(t)\}$ can also be found from

$$\{F(t)\}=[B]\{P(t)\}. \tag{6}$$

Rewriting Eq. (5) in a simple form

$$\{d_T\}_{n\times 1}T+[B]_{n\times N_p}\{P(t)\}_{N_p\times 1}=\{r\}_{n\times 1} \tag{7}$$

and vector $[r]$ contains all the terms on the right-hand side of (5). In (7) T is the unknown prestress force of the beam and $\{P(t)\}$ is the unknown moving load vector to be identified.

The inverse problem is to solve (7) in the time domain. Since both T and $\{P(t)\}$ are uncoupled, (7) can be further simplified into

$$[B_d]X=\{r\} \tag{8}$$

where $[B_d]=[\{d\},[B]]$, $X=\{^T_{\{P\}}\}$. The prestress force T and the moving load $\{P\}$ can be calculated directly using the least-squares method.

In order to have bounds on the ill-conditioned solution, the damped least-squares method is adopted, [11], and singular value decomposition is used in the pseudo-inverse computation. The prestress force and moving loads can also be identified from the measured strains. The strain at the bottom of the beam at a point x_m from the left support can be expressed similar to (4) in terms of the generalized coordinates as

$$\{\tilde{\varepsilon}(x_m,t)\}=-\frac{h_0}{2}[Y''(x_m)]\{q_i(t)\}\quad (m=1,2,\ldots,N_m). \tag{9}$$

The rest of the computation for the identification is similar to that for identification from measured displacements mentioned above.

4 Identification of the Prestress, Moving Forces, and the Flexural Rigidity of the Beam

Other variables in the system should also be included in the identification for a real application. One parameter that is subjected to variation is the flexural rigidity EI_0 of the beam section.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 2, 2003; final revision, Jan. 21, 2004. Associate Editor: O. O'Reilly.

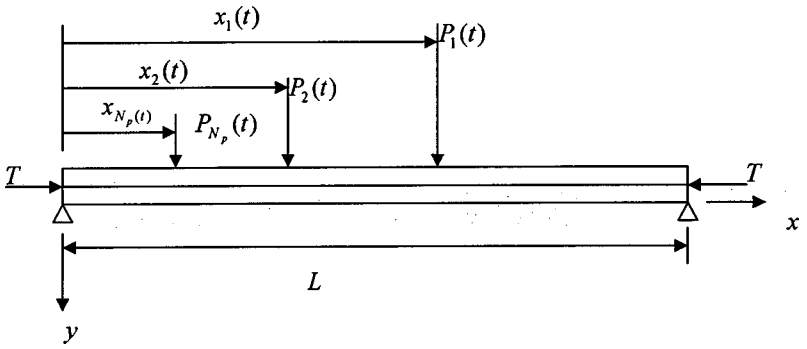


Fig. 1 A simply supported prestressed beam under N_p moving loads

Table 1 Errors in the identified single moving load and prestress force (in percent)

Number of Vibration Modes	Noise Level		
	1%	5%	10%
3	5.8/27.2	6.0/28.9	7.6/36.0
4	5.1/26.3	5.0/28.2	6.3/34.8
5	5.1/26.1	5.1/28.0	5.7/34.2
6	4.8/25.8	4.9/27.6	5.4/32.7

Note: •/• denotes errors for the moving load and the prestress force, respectively.

If we have a uniform uncracked beam in the problem, we have prestress force T, flexural rigidity EI_0, and moving load $P(t)$ as the three variables in the identification. Rewriting (7) as

$$\{d_T\}_{n\times 1}T+[B]_{n\times N_p}\{P(t)\}_{N_p\times 1}-\{d_{EI}\}EI_0 = \{r'\}_{n\times 1} \tag{10}$$

$$\{r'\}=[I]\{\ddot{Q}(t)\}+[C]\{\dot{Q}(t)\} \tag{11}$$

The rest is similar as Section 3.

5 Simulations

A 30-meters long simply supported Euler-Bernoulli beam with an axial prestress force of $0.3T_{cr}=8.2247\times 10^6\,N$ is studied, where T_{cr} is the critical buckling load of the beam and $T_{cr}=\pi^2EI_0/L^2$. The first six natural frequencies of the beam are: 1.03, 4.75, 10.11, 19.56, 30.67, and 44.25 Hz. The damping ratios for these six modes are all equal to 0.02. The prestress force is assumed constant along the beam. The moving load is taken as $P(t)=40{,}000[1+0.5\sin(2\pi t)+0.3\sin(10\pi t)]N$, which moves along the axial direction of the beam at 30 m/s from the left to right supports. The parameters of the beam are: $\rho A=5.0\times 10^3$ kg/m, $E=5\times 10^{10}$ N/m^2, $L=30$ m, $b=0.6$ m, and $h_0=1.0$ m. The flexural rigidity EI_0 of the beam is calculated as 2.5×10^9 Nm^2. White noise is added to the calculated displacements and strains to simulate the polluted measurements. 5% and

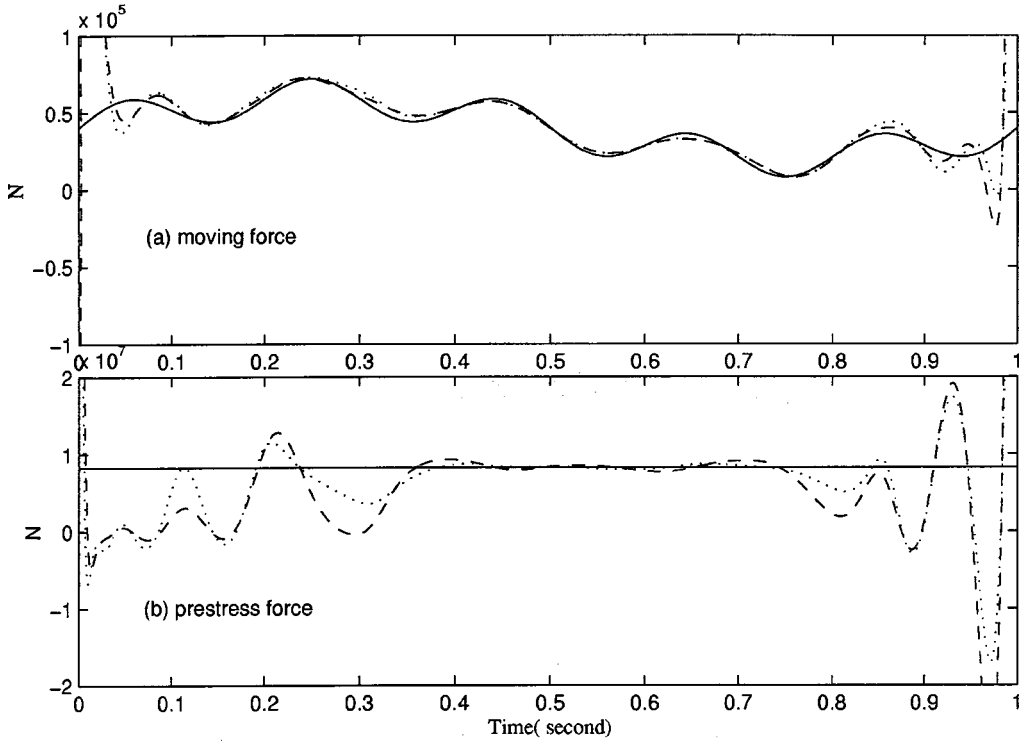


Fig. 2 Identification of single moving force and prestress force (— true, . . . 5% noise, - - - 10% noise)

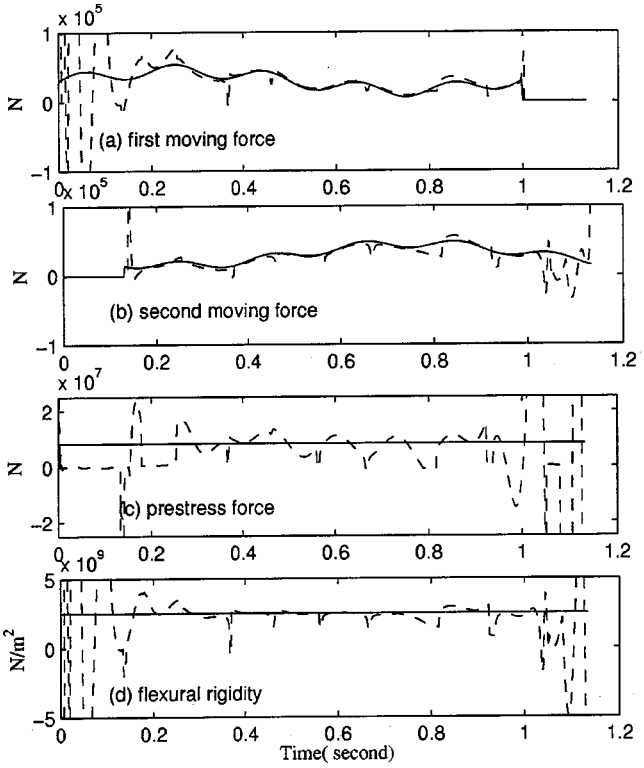


Fig. 3 Identification of two moving forces, prestress force and flexural rigidity of a beam (— true, - - - identified)

10% noise levels are studied in this brief note. The number of measured location is equal to the number of modes and they are evenly distributed on the beam.

Case 1—Identification of both the single moving load and the prestress force.

The first three to six modes and measured strains are used in the calculation. The sampling frequency is 1000 Hz. Table 1 shows the errors in the identified single moving load and the prestress force from using different number of vibration modes. It is found that: (a) the error decreases with the increase in the number of the vibration modes used in the identification; (b) the error in the moving load is much less than the error in the prestress force. This may be due to the reason that the response is not sensitive to the prestress force; and (c) the error is not sensitive to the noise level in the response measurements.

Figure 2 shows the identified results from measured strains with 5% and 10% noise. There is only a slight difference in the time histories of the identified moving load with the two noise levels. This is because the measurement noise has been substantially removed by the orthogonal function approximation.

Case 2—Identification of two moving loads, prestress force and the flexural rigidity of the beam.

The two moving loads are

$$P_1(t) = 20{,}000[1 + 0.5\sin(2\pi t) - 0.3\sin(10\pi t)] \quad N,$$

$$P_2(t) = 20{,}000[1 + 0.5\sin(2\pi t) + 0.2\sin(10\pi t)] \quad N.$$

moving as a group at 4-meters spacing at 30 m/s from left to right. The sampling rate is 500 Hz and four modes are used for the identification. 5% noise is included in the response measurements. The same system as for Case 1 is studied. Figure 3 shows that the identified prestress force, two moving loads and the flexural stiffness are fluctuating around the true values, especially around the middle half of the time histories.

6 Conclusion

This paper includes the load environment in the system identification of a structure. A method is presented in the time domain to identify the prestress force and the flexural rigidity of a Euler-Bernoulli beam under moving loads with regularized solution. The noise effect is minimized using orthogonal polynomial functions. The method gives some good results from the first few measured modes and data obtained from several measurement points. Numerical results demonstrate the feasibility of indirectly identifying both the moving loads on the beam and prestress force as well as the flexural rigidity of the beam from noisy measurements.

References

[1] Normann, O. K., and Hopkins, R. C., 1952, "Weighing Vehicles in Motion," Highway Research Board Bulletin 50, National Research Council, Washington, DC.
[2] Morses, F., 1984, "Weigh-in-Motion System Using Instrumented Bridge," J. Transp. Eng. ASCE, **105**, pp. 233–249.
[3] Davis, P., and Sommerville, F., 1986, "Low-Cost Axle Load Determination," *Proceedings of the 13th ARRB and 5th REAAA Combined Conference*, Part 6, pp. 1703–1723.
[4] Peters, R. J., 1986, "An Unmanned and Undetectable Highway Speed Vehicle Weighing System," *Proceedings of the 13th ARRB and 5th REAAA Combined Conference*, Part 6, pp. 70–83.
[5] Tritt, B., and Richards, B., 1978, "Determination of Vehicle Axle Mass Description and Demonstration of the ARRB System," *Proceedings of Axle Mass Determination Workshop*, ARRB.
[6] Zhu, X. Q., and Law, S. S., 1999, "Moving Loads Identification on a Multi-Span Continuous Bridge," J. Sound Vib., **228**, pp. 377–396.
[7] Zhu, X. Q., and Law, S. S., 2001, "Orthogonal Function in Moving Loads Identification on a Multi-Span Bridge," J. Sound Vib., **245**, pp. 329–345.
[8] Law, S. S., and Zhu, X. Q., 2000, "Study on Different Beam Models in Moving Load Identification," J. Sound Vib., **234**, pp. 661–679.
[9] Abraham, M. A., Park, S. Y., and Stubbs, N., 1995, "Loss of Prestress Prediction on Nondestructive Damage Location Algorithms," SPIE Smart Struct. Mater., **2446**, pp. 60–67.
[10] Newmark, N. W., 1959, "A Method of Computation for Structural Dynamics," J. Eng. Mech., **85**(3), pp. 67–94.
[11] Tikhonov, A. M., 1963, "On the Solution of Ill-Posed Problems and the Method of Regularization," Sov. Math., **4**, pp. 1035–1038.

The Effect of Warping Stress on the Lateral-Torsion Buckling of Cold-Formed Zed-Purlins

Xiao-ting Chu, Long-yuan Li, and Roger Kettle
School of Engineering and Applied Science, Aston University, Aston Triangle, Birmingham B4 7ET, UK

An analytical model is developed for analysing the elastic lateral-torsion buckling of cold-formed zed-purlins with partially lateral restraint from metal sheeting. The model is used to estimate the effects of warping stresses on the lateral-torsion buckling behavior of cold-formed zed-purlins with various boundaries and interval braces provided by antisag bars. The results show that the warping stress only has remarkable influence on the lateral-torsion buckling when the boundary of the member is fixed and there is no antisag bar present. [DOI: 10.1115/1.1781178]

1 Introduction

Cold-formed sections are widely used as purlins or rails, the intermediate members between the main structural frame and the corrugated roof or wall sheeting in light gauge steel construction. These cold-formed sections are produced in a variety of forms, such as zed, channel, and sigma, which are inherently sensitive to local, distortional, and lateral-torsion buckling, [1,2]. In the codes and standards local buckling is taken into account by using effective widths for plane elements, [3–6], originally developed by Winter [7] from the post-buckling analysis of plates under compression. This concept has now been extended to account for the effect of distortional buckling of edge-stiffened elements or of stiffened elements with an intermediate stiffener, [5,6]. The lateral-torsion buckling of cold-formed members is generally calculated from the theory of the lateral buckling of detached beams in which the effects of warping torsion, prior to buckling, is normally not considered, [8].

As most cold-formed sections are restrained by cladding or metal sheeting, the loads acting on the sections not only cause bending of the member about its two principal axes but also torsion of the cross section. Recently, Li [8] developed an analytical model for predicting the lateral-torsion buckling of cold-formed zed-purlins partially restrained by the sheeting. The model considers bending and torsion for both pre-buckling and buckling analyses. However, the torsion considered in the model is only non-warping torsion. In other words, the longitudinal stress generated by warping torsion is ignored in the model, which is the same as that treated in the lateral buckling of beams.

This paper is a further development of Li's model to consider the effect of warping stress on the lateral-torsion buckling of cold-formed members. It is well known that, when a thin-walled beam has one or more cross sections that are constrained against warping, a complex distribution of longitudinal warping stresses is developed. These longitudinal warping stresses, together with the longitudinal stresses generated by the two bending moments, may cause the beam to have local, distortional, or lateral-torsion buckling. In the present paper, the focus will only be on lateral-torsion buckling. Local and distortional buckling only use the local compressive stress and so the warping stresses can simply be added into the local compressive stress caused by bending to conduct the local and distortional buckling analyses.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 22, 2003, final revision, November 25, 2004. Associate Editor: S. Mukherjee.

2 Analytical Model

Consider a purlin, the section of which is shown in Fig. 1, that is laterally restrained in the translational direction but free in the rotational direction, which is the case for most practical applications. Let the origin of the coordinate system (x,y,z) be the centroid of the cross section, with the x-axis being along the longitudinal direction of the beam, and the y and z-axes taken in the plane of the cross section. The strain energy of the beam due to deflections and rotation can be expressed as

$$U_o=\frac{E}{2}\int_o^l\left[I_y\left(\frac{d^2w}{dx^2}\right)^2+2I_{yz}\frac{d^2w}{dx^2}\frac{d^2v}{dx^2}+I_z\left(\frac{d^2v}{dx^2}\right)^2\right]dx+\frac{GJ}{2}\int_o^l\left(\frac{d\phi}{dx}\right)^2dx+\frac{EC_w}{2}\int_o^l\left(\frac{d^2\phi}{dx^2}\right)^2dx \quad (1)$$

where v and w are the deflections of the beam in the y and z-directions, ϕ is the angle of twist, I_y and I_z are the second moments of the cross-section area about the y and z-axes, I_{yz} is the product moment of the cross-section area, E is the Young's modulus, G is the shear modulus, J is the torsion constant, C_w is the warping constant, and l is the span length of the beam. The corresponding work done by the uniformly distributed uplift load, q_y, is

$$W_o=\int_o^l q_y(a\phi-v)dx \quad (2)$$

where a is the distance between the loading line and the web central line. Since the purlin, in the present case, is assumed to be laterally restrained in its upper flange, the angle of rotation can be expressed in terms of the horizontal displacement, i.e., $\phi=2w/d$, where d is the midline depth of the section. The deflections, $v(x)$ and $w(x)$, and the angle of twist, $\phi(x)$ due to the externally applied load can be determined by employing the principle of minimum potential energy as follows:

$$\delta(U_o-W_o)=0. \quad (3)$$

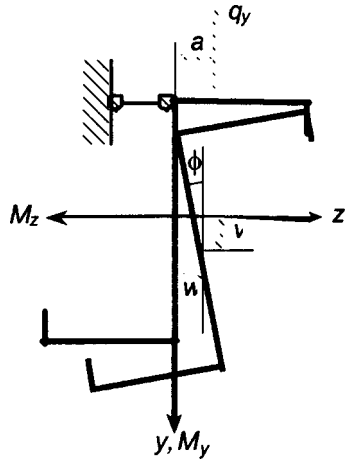


Fig. 1 Analytical model for lateral-torsion buckling analysis

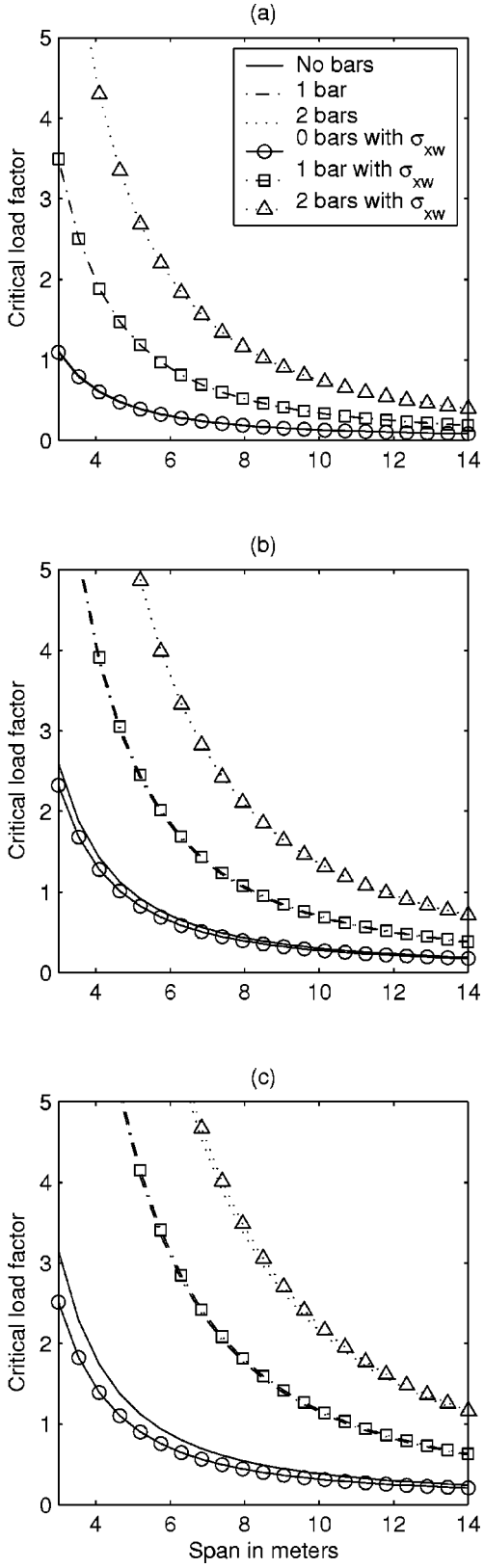


Fig. 2 Comparison of critical loads ($a=b/2$). (*a*) Both ends simply supported. (*b*) One end simply supported and the other end fixed. (*c*) Both ends fixed.

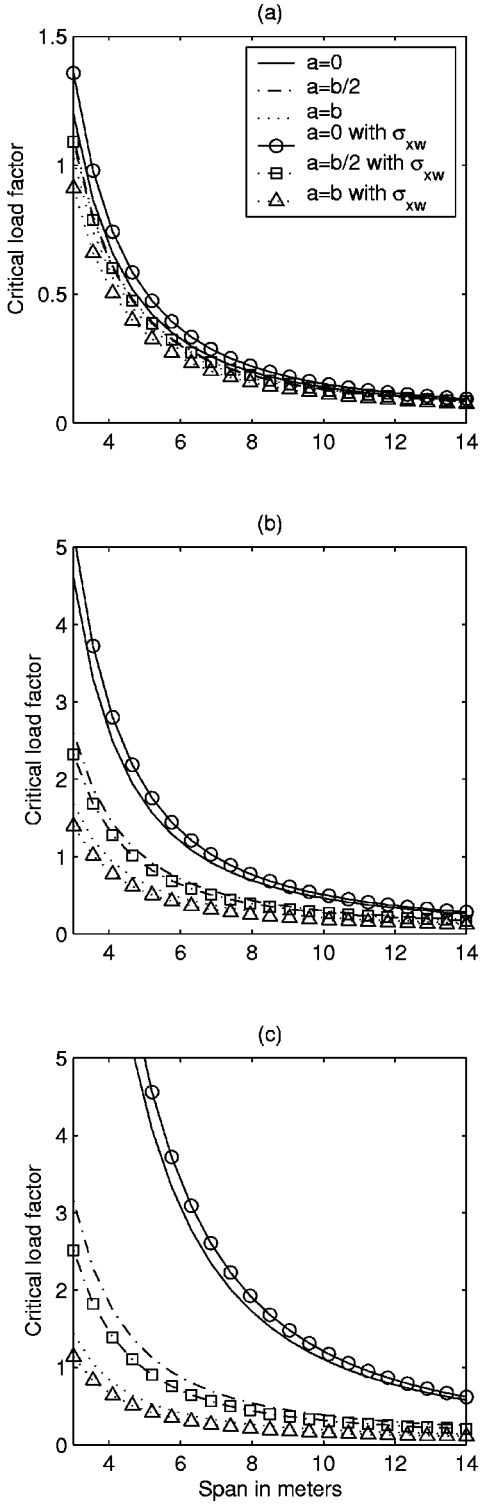


Fig. 3 Comparison of critical loads for no antisag bar purlin. (*a*) Both ends simply supported. (*b*) One end simply supported and the other end fixed. (*c*) Both ends fixed.

After the deflections and rotation are determined, the pre-buckling longitudinal stresses can be calculated which are then used to calculate the work done by the stresses through the buckling displacements as follows, [8]:

$$W_1 = -\int_o^l \int_A \sigma_{xb}\left[\phi_b\left(z\frac{d^2v_b}{dx^2} - y\frac{d^2w_b}{dx^2}\right)\right]dAdx$$
$$-\frac{1}{2}\int_o^l \int_A \sigma_{xw}\left[(y^2+z^2)\left(\frac{d\phi_b}{dx}\right)^2\right]dAdx - \frac{d}{4}\int_o^l q_y\phi_b^2 dx \quad (4)$$

where v_b and w_b are the buckling deflections of the beam in the y and z directions, and $\phi_b = 2w_b/d$ is the buckling angle of twist, σ_{xb} and σ_{xw} are the bending and warping stresses generated by the pre-buckling bending and twisting moments, respectively, which can be calculated directly from the obtained pre-buckling displacements, v, w, and ϕ. In Eq. (4), the first term represents the work done by the pre-buckling bending stress, the second term represents the work done by the pre-buckling warping stress and the third term represents the work done by the distributed load due to the load that is acting above the shear center, [9]. The expression used for calculating the strain energy U_1 generated by the buckling displacements is the same as Eq. (1) but v and w need to be replaced by v_b and w_b. The minimum critical buckling load and the corresponding buckling mode displacements can thus be determined by the following variational equation:

$$\delta(U_1 - \lambda W_1) = 0 \tag{5}$$

where λ is the loading proportional factor. If the work done by the warping stress is neglected in Eq. (4), then the present model reduces to the conventional lateral buckling models, [8,10–12].

In the present study, both Eqs. (3) and (5) are solved using a numerical method in which cubic spline interpolations are used to construct the deflection distribution with seven nodal displacements as the unknowns. The numerical solutions are not sensitive to the number of points chosen, as long as there is at least one point between constraint positions. The assumed deflection functions are required to pre-satisfy all of the required displacement boundary conditions. In this way, the variational Eqs. (3) and (5) are reduced to the matrix forms of a set of linear algebraic equations and a set of eigenvalue equations. The details of the numerical treatment can be found in our papers, [8], and thus are not presented further here.

3 Numerical Example

The purlin considered here is that: web depth $d = 202$ mm, flange width $b = 75$ mm, lip length $c = 20$ mm, thickness $t = 2$ mm, Young's modulus $E = 205$ Gpa, Poisson's ratio $\nu = 0.3$, and yield stress $\sigma_y = 390$ Mpa. The boundary conditions for the lateral displacement, w and w_b, are simply supported whereas, for the vertical displacement, v and v_b, are specified in each individual case. For cases where the purlin is laterally restrained by anti-sag bars it is assumed that the position of the bars be at the center of the span (for one bar) or at 3/8 and 5/8 of the span length (for two bars). The loading density is assumed as that $q_y = 12M_y/l^2$ for purlins with both ends fixed and $q_y = 8M_y/l^2$ for purlins with other boundary conditions (where $M_y = 2\sigma_y I_z/d$ is the yield moment).

The present pre-buckling stress analysis showed that the warping stress is considerably lower than the two bending stresses and is strongly dependent on the position of the applied load, boundary conditions on the beam and whether there are any antisag bars. Therefore, the influence of the warping stress on the lateral-torsion buckling is examined in terms of the boundary conditions, loading positions, and the number of antisag bars. Figure 2 shows the comparison of critical loads of the purlin with and without considering the warping stress, for three different boundary conditions. The results show that, the warping stress has almost no influence on the critical load for the purlin that is simply supported or that has one or two antisag bars. The influence of the warping stress on the critical load is noticed only in the case where the purlin has no antisag bars and is fixed at least at one end. The influence is found to decrease with the increased span length.

The influence of the warping stress on the critical load due to various loading positions is shown in Fig. 3 for the purlin without antisag bars. Again, the influence of the warping stress on the critical load is found merely in the purlin with one or two fixed ends. It is interesting to notice that the worst case associated with the lowest critical load occurs when the load is placed at the corner between the flange and the lip, whereas the highest critical load occurs when the load is applied through the central line of the web. The influence of the warping stress on the critical load is variable. When the warping stress is considered, the critical load factor is reduced for $a = b/2$ and $a = b$, but increases when $a = 0$.

4 Conclusions

An analytical model based on the energy method has been developed to estimate the effect of the warping stress on the lateral-torsion buckling of cold-formed zed-purlins subject to partial-lateral restraint from the metal sheeting under a uniformly distributed uplift load. From the present numerical studies the following conclusions are drawn:

- The remarkable influence of the warping stress on the critical load for lateral-torsion buckling of the partial-laterally restrained purlin was found only when the purlin was fixed at least at one end and no antisag bars were present.
- For simply supported purlins or purlins with antisag bars, the effect of the warping stress on the critical load for lateral-torsion buckling is almost negligible.
- As far as lateral-torsion buckling is concerned, the ideal loading place is at the web central line as this leads to the highest values of the critical load.

References

[1] Hancock, G., 1997, "The Behaviour and Design of Cold-Formed Purlins," Steel Construction, **15**(3), pp. 2–16.
[2] Davies, J. M., 2000, "Recent Research Advances in Cold-Formed Steel Structures," Journal of Constructional Steel Research, **55**, pp. 267–288.
[3] British Standards Institution, 1987, BS5950, Structural Use of Steel in Building, Part 5. Code of Practice for Design of Cold-Formed Sections, BCI.
[4] European Committee for Standardization, 1996, EuroCode 3, Design of Steel Structures, Part 1.3: General Rules for Cold Formed Thin Gauge Members and Sheeting.
[5] American Iron and Steel Institute, 2001, North American Specification for the Design of Cold-Formed Steel Structural Members (draft edition).
[6] Standards Australia/Standards New Zealand, 1996, Cold-Formed Steel Structures, AS/NZS 4600.
[7] Winter, G., 1947, Strength of Thin Steel Compression Flanges, Cornell University Engineering Experimental Station, Print No. 32.
[8] Li, L. Y., 2003, "Lateral-Torsion Buckling of Cold-Formed Zed-Purlins Partial-Laterally Restrained by Metal Sheeting," Thin-Walled Struct., (in press).
[9] Timoshenko, S. P., and Gere, J. M., 1961, *Theory of Elastic Stability*, McGraw-Hill, New York.
[10] Tarnai, T., 1979, "Variational Methods for Analysis of Lateral Buckling of Beams Hung at Both Ends," Int. J. Mech. Sci., **21**, pp. 329–337.
[11] Reissner, E., 1989, "Lateral Buckling of Beams," Comput. Struct., **33**(5), pp. 1289–1306.
[12] Lee, J., and Kim, S. E., 2002, "Lateral Buckling Analysis of Thin-Walled Laminated Channel-Section Beams," Compos. Struct., **56**, pp. 391–399.

Effect of Loop Shape on the Drag-Induced Lift of Fly Line

Caroline Gatti-Bono
e-mail: gattic@umich.edu
Applied Numerical Algorithms, Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

N. C. Perkins
e-mail: ncp@umich.edu
Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109-2125, Fellow ASME

This note explains why casting a loop with a positive angle of attack is advantageous in distance fly casting. Several loop shapes, one with a positive angle of attack, one with a negative angle of attack, and two symmetrical loops with zero angle of attack are studied. For each loop, we compute the vertical drag component, i.e., the "lift." It is found that a loop with a positive angle of attack generates lift about four times larger than a symmetrical loop. Thus, loops with positive angles of attack stay "aerialized longer" which is consistent with observations made by (competition) distance fly casters. [DOI: 10.1115/1.1778414]

1 Introduction

Fly casting involves considerable mechanics of both the fly rod and fly line as described in several studies, [1–9]. For instance, the angler imparts both rigid and flexible body motions of the fly rod in accelerating the fly line during the forward and back casting strokes. These strokes end with an abrupt deceleration of the fly rod, often referred to as the ''stop,'' after which a ''loop'' of fly line is formed as shown in the photograph below. This loop propagates as a nonlinear wave under the action of fly line tension, air drag and gravity. The initial conditions that form the initial shape and velocity of this loop are generated during a short time interval following the stop, [10]. Eventually the loop propagates to the end of the fly line and the attached ''leader'' and ''fly'' turn over as the line straightens at the end of the cast.

The dynamics of this loop is the subject of a number of studies beginning with those that assume idealized semi-circular or square loop shapes, [1–3]. Further studies, [5–9], relax these assumptions and compute the loop shape from the kinematics of the attached fly rod and the equations of motion of the flexible fly line. Fly casting experts (see, for example, [11–13]) are fully aware that the shape of the loop has considerable influence on its dynamics. For instance, it is well understood that loops with smaller diameters propagate farther as they provide less projected area in the flow, hence less air drag. As a result, casting small loops are a distinct advantage when casting longer distances or into a head wind.

The purpose of this note is to explain a second advantage that results from casting loops that are asymmetrical and with a positive angle of attack such as shown in Fig. 1. Fly casting loops of this form have a ''pointed'' top portion followed by a larger and more rounded ''belly'' beneath. The belly forms a positive angle of attack in the flow and the air drag along this portion of the loop generates a component that is vertically upwards. The upward drag component acts opposite gravity and allows these loops to propagate farther. Fly casters sometimes say these loops ''stay aerialized'' longer and also use the term ''climbing loops'' to describe this effect. The purpose of this note is to explain this observed phenomenon by analyzing the drag on a loop as a function of its shape. Below, we consider four qualitatively distinct loop shapes and compare their vertical drag components.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, August 28, 2003; final revision, December 16, 2003. Associate Editor: O. M. O'Reilly.

2 Analysis of Loop Drag

Four loop shapes are illustrated in Fig. 2. These loops have identical length $l=\pi R$ and diameter $2R$, and are assumed to propagate to the right without changing shape. Two loops are asymmetrical (Figs. 2(*a*) and 2(*b*)) and two are symmetrical (Figs. 2(*c*) and 2(*d*)). The first asymmetrical loop (Fig. 2(*a*)) has a positive angle of attack and is termed a ''climbing loop,'' while the second (Fig. 2(*b*)) has a negative angle of attack and is termed a ''falling loop.'' The two symmetrical loops consist of a ''pointed loop'' and a ''circular loop'' as illustrated in Figs. 2(*c*) and 2(*d*), respectively. It should be noted that the shapes of these loops do not satisfy the steady-state conditions. However, they are close to the shapes observed in real casts and, therefore, they do provide good insight into the mechanics of a cast. Below we demonstrate how the loop shape significantly affects the vertical component of drag on fly line. We begin by computing the velocity field for an arbitrary loop shape.

2.1 Velocity Field. Figure 3 illustrates an arbitrary loop shape and a control volume that travels with the loop with velocity $\mathbf{v}_{cv}=v_{cv}\mathbf{i}$. The upper portion of the loop (assumed horizontal) travels with velocity $\mathbf{v}_o=v_o\mathbf{i}$ and the velocity of the bottom portion of the loop (assumed horizontal) is zero as it is attached to the end of the stationary fly rod (and the effects of gravity are neglected). The velocity of an arbitrary material point P relative to the control volume is denoted as $\mathbf{v}_r$ and its magnitude is uniform along the loop since the loop does not deform as it propagates. The absolute velocity of point P is

$$\mathbf{v}_p=\mathbf{v}_{cv}+\mathbf{v}_r \tag{1}$$

The velocities of the material points coincident with the top point (AA) and bottom (point BB) of the loop are

$$\mathbf{v}_{AA}=v_o\mathbf{i} \tag{2}$$

$$\mathbf{v}_{BB}=\mathbf{0}, \tag{3}$$

respectively. Using Eqs. (2) and (3) in Eq. (1) leads to the conclusion that

$$v_{cv}=v_r=\frac{v_o}{2} \tag{4}$$

Therefore,

$$\mathbf{v}_p=\frac{v_o}{2}\mathbf{i}+\frac{v_o}{2}\mathbf{e}_t \tag{5}$$

where $\mathbf{e}_t$ is the unit tangent vector to the loop at P. Thus, given the shape of the loop, Eq. (5) can be used to evaluate the velocity of an arbitrary material point. Resolving this velocity into components tangential and normal to the loop allows one to compute the drag components due to skin friction and form drag as follows.

2.2 Vertical Drag Component. The drag on an element of fly line derives from skin friction (tangent to the element) and form drag (normal to the element). The drag coefficients for skin friction and form drag are denoted by C_{dt} and C_{dn}, respectively. Let ρ_a denote the density of air and let d denote the diameter of the fly line (considered uniform for this example).

The loops shown in Fig. 2 are composed of straight segments and circular segments that, in total, subtend a semi-circle. For a straight segment of length l_s, the vertical component of drag is

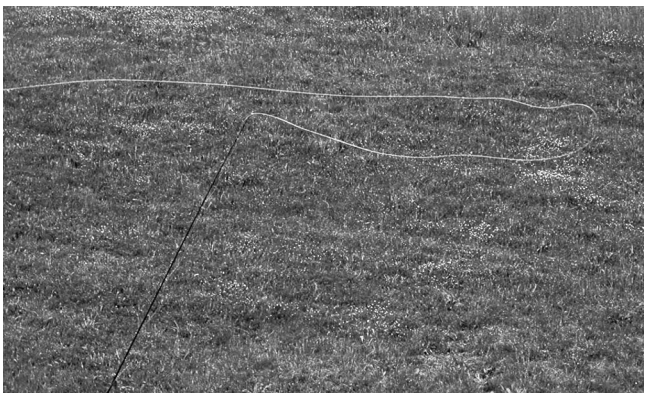


Fig. 1 The fly line "loop" is formed after the "stop" in a casting stroke and propagates as a nonlinear wave. This loop is asymmetrical and possesses a positive angle of attack. Such loops are a hallmark of expert fly casters.

$$D_Y = -\frac{1}{2}\rho_a \pi d l_s C_{dt} v_t |v_t| \mathbf{e}_t \cdot \mathbf{j} - \frac{1}{2}\rho_a d l_s C_{dn} v_n |v_n| \mathbf{e}_n \cdot \mathbf{j} \quad (6)$$

where v_t and v_n are the velocity components tangential and normal to the fly line, respectively (see, for example, [14]). Note the contributions of both skin friction and form drag to this result. For the (sum of) circular segments of radius R, the vertical component of drag is

$$D_Y = \int_0^{\pi} \left[-\frac{1}{2}\rho_a \pi d R C_{dt} v_t |v_t| \mathbf{e}_t \cdot \mathbf{j} - \frac{1}{2}\rho_a d R C_{dn} v_n |v_n| \mathbf{e}_n \cdot \mathbf{j} \right] d\theta$$

$$= \frac{\pi}{3}\rho_a d R v_o^2 C_{dt} \quad (7)$$

and this drag contribution depends only on the skin friction. We now employ Eqs. (6) and (7) to compute the vertical drag component for the four loops of Fig. 2.

Climbing Loop: For this loop, the straight segment $CD = (\pi^2 + 4/4\pi)R$ and the semi-circle of radius $R/2$, composed of the arcs BC and DE, are subject to drag. The total drag is expressed as

$$D_Y = \frac{1}{2}\rho_a d R v_o^2 \left(\frac{(\pi^4 + 8\pi^2 + 64)\pi}{3(\pi^2+4)^2} C_{dt} + \frac{\pi(\pi^2-4)}{(\pi^2+4)^2} C_{dn} \right) \quad (8)$$

and the contribution due to form drag on the straight segment (second term) is positive due to the positive angle of attack (Fig. 2(*a*)).

Falling Loop: For this loop, the straight segment $BC = (\pi^2 + 4/4\pi)R$ and the semi-circle of radius $R/2$, composed of the arcs AB and CD, are subject to drag. The total drag is expressed as

$$D_Y = \frac{1}{2}\rho_a d R v_o^2 \left(4\pi \frac{(\pi^4 + 2\pi^2 + 4)}{3(\pi^2+4)^2} C_{dt} - \frac{\pi(\pi^2-4)}{(\pi^2+4)^2} C_{dn} \right) \quad (9)$$

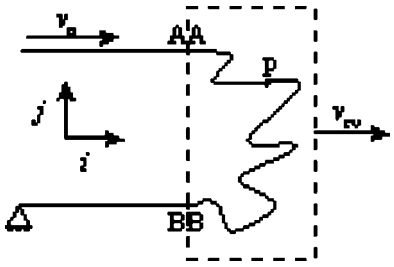


Fig. 3 Kinematics of arbitrary loop

and the contribution due to form drag on the straight segment (second term) is negative due to the negative angle of attack (Fig. 2(*b*)).

Pointed Loop: For this loop, the straight segments $BC = DE = \pi R/4$ and the semi-circle of radius $R/2$, composed of the arcs AB, CD, and EF, are subject to drag. The total vertical drag becomes

$$D_Y = \frac{1}{16}\rho_a d R v_o^2 \left(\frac{14\pi}{3} + \frac{2\pi(\pi^2-4)}{\pi^2} \right) C_{dt} \quad (10)$$

and the contributions due to form drag cancel due to symmetry.

Circular Loop: The result for this loop follows directly from Eq. (7)

$$D_Y = \frac{\pi}{3}\rho_a d R v_o^2 C_{dt} \quad (11)$$

and it is independent of form drag as mentioned above.

3 Example and Conclusions

Prior studies of fly line dynamics have used slightly different values for drag coefficients for skin friction and form drag, [1–3,5–9]. Here, we shall assume values $C_{dt} = 0.015$ and $C_{dn} = 1$ that are typical of those used in prior studies. We also recognize that these values depend, in general, on fly line speed (Reynold's number), [2]. Using these drag coefficients and the results above leads to the following table that compares the lift on the four loops shown in Fig. 2.

The results of Table 1 show that the lift generated by a climbing loop is approximately *four times greater* than that of a semi-circular loop with the same characteristic dimensions. The source of this additional lift is the contribution of form drag on the "belly" of the fly line that has a positive angle of attack. The negative angle of attack for the falling loop shape results in a net negative "lift," again due to the form drag on the belly. The symmetrical loops (circular and pointed) generate approximately the same lift. These results may be readily generalized to other loops shapes.

This note explains a fact observed by fly casting experts, namely, that a climbing loop is advantageous in distance casting. *How* to generate a climbing loop through control of the casting stroke is left as a (considerable) exercise to the reader.

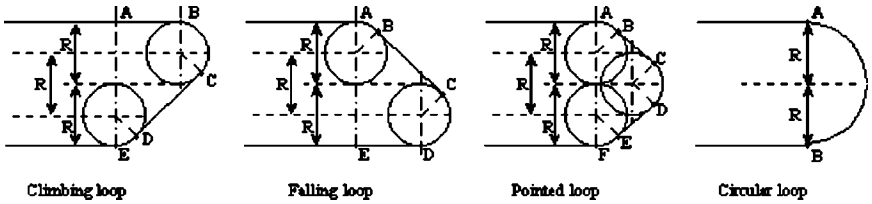


Fig. 2 Four qualitatively different loop shapes. (*a*) and (*b*) *l*=*AE*, (*c*) *l*=*AF*, and (*d*) *l*=*AB*

Table 1 Comparison of the vertical component of drag for the four loops shown in Fig. 2.

	Climbing loop	Falling loop	Pointed loop	Circular loop
$\frac{16D_y}{\rho_a d R v_o^2}$	0.924	−0.450	0.276	0.251

Acknowledgment

The authors wish to acknowledge the input of Mr. Bruce Richards from Scientific Anglers (3M) who first posed this interesting question to us.

References

[1] Spolek, G. A., 1986, ''The Mechanics of Flycasting: The Fly Line,'' Am. J. Phys., **54**(9), pp. 832–835.

[2] Lingard, S., 1988, ''Note on the Aerodynamics of a Fly Line,'' Am. J. Phys., **56**(8), pp. 756–757.

[3] Robson, J. M., 1990, ''The Physics of Flycasting,'' Am. J. Phys., **58**(3), pp. 234–240.

[4] Hoffmann, J. A., and Hooper, M. R., 1998, ''Fly Rod Response,'' J. Sound Vib., **209**(3), pp. 537–541.

[5] Hendry, M. A., and Hubbard, M., 2000, ''Dynamic Finite Element Simulation of Fly Casting and Its Potential Use in Fly Rod Design,'' *Proceedings: The Engineering of Sport*, Blackwell Science, London, pp. 407–414.

[6] Gatti-Bono, C., and Perkins, N. C., 2002, ''Physical and Numerical Modelling of the Dynamic Behavior of Fly Line,'' J. Sound Vib., **255**(3), pp. 555–577.

[7] Watanabe, T., and Tanaka, K., 2002, ''Modelling the Dynamics of a Fly Line,'' *Proceedings The Engineering of Sport*, Blackwell Science, London, pp. 353–359.

[8] Gatti-Bono, C., and Perkins, N. C., 2003, ''Comparison of Numerical and Analytical Solutions for Fly Casting Dynamics,'' J. Sports Eng., **6**(3), pp. 165–176.

[9] Gatti-Bono, C., and Perkins, N. C., 2003, ''Numerical Model for the Dynamics of a Coupled Fly Line/Fly Rod System and Experimental Validation,'' J. Sound Vib., in press.

[10] Perkins, N. C., and Richards, B., 2003, ''Dissecting Your Casting Stroke,'' Fly Fisherman, Primedia Magazines, Dec.

[11] Krieger, M. K., 1987, *The Essence of Flycasting*, Club Pacific, San Francisco, CA.

[12] Wulff, J., 1987, *Fly Casting Techniques*, The Lyons Press, Guilford, CT.

[13] Kreh, B., 1991, *Modern Fly Casting Method*, Odysseus Editions, Birmingham, AL.

[14] Sarpkaya T., and Isaacson M., 1981, *Mechanics of Wave Forces on Offshore Structures*, Van Nostrand Reinhold, New York.

Mitigating the Effects of Local Flexibility at the Built-In Ends of Cantilever Beams

Jonathan W. Wittwer

Larry L. Howell
e-mail: lhowell@byu.edu

Department of Mechanical Engineering, Brigham Young University, 435 CTB, Provo, UT 84602

Local distortion at the built-in ends of cantilever beams can lead to significant errors when models assume the support to be perfectly rigid. This paper presents a novel approach for mitigating this effect, using appropriately sized fillets to provide the additional stiffness needed to make simplified models more accurate and reduce stress concentrations. The optimal nondimensional fillet radius, called the optimal fillet ratio, is shown to be nearly constant for a wide range of geometries under predominantly bending loads, making it a useful parameter in the design of planar monolithic flexible mechanisms. [DOI: 10.1115/1.1782913]

1 Introduction

Flexible or compliant mechanisms are popular components in a large variety of precision machinery and instruments, [1]. While the use of compliance in precision machine design it not a new concept, [2], the past decade has seen a rapid expansion of methods for designing and analyzing flexible or compliant mechanisms, [3–5]. This paper focuses on analysis of monolithic planar compliant mechanisms, which are common components in precision devices and microelectromechanical systems (MEMS), such as folded-beam linear suspension springs and micro force gauges, [6–8].

It is still common to assume that flexible members are attached to perfectly rigid supports. However, the seminal works by O'Donnell [9] and Small [10] demonstrate through analysis and experiment that the local flexibility at the juncture of a support and a cantilever beam or plate can lead to a significantly larger deflection for a given load. The stress distributions in Fig. 1 show the local distortion occurring at the juncture. To account for this additional deflection, O'Donnell [11] and Matusz et al. [12] developed flexibility coefficients to use in a variety of classical equations for the deflections and stresses in beams.

This paper derives a novel approach to the analysis of planar loading of cantilever beams by using appropriately sized fillets to mitigate the effects of local elasticity in the support. It stands to reason that if the local flexibility of the support results in additional deflection from bending and shear loading, then adding material to the beam in some optimal geometry would provide the additional stiffness to mitigate the effect. It is common practice to use fillets to reduce stress concentrations and improve manufacturability, so the additional stiffness will be applied by optimizing the size of the fillet at the built-in end of the beam.

The optimal fillet radius is specified in terms of a nondimensional parameter called the *optimal fillet ratio*, ρ^*, which is equal to the fillet radius divided by the beam width (r/h). This parameter is shown to be nearly constant for a wide range of beam geometries under predominantly bending loads. This discovery makes it a useful parameter in the design of planar monolithic flexible mechanisms. It enables the designer to use simplified analytical models during design that do not account for localized distortion, while achieving accurate predictions by specifying the appropriate fillet prior to manufacture.

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 4, 2003, final revision, April 23, 2004. Associate Editor: S. Govindjee.

2 FEA Model Setup

Figure 2(*a*) shows the FEA model used to simulate pure bending at the juncture of the beam and an elastic half-plane. The deflection at the reference point (point 1) was compared to the vertical deflection of a moment end-loaded cantilever beam with a perfectly rigid support (Fig. 2(*b*)). A two-dimensional eight-node structural solid element was used in all the FEA models. To simulate a moment applied at point 1, a couple was applied to an extension of the beam far enough from the point of interest so that the local distortions at the application of the forces are insignificant. A similar setup was used for pure bending of a beam attached to an elastic quarter-plane (not shown).

To ensure that the loading conditions were appropriate and to determine an appropriate element size, FEA models were tested using the same boundary condition as in Fig. 2(*b*). Using a mapped mesh with a basic element size of $h/6$ was sufficient to make the systematic error insignificant.

It is common in compliant mechanisms to have both bending and shear loading of beams, so models were created to simulate vertically end-loaded cantilever beams. The quarter-plane model for this loading condition is shown in Fig. 3.

A good approximation to an infinite plane can be achieved using a large finite plane size, a and b (see Fig. 2(*a*) and Fig. 3(*a*)). Following the procedure used in [12], an appropriate value was determined by increasing the plane size until the sensitivity of the results to the plane size was insignificant. As in [12], a value of $a/h \geqslant 5$ was found to be sufficient for the quarter-plane, but the value used in this study was $a/h=15$ in order to reduce the systematic error to less than 0.1%. A value of $a/h \geqslant 9$ with $b=(a/2)-h$ was found to be sufficient for the half-plane (Fig. 3(*a*)), but a conservative value of $a/h=20$ was used. In each model, the mesh was refined around the points of high stress in order to obtain a more accurate determination of the stress concentration factor.

3 Formulation of Simplified Analytical Model

Using Castigliano's displacement theorem for the analysis of an end-loaded rectangular cantilever beam, the vertical deflection, δ, due to an applied moment or shear force are given by the following equations:

$$\delta=\left(\frac{6L^2}{Ewh^3}\right)M=C_M M \tag{1}$$

$$\delta=\left(\frac{4L^3}{Ewh^3}+\frac{12(1+v)}{5Ewh}\right)F=C_F F \tag{2}$$

where E is the elastic modulus, v is Poisson's ratio, w is the beam width, h is the beam thickness, and L is the beam length. The general assumptions are that deflections are small, cross sections remain plane, and the material is linearly elastic, isotropic, and homogenous. For small deflections, the deflection is proportional to the applied force, with a proportionality constant, C, commonly termed the *compliance*, which is the inverse of the *spring constant* or *spring rate*.

The optimal fillet is determined based upon a comparison of Eq. (1) or (2) with the results obtained from finite element analysis of a cantilever attached to an elastic half-space or quarter-space as shown in Fig. 2 and Fig. 3. The results are given in terms of an error in the compliance, $\varepsilon=(C-C_a)/C_a$ or spring rate $\varepsilon=(k_a-k)/k$, where the subscript, a, refers to the ''actual'' value as simulated using the FEA model. The error ε will often be reported as a percent. The *optimal fillet* is defined herein as the fillet that reduces this error to zero.

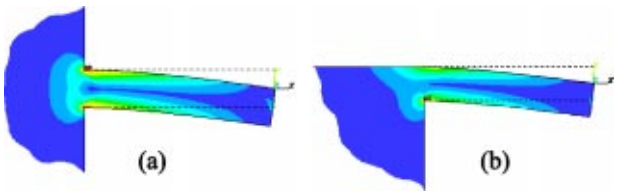


Fig. 1 Stress distribution at the juncture of a flexible beam and (*a*) an elastic half-plane and (*b*) an elastic quarter-plane

4 Discussion of Results

The results found that the percent error was only significantly affected by two non-dimensional geometric parameters: the slenderness ratio (L/h) and the fillet ratio $\rho=(r/h)$.

Figure 4 shows graphs of error versus fillet ratio for the cases involving pure bending for both the half-plane (Fig. 4(*a*)) and quarter-plane (Fig. 4(*b*)) models. Each line on the graph represents a specific slenderness ratio. A positive percent error means that the beam will have a larger deflection than predicted when using Eq. (1). The point where a line crosses the *x*-axis is defined as the *optimal fillet ratio*, ρ^*, for that slenderness ratio. At this point, the results from the FEA model match the simplified analytical equation.

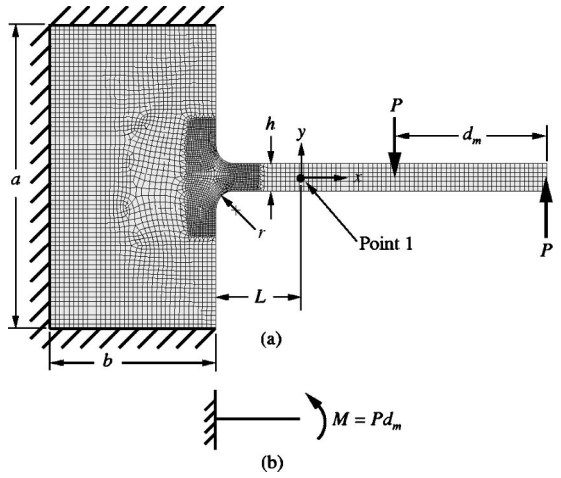


Fig. 2 (*a*) FEA model and (*b*) simplified model for simulating a constant-moment end-loaded cantilever beam of length, *L*, attached to an elastic half-plane

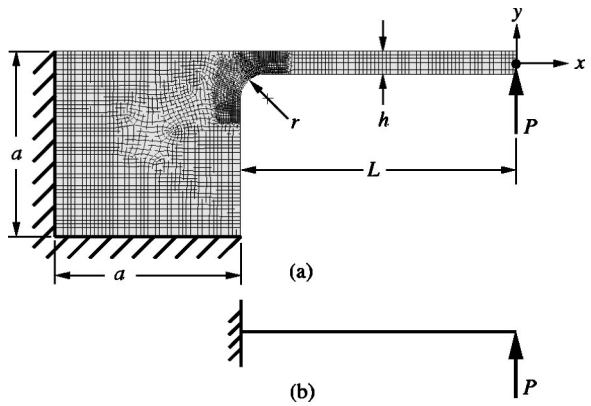


Fig. 3 (*a*) FEA model and (*b*) simplified model for simulating a vertically end-loaded cantilever beam of length, *L*, attached to an elastic quarter-plane

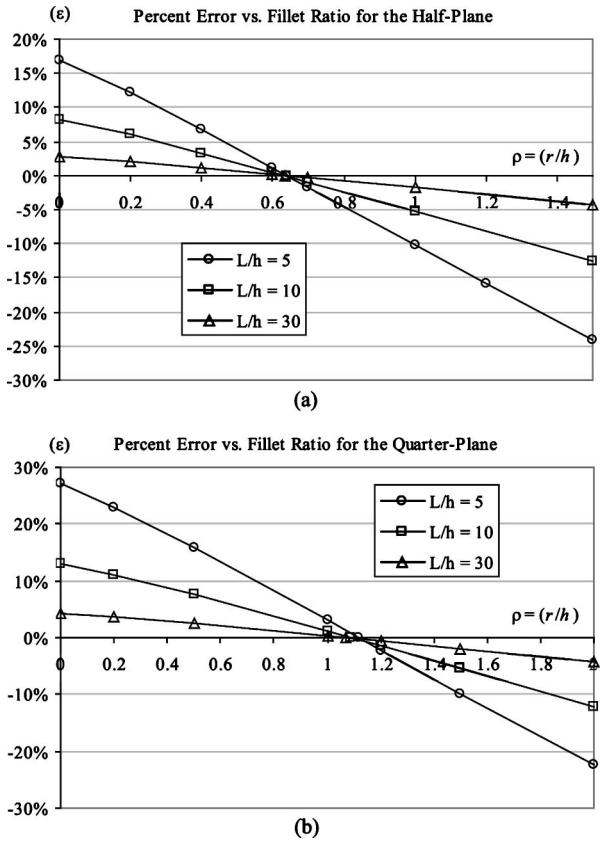


Fig. 4 Plot of the percent error versus the fillet ratio and slenderness for the (*a*) half-plane model, and (*b*) the quarter-plane model, under pure bending

These graphs show that for relatively small fillets, the error due to local elasticity in the support can be highly significant, especially for non-slender beams. The serendipitous discovery was that each of the lines in the plots intersected the *x*-axis at nearly the same point. Or, in other words, ρ^* is nearly identical for each geometry.

When the sufficient conditions are met for the plane size (as discussed earlier), ρ^* has only a slight dependence on the slenderness as shown in Fig. 4. Because the optimal values are very similar and the sensitivity to variation is small, a sufficient approach is to use the same fillet ratio regardless of the slenderness. Using the graphs in Fig. 4, appropriate approximations of ρ^* for the half-plane and quarter-plane are $\rho_H=0.64$ and $\rho_Q=1.1$, as summarized in Table 1.

The geometric stress concentration factor is usually a concern in design for determining both a static safety factor, and for calculating an estimate of the fatigue life. Most charts of geometric stress concentrations are only for fillet ratios below 0.3, since values larger than that result in very small or negligible stress concentrations, [13]. When using the ratios $\rho_H=0.64$ and $\rho_Q=1.1$, the results obtained from the finite element model for pure bending show essentially no stress concentration ($K_t=1.0$). This shows that not only does using the optimal fillet ratio miti-

Table 1 Approximations of ρ^* for two juncture types

Juncture Type	$\rho=r/h$
Half-plane	0.64
Quarter-plane	1.1

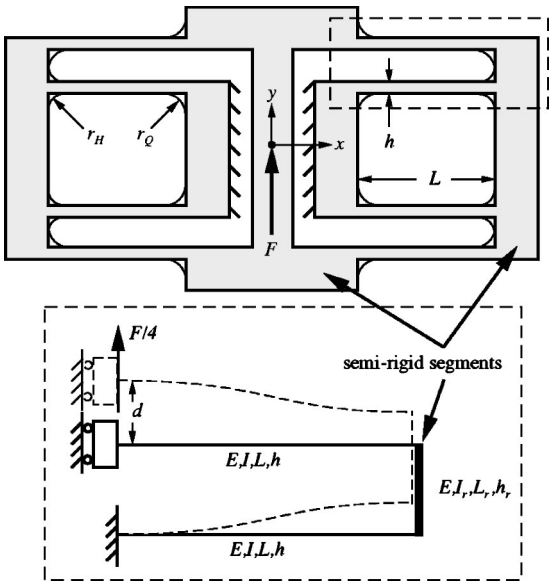


Fig. 5 Schematic for a folded-beam linear suspension and the simplified model for applying Castigliano's method to obtain the spring constant

gate the effects of the flexibility of the juncture, but it also practically eliminates the geometric stress concentration.

For the cases involving nonuniform bending (Fig. 3), the results showed the same trends as those for pure bending in that the intersections of the lines and the *x*-axis were nearly identical. The important result was that the optimal fillet ratios were all fairly close to the same values as those for pure bending, namely $\rho_H = 0.64$ and $\rho_Q = 1.1$. Combined loading involving bending and shear makes up a large portion of problems in compliant mechanisms analysis, so the fact that the same fillet ratio works to mitigate the effect of local elasticity for *both* loading conditions is advantageous. In addition, the geometric stress concentration for vertical end-loading was less than 1.06 when the optimum fillet ratio was used.

5 Example

A device used in both precision instrumentation and in microelectromechanical systems (MEMS) is a folded-beam linear suspension, [3,7,8]. A schematic for the suspension is shown in Fig. 5 along with the corresponding simplified model that can be used to obtain the spring constant. This suspension uses a combination of half-plane and quarter-plane junctures. These junctures do not represent beams attached to infinite half-planes or quarter-planes, but the method of using optimal fillets is fairly robust to these boundary conditions when the rigid members are over five times the thickness of the beams ($h_r/h > 5$). From Castigliano's method, the compliance of this linear suspension $C_{d,F}$ for a load F and deflection d, applied in the *y*-direction (see Fig. 5), is

Table 2 Parameter values used in Example 1

Variable	Value	Units
E	162000	MPa
v	0.22	—
w	3.5	μm
L	75	μm
h	3	μm
L_r	9	μm
h_r	18	μm
r_H	1.9	μm
r_Q	3.3	μm

Table 3 Percent difference between analytical and finite element results for Example 1. Units for the spring constant are μN/μm.

Model		FEA w/out Fillets	FEA Using ρ^*
	$k_{d,F}$	63.1	72.0
Simplified, Eq. (3)	72.1	−12.5%	−0.14%
Beam element	71.6	−11.9%	0.56%

$$C_{d,F} = \frac{1}{wE}\left(\frac{L^3}{2h^3} + \frac{6(1+v)L}{5h} + \frac{L_r}{2h_r} + \frac{3L^2L_r}{2h_r^3}\right) = \frac{1}{k_{d,F}} \quad (3)$$

where w is the out-of-plane thickness of the suspension, and the last two terms represent the compliance of the semi-rigid segment as shown in Fig. 5. A complete beam-element model including shear effects was used to validate the assumptions made in deriving Eq. (3).

Two detailed finite element models using eight-node structural solid elements were made for the linear suspension. The first was modeled without fillets, and the second was modeled using fillets based upon the values $\rho_H = 0.64$ and $\rho_Q = 1.1$. These FEA models, which *can* account for local elasticity, serve as benchmarks for comparison to Eq. (3) and the beam-element model. Table 2 lists the values of the variables used in this example.

The results for this example are summarized in Table 3. The results indicate that when designing the mechanism using models that *cannot* account for local elasticity, manufacturing the mechanism using the estimated optimal fillet ratios $\rho_H = 0.64$ and $\rho_Q = 1.1$ may lead to a nearly insignificant percent error. If the mechanism were to be made without fillets, the FEA simulation indicates that the percent error would be as much as two orders of magnitude higher. However, one should also consider that manufacturing limitations usually result in a minimum fillet radius, which for surface micromachined MEMS is usually about 1 micron (making the fillet ratio $r/h = 0.33$ for this example). This would still result in a significant percent error, but instead of building a more complicated model to predict the behavior of the real device, the appropriate fillet can be designed into the actual mechanism to make it behave more like the simplified or beam-element model. In the process, the geometric stress concentrations at these junctures are eliminated through the use of relatively large fillets.

Acknowledgments

The authors gratefully acknowledge the support of the National Science Foundation through NSF grant CMS-9978737, and an NSF graduate research fellowship.

References

[1] Smith, S. T., and Chetwynd, D. G., 1992, *Foundations of Ultraprecision Mechanism Design*, Gordon and Breach Science Publishers, New York.
[2] Eastman, F. S., 1937, "The Design of Flexure Pivots," J. Aerosp. Sci., **5**(1), pp. 16–21.
[3] Howell, L. L., 2001, *Compliant Mechanisms*, John Wiley and Sons, New York.
[4] Lobontiu, N., 2003, *Compliant Mechanisms: Design of Flexure Hinges*, CRC Press, Boca Raton, FL.
[5] Smith, S. T., 2000, Flexures: *Elements of Elastic Mechanisms*, Gordon and Breach Science Publishers, New York.
[6] Wittwer, J. W., Gomm, T., and Howell, L. L., 2002, "Surface Micromachined Force Gauges: Uncertainty and Reliability," J. Micromech. Microeng., **12**(1), pp. 13–20.
[7] Jaecklin, V. P., Linder, C., De Rooji, N. F., and Moret, J.-M., 1993, "Comb Actuators for XY-Microstages," Sens. Actuators, **39**(1), pp. 83–89.
[8] Zhou, G., Low, D., and Dowd, P., 2001, "Method to Achieve Large Displacements Using Comb Drive Actuators," Proc. SPIE, Bellingham, WA, **4557**, pp. 428–435.
[9] O'Donnell, W. J., 1960, "The Additional Deflection of a Cantilever due to the

Elasticity of the Support,'' ASME J. Appl. Mech., **27**, pp. 461–464.
[10] Small, N. C., 1961, ''Bending of a Cantilever Plate Supported From an Elastic Half-Space,'' ASME J. Appl. Mech., **28**, pp. 387–394.
[11] O'Donnell, W. J., 1963, ''Stresses and Deflections in Built-in Beams,'' J. Eng. Ind., **85**(3), pp. 265–273.
[12] Matusz, J. M., O'Donnell, W. J., and Erdlac, R. J., 1969, ''Local Flexibility Coefficients for the Built-In Ends of Beams and Plates,'' J. Eng. Ind., **91**(3), pp. 607–614.
[13] Pilkey, W. D., 1997, *Peterson's Stress Concentration Factors*, John Wiley and Sons, New York.