Transactions of the ASME®

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published quarterly (Mar., June, Sept., Dec.)
The American Society of Mechanical Engineers,
Three Park Avenue, New York, NY 10016.
Periodicals postage paid at New York, NY and additional mailing office. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.
CHANGES OF ADDRESS must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address.
STATEMENT from By-Laws. The Society shall not be responsible for statements or opinions advanced in papers or . . . printed in its publications (B7.1, Para. 3).

Journal of Applied Mechanics

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 67 • NUMBER 2 • JUNE 2000

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 67, Number 2 JUNE 2000

BRIEF NOTES

BOOK REVIEW

ANNOUNCEMENTS AND SPECIAL NOTICES

Shell Stability Related to Pattern Formation in Plants

C. R. Steele
Professor,
Fellow ASME
Division of Mechanics and Computation,
Department of Mechanical Engineering,
Stanford University,
Stanford, CA 94305

In the last few years we have studied the possible relation of instability of a shell surface to the patterns that develop in plants. In the present work, it is found that there is a linear relation between the epidermis (tunica) thickness and the wavelength between new leaves (primordia). This relation is near the buckling wavelength calculated from the geometry of the tunica and interior (corpus) cells. The main focus is on the changes in pattern that occur. (1) The wild variety of snapdragon has primordia that bulge out of plane, while a mutant has in-plane folding. A crude mechanical model is an elastic ring constrained at the outer diameter and subjected to uniform growth, represented by thermal expansion. It is found that the difference in the in-plane and out-of-plane buckling can be accounted for by a modest change in one geometric parameter. (2) The second change is that in the unicellular alga Acetabularia. *The geometry consists of a standard cylindrical pressure vessel with a nearly hemispherical end cap. At a point in time, the end cap flattens and a uniform circumferential array of new shoots forms. A mechanical model for the growth is proposed, in which the wall consists of a viscous material with a locally linear relation between mean stress and creep (growth) rate. The result is that the elliptical shape for stable growth can be regulated by one parameter of viscosity. The results reinforce the suggestion that the stability of the surface is instrumental in the generation of plant patterns, and that substantial change in pattern can be controlled by the modification of few mechanical parameters.* [S0021-8936(00)03002-6]

1 Introduction

The pattern of plant leaves and other organs (phyllotaxis) is intriguing. The exact number of different phyllotaxes occurring in nature is not known, because new patterns are continuously reported. However, every pattern is one of two types. In the one, the primordia compose a double set of spirals, such as seen in the pine cone, pineapple, or sunflower head. A most interesting feature is the number of spirals crossing a fixed radius. The numbers in the two directions are most commonly two successive terms in the Fibonacci series. Each new leaf is at the golden section of the angle between the nearest older primordia. In the second type of pattern, the whorl, the new primordium forms exactly between two older ones. Details of the classification and mathematical analysis of these patterns are in Jean [1], and much of the current thought on morphogenesis is in the collection of articles ([2]).

These patterns have aroused the scientific urge since antiquity. Prevalent today is the reaction-diffusion theory of Turing [3], which forms the basis of the discussion of morphogenesis by Harrison [4]. A disadvantage of the theory is that a fundamental ingredient, the *morphogens*, has not yet been identified. However, many interesting patterns resembling those in plants can be generated with solutions to the equations.

The view we have had is that the mechanical aspects of the cells and tissues may play a significant role in pattern formation. Surprising for most engineers is that the internal pressure (turgor pressure) inside plant cells is from 7 to 10 atmospheres. This provides important stiffening and can cause stability problems. It is difficult to imagine that nature would ignore this tremendous driving force available for pattern generation. Apparently Schwendender [5] first recognized this possibility. A recent survey of shell stability including several examples from nature, but not morphogenesis, is by Karam and Gibson [6,7]. Our basic premise is that for mechanical behavior both unicellular algae, such as *Acetabularia*, and multicellular plants, such as the sunflower, are equivalent to a pressure vessel, as indicated in Table 1.

Figure 1(*a*) shows the tip in a single cell plant *Acetabularia*. This is similar to many root hairs as well, in that growth takes place in the cap region as more of the cylindrical region is formed. At a certain point, *Acetabularia* has a transformation of the tip region from nearly hemispherical to elliptical, as shown in Fig. 1(*b*). When the ratio of the width to height a/b is greater then about 1.4, an array of lateral shoots is initiated, as shown in Fig. 1(*c*). Each of these will grow as in Fig. 1(*a*) and then perform the change in Fig. 1(*b*) to produce more lateral shoots. The main stem also repeats the cycle several times. Every pressure vessel engineer knows what happens when an elliptic head on a cylinder has a/b greater than 1.4. The circumferential membrane stress is compressive, and a ring of circumferential buckles, just as in Fig. 1(*c*), may occur. In pioneering work, Martynov [8] quantified the relation of elastic buckling to pattern formation. He showed that for *Acetabularia* the physical properties and dimensions are sufficient to produce elastic buckles that are predictive of the number of shoots. Just by flattening the end cap, one obtains pattern from no pattern because of the buckling.

So *Acetabularia* is a great motivation for considering the role of stability for other plant patterns. However, *Acetabularia* is also a great success for the reaction-diffusion theory. Harrison et al. [9] and Harrison and Hillier [10] find that the dependence of the pattern on temperature and calcium is close to that predicted by the theory. Dumais and Harrison [11] summarize the known facts and theories on *Acetabularia* and related algae and indicate that the strongest case can be made for the diffusion theory. However, the mechanical consequences of the change from Fig. 1(*a*) to Fig. 1(*b*) cannot be ignored. The complete theory will undoubtedly include both mechanical and electro-chemical effects.

A more complex shape is the sunflower, shown in Fig. 2(*a*). The double rows of spirals are clear. Typically in the sunflower, there are 55 spirals in the one direction and 89 in the other (two successive terms in the Fibonacci series). The distance between the primordia is around ten cells, so the pattern concerns the tissue and not individual cells. Hernández [12] shows a change of the

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

Table 1 Relation of pressure vessel to alga and plant

Pressure Vessel	*Acetabularia* (one cell)	Sunflower Head (many cells)
shell wall	cell wall	tunica layer
internal liquid/soft elastic material	cytoplasm/vacuole	corpus
viscous creep/thermal expansion	growth	growth

cross section of the sunflower head, which is similar to that of *Acetabularia*. An almost hemispherical dome flattens to the shape indicated in Fig. 2(*b*). Of interest is the annular region of negative Gaussian curvature. The calculations for this by Wu [13], using the Fast4 shell of revolution program, produced the stress resultants shown in Fig. 2(*c*), with a significant compressive circumferential stress in the negative curvature region. This is exactly the region in which the new primordia are forming. So the same principle seems to be working for the complex multicellular structure as for one cell. Compression is needed for the generation of pattern. A confirmation of the compression is shown in Fig. 2(*d*). Dumais [14] made diagonal cuts across the primordial sunflower head. The sides of the cuts in the center dome gape apart, indicating a region of tension, while the sides of the cuts in the annular region of new primordia formation remain pressed together, consistent with the calculation.

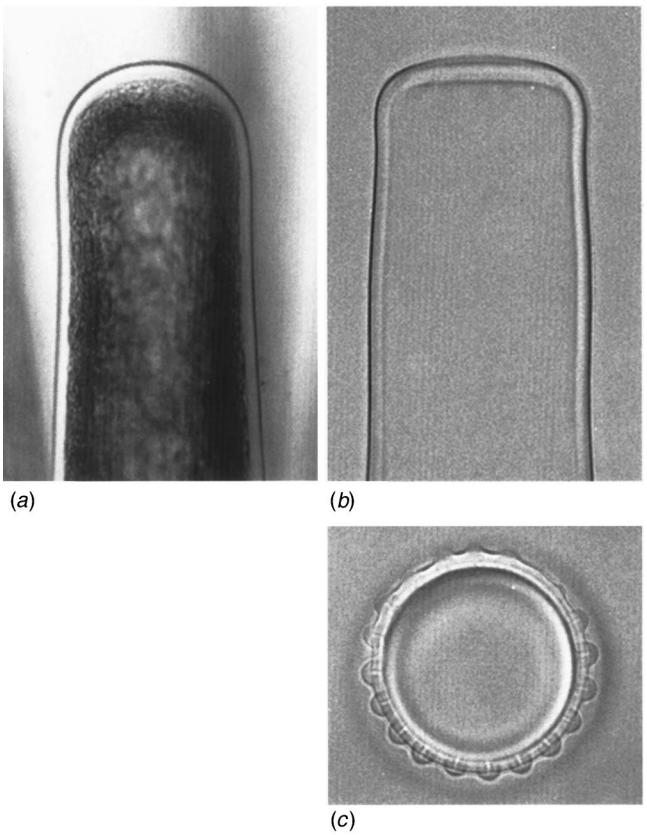
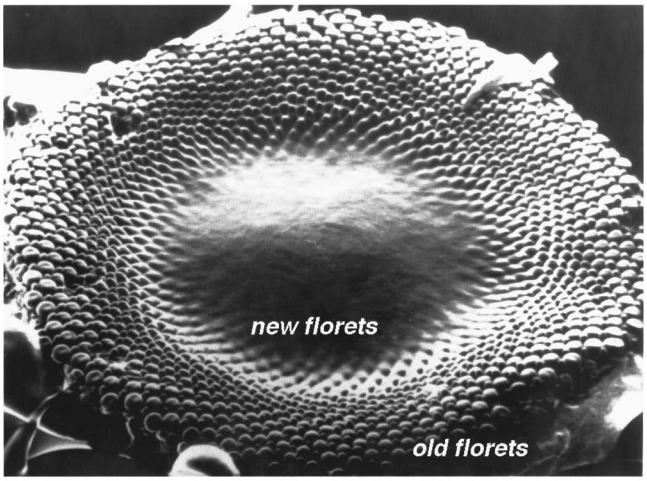


Fig. 1 (*a*) Tip growth in the unicellular alga *Acetabularia*. The growth takes place in the end cap, which is nearly spherical (elliptical with $a/b=1.2$ for this example). The cylindrical portion remains of constant diameter, equal to about 50 μm. The load-bearing wall, composed mostly of mannan polymers, can be seen as a thin transparent layer surrounding the tip. The cytoplasm (dark granular region) and the central vacuole exert a pressure of 7–10 atmospheres on the wall and thus provide the driving force for elongation of the cell. (Photograph from Dumais [14].) (*b*) In *Acetabularia*, the elongation stage stops at regular intervals and the end cap changes from nearly hemispherical to ellipsoidal ($a/b=1.9$ for this example). Here only the wall is shown. When the ratio of radial to axial semi-major axes of the ellipse reaches a value near 1.5, significant circumferential compression occurs, which causes buckling of the surface. (From Dumais and Harrison, [11].) This is just as in a standard thin-walled pressure vessel. (*c*) Axial view of *Acetabularia* after buckling. It appears that the compressive circumferential stress causes a buckling pattern that initiates the development of an equally spaced array of lateral hairs. (Such a symmetric pattern is called a whorl.) Subsequently, each lateral hair elongates as in (*a*). (From Dumais and Harrison [11].)

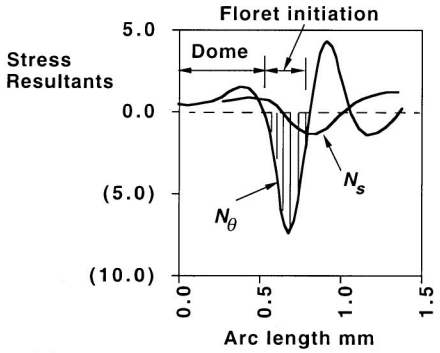
Wu [13] carried out a number of calculations for local buckling, some of which are reported in Green [15]. These were with Fast4 for the complete shell of revolution and for an elastic plate on an elastic foundation. The plate equations are the well-known von Kármán equations ([16]), but with an initial displacement that produces the nonlinear shallow-shell equations. Of interest is the result for a region with two initial hills, representing two primordia from the preceding generation. Adding sufficient compressive stress causes a buckle that appears approximately at the golden section between the two initial hills. Thus buckling seems to provide the local mechanism for generating the complex spiral patterns such as in Fig. 2(*a*). To see if this is effective on the global scale, subsequent calculations have been carried out by Steve Rennich on the complete circular plate, as discussed in Green et al. [17,18]. The spiral patterns can indeed be generated by successive buckling far into the post buckling regime, as shown in Fig. 3(*a*). The calculation involves a standard perturbation procedure with suitable scaling of the perturbation steps to restrict numerical instability. For the linear calculation at each step, a Fourier series in the circumferential direction and finite difference in the radial direction is used. Since the steps must be small, the total calculation requires substantial time (hours on a work station). Sometimes in the calculation, one set of the spirals would be lost and the pattern would degenerate into the ridges shown in Fig. 3(*b*). This seemed to be a severe defect of the simple plate model until the work by Carpenter et al. [19] appeared, that shows a mutant in snapdragon with this ridge behavior. So the stability calculation appears to have some predictive capability. A particular challenge to the stability analysis is the extreme pattern in *Costus*, described by Kirchoff and Rutishauser [20], which is an unusual ''spiral'' that looks more like a staircase. This is still a spiral in the generalized Fibonacci series (1,7,8, . . .), according to Jean [1].

The main lack is that we have not shown how the spiral pattern can originate either de novo or as a transition from a whorl, which is the normal behavior in plants. In calculations with an initial random displacement of the surface, the whorled pattern will appear but not spirals. Spirals can be obtained when appropriate boundary conditions are prescribed. Indeed, in the calculations, there seems to be nothing special about the Fibonacci pattern; other numbers of spirals can propagate equally well. In nature, there is an indication that if the spiral pattern is lost, it is difficult to reestablish. In the experiments of Hernández and Palmer [21], a circular cut was made in the central dome of the sunflower at some distance from the ring of the new growth. Fast4 calculations ([13]) show that circumferential compression occurs near the edges of the cut. So it is consistent that new growth occurs on both sides of the cut, but the spiral pattern seems to be lost. In his present work, Dumais has repeated the experiment and found a tendency for the spiral pattern finally to reestablish, although the number of spirals is not a Fibonacci number.

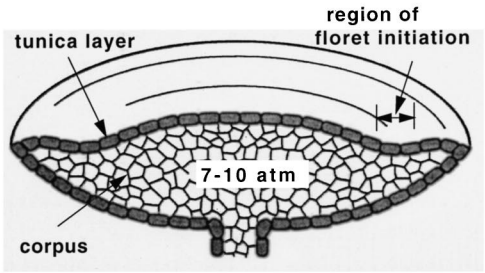
What triggers and controls these changes? The change from whorl to spiral is common, and a change from one type of spiral to another occurs in several plants. Kwiatkowska and Florek-Marwitz [22] catalog such transitions and show that they are related to changes in the area of the central dome. Thus the geometry is of importance. We wish to find to what extent such changes in form can be related to changes in material properties. As a beginning to the question, we address in this paper a mutation in



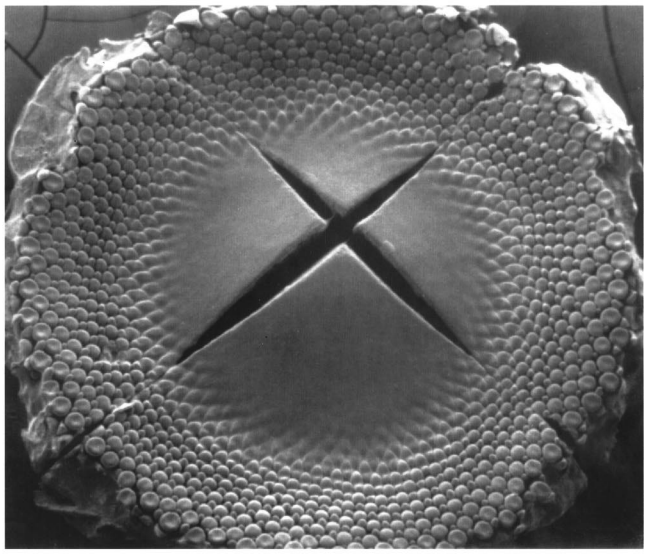
(*a*)



(*c*)



(*b*)



(*d*)

Fig. 2 (*a*) **Sunflower during the pattern development. The older florets are at the outer region, while the new florets are generated at the rim of the inner smooth dome. Each new floret is at the ''golden section'' between those of the older generation, which generates the spirals. (See Jean, [1].) At this stage, the sunflower head has a diameter of around 3–4 mm.** (*b*) **Shell model for the sunflower. The outer layer of cells (tunica) have substantially thicker walls to withstand the internal pressure, in comparison with the inner cells (corpus). The region of the floret initiation has negative Gaussian curvature of the tunica. (Geometry from Hernández [12].)** (*c*) **Stress calculated for the shell model of the sunflower. The region of floret initiation has substantial compressive stress in the circumferential direction shown by the shaded area. (From Wu [13].)** (*d*) **Effects of cuts across the sunflower capitulum. The tension regions gape open, but the region of the floret initiation is pressed together because of the compressive circumferential stress. (From Dumais [14].)**

snapdragon, and the change in *Acetabularia* from Fig. 1(*a*) to Fig. 1(*b*). As a preliminary, however, we consider the basic question of whether the geometry of the plants is consistent with the possibility of mechanical instability.

2 Local Stability of a Layer

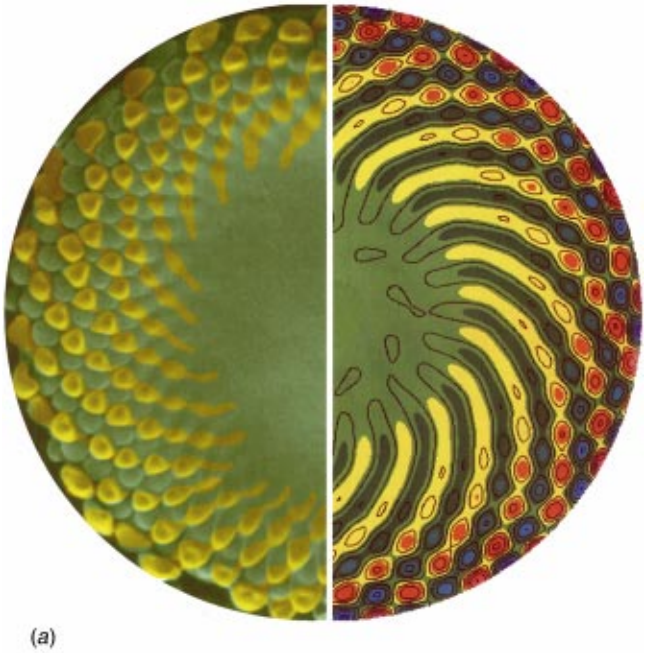
The emphasis in the preceding stability calculations was on the pattern and not on the physical relevance of the parameters. The key parameter is referred to in Green et al. [23,18] as the ''natural'' wavelength. This is the wavelength for minimum buckling load of a flat plate on an elastic foundation. Now, we wish to establish the physical basis for this natural wavelength. The model consists of a sandwich plate, representing the tunica, on a half-space, representing the corpus, as indicated in Fig. 4.

2.1 Half-Space Stiffness. In the preceding calculations the foundation stiffness was taken as constant. Since the buckling wavelength is usually small in comparison with the thickness of the corpus, the plant interior can be represented as a half-space. For a half-space with a sinusoidal deformation on the surface with the wavelength λ, the linear elastic equations can be solved for the relation between the surface stress components and the displacement components. The relation can be written as a matrix of surface stiffness coefficients. For the *z*-axis normal to the surface, the relation is for plane strain:

$$\begin{bmatrix}\sigma_{zz}\\ \sigma_{zx}\end{bmatrix}=\frac{2\pi E}{\lambda(1+\nu)(3-4\nu)}\begin{bmatrix}2(1-\nu) & 1-2\nu\\ 1-2\nu & 2(1-\nu)\end{bmatrix}\begin{bmatrix}w\\ u\end{bmatrix} \tag{1}$$

in which E is Young's modulus, and ν is Poisson's ratio. For plane stress the result is

$$\begin{bmatrix}\sigma_{zz}\\ \sigma_{zx}\end{bmatrix}=\frac{2\pi E}{\lambda(1+\nu)(3-\nu)}\begin{bmatrix}2 & 1-\nu\\ 1-\nu & 2\end{bmatrix}\begin{bmatrix}w\\ u\end{bmatrix}. \tag{2}$$



(a)

(b)

Fig. 3 (*a*) **Comparison of the actual sunflower (left) and computations (right). The computation is with the use of the von Kármán equations for a plate on an elastic foundation. The plate is compressed by a uniform radial edge force. The plate is initially flat, but edge conditions on the rotation are prescribed that have the spiral pattern. For subsequent buckling, the spiral pattern propagates toward the center, as shown. Typically, the distance between calculated buckles is about twice the natural wavelength. (Calculation by S. Rennich, color enhancement of figure by J. Dumais, from the cover of the *American Journal of Botany*, July, 1999.) (*b*) Calculation of plate post-buckling, in which spiral pattern degenerates into ridges. The distance between ridges is about equal to the natural wavelength. (Calculation by S. Rennich.)**

For an incompressible material $\nu = 0.5$ and for the tangential displacement u set to zero, Eq. (1) gives the ratio of the normal stress and the normal displacement, which is the equivalent elastic foundation stiffness k for the half-space:

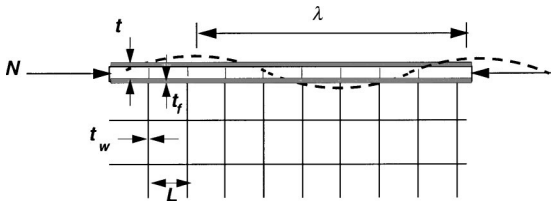


Fig. 4 Representation of a plant by an elastic layer, consisting of a sandwich plate representing the tunica, attached to an elastic half-space, representing the interior cells (corpus). The compressive force *N* in the tunica causes a buckling deformation, indicated by the dashed line, with the wavelength λ. The thickness of the tunica is *t*, the thickness of the walls of the tunica is t_f, the thickness of the interior cell walls is t_w and the cell diameter is *L*.

$$\frac{\sigma_{zz}}{w} = \frac{4\pi E}{3\lambda} = k. \tag{3}$$

Of course, the value of k is dependent on the wavelength λ.

2.2 Plate on Half-Space. For a sinusoidal displacement of a plate on an elastic foundation, the critical compressive force resultant is

$$N = D\left(\frac{2\pi}{\lambda}\right)^2 + \frac{k}{\left(\frac{2\pi}{\lambda}\right)^2} \tag{4}$$

where D is the effective bending stiffness of the plate. For the foundation stiffness k, a known value, the minimum of Eq. (4) occurs at the wavelength $\lambda = 2\pi\sqrt{D/k}$. If the foundation for the half-space Eq. (3) is used, then the minimum N occurs at the wavelength

$$\lambda = 2\pi\left(\frac{3D}{E_H}\right)^{1/3} \tag{5}$$

and has the value

$$N_{cr} = (3DE_H^2)^{1/3}. \tag{6}$$

For the dimensions in Fig. 4, and if the structural material in the walls of the cells and the plate is the same (cellulose), then the bending stiffness and half-space modulus are

$$D \approx \frac{2}{3}Et^2t_f \quad ; \quad E_H \approx E2\frac{t_w}{L} \tag{7}$$

and Eq. (5) gives the critical wavelength

$$\frac{\lambda}{t} = 2\pi\left(\frac{f}{3}\right)^{1/3} \tag{8}$$

where the geometric factor is

$$f \approx \frac{6t_f}{t}\frac{L}{t_w} \tag{9}$$

and Eq. (6) gives the critical compressive strain

$$\frac{N_{cr}}{E2t_f} = \left(\frac{36}{f^2}\right)^{1/3}. \tag{10}$$

2.3 Relation to Plant Patterns. In an unpublished study initiated by C. Schmid and continued by J. Dumais, micrographs available in the literature for a wide variety of plants were examined. The result for the wavelength, i.e., the distance between

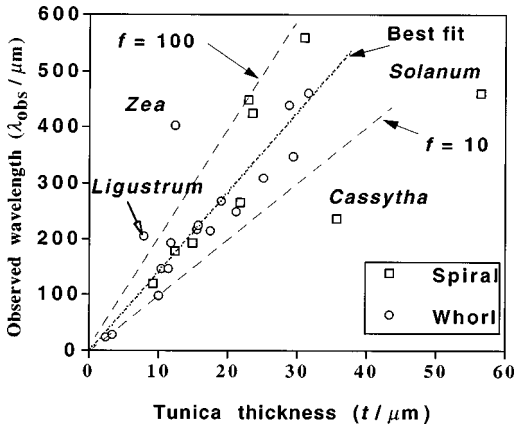


Fig. 5 Relation of observed peak to peak distance between primordia and the tunica thickness. The dotted line is the best fit, with a slope 14.1. The dashed lines show the natural wavelength for a plate on an elastic substrate with relative volume fractions of $f=10$ and 100. Generally, there seems to be little difference in the whorl and spiral patterns. (Data collected from literature by C. Schmid and J. Dumais.)

primordia, as a function of the total tunica thickness is shown in Fig. 5. Except for four plants, there is a remarkably linear correlation. A precise determination of thickness and the cellulose content of walls is impossible from the general micrographs of the cross section. However, it appears that the ratio of tunica thickness to wall thickness is greater than the ratio of interior cell wall to cell diameter, so a lower bound on the factor Eq. (9) is $f=6$. Added to Fig. 5 is the relation Eq. (8) for the ''reasonable'' value of $f=10$ and for a large value of $f=100$. These bound most of the measured values.

Generally, the wavelength between buckles in the calculation Fig. 3(*a*) is roughly twice the natural wavelength. So it seems that the results in Fig. 5 showing the actual pattern to be at a longer wavelength than the ''reasonable'' value of $f=10$ is consistent with the calculations. However, when the buckles degenerate into ridges such as shown in Fig. 3(*b*), the ridges are at the natural wavelength, since the solution is nearly one-dimensional between ridges. The actual dimensions for the developing plants Fig. 5 seem consistent with the concept that the pattern is a consequence of the stability of the surface.

3 Stability of Constrained Ring

As a first study of the possible relation of mechanical properties to form, the snapdragon is considered. In snapdragons, stamens originate as five symmetric vertical undulations in an annulus (Fig. 6(*a*)). In the *deficiens* mutation, a comparable annulus undulates horizontally, making a wavy ribbon with five folds (Fig. 6(*b*), as discussed by Green [24]). To see whether this change in form might be related to a change in the mechanical properties, we consider the stability of an elastic ring constrained from radial expansion, similar to a doughnut inside a rigid coffee cup. The wall of the cup restricts any outward radial displacement of the ring. In addition, a uniform elastic foundation connects the ring to the bottom of the cup, which constraints axial displacement. Growth is simulated by a uniform heating of the ring. As the temperature increases, the stress in the ring increases until a critical condition is reached, after which the ring will no longer remain in the original shape. The stress will be released by a deformation of the ring that could be out of the initial plane of the ring, as shown in Fig. 6(*c*) or that could consist of an in-plane deformation, as shown in Fig. 6(*d*). The following analysis is routine, but the results may be of significance in understanding the behavior of plants, so the details are included.

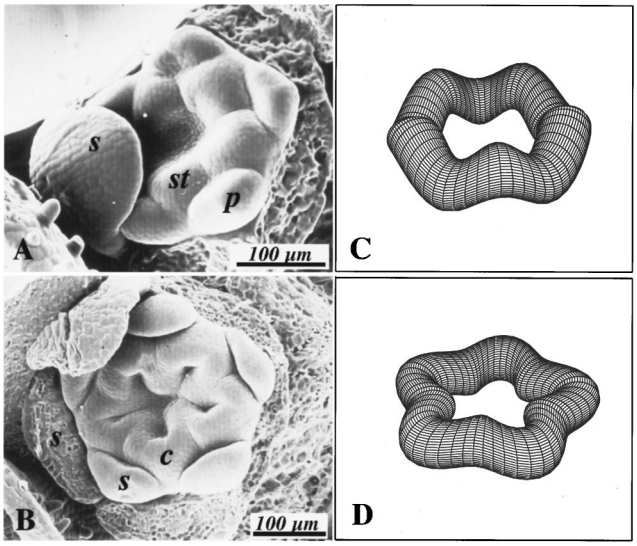


Fig. 6 Buckling and organ shape. (A) Wild-type stamens (st) in snapdragon. Creases traverse the formative zone, making a ring of humps. (B) The *deficiens* mutant of snapdragon. The formative region undulates in a plane as a ribbon, forming five cup-like primordia. (C) Out-of-plane buckling mode for a thick ring. (D) In-plane buckling of an elastic ring constrained by a wall at the outer margin. (From Green [24].)

3.1 Out-of-Plane Buckling. With the thermal heating of the constrained ring, it is possible for the ring to accommodate the excess strain by deforming out of the plane of the ring, as indicated in Fig. 6(*c*). The analysis is based on the total potential energy:

$$\Pi=\int_0^{2\pi}\left[\frac{1}{2}(EI_2\kappa_2^2+GJ\kappa_1^2+ku_3^2)-\varepsilon EA\alpha T\right]Rd\theta \qquad (11)$$

in which R is the radius of the ring, θ is the circumferential angle, E is the elastic Young's modulus, A is the area, α is the coefficient of thermal expansion, T is the temperature change, I_2 is the geometric moment of inertia of the cross section about the axis through the centroid parallel to the plane of the ring, and GJ is the torsional stiffness. The stiffness of the elastic foundation is k. The twist is given by

$$\kappa_1=\frac{1}{R}\left(\frac{d\chi_1}{d\theta}-\chi_2\right) \qquad (12)$$

in which χ_1 is the rotation of the cross section, and the rotation in the orthogonal direction is

$$\chi_2=-\frac{1}{R}\frac{du_3}{d\theta} \qquad (13)$$

where u_3 is the out-of-plane displacement. The change of curvature measure is

$$\kappa_2=\frac{1}{R}\left(\frac{d\chi_2}{d\theta}+\chi_1\right) \qquad (14)$$

and the strain is from moderate rotation theory

$$\varepsilon=\frac{1}{2}\chi_2^2. \qquad (15)$$

The assumed displacement is

$$u_3=U\cos n\theta \qquad (16)$$

with the rotation

$$\chi_1 = -\frac{U}{R}\cos n\theta \tag{17}$$

in which n is the harmonic index. The choice of χ_1 reduces to zero the high energy torsion term in Eq. (11), leaving the potential

$$\Pi = \pi R\left[EI_2(n^2-1)^2\frac{1}{R^4} + k - \frac{EA\alpha T n^2}{R^2}\right]\frac{U^2}{2}. \tag{18}$$

So for nonzero U, the critical temperature is

$$\alpha T = \frac{R^2}{EA}\left[\frac{k}{n^2} + \frac{EI_2}{R^4}\frac{(n^2-1)^2}{n^2}\right]. \tag{19}$$

For $n>1$, the approximate minimum is at the harmonic index

$$n = R\left(\frac{k}{EI_2}\right)^{1/4} = \left(\frac{R}{R_g}\right)^{1/2}\eta^{1/4} \tag{20}$$

which gives the minimum out-of-plane buckling temperature

$$\alpha T_{\text{op}} = \frac{2(kEI_2)^{1/2}}{EA} = 2\frac{R_g}{R}\eta^{1/2} \tag{21}$$

where R_g is the radius of gyration for the cross section and the dimensionless foundation stiffness factor is

$$\eta = \frac{kR^2}{EA} \tag{22}$$

3.2 In-Plane-Buckling. The analysis is based on the total potential energy given by

$$\Pi = \int_0^{2\pi}\left[\frac{1}{2}(EI_3\kappa_3^2 + EA\varepsilon^2) - \varepsilon EA\alpha T\right]Rd\theta \tag{23}$$

in which I_3 is the geometric moment of inertia of the cross section about the vertical axis. The strain from the moderate rotation theory is given by

$$\varepsilon = \frac{1}{R}\frac{du}{d\theta} + \frac{w}{R} + \frac{1}{2}\chi_3^2 \tag{24}$$

in which u is the displacement in the circumferential direction, w is the displacement in the radial direction, and the rotation is

$$\chi_3 = -\frac{1}{R}\frac{dw}{d\theta} + \frac{u}{R}. \tag{25}$$

The change of curvature measure is

$$\kappa_3 = \frac{1}{R}\frac{d\chi_3}{d\theta}. \tag{26}$$

The ring attempts to expand with increasing T, but is constrained by the wall $w\leqslant 0$. An approximation for the stability limit can be obtained by assuming a reasonable displacement shape function. The following consists of a uniform compression of the ring and an inextensional sinusoidal deformation:

$$w = W(-1+\cos n\theta) \tag{27}$$

$$u = -\frac{W}{n}\sin n\theta \tag{28}$$

in which W is the unknown amplitude. Substituting this deformation into the potential Eq. (23) yields the result

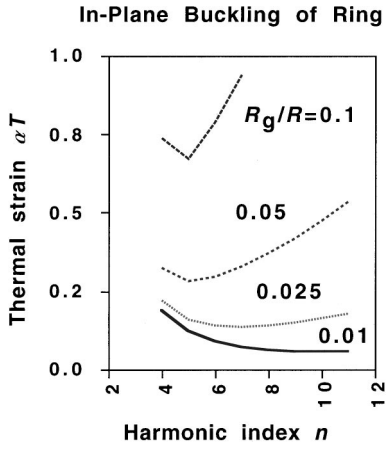


Fig. 7 Thermal strain to cause in-plane buckling as a function of the harmonic index. For a given value of geometry R_g/R, the minimum gives the critical condition. For the thicker rings, the critical condition stabilizes at the harmonic $n=5$.

$$\Pi = \pi R\left\{\frac{1}{2}EI_3(n^2-1)^2\left(\frac{W}{R^2}\right)^2 + EA\left[\left(\frac{W}{R}\right)^2 - \left(\frac{W}{R}\right)^3\frac{(n^2-1)^2}{2n^2} + 3\left(\frac{W}{R}\right)^4\frac{(n^2-1)^4}{32n^4}\right] - 2EA\alpha T\left[-\frac{W}{R} + (n^2-1)^2\left(\frac{W}{R^2}\right)^2\frac{1}{4n^2}\right]\right\}. \tag{29}$$

The potential should have a minimum value for the equilibrium solution. Setting the derivative of Π with respect to W to zero yields the relation between amplitude and temperature:

$$\frac{W}{R}\left[\frac{EI_3}{EAR^2}\frac{(n^2-1)^2}{2} + 1\right] - \frac{3}{4}\left(\frac{W}{R}\right)^2\frac{(n^2-1)^2}{n^2} + \frac{3}{16}\left(\frac{W}{R}\right)^3\frac{(n^2-1)^4}{n^4} = \alpha T\left[\frac{W}{R}\frac{(n^2-1)^2}{2n^2} - 1\right]. \tag{30}$$

The solution procedure is to fix the harmonic index n, then compute the thermal strain αT as a function of the displacement w/R. The minimum αT is the critical value at which the ring will change from the compressed state to the buckled state. In Fig. 7 is shown that minimum thermal strain for each harmonic index. The only parameter is the ratio of bending to stretching stiffness:

$$\frac{R_g}{R} = \left(\frac{EI_3}{EAR^2}\right)^{1/2}. \tag{31}$$

For a thin-walled ring of circular cross section with radius r, the ratio is

$$\frac{R_g}{R} = \frac{r}{R\sqrt{2}}. \tag{32}$$

The results in Fig. 7 show that the thicker rings, with larger values of the ratio, have a buckling mode with the harmonic $n=5$. For thinner rings, $R_g/R<0.025$, the critical value of thermal strain decreases and the harmonic increases. The values of thermal strain are excessive for the usual engineering problem, but for the analogous problem of growth, such values may be reasonable.

3.3 Condition for Buckling Mode Change. For simplicity, we consider the ring with circular cross section, for which I_2 and I_3 are equal. The critical temperature for in-plane instability Fig. 7 depends on only the parameter R_g/R, while the critical tempera-

ture for out-of-plane instability Eq. (21) depends on the additional parameter η. However, if the critical harmonic index n is prescribed, then Eq. (29) gives the necessary value of η, and the critical out-of-plane temperature also depends only on R_g/R:

$$\alpha T_{\text{op}} = 2n^2\left(\frac{R_g}{R}\right)^2. \tag{33}$$

For the harmonic $n=5$, the values of the out-of-plane buckling Eq. (33) are lower than the in-plane values Fig. 7. Thus the out-of-plane mode Fig. 6(*a*, *c*) will be observed. If η is increased with R_g/R held fixed, then the critical harmonic index n and temperature for out-of-plane buckling will increase. For the case of $R_g/R=0.05$, an increase in η of a factor of 4 is necessary to change from the out-of-plane buckling to the in-plane buckling at $n=5$. Similarly, for $R_g/R=0.1$, η must increase by the factor of 2. Thus a modest change in the ratio of ring stretching stiffness to foundation stiffness changes the buckling mode.

4 Model for Growth in *Acetabularia*

Now we show that the transition from Fig. 1(*a*) to Fig. 1(*b*) can also be obtained by changing one physical parameter. We consider the transformation of a shell from a hemispherical configuration, shown in Fig. 8(*a*), to the shape shown in Fig. 8(*b*), consisting of a cylindrical portion with an end cap of the same geometry as the initial. The initial and current configurations are indicated by the capital and lower case letters. This represents the steady-state phase of the growth of *Acetabularia* as well as typical root hairs ([25]). Hejnowicz et al. [26] provide an Eulerian formulation for the problem of the axisymmetric growth. However, we use a Lagrangian formulation. The current arclength s, measured from the bottom, is a function of the initial arclength S. The stretch ratio (growth) in the circumferential direction is

$$\lambda_\theta = \frac{r(s(S))}{R(S)} \tag{34a}$$

and the meridional stretch ratio is

$$\lambda_s = \frac{ds}{dS}. \tag{34b}$$

Denote the ratio of the stretches by

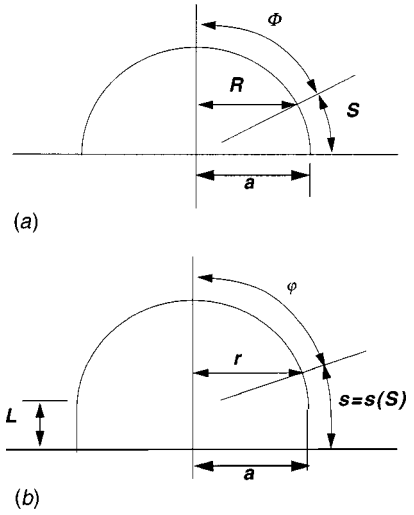


Fig. 8 (*a*) Initial hemispherical shell of radius *a*. The angle Φ is measured from the apex and the arc length *S* is measured from the equator. (*b*) Current configuration consisting of hemispherical shell with a cylindrical extension added.

$$f = \frac{\lambda_s}{\lambda_\theta}. \tag{34c}$$

Then Eq. (34*a*) and Eq. (34*b*) give the relation

$$\int_0^s \frac{ds}{f(s)r(s)} = \int_0^S \frac{dS}{R(S)}. \tag{35}$$

Since the radius as a function of arclength $r=r(s)$ completely defines a surface of revolution, Eq. (35) yields a unique solution for $s(S)$. Thus for a given ratio of stretches, there is a unique mapping from one surface of revolution to another.

If the ratio of stretches has the constant value f_c in the cylindrical portion of the mapping in Fig. 8(*b*), and instead of arclength, the angle from the apex is used, Eq. (35) becomes

$$\frac{L}{af_c} + \int_\varphi^{\pi/2} \frac{r_1 d\varphi}{fr} = \int_\Phi^{\pi/2} \frac{R_1 d\Phi}{R} \tag{36}$$

in which the subscript 1 denotes the meridional radius of curvature. Since $\varphi=\varphi(\Phi,L)$, the derivative of Eq. (36) with respect to L gives

$$\frac{\partial\varphi}{\partial L} = \frac{rf}{r_1 f_c a} \tag{37}$$

and the derivative of Eq. (34*a*) gives the result

$$\frac{\dot\lambda_\theta}{\lambda_\theta} = \frac{\dot L}{a}\frac{f}{f_c}\cos\varphi \tag{38}$$

$$\frac{\dot\lambda_s}{\lambda_s} = \frac{\dot L}{a}\frac{f}{f_c}\left(\cos\varphi + \frac{r}{f}\frac{df}{ds}\right) \tag{39}$$

where the dots denote the time derivative. So for any shape of the cap in steady-state growth, there is a simple dependence of the local grow rate on the angle from the apex.

Measurements of the tip growth by Chen [25] show a distribution that is roughly isotropic and approximated by the cosine variation. In the formulation of Hejnowicz et al. [26], the growth is in terms of the radius as a function of the axial length, for which the simple relations Eqs. (38) and (39) are not apparent.

4.1 Hemisphere. For the shape of the hemisphere in Fig. 8(*a*) and Fig. 8(*b*), the current angle as a function of arclength is

$$\varphi = \begin{cases} \dfrac{\pi}{2} & \text{for } 0 \leqslant s \leqslant L \\ \dfrac{\pi}{2} - \dfrac{s-L}{a} & \text{for } L < s \leqslant L + a\dfrac{\pi}{2} \end{cases} \tag{40}$$

and the radius is

$$r = a\sin\varphi. \tag{41}$$

Thus Eq. (35) yields the solution for the current angle in terms of the initial:

$$\varphi = \begin{cases} \dfrac{\pi}{2} & \text{for } 0 < \Phi \leqslant \Phi_L \\ 2\,\text{acot}\dfrac{\cot\Phi/2}{\cot\Phi_L/2} & \text{for } \Phi_L < \Phi \leqslant \dfrac{\pi}{2} \end{cases} \tag{42}$$

where Φ_L is the value of the initial angle that maps to the equator of the current configuration:

$$\Phi_L = 2\,\text{acot}\left(\exp\frac{L}{a}\right). \tag{43}$$

So for large L, the transition angle Φ_L becomes small, and most of the initial hemisphere has mapped into the cylindrical portion of the current configuration.

The results can be expanded for small values of L/a, which yields the results

$$\varphi \approx \Phi + \frac{L}{a}\sin\Phi \tag{44a}$$

$$\sin\varphi \approx \sin\Phi\left(1+\frac{L}{a}\cos\Phi\right) \tag{44b}$$

$$\lambda \approx 1+\frac{L}{a}\cos\Phi. \tag{44c}$$

So the rate of growth for increasing length L has a simple cosine distribution:

$$\dot{\lambda} \approx \frac{\dot{L}}{a}\cos\Phi, \tag{45}$$

which agrees with the general result Eq. (38)

Due to the turgor pressure of 7–10 atmospheres, there is substantial membrane stress in the surface. The stress is constant σ_0 in the hemisphere, and the circumferential stress in the cylinder is twice as much. A simple mechanical model would have material with the same properties at every point of the wall. The growth rate due to the addition of material to the wall from the cytoplasm would be a function of the local stress. So, the cosine distribution Eq. (45) in a region of constant stress seems to be inconsistent with such a simple model. However, the matter deserves further study.

4.2 Perturbed Shapes. Small changes in the initial and current shapes can be considered in the form

$$r(\varphi) = (a\sin\varphi)[1+\eta(\varphi)] \tag{46a}$$

$$R(\Phi) = (a\sin\Phi)[1+\aleph(\Phi)]. \tag{46b}$$

Equation (35) gives the perturbation result for the angle

$$\varphi \approx \Phi + \beta \tag{47}$$

in which

$$\beta = \left[\frac{L}{a}+\int_{\Phi}^{\pi/2}\frac{1}{\cos\Phi}\frac{d}{d\Phi}(\eta-\aleph)d\Phi\right]\sin\Phi$$

and the stretch ratio

$$\lambda \approx 1+\left[\frac{L}{a}-\int_{\Phi}^{\pi/2}\frac{\sin\Phi}{(\cos\Phi)^2}(\eta-\aleph)d\Phi\right]\cos\Phi. \tag{48}$$

This also agrees with the general result Eq. (38) when the current shape is the same as the initial $\eta=\aleph$.

For the stress, additional geometry is needed. The circumferential and meridional radii of curvature for the perturbed shape Eq. (46*a*) are

$$r_2 = \frac{r}{\sin\varphi} \approx a[1+\eta(\varphi)] \tag{49a}$$

$$r_1 = \frac{1}{\cos\varphi}\frac{dr}{d\sigma} \approx a\left[1+\eta(\varphi)+\tan\varphi\frac{d\eta}{d\varphi}\right]. \tag{49b}$$

Therefore the current membrane stress resultants in the circumferential and meridional directions are

$$N_\theta = \frac{pr_2}{2}\left(2-\frac{r_2}{r_1}\right) \approx \frac{pa}{2}\left[1+\eta+\tan\varphi\frac{d\eta}{d\varphi}\right] \tag{50a}$$

$$N_s = \frac{pr_2}{2} \approx \frac{pa}{2}[1+\eta] \tag{50b}$$

in which p is the turgor pressure. The resultant for the unperturbed hemisphere is

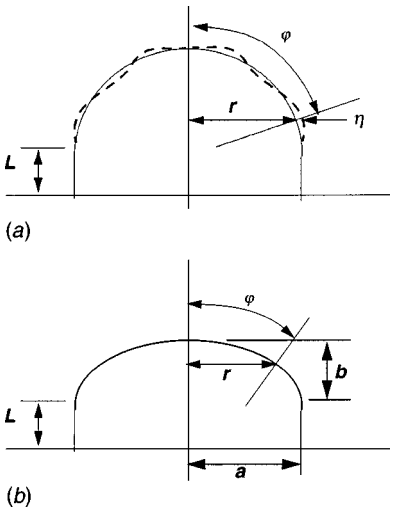


Fig. 9 (*a*) Perturbation of hemispherical shell by the radial displacement of magnitude η. (*b*) Elliptical end cap with semi-major and minor axes *a* and *b*.

$$N_0 = \frac{pa}{2}. \tag{50c}$$

Thus the perturbation of the hemispherical shape causes a non isotropic stress state. So, consistent with an isotropic growth, the mean stress is used:

$$\overline{N} = \frac{N_s+N_\theta}{2} \approx N_0\left[1+\eta+\frac{1}{2}\tan\varphi\frac{d\eta}{d\varphi}\right]. \tag{50d}$$

Thus the shear stress is assumed to not effect the growth. (See Fig. 9.)

4.3 Rate Equation. A linear relation between growth rate and mean stress is

$$\dot{\lambda} = \dot{\lambda}_0\left[1+\gamma\left(\frac{\overline{N}}{N_0}-1\right)\right] \tag{51}$$

where $\dot{\lambda}_0$ and γ are constants as indicated in Fig. 10. For an interpretation, the growth is considered to be resisted by a friction term and a viscosity, inversely proportional to γ. For high values of the mean stress, saturation occurs. The time derivative of Eq. (48) and the stress from Eq. (50*d*) substituted into Eq. (51) yields the linear differential equation for the perturbation shape function η:

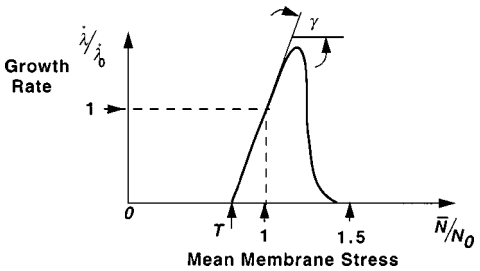


Fig. 10 Growth rate as a function of local mean stress for a simple, purely mechanical model. Growth occurs between a threshold value *T* and a saturation value of 1.5. For the spherical and elliptic caps, the stress is near N_0, and the slope of the curve has the positive value γ. For the cylindrical region, the effective stress is higher by a factor of 1.5, and growth does not occur. The value *T* could represent an initial frictional resistance, while γ is the inverse of the effective viscosity.

$$\left[\frac{\dot{L}}{a}-\int_{\varphi}^{\pi/2}\frac{\sin\varphi}{(\cos\varphi)^2}\,\dot{\eta}\,d\varphi\right]\cos\varphi=\dot{\lambda}_0\left[1+\gamma\left(\eta+\frac{1}{2}\tan\varphi\,\frac{d\eta}{d\varphi}\right)\right]. \tag{52}$$

4.4 Steady-State Growth. If the time variation in η is deleted from Eq. (52), then the equation for steady-state growth is obtained:

$$\frac{\dot{L}}{a}\cos\varphi=\dot{\lambda}_0\left[1+\gamma\left(\eta+\frac{1}{2}\tan\varphi\,\frac{d\eta}{d\varphi}\right)\right] \tag{53}$$

which has the exact solution

$$\eta=\frac{1}{\gamma}\left[\frac{2\dot{L}}{3\dot{\lambda}_0 a}g(\varphi)-1\right] \tag{54}$$

where the function g is

$$g(\varphi)=\frac{1+\cos\varphi+(\cos\varphi)^2}{1+\cos\varphi} \tag{55a}$$

which can be expanded in terms of cosines of even multiples of φ

$$g(\varphi)=\frac{4}{\pi}+4\left(\frac{10}{3\pi}-1\right)\cos 2\varphi-8\left(1-\frac{47}{15\pi}\right)\cos 4\varphi+\ldots \tag{55b}$$

$$=1.2732+0.2441\cos 2\varphi-0.02103\cos 4\varphi+0.00483\cos 6\varphi+\ldots \tag{55c}$$

Thus the function is reasonably approximated by the first two terms. After dropping the others, then

$$g(\varphi)\approx 4\left(1-\frac{7}{3\pi}\right)+8\left(\frac{10}{3\pi}-1\right)(\cos\varphi)^2. \tag{55d}$$

Now the rate can be chosen to make the constant term in the solution Eq. (53) equal to zero:

$$\dot{\lambda}_0=\frac{8}{3}\left(1-\frac{7}{3\pi}\right)\frac{\dot{L}}{a}=0.686\,\frac{\dot{L}}{R_0}, \tag{56}$$

and the only term remaining in the solution is

$$\eta=\frac{1}{\gamma}\,\frac{\dot{L}}{\dot{\lambda}_0 a}\,\frac{16}{3}\left(\frac{10}{3\pi}-1\right)(\cos\varphi)^2=\frac{k}{2}(\cos\varphi)^2 \tag{57}$$

which gives exactly the perturbation from a spherical to an elliptical surface Fig. 9(*a*), for which the parameter k is related to the major and minor semi-axes

$$k=\frac{a^2}{b^2}-1\approx 2\,\frac{(a-b)}{a}. \tag{58}$$

Using Eq. (56) in Eq. (57), this gives the product of the parameters:

$$k\gamma=4\,\frac{\left(\frac{10}{3\pi}-1\right)}{\left(1-\frac{7}{3\pi}\right)}=0.949. \tag{59}$$

From these results the conclusion is that a steady-state growth is possible for an elliptic perturbation from the hemisphere. Since there is no growth in the cylindrical region, the relation between growth rate and stress must have the form indicated in Fig. 10. The form Eq. (51) is the region near N_0 with the positive slope γ. The amplitude of the ellipticity k is inversely proportional to γ, Eq. (59). Thus for high values of γ, the deviation from the hemispherical shape is small. The steep slope means that the growth rate can vary considerably at different points with little change in the stress, which justifies the hemispherical result Eq. (45).

The shape of the end cap is independent of any vertical shift in the curve, since this is done with a change in $\dot{\lambda}_0$, which only affects the elongation rate $\dot{L}$ as given by Eq. (56). In fact, $\dot{\lambda}_0$ can be an arbitrary function of time and the shape given by η remains the same. To change the shape, the slope γ must be changed. So a change from a nearly spherical cap to an elliptic cap can be accomplished by simply decreasing the value of γ. The actual tip in Fig. 1(*a*) has $a/b=1.2$, which is a little large for quantitative accuracy of the perturbation analysis. However, this shape gives $k=0.49$ and from Eq. (59) the value $\gamma=1.9$. The flattened shape in Fig. 1(*b*) is beyond the perturbation analysis but would correspond to a small value of γ.

4.5 Stability of Solution. By dividing Eq. (52), by $\cos\varphi$ and taking the derivative, the equation can be written in the standard form

$$p\,\frac{d}{d\varphi}\left(\frac{1}{p}\,\frac{dy}{d\varphi}\right)-\frac{2}{\dot{\lambda}_0\gamma}\dot{y}=-\frac{2}{\gamma}(\sin\varphi)^2 \tag{60}$$

in which

$$p=\sin\varphi(\cos\varphi)^2$$
$$y=(\sin\varphi)^2\eta.$$

The nonhomogeneous term in Eq. (60) yields the steady-state solution Eq. (54). For p constant this is the same equation as for a string under tension attached to a viscous foundation. The variation in the coefficient does not affect the stability of the system. This can be most easily seen by a ''WKB'' asymptotic approximation for the homogeneous equation, valid for small values of γ, which is

$$y\approx p^{1/2}\exp\left(\pm i\,\frac{\varphi}{\sqrt{\dot{\lambda}_0\gamma\tau/2}}-\frac{t}{\tau}\right) \tag{61}$$

in which τ is an unknown decay time. This approximation is singular at $\varphi=0$. However, a uniformly valid solution can be obtained in terms of the Bessel function

$$\eta\approx\frac{\varphi^{1/2}\cos\varphi}{(\sin\varphi)^{3/2}}J_1\left(\frac{\varphi}{\sqrt{\dot{\lambda}_0\gamma\tau/2}}\right)\exp\left(-\frac{t}{\tau}\right). \tag{62}$$

With Eq. (62), the approximation for the spectrum of eigenvalues of the decay time is

$$\tau_n\approx\frac{1}{2n^2\dot{\lambda}_0\gamma}\quad\text{for}\quad n=3,4,\ldots. \tag{63}$$

Each mode has a positive decay time. Since the set of modes is complete, any perturbation will have a positive decay time and the system is stable for positive γ.

Measurements of the tip of the root hair by Chen [25] indicate a shape that is elongated, i.e., with $k<0$. The present model even with various extensions fails to yield a stable growth with $k<0$, unless a variation along the meridian in the material properties is prescribed. So this is an unresolved feature.

4.6 Change in Turgor Pressure. An approximate solution can be obtained for a small change in the internal pressure. The solution is taken in the form Eq. (57) with k as a function of time. Then Eq. (52) splits into terms that are constant in φ and those varying with $\cos 2\varphi$, giving the coupled equations

$$\frac{\dot{L}}{a}\,\frac{2}{\pi}-\frac{\dot{k}}{4}=\dot{\lambda}_0\left[1+\gamma\left(\frac{p}{p_0}-1\right)\right] \tag{64}$$

$$\frac{\dot{L}}{a}\frac{4}{3\pi}-\frac{\dot{k}}{4}=\dot{\lambda}_0\gamma\frac{k}{2} \tag{65}$$

in which p is the new value of internal pressure. Eliminating the axial length change L gives the equation for k:

$$\tau\dot{k}+k=\frac{4}{3\gamma}\left[1+\gamma\left(\frac{p}{p_0}-1\right)\right] \tag{66}$$

in which the decay time is $\tau=1/(6\dot{\lambda}_0\gamma)$. The solution is

$$k=\frac{4}{3\gamma}\left[1+\gamma\left(\frac{p}{p_0}-1\right)(1-e^{-t/\tau})\right]. \tag{67}$$

This indicates that the ellipticity increases with an increase in the internal pressure. Interesting is the axial growth rate $\dot{L}$:

$$\frac{\dot{L}}{a}\frac{2}{\pi}=\dot{\lambda}_0\left[1+\gamma\left(\frac{p}{p_0}-1\right)(1+2e^{-t/\tau})\right] \tag{68}$$

which has a jump due to the step in pressure. This is exactly the behavior reported in Green [27] in a different type of cell that grows uniformly along its length, rather than at the tip. It is interesting that the present analysis of the tip growth with the locally linear relation between growth rate and stress in Fig. 10 produces similar results. Indeed a good fit of the results in Green [27] can be obtained with Eq. (68) with reasonable values of γ about 10 and τ about 30 minutes.

4.7 Change in Material Property. A similar analysis can be carried out for the response of the cell to a sudden change in the parameter γ without a change in turgor pressure. The results for the ellipticity and the axial growth rate are

$$k=\frac{4}{3\gamma}\left[1+\left(\frac{\gamma}{\gamma_0}-1\right)e^{-t/\tau}\right] \tag{69}$$

$$\frac{\dot{L}}{a}\frac{2}{\pi}=\dot{\lambda}_0\left[1-2\left(\frac{\gamma}{\gamma_0}-1\right)e^{-t/\tau}\right]. \tag{70}$$

Thus, as before, the ellipticity changes continuously with the decay time τ, and the axial growth rate has a jump. For a decrease in γ from the previous value γ_0, the axial rate has a negative jump and then resumes the old value.

5 Conclusion

Pattern formation in plants is an exceedingly complex subject. However, it appears that the stability of the surface may play a fundamental role. With the high turgor pressure, compression occurs in the formative zone due to the geometry, as in *Acetabularia* Fig. 1(*b*) and the sunflower Fig. 2(*d*). In other situations, growth of the surface with peripheral constraint produces compression Fig. 6. The physical properties of geometry and elastic moduli are consistent with the production of elastic instability, as shown by Martynov [8] for *Acetabularia* and by Fig. 5 for a wide variety of plants.

The analysis of a constrained elastic ring subject to an increase in temperature, which simulates uniform growth of a plant in an annular region, shows that both in-plane (horizontal) and out-of-plane (vertical) modes of instability are possible. One parameter controls the choice for the preferred mode. This parameter is the ratio of stiffnesses of the elastic foundation and the ring wall. For the plant, the elastic foundation simulates the corpus and the wall simulates the tunica. For the relatively modest change in this ratio of a factor 2–4, the mode changes from the out-of-plane to the in-plane. Thus a prominent shape alternative can be accounted for by changing the value of a single material property, which could be genetically determined.

The change in *Acetabularia* from an efficient pressure vessel shape Fig. 1(*a*) to a less efficient shape Fig. 1(*b*), which has higher stress and a compressive region, is at first hard to understand. However, an analysis of the growth process, with a locally linear relation between growth rate and stress, shows that the ellipticity of the tip region depends on a single mechanical property. A dynamic stability analysis indicates that the steady-state growth is stable. The equations couple the tip shape and the axial growth rate. With this coupling, a step decrease in internal pressure causes a ''shock,'' consisting of a step decrease in the axial rate that recovers with time. This is similar to the measurements by Green [27] on the diffuse growth of the internodal cell of the green alga *Nitella*. We predict that similar behavior will occur with tip growth.

There are many things falling into place which reinforce the conjecture that the physical stability of the surface plays a key role in plant pattern formation. It is not just a matter of providing a set of equations that can produce a pattern. As argued by Green [24], there is a close interaction between the mechanics and gene expression to produce the stunning variety of plant patterns. By using the mechanics, the task for the genes is simplified in that only the physical properties and boundary conditions need be set up; the pattern comes out as a consequence. The buckling modifies the stress field which can trigger the expression of genes that are important for the subsequent stages of growth and pattern formation.

Acknowledgment

The great loss to this line of research was the untimely death of Paul B. Green in 1998. He possessed such enthusiasm that was infectious for colleagues and students and insight that overcame the lack of a deep understanding of mechanics. Hope for continuation rests heavily on his student Jacques Dumais, who is a biologist with the ability and interest in mechanics and no aversion to hard work. He has motivated the present effort. I also appreciate the discussions and contributions of biologists Luis Hernández and, recently, Dorota Kwiatkowska, as well as mechanicians Cheng-Hsiu Wu, Steve Rennich, Jonathan Fay, and Arif Karabeyoglu. Finally, I acknowledge the observation made by the late Richard Skalak to Paul Green that the plant patterns looked like plate buckling, that initiated this work. I suspect that Werner T. Koiter would have made a similar suggestion.

References

[1] Jean, R. V., 1994, *Phyllotaxis: A Systemic Study in Plant Morphogenesis*, Cambridge University Press, Cambridge, UK.
[2] Jean, R. V., and Barabé, D., eds., 1998, *Symmetry in Plants*, World Scientific, Singapore.
[3] Turing, A. M., 1952, ''The Chemical Basis of Morphogenesis,'' Philos. Trans. R. Soc. London, Ser. B, **B237**, pp. 37–72.
[4] Harrison, L. G., 1993, *Kinetic Theory of Living Pattern*, Cambridge University Press, Cambridge, UK.
[5] Schwendener, S., 1874, *Das mechanische Prinzip im anatomischen Bau der Monocotyledonen*., Engelmann, Liepzig.
[6] Karam, G. N., and Gibson, L. J., 1995, ''Elastic Buckling of Cylindrical Shells With Elastic Cores—I, Analysis,'' Int. J. Solids Struct., **32**, Nos. 8–9, pp. 1259–1283.
[7] Karam, G. N., and Gibson, L. J., 1995, ''Elastic Buckling of Cylindrical Shells With Elastic Cores—II, Experiments,'' Int. J. Solids Struct., **32**, Nos. 8–9, pp. 1285–1306.
[8] Martynov, L. A., 1975, ''A Morphogenetic Mechanism Involving Instability of Initial Form,'' J. Theor. Biol., **52**, pp. 471–480.
[9] Harrison, L. G., Snell, J., Verdi, R., Vogt, D. E., Zeiss, G. D., and Green, B. R., 1981, ''Hair Morphogenesis in *Acetabularia mediterranea*: Temperature-Dependent Spacing and Models of Morphogen Waves,'' Protoplasma, **105**, pp. 211–221.
[10] Harrison, L. G., and Hillier, N. A., 1985, ''Quantitative Control of *Acetabularia* Morphogenesis by Extracellular Calcium: A Test of Kinetic Theory,'' J. Theor. Biol., **114**, pp. 177–192.
[11] Dumais, J., and Harrison, L. G., 2000, ''Whorl Morphogenesis in the Dasycladalean Algae: The Pattern Formation Viewpoint,'' Philos. Trans. R. Soc. London, Ser. B, **355**, pp. 281–305.
[12] Hernández, L. F., 1991, ''Morphometry and Surface Growth Dynamics of the Sunflower (*Helianthus annuus* L.) Receptacle, Its Importance in the Determination of Yield,'' Suelo y Planta, **1**, pp. 91–103.
[13] Wu, C.-H., 1993, ''Fourier Series for Stability of Shallow Shells With One Large Element: Application to Plant Morphogenesis,'' Ph.D. thesis, Stanford University, Stanford, CA.

[14] Dumais, J., 1999, personal communication.
[15] Green, P. B., 1992, ''Pattern Formation in Shoots: A Likely Role for Minimal Energy Configurations of the Tunica,'' Int. J. Plant Sci., **153**, No. 3, pp. S59–S75.
[16] Szilard, R., 1974, *Theory and Analysis of Plates: Classical and Numerical Methods*, Prentice-Hall, Engelwood Cliffs, NJ.
[17] Green, P. B., 1996, ''Transductions to Generate Plant Form and Pattern: An Essay on Cause and Effect,'' Ann. Bot. (London), **78**, pp. 279–281.
[18] Green, P. B., Steele, C. R., and Rennich, S. C., 1998, ''How Plants Produce Pattern: A Review and a Proposal That Undulating Field Behavior is the Mechanism,'' *Symmetry in Plants*, Jean, R. V., and Barabé, D., eds., World Scientific, Singapore, pp. 359–392.
[19] Carpenter, R., Copsey, L., Vincent, C., Doyle, S., Magrath, R., and Coen, E., 1995, ''Control of Flower Development and Phyllotaxy by Meristem Identity Genes in *Antirrhinum*,'' Plant Cell, **7**, pp. 2001–2011.
[20] Kirchoff, B. K., and Rutishauser, R., 1990, ''The Phyllotaxy of *Costus* (Costaceae),'' Botanical Gaz, **151**, No. 1, pp. 88–105.
[21] Hernández, L. F., and Palmer, J. H., 1988, ''Regeneration of the Sunflower Capitulum After Cylindrical Wounding of the Receptacle,'' Am. J. Botany, **75**, No. 9, pp. 1253–1261.
[22] Kwiatkowska, D., and Florek-Marwitz, J., 1999, ''Ontogenetic Variation of Phyllotaxis and Apex Geometry in Vegetative Shoots of *Sedum Maximum* (L.) Hoffm.,'' Acta Soc. Botan. Polon., **68**, No. 2, pp. 85–95.
[23] Green, P. B., Steele, C. R., and Rennich, S. C., 1996, ''Phyllotactic Patterns: A Biophysical Mechanism for Their Origin,'' Ann. Bot. (London), **77**, pp. 515–527.
[24] Green, P. B., 1999, ''Expression of Pattern in Plants: Combining Molecular and Calculus-Based Biophysical Paradigms,'' Am. J. Botany, **86**, No. 8, pp. 1059–1076.
[25] Chen, J. C. W., 1973, ''The Kinetics of Tip Growth in the *Nitella* Rhizoid,'' Plant Cell Physiol., **14**, pp. 631–640.
[26] Hejnowicz, Z., Heinenann, B., and Sievers, A., 1971, ''Tip Growth: Patterns of Growth Rate and Stress in the *Chara* Rhizoids,'' Z. Pflanzenphysiol., **81**, pp. 409–424.
[27] Green, P., 1968, ''Growth Physics in *Nitella*: A Method for Continuous in vivo Analysis of Extensibility Based on a Micro-Manometer Technique for Turgor Pressure,'' Plant Physiol., **43**, No. 8, pp. 1169–1184.

Warner T. Koiter Medal Recipient

Conferral at the Applied Mechanics Dinner,
1999 International Mechanical Engineering Congress and Exposition

THE WARNER T. KOITER MEDAL was established in 1996 to recognize distinguished contributions to the field of solid mechanics with emphasis on the effective blending of theoretical and applied elements, and on a high degree of leadership in the international solid mechanics community.

The medal honors the late Dr. Warner T. Koiter (1914–1997), world-renowned authority in the field of solid mechanics, and it commemorates his vast contributions as research engineer and teacher.

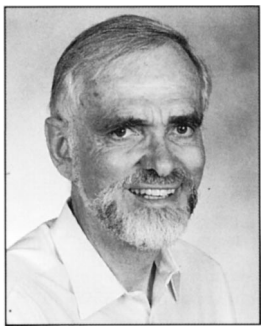
CHARLES R. STEELE, Ph.D., professor, department of mechanical engineering, Stanford University, California, *for leading research accomplishments in solid mechanics and pioneering analytical work in the biomechanics of the inner ear, which have forged a larger understanding of auditory response in humans*.

Dr. Steele, a professor of applied mechanics at Stanford University, is recognized worldwide for his contributions to the theory, practice and applications of applied mechanics. His efforts include shell theory, dynamics of the inner ear, mechanical properties of bone and the development of plant morphology.

Dr. Steele has made many contributions to the theory of shells. He is the master of shell asymptotics which accurately characterize the highly fluctuating nature of shell response in the vicinity of boundaries and intersections. In practice, these solutions invariably determine the areas of critical stress which are almost impossible to calculate accurately by traditional numerical approaches. Dr. Steele has synthesized his asymptotic solutions with finite element concepts to develop computational procedures for problems of shell intersections that are accurate, general and widely used in engineering practice. His insight into the behavior of shells provided the stepping stone for his work in modeling and explaining hearing mechanisms.

Among his seminal contributions, Dr. Steele developed the first truly three-dimensional model of the cochlea. His work on this and the basilar membrane are at the forefront of the field. He is also working on the complicated motions of the tympanic membrane, commonly known as the eardrum. Since there has been limited success during the past 20 years using finite element methods to model the eardrum, Dr. Steele is using analytical methods.

Dr. Steele has been involved in the development of a noninvasive technique for determining the mechanical characteristics of bone by measuring its vibrational response. This approach has been used in a range of applications, from examining the effects of osteoporosis on the strength and stiffness of bones to more fundamental investigations on the effects of geometry and bone mineral distribution on the structural integrity of bone.


Charles R. Steele

The morphogenesis of plants is another area Dr. Steele researched. The first strong evidence that the pattern formation in plants is governed by the elastic stability of the surface in the growth region was obtained in 1996.

Dr. Steele has published 88 scholarly articles in archive journals and has produced a host of subsidiary reports, articles and surveys. He is editor-in-chief of the *International Journal of Solids and Structures*.

A Fellow of ASME, Dr. Steele was on the Applied Mechanics Division Executive Committee (1979–1984) and served as chair (1983–1984). He also was chair of the Applied Mechanics Reviews Advisory Board (1985–1995) and member of the ASME Board on Publications (1989–1995).

Dr. Steele is a Fellow of the American Academy of Mechanics, which has promoted a series of meetings between North America and South America. He is a member of the American Institute of Aeronautics and Astronautics, the Acoustical Society of America, the American Society of Biomechanics, the American Academy of Engineering and the American Society for Gravitational and Space Biology.

His honors and awards include a Certificate of Recognition from the National Aeronautics and Space Administration (1987) for bone tissue analyzer and method; the Humboldt Senior Fellowship Award from Germany (1994); an honorary doctorate from Zaporozhye State University, Ukraine (1997); and Eminent Academician of the Ukrainian Academy of Higher Education (1998).

Dr. Steele received his bachelor's degree in mechanical engineering at Texas A&M, College Station, in 1956. He earned his doctorate in engineering mechanics at Stanford University in 1960.

A Parametric Model for a Class of Foam-Like Isotropic Hyperelastic Materials

S. Jemiolo
Institute of Structural Mechanics,
Warsaw University of Technology,
PL-00-632 Warsaw, Poland

S. Turteltaub[1]
Assoc. Mem. ASME,
Department of Mechanical Engineering
and Applied Mechanics,
University of Pennsylvania,
Philadelphia, PA 19104

A parametric model for foam-like materials is proposed and its correlation with experimental results is analyzed. The class of foam-like materials is assumed to be described by an isotropic elastic potential based on a general model proposed by Ogden. The class is parametrized using the relative mass density of the material. Functional relations between material parameters and the relative mass density are obtained from experimental data. A simple application problem, namely the optimization of a foam for maximum energy absorption under homogeneous compression, is formulated and solved numerically. [S0021-8936(00)02802-6]

1 Introduction

As pointed out by Ashby [1], structural materials commonly encountered in nature, such as wood or bone, belong to the class of cellular solids, i.e., materials that at a relatively small length scale are composed of cells whose geometry consists of ''walls'' made out of a ''dense'' solid material and the rest is ''void'' (filled with some fluid). Ashby also highlights the connection between this type of geometry and its optimality in terms of strength and weight. Other examples of cellular solids include synthetic materials such as open-cell elastomeric foams, which are commonly used in a variety of practical applications. In particular, foams are often used in energy-absorbing devices and their range of operation often includes large deformations (see, e.g., Maiti et al. [2]). Roughly speaking, models for foams can be divided into two large groups: models based on the foam's microstructure and phenomenological models. A model based on the foam's microstructure that has received wide attention is the one developed by Gibson, Ashby, and co-workers (see Gibson and Ashby [3]). The Gibson-Ashby model comprises two distinct functional relations for the range of deformations (linearly elastic and nonlinearly elastic—the latter referred to as the plateau/densification region). This is essentially a one-dimensional model (for uniaxial compression), however, it is not clear how to extend it to a three-dimensional setting for the nonlinear range. Also, from a computational point of view, it would be difficult to keep track of each region in a three-dimensional deformation.

Commonly used phenomenological models in finite elasticity are based on the stored energy potential proposed by Blatz and Ko [4] or, more generally, in expressions for compressible hyperelastic materials developed by Ogden [5]. However, neither the Blatz-Ko nor the Ogden potentials include any explicit information related to the microstructure of the foam. Nonetheless, this information, as shown by Maiti et al. [2], can be useful for design purposes. The aim of the present work is to introduce a characteristic parameter of the microstructure into the phenomenological constitutive information. Experimental results show that the relative mass density of an *isotropic* elastic foam plays an essential role in the response function. This parameter is perhaps the simplest characteristic feature of the microstructure and can be used to describe a whole class of materials with a single functional relation. The model proposed here is essentially an extension of Ogden's potential and is applicable for large three-dimensional deformations. It comprises a single functional relation that can describe the behavior of foam-like materials. One of the motivations for the present model is to analyze design problems where one anticipates large three-dimensional deformations and where the objective is to determine an optimal distribution of relative mass density throughout an energy-absorbing device. Since the model includes the relative mass density r as a parameter, then r can be naturally used as an optimization variable.

The outline of the paper is as follows: In Section 2 some basic notation and definitions are introduced; in Section 3, a specific class of stored energy functions—used to describe the behavior of elastic foams—is considered; the asymptotic behavior of the model for small deformations is analyzed in Section 4. Uniaxial homogeneous deformations are included in Section 5 and the validity of the model is studied in Section 6 where a parametrization of the model with respect to the relative mass density is derived based on experimental data. A simple application problem—the optimization of an energy-absorbing foam—is analyzed in Section 7. Some final remarks and conclusions follow in the last section.

2 Preliminaries

An essential parameter that describes the microstructure of a foam is its relative mass density that is defined by $r(\mathbf{X}) = \rho_0(\mathbf{X})/\rho_s$, where $\rho_0(\mathbf{X})$ is the mass density of the foam at a point $\mathbf{X}$ in an undeformed stress-free reference configuration and ρ_s is the density of the solid material from which the foam is made. The parameter r, $0<r\leq 1$, describes at a continuum level the (average) volume fraction occupied by the dense solid material in a representative unit cell at a microscopic level. Now, consider a (possibly) nonhomogeneous body that occupies a region Ω_0 in its reference configuration and let $\boldsymbol{\chi}(\mathbf{X})$ be a smooth deformation. The deformation gradient is defined as $\mathbf{F}=\mathbf{F}(\mathbf{X})\equiv \mathrm{Grad}\boldsymbol{\chi}(\mathbf{X})$ and it is assumed that the Jacobian of the deformation is positive, i.e., $J=J(\mathbf{X})\equiv \det\mathbf{F}(\mathbf{X})>0$, $\forall \mathbf{X}\in\Omega_0$. The polar decomposition of the deformation gradient is $\mathbf{F}=\mathbf{RU}=\mathbf{VR}$, where $\mathbf{R}$ is the rotation tensor ($\det\mathbf{R}=1$) and $\mathbf{U}$, $\mathbf{V}$ are, respectively, the right and left stretch tensors. The class of foams considered here is assumed to be *isotropic* and represented by a stored energy density that depends on the stretch tensor and the relative mass density, i.e., $W=\hat{W}(\mathbf{U};r)$, with $\hat{W}(\mathbf{I};r)=0$. Observe that the nonhomogeneity of the body enters this formulation only implicitly via the relative mass density. From the requirement of objectivity (and from the assumption of isotropy), the stored energy density must satisfy $\hat{W}(\mathbf{U};r)=\hat{W}(\mathbf{QUQ}^T;r)$, $\forall \mathbf{Q}\in O(3)$,

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 30, 1998; final revision, Oct. 30, 1999. Associate Technical Editor: J. T. Jenkins. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

where $O(3)$ is the group of orthogonal tensors. As a consequence of this, $\hat{W}$ depends on $\mathbf{U}$ only via its isotropic invariants. Hence, the arguments of the stored energy function are: the relative mass density r and any set of three independent invariants of $\mathbf{U}$ (or, equivalently, of $\mathbf{V}$). In particular, a useful functional expression for W is provided by $W=\overline{W}(\lambda_1,\lambda_2,\lambda_3;r)$, where $\overline{W}$ is a symmetric function of the principal stretches. In terms of the stored energy function $\overline{W}$, the Cauchy stress tensor is given by

$$\boldsymbol{\sigma}=\frac{1}{J}\sum_{k=1}^{K}\lambda_k\frac{\partial\overline{W}}{\partial\lambda_k}\mathbf{R}\mathbf{P}_k\mathbf{R}^T, \tag{1}$$

where λ_k, $(k=1,\ldots,K$, and $K\leqslant 3)$, are the *distinct* eigenvalues of the stretch tensors (i.e., $\mathbf{U}=\Sigma_{k=1}^{K}\lambda_k\mathbf{P}_k$, $\mathbf{V}=\Sigma_{k=1}^{K}\lambda_k\mathbf{R}\mathbf{P}_k\mathbf{R}^T$, and the corresponding $\mathbf{P}_k$ are projectors[2] of $\mathbf{U}$). The first Piola-Kirchhoff stress tensor is given by $\mathbf{S}=J\boldsymbol{\sigma}\mathbf{F}^{-T}$.

3 Specific Constitutive Relationships for Foams

The form of $\overline{W}$ assumed here is based on a general expression for the stored energy function of compressible materials proposed by Ogden [5], i.e.,

$$\begin{aligned}\overline{W}(\lambda_1,\lambda_2,\lambda_3;r)=\sum_{i=1}^{N}\{&2m_i(\lambda_1^{\alpha_i}+\lambda_2^{\alpha_i}+\lambda_3^{\alpha_i}-3)\\&+n_i[(\lambda_1\lambda_2)^{\beta_i}+(\lambda_2\lambda_3)^{\beta_i}+(\lambda_3\lambda_1)^{\beta_i}-3]\\&+p_i(J^{-\gamma_i}-1)\},\end{aligned} \tag{2}$$

where $J=\lambda_1\lambda_2\lambda_3$ and

$$p_i=\frac{2(m_i\alpha_i+n_i\beta_i)}{\gamma_i},\quad \forall i\in[1,N]. \tag{3}$$

In (2), $N\geqslant 1$ is a number chosen depending on the required accuracy of the model. The form of the parameters p_i follow from the requirement of a stress-free undeformed (reference) configuration. In order to introduce a functional dependence on the relative mass density, it is assumed that the material parameters α_i, β_i, γ_i, m_i, and n_i are functions of r, i.e., $\alpha_i=\hat{\alpha}_i(r)$, $\beta_i=\hat{\beta}_i(r)$, $\gamma_i=\hat{\gamma}_i(r)$, $m_i=\hat{m}_i(r)$, $n_i=\hat{n}_i(r)$. It is worth mentioning that even though the above functions are referred to as material parameters, the relevant mechanical behavior of the material described by (2) is characterized by the derivative of the stored energy (i.e., the functions $\hat{\alpha}_i$, etc., do not necessarily have a physical meaning per se). The derivative of $\overline{W}$ with respect to the principal stretches is

$$\begin{aligned}\frac{\partial\overline{W}}{\partial\lambda_k}=\sum_{i=1}^{N}\Bigg\{&\frac{1}{\lambda_k}[2m_i\alpha_i\lambda_k^{\alpha_i}+n_i\beta_i\lambda_k^{\beta_i}(\lambda_l^{\beta_i}+\lambda_m^{\beta_i})\\&-2(m_i\alpha_i+n_i\beta_i)J^{-\gamma_i}]\Bigg\}\quad l,m\neq k,\end{aligned} \tag{4}$$

where k, l, m range in {1,2,3}. The stored energy (2) can be expressed alternatively as

$$\begin{aligned}\hat{W}(\mathbf{U};r)=\sum_{i=1}^{N}[&2m_i(\operatorname{tr}\mathbf{U}^{\alpha_i}-3)+n_i(\operatorname{tr}\operatorname{Cof}\mathbf{U}^{\beta_i}-3)\\&+p_i(J^{-\gamma_i}-1)],\end{aligned} \tag{5}$$

where, since $\mathbf{U}$ is symmetric positive definite, $\operatorname{Cof}\mathbf{U}^{\beta_i}=(\operatorname{Cof}\mathbf{U})^{\beta_i}=J^{\beta_i}\mathbf{U}^{-\beta_i}$.

Specific functions $\hat{\alpha}_i$, $\hat{\beta}_i$, $\hat{\gamma}_i$, $\hat{m}_i$, and $\hat{n}_i$ appropriate for foams are described in Sections 4 and 6. However, it is worth noting a few facts here: In order to guarantee the existence of solutions to boundary value problems, it is useful to consider some restrictions on the form of the stored energy (5). Specifically, it is assumed that the stored energy is polyconvex and also satisfies a growth condition (see, e.g., Ball [6] and Dacorogna [7]). Polyconvexity is assured if there exists a function W_p such that $W_p(\mathbf{U},\operatorname{Cof}\mathbf{U},J;r)=\hat{W}(\mathbf{U};r)$ and W_p is convex in $\mathbf{U}$, $\operatorname{Cof}\mathbf{U}$ and J for every r. Thus, for the potential (5), polyconvexity is satisfied if, $\forall i\in[1,N]$, one has $m_i\alpha_i(\alpha_i-1)\geqslant 0$, $n_i\beta_i(\beta_i-1)\geqslant 0$ and $p_i\gamma_i(\gamma_i+1)=2(m_i\alpha_i+n_i\beta_i)(\gamma_i+1)\geqslant 0$. Moreover, to guarantee existence of solutions, it is sufficient to impose a growth condition of the form $W_p(\mathbf{U},\operatorname{Cof}\mathbf{U},J;r)\geqslant c_1+c_2(\|\mathbf{U}\|^{q_1}+\|\operatorname{Cof}\mathbf{U}\|^{q_2})$, where $c_1\in\mathbb{R}$, $c_2>0$, $q_1\geqslant 2$, $q_2\geqslant q_1/(q_1-1)$ and $\|\mathbf{U}\|\equiv\sqrt{\mathbf{U}\cdot\mathbf{U}}=\sqrt{\operatorname{tr}\mathbf{U}^2}$ (see Ball and Murat [8]). In this case, the growth condition is satisfied if

$$\alpha_i\geqslant 2,\quad \beta_i\geqslant\frac{\alpha_i}{\alpha_i-1},\quad m_i>0,\quad n_i>0,\quad \forall i\in[1,N], \tag{6}$$

which implies polyconvexity if $\gamma_i\geqslant -1$. In (5), the term $\Sigma_{i=1}^{N}p_i(J^{-\gamma_i}-1)$—which measures the change of energy due to a change of volume—is required to tend to $+\infty$ as $J\to 0^+$. This condition is imposed so as to assign an infinite amount of energy in order to compress a body to a single point. Thus, it is assumed that

$$\exists j\in[1,N]:\gamma_j>0. \tag{7}$$

Observe that if the material is polyconvex, then $\hat{W}\to+\infty$ as $J=\lambda_1\lambda_2\lambda_3\to+\infty$ for *smooth* deformations, even though $\Sigma_{i=1}^{N}p_i(J^{-\gamma_i}-1)\to-\Sigma_{i=1}^{N}p_i$ when all γ_i are strictly positive. It is noted that the generalized Blatz-Ko material is a special case of (5) with $N=2$ and $m_1=1/4\mu_0 f$, $m_2=1/4\mu_0(1-f)$, $n_1=n_2=0$, $\alpha_1=2$, $\alpha_2=-2$, $\gamma_1=2\nu_0/(1-2\nu_0)$ and $\gamma_2=-\gamma_1$, where μ_0 and ν_0 are, respectively, the shear modulus and Poisson's ratio for small deformations and f is a parameter presumably related to the volume fraction of voids—although not equal to it (see Blatz and Ko [4]). Hence, for $f\neq 1$, the generalized Blatz-Ko potential is polyconvex if $1/4\leqslant\nu_0<1/2$ but does not satisfy the growth condition since $\alpha_2<0$. For the special case $f=1$, the Blatz-Ko potential is polyconvex for any value of $\nu_0\in(-1,1/2)$ and satisfies the growth condition (see Horgan [9] for a discussion on the loss of ellipticity of the Blatz-Ko potential). It is noted in passing that the special case $n_i=0$, $\gamma_i=\alpha_i\delta$, $\forall i\in[1,N]$ can be viewed as a generalization of the Blatz-Ko potential. In this case, the stored energy becomes

$$\hat{W}_0(\mathbf{U};r)=\sum_{i=1}^{N}2m_i\left[(\operatorname{tr}\mathbf{U}^{\alpha_i}-3)+\frac{1}{\delta}(J^{-\alpha_i\delta}-1)\right].$$

The limit case $\delta\to 0$ for this special potential can be useful for materials that, for example, have a negligible lateral displacement for uniaxial compression.

4 Asymptotic Analysis for Small Deformations

Experimental results for foam-like materials in the range of small deformations are relatively well characterized. Hence, it is convenient to describe the behavior of a material given by (5) under small deformations since it must match the experimental results. To this end, consider the expansion of the stored energy in terms of the Lagrange strain tensor $\mathbf{E}=1/2(\mathbf{C}-\mathbf{I})$. From (5), the stored energy is given by $W(\mathbf{E};r)=1/2\lambda_0(\operatorname{tr}\mathbf{E})^2+\mu_0\operatorname{tr}\mathbf{E}^2+O(\|\mathbf{E}\|^3)$, where the Lamé moduli are given by

$$\lambda_0=\sum_{i=1}^{N}(n_i\beta_i^2+p_i\gamma_i^2),\quad 2\mu_0=\sum_{i=1}^{N}(2m_i\alpha_i^2+n_i\beta_i^2), \tag{8}$$

hence the corresponding bulk modulus κ_0 is, on using (3),

$$3\kappa_0=\sum_{i=1}^{N}2[m_i\alpha_i(\alpha_i+3\gamma_i)+n_i\beta_i(2\beta_i+3\gamma_i)]. \tag{9}$$

[2]The projectors satisfy the following relations: $\Sigma_{k=1}^{K}\mathbf{P}_k=\mathbf{I}$ (where $\mathbf{I}$ is the second-order identity tensor), $\mathbf{P}_k\cdot\mathbf{P}_l\equiv\operatorname{tr}(\mathbf{P}_k\mathbf{P}_l)=0$ for $k\neq l$ and $\operatorname{tr}\mathbf{P}_k=$multiplicity of λ_k.

The corresponding Saint-Venant–Kirchhoff approximation is obtained by neglecting higher order terms of $W(\mathbf{E};r)$. The second Piola-Kirchhoff stress tensor ($\mathbf{T}\equiv\mathbf{F}^{-1}\mathbf{S}$) for the Saint-Venant–Kirchhoff material can be expressed as $\mathbf{T}=\partial W/\partial\mathbf{E}=\lambda_0(\mathrm{tr}\,\mathbf{E})\mathbf{I}+2\mu_0\mathbf{E}$. Formally, the linearization of the Saint-Venant–Kirchhoff material (Hooke's law) can be obtained upon replacing the Lagrange strain by the infinitesimal strain in the previous expression for $\mathbf{T}$ (in which case all stress measures are equivalent). In order to guarantee that the material is stable under linear perturbations from the reference state it is required that $3\kappa_0>0$, $2\mu_0>0$. These conditions naturally impose restrictions on α_i, β_i, γ_i, m_i, and n_i, however, it is noted that such requirements do not guarantee (nor it is assumed) stability for large deformations. It is also worth to point out that, for the special case $n_i=0$, $\gamma_i=\alpha_i\delta$, $\forall i\in[1,N]$ and since $\nu_0=\lambda_0/(2(\lambda_0+\mu_0))$, it follows from (8) that $\nu_0=\delta/(1+2\delta)$, which is *independent* of α_i and m_i.

Experimental results and models based on simple microgeometries suggest that for open-cell elastomeric foams in the *linearly elastic range*, Young's modulus E_0 and Poisson's ratio ν_0 depend on the relative mass density as follows: $E_0=\hat{E}_0(r)=E_s r^2$, $\nu_0=\hat{\nu}_0(r)=1/3$, where E_s is equal to the dense solid material's Young modulus (see Gibson and Ashby [3]). Observe that, as a first approximation, the Gibson-Ashby model assumes that Poisson's ratio does not depend on the relative mass density, which is equivalent to assume that both κ_0 and μ_0 depend on the *same* power of r. Since $2\mu_0=E_0/(1+\nu_0)$ and $3\kappa_0=E_0/(1-2\nu_0)$, the Gibson-Ashby model corresponds to

$$\hat{K}_1(r)\equiv 3\hat{\kappa}_0(r)=3E_s r^2,\quad \hat{K}_2(r)\equiv 2\hat{\mu}_0(r)=\tfrac{3}{4}E_s r^2, \tag{10}$$

where $K_1=\hat{K}_1(r)$ and $K_2=\hat{K}_2(r)$ are the so-called Kelvin moduli. It is also noted that, from a theoretical point of view, it would be more appropriate to find from two *independent* experimental tests the dependence with respect to the relative mass density of the Kelvin moduli.

The functions $\hat{m}_i$ and $\hat{n}_i$ can be suitably prescribed in order to *asymptotically* match the present model with these experimental results (whereas the functions $\hat{\alpha}_i$, $\hat{\beta}_i$, and $\hat{\gamma}_i$ can be used to describe the nonlinear range). In view of $(8)_2$, (9) and (10), the functions $\hat{m}_i$ and $\hat{n}_i$ can be chosen in various ways (in terms of α_i, β_i, γ_i, K_1, and K_2) in order to have an asymptotic agreement. However, for the special case $N=1$, the correspondence is uniquely determined as follows:

$$m=\frac{1}{6\alpha}\left[\frac{2(2\beta+3\gamma)K_2-\beta K_1}{\alpha\beta-\gamma(\beta-2\alpha)}\right],\qquad n=\frac{1}{3\beta}\left[\frac{\alpha K_1-(\alpha+3\gamma)K_2}{\alpha\beta-\gamma(\beta-2\alpha)}\right]. \tag{11}$$

For the case $N=1$, one can determine explicitly the ranges of the values α, β, and γ that satisfy *both* the growth conditions (6), (7) and linear stability about the undeformed configuration. (The growth conditions are sufficient in order to guarantee polyconvexity.) The admissible regions can be obtained from the requirement K_1, $K_2>0$ and Eqs. (6), (7), and (11). To determine these regions, it is convenient to study three cases: $K\in(0,1]$, $K\in(1,4]$, and $K\in(4,+\infty)$ where $K=K_1/K_2$. Note that the limits of the intervals for K correspond to, respectively, $\nu_0=-1$, $\nu_0=0$, $\nu_0=1/3$ and $\nu_0=1/2$. It is found that for the range $K\in(0,1]$ (i.e., $\nu_0\in(-1,0]$) it is not possible to satisfy the growth condition. For the other cases, the admissible regions are as follows: for $K\in(1,4]$: $\alpha\geq 2$, $\beta\geq\alpha/(\alpha-1)$ and $0<\gamma<\gamma_A$; for $K\in(4,+\infty)$: $\alpha\geq 2$, $\beta\geq\alpha/(\alpha-1)$ and $\gamma_A<\gamma<\gamma_B$ if $2\alpha/\beta<(K-4)/(K-1)$ or $\gamma_B<\gamma<\gamma_A$ if $2\alpha/\beta>(K-4)/(K-1)$, where $\gamma_A=(\alpha/3)(K-1)$ and $\gamma_B=(\beta/6)(K-4)$.

It is worth mentioning that one can use other models for open-cell elastomeric foams in the linearly elastic range instead of Gibson and Ashby's model (see, e.g., Warren and Kraynik [10] where the Kelvin moduli do not depend on the same power of r). If one wishes to use Warren and Kraynik's model, the asymptotic matching presented in this section should be modified accordingly. Furthermore, in view of the previous analysis, if negative Poisson's ratio values are considered in the model for small deformations (which has been observed in certain foams), one has to keep in mind that polyconvexity (and existence of solutions) is not necessarily guaranteed for $N=1$.

5 Uniaxial Deformations

In preparation for the analysis of experimental data it is useful to derive explicit formulas of the stress-stretch relations for simple homogeneous uniaxial deformations. Formulas for other modes of homogeneous deformations such as bi-axial deformation, dilatation, or simple shear can be derived in a similar way. However, the most commonly available experimental data correspond to simple compression or tension. Throughout this section it is assumed that the body is homogeneous (i.e., $r=r(\mathbf{X})=$ constant). In view of (1) and (4), the Cauchy stress tensor is given by

$$\boldsymbol{\sigma}=\frac{1}{J}\sum_{i=1}^{N}\{2m_i\alpha_i\mathbf{V}^{\alpha_i}+n_i\beta_i[(\mathrm{tr}\,\mathrm{Cof}\mathbf{V}^{\beta_i})\mathbf{I}-\mathrm{Cof}\mathbf{V}^{\beta_i}]-2(m_i\alpha_i+n_i\beta_i)J^{-\gamma_i}\mathbf{I}\}. \tag{12}$$

Let λ_1 be the stretch in the axial direction. In the case of simple compression or tension, $\lambda_2=\lambda_3=\lambda$ and $\mathbf{R}=\mathbf{I}$. A functional relation between λ and λ_1 can be established by requiring that the principal Cauchy stresses in directions perpendicular to the axial direction must vanish, i.e., in view of (12),

$$\sum_{i=1}^{N}[2m_i\alpha_i\lambda^{\alpha_i}+n_i\beta_i\lambda^{\beta_i}(\lambda_1^{\beta_i}+\lambda^{\beta_i})-2(m_i\alpha_i+n_i\beta_i)(\lambda_1\lambda^2)^{-\gamma_i}]=0. \tag{13}$$

The first Piola-Kirchhoff stress in the axial direction is

$$S_1=\sum_{i=1}^{N}\frac{2}{\lambda_1}[m_i\alpha_i\lambda_1^{\alpha_i}+n_i\beta_i\lambda_1^{\beta_i}\lambda^{\beta_i}-(m_i\alpha_i+n_i\beta_i)(\lambda_1\lambda^2)^{-\gamma_i}], \tag{14}$$

which is *formally* a function of λ_1 only since, in view of (13), the lateral stretch λ can be interpreted as an implicit function of λ_1. The above expression corresponds to the stress-stretch relation that should be used when comparing the present model with experimental data obtained from uniaxial compression tests (i.e., nominal stress versus axial stretch) under the assumption that the test corresponds to a *homogeneous* deformation. In general, Eq. (13) cannot be solved in closed form for λ as a function of λ_1. However, for the special case $n_i=0$ and $\gamma_i=\alpha_i\delta$, one can show that $\lambda=\lambda_2=\lambda_3=\lambda_1^{-\nu}$, hence $J=\lambda_1^{1-2\nu}$, where $\nu=\delta/(1+2\delta)$. Observe that

$$\nu=-\frac{\log\lambda}{\log\lambda_1},\quad (n_i=0), \tag{15}$$

which provides an interpretation of ν as a generalized Poisson's ratio for large deformations. The limit case $\delta\to 0$ (i.e., $\nu\to 0$) corresponds to an axial deformation with negligible lateral deformation. The principal value of the first Piola-Kirchhoff stress tensor in the axial direction is, when $n_i=0$,

$$S_1=\sum_{i=1}^{N}\frac{2m_i\alpha_i}{\lambda_1}[\lambda_1^{\alpha_i}-\lambda_1^{-\alpha_i\nu}],\quad (n_i=0). \tag{16}$$

From (15), it follows that the special potential $\hat{W}_0$ (i.e., a generalization of Blatz-Ko's potential) predicts a *linear* relation be-

tween $\log\lambda$ and $\log\lambda_1$ for uniaxial deformation. This could be a reasonable assumption in *compression* since experimental results for elastomeric foams reported by Maiti et al. [2] show that $\lambda\sim 1$, hence $\nu\sim 0$. However, experimental measurements for polyurethane foams in *tension* conducted by El-Ratal and Mallick [11] reveal that $\log\lambda$ and $\log\lambda_1$ do not depend linearly and that the lateral stretch deviates considerably from 1. This point is analyzed quantitatively in Section 6.

6 Applicability of the Model

In order to investigate the applicability of the present model, one has to determine numerically the material functions $\hat{\alpha}_i$, $\hat{\beta}_i$, and $\hat{\gamma}_i$ based on experimental data. This can be achieved upon applying a modified version of the method used by Twizell and Ogden [12] (essentially a least-squares error minimization using the so-called Levenberg-Marquardt algorithm). Ideally, one should consider data from independent modes of deformation and then minimize the difference between experimental and theoretical values simultaneously. However, the experimental data available for this work are limited to uniaxial compression (and tension) only.

In order to determine the *functions* $\hat{\alpha}_i(r)$, $\hat{\beta}_i(r)$ and $\hat{\gamma}_i(r)$ one can proceed as follows: experimental values for homogeneous compression are usually points of the form (λ_1, S_1) where λ_1 is the axial stretch and S_1 is the first Piola-Kirchhoff stress in the axial direction. Ideally, the value of the lateral stretch λ is also reported but this is sometimes not the case. Assuming the latter, the first step is to consider experimental data sets $\{(\lambda_1^{(k,j)}, S_1^{(k,j)})\}$ from a representative number D of relative densities (i.e., $r_1, \ldots, r_D$), where $1<k\leqslant M_j$, $1<j\leqslant D$ and M_j is the number of experimental points for each relative mass density r_j. Then, for each subset of experimental points corresponding to a *fixed* r_j, the corresponding values of $\alpha_i^{(j)}$, $\beta_i^{(j)}$, and $\gamma_i^{(j)}$ $(1\leqslant i\leqslant N)$ that minimize the error for each data subset are determined using the method described by Twizell and Ogden. Subsequently, functional forms for $\hat{\alpha}_i(r)$, $\hat{\beta}_i(r)$ and $\hat{\gamma}_i(r)$ are assumed (e.g., a polynomial or a power law) and a second least-squares minimization is carried out for the points $(r_j, \alpha_i^{(j)})$, etc. At the end of this process the model is fully determined in terms of its dependence on r. For simple homogeneous compression, in view of (13), in general there is no closed-form expression for the lateral stretch λ in terms of λ_1 (except, e.g., for the special case of Section 3). Hence, for the general case, the following preliminary steps need to be implemented before applying the method used by Twizell and Ogden (the index i ranges in $[1,N]$, the superscript j is dropped since the following applies for given r_j):

(a) Prescribe the Kelvin moduli $\hat{K}_1=3\hat{\kappa}_0$ and $\hat{K}_2=2\hat{\mu}_0$ as functions of r (or, equivalently $\hat{E}_0(r)$ and $\hat{\nu}_0(r)$) consistent with experimental data for small deformations. For example, use the functions (10).

(b) Obtain the functions $\hat{m}_i$ and $\hat{n}_i$ (use (11) for $N=1$).

(c) From the equilibrium Eq. (13) compute (symbolically) the following derivatives: $\partial\lambda/\partial\alpha_i$, $\partial\lambda/\partial\beta_i$, $\partial\lambda/\partial\gamma_i$. These derivatives are given as functions of α_i, β_i, γ_i, λ_1, λ and its functional form is relatively complex but can be easily obtained with a symbolic manipulator.

(d) Following Twizell and Ogden [12], define $E^{(k)}=\overline{S_1^{(k)}}-S_1^{(k)}$, where $\overline{S_1^{(k)}}$ is the experimental value of the nominal stress and $S_1^{(k)}$ is the stress given by (14) as a function of α_i, β_i, γ_i, λ_1 and λ (there is one $E^{(k)}$ per each experimental point k).

(e) Compute symbolically the matrix of derivatives $P_{ki}=\{\partial E^{(k)}/\partial\alpha_i, \partial E^{(k)}/\partial\beta_i, \partial E^{(k)}/\partial\gamma_i\}_{i=1}^{N}$ using the expressions for $\partial\lambda/\partial\alpha_i$, $\partial\lambda/\partial\beta_i$, $\partial\lambda/\partial\gamma_i$.

For a *fixed* relative mass density r_j, the unknown vector is $\mathbf{x}=\{(\alpha_i,\beta_i,\gamma_i)\}_{i=1}^{N}$. Since it is assumed that the experimental data do not include the lateral stretches $\lambda^{(k)}$, they need to be computed numerically for each $\lambda_1^{(k)}$ *and* each $\mathbf{x}$. The rest of the algorithm, briefly reviewed below for the sake of completeness, is similar to the one described by Twizell and Ogden [12] except that in order to update $\mathbf{x}$ the new values of $\lambda^{(k)}$ and the corresponding gradients P_{ki} and errors $E^{(k)}$ need to be recalculated every time $\mathbf{x}$ is modified.

1 Assume an initial value $\mathbf{x}_0=\{(\alpha_i^0,\beta_i^0,\gamma_i^0)\}_{i=1}^{N}$.
2 From (13), compute $\lambda^{(k)}$ for each $\lambda_1^{(k)}$ (for $1<k\leqslant M_j$).
3 From (14), compute $\overline{S_1^{(k)}}$ to obtain the vector of errors $\mathbf{E}=\{E^{(k)}\}_{k=1}^{M_j}$ and thus the matrix of gradients $\mathbf{P}$ using the expressions derived in steps (d) and (e) above.
4 Update $\mathbf{x}$ from iteration q to $q+1$ as follows:

$$\mathbf{x}_{q+1}=\mathbf{x}_q-[\mathbf{P}_q^T\mathbf{P}_q+\gamma_q\mathbf{I}]^{-1}\mathbf{P}_q^T\mathbf{E}_q,$$

where $\mathbf{I}$ is the identity matrix (of suitable size depending on the number of data and unknowns) and the scalar parameter γ_q is essential to avoid singularities in the algorithm (see Twizell and Ogden [12] for a detailed description of it). Steps 2 to 4 are repeated until the method converges.

On applying the above algorithm with $N=1$ to the experimental data reported by Maiti et al. [2] (see also Gibson and Ashby [3]), one can determine the parameters α, β, and γ for five relative mass densities of polyethylene (in the range $r\in[0.0245,0.3]$) and three densities of polyurethane ($r\in[0.012,0.043]$). Subsequently, the functions $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ can be obtained from a simple curve fitting. A parabolic function was chosen for polyethylene and a linear relation for polyurethane due to the limited number of relative mass densities. It is noted that polyethylene foams typically have *closed cells* and a nonelastic behavior upon unloading. However, Maiti et al. [2] propose a similar constitutive model for polyethylene foams than for polyurethane foams in compression (the present work makes a similar assumption, though clearly the model would not be valid if the material has a hysteretic behavior). Nonetheless, the effectiveness of the numerical procedure for results in compression is best illustrated with data for polyethylene foams since they cover a wider range of relative mass densities.

The nominal stress versus axial stretch curves obtained *after* correlating the functions $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are shown in Figs. 1 (polyethylene, $E_s=0.7$ GPa) and 2 (polyuretane, $E_s=0.045$ GPa). Figure 3 corresponds to the functions $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$. It is noted that these functions are in the admissible range defined in Section 4 (recall that the Gibson-Ashby model assumes that $\overline{K}=K_1/K_2=4$). It is worth noting that these curves should not be used for extrapolation, but rather for interpolation (as a function of r).

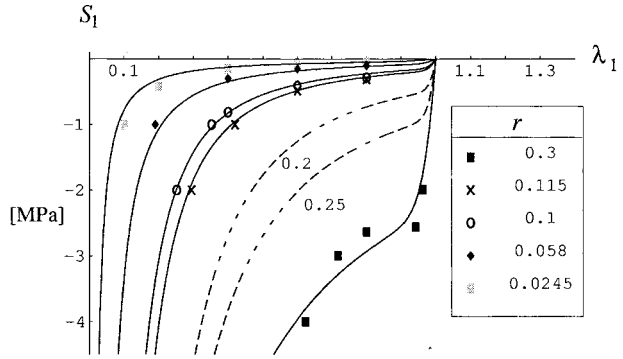


Fig. 1 Nominal axial stress versus axial stretch for simple compression (polyethylene). The experimental data are taken from Maiti et al. [2]; the solid lines correspond to the theoretical model. The dashed lines are obtained by interpolation.

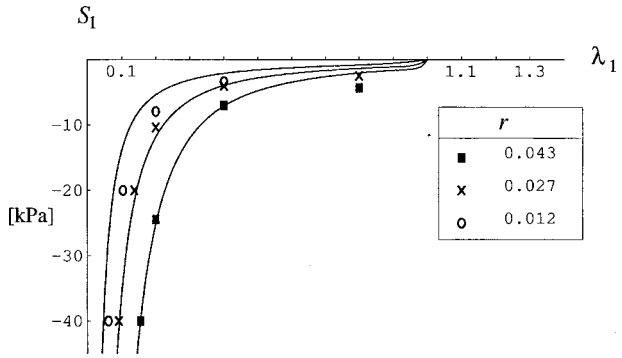


Fig. 2 Nominal axial stress versus axial stretch for simple compression (polyurethane). The experimental data are taken from Maiti et al. [2]; the solid lines correspond to the theoretical model.

There is a relatively good agreement with the experimental data using a one-term stored energy ($N=1$). More accurate curve fittings can be obtained by using more than one term. Observe that the functions α, β, and γ for polyethylene shown in Fig. 3 are nonmonotonic. However, this does not result in a nonunique model as illustrated by the two interpolated stress-stretch curves for $r=0.2$ and $r=0.25$ shown in Fig. 1 (dashed lines). It is also important to notice that these curves are likely to be different if more data are used for curve fittings (i.e., additional relative densities for compression or other types of deformation).

To further investigate the applicability of the model, experimental results reported by El-Ratal and Mallick [11] were analyzed. Their results consist of uniaxial *tension* tests for two different open-cell polyurethane foams with densities $\rho_1 = 0.017\,\mathrm{kg\cdot m^{-3}}$ and $\rho_2 = 0.035\,\mathrm{kg\cdot m^{-3}}$. The experimental data include both axial and lateral stretches. As opposed to the previous case (compression), a different approach is used here since there is no information about the foam's behavior in the linearly elastic range (i.e., Young's modulus E_s is not reported for the specific foam samples except for values derived from the numerical analysis of the nonlinear behavior). In this case, the parameters m and n are not specified using (11) but computed from the nonlinear data in a similar way as for α, β, and γ. Also, the derivatives $\partial\lambda/\partial\alpha_i$, etc., are not required since the experimental values of the lateral stretch λ are known. However, in addition to the errors $E=\overline{S}_1-S_1$ for axial stress, the errors $E=\overline{S}-S=-S$ for (zero) lateral stresses—i.e., Eq. (13)—need to be included in the minimization procedure to guarantee a good agreement between theoretical and experimental values. Using both (13) and (14), the parameters m, n, α, β, and γ are determined—independently for ρ_1 and ρ_2—in order to minimize the sum of the errors E. One can obtain a very good correlation with only one term ($N=1$); however, the resulting power γ turns out to be negative, hence the predicted behavior in compression is unrealistic. To overcome this difficulty, the approach taken here is as follows: Several experimental points from *compression* tests for the same material (polyurethane) were taken from the previous data set (Maiti et al. [2])—interpolated to the same relative densities as in El-Ratal and Mallick [11] tension tests using $\rho_s = 1200\,\mathrm{kg\cdot m^{-3}}$. Then, this

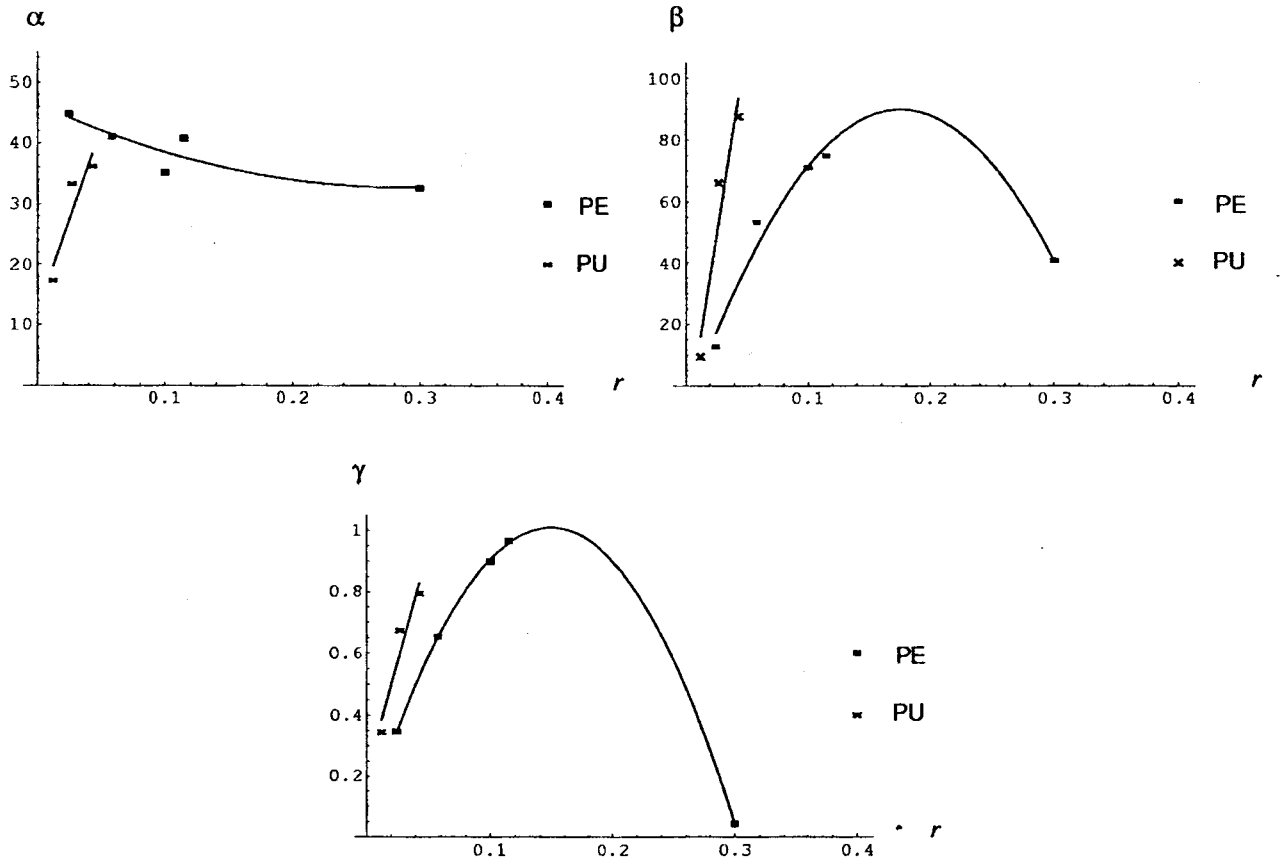


Fig. 3 Parameters α, β, and γ for polyethylene (PE) and polyurethane (PU) as functions of the relative mass density r

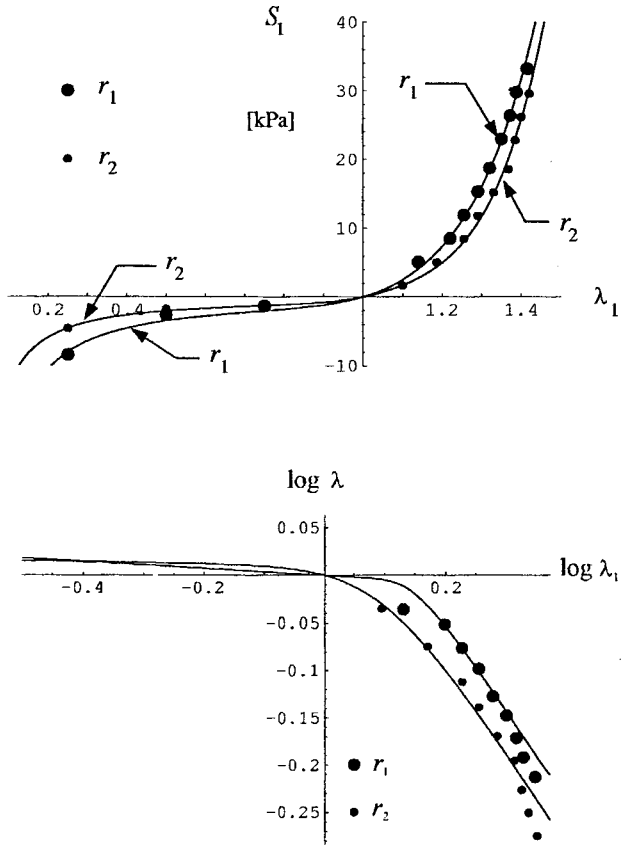


Fig. 4 Top: Nominal axial stress versus axial stretch from uniaxial tension tests of polyurethane (El-Ratal and Mallick [11]). The compression data ($\lambda_1<1$) were taken from Maiti et al. [2]. The solid lines correspond to the theoretical model. Bottom: Logarithmic measures of lateral versus axial stretch for uniaxial tension tests (El-Ratal and Mallick [11]). Observe the nonlinearity between log λ_1 and log λ.

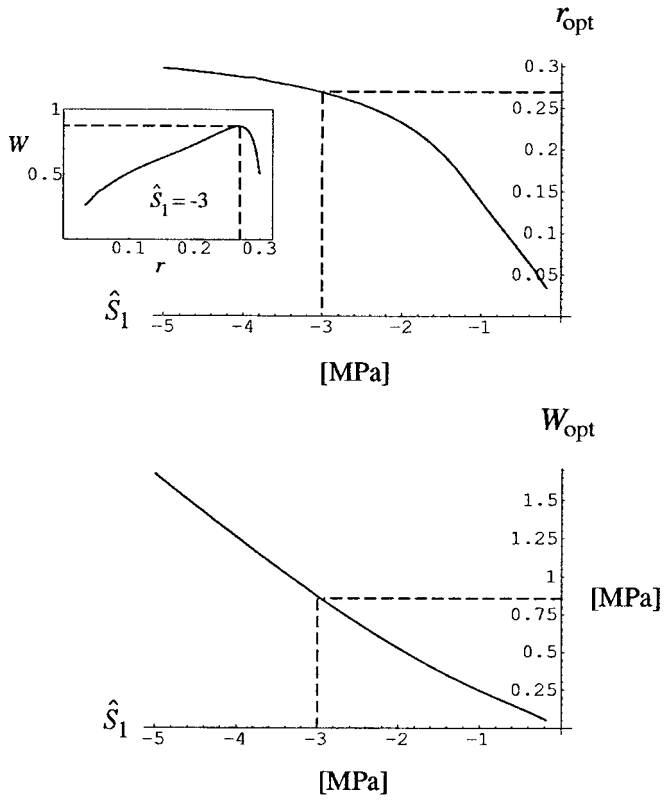


Fig. 5 Top: Optimal polyethylene foam density r_{opt} (for maximum stored energy) as a function of the prescribed load $\hat{S}_1$ in uniaxial homogeneous compression. As an example, the inset shows the stored energy as a function of r for $\hat{S}_1=-3$ MPa. Bottom: Maximum stored energy W_{opt} per unit reference (underformed) volume of polyethylene (for the optimal relative mass density) as a function of the prescribed compressive load $\hat{S}_1$.

''additional'' data were used in conjunction with the tension tests' data in order to obtain the values of m, n, α, β, and γ (with $\gamma>0$). The corresponding stress-stretch curves and the relation between the axial and lateral stretches (λ_1,λ) are shown in Fig. 4 and show a good agreement between theoretical and experimental values.

Since there are only two different densities, no correlation with r is proposed in this case. However, this example highlights an interesting point: As shown in Fig. 4, the relation between the logarithmic measures of the axial and lateral stretches in tension is not linear as predicted by Blatz-Ko's potential (or its generalization shown in Section 3). The potential (5) provides a better correlation (solid lines in Fig. 4 (bottom)). Furthermore, even though there are no experimental values for the lateral stretch λ reported in Maiti et al. [2] the behavior predicted in compression, as shown in Fig. 4, seems reasonable. Observe that λ is a monotonically decreasing function of λ_1 with values close to 1 when $\lambda_1\in(0,1]$ (except for $\lambda_1\ll 1$), which means that the material expands laterally by a small amount when it is compressed from $\lambda_1=1$. Although not shown in Figs. 1 and 2, the (computed) lateral stretches follow the same trend.

7 Optimization Problem

As an application of the model developed in the previous sections, consider an optimal design problem, namely the determination of the relative mass density of a foam that provides the maximum stored energy under a given loading condition. Suppose that a *homogeneous* prismatic body made out of foam is subject to *homogeneous* uniaxial compression. Let $\hat{S}_1$ be the given first Piola-Kirchhoff stress in the axial direction and suppose that the objective is to maximize the stored energy viewed as a function of the relative mass density r, subject to axial and lateral equilibrium. The optimization problem can be expressed as: *find* $r_{opt}\in[r_m,r_M]$ *that maximizes* W *subject to* (13) *and* (14) *with* $S_1=\hat{S}_1$, *where*

$$\overline{W}=2\hat{m}(r)(\lambda_1^{\hat{\alpha}(r)}+2\lambda^{\hat{\alpha}(r)}-3)+\hat{n}(r)[2(\lambda_1\lambda)^{\hat{\beta}(r)}+\lambda^{2\hat{\beta}(r)}-3]+\hat{p}(r)[(\lambda_1\lambda^2)^{-\hat{\gamma}(r)}-1].$$

As an example, this problem is solved numerically for polyethylene with $r_m=0.0245$ and $r_M=0.3$. The proposed method to solve this problem is actually similar to the algorithm used in Section 6: From the constraints (13) and (14) one can interpret the axial and lateral stretches as being implicit functions of r (i.e., for given $\hat{S}_1$, it is possible to solve (13) and (14) as a function of r, though not in closed form). However, it is straightforward to compute symbolically the derivatives $d\lambda/dr$ and $d\lambda_1/dr$ as functions of λ, λ_1 and r. Thus, the derivative of $\overline{W}$ with respect to r—taking into account the equilibrium contraints (13) and (14)—can be obtained as a function of λ, λ_1 and r. A modified version of the algorithm described by Twizell and Ogden [12] can be once again implemented with r as the unknown.

The optimal densities and the corresponding maximum value of the stored energy per unit reference volume are shown in Fig. 5

for different compressive loads. Clearly, there is a substantial difference between the stored energy for different relative mass densities.

8 Concluding Remarks

As shown in Section 6, the behavior of isotropic elastomeric foams made out of a given material can be described in parametric form for different relative mass densities using a single functional relation. In principle, the same procedure can be applied for other foam-like elastic materials. It is important to point out that the potential proposed here can only capture the behavior in compression in some average sense since typically strain localization occurs in experiments, hence the deformation might not be homogeneous as assumed in Section 6. In principle, this means that the potential, even though it is assumed to represent the foam's behavior locally in compression, might in fact not be able to describe the real deformation in detail. The reader is referred to the work of Lakes et al. [13] and Abeyaratne and Triantafyllidis [14] for analyses of microbuckling in compression.

Although the optimization problem described in Section 7 is relevant per se, more interesting problems can be considered. Specifically, one can analyze optimization problems where the deformation is nonhomogeneous. Also, in connection with recent methods developed in topology optimization for linearly elastic structures (see Bendsøe et al. [15]), one could generalize the nonlinear problem and allow r to be a function of position (i.e., one could consider nonhomogeneous bodies). However, it is important to have a well-posed formulation of the corresponding design problem since it is likely that the solution, if it exists, would *not* correspond to an *isotropic* foam. Consequently, it seems appropriate to extend the present formulation to anisotropic foams in order to identify the closure of the design problem.

Acknowledgments

The authors are pleased to acknowledge John Taylor (University of Michigan), Ciro Soto (Ford-SRL) and Robert Lipton (Worcester Polytechnic Institute) for valuable discussions and recommendations. This work was supported in part by the Ford Motor Company (SRL) under project No. 95-106R (S. T. and S. J.) and in part by the Dekaban Fund (S. J.). This work was conducted while one of the authors (S. J.) was on sabbatical at the Department of Aerospace Engineering of The University of Michigan (Ann Arbor, MI) and the other author (S. T.) was a research fellow at the same institution. They wish to thank Prof. J. Taylor for his kind hospitality and support.

References

[1] Ashby, M. F., 1983, ''The Mechanical Properties of Cellular Solids,'' Metall. Trans. A, **14A**, pp. 1755–1769.
[2] Maiti, S. K., Gibson, L. J., and Ashby, M. F., 1984, ''Deformation and Energy Absorption Diagrams for Cellular Solids,'' Acta Metall., **32**, No. 11, pp. 1963–1975.
[3] Gibson, L. J., and Ashby, M. F., 1997, *Cellular Solids: Structure and Properties*, Cambridge University Press, Cambridge, UK.
[4] Blatz, P. J., and Ko, W. L., 1962, ''Application of Finite Elastic Theory to the Deformation of Rubbery Materials,'' Trans. Soc. Rheol., **VI**, pp. 223–251.
[5] Ogden, R. W., 1972, ''Large Deformation Isotropic Elasticity: On the Correlation of Theory and Experiment for Compressible Rubberlike Solids,'' Proc. R. Soc. London, Ser. A, **328**, pp. 567–583.
[6] Ball, J. M., 1977, ''Convexity Conditions and Existence Theorems in Nonlinear Elasticity,'' Arch. Ration. Mech. Anal., **63**, pp. 337–403.
[7] Dacorogna, B., 1989, *Direct Methods in the Calculus of Variations*, Springer-Verlag, New York.
[8] Ball, J. M., and Murat, F., 1984, ''$W^{1,p}$-quasiconvexity and Variational Problems for Multiple Integrals,'' J. Funct. Anal., **58**, pp. 225–253.
[9] Horgan, C. O., 1996, ''Remarks on Ellipticity for the Generalized Blatz-Ko Constitutive Model for a Compressible Nonlinearly Elastic Solid,'' J. Elast., **42**, pp. 165–176.
[10] Warren, W. E., and Kraynik, A. M., 1997, ''Linear Elastic Behavior of a Low-Density Kelvin Foam With Open Cells,'' ASME J. Appl. Mech., **64**, pp. 787–794.
[11] El-Ratal, W. H., and Mallick, P. K., 1996, ''Elastic Response of Flexible Polyurethane Foams in Uniaxial Tension,'' J. Eng. Mater. Technol., **118**, pp. 157–161.
[12] Twizell, E. H., and Ogden, R. W., 1983, ''Nonlinear Optimization of the Material Constants in Ogden's Stress-Deformation Function for Incompressible Isotropic Elastic Materials,'' J. Aust. Math. Soc. B, Appl. Math., **24**, pp. 424–434.
[13] Lakes, R., Rosakis, P., and Ruina, A., 1993, ''Microbuckling Instability in Elastomeric Cellular Solids,'' J. Mater. Sci., **28**, No. 17, pp. 4667–4672.
[14] Abeyaratne, R., and Triantafyllidis, N., 1984, ''An Investigation of Localization in a Porous Elastic Material Using Homogenization Theory,'' ASME J. Appl. Mech., **51**, No. 3, pp. 481–486.
[15] Bendsøe, M. P., Guedes, J. M., Haber, R. B., Pedersen, P., and Taylor, J. E., 1994, ''An Analytical Model to Predict Optimal Properties in the Context of Optimal Structural Design,'' ASME J. Appl. Mech., **61**, pp. 930–937.

Electrode-Ceramic Interfacial Cracks in Piezoelectric Multilayer Materials

C. Ru
Department of Mechanical Engineering,
University of Alberta,
Edmonton, AB T6G 2G8, Canada

A thin electrode layer embedded at the interface of two piezoelectric materials represents a common feature of many electroceramic multilayer devices. The analysis of interface cracks between the embedded electrode layer and piezoelectric ceramic leads to a nonstandard mixed boundary value problem which likely prevents a general analytical solution. The present work shows that the associated mixed boundary value problem does indeed admit an exact elementary solution for a special case of major practical interest in which the two piezoelectric half-planes are poled in opposite directions perpendicular to the electrode layer. In this case, it is found that oscillatory singularity disappears, in spite of the unsymmetric characters of the problem, and electroelastic fields exhibit power singularities. Particular emphasis is placed on the near-tip singular stresses along the bonded interface. The results show that tensile stress exhibits the square root singularity along the interface whereas shear stress exhibits the dominant-order nonsquare root singularity. In addition, the present model indicates that a pure electric-field loading could induce the dominant-order singular shear stress directly ahead of the interface crack tip. [S0021-8936(00)00602-4]

1 Introduction

Ferroelectric/piezoelectric ceramics have widely been used in design of various modern electromechanical multilayer devices, such as transducers, capacitors, sensors, and actuators [1–6]. It has been observed that interfacial cracking between embedded thin electrode layers and ceramic matrix is a common cause of failure in many electroceramic multilayer devices (see [5,7–9]). On the other hand, existing theoretical works on interfacial cracks in piezoelectric materials (see, e.g., [10–11]) have been limited to interface cracks between two bonded piezoelectric half-planes in the absence of an intermediate electrode layer, which lead to conventional generalized displacement or traction boundary conditions and can be solved by the standard techniques of analytical continuation established for interface cracks in anisotropic media (see [12]). Evidently, these existing solutions are not applicable to electrode-ceramic interfacial cracks in electroceramic multilayer systems because the presence of a thin electrode layer changes electrical interface conditions while it does not affect mechanical interface conditions. Recently, electroelastic field around discontinuous electrode layers embedded between two piezoelectric half-planes has been studied ([13]) in the absence of any interface crack. To our knowledge, however, no effort has been made to analyze electrode-ceramic interfacial cracks for piezoelectric bimaterials bonded through a thin electrode layer, despite its obvious relevance to reliability mechanics of electroceramic multilayered devices.

The present work is devoted to plane-strain analysis of electrode-ceramic interface cracks in piezoelectric multilayer materials. According to common practice of multilayered electroceramic devices, the thickness of thin electrode layer is negligible and the ceramic materials on two sides of the electrode layer are assumed to be semi-infinite. As shown later, this leads to a nontrivial mixed boundary value problem which likely does not admit a general closed-form solution. Inspirited by the role of embedded electrode layers in multilayered electromechanical devices, then we consider a special case of major practical interest in which the upper and lower piezoelectric half-planes are poled in opposite directions perpendicular to the electrode layer (see Fig. 1). Our major finding is that the mixed boundary value problem can be solved explicitly for this practically significant case. In Sections 3 and 4, the exact solutions are given for a single interface crack situated on one side of the electrode layer (Fig. 1(*a*)), and for a pair of interface cracks situated symmetrically on the opposite sides of the electrode layer (Fig. 1(*b*)), respectively. In both cases, the crack-tip fields exhibit power singularities $r^{(\rho_i-1/2)}$ $(i=1,2,3)$, where $(-1/2)<\rho_1<\rho_2=0<\rho_3=(-\rho_1)$, and the oscillatory singularity disappears. In particular, the present linear piezoelectric model predicts that a pure electric-field loading could produce the dominant-order singular shear stress directly ahead of the electrode-ceramic interface crack tip. This gives a possible explanation for electrically induced interfacial debonding observed in many piezoelectric multilayer devices.

2 Formulation of a Linear Piezoelectric

The basic equations for a linear piezoelectric are

$$\sigma_{ij,j}=0, \quad D_{i,i}=0$$

$$\gamma_{ij}=\frac{1}{2}[u_{i,j}+u_{j,i}], \quad E_i=-\varphi_i, \tag{1}$$

$$\sigma_{ij}=C_{ijkl}\gamma_{kl}-e_{kij}E_k, \quad D_k=e_{kij}\gamma_{ij}+\varepsilon_{kl}E_l$$

where u_i and φ denote the displacement and electrical potential, σ_{ij}, γ_{ij}, E_i, and D_i are the stress, strain, electrical field, and electrical displacement, and C_{ijkl}, e_{ijk}, and ε_{ij} are the elastic, piezoelectric, and dielectric constants, respectively. In two-dimensional case (see [14,15,11,16]) let us consider the solution of the form

$$u(x,y)\equiv\begin{pmatrix} u_1(x,y) \\ u_2(x,y) \\ u_3(x,y) \\ \varphi(x,y) \end{pmatrix}=af(x+py) \tag{2}$$

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 4, 1997; final revision, November 26, 1999. Associate Technical Editor: M. M. Carroll. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

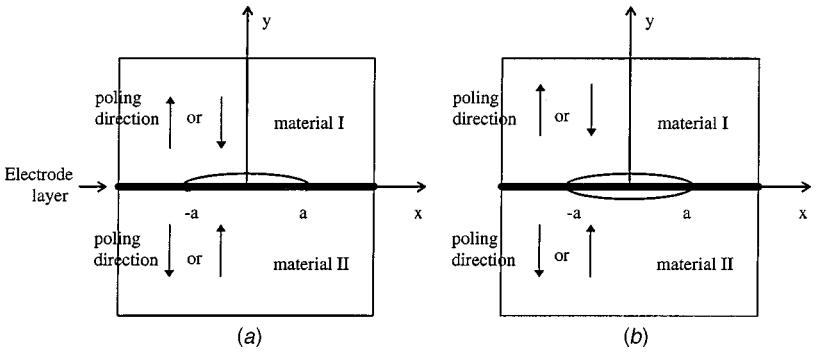


Fig. 1 The electrode-ceramic interface cracks lying between an electrode layer and ceramic matrix in a piezoelectric multilayer material; (*a*) a single interface crack, (*b*) two parallel interface cracks

where $f(*)$ is an analytic function, p a complex number, and a a constant four-element column. All equations of (1) are satisfied by (2) for arbitrary $f(*)$ if

$$[Q+p(R+R^T)+p^2T]a=0 \tag{3}$$

where the matrix R and the symmetric matrices Q and T are defined by the material constants (for details, see [16]). For existence of a nonzero vector a, p has to satisfy an eigenequation. For a stable material, eight eigenroots form four conjugate pairs with nonzero imaginary parts. Assume that p_α are four distinct roots with positive imaginary parts and $a_\alpha(\alpha=1,2,3,4)$ are the associated eigenvectors, the general solution of (1) can be given in the form

$$u(x,y)=\sum_{a=1}^{4}[a_\alpha f_\alpha(z_\alpha)+\overline{a_\alpha}\,\overline{f_\alpha(z_\alpha)}]=Af(z)+\overline{Af(z)},$$
$$f(z)=(f_1(z_1)f_2(z_2),f_3(z_3),f_4(z_4))^T, \tag{4}$$
$$z_\alpha=x+p_\alpha y,\quad \alpha=1,2,3,4$$

and the associated stresses and electrical displacements are given by

$$(\sigma_{2i},D_2)=\sum_{\alpha=1}^{4}[b_\alpha f'_\alpha(z_\alpha)+\overline{b_\alpha}\,\overline{f'_\alpha(z_\alpha)}]=Bf'(z)+\overline{Bf'(z)}, \tag{5}$$
$$(\sigma_{1i},D_1)=-\sum_{\alpha=1}^{4}[b_\alpha p_\alpha f'_\alpha(z_\alpha)+\overline{b_\alpha}\,\overline{p_\alpha}\,\overline{f'_\alpha(z_\alpha)}],\quad i=1,2,3$$

where $f_\alpha(*)$ are four arbitrary analytic functions, the column vectors $b_\alpha(\alpha=1,2,3,4)$ is determined by the corresponding pair (p_α,a_α) through

$$b_\alpha=(R^T+p_\alpha T)a_\alpha=\frac{-1}{p_\alpha}(Q+p_\alpha R)a_\alpha,\quad \alpha=1,2,3,4 \tag{6}$$

and the constant matrices A, B, and Y are defined by

$$A=(a_1,a_2,a_2,a_4),\quad B=(b_1,b_2,b_3,b_4),\quad Y\equiv iAB^{-1}. \tag{7}$$

3 An Electrode-Ceramic Interface Crack

Consider two piezoelectric half-planes bonded through a thin electrode layer. In view of the role of embedded electrode layers in electroceramic multilayer devices, a special case in which the two piezoelectric half-spaces are poled in opposite directions perpendicular to the electrode layer is of particular interest. In this case, plane deformation is decoupled from antiplane shear (see the Appendix). Here, we first consider a single electrode-ceramic interface crack, of length $2a$, situated on one side of the electrode layer, as shown in Fig. 1(*a*). The corresponding boundary value problem in plane strain is of the form

$$\sigma_{2i}^+=\sigma_{2i}^-=\sigma_{2i}^*,\quad D_2^+=D_2^*,E_1^-=0,\quad z\in L$$
$$\sigma_{2i}^+-\sigma_{2i}^-=0,\quad u_i^+-u_i^-=0,\quad E_1^+=E_1^-=0,\quad y=0,\quad |x|>a \tag{8}$$
$$\sigma_{ij}\to 0,\quad D_i\to 0,\quad |z|-\infty,\quad i,j=1,2$$

where $L=[-a,a]$, the superscripts ''+'' and ''−'' indicate the limit values from the upper and lower half-planes, respectively, and σ_{21}^*, σ_{22}^*, and D_2^* are the loadings parameters prescribed on the interface crack L.

3.1 Reduction to a Hilbert Problem. We first reduce the problem (8) to a standard Hilbert problem. According to the standard techniques of analytical continuation developed for interface cracks in anisotropic media (see [12]), on using (5), (7) the continuities of traction along the whole real axis can be written as

$$[B_If'_I(x)-\overline{B}_{II}\overline{f'_{II}}(x)]^+-[B_{II}f'_{II}(x)-\overline{B_I}\overline{f'_I}(x)]^-=(0,0,\Delta D_2(x))^T \tag{9}$$

where the subscript I or II denotes the quantities associated with the upper or lower half-planes, and $\Delta D_2=(D_2^+-D_2^-)$ denotes the unknown discontinuity of D_2 across the interface, which approaches zero at infinity. It then follows from (9) that

$$B_If'_I(z)-\overline{B}_{II}\overline{f'_{II}}(z)=C(z),\quad y>0$$
$$B_{II}f'_{II}(z)-\overline{B_I}\overline{f'_I}(z)=C(z),\quad y<0 \tag{10}$$

where $C(z)$ is defined by

$$C(z)\equiv\frac{1}{2\pi i}\int_{-\infty}^{+\infty}\frac{(0,0,\Delta D_2(t))^T}{t-z}dt \tag{11}$$

which approaches zero as quickly as $1/z^2$ at infinity. On the other hand, the continuities of displacement and tangential electrical field along the bonded interface give

$$[Y_IB_If'_I(x)+\overline{Y}_{II}\overline{B_{II}f'_{II}}(x)]^+-[\overline{Y}_I\overline{B_If'_I}(x)+Y_{II}B_{II}f'_{II}(x)]^-=0, \tag{12}$$
$$y=0,\quad |x|>a$$

then it follows that

$$[Y_I B_I f_I'(z)+\overline{Y_{II}}\,\overline{B_{II} f_{II}'}(z)]=G(z), \quad y>0$$
$$[\overline{Y_I}\,\overline{B_I f_I'}(z)+Y_{II}B_{II} f_{II}'(z)]=G(z), \quad y<0 \tag{13}$$

where $G(z)$ is an unknown function which is analytical in the entire plane except L and approaches zero at infinity as quickly as $1/z^2$. Finally, the condition $E_1=0$ along the lower real axis gives

$$2Y_{II}\left[B_{II} f_{II}'(x)-\frac{1}{2}C(x)\right]^- -2\overline{Y_{II}}\left[B_I f_I'(x)-\frac{1}{2}C(x)\right]^+ + \overline{Y_{II}}C(z^+)+Y_{II}C(z^-)=(*,*,0)^T, \quad y=0 \tag{14}$$

where "*" denotes some arbitrary quantities, and the boundary condition on the upper crack face gives

$$\left[B_I f_I'(x)-\frac{1}{2}C(z)\right]^+ +\left[B_{II} f_{II}'(x)-\frac{1}{2}C(z)\right]^- =\begin{pmatrix}\sigma_{21}^*\\ \sigma_{22}^*\\ D_2^*\end{pmatrix}-\frac{1}{2}(0,0,\Delta D_2(x))^T, \tag{15}$$

$$y=0, \quad |x|<a.$$

Note that (10) and (13) give

$$(Y_I+\overline{Y_{II}})\left[B_I f_I'(z)-\frac{1}{2}C(z)\right]=\frac{1}{2}(\overline{Y_{II}}-Y_I)C(z)+G(z), \quad y>0$$
$$(Y_{II}+\overline{Y_I})\left[B_{II} f_{II}'(z)-\frac{1}{2}C(z)\right]=\frac{1}{2}(\overline{Y_I}-Y_{II})C(z)+G(z), \quad y<0. \tag{16}$$

Hence, substitution of (16) into (14) and (15) yields a nontrivial mixed boundary value problem for the two unknown functions, $\Delta D_2(x)$, defined on the whole real axis, and the vector function $G(z)$, analytical in the entire plane except L. Unfortunately, a general closed-form solution seems unavailable for this mixed boundary value problem.

The major finding of this work is that the above problem does indeed admit an exact elementary solution when the two piezoelectric half-planes are poled in opposite directions perpendicular to the intermediate electrode layer. To demonstrate this, from now on, we shall use the assumption that the upper and lower piezoelectric half-planes are poled in opposite directions parallel to the y-axis. Under this condition, the condition (14) becomes

$$\left([Y_{II}]_{33}-\frac{Y_{23}^2}{Y_{22}}\right)(C^+(z)+C^-(z)) =[G^+(z)-G^-(z)]_3+\frac{[Y_{II}]_{23}}{Y_{22}}[G^+(z)-G^-(z)]_2, \quad y=0 \tag{17}$$

here $C(z)$ denotes its third component, the subscripts 2 and 3 denotes the components of $G(z)$, and the subscript I or II is omitted when the associated quantity has the same value in the two half-planes (see Appendix). Hence, along the bonded interface across which $G(z)$ is continuous, (17) gives

$$C(x^+)+C(x^-)=0, \quad y=0, \quad |x|>a. \tag{18}$$

Let us define

$$C(z)=\frac{C_0(z)}{\sqrt{a^2-z^2}}, \quad G(z)=\frac{G_0(z)}{\sqrt{z^2-a^2}} \tag{19}$$

where all multivalued functions in (19) are defined by the single-valued branch cut along the negative real-axis (then $\sqrt{(a^2-z^2)}$ is an analytic function with the branch cuts $[-\infty,-a]$ and $[a,\infty]$, while $\sqrt{(z^2-a^2)}$ is an analytic function with the branch cut $[-a,a]$), and $G_0(z)$ is analytic in the entire plane except L. It follows from (18) that $C_0(z)$ is analytic in the entire plane except L. Thus, the problem can now be reduced to a boundary value problem on L.

In fact, in terms of $C_0(z)$ and $G_0(z)$, the condition (17) on L has a tractable form

$$i\left([Y_{II}]_{33}-\frac{Y_{23}^2}{Y_{22}}\right)(C_0^+(z)+C_0^-(z)) =[G_0^+(z)+G_0^-(z)]_3+\frac{[Y_{II}]_{23}}{Y_{22}}[G_0^+(z)+G_0^-(z)]_2, \quad x\in L. \tag{20}$$

Because $C_0(z)$ and $G_0(z)$ approach zero at infinity and cannot exhibit square-root singularity at the crack tip, it follows from (20) that

$$i\left([Y_{II}]_{33}-\frac{Y_{23}^2}{Y_{22}}\right)C_0(z)=[G_0(z)]_3+\frac{[Y_{II}]_{23}}{Y_{22}}[G_0(z)]_2 \tag{21}$$

in the whole z-plane. Now, the expression (16) has the form

$$B_I f_I'(z)-\frac{1}{2}C(z)=\begin{pmatrix}0\\ -\dfrac{[Y_I]_{23}}{2Y_{22}}\\ 0\end{pmatrix}\frac{C_0(z)}{\sqrt{a^2-z^2}} +(Y_I+\overline{Y_{II}})^{-1}\frac{G_0(z)}{\sqrt{z^2-a^2}}, \quad y>0$$

$$B_{II} f_{II}'(z)-\frac{1}{2}C(z)=\begin{pmatrix}0\\ \dfrac{[Y_I]_{23}}{2Y_{22}}\\ 0\end{pmatrix}\frac{C_0(z)}{\sqrt{a^2-z^2}} +(Y_{II}+\overline{Y_I})^{-1}\frac{G_0(z)}{\sqrt{z^2-a^2}}, \quad y<0. \tag{22}$$

Substituting (21) and (22) into the remaining boundary condition (15), we obtain a standard Hilbert condition for $G_0(z)$ on L

$$MG_0^+-\overline{M}G_0^-=i\begin{pmatrix}\sigma_{12}^*\\ \sigma_{22}^*\\ D^*\end{pmatrix}\sqrt{|x^2-a^2|}, \quad x\in L \tag{23}$$

where M is a constant Hermitian matrix defined by

$$M=\frac{1}{2}\begin{pmatrix} \dfrac{Y_{33}}{(Y_{11}Y_{33}-Y_{13}^2)} & 0 & \dfrac{i[Y_{II}]_{13}}{(Y_{11}Y_{33}-Y_{13}^2)} \\ 0 & \dfrac{1}{Y_{22}}+\dfrac{Y_{23}^2}{Y_{22}^2\left(Y_{33}-\dfrac{Y_{23}^2}{Y_{22}}\right)} & \dfrac{[Y_{II}]_{23}}{Y_{22}\left(Y_{33}-\dfrac{Y_{23}^2}{Y_{22}}\right)} \\ \dfrac{-i[Y_{II}]_{13}}{(Y_{11}Y_{33}-Y_{13}^2)} & \dfrac{[Y_{II}]_{23}}{Y_{22}\left(Y_{33}-\dfrac{Y_{23}^2}{Y_{22}}\right)} & \dfrac{Y_{11}}{(Y_{11}Y_{33}-Y_{13}^2)}+\dfrac{1}{\left(Y_{33}-\dfrac{Y_{23}^2}{Y_{22}}\right)} \end{pmatrix} \tag{24}$$

and Y_{ij} $(i,j=1,2,3)$ are all real numbers associated with a piezoelectric poled in the positive y-direction, as defined by $(A1)$ and $(A2)$ (see Appendix).

3.2 General Solution. The solution of (23) can be obtained using the standard method. Substituting

$$G_0(z)=z^\rho g$$

into the homogeneous equation of (23), we obtain the eigenvalue problem

$$[\mathrm{Re}(M)]^{-1}i\ \mathrm{Im}[M]g-\lambda g=0,\quad \lambda=\frac{1-e^{2i\rho\pi}}{1+e^{2i\rho\pi}} \tag{25}$$

where the branch cut is made from the crack tip along the negative real axis, and g is a nonzero constant vector. It turns out that three distinct eigenroots λ_i $(i=1,2,3)$ are

$$\lambda_2=0,\quad \lambda_{1,3}=\pm i\sqrt{\left|\frac{M_{22}M_{13}^2}{M_{11}(M_{22}M_{33}-M_{23}^2)}\right|} \tag{26}$$

which correspond to three distinct *real* singularity indexes

$$\rho_1<\rho_2=0<\rho_3=-\rho_1,\quad i\sin[2\rho_k\pi]=\frac{-2\lambda_k}{1-\lambda_k^2},\qquad |\rho_k|<\frac{1}{2},\quad k=1,2,3 \tag{27}$$

where M_{ij} denote the elements of the matrix M defined by (24). Obviously, the singularity indexes remain unchanged when the poling directions of the two piezoelectric half-spaces reverse simultaneously. In particular, because all singularity indexes are real, there is no oscillatory singularity. Here, it should be stressed that, due to the electrode layer between two symmetric piezoelectric half-planes, the present interface crack problem is not symmetric about the interface. Hence, the nonexistence of oscillatory singularity at the electrode-ceramic interface crack is not self-evident.

Three linearly independent eigenvectors corresponding to three distinct eigenroots can be given by

$$g_1=\begin{pmatrix} 1 \\ \dfrac{-\lambda_1 M_{11}M_{23}}{M_{13}M_{22}} \\ \dfrac{\lambda_1 M_{11}}{M_{13}} \end{pmatrix},\quad g_2=\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\quad g_3=\begin{pmatrix} 1 \\ \dfrac{\lambda_1 M_{11}M_{23}}{M_{13}M_{22}} \\ \dfrac{-\lambda_1 M_{11}}{M_{13}} \end{pmatrix}. \tag{28}$$

Thus, the general solution of $G_0(z)$ can be sought in the form (see [17])

$$G_0(z)=\sum_{i=1}^{3} f_i(z)g_i \tag{29}$$

where $f_i(z)$ $(i=1,2,3)$ are three unknown functions which are analytic in the entire plane except L and approach zero at infinity. Substituting (29) into (23) yields

$$f_k(x)^+-e^{2i\rho_k\pi}f_k(x)^-=\frac{id_k}{1+\lambda_k}\sqrt{|(x^2-a^2)|},\quad x\in L,\quad k=1,2,3 \tag{30}$$

where the constants d_i $(i=1,2,3)$ are determined by

$$[\mathrm{Re}(M)]^{-1}\begin{pmatrix} \sigma_{12}^* \\ \sigma_{22}^* \\ D_2^* \end{pmatrix}=\sum_{i=1}^{3} d_i g_i \tag{31}$$

with the results

$$d_1=\frac{\sigma_{12}^*}{2M_{11}}+\frac{M_{13}(D_2^*M_{22}-M_{23}\sigma_{22}^*)}{2\lambda_1 M_{11}(M_{22}M_{33}-M_{23}^2)},\quad d_2=\frac{\sigma_{22}^*}{M_{22}},\qquad d_3=\frac{\sigma_{12}^*}{2M_{11}}-\frac{M_{13}(D_2^*M_{22}-M_{23}\sigma_{22}^*)}{2\lambda_1 M_{11}(M_{22}M_{33}-M_{23}^2)}. \tag{32}$$

Since $f_i(z)$ $(i=1,2,3)$ cannot exhibit singularity of order higher than $(-1/2)$ at the crack tip, the condition (30) gives

$$f_k(z)=\frac{d_k}{2}\left[\sqrt{z^2-a^2}-(z+2a\rho_k)\left(\frac{z-a}{z+a}\right)^{\rho_k}\right],\quad k=1,2,3. \tag{33}$$

Hence, once $G_0(z)$ is obtained from (29), $C_0(z)$ and $Bf'(z)$ can be found from (21) and (22), and then electroelastic fields can be calculated by substituting z in the Stroh's functions $f'(z)$ by the respective variables z_α $(\alpha=1,2,3)$.

4 Two Parallel Interface Cracks

Although a single interfacial crack examined in Section 3 is of basic importance, a pair of parallel cracks caused by symmetric debonding of the electrode layer from ceramic matrix, as shown in Fig. 1(*b*), is also of practical interest. In this case, the boundary value problem is of the form

$$\sigma_{2i}^+=\sigma_{2i}^-=\sigma_{2i}^*,\quad D_2^+=D_2^*,\quad D_2^-=-D_2^*,\quad z\in L$$
$$\sigma_{2i}^+-\sigma_{2i}^-=0,\quad u_i^+-u_i^-=0,\quad E_1^+=E_1^-=0,\quad y=0,\quad |x|>a$$
$$\sigma_{ij}\to 0,\quad D_i\to 0,\quad |z|\to\infty,\quad i,j=1,2. \tag{34}$$

It is not difficult to verify that all formulas given in Section 3 up to (18) remain true provided that (i) $\Delta D_2(x)$ appearing in (11) is known on the segment $L=[-a,a]$ as

$$\Delta D_2(t)=C(z)^+-C(z)^-=2D_2^*,\quad y=0,\quad |x|<a \tag{35}$$

and (ii) the condition (14) holds only on the bonded part of the interface and then reduces to the form (18). Consequently, from (18) and (35), it is found that

$$C(z)=\frac{-iD_2^*}{\sqrt{a^2-z^2}}[\sqrt{z^2-a^2}-z]. \tag{36}$$

Thus, the remaining boundary condition (15) leads to a standard Hilbert condition for the unknown function $G(z)$ on L as follows:

$$NG(x)^+ + \overline{N}G(x)^- = \begin{pmatrix} \sigma_{12}^* \\ \sigma_{22}^* + \dfrac{[Y_I]_{23}}{Y_{22}} D_2^* \\ 0 \end{pmatrix}, \quad x \in L,$$
$$N \equiv (Y_I + \overline{Y_{II}})^{-1}. \tag{37}$$

Similar to Section 3.2, the general solution of (37) is of the form

$$G(z)=\sum_{i=1}^{3} q_i(z) h_i \tag{38}$$

where h_i and $q_i(z)$ $(i=1,2,3)$ are given by

$$h_1=\begin{pmatrix} 1 \\ 0 \\ \dfrac{-i\lambda_1 Y_{33}}{[Y_{II}]_{13}} \end{pmatrix}, \quad h_2=\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad h_3=\begin{pmatrix} 1 \\ 0 \\ \dfrac{i\lambda_1 Y_{33}}{[Y_{II}]_{13}} \end{pmatrix}$$
$$q_i(z)=\frac{b_i}{2}\left[1-\frac{(z+2a\rho_i)}{\sqrt{z^2-a^2}}\left(\frac{z-a}{z+a}\right)^{\rho_i}\right], \quad i=1,2,3 \tag{39}$$

and b_i $(i=1,2,3)$ are three constants determined by

$$b_1=b_3=\sigma_{12}^*\frac{Y_{11}Y_{33}-Y_{13}^2}{Y_{33}}, \quad b_2=2Y_{22}\left(\sigma_{22}^*+\frac{[Y_I]_{23}}{Y_{22}}D_2^*\right) \tag{40}$$

and three distinct singularity indexes are given by

$$e^{2i\rho_k k\pi}=\frac{1-\lambda_k}{1+\lambda_k}, \quad k=1,2,3, \quad \lambda_2=0, \quad \lambda_{1,3}=\pm i\sqrt{\left|\frac{[Y_{II}]_{13}^2}{Y_{11}Y_{33}}\right|}. \tag{41}$$

Note that

$$\left|M_{33}-\frac{M_{23}^2}{M_{22}}\right|>|M_{33}|, \quad \left|\frac{M_{13}^2}{M_{11}M_{33}}\right|<\left|\frac{Y_{13}^2}{Y_{11}Y_{33}}\right|.$$

It is seen that the absolute value of the singularity index ρ_1 given by (41) is bigger than that obtained from (27) for a single interface crack. On the other hand, similar to a single interface crack in Section 3, all singularity indexes in (41) are real and then there is no oscillatory singularity.

5 Singular Stresses Along the Bonded Interface

The complete solutions obtained in Sections 3 and 4 are used to study the near-tip singular field, with an emphasis on the singular stresses ahead of the interface crack tip.

5.1 A Single Interface Crack. From (19), (21), (22), the upper limit of Stroh's function ahead of the crack tip $x=a$ is

$$B_I f_I'(x^+)=\frac{1}{\sqrt{|x^2-a^2|}}MG_0(x), \quad x>a. \tag{42}$$

Thus, the singular parts of the tensile and shear stresses ahead of the crack tip $x=a$ are

$$\sigma_{22}=\frac{-\sigma_{22}^*(2a)^{1/2}}{2\sqrt{|x-a|}}, \tag{43}$$

$$\sigma_{12}=\frac{-\sigma_{12}^*(2a)^{1/2}}{4\sqrt{|x-a|}}\left[(1+2\rho_1)\left(\frac{|x-a|}{2a}\right)^{\rho_1}+(1-2\rho_1)\left(\frac{2a}{|x-a|}\right)^{\rho_1}\right]$$
$$+\frac{-\lambda_1 M_{11}(M_{22}D_2^*-M_{23}\sigma_{22}^*)(2a)^{1/2}}{4M_{13}M_{22}\sqrt{|x-a|}}$$
$$\times\left[(1-2\rho_1)\left(\frac{2a}{|x-a|}\right)^{\rho_1}-(1+2\rho_1)\left(\frac{|x-a|}{2a}\right)^{\rho_1}\right], \tag{44}$$

and the upper limit of the normal electrical displacement for $x>a$ is

$$D_2(x^+)=\frac{-\sigma_{22}^*M_{23}(2a)^{1/2}}{2M_{22}\sqrt{|x-a|}}-\frac{\sigma_{12}^*M_{13}(2a)^{1/2}}{4\lambda_1 M_{11}\sqrt{|x-a|}}$$
$$\times\left[(1-2\rho_1)\left(\frac{2a}{|x-a|}\right)^{\rho_1}-(1+2\rho_1)\left(\frac{|x-a|}{2a}\right)^{\rho_1}\right]$$
$$-\frac{(M_{22}D_2^*-M_{23}\sigma_{22}^*)(2a)^{1/2}}{4M_{22}\sqrt{|x-a|}}$$
$$\times\left[(1-2\rho_1)\left(\frac{2a}{|x-a|}\right)^{\rho_1}+(1+2\rho_1)\left(\frac{|x-a|}{2a}\right)^{\rho_1}\right]. \tag{45}$$

Further, on using (11), (19), (21), it can be verified that the lower limit of the normal electrical displacement for $x>a$ is

$$D_2(x^-)=\frac{\sigma_{12}^*\lambda_1\left(M_{33}-\dfrac{Y_{22}}{Y_{22}Y_{33}-Y_{23}^2}\right)(2a)^{1/2}}{4M_{13}\sqrt{|x-a|}}$$
$$\times\left[(1-2\rho_1)\left(\frac{2a}{|x-a|}\right)^{\rho_1}-(1+2\rho_1)\left(\frac{|x-a|}{2a}\right)^{\rho_1}\right]$$
$$-\frac{(M_{22}D_2^*-M_{23}\sigma_{22}^*)(2a)^{1/2}}{4\sqrt{|x-a|}}\frac{\left(M_{33}-\dfrac{Y_{22}}{Y_{22}Y_{33}-Y_{23}^2}\right)}{(M_{22}M_{33}-M_{23}^2)}$$
$$\times\left[(1-2\rho_1)\left(\frac{2a}{|x-a|}\right)^{\rho_1}+(1+2\rho_1)\left(\frac{|x-a|}{2a}\right)^{\rho_1}\right]. \tag{46}$$

It is seen from (43), (44) that (i) the tensile stress exhibits a square root singularity and the corresponding stress intensity factor is determined by the remote tensile stress. In particular, this implies that a pure electric-field loading does not induce any singular tensile stress ahead of the electrode-ceramic interface crack tip. (ii) the shear stress exhibits the dominant-order nonsquare root singularity. In addition, both σ_{22}^* and D_2^* could give rise to a dominant-order singular shear stress at the bonded interface. These results suggest that the interface shear stress could play a significant role in debonding of the electrode layers from ceramic matrix in piezoelectric multilayer materials.

5.2 Two Parallel Interface Cracks. If a pair of parallel electrode-ceramic cracks (as shown in Fig. 1(*b*)) are considered, it follows from (19), (21), (22) that

$$B_I f_I'(x^+)=\begin{pmatrix} 0 \\ \dfrac{[Y_I]_{23}}{Y_{22}} \\ -1 \end{pmatrix}\frac{D_2^*}{4}\frac{(2a)^{1/2}}{\sqrt{|x-a|}}+NG(x), \qquad x>a. \tag{47}$$

Thus, the singular stresses directly ahead of the crack tip $x=a$ are

$$\sigma_{22}(x)=\frac{-\sigma_{22}^{*}(2a)^{1/2}}{2|x-a|^{1/2}},\qquad x>a \tag{48}$$

$$\sigma_{12}(x)=-\sigma_{12}^{*}\frac{(2a)^{1/2}}{4|x-a|^{1/2}}\left[(1+2\rho_1)\left(\frac{|x-a|}{2a}\right)^{\rho_1}+(1-2\rho_1)\left(\frac{2a}{|x-a|}\right)^{\rho_1}\right],\quad x>a \tag{49}$$

and the upper normal electrical displacement (the lower one can be obtained immediately from the symmetry of the present case) is given by

$$D_2(x^{+})=\frac{-D_2^{*}(2a)^{1/2}}{2\sqrt{|x-a|}}+\sigma_{12}^{*}\frac{i\lambda_1 Y_{\mathbf{II}}(2a)^{1/2}}{4[Y_{\mathrm{II}}]_{13}|x-a|^{1/2}}\times\left[(1+2\rho_1)\left(\frac{|x-a|}{2a}\right)^{\rho_1}-(1-2\rho_1)\left(\frac{2a}{|x-a|}\right)^{\rho_1}\right]. \tag{50}$$

Similar to Section 5.1, the tensile stress given by (48) exhibits a square root singularity while the shear stress given by (49) exhibits the dominant-order nonsquare root singularity. In the present case, however, the loading parameters σ_{22}^{*} and D_2^{*} do not induce any singular shear stress at the bonded interface.

5.3 Effects of an Electrical Field on Singular Shear Stress Electroceramic multilayer devices are used usually under electrical or electrical/mechanical loading. Hence, it is of great interest to study the effects of an electrical field on interfacial crack growth in piezoelectric multilayer materials. As stated above, only shear stress exhibits the dominant-order singularity ahead of the interface crack tip. Therefore, let us examine the effects of an electrical field on the dominant-order singular shear stress.

Electrically Induced Singular Shear Stress. First, we consider a pure electrical loading. Because electrically induced interfacial cracking has been observed as one of main failure models in many electroceramic multilayer systems (see [5,7,8,9]), it is of particular interest to examine whether or not the linear piezoelectric model predicts an electric-field induced singular stress ahead of the electrode-ceramic interface crack tip. Recently, Ru et al. [18] have examined interfacial cracking in electrostrictive multilayer systems. Similar issue for piezoelectric multilayer materials has yet to be investigated.

Under a pure electrical loading, the tensile stress σ_{22} does not exhibit any singularity along the bonded interface, and the shear stress σ_{12} along the bonded interface is given by

$$\sigma_{12}=\frac{-i\lambda_1(1+2\rho_1)Y_{33}D_2^{*}(2a)^{1/2-\rho_1}}{4[Y_{\mathrm{II}}]_{13}|x-a|^{1/2-\rho_1}},\quad \rho_1<0. \tag{51}$$

Evidently, the electrically induced singular shear stress (51) changes sign when the direction of applied electrical field reverses. Hence, if the singular interface shear stress plays a dominant role in electrically induced interfacial debonding, the present model appears to predict that the electrically induced interfacial debonding is not sensitive to reversal of the applied electrical field. This conclusion distinguishes the electrode-ceramic interfacial debonding from crack growth in a homogeneous piezoelectric medium, for the latter some experiments (see [19–20]) and nonlinear theoretical models (see, e.g., [21]) have showed that the effect of an electric field on crack growth essentially depends on the direction of the applied electrical loading (although the linear piezoelectric model does not predicts any electric-field induced stress intensity factor). To our knowledge, it seems that no clear experimental result has been reported in the literature on the role of an electric-field in interfacial cracking in piezoelectric media.

Effect of Electrical Field on Singular Shear Stress Induced by a Remote Tensile Stress. Next, we consider a combined electrical/mechanical loading. Of particular interest is the effect of an electrical field on the singular shear stress induced by a remote tensile stress $(-\sigma_{22}^{*})>0$. In this case, the dominant-order singular shear stress is

$$\sigma_{12}=\frac{\lambda_1(1+2\rho_1)M_{11}(M_{22}D_2^{*}-M_{23}\sigma_{22}^{*})(2a)^{(1/2)-\rho_1}}{4M_{13}M_{22}|x-a|^{(1/2)-\rho_1}}. \tag{52}$$

Note that

$$M_{22}>0,\quad M_{23}\propto-[Y_{\mathrm{II}}]_{23}\propto[Y_{\mathrm{I}}]_{23}.$$

It is concluded that electrical field enhances the dominant-order singular shear stress caused by the remote tensile stress if

$$D_2^{*}[Y_{\mathrm{I}}]_{23}>0. \tag{53}$$

On the other hand, electrical field reduces the dominant-order singular shear stress caused by the remote tensile stress if

$$D_2^{*}[Y_{\mathrm{I}}]_{23}<0. \tag{54}$$

It turns out from (52)–(54) that an electrical field enhances (or reduces) the singular shear stress caused by a remote tensile stress if the electrical field is applied opposite to (or in the same direction as) the poling direction. This conclusion is contrary to the known effect of an electric field on crack growth in a homogeneous piezoelectric medium under combined electrical/mechanical loading (see [19–20]) where these authors concluded that crack growth is enhanced (or impeded) by an electrical field applied in the same direction as (or opposite to) the poling direction.

6 Conclusions

In view of practical importance of interfacial cracking observed in many piezoelectric multilayer devices, the problem of electrode-ceramic interface cracks is studied in the paper. Owing to the presence of an intermediate electrode layer, the analysis of interface cracks between two piezoelectric materials is led to a nonstandard mixed boundary value problem which likely does not admit a general analytical solution. Our major finding is that this mixed boundary value problem can be solved explicitly when the two piezoelectric half-planes are poled in opposite directions perpendicular to the electrode layer. In these cases, there is no oscillatory singularity in spite of the lack of a symmetry about the interface when a single interface crack is considered. Furthermore, the tensile stress is found to exhibit a square-root singularity ahead of the interfacial crack tip, while the shear stress exhibits the dominant-order power singularity. In Section 5.3, the effect of an electrical field on interfacial debonding is discussed in terms of the dominant-order singular shear stress. In particular, the present model predicts that a pure electric-field loading could induce a dominant-order singular shear stress ahead of the interface crack tip. This provides a possible explanation for electrically induced interfacial debonding observed in many piezoelectric multilayer devices.

Acknowledgment

The author acknowledges the financial support of the Natural Science and Engineering Research Council of Canada through a grant awarded to Dr. David Steigmann of the University of Alberta.

Appendix

If the poling direction of a piezoelectric is parallel to the *y*-direction, plane deformation in the *x*-*y* plane is decoupled from antiplane shearing. In this case, if the poling axis is along the positive *y*-axis, we have (see [11])

$$\mathrm{Re}[Y]=\begin{pmatrix} 1/C_L & 0 & 0 \\ 0 & 1/C_T & 1/e \\ 0 & 1/e & -1/\varepsilon \end{pmatrix},$$

$$C_L>0, \quad C_T>0, \quad e>0, \quad \varepsilon>0. \tag{A1}$$

Further, according to some known numerical solutions ([22,23]) it can be verified ([21,24]) that

$$\mathrm{Im}[Y]=\begin{pmatrix} 0, & Y_{12}, & Y_{13} \\ -Y_{12}, & 0 & 0 \\ -Y_{13}, & 0, & 0 \end{pmatrix}, \quad Y_{12}>0, \quad Y_{13}<0 \tag{A2}$$

where Y_{12} and Y_{13} are two real numbers. Hence, the matrix Y (see (7)) for a piezoelectric poled in the positive y-direction is of the form

$$Y^{+}=\begin{pmatrix} Y_{11}, & iY_{12}, & iY_{13} \\ -iY_{12}, & Y_{22}, & Y_{23} \\ -iY_{13}, & Y_{23}, & Y_{33} \end{pmatrix} \tag{A3}$$

where all other real numbers, Y_{11}, Y_{22}, Y_{23}, and Y_{33}, can be obtained by comparing (A1) with (A3). Now, through a rotation of the coordinate system, the matrix Y for a piezoelectric poled in the negative y-direction is found to have the form

$$Y^{-}=\begin{pmatrix} Y_{11}, & iY_{12}, & -iY_{13} \\ -iY_{12}, & Y_{22}, & -Y_{23} \\ iY_{13}, & -Y_{23}, & Y_{33} \end{pmatrix} \tag{A4}$$

where the superscript ''$-$'' indicates the poling direction. Further, it follows that

$$[\overline{Y^{+}}+Y^{-}]^{-1}=\frac{1}{2(Y_{11}Y_{33}-Y_{13}^{2})} \times\begin{pmatrix} Y_{33} & 0 & iY_{13} \\ 0 & \dfrac{(Y_{11}Y_{33}-Y_{13}^{2})}{Y_{22}}, & 0 \\ -iY_{13} & 0 & Y_{11} \end{pmatrix}. \tag{A5}$$

$$(\overline{Y^{+}}+Y^{-})^{-1}(\overline{Y^{+}}-Y^{-})=\begin{pmatrix} 0 & i\dfrac{Y_{13}Y_{23}-Y_{12}Y_{33}}{Y_{33}Y_{11}-Y_{13}^{2}} & 0 \\ i\dfrac{Y_{12}}{Y_{22}} & 0 & \dfrac{Y_{23}}{Y_{22}} \\ 0 & \dfrac{Y_{11}Y_{23}-Y_{12}Y_{13}}{Y_{33}Y_{11}-Y_{13}^{2}} & 0 \end{pmatrix}. \tag{A6}$$

References

[1] Berlincourt, D., 1981, ''Piezoelectric Ceramics: Characteristics and applications,'' J. Am. Ceram. Soc., **70**, pp. 1506–1595.

[2] Pohanka, R. C., and Smith, P., 1987, ''Recent Advances in Piezoelectric Ceramics,'' *Electronic Ceramics*, L. M. Levinson, ed., Marcel Dekker, New York, pp. 45–146.

[3] Freiman, S. W., 1989, ''Review of Mechanically Related Failures of Ceramic Capacitors and Capacitor Materials,'' J. Am. Ceram. Soc., **72**, pp. 2258–2263.

[4] Newnham, R. E., 1989, ''Electroceramics,'' Rep. Prog. Phys., **52**, pp. 123–156.

[5] Winzer, S. R., Shankar, N., and Ritter, A., 1989, ''Desiging Cofired Multilayer Electrostrictive Actuators for Reliability,'' J. Am. Ceram. Soc., **72**, pp. 2246–2257.

[6] Cao, H. C., and Evans, A. G., 1994, ''Electric-Field Induced Fatigue Crack Growth in Piezoelectrics,'' J. Am. Ceram. Soc., **77**, pp. 1783–1786.

[7] Furuta, A., and Uchino, K., 1993, ''Dynamic Observation of Crack Propagation in Piezoelectric Multilayer Ceramic Actuators,'' J. Am. Ceram. Soc., **76**, pp. 1615–1617.

[8] Aburatani, H., Harada, S., Uchino, K., and Furuta, A., 1994, ''Destruction Mechanism of Ceramic Multilayer Actuators,'' Jpn. J. Appl. Phys., **33**, pp. 3091–3094.

[9] Freiman, S. W., and White, G. S., 1994, ''Intelligent Ceramic Materials: Issues of Brittle Fracture,'' *Proc. 2nd Int. Conf. Intelligent Materials*, pp. 52–62.

[10] Kuo, C. M., and Barnett, D. M., 1991, ''Stress Singularities of Interfacial Cracks in Bonded Piezoelectric Half-Spaces,'' *Modern Theory of Anisotropic Elasticity and Applications*, J. J. Wu, T. C. T. Ting and D. M. Barnett, eds., SIAM, Philadelphai, pp. 33–50.

[11] Suo, Z., Kuo, C.-M., Barnett, D. M., and Willis, J. R., 1992, ''Fracture Mechanics for Piezoelectric Ceramics,'' J. Mech. Phys. Solids, **40**, pp. 739–765.

[12] Clements, D. L., 1971, ''A Crack Between Dissimilar Anisotropic Media,'' Int. J. Eng. Sci., **9**, pp. 257–265.

[13] Ru, C. Q., 1999, ''Exact Solution for Finite Electrode Layers Embedded at the Interface of Two Piezoelectric Half-Planes,'' J. Mech. Phys. Solids, **48**, pp. 693–708.

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

[15] Lothe, J., and Barnett, D. M., 1976, ''Integral Formalism for Surface Waves in Piezoelectric Crystals,'' J. Appl. Phys., **47**, pp. 1799–1807.

[16] Chung, M. Y., and Ting, T. C. T., 1996, ''Piezoelectric Solid with an Elliptic Inclusion or Hole,'' Int. J. Solids Struct., **33**, pp. 3343–3361.

[17] Muskhelishvili, I. N., 1963, *Some Basic Problems of the Mathematical Theory of Elasticity*, P. Noordhoff Ltd, Amsterdam, The Netherlands.

[18] Ru, C. Q., Mao, X., and Epstein, M., 1998, ''Electric-Field Induced Interfacial Cracking in Multilayer Electrostrictive Actuators,'' J. Mech. Phys. Solids **46**, pp. 1301–1318.

[19] Tobin, A. G., and Pak, Y. E., 1993, ''Effect of Electric Field on Fracture Behavior of PZT Ceramics,'' Proc. SPIE, No. 1916, pp. 78–86.

[20] Park, S., and Sun, C. T., 1996, ''Fracture Criteria for Piezoelectric Ceramics,'' J. Am. Ceram. Soc., **78**, pp. 1475–1480.

[21] Ru, C. Q., 1999, ''Effect of Electrical Polarization Saturation on Stress Intensity Factors in a Piezoelectric Ceramic,'' Int. J. Solids Struct., **36**, pp. 869–883.

[22] Sosa, H., 1992, ''On the Fracture Mechanics of Piezoelectric Solids,'' Int. J. Solids Struct., **29**, pp. 2613–2622.

[23] Pak, Y. E., 1992, ''Linear Electro-Elastic Fracture Mechanics of Piezoelectric Materials,'' Int. J. Fract., **54**, pp. 79–100.

[24] Ru, C. Q., 1999, ''Electric-Field Induced Crack Closure in Linear Piezoelectric Media,'' Acta Mater., **47**, pp. 4683–4693.

Surface Waves in Coated Anisotropic Medium Loaded With Viscous Liquid

T.-T. Wu[1]

T. Y. Wu

Institute of Applied Mechanics,
College of Engineering
National Taiwan University,
Taipei, Taiwan, R.O.C.

The development of micro-acoustic wave sensor in biosensing created the need for further investigations of the surface wave propagation in a viscous liquid loaded layered medium. In this paper, we employed the sextic formalism of surface waves to study the viscous effect on the dispersion and attenuation characteristics of surface waves in a viscous liquid loaded layered medium. The dispersion relation for the viscous liquid loaded single-layered anisotropic half-space is given. Numerical examples of the Rayleigh wave and Love wave dispersion for the cases of a Cu/Fe layered half-space (isotropic) and of a SiO_2/Si layered half-space (anisotropic) loaded with viscous liquid are calculated and discussed. [S0021-8936(00)01902-4]

1 Introduction

Surface waves have been applied successfully in many of the technological fields, such as NDE of materials, seismological exploration, and SAW devices in electronic industry. Theoretical analyses of the propagation of surface acoustic waves in layered media have been reported in the literature. A review of the early analyses on the dispersion of surface waves in an isotropic layered medium can be found in the book by Aki and Richards [1]. In the last decade, the applications of acoustic microscopy and fiber-reinforced composites have initiated the interest in studying the wave propagation in layered isotropic or anisotropic media ([2–4]). Experimental and inverse analyses of surface waves in an anisotropic medium or layered medium have also been reported ([5,6]). On the other hand, the development of the micro-acoustic wave sensor in biosensing created the need for further investigations of the surface wave propagation in fluid loaded layered medium. A detailed experimental study of a Love wave sensor for biochemical sensing in liquids was given by Kovacs et al. [7]. They showed that, for small viscosity, the interaction of an acoustic Love wave with a viscous liquid can be described by a Newtonian liquid model.

In the field of nondestructive evaluation using elastic waves, several investigations on the viscosity-induced attenuation have also been reported in recent years. On neglecting the heat conduction effect, Wu and Zhu [8] proposed an approach for studying attenuated leaky Rayleigh waves due to viscous damping. In a subsequent paper, Zhu and Wu [9] employed the same approach to study Lamb wave propagation in a plate bordered with a viscous fluid layer. Recently, Nagy and Nayfeh [10] investigated the viscosity-induced attenuation of longitudinal guided waves in rods loaded with a fluid layer. In a subsequent paper, on including the viscous effect on the longitudinal wave in a fluid, Nayfeh and Nagy [11] derived a formal solution and examined the effects of fluid viscosity on the Lamb wave as well as leaky the Rayleigh wave. The viscous liquid loaded substrates (layered half-space or rod) of the above-mentioned investigations are assumed isotropic. Further, the investigations presented so far are either for a Lamb wave (and Rayleigh wave) or Love wave propagation.

In this paper, based on the sextic formalism ([12,13]), we present an approach that is suitable for studying both the Rayleigh and Love wave propagation in a viscous liquid loaded anisotropic layered half-space. To demonstrate the utilization of this approach, dispersion relations for the case of a single-layered anisotropic half-space loaded with a viscous liquid are presented.

2 Surface Waves in Viscous Liquid Loaded Anisotropic Layered Solids

In the conventional way of studying the propagation of surface waves in an isotropic layered half-space, the element number of the determinant, which results from satisfying the appropriate interface and boundary conditions, increases rapidly with the number of layers overlaying the half-space. For the case of an anisotropic layered half-space, an alternative way is the application of the sextic formalism with special treatment on the numerical stability. A stable sextic formalism for the solution of surface waves was given by Mal to study the anisotropic composite laminate under periodic surface loads ([12]). Later, in [13], another stable sextic formalism for the anisotropic surface wave solution based on the invariant imbedding technique was also given. In the sextic formalism, the equation of motion and the constitutive equation are combined and arranged to form a first-order matrix differential equation. The displacement and the traction acting across the planes normal to the layering surfaces are grouped into a six-dimensional vector. In each layer, the solution of the matrix ODE forms a transfer matrix that can be utilized to map the variables from one surface to the next layering surface. With this formulation, the size of the matrix encountered in the computation is independent of the number of layers. It is worth noting that to avoid the numerical instability in the calculation, special treatments must be taken. Details of the respective special treatments can be found in [12] and [13].

In the literature, Wu and Zhu [8] utilized the Lamb's viscous liquid model to solve the solid-viscous liquid interaction problems. In a later paper, Nayfeh and Nagy [11] pointed out that the model utilized by Wu and Zhu for a viscous liquid has the deficiency of incorporating the attenuation of a longitudinal wave. They suggested possible models to improve the above deficiency, one is modeling the viscous liquid (with the viscous coefficient denoted as μ_L) as a hypothetical solid whose shear rigidity equals $i\omega\mu_L$. The other one is the use of the so-called Stokes model which split the viscosity parameter between C_{11} and C_{13}. Accord-

[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, September 30, 1998; final revision, December 7, 1999. Associate Technical Editor: A. K. Mal. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

ing to their conclusion and our calculations, we note that the differences induced by adopting the suggested two different models are very small for a low-frequency range.

In the Stokes model for isotropic viscous liquids, we set $C_{11}=\kappa+(4/3)i\omega\mu_L$ and $C_{13}=\kappa-(2/3)i\omega\mu_L$. $\kappa=\rho_L c_L^2$ is the bulk modulus of the liquid, where ρ_L and c_L are the density and longitudinal wave velocity of the viscous liquid. Similar to the wave propagation in an isotropic elastic solid, the waves in an isotropic viscous liquid can be divided into the in-plane (x-z plane with displacements u, w) and the antiplane (with displacement v) motion. For the in-plane motion, the scalar and vector displacement potential of the viscous liquid φ, ψ satisfy the Helmholtz equation

$$\nabla^2\varphi+k_L^2\varphi=0 \tag{1}$$

with $k_L^2=\omega^2\rho_L/(\kappa+(4/3)i\omega\mu_L)$, and the diffusion equation

$$\frac{\partial\psi}{\partial t}-\left(\frac{\mu_L}{\rho_L}\right)\nabla^2\psi=0 \tag{2}$$

where $\nabla^2=(\partial^2/\partial x^2+\partial^2/\partial z^2)$.

We note that if the ratio $\omega\mu_L/\kappa$ is very small, then k_L^2 in Eq. (1) can be approximated as $k_L^2=\omega^2/c_L^2$. This approximation leads to a solution similar to the result of Wu and Zhu [8].

For the case of antiplane motion, the material is only subjected to the shear deformation, and therefore, the deformed volume remains unchanged. From the Navier-Stokes equation, the antiplane displacement v satisfies the diffusion equation as

$$\frac{\partial v}{\partial t}-\left(\frac{\mu_L}{\rho_L}\right)\nabla^2 v=0. \tag{3}$$

For a harmonic plane progressive wave propagating in the x-z plane, the scalar potential φ, the vector potential ψ, and the antiplane displacement v can be assumed in the form as

$$\varphi=(a_1e^{-ik_{z1}z}+a_4e^{-ik_{z4}z})e^{i(\omega t-k_x x)} \tag{4}$$

$$\psi=(a_3,e^{-ik_{z3}z}+a_6e^{-ik_{z6}z})e^{i(\omega t-k_x x)} \tag{5}$$

$$v=(a_2e^{-ik_{z2}z}+a_5e^{-ik_{z5}z})e^{i(\omega t-k_x x)} \tag{6}$$

where a_1,a_2,a_3,a_4,a_5,a_6 are unknown constants and k_{z1}, k_{z2}, and k_{z3} are the wave numbers of the up-going waves along the positive z-direction, while k_{z4}, k_{z5}, and k_{z6} are those for the down-going waves. The relations between k_{z1}, k_{z2}, k_{z3}, k_{z4}, k_{z5}, k_{z6}, and k_x can be obtained by substituting Eqs. (4)–(6) into Eqs. (1)–(3).

Similar to the derivation of the sextic formalism ([13]), the relationships between the traction $\hat{\mathbf{t}}_\alpha$ and the velocity $\hat{\mathbf{v}}_\alpha$ of a viscous liquid for the up-going wave ($\alpha=1$) and down-going wave ($\alpha=2$) can be obtained as

$$\hat{\mathbf{t}}_\alpha(z)=\mathbf{Z}_{\alpha L}\hat{\mathbf{v}}_\alpha(z)\quad \alpha=1,2 \tag{7}$$

where $\mathbf{Z}_{1L},\mathbf{Z}_{2L}$ are the local impedance of the up-going wave and down-going waves, respectively, and are defined as

$$\mathbf{Z}_{\alpha L}=\mathbf{L}_{\alpha L}\mathbf{A}_{\alpha L}^{-1}\quad \alpha=1,2 \tag{8}$$

where

$$\mathbf{A}_{1L}=\begin{bmatrix}\omega k_x & 0 & -\omega k_{z3}\\ 0 & i\omega & 0\\ \omega k_{z1} & 0 & \omega k_x\end{bmatrix},\quad \mathbf{A}_{2L}=\begin{bmatrix}\omega k_x & 0 & \omega k_{z3}\\ 0 & i\omega & 0\\ -\omega k_{z1} & 0 & \omega k_x\end{bmatrix} \tag{9}$$

$$\mathbf{L}_{1L}=\begin{bmatrix}-2i\omega\mu_L k_x k_{z1} & 0 & i\omega\mu_L(k_{z3}^2-k_x^2)\\ 0 & \omega\mu_L k_{z2} & 0\\ -\kappa(k_{z1}^2+k_x^2)+\frac{2}{3}i\omega\mu_L k_x^2-\frac{4}{3}i\omega\mu_L k_{z1}^2 & 0 & -2i\omega\mu_L k_x k_{z3}\end{bmatrix} \tag{10}$$

$$\mathbf{L}_{2L}=\begin{bmatrix}2i\omega\mu_L k_x k_{z1} & 0 & i\omega\mu_L(k_{z3}^2-k_x^2)\\ 0 & -\omega\mu_L k_{z2} & 0\\ -\kappa(k_{z1}^2+k_x^2)+\frac{2}{3}i\omega\mu_L k_x^2-\frac{4}{3}i\omega\mu_L k_{z1}^2 & 0 & 2i\omega\mu_L k_x k_{z3}\end{bmatrix}. \tag{11}$$

3 Dispersion Relation

Consider an anisotropic layered half-space with elastic properties varied only along the z-axis. If, in each layer (e.g., medium B) the material is assumed to be homogeneous, then the total traction $\hat{\mathbf{t}}(z)$ in medium B can be expressed as ([13])

$$\hat{\mathbf{t}}(z)=\mathbf{G}_B(z)[i\omega\hat{\mathbf{u}}(z)] \tag{12}$$

where ω is the circular frequency and $\mathbf{G}_B(z)$ is the global impedance which relates the velocity field $i\omega\hat{\mathbf{u}}(z)$ to the traction field. The expression of $\mathbf{G}_B(z)$ is given in the Appendix for convenience. For a single-layered anisotropic half-space (Fig. 1) with a traction-free surface, the traction at $z=h$ vanishes, then from Eq. (12) and with the existence of a nontrivial solution we have the dispersion equation, which relates k_x to the circular frequency of the plane wave ω as

$$\det|\mathbf{G}_B|=0. \tag{13}$$

For the case of a viscous liquid half-space on top of a single-layered half-space, from Eq. (12), the traction at the solid-liquid interface can be written in terms of the particle velocity $\hat{\mathbf{v}}(h^-)$ as

$$\hat{\mathbf{t}}(h^-)=\mathbf{G}_B\hat{\mathbf{v}}(h^-). \tag{14}$$

On the other hand, in the viscous liquid half-space, there is no existing down-going wave; therefore, the global impedance is equivalent to the local impedance of the up-going wave $\mathbf{Z}_{1L}$, which is given in Eq. (8). The traction $\hat{\mathbf{t}}(h^+)$ at the interface is then written as

$$\hat{\mathbf{t}}(h^+)=\mathbf{Z}_{1L}\hat{\mathbf{v}}(h^+). \tag{15}$$

From the continuity conditions of the traction and the particle velocity at the solid-viscous liquid interface, and Eqs. (14), (15), we find that

$$(\mathbf{G}_B-\mathbf{Z}_{1L})\hat{\mathbf{v}}(h)=0. \tag{16}$$

For the existence of a nontrivial solution of the particle velocity at the interface, the following condition must be satisfied, i.e.,

$$\det|\mathbf{G}_B-\mathbf{Z}_{1L}|=0. \tag{17}$$

Equation (17) is the dispersion relation for a viscous liquid loaded single-layered anisotropic half-space. The relative magnitude of the interface velocity vector $\hat{\mathbf{v}}(h)$ can be obtained by sub-

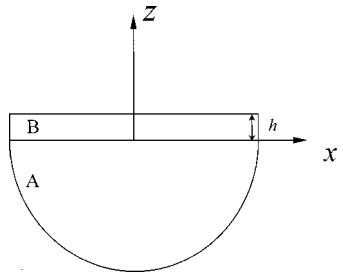


Fig. 1 The coordinates of the single-layered half-space

stituting k_x and ω (which satisfy Eq. (17)) into Eq. (16). Once the interface velocity vector is obtained, the traction vector at the interface $\hat{\mathbf{t}}(h)$ can be obtained from Eq. (14) or (15). The particle velocity and the traction vectors at any position in the viscous liquid half-space, interface layer, and the solid half-space can then be evaluated, in a straightforward way, from the known vectors $\hat{\mathbf{v}}(h), \hat{\mathbf{t}}(h)$.

4 Numerical Calculations

In the following numerical simulations, both the Rayleigh wave and Love wave in a single-layered isotropic as well as anisotropic half-spaces loaded with viscous liquid are considered. In particular, the numerical result for the Love wave in a viscous liquid loaded single-layered half-space was simulated and compared with those in [7].

4.1 Isotropic Single-Layered Half-Space Loaded With Viscous Liquid. In [7], the propagation of a Love wave in a viscous liquid loaded single-layered half-space was given. In their calculation, the substrate (ST-cut quartz) was approximated as an isotropic substrate with similar properties ($\rho = 2200\,\text{kg/m}^3$, $\mu = 1.74\times 10^{10}\,\text{N/m}^2$). The properties of the surface SiO_2 layer ($h = 1.46\,\mu\text{m}$) were $\rho = 2650\,\text{kg/m}^3$, $\mu = 6.6\times 10^{10}\,\text{N/m}^2$. The operating frequency of the Love wave sensor assumed was $f = 123.5\,\text{MHz}$. Figure 2 shows the frequency shift ($|\Delta f|/f$ in percentage) and the attenuation ($|\text{Im}\,k_x|/\text{Re}\,k_x$ in percentage) as functions of $\sqrt{\omega\eta\rho}$, which was calculated based on the current formulation. The results shown in Fig. 2 are exactly the same as those shown in Fig. 5 of [7].

In the following calculations, the case of an isotropic Fe half-space with Cu surface layer (20 μm in thickness) is considered. The material properties of polycrystalline Cu and Fe are

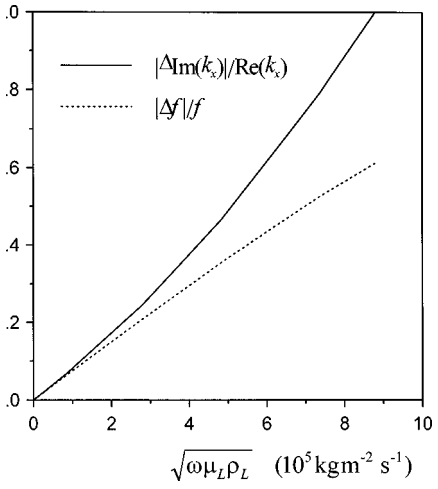


Fig. 2 The frequency shift and attenuation of Love wave in a viscous liquid loaded single-layered half-space (isotropic ST-cut quartz)

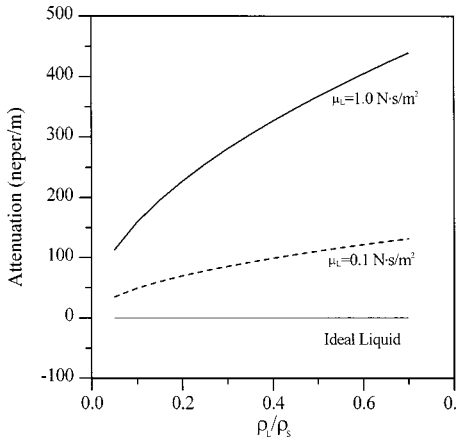


Fig. 3 The Love wave attenuation as a function of the density ratio ρ_L/ρ_s in a liquid loaded Cu-Fe layered half-space (isotropic), the frequency is equal to 20 MHz

$$\text{Cu:}\quad \rho = 8500\,\text{kg/m}^3,\quad \lambda = 11.2\times 10^{10}\,\text{N/m}^2,$$
$$\mu = 4.39\times 10^{10}\,\text{N/m}^2.$$
$$\text{Fe:}\quad \rho = 7870\,\text{kg/m}^3,\quad \lambda = 11.3\times 10^{10}\,\text{N/m}^2,$$
$$\mu = 8.2\times 10^{10}\,\text{N/m}^2.$$

To understand the influence of the liquid density on the attenuation of the fundamental Love wave, two frequencies (20 MHz and 300 MHz) were chosen for the following simulations (Figs. 3 and 4). In the figures, ρ_S is the density of the substrate (Fe). The dotted line represents an ideal liquid loaded on the Cu-Fe layered half-space, and the solid and bold solid lines represent the cases of moderate viscous liquid loading ($\mu_L = 0.1\,\text{N.s/m}^2$) and highly viscous liquid loading ($\mu_L = 1\,\text{N.s/m}^2$), respectively. The result shows that the bigger the ρ_L/ρ_S ratio, the higher the attenuation of the Love waves. The attenuation of the Love wave is dependent on the magnitude of the viscosity as well as the frequency. Results also showed that there is more than a one order difference between the Love wave attenuation for operating on 20 MHz and 300 MHz.

On the phase velocity dispersion of the fundamental Rayleigh surface wave mode, numerical results showed that the increase of viscosity results in a slight decrease of the Rayleigh wave velocity. For example, at a frequency equal to 400 MHz, the Rayleigh

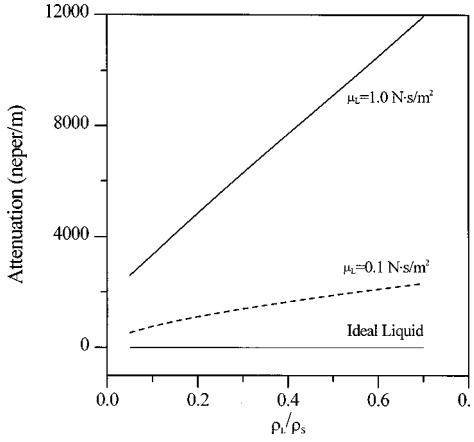


Fig. 4 The Love wave attenuation as a function of the density ratio ρ_L/ρ_S in a liquid loaded Cu-Fe layered half-space (isotropic), the frequency is equal to 300 MHz

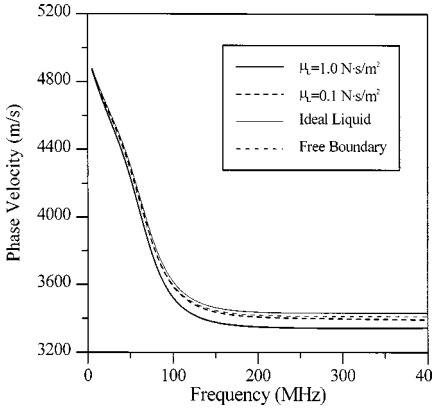


Fig. 5 The phase velocity dispersion of the Rayleigh surface wave in a liquid loaded SiO_2−Si layered half-space (anisotropic). The Rayleigh wave is propagating on the [001] surface and along the direction with 15 deg away from [100] axis.

wave velocity of the ideal liquid case is 2129.8 m/s, while that of the highly viscous liquid case is 2113.8 m/s. For the case of Love wave velocity, we note that the increase of viscosity results in a negligible change in wave velocity.

4.2 Anisotropic Single-Layered Half-Space Loaded With Viscous Liquid. In this subsection, we consider propagation of surface waves in an anisotropic single-layered half space loaded with viscous liquid. The viscous liquid loadings considered are the moderate viscosity ($\mu_L = 0.1$ N.s/m^2) and high viscosity ($\mu_L = 1$ N.s/m^2) cases. The properties of the anisotropic silicon half-space and the surface layer (isotropic SiO_2, 20 μm) are given as SiO_2:

$$\rho = 2332 \text{ kg/m}^3, \quad C_{11} = 16.6\times 10^{10} \text{ N/m}^2,$$

$$C_{12} = 6.4\times 10^{10} \text{ N/m}^2, \quad C_{44} = 6.4\times 10^{10} \text{ N/m}^2$$

Si:

$$\rho = 2200 \text{ kg/m}^3, \quad C_{11} = 7.85\times 10^{10} \text{ N/m}^2,$$

$$C_{12} = 1.61\times 10^{10} \text{ N/m}^2, \quad C_{44} = 3.12\times 10^{10} \text{ N/m}^2.$$

Figure 5 shows the calculated results for the fundamental Rayleigh wave propagating on the [001] surface along the direction with 15 deg away from [100] axis. From the figure, as compared with the case of free single-layered half-space (dotted line), one finds that at a fixed frequency, an ideal liquid loading results in a slight increase of the Rayleigh wave velocity. Similar to that of an isotropic single-layered half-space, the Rayleigh wave velocity

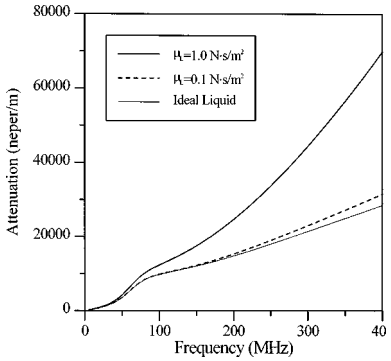


Fig. 6 The attenuation of the Rayleigh surface wave in a liquid loaded SiO_2−Si layered half-space (anisotropic). The Rayleigh wave is propagating on the [001] surface and along the direction with 15 deg away from [100] axis.

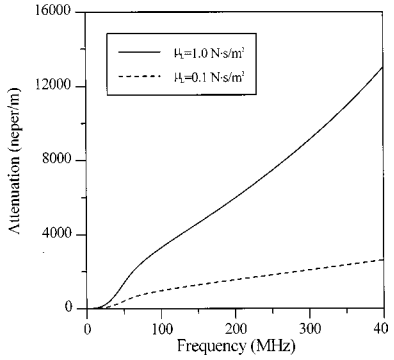


Fig. 7 The attenuation of the Love wave in a liquid loaded SiO_2−Si layered half-space (anisotropic). The Love wave is propagating on the [001] surface and along the direction with 15 deg away from [100] axis.

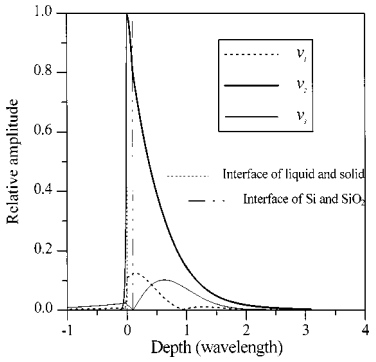


Fig. 8 The distribution of the particle velocity components for the Love wave propagating along the direction with 15 deg away from [100] axis (f=25 MHz and μ_L=1 N.s/m^2)

decreases as the frequency increases. For frequency high enough, the dispersion of the phase velocities approach that of a liquid loaded SiO_2 half-space. Figure 6 presents the corresponding attenuation of the cases shown in Fig. 5. We note that the big attenuation of the ideal liquid case is due to the leak of energy into the liquid half-space.

Figure 7 shows the calculated results for the attenuation of the fundamental Love wave propagating along the direction 15 deg away from the [100] axis. It is noted that there is no attenuation due to an ideal liquid loading. On examining the magnitudes of the attenuation of the Rayleigh wave (Fig. 6) and Love wave (Fig. 7), we found that the Love wave attenuation is much smaller than that of the Rayleigh wave.

Figure 8 shows the distribution of the particle velocity components along the depth. The calculated results are for the fundamental Love wave propagating along the direction with 15 deg away from the [100] axis and with the frequency equal to 25 MHz and viscosity equal to $\mu_L = 1$ N.s/m^2. Due to the anisotropy of the substrate, we note that in-plane particle velocity components (v_1, v_3) are not vanishing. However, the amplitude of the anti-plane velocity component v_2 is much larger than that of the in-plane components. In addition, most of the energy of the Love wave is confined around the surface SiO_2 layer (about 0.1 to 0.2 wavelengths).

5 Concluding Remarks

In this paper we have employed the sextic formalism of surface waves to study the viscous effect on the dispersion relation of surface wave propagation in a viscous liquid loaded layered medium. The dispersion relations for a viscous liquid loaded single-layered anisotropic half-space are given. Numerical examples for

both the cases of isotropic and anisotropic substrates were calculated. The results showed that, for the isotropic case, the calculated results are in agreement with those of the existing reference. Due to the leaky feature of the Rayleigh wave in a liquid jointed half-space, the attenuation of the Rayleigh wave is much bigger than that of the Love wave. The attenuation of the propagation of the Love wave is solely due to the viscous effect of the loaded viscous liquid. Finally, we note that the current formulation can be utilized to study the Rayleigh wave and the Love wave propagation in a viscous liquid loaded anisotropic layered structure.

Acknowledgment

The authors thank the financial support of this research from the National Science Council of ROC through the grant NSC87-2212-E-002-029 and 88-2218-E-002-020. The authors thank Prof. A. K. Mal for his helpful comments in the completion of this paper.

Appendix

$$\mathbf{G}_B(z)=[\mathbf{Z}_1\mathbf{W}_1(z)\mathbf{R}_{BA}\mathbf{W}_2^{-1}(z)+\mathbf{Z}_2]\times[\mathbf{W}_1(z)\mathbf{R}_{BA}\mathbf{W}_2^{-1}(z)+\mathbf{I}]^{-1}$$

$$\mathbf{Z}_\alpha=-\frac{1}{\omega}\mathbf{L}_\alpha\mathbf{A}_\alpha^{-1},\quad \mathbf{W}_\alpha(z)=\mathbf{A}_\alpha\mathbf{\Phi}_\alpha(z)\mathbf{A}_\alpha^{-1},\quad \alpha=1,2$$

$$\mathbf{\Phi}_1(z)=\mathrm{diag}(e^{-ik_{z1}z},e^{-ik_{z2}z},e^{-ik_{z3}z})$$

$$\mathbf{\Phi}_2(z)=\mathrm{diag}(e^{-ik_{z4}z},e^{-ik_{z5}z},e^{-ik_{z6}z})$$

$$\mathbf{R}_{BA}=[\mathbf{Z}_1-\mathbf{G}_A]^{-1}[\mathbf{G}_A-\mathbf{Z}_2]$$

where $\mathbf{G}_A$ is the global impedance of the adjacent medium A. $\mathbf{L}_\alpha$, $\mathbf{A}_\alpha$ are 3×3 matrices, which are generated from the six eigenvectors.

References

[1] Aki, K., and Richards, P. G., 1980, *Quantitative Seismology: Theory and Methods*, Vol. 1, W. H. Freeman, San Francisco.
[2] Kundu, T., and Mal, A. K., 1986, ''Acoustic Material Signature of a Layered Plate,'' Int. J. Eng. Sci., **24**, pp. 1819–1829.
[3] Nayfeh, H., and Taylor, T. W., 1988, ''Surface Wave Characteristics of Fluid-Loaded Multilayered Media,'' J. Acoust. Soc. Am., **84**, No 6, pp. 2187–2191.
[4] Bouden, M., and Datta, S. K., 1990, ''Rayleigh and Love Waves in Cladded Anisotropic Medium,'' ASME J. Appl. Mech., **57**, pp. 398–403.
[5] Chai, J.-F., and Wu, T.-T., 1994, ''Determinations of Anisotropic Elastic Constants Using Laser Generated Surface Waves,'' J. Acoust. Soc. Am., **95**, No 6, pp. 3232–3241.
[6] Wu, T.-T., and Liu, Y.-H., 1999, ''Inverse Analyses of Thickness and Elastic Properties of a Bonding Layer Using Laser Generated Surface Waves,'' Ultrasonics, **37**, pp. 23–30.
[7] Kovacs, G., Vellekoop, M. J., Haueis, R., Lubking, G. W., and Venema, A., 1994, ''A Love Wave Sensor for (Bio)chemical Sensing in Liquids,'' Sens. Actuators A, **43**, pp. 38–43.
[8] Wu, J., and Zhu, Z., 1995, ''An Alternative Approach for Solving Attenuated Leaky Rayleigh Waves,'' J. Acoust. Soc. Am., **97**, No 5, pp. 3191–3193.
[9] Zhu, Z., and Wu, J., 1995, ''The Propagation of Lamb Waves in a Plate Bordered With a Viscous Liquid,'' J. Acoust. Soc. Am., **98**, pp. 1057–1064.
[10] Nagy, P. B., and Nayfeh, A. H., 1996, ''Viscosity-Induced Attenuation of Longitudinal Guided Waves in Fluid-Loaded Rods,'' J. Acoust. Soc. Am., **100**, No 3, pp. 1501–1508.
[11] Nayfeh, A. H., and Nagy, P. B., 1996, ''Excess Attenuation of Leaky Lamb Waves due to Viscous Fluid Loading,'' J. Acoust. Soc. Am., **101**, No 5, pp. 2649–2658.
[12] Mal, A. K., 1988, ''Wave Propagation in Layered Composite Laminates Under Periodic Surface Loads,'' Wave Motion, **10**, pp. 257–266.
[13] Braga, M. B., 1990, ''Wave Propagation in Anisotropic Layered Composites,'' Ph.D. dissertation, Stanford University, Stanford, CA.

Study of Frictional Impact Using a Nonsmooth Equations Solver

L. Johansson

A. Klarbring

Division of Mechanics,
Department of Mechanical Engineering,
Linköping University,
S581 83 Linköping, Sweden

In this paper a mathematical formulation and a numerical algorithm for the analysis of impact of rigid bodies against rigid obstacles are developed. The paper concentrates on three-dimensional motion using a direct approach where the impenetrability condition and Coulomb's law of friction are formulated as equations, which are not differentiable in the usual sense, and solved together with the equations of motion and necessary kinematical relations using Newton's method. An experiment has also been performed and compared with predictions of the algorithm, with favorable results.
[S0021-8936(00)01402-1]

1 Introduction

This paper is concerned with a method for the analysis of impact of rigid bodies against rigid obstacles. Some problems of this kind can be treated by specifying the quotient between the relative normal velocity of approach and separation ([1]); i.e., by introducing the classical coefficient of restitution. This is sometimes generalized to what is known as Poisson's hypothesis (see [2]). In many cases, however, such as in the experiment described in Section 3 below, it is necessary to take both the normal and the tangential impulse at the impact into account to achieve reasonable agreement with observations. This class of problems has been the object of several recent studies, invariably leading to the introduction of one or more additional constitutive parameters, such as the coefficient of friction. Thus, in Brach [3], a quotient between normal and tangential impulses are introduced and several bounds, based on physical assumptions, are derived for this quotient. In Stronge [4], the division of the impact process into a compression and an expansion phase is analyzed, and the problem is treated using a coefficient of restitution relating energies rather than velocities. Walton [5] suggests a model involving three parameters which for the special case of spheres is equivalent to the model used in the present paper.

The approach in the references cited above, as well as in the present paper, is to use rigid-body dynamics combined with point contact laws. An important question is, of course, how accurately the impact behavior of a physical body can be modeled using a theory based on rigid-body motion and point contact laws with a few constitutive parameters. Recent work addressing this issue include Stoianovici and Hurmuzlu [6], where a slender bar is dropped onto a massive surface. The (classical) coefficient of restitution is found to depend strongly on the orientation of the bar, something which is ascribed to the onset of vibrations in the bar, and that the impacts are actually divided into a series of microimpacts. In Calsamiglia et al. [7] the coefficient of friction for disks impacting a massive surface is found to depend on the inclination angle. This is attributed to an elastic mechanism where nominally sliding contacts actually stops sliding and then resumes sliding in the same direction during impact.

In the present paper the impenetrability condition and Coulomb's law of friction are formulated in terms of velocities and impulses rather than displacements and forces. Since the velocities are not necessarily continuous, it is assumed that these laws apply to a linear combination of the left and right limits of the velocities, which are assumed to exist. These velocities are, as an approximation, replaced with the velocities at the beginning and end of a time step in the overall algorithm. This is the approach of Moreau, as described in Moreau [8] a paper which is much concerned with the mathematical setting of the problem, but where a very elegant algorithm for the case of completely inelastic impacts is also given. See also Moreau [9], where granular materials are simulated with an algorithm involving iteration between the contact laws and the equations of motion.

In a previous paper [10], the present problem, in the two-dimensional case, was solved using an algorithm based on a formulation as a linear complementarity problem (LCP). This approach seems to be less attractive for three-dimensional problems, and in this paper the problem is instead formulated in the form of a system of nonsmooth equations, to which Newton's method is applied directly. This is an adaption to the present class of problems of the method developed for elastostatic contact problems by Christensen et al. [11], which, in turn, is a development of the method of Alart and Curnier [12].

The predictions of the model used in this paper are also compared with an experiment performed for this purpose, and with an experiment discussed in more detail in an earlier paper, [10].

2 Governing Equations

In this section, relations governing the three-dimensional motion of a rigid body that comes into frictional contact with a rigid wall will be stated (Fig. 1). The rigid wall is assumed to be flat.

The contact conditions are most easily written in a coordinate frame attached to the wall, since the normal to the wall is constant in such a frame. On the other hand, an inconvenience is that the inertia tensor is then nonconstant. In the present work it was decided to write the rotational equation of motion with reference to a body-fixed frame where the inertia tensor is constant, and the translational equation of motion and contact conditions in a frame fixed to the rigid wall. This is one of several possible choices, but it seems unavoidable that the transformation between a frame attached to the wall and one attached to the body enters the formulation at some point.

We also note in this respect that the orientation of the body was described using quaternions ([13]), and these are most commonly, although not necessarily integrated from the angular velocity components in terms of a body attached frame. Thus, two coordinate systems are introduced, moving $x'y'z'$-coordinates fixed in the body and oriented along its principal axes of inertia and inertial xyz-coordinates with the y-axis normal and x and z-axes parallel to the rigid wall. The equations introduced below can be regarded as vector-matrix equations, where vectors are three-dimensional column vectors consisting of components of physical vectors relative to one of these coordinate frames while matrices are operators relating such vectors. The contact impulses $\mathbf{P}$, external impulses $\mathbf{F}$, the velocity $\dot{\mathbf{x}}$ of the contact point, the velocity $\dot{\mathbf{x}}_G$ of the center

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 13, 1999; final revision, Nov. 1, 1999. Associate Technical Editor: N. C. Perkins. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

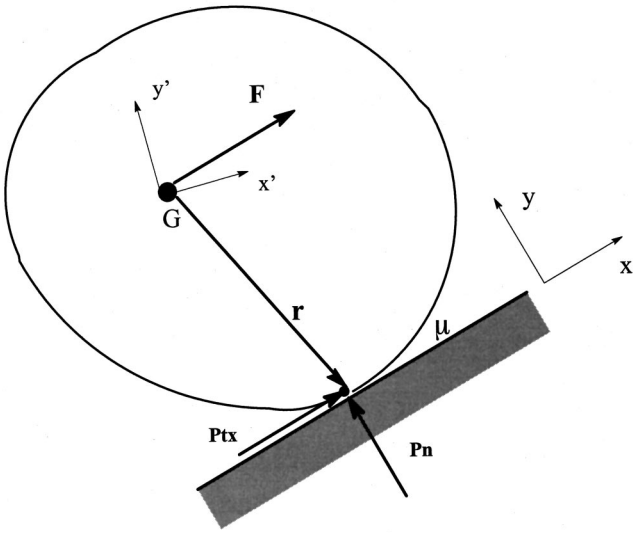


Fig. 1 Geometry of the contact

of mass and the position vector $\mathbf{r}$ and related matrix $\mathbf{R}$ expressing the location of the contact point, are all written in terms of their xyz-components, while the inertia tensor $\mathbf{I}'$ and the angular velocity $\boldsymbol{\omega}'$ are written in terms of their $x'y'z'$-components.

2.1 Equations of Motion and Kinematical Equations. The linear equations of motion are written in integrated form in terms of velocities and impulses and expressed in xyz-components as

$$\mathbf{m}\dot{\mathbf{x}}_G^2=\mathbf{m}\dot{\mathbf{x}}_G^1+\mathbf{P}+\mathbf{F}. \tag{1}$$

Here, $\mathbf{m}=m\mathbf{1}$, m being the mass of the body and $\mathbf{1}$ the identity matrix. $\dot{\mathbf{x}}_G^1$ and $\dot{\mathbf{x}}_G^2$ are the velocities of the center of mass at times t_1 and t_2, $t_1<t_2$, respectively. Here and elsewhere in this paper superscripts 1 and 2 denote the value of a quantity at times t_1 and t_2, respectively. $\mathbf{P}=(P_{tx},P_n,P_{tz})^t$ are the contact impulses during the time interval $[t_1,t_2]$. A superscript t indicates transpose of a vector or matrix. These contact impulses are obtained from integration of a vector valued measure $d\mathbf{P}$:

$$\mathbf{P}=\int_{[t_1,t_2]} d\mathbf{P}.$$

For smooth contact situations, $d\mathbf{P}=\overline{\mathbf{P}}dt$, where $\overline{\mathbf{P}}$ is a contact force vector, and $\int_{[t_1,t_2]}$ is the standard Lebesgue integral. In case of impact, when the velocities are discontinuous, $d\mathbf{P}$ is a singular measure of ''Dirac'' type, a so-called percussion. We refer to Moreau [9] for a detailed discussion of these issues. $\mathbf{F}$ are known impulses from external forces. These forces are assumed to act at the center of mass of the body and are therefore not present in the moment equation given below.

The moment equations are written in $x'y'z'$-components:

$$\mathbf{I}'\boldsymbol{\omega}'^2=\mathbf{I}'\boldsymbol{\omega}'^1-\int_{t_1}^{t_2}\mathbf{G}(\boldsymbol{\omega}')dt+\int_{[t_1,t_2]}\mathbf{AR}d\mathbf{P}. \tag{2}$$

Here, $\mathbf{I}'$ is the inertia tensor, which in the $x'y'z'$-system becomes a constant diagonal matrix with components $I_{x'}$, $I_{y'}$, and $I_{z'}$. $\mathbf{A}$ is the transformation matrix from xyz-coordinates to $x'y'z'$-coordinates such that $\mathbf{P}'=\mathbf{AP}$, where $\mathbf{P}'$ and $\mathbf{P}$ are the two representations of a geometrical vector. $\mathbf{R}$ is the antisymmetric matrix of components of $\mathbf{r}$ such that $\mathbf{RP}=\mathbf{r}\times\mathbf{P}$, where $\times$ is the ordinary vector product. The vector of gyroscopic terms can be written in $x'y'z'$-components as

$$\mathbf{G}(\boldsymbol{\omega}')=\begin{bmatrix}\omega_{y'}\omega_{z'}(I_{z'}-I_{y'})\\ \omega_{x'}\omega_{z'}(I_{x'}-I_{z'})\\ \omega_{x'}\omega_{y'}(I_{y'}-I_{x'})\end{bmatrix}. \tag{3}$$

The kinematic relations relating the velocities at the contact interface with the velocity at the center of mass are written as

$$\dot{\mathbf{x}}=\dot{\mathbf{x}}_G+(\mathbf{A}^t\boldsymbol{\omega}')\times\mathbf{r}=\dot{\mathbf{x}}_G+(\mathbf{AR})^t\boldsymbol{\omega}'. \tag{4}$$

2.2 Contact Conditions. In this section we assume that contact between the rigid body and the obstacle wall has been detected, i.e., the impenetrability condition $y\geqslant 0$ is fulfilled with equality, and conditions will be formulated to ensure that the body either stays in contact or breaks away, but does not penetrate. In the numerical implementation a small penetration is accepted without correction (see Section 2.4 below). The velocity of the rigid body at the contact interface is $\dot{\mathbf{x}}=(\dot{x},\dot{y},\dot{z})^t$ and expressed by (4). These velocities together with the contact impulses will be used to formulate contact-impact constitutive laws. Note that in the present situation it cannot be assumed that $\dot{\mathbf{x}}$ is continuous or even that it exists at all times. This is obvious for the case of an impact, but in fact other cases of discontinuous velocities can occur, as is shown by the Painlevé example (see Brogliato [14]). To resolve this difficulty it is assumed, following Moreau [9], that the right and left limits of $\dot{\mathbf{x}}$ exist, and that as a generalization of the Signorini contact condition from static or quasi-static situations to the impact case, we may state

$$\dot{y}^a\geqslant 0,\quad P_n\geqslant 0,\quad P_n\dot{y}^a=0, \tag{5}$$

where

$$\dot{y}^a=\frac{e_n}{1+e_n}\dot{y}^-(t_1)+\frac{1}{1+e_n}\dot{y}^+(t_2)$$

is an average velocity, e_n is a constitutive parameter discussed in Section 2.5, and $\dot{y}^-(t_1)$ and $\dot{y}^+(t_2)$ are the left limit of $\dot{y}$ at time t_1 and the right limit of $\dot{y}$ at time t_2, respectively. Note that (5) is stated for a time interval, but makes sense also if the time interval is shrunk to time a instant. Furthermore, it can be shown that for smooth motion, when $\dot{\mathbf{x}}$ is continuous, (5) simplifies to the standard Signorini conditions. A further rationale behind (5) is that if the first inequality is activated at all times when contact occurs, then impenetrability will be enforced ([9]). Thus, the contact law as stated here is applicable both to cases of impact and cases where the contact forces are smooth.

Relation (5) is a complementarity condition and this is a mathematical system that has been much studied in the area of mathematical programming. The number of numerical methods devised to deal with such and related systems is quite extensive, both in the mathematical and the engineering literature. Recently, however, it has become clear for the case of quasi-static frictional contact problems (Alart and Curnier [12] and Christensen et al. [11]) that a most effective method is based on writing the problem as a system of nondifferentiable equations and applying a Newton method specially devised for such equations ([15]). Here, we will use a similar approach for frictional impact problems. To that end, we need to rewrite (5) as an equation and, following [16], a first step is to state (5) as a variational inequality:

$$P_n\in K_n:\quad \dot{y}^a(P_n^*-P_n)\geqslant 0\quad \forall P_n^*\in K_n, \tag{6}$$

where

$$K_n=\{P_n|P_n\geqslant 0\}.$$

This variational inequality can be rewritten as

$$P_n\in K_n:\quad (P_n-(P_n-r_n\dot{y}^a))(P_n^*-P_n)\geqslant 0\quad \forall P_n^*\in K_n, \tag{7}$$

where $r_n>0$ is a parameter which we eventually will be able to adjust for best numerical performance; note, however, that (7) is equivalent to (5) for any positive value of r_n. It is well known

that (7) indicates a projection onto the set K_n, and due to the simplicity of this set we end up with a most useful representation of (5):

$$P_n = \mathrm{Proj}[P_n - r_n \dot{y}^a, K_n] = (P_n - r_n \dot{y}^a)_+, \tag{8}$$

where $(x)_+ = \max(x,0)$. Thus, we are able to write the complementarity condition (5) as an equation. Even though this equation is not differentiable, it opens the way for the use of a Newton-type algorithm.

Next we need to write Coulomb's law of friction in a way such that discontinuous velocities and percussions are included. Coulomb's friction law implies firstly that the friction force belongs to a cone of forces, the friction coefficient being the measure of the pointedness of this cone, and secondly that when sliding occurs it is in the opposite direction of the force. These principles can, as is well known (see, e.g., Christensen et al. [11]), be put in the mathematical form of a variational inequality. By choosing to write this variational inequality in terms of an average velocity and impulses, instead of velocity and forces as in the classical smooth case, we obtain a formulation of friction which includes impact and which reduces to the classical Coulomb friction law for sliding contact without impact. We thus put down our law of friction as follows:

$$\mathbf{P}_t \in K_t(P_n): \quad (\dot{\mathbf{w}}^a)^t(\mathbf{P}_t^* - \mathbf{P}_t) \geqslant 0 \quad \forall \mathbf{P}_t^* \in K_t(P_n), \tag{9}$$

where $\mathbf{P}_t = (P_{tx}, P_{tz})^t$,

$$K_t(P_n) = \{\mathbf{P}_t \,|\, |\mathbf{P}_t| \leqslant \mu P_n\},$$

$$\dot{\mathbf{w}}^a = \frac{e_t}{1+e_t}\dot{\mathbf{w}}^-(t_1) + \frac{1}{1+e_t}\dot{\mathbf{w}}^+(t_2),$$

$\dot{\mathbf{w}}^+$ and $\dot{\mathbf{w}}^-$ are right and left-hand limits of $\dot{\mathbf{w}} = (\dot{x}, \dot{z})^t$ and e_t is a constitutive parameter discussed in Section 2.5. Similarly to (6) this variational inequality may be rewritten as a projection:

$$\mathbf{P}_t = \mathrm{Proj}[\mathbf{P}_t - r_t \dot{\mathbf{w}}^a, K_t(P_n)], \tag{10}$$

where $r_t > 0$ is a parameter, which can be adjusted to improve numerical performance. It is possible to conclude that (10) is a continuous equation which, however, is not differentiable everywhere.

2.3 Discretization and Formulation as a System of Nonlinear Equations. For the discretization, a sequence of times $[t_1, \ldots, t_j, t_{j+1}, \ldots, t_n]$ is introduced. Equations (1), (2), (4), (8), and (10) will be discretized to form a system of 12 nonlinear equations for 12 unknowns at time t_{j+1}, assuming that everything is known at time t_j. These equations can be solved for the unknown quantities at time t_{j+1}. The solution can then be advanced in time by repeating this process until the desired time interval is covered.

First, assuming that $\mathbf{F}$ or a suitable approximation of $\mathbf{F}$ is known, no further discretization is necessary in Eq. (1) and we have

$$\mathbf{f}_1 = \mathbf{m}\dot{\mathbf{x}}_G^{j+1} - \mathbf{m}\dot{\mathbf{x}}_G^{j} - \mathbf{P} - \mathbf{F} = \mathbf{0}. \tag{11}$$

To discretize Eq. (2) we put

$$\int_{t_j}^{t_{j+1}} \mathbf{G}(\boldsymbol{\omega}')\,dt = \Delta t[\xi \mathbf{G}(\boldsymbol{\omega}'^j) + (1-\xi)\mathbf{G}(\boldsymbol{\omega}'^{j+1})], \tag{12}$$

where ξ is a discretization parameter usually set to $\xi = 0.5$ and $\Delta t = t_{j+1} - t_j$. It is further assumed that $\mathbf{A} = \mathbf{A}(t)$ and $\mathbf{R} = \mathbf{R}(t)$ are constant and known: In practice these are computed from the position of the body at the beginning of the time step, i.e., $\mathbf{A}(t) = \mathbf{A}(t_j)$ and $\mathbf{R}(t) = \mathbf{R}(t_j)$. Equation (2) then gives

$$\mathbf{f}_2 = \mathbf{I}'\boldsymbol{\omega}'^{j+1} - \mathbf{I}'\boldsymbol{\omega}'^{j} + \xi\Delta t\mathbf{G}(\boldsymbol{\omega}'^{j}) + (1-\xi)\Delta t\mathbf{G}(\boldsymbol{\omega}'^{j+1}) - \mathbf{ARP} = \mathbf{0}. \tag{13}$$

It would perhaps be appropriate to discretize $\mathbf{A}$ and $\mathbf{R}$ in the same way as $\mathbf{G}$. However, $\mathbf{G}$ depends on the angular velocity, which cannot be assumed to be continuous, whereas $\mathbf{A}$ and $\mathbf{R}$ depend on position, which is assumed to be continuous. It was therefore deemed less important to include the value of the quantity at time t_{j+1} in the discretization of $\mathbf{A}$ and $\mathbf{R}$ than in the discretization of $\mathbf{G}$. Further, note that we aim at a system of equations with velocities, angular velocities, and contact impulses as unknowns, and $\mathbf{A}$ and $\mathbf{R}$ depend on position. Using a discretization involving $\mathbf{A}(t_{j+1})$ and $\mathbf{R}(t_{j+1})$ would require either adding the equations for the position to the system of equations instead of computing the position in a second stage, or else using an iterative procedure. Although this would be quite possible, it was decided to sacrifice this improvement of accuracy in the interest of simplicity.

To discretize (4) it is only necessary to assume that $\mathbf{A}$ and $\mathbf{R}$ are constant and known. Writing the equation at time t_{j+1} then gives

$$\mathbf{f}_3 = \dot{\mathbf{x}}^{j+1} - \dot{\mathbf{x}}_G^{j+1} - (\mathbf{AR})^t\boldsymbol{\omega}'^{j+1} = \mathbf{0}. \tag{14}$$

Finally, to discretize (8) and (10) it is assumed that the right and left limits of the velocities can be approximated with the velocities at the beginning and end of a time step of the discretization. Equations (8) and (10) then give:

$$f_4 = P_n - \left(P_n - r_n\left(\frac{e_n}{1+e_n}\dot{y}^j + \frac{1}{1+e_n}\dot{y}^{j+1}\right)\right)_+ = 0, \tag{15}$$

$$\mathbf{f}_5 = \mathbf{P}_t - \mathrm{Proj}\left[\mathbf{P}_t - r_t\left(\frac{e_t}{1+e_t}\dot{\boldsymbol{\omega}}^j + \frac{1}{1+e_t}\dot{\boldsymbol{\omega}}^{j+1}\right), K_t((P_n)_+)\right] = \mathbf{0}, \tag{16}$$

where $K_t(P_n)$ has been replaced with $K_t((P_n)_+)$, following Christensen et al. [11].

2.4 Solution of the Equations. Equations (11), (13), (14), (15), and (16) form a system of 12 nonlinear equations for the 12 unknowns $\dot{\mathbf{x}}^{j+1}$, $\dot{\mathbf{x}}_G^{j+1}$, $\boldsymbol{\omega}'^{j+1}$, and $\mathbf{P}$. Putting $\mathbf{f} = (\mathbf{f}_1^t, \mathbf{f}_2^t, \mathbf{f}_3^t, f_4, \mathbf{f}_5^t)^t$ the solution is found by a direct application of Newton's method to the equation $\mathbf{f} = \mathbf{0}$. This equation is not differentiable in the usual sense, but has the property of being B-differentiable, meaning essentially that directional derivatives exist at each point. This makes it possible to apply the Newton method deviced for such equations by Pang [15]. Here, a somewhat simplified version of this method was used where the necessary derivatives were computed by arbitrarily picking one such directional derivative at nondifferentiable points (see Christensen et al. [11]).

The necessary derivatives of $\mathbf{f}_1$, $\mathbf{f}_2$, and $\mathbf{f}_3$ are easily calculated; those of f_4 and $\mathbf{f}_5$ are considerably more complicated. However, these equations can be formally identified with the corresponding equations for the quasistatic case discussed in Christensen et al. [11] where the derivatives are given. The solution requires some tuning of the numerical parameters r_n and r_t. The line search suggested by Pang (see Christensen et al. [11]) was tested. It did not show any obvious advantage for the computations for the present paper, but might be necessary for cases with multiple simultaneous contact points.

The overall aim of the algorithm is to compute the position, angular orientation, velocity, and angular velocity at time t_{j+1}, assuming that these quantities are known at an earlier time t_j. To compute the velocities, it is first checked if the body is in contact, i.e., if the impenetrability condition $y \geqslant 0$ is fulfilled with equality, or violated by some small amount. If it is, the velocities at time t_{j+1} are computed by solving all the equations in Section 2.3 simultaneously. This calculation proceeds from the position the body had at the end of the previous time step, accepting without correction a small violation of the impenetrability condition and assuming that $\mathbf{A}$ and $\mathbf{R}$ are constant throughout the time step at their values at the beginning of the time step. If no contact condition is violated, only Eqs. (11) and (13) need to be solved and the terms involving $\mathbf{A}$ and $\mathbf{R}$ disappear.

When the velocities and angular velocities have been computed, the position and orientation of the body can be computed. For the orientation of the body quaternions were used (see Stevens and Lewis [11]).

If the motion of a system of bodies interconnected with springs and dashpots is sought, the applied loads cannot be assumed to be known at time t_{j+1}, since, in order to incorporate the forces from the connecting springs and dashpots, they must be assumed to depend on positions and velocities at time t_{j+1}. This difficulty can be resolved iteratively and if such an iteration is introduced it is also possible to use better approximations to **A** and **R** if so desired.

2.5 Physical Interpretation of the Parameters e_n and e_t. It should be noted (Moreau [9]), that the constitutive parameters e_n and e_t introduced above will determine the nature of the contact. Considering the case of impact there is a nonzero normal impulse, $P_n > 0$, and (5) implies

$$\dot{y}^+ = -e_n \dot{y}^- . \tag{17}$$

Thus, the parameter e_n can be interpreted as a restitution coefficient with $e_n = 0$ corresponding to purely plastic impact and $e_n = 1$ to purely elastic impact.

In the tangential direction there can be either stick or slip. If the coefficient of friction is sufficiently large that there is no slip, each component of $\mathbf{P}_t^* - \mathbf{P}_t$ can be either positive or negative, and (9) implies

$$\dot{\mathbf{w}}^+ = -e_t \dot{\mathbf{w}}^- , \tag{18}$$

and e_t can be interpreted as a tangential restitution coefficient. It should be noted, however, that the interpretation of e_t as a tangential restitution coefficient is possible only for large values of μ, since the tangential impulse is limited by Coulomb's law of friction, and in the sliding case, for a nonspherical body, there is a coupling between normal and tangential contact impulses which complicates things further. It should also be noted that the opposite direction of tangential velocity and tangential contact impulse as required by Coulomb's law is applied to the linear combination $e_t/(1+e_t)\dot{\mathbf{w}}^- + 1/(1+e_t)\dot{\mathbf{w}}^+$, which does not necessarily imply that the tangential contact impulse is opposite to $\dot{\mathbf{w}}^-$.

It should be noted that the parameters e_n and e_t follows as a consequence from using contact laws of the form (5) and (9). Since these equations should be valid also at a time instant where there is a discontinuity in velocity, the velocity components themselves cannot be used. When this issue is resolved by using averages of the left and right limits of the velocity components, the parameters e_n and e_t appear naturally.

With the introduction of these restitution coefficients, there is a total number of three parameters describing the impact process: μ, e_n, and e_t. This is somewhat problematic, since these parameters must be determined experimentally. In particular there is no obvious simple method for measuring e_t. This topic is discussed further in Section 3.

3 Comparison With Experiments

An experiment was performed where a body, consisting of two rubber balls (superballs) glued to a cylinder, was thrown so as to bounce from a wooden surface. The process was photographed in stroboscopic light and the positions measured from the photographs were compared to calculations with the above algorithm. One such photograph is shown in Fig 2. This is the photograph actually used for the comparisons below, and was selected from a series of similar photographs because, as far as could be judged from a 29 cm by 23 cm blowup, the motion occurs in a single plane. The motion occurs in a plane about 3 cm in front of the ruler also seen on the photograph.

In a previous paper (Johansson [10]) an experiment is reported where a single ball is thrown into the space between two surfaces. The ball bounces first on the lower surface, then on the upper and then a second time on the lower, before exiting in approximately the same direction as it was entering. The mechanism is that the ball starts to rotate rapidly at the first impact. This causes the relative tangential velocity at the second impact to be large enough so that the ball changes direction and returns again in the same general direction as it was thrown from. The process was photographed in stroboscopic light, to produce photographs of the same type as that shown in Fig. 2. The motion was compared to calculations performed with a different algorithm than in the present work, but the model is the same, so that the results for cases where both algorithms can be used are identical except for differences in numerical error. To perform the comparison the normal coefficient of restitution and the coefficient of friction were measured directly, but no direct method of measuring the tangential coefficient of restitution was available. However, adjusting e_t to obtain a good fit resulted in the good agreement between experiments and calculations shown in Fig. 3, which shows the computed and measured positions of the ball at different times. Here the crosses are experimental points, the filled circles are calculated points and the circles are numbered in order of increasing time.

Fig. 2 Photograph from the experiments

The present experiment is a companion experiment to the one in Johansson [10] outlined above, and was performed for two main reasons. First, since there is no obvious simple experiment to measure the tangential coefficient of restitution, it was desired to test if the value of e_t obtained by adjusting computations to the experimental findings in the single ball experiments could be used to predict the motion of a composite body as used in the present experiment, where the contact surface is on one of the balls used in the single ball experiment. Secondly, when using balls, the geometry results in a decoupling of the tangential and normal directions. This is because a normal contact impulse does not influence the rotation of a ball and therefore does not affect the tangential velocity, and a tangential contact impulse does not influence the normal contact velocity since the rotation of a ball does not influence the normal contact velocity. Thus, it was desired to test a case where this decoupling does not occur.

Next the experimental setup is described. Measured values are given with more significant figures than actually warranted by the precision of measurements, and the last digit should not be trusted. For the coefficient of friction an approximate interval is given, giving some indication of the precision of these measurements.

The body used consisted of two rubber balls, of the type available in toy stores glued to a paper cylinder, as seen in Fig. 2. The mass of the composite body was 0.02315 kg the moment of inertia about the axis where rotation occurs in the present experiment was $3.968 \cdot 10^{-6}$ kgm^2. The distance between the centers of the balls

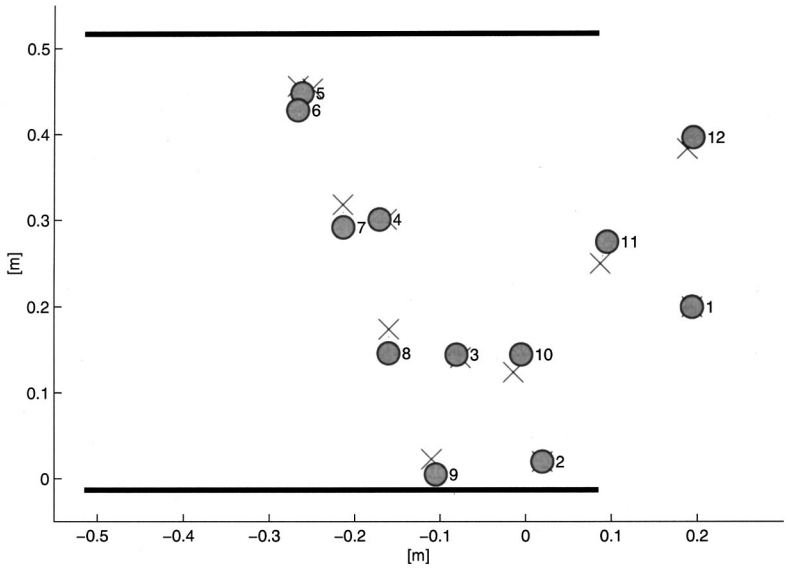


Fig. 3 Comparison with experiment and computations for single ball case

was 0.0914 m and the radius of the balls 0.0131 m. The body was thrown on a wooden table, with width 0.6000 m. The length scale was inferred from the length of the paper tube holding the balls together, which was 0.0702 m.

The flight of the body was photographed from a distance of about 0.9 meters with a Nikon SLR camera with a 35 mm lens set at aperture f/5.6, using Kodak TMAX 400 film. The body was thrown by hand after releasing the shutter, which was set to 1 s. This was repeated for a series of photographs and one where the motion of the body was in a single plane, as far as could be judged from a large scale blowup, was selected. The setup was illuminated by a Bruel & Kjaer strobe light. The frequency was set to 30 Hz, but a more accurate time interval of $\Delta t = 0.0313$ s between flashes was computed by analyzing a photograph of a ball in free fall taken for this purpose.

The normal coefficient of restitution was determined by measuring the rebound from a 1 m drop of a single ball and was found to be $e_n = 0.91$. The coefficient of friction was measured by pulling four balls glued to an aluminum weight along the surfaces. These measurements were made independently by three different persons, not involved in the computations or in other parts of the experiment, in order to avoid any bias. It was found that the dynamic coefficient of friction was $\mu = 1.13 \pm 0.2$. There was an appreciable difference between static and dynamic friction, the static value being about 1.6. It turned out in the calculations that the contact was well into the stick region so the exact value of μ does not actually matter, and the value $\mu = 1.13$ was used.

To compare computations with experiment, initial conditions were obtained by measuring the first and third positions of the body in Fig. 2, which are before the bounce, knowing the time in

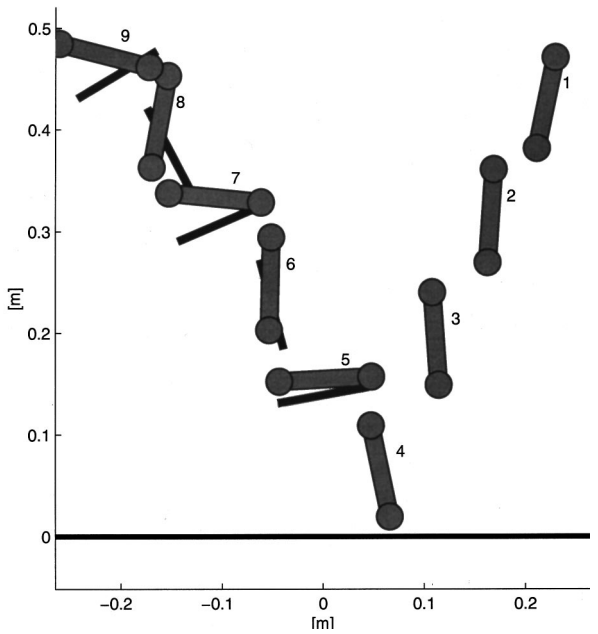


Fig. 4 Comparison between experiments and computations using $e_t = 0.68$ as obtained from single ball experiment

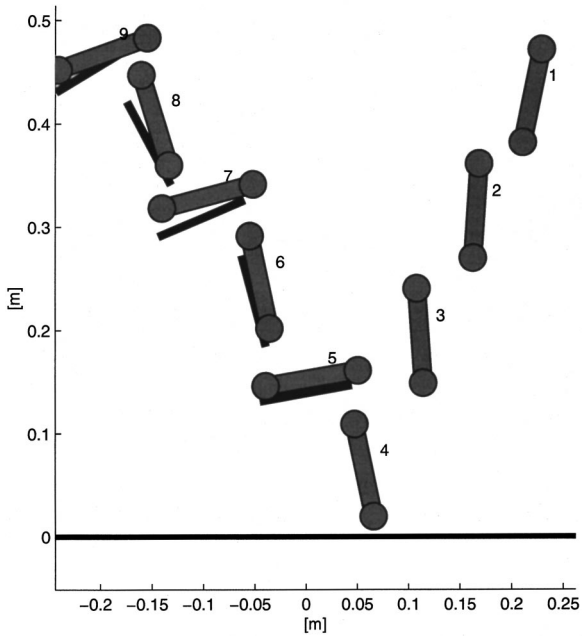


Fig. 5 Comparison between experiments and computations using $e_t = e_n = 0.91$

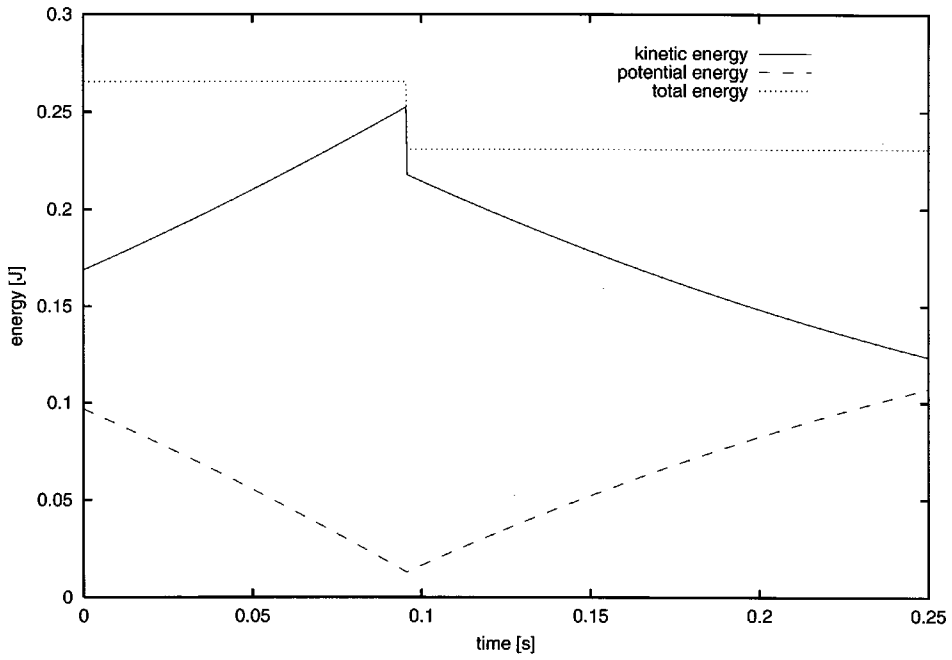


Fig. 6 Energy balance corresponding to Fig. 4

between. The time interval used in the computations was taken as 0.000313 s, which is one hundredth of the time interval of the strobe light.

Figures 4 and 5 show a comparison of the measured and computed positions of the body at different times. Here, the figures mimicking the actual appearance of the body shows the computed positions, whereas the solid black rectangle shows the experimental positions. The positions are numbered in order of increasing time. Figure 4 shows the computations using $e_t=0.68$ as obtained from the single ball experiment in Johansson [10] as discussed above, whereas in the computations for Fig. 5 the tangential coefficient was set equal to the normal coefficient of restitution (as obtained from dropping a single ball), $e_n=e_t=0.91$.

It is interesting to note that both computations give a reasonable agreement with experiment, but that the computation with $e_n=e_t$ is better. It is also interesting that the computed tangential contact impulses are actually in opposite directions in the two computations. In Fig. 4 the contact impulse is in the same direction as the gross motion of the body, whereas in Fig. 5 it is in the opposite direction, as expected intuitively. The mechanism is as follows: The normal contact impulse will cause the body to rotate, giving the contact point a velocity opposite to that of the center of mass. With a high coefficient of friction the contact is in the stick region, and the tangential coefficient of restitution places a limit on the allowed post-impact tangential velocity of the contact point. Thus, if e_t is small, a tangential impulse in the direction of the pre-impact motion, opposite to what is intuitively expected, is necessary to keep the post-impact tangential velocity from exceeding the value given by (18). Actually, this mechanism can be exploited to construct examples where energy is gained at the impact; indeed, this will happen in the present case if we change the coefficients of restitution to $e_n=1$ and $e_t=0$. It has, however, been shown by Moreau [9] that this will never occur if $e_n=e_t$. The energy balance corresponding to the motion in Fig. 4 is shown in Fig. 6 and it is seen that in this case the body looses energy at impact as expected.

In conclusion, it seems that putting $e_n=e_t$ is the prudent choice to make if no value from a situation very close to that which one wants to model is at hand. This also seems natural if the use of a linear combination of the left and right limits of the velocities in Eqs. (5) and (9) is interpreted as specifying an intermediate time during the contact process at which the contact laws are applied.

4 A Numerical Example

Since the section on experimental verification above only shows calculations with motion in a plane and with stick conditions prevailing at the contact, an example involving three-dimensional motion and slip at the contact is given here. The body used in this calculation consists of two solid spheres with 1 kg mass and 0.5 m radius connected with a 4 m massless rod between their centers of mass. The coefficients of restitution are $e_n=e_t=1$ and the coefficient of friction is $\mu=0.2$. The body is subject to a downwards gravitational force and is released at 20 m height

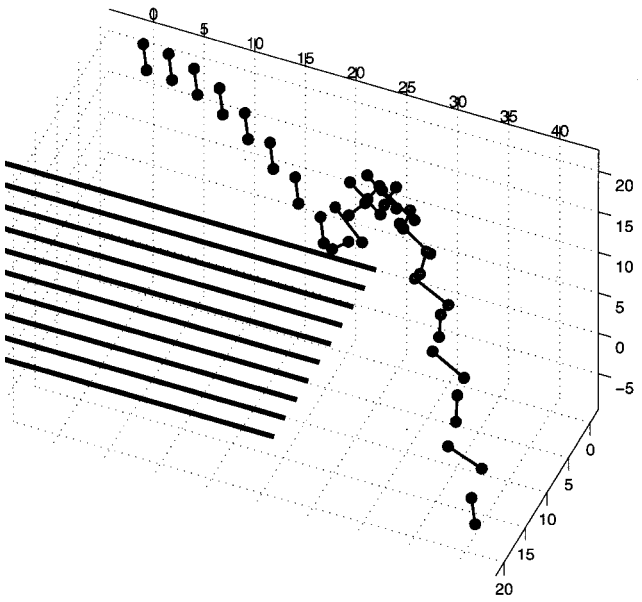


Fig. 7 Motion in the three dimensional example

with a horizontal 10 m/s velocity. The axis of rotational symmetry is tilted 0.2 rad in the plane perpendicular to the direction of the initial velocity. The time step used in the calculation is 0.008333333 s.

The resulting motion is shown in Fig. 7, where the body is plotted four times each second. When the body hits the ground there is a tangential contact impulse which is not in the direction opposite to the initial velocity, because of the tilt. The body begins to rotate, and the subsequent motion is not in the plane spanned by the initial velocity and gravitational direction.

5 Conclusions

In this paper a mathematical formulation and an algorithm for rigid-body impact with friction has been developed, where the governing laws are formulated as a system of B-differentiable equations which are solved directly with Newton's method. The method has performed very satisfactorily numerically. An interesting extension of the present work would be to test the performance of the method in cases of large numbers of simultaneous contacts, such as granular flow.

The predictions of the algorithm has also been compared to experiments, showing good agreement. An interesting point is that the agreement between computations and experiment was actually found to be better when putting $e_n = e_t$ rather than relying on values for e_t previously found by analyzing single ball experiments. Since this choice will guarantee the energy consistency of the algorithm, it seems that this is the prudent choice to make if no experimental value from a situation very close to that which one wants to model is at hand.

Acknowledgment

The authors wish to thank Mr. Bo Skoog for his invaluable help with the experiments.

References

[1] Newton, I., 1687, *Philosophiae Naturalis Principia Mathematica*, Swedish translation by C. V. L. Charlier, Gleerups, Lund, 1927–1931 and Liber, Malmö, 1986.
[2] Kilmister, C. W., and Reeve, J. E., 1966, *Rational Mechanics*, Longmans, London.
[3] Brach, R. M., 1989, ''Rigid Body Collisions,'' ASME J. Appl. Mech., **56**, pp. 133–138.
[4] Stronge, W. J., 1990, ''Rigid Body Collisions With Friction,'' Proc. R. Soc. London, Ser. A, **431**, pp. 169–181.
[5] Walton, O. R., 1993, ''Numerical Simulation of Inelastic, Frictional Particle-Particle Interactions,'' *Particulate Two-Phase Flow*, M. C. Roco, ed., Butterworth-Heinemann, Stoneham, pp. 884–911.
[6] Stoianovici, D., and Hurmuzlu, Y., 1996, ''A Critical Study of the Applicability of Rigid-Body Collision Theory,'' ASME J. Appl. Mech., **63**, pp. 307–316.
[7] Calsamiglia, J., Kennedy, S. W., Chatterjee, A., Ruina, A., and Jenkins, J. T., 1999, ''Anomalous Frictional Behavior in Collisions of Thin Disks,'' ASME J. Appl. Mech., **66**, pp. 146–152.
[8] Moreau, J. J., 1988, ''Unilateral Contact and Dry Friction,'' *Nonsmooth Mechanics and Applications* (CISM Courses and Lectures No. 302), J. J. Moreau and P. D. Panagiotopoulos, eds., Springer, Wien, pp. 1–82.
[9] Moreau, J. J., 1994, ''Some Numerical Methods in Multibody Dynamics: Application to Granular Materials,'' Eur. J. Mech. A/Solids, **13**, pp. 93–114.
[10] Johansson, L., 1999, ''A Linear Complementarity Algorithm for Rigid Body Impact With Friction,'' Eur. J. Mech. A/Solids, **18**, pp. 703–717.
[11] Christensen, P. W., Klarbring, A., Pang, J. S., and Strömberg, N., 1998, ''Formulation and Comparison of Algorithms for Frictional Contact Problems,'' Int. J. Numer. Methods Eng., **42**, pp. 145–173.
[12] Alart, P., and Curnier, A., 1991, ''A Mixed Formulation for Frictional Contact Problems Prone to Newton Like Solution Methods,'' Comput. Methods Appl. Mech. Eng., **92**, pp. 353–375.
[13] Stevens, B. L., and Lewis, F. L., 1992, *Aircraft Control and Simulation*, John Wiley and Sons, New York.
[14] Brogliato, B., 1996, *Nonsmooth Impact Mechanics: Models, Dynamics and Control*, Springer, Berlin.
[15] Pang, J. S., 1990, ''Newton's Method for B-Differentiable Equations,'' Math. Op. Res., **15**, pp. 311–341.
[16] Klarbring, A., 1992, ''Mathematical Programming and Augmented Lagrangian Methods for Frictional Contact Problems,'' *Proceedings Contact Mechanics International Symposium*, A. Curnier, ed., Presses Polytechniques et Universitaires Romandes, Lausanne, pp. 409–422.

Unsteady Laminar Duct Flow With a Given Volume Flow Rate Variation

D. Das

J. H. Arakeri

Department of Mechanical Engineering,
Indian Institute of Science,
Bangalore 560 012, India

In this paper we give a procedure to obtain analytical solutions for unsteady laminar flow in an infinitely long pipe with circular cross section, and in an infinitely long two-dimensional channel, created by an arbitrary but given volume flow rate with time. In the literature, solutions have been reported when the pressure gradient variation with time is prescribed but not when the volume flow rate variation is. We present some examples: (a) the flow rate has a trapezoidal variation with time, (b) impulsively started flow, (c) fully developed flow in a pipe is impulsively blocked, and (d) starting from rest the volume flow rate oscillates sinusoidally. [S0021-8936(00)01702-5]

1 Introduction

The solution for laminar flow in a pipe for a prescribed time variation of an axial pressure gradient is well known (see, for example, Szymanski [1] for impulsively imposed pressure gradient and Uchida [2] for sinusoidally varying pressure gradient). However, an analytical solution when the volume flow rate is prescribed as a function of time has not been reported in the literature. The temporal decay of a fully developed pipe and channel flow following a sudden blockage is discussed by Weinbaum and Parker [3]. They obtain the velocity using an approximate Pohlhausen-type analysis. In this paper we give a method to obtain analytically the velocity for an arbitrary, but given, flow rate $Q(t)$, in a pipe of circular cross section and in a two-dimensional channel.

Cases where the volume flow rate variation is known, rather than the pressure gradient variation, are frequently encountered. Sudden blockage of flow, as by a valve, in a pipe is a common experience and of obvious importance. Unsteady flow in a long pipe where the volume flow rate is measured (say with an orifice meter), flow in a pipe driven by a known motion of a piston, and decaying oscillation of a liquid column in a U-tube are other examples. Generally the velocity profiles in these flows, at least during the deceleration phases, are inflectional in nature and hence unstable at low Reynolds numbers. Hence the bidirectional solution is valid for a short time, but accurate determination of these profiles is important in determining the stability characteristics of these flows.

2 Problem Formulation

We use the cylindrical polar coordinates (r,θ,x), where, r is radial distance from the center of the pipe, θ is the angular direction, and x is the axial direction. Velocities in the r, θ and x-direction are u_r, u_θ, and u, respectively. We consider incompressible bidirectional flow in an infinitely long circular pipe (radius$=R$) with zero swirl: $u_\theta=0$, $u_r=0$ and $u=u(r,t)$. The condition of incompressibility implies that any pressure change is felt instantaneously everywhere.

The governing equation of motion in the x-direction is

$$\frac{\partial u}{\partial t}=-\frac{1}{\rho}\frac{\partial P}{\partial x}+\nu\left(\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}\right) \tag{1}$$

where P is pressure.

The radial momentum equation in the absence of any body force is

$$\frac{\partial P}{\partial r}=0. \tag{2}$$

This implies that the pressure gradient is a function of time alone.

The boundary conditions are

$$u(R,t)=0, \tag{3}$$

and

$$\frac{\partial u(0,t)}{\partial r}=0. \tag{4}$$

We need an initial condition for the velocity, $u(r,0)$, which depends on the problem we are considering. The solution of these equations is possible if pressure, as a function of time, is known. In our case, however, pressure is unknown and determined indirectly by the volume flow rate, which is given. The velocity is related to the volume flow rate by

$$\int_0^R 2\pi r u(r,t)dr=u_p(t)\pi R^2=Q(t). \tag{5}$$

u_p is the velocity averaged over the cross section; it can be considered as the velocity of a piston which would cause the flow rate Q.

With this additional condition the analytical solution of Eq. (1) is possible using the Laplace transform technique. The Laplace transform of Eqs. (1), (3), (4), and (5) gives

$$\frac{d^2\overline{u}(r,s)}{dr^2}+\frac{1}{r}\frac{d\overline{u}(r,s)}{dr}-\frac{s}{\nu}\overline{u}(r,s)=\frac{1}{\mu}\frac{d\overline{P}(x,s)}{dx}+\frac{1}{\nu}u(r,0) \tag{6}$$

$$\overline{u}(R,s)=0 \tag{7}$$

$$\left.\frac{d\overline{u}}{dr}\right|_{r=0}=0 \tag{8}$$

$$\int_0^R 2\pi r\overline{u}dr=\overline{u}_p(s)\pi R^2 \tag{9}$$

where $\overline{u}(r,s)=\int_0^\infty e^{-st}u(r,t)dt$, $\overline{P}(r,s)=\int_0^\infty e^{-st}P(r,t)dt$.

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 17, 1998; final revision, September 15, 1999. Associate Technical Editor: D. A. Siginer. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Equation (6) is a second-order inhomogeneous ordinary differential equation. The homogeneous part is the modified Bessel's equation of zeroth order and assuming the particular integral as ϕ_p, the complete solution is

$$\overline{u}(r,s)=C_1I_o(pr)+C_2K_o(pr)+\phi_p. \tag{10}$$

Here $p=\sqrt{s/\nu}$, I_o and K_o are modified Bessel's functions of the first and second kinds respectively; C_1 and C_2 are arbitrary constants.

Using the boundary conditions (7) and (8) in Eq. (10) we determine the two unknown coefficients C_1 and C_2 (for details see Das [4]). Substituting for C_1 and C_2 in Eq. (10) we get

$$\overline{u}(r,s)=\phi_p\left[1-\frac{I_o(pr)}{I_o(pR)}\right]. \tag{11}$$

To get the unknown ϕ_p we substitute the expression for $\overline{u}(r,s)$ (Eq. (11)) in Eq. (9)

$$2\phi_p\left[\int_0^R rdr-\frac{1}{I_o(pR)}\int_0^R rI_o(pr)dr\right]=\overline{u_p}(s)R^2 \tag{12}$$

to obtain

$$\phi_p=\frac{\overline{u_p}(s)}{\left[1-\dfrac{2I_1(pR)}{pRI_o(pR)}\right]}. \tag{13}$$

Putting the value of ϕ_p in Eq. (11) we obtain

$$\overline{u}(r,s)=\frac{pR\overline{u_p}(s)[I_o(pR)-I_o(pr)]}{[pRI_o(pR)-2I_1(pR)]} \tag{14}$$

or

$$\overline{u}(r,s)=\overline{u_p}(s)\cdot\overline{G}(r,s), \tag{15}$$

where

$$\overline{G}(r,s)=\frac{pR[I_o(pR)-I_o(pr)]}{[pRI_o(pR)-2I_1(pR)]}. \tag{16}$$

As $p=\sqrt{(s/\nu)}$, we can write $pr=A\sqrt{(s)}$ and $pR=B\sqrt{(s)}$ where $A=r/\sqrt{\nu}$ and $B=R/\sqrt{\nu}$ and

$$\overline{G}(r,s)=\frac{[I_o(B\sqrt{s})-I_o(A\sqrt{s})]}{\left[I_o(B\sqrt{s})-\dfrac{2I_1(B\sqrt{s})}{B\sqrt{s}}\right]} \tag{17}$$

The inversion of the Eq. (15) cannot be obtained using the convolution theorem. Applying the convolution theorem to Eq. (15) we obtain

$$u(r,t)=\int_0^t u_p(t)G(r,t-\tau)dt. \tag{18}$$

One needs to evaluate

$$G(r,t)=\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\overline{G}(r,s)e^{st}ds. \tag{19}$$

We can write the integrand in the form of $a\Gamma^{n+1}/b\Gamma^n$ where Γ is the radius of the contour taken such that all the poles lie in the left of the contour. The integrand diverges (as $\Gamma\to\infty$), preventing the application of the convolution theorem.

Hence we take the inverse transform of Eq. (15) and obtain

$$u(r,t)=\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\overline{u_p}(s)\overline{G}(r,s)e^{st}ds. \tag{20}$$

The solution of Eq. (20) can be obtained for any arbitrary piston motion provided $u_p(t)$ is Laplace transformable.

3 Examples

We consider some examples of piston motion which may be of interest. We consider a trapezoidal piston motion, i.e., the piston motion has three stages: constant acceleration of the piston starting from rest, a period of constant velocity, and a constant deceleration of the piston to a stop. An experimental study of the evolution and stability of such flows has been carried out by Das [4] and Das and Arakeri [5]. Often vortex rings are generated using such a trapezoidal piston motion. The flow due to linearly accelerated piston motion, constant piston motion, and impulsively started flow are discussed as special cases of the trapezoidal piston motion. The solution due to a sinusoidally varying piston motion is also discussed. A few new properties of Bessel's functions have been obtained which are discussed in the Appendix.

3.1 Trapezoidal Piston Motion. We obtain here the solution for a piston velocity which has a trapezoidal variation with time as given below:

$$\begin{aligned} u_p(t) &= \frac{U_pt}{t_o} && \text{for } 0\leqslant t\leqslant t_o\\ &= U_p && \text{for } t_o\leqslant t\leqslant t_1\\ &= \frac{U_p(t_2-t)}{(t_2-t_1)} && \text{for } t_1\leqslant t\leqslant t_2\\ &= 0 && \text{for } t_2\leqslant t\leqslant\infty. \end{aligned} \tag{21}$$

For $0\leqslant t\leqslant t_o$, taking the Laplace transform of Eq. (21) we get

$$\overline{u_p}(s)=\frac{U_p}{t_os^2}. \tag{22}$$

From Eq. (20) we obtain

$$u(r,t)=\frac{1}{2\pi i}\left[2\pi i\Sigma\ \text{residues of poles of}\ \left\{\frac{U_p}{t_os^2}e^{st}\overline{G}(r,s)\right\}\right]. \tag{23}$$

We have to determine the poles of the above expression and it can be easily shown that $s=0$ is a pole of order 2. The residue at $s=0$ is

$$Res_o=\frac{U_p}{t_o}\left[2t\left\{1-\left(\frac{r}{R}\right)^2\right\}+\frac{R^2}{2\nu}\left(\frac{1}{4}\left\{1-\left(\frac{r}{R}\right)^4\right\}-\frac{1}{3}\left\{1-\left(\frac{r}{R}\right)^2\right\}\right)\right]. \tag{24}$$

The other singular points are the zeroes of

$$I_o(B\sqrt{s})-\frac{2I_1(B\sqrt{s})}{B\sqrt{s}}.$$

Writing $B\sqrt{s}=-iv$ we find that

$$I_o(B\sqrt{s})-\frac{2I_1(B\sqrt{s})}{B\sqrt{s}}=J_o(v)-\frac{2}{v}J_1(v)=J_2(v). \tag{25}$$

Hence the poles are the zeroes of the Bessel's function of second order. If v_n, $n=1,2,\ldots,\infty$ are zeroes of $J_2(v)$ then $s_n=-v_n^2/B^2$, $n=1,2,\ldots,\infty$ are the poles. Since all v_n are symmetrically placed about zero on the real axis, all the poles (s_n) lie on the negative real axis. These are simple poles, and residues at all these poles can be obtained as

$$Res_n=\frac{2U_pB^2}{t_o}e^{-(v_n/B)^2t}\frac{\left[J_o(v_n)-J_o\left(\dfrac{r}{R}v_n\right)\right]}{[v_n^3J_1(v_n)]}. \tag{26}$$

Adding Res_o and Res_n we obtain solution for $0\leqslant t\leqslant t_o$ and substituting for $B(=R/\sqrt{\nu})$

$$\frac{u}{U_p}=\frac{1}{t_o}\left[2t(1-c^2)+\frac{R^2}{2\nu}\left(\frac{1}{4}(1-c^4)-\frac{1}{3}(1-c^2)\right)\right]+\frac{2R^2}{t_o\nu}\Sigma_{n=1}^{\infty}e^{(-v_n^2\nu/R^2)t}\left[\frac{J_0(v_n)-J_0(cv_n)}{v_n^3J_1(v_n)}\right]. \quad (27)$$

Here $c=r/R$. It may be noted that U_p/t_o is the piston acceleration. Transition in a pipe flow started from rest with the linear increase in the mean velocity was studied by Lefebvre and White [6]. In their analysis of the data they assumed that the velocity profiles given by Szymanski [1] which is the solution of laminar pipe flow caused by a suddenly applied pressure gradient, is valid.

The same method is used (for details see Das [4]) to obtain the solution during constant piston velocity phase ($t_o\leqslant t\leqslant t_1$),

$$\frac{u}{U_p}=2(1-c^2)+\frac{2R^2}{\nu t_o}\Sigma_{n=1}^{\infty}(e^{(-v_n^2\nu/R^2)t}-e^{(-v_n^2\nu/R^2)(t-t_o)})\times\left[\frac{J_0(v_n)-J_0(cv_n)}{v_n^3J_1(v_n)}\right], \quad (28)$$

during the deceleration of the piston motion ($t_1\leqslant t\leqslant t_2$),

$$\frac{u}{U_p}=2\frac{t_2-t}{t_2-t_1}(1-c^2)-\frac{R^2}{2\nu(t_2-t_1)}\left(\frac{1}{4}(1-c^4)-\frac{1}{3}(1-c^2)\right)+\frac{2R^2}{\nu}\Sigma_{n=1}^{\infty}\left(\frac{e^{(-v_n^2\nu/R^2)t}-e^{(-v_n^2\nu/R^2)(t-t_o)}}{t_o}-\frac{e^{(-v_n^2\nu/R^2)(t-t_1)}}{t_2-t_1}\right)\left[\frac{J_0(v_n)-J_0(cn_n)}{v_n^3J_1(v_n)}\right], \quad (29)$$

and after the piston has stopped, i.e., $t_2\leqslant t\leqslant\infty$,

$$\frac{u}{U_p}=\frac{2R^2}{\nu}\Sigma_{n=1}^{\infty}\left(\frac{e^{(-v_n^2\nu/R^2)t}-e^{(-v_n^2\nu/R^2)(t-t_o)}}{t_o}+\frac{e^{(-v_n^2\nu/R^2)(t-t_2)}-e^{(-v_n^2\nu/R^2)(t-t_1)}}{t_2-t_1}\right)\left[\frac{J_0(v_n)-J_0(cv_n)}{v_n^3J_1(v_n)}\right]. \quad (30)$$

From Eq. (6) we get

$$\phi_p=-\frac{1}{s\rho}\frac{d\overline{P}}{dx}. \quad (31)$$

Here $d\overline{P}/dx$ is unknown and can be determined from Eq. (9). Hence using equation (31) we obtain the pressure gradient as

$$\frac{d\overline{P}}{dx}=-\frac{\rho s\overline{u_p}(s)}{\left[1-\frac{2I_1(pR)}{pRI_o(pR)}\right]}. \quad (32)$$

We obtain the expressions for the variation of nondimensional pressure gradient with time by taking the inverse transform of Eq. (32). During the piston acceleration, $0\leqslant t\leqslant t_o$,

$$\frac{dP^*}{dx^*}=-\frac{1}{6}\frac{R^2}{\nu t_o}-\frac{t}{t_o}+\frac{R^2}{4\nu t_o}\sum_{n=1}^{\infty}e^{(-v_n^2\nu/R^2)t}\left[\frac{J_0(v_n)}{v_nJ_1(v_n)}\right], \quad (33)$$

where $P^*=P/(8\mu U_p/R)$ is nondimensional pressure and $x^*=x/R$. During the time when the piston velocity is constant ($t_o\leqslant t\leqslant t_1$),

$$\frac{dP^*}{dx^*}=-1+\frac{R^2}{4\nu t_o}\sum_{n=1}^{\infty}(1-e^{(v_n^2\nu/R^2)t_o})e^{(-v_n^2\nu/R^2)t}\left[\frac{J_0(v_n)}{v_nJ_1(v_n)}\right], \quad (34)$$

during the piston deceleration ($t_1\leqslant t\leqslant t_2$),

$$\frac{dP^*}{dx^*}=-\frac{1}{6}\frac{R^2}{\nu(t_2-t_1)}-\frac{(t_2-t)}{(t_2-t_1)}+\frac{R^2}{4\nu}\sum_{n=1}^{\infty}\left(\frac{e^{-(v_n^2\nu/R^2t)}-e^{(-v_n^2\nu/R^2)(t-t_o)}}{t_o}-\frac{e^{-(v_n^2\nu/R^2)(t-t_1)}}{(t_2-t_1)}\right)\times\left[\frac{J_0(v_n)}{v_nJ_1(v_n)}\right], \quad (35)$$

and after the piston has stopped ($t_2\leqslant t\leqslant\infty$),

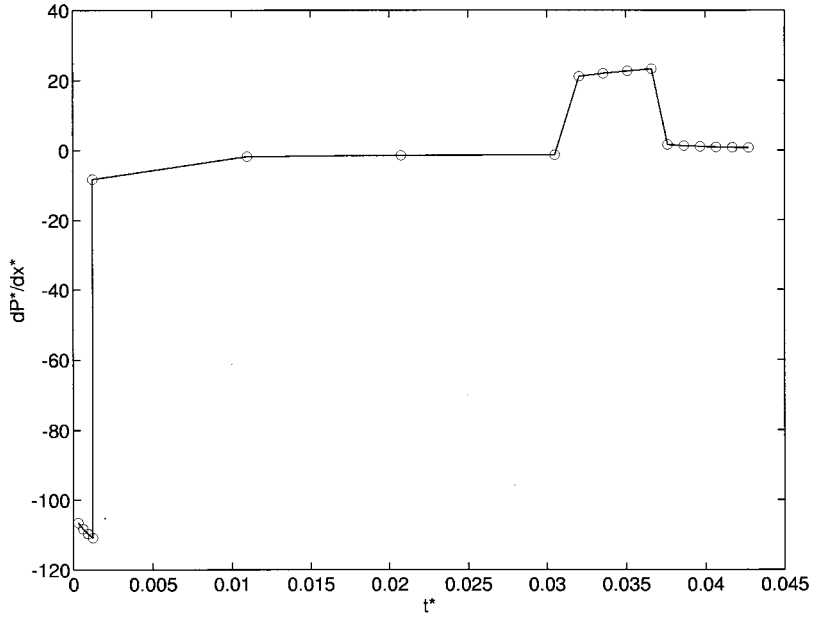


Fig. 1 The variation of pressure gradient with time for trapezoidal piston motion

$$\frac{dP^*}{dx^*}=\frac{R^2}{4\nu}\sum_{n=1}^{\infty}\left(\frac{e^{(-v_n^2\nu/R^2)t}-e^{(-v_n^2\nu/R^2)(t-t_o)}}{t_o}+\frac{e^{(-v_n^2\nu/R^2)(t-t_2)}-e^{(-v_n^2\nu/R^2)(t-t_1)}}{t_2-t_1}\right)\left[\frac{J_0(v_n)}{v_nJ_1(v_n)}\right]. \tag{36}$$

The infinite series in Eqs. (27) to (30) are convergent and $n=50$ is enough for the cases we are considering. The solution depends on t_o, t_1, and t_2. Velocity profiles calculated for trapezoidal piston motion are plotted in Fig. 2 for different nondimensional times ($t^*=t\nu/R^2$) with $t_o\nu/R^2$, $t_1\nu/R^2$ and $t_2\nu/R^2=0.0012$, 0.0305, and 0.0366, respectively. Figure 1 shows the nondimensional pressure gradient with time. During the acceleration and deceleration phases the pressure gradients are large mainly because of fluid inertia. After the piston stops the pressure gradient slowly decays to zero. The centerline velocity changes with the piston velocity during the acceleration phase and boundary layer near the wall starts growing (see Fig. 2(*a*)). In Fig. 2(*b*) velocity profiles are shown when the piston moves at constant velocity. Diffusion of vorticity causes the boundary layer to grow with time and the center line velocity increases to satisfy the mass flow condition. During the time period when the piston decelerates and stops at time t_2, an adverse pressure gradient causes the flow to reverse its direction near the wall. This can be seen in Fig. 2(*c*). It is observed that the reverse flow first appears near wall and starts moving towards the centerline. After the piston motion ceases, the velocity profile (Fig. 2(*d*)) continues to have reverse flow near the wall to satisfy the zero mass flow condition. With time the velocity decays to zero everywhere. Velocity profiles beyond t_1 contain an inflection point and have reverse flow. The solution is clearly valid only as long as the flow is laminar, i.e., until instability sets in.

When the piston moves at a constant speed after initial acceleration we obtain the solution from (28). We observe that at infinite time the velocity is the Hagen-Poiseuille solution. Similarly the solution for the pressure shows that the pressure gradient ($=-8\nu U_p/R^2$) for the Hagen-Poiseuille solution is obtained as time tends to infinity.

3.2 Constant Acceleration Case. For a piston with constant acceleration ($=a_p; u_p=a_pt$) the solution is obtained from Eq. (27) by putting $t_o=t$,

$$\frac{u(r,t)}{u_p}=2\{1-c^2\}+\frac{R^2}{2\nu t}\left(\frac{1}{4}\{1-c^4\}-\frac{1}{3}\{1-c^2\}\right)+\frac{2R^2}{t\nu}\times\sum_{n=1}^{\infty}e^{(-v_n^2\nu/R^2)t}\left[\frac{J_0(v_n)-J_0(cv_n)}{v_n^3J_1(v_n)}\right]. \tag{37}$$

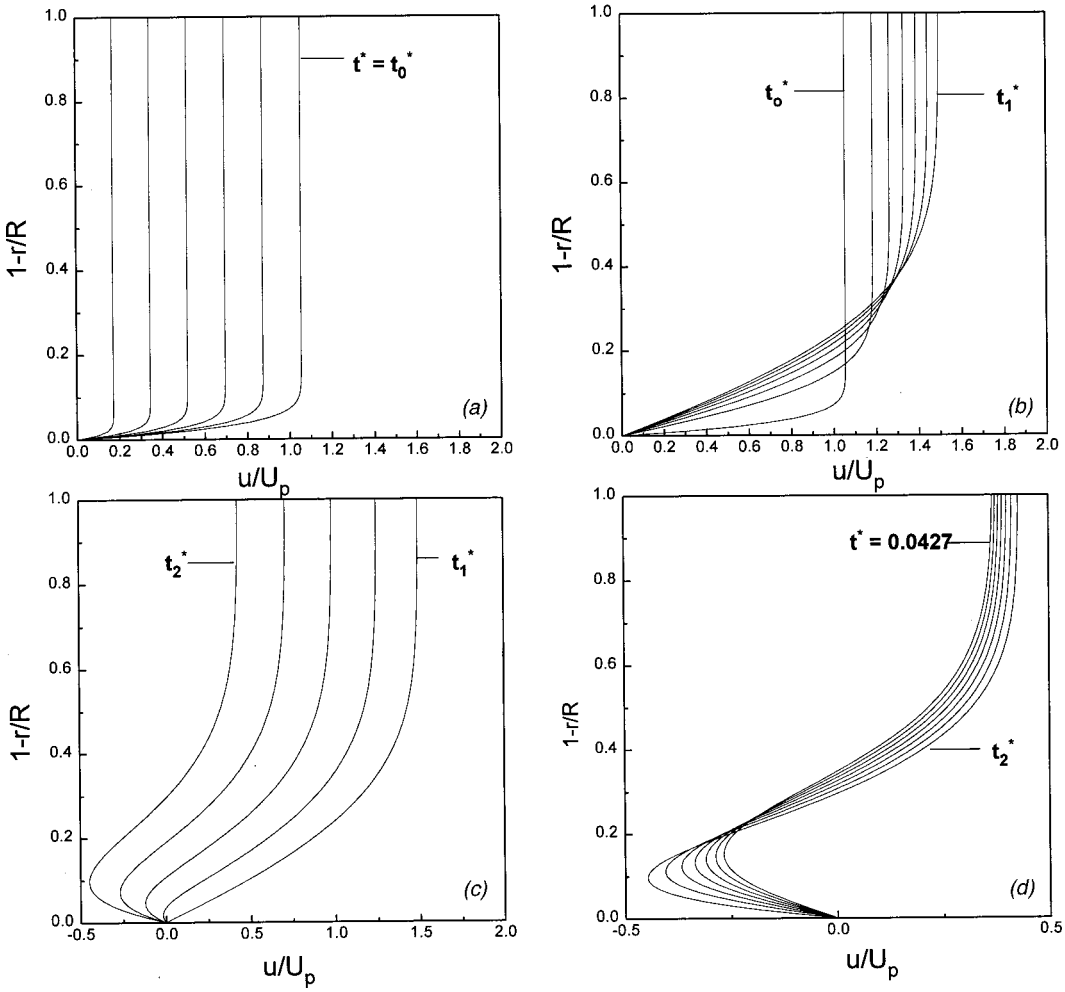


Fig. 2 Velocity profiles at different times (*a*) during the acceleration of the piston motion (profiles are shown at time intervals of $t_o^*/6$), (*b*) when the piston velocity is constant (profiles are at time intervals of $(t_1^*-t_o^*)/6$), (*c*) during the deceleration of the piston velocity (profiles are at time intervals of $(t_2^*-t_1^*)/4$), and (*d*) after the piston motion has stopped (profiles are at time intervals of $(0.0427-t_2^*)/6$)

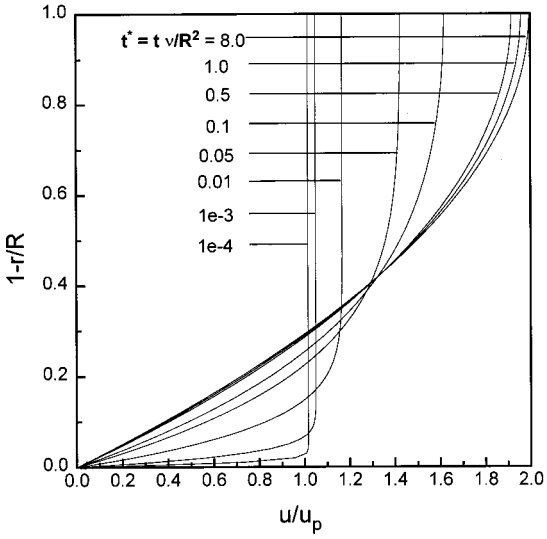


Fig. 3 Velocity profiles at different nondimensional times for constant acceleration of the piston motion

Figure 3 shows the velocity profiles at different times.

Here the solution for pressure gradient as time tends to infinity is

$$\frac{dP^*}{dx^*} = -1 - \frac{1}{6}\frac{a_p R^2}{\nu u_p}. \tag{38}$$

The first terms in (37) and (38) corresponding to the Hagen-Poiseuille solution (corresponding to zero piston acceleration) and the second terms are due to the piston acceleration.

3.3 Impulsively Started Flow. The solution of an impulsively started flow in a pipe,

$$\begin{aligned} u_p &= 0, \qquad \text{for } t \leqslant 0 \\ &= U_p, \qquad \text{for } t > 0, \end{aligned} \tag{39}$$

is given by

$$\frac{u}{U_p} = 2(1-c^2) - 2\sum_{n=1}^{\infty} e^{(-v_n^2 \nu/R^2)t}\left[\frac{J_0(v_n) - J_0(cv_n)}{v_n J_1(v_n)}\right]. \tag{40}$$

Figure 4(*a*) shows the velocity profiles at different times. The corresponding pressure gradient, plotted in Fig. 4(*b*), is

$$\frac{dP^*}{dx^*} = -1 - \frac{1}{4}\sum_{n=1}^{\infty} e^{(-v_n^2 \nu/R^2)t}\left[\frac{(v_n J_0(v_n))}{J_1(v_n)}\right]. \tag{41}$$

At $t=0$ an infinite pressure gradient is required to counter the infinite wall shear stress.

Note that the velocity profile looks different from those corresponding to an impulsively applied pressure gradient. Also in the present case the solution tends to the final asymptotic state at $\nu t/R^2 \simeq 0.2$ compared to the Szymanski [1] solution value of $\nu t/R^2 \simeq 0.8$.

3.4 Impulsively Blocked Fully Developed Flow. The exact solution of the problem considered by Weinbaum and Parker [3] is

$$\frac{u}{U_p} = -2\sum_{n=1}^{\infty} e^{(-v_n^2 \nu/R^2)t}\left[\frac{J_0(v_n) - J_0(cv_n)}{v_n J_1(v_n)}\right]. \tag{42}$$

The initial condition for this problem is $u(r,0) = 1-c^2$, and the mass flow condition is $\int_0^R 2\pi r u dr = 0$.

3.5 Oscillatory Flow. The asymptotic solution for flow in a pipe or in a two-dimensional channel due to a sinusoidally oscillating piston is known (see Uchida [2] and Wembersley [7]). The solution is obtained by using the fact that the pressure gradient also oscillates sinusoidally. We consider the case of oscillating piston motion starting from rest, i.e., $u_p = 0$ for $t<0$ and $u_p = U_o \sin(\omega t)$ for $t>0$. The poles are simple poles at $s = i\omega$ and $s = -i\omega$, and are the zeroes of $J_2(v)$. The solution is

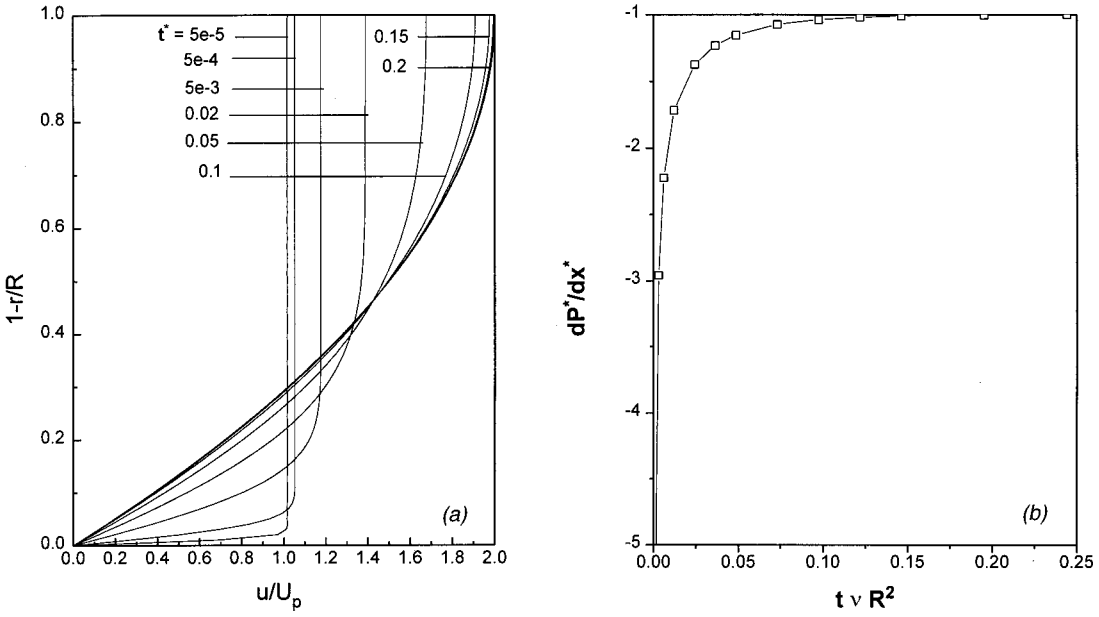


Fig. 4 Velocity profiles (*a*) and pressure gradient (*b*) for impulsively started motion

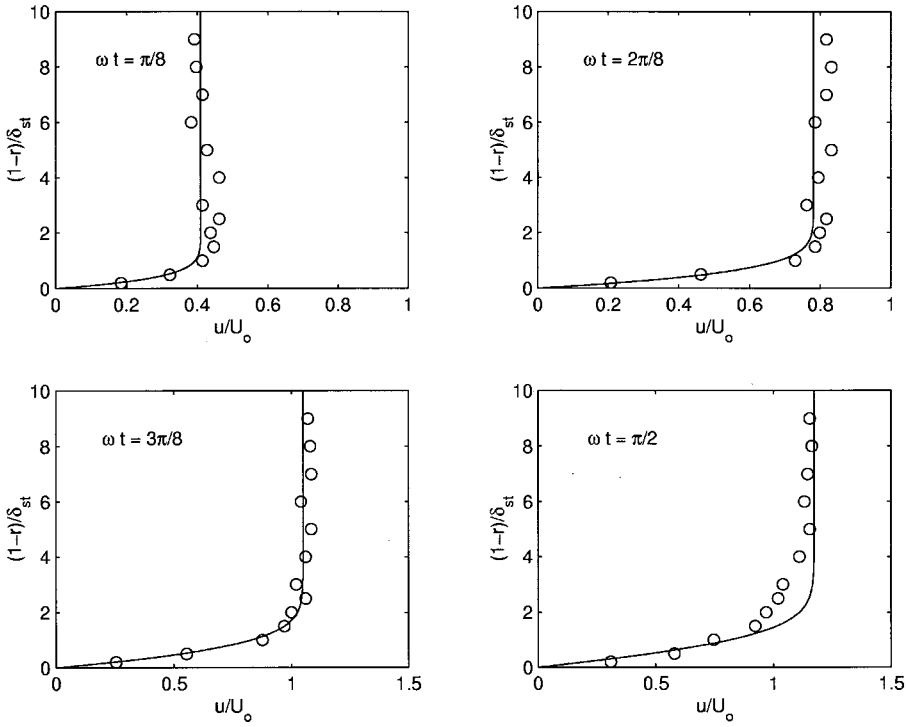


Fig. 5 Comparison of the present solution (solid line) for the unidirectional oscillatory flow starting from rest with the experimental data of Akhavan et al. [8] (symbols)

$$u=\frac{iU_o}{2}e^{pt}(G(r,-i\omega)-G(r,i\omega))$$
$$+2U_oB^2\omega\sum_{n=1}^{\infty}\frac{v_n}{(v_n^4+\omega^2B^4)}e^{(-v_n^2/B^2)t}\left[\frac{J_0(v_n)-J_0(cv_n)}{J_1(v_n)}\right] \quad (43)$$

where

$$G=\frac{I_o(Bi\omega)-I_o(Ai\omega)}{I_o(Bi\omega)-\dfrac{2I_1(Bi\omega)}{Bi\omega}}.$$

Substituting the residues we obtain the final expression for velocity as

$$\frac{u}{U_o}=\frac{i}{2}\left[\frac{J_0(v_a)-J_0(cv_a)}{J_2(v_a)}e^{i\omega t}-\frac{J_0(v_b)-J_0(cv_b)}{J_2(v_b)}e^{-i\omega t}\right]$$
$$+\Sigma_{n=1}^{\infty}e^{(-v_n^2\nu/R^2)t}\frac{R^2\nu\omega}{v_n^4\nu^2+R^4\omega^2}\left[\frac{2v_n(J_0(v_n)-J_0(cv_n))}{J_1(v_n)}\right]. \quad (44)$$

Here, $iv_a=R\sqrt{i\omega/\nu}$ and $iv_b=R\sqrt{-i\omega/\nu}$.

Using Eq. (31) we obtain the pressure gradient as

$$\frac{d\overline{P}}{dx}=-\frac{\rho sU_o\omega}{s^2+\omega^2}\frac{I_o(B\sqrt{(s)})}{\left[1-\dfrac{2I_1(B\sqrt{(s)})}{B\sqrt{(s)}I_o(B\sqrt{(s)})}\right]}. \quad (45)$$

We obtain the expressions for the variation of pressure gradient with time by taking the inverse transform of Eq. (45)

$$\frac{dP^*}{dx^*}=\frac{i\omega R^2}{16\nu}\left[\frac{J_0(v_a)}{J_2(v_a)}e^{i\omega t}-\frac{J_0(v_b)}{J_2(v_b)}e^{-i\omega t}\right]$$
$$+\frac{1}{8}\Sigma_{n=1}^{\infty}e^{(-v_n^2\nu/R^2)t}\frac{R^2\nu\omega}{v_n^4\nu^2+R^4\omega^2}\left[\frac{2v_nJ_0(v_n)}{J_1(v_n)}\right]. \quad (46)$$

This solution is applicable from starting of the piston motion and reaches the conventional oscillating pipe flow solution (similar to the solution of Uchida [2] and Wembersley [7]) after few cycles of acceleration and deceleration depending on the value of the Wembersely parameter (ratio of pipe diameter (R) and Stokes layer thickness, $\delta_{st}=\sqrt{2\nu/\omega}$). Thus the solution obtained through the present method will have more applicability compared to the existing solutions of fully developed oscillating pipe flow. The solution is compared with the experimental results of Akhavan [8] in Fig. 5. Figure 6 shows velocity profiles at different phases for two cycles of oscillation starting from rest at $t=0$. Note the changes in shape of the velocity profiles at a particular ωt in the first two cycles of oscillation.

4 Channel Flow

In case of flow in a two-dimensional channel similar equations are solved and the solutions for a trapezoidal piston motion (as given by Eq. (21)) at different times are
for $0\leqslant t\leqslant t_o$

$$\frac{u}{U_p}=\frac{1}{t_o}\left(\frac{3t}{2}(1-c_h^2)-\frac{h^2}{40\nu}(5c_h^4-6c_h^2+1)\right)$$
$$-\frac{2h^2}{\nu t_o}\Sigma_{nh=1}^{\infty}e^{-v_{nh}^2\nu t/h^2}\left[\frac{\cos(c_hv_{nh})-\cos(v_{nh})}{v_{nh}^3\sin(v_{nh})}\right], \quad (47)$$

for $t_o\leqslant t\leqslant t_1$

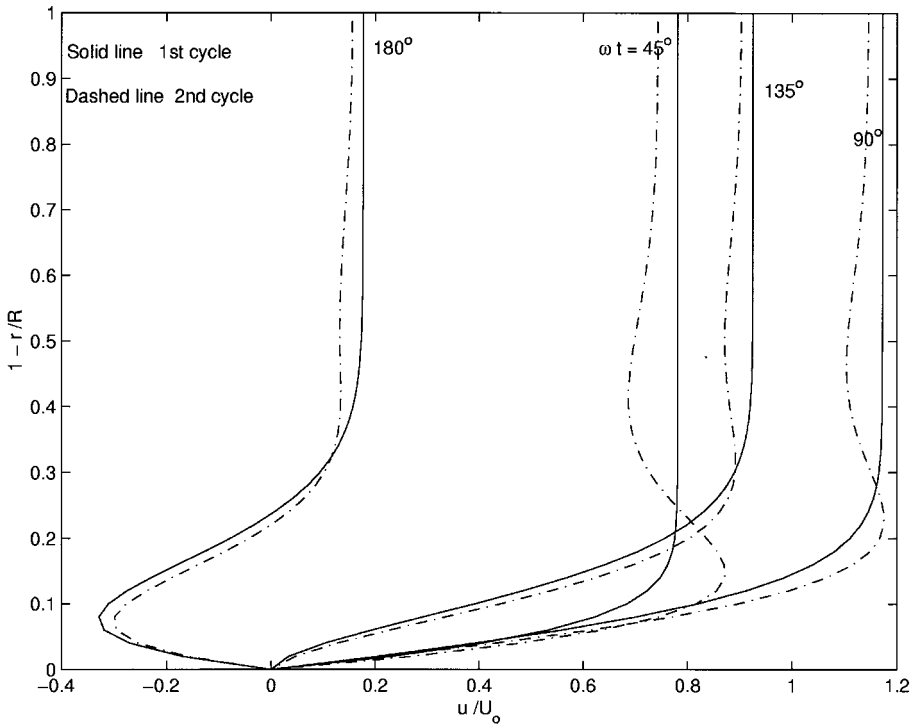


Fig. 6 Velocity profiles at different phases for two cycles, for oscillating pipe flow starting from rest at $t=0$ ($R/\sqrt{2\nu/\omega}=10$)

$$\frac{u}{U_p}=\frac{3}{2}(1-c_h^2)-\frac{2h^2}{\nu t_o}\Sigma_{nh=1}^{\infty}(e^{(-v_{nh}^2\nu/h^2)t}(-e^{-v_{nh}^2\nu/h^2})(t-t_o))\times\left[\frac{\cos(c_h v_{nh})-\cos(v_{nh})}{v_{nh}^3\sin(v_{nh})}\right], \quad (48)$$

for $t_1\leqslant t\leqslant t_2$

$$\frac{u}{U_p}=\frac{3}{2}(1-c_h^2)\left(\frac{t_2-t}{t_2-t_1}\right)+\frac{1}{t_2-t_1}\frac{h^2}{40\nu}(5c_h^4-6c_h^2+1)-\frac{2h^2}{\nu}\Sigma_{nh=1}^{\infty}\left(\frac{e^{(-v_{nh}^2\nu/h^2)t}-e^{(-v_{nh}^2\nu/h^2)(t-t_o)}}{t_o}-\frac{e^{(-v_{nh}^2\nu/h^2)(t-t_1)}}{t_2-t_1}\right)\left[\frac{\cos(c_h v_{nh})-\cos(v_{nh})}{v_{nh}^3\sin(v_{nh})}\right], \quad (49)$$

and for $t_2\leqslant t\leqslant\infty$

$$\frac{u}{U_p}=-\frac{2h^2}{\nu}\Sigma_{nh=1}^{\infty}\left(\frac{e^{(-v_{nh}^2\nu/h^2)t}-e^{(-v_{nh}^2\nu/h^2)(t-t_o)}}{t_o}+\frac{e^{(-v_{nh}^2\nu/h^2)(t-t_2)}-e^{(-v_{nh}^2\nu/h^2)(t-t_1)}}{t_2-t_1}\right)\times\left[\frac{\cos(c_h v_{nh})-\cos(v_{nh})}{v_{nh}^3\sin(v_{nh})}\right] \quad (50)$$

where subscript h indicates the quantities for channel. h is channel half height, y is the distance from the centerline towards the wall of the channel, $c=y/h$, and v_{nh}, $nh=1,2\ldots\ldots\ldots\infty$, are roots of $\tan(v)=0$. The solution for the problem impulsively blocked fully developed channel flow, considered by Weinbaum and Parker [3], is

$$\frac{u}{U_p}=2\Sigma_{n=1}^{\infty}e^{(-v_n^2\nu/h^2)t}\left[\frac{\cos(c_h v_{nh})-\cos(v_{nh})}{v_{nh}\sin(v_{nh})}\right]. \quad (51)$$

5 Conclusions

We have obtained analytical solutions for unsteady bidirectional flows when the volume flow rate is given as a function of time. These solutions are applicable in those cases where the volume flow rate is known as a function of time either by measurement or by when the flow is driven, for example, by the controlled motion of a piston. Solution procedure is given for arbitrary piston motions. Different examples are considered.

Appendix

Some Properties of Zeros of Bessel's Function of Second Kind and Zeros of tan(v). From the solutions of Eqs. (1) to (5) for the trapezoidal piston motion it is possible to obtain some properties of the zeros of Bessel's function of second kind. We match the solutions obtained from (27) and (28) at $t=t_o$. This matching leads to the relation

$$\Sigma_{n=1}^{\infty}\left[\frac{J_0(cv_n)-J_0(v_n)}{v_n^3J_1(v_n)}\right]=\frac{1}{12}(1-c^2)-\frac{1}{16}(1-c^4). \quad (A1)$$

Here, v_n, $n=1,2,\ldots,\infty$ are the zeros of Bessel's function of the second kind and c is the nondimensional distance from the center of the pipe. Matching of solutions obtained from (28) and (29) at $t=t_1$ and from (29) and (30) at $t=t_2$ leads to the same relation as shown in Eq. (*A*1).

It is possible to obtain other relations from derivatives of relation (*A*1) with respect to c.

For $c=0$ the relation (*A*1) becomes

$$\Sigma_{n=1}^{\infty}\left[\frac{J_0(v_n)-1}{v_n^3J_1(v_n)}\right]=\frac{1}{48}. \quad (A2)$$

Similarly the matching of solutions (47) and (48) at $t=t_0$ leads to the relation,

$$\frac{1}{16}c_h^4-\frac{3}{40}c_h^2+\frac{1}{80}=\Sigma_{nh=1}^{\infty}\left[\frac{\cos(v_{nh})-\cos(c_h v_{nh})}{v_{nh}^3\sin(v_{nh})}\right]. \quad (A3)$$

For $c_h=0$ the relation (*A*3) becomes

$$\Sigma_{nh=1}^{\infty}\left[\frac{\cos(v_{nh})-1}{v_{nh}^3\sin(v_{nh})}\right]=\frac{1}{80}. \quad (A4)$$

References

[1] Szymanski, P., 1932, ''Some Exact Solutions of the Hydrodynamic Equations of a Viscous Fluid in the Case of a Cylindrical Tube,'' J. Math. Pures Appl., **11**, pp. 67–107.

[2] Uchida, S., 1956, ''The Pulsating Viscous Flow Superposed on the Steady Laminar Motion of Incompressible Fluids in a Circular Pipe,'' Z. Angew. Math. Phys., **7**, pp. 403–422.

[3] Weinbaum, S., and Parker, K., 1975, ''The Laminar Decay of Suddenly Blocked Channel and Pipe Flows,'' J. Fluid Mech., **69**, pp. 729–752.

[4] Das, D., 1998, Ph.D. thesis, Department of Mechanical Engineering, I.I.Sc., Bangalore, India.

[5] Das, D., and Arakeri, J. H., 1998, ''Transition of Inflectional Velocity Profiles With Reverse Flow,'' J. Fluid Mech., **374**, pp. 251–283.

[6] Lefebvre, P. J., and White, F. M., 1989, ''Experiments on Transition to Turbulence in a Constant-Acceleration Pipe Flow,'' ASME J. Fluids Eng., **111**, pp. 428–432.

[7] Wemersley, J. R., 1955, ''Method for the Calculation of Velocity, Rate of Flow and Viscous Drag in Arteries When the Pressure Gradient is Known,'' J. Physiol., **127**, pp. 553–563.

[8] Akhavan, R., Kamm, R. D., and Shapiro, A. H., 1991, ''An Investigation of Transition to Turbulence in Bounded Oscillatory Stokes Flows, Part 1—Experiments,'' J. Fluid Mech., **225**, pp. 395–422.

U. Zakout
Department of Civil Engineering,
Middle East Technical University,
0s6531 Ankara, Turkey

Z. Akkas
Department of Engineering Sciences,
Middle East Technical University,
0s6531 Ankara, Turkey

G. E. Tupholme
School of Computing and Mathematics,
University of Bradford,
Bradford, West Yorkshire BD7 1DP, UK

Transient Response of an Infinite Elastic Medium Containing a Spherical Cavity Subjected to Torsion

An exact closed-form solution is obtained for the transient response of an infinite isotropic elastic medium containing a spherical cavity subjected to torsional surface loading using the residual variable method. The main advantage of the present approach is that it eliminates the computational problems arising in the existing methods which are primarily based on Fourier or Laplace transformation techniques. Extensive numerical results for the circumferential displacements and shear stresses at various locations are presented graphically for Heaviside loadings. [S0021-8936(00)01102-8]

1 Introduction

The analysis of wave propagation in infinite elastic media containing cavities has been the topic of numerous studies in the fields of acoustics, geophysics, and seismology. In general, embedded spherical cavities subjected to specified loads generate both dilatational and transverse waves. Extensive reviews of such situations have been given by Achenbach [1], Eringen and Suhubi [2], Miklowitz [3], and Gaunaurd [4].

In this paper, we investigate the transient response of an infinite, elastic, isotropic medium resulting from the application of torsional tractions on the surface of an embedded spherical cavity.

Using Fourier transform techniques, Sato et al. [5] and Eringen and Suhubi [2] found the response of an elastic sphere subjected to a ring load and recently Godin [6] analyzed a spherical crack problem in an infinite medium under static torsional stresses. Chadwick and Trowbridge [7,8] have given detailed studies of both the torsional and the rectilinear oscillations of a rigid sphere embedded in an infinite elastic solid. They also presented ([9]) a formal solution of the elastic pulses generated within an infinite body by the application of time-dependent tractions to the surface of a spherical cavity, as an application of the three scalar wave functions representation which they use. The detailed structure of the developing ''poloidal'' and ''toroidal'' constituents of the wavefronts was analyzed by Tupholme [10] with the solutions for situations in which the spherical surface is loaded in torsion given. Tupholme [10] showed that a purely toroidal pulse can be generated by a torsional loading, and that its structure bears a close similarity to the acoustic pulses which he describes elsewhere ([11]). More generally, by quite different methods, Chadwick and Johnson [12] discussed the torsional oscillations of axisymmetric rigid convex inclusions embedded in an infinite elastic solid using ''ray theory.'' The cases of oblate and prolate spheroidal (and in particular spherical) inclusions are considered in detail.

Here the residual variable method, which is based on modified wave equations, is used to obtain exact closed-form solutions which are more amenable to numerical evaluation. The residual variable method has been used previously by Geers [13], Akkas [14,15], Akkas and Zakout [16], and Zakout and Akkas [17] to solve various problems in acoustic and elastic media. The solutions are valid for any time and location inside the medium. Numerical results are presented for the modal responses of the shear stresses and circumferential displacements in the case of modal Heaviside stresses acting on the surface of the cavity.

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 20, 1998; final revision, December 7, 1999. Associate Technical Editor: A. K. Mal. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

2 Formulation

Consider an infinite homogeneous isotropic elastic medium containing a spherical cavity of radius a as shown in Fig. 1. The spherical coordinates are (r,θ,φ), where r is the nondimensional radial coordinate related to the dimensional radial coordinate $\bar{r}$ through $r=\bar{r}/a$ and θ and φ are meridional and circumferential angles, respectively. Let $\phi(r,\theta,t)$, $\chi(r,\theta,t)$, and $\psi(r,\theta,t)$ be the displacement potentials for the elastic medium. The spherical components of the displacement vector are u_r, u_θ, and u_φ and those of the stress tensor are σ_{rr}, $\sigma_{r\theta}$, $\sigma_{r\varphi}$, $\sigma_{\theta\theta}$, $\sigma_{\varphi\varphi}$, and $\sigma_{\theta\varphi}$. The nondimensional time t is related to the dimensional time through $t=\bar{t}c_1/a$, c_1 being the dilatational wave speed in the medium. If the only shear stress component acting on the surface of the cavity is $\sigma_{r\varphi}$ and $\sigma_{r\varphi}$ is independent of φ, then no dilatational waves are generated. In this case, we also have $u_r=u_\theta=0$, $\phi=\chi=0$, and $\sigma_{rr}=\sigma_{\theta\theta}=\sigma_{\varphi\varphi}=\sigma_{r\theta}=0$. Expanding the nonzero displacement potential $\psi(r,\theta,t)$ in terms of Legendre polynomials as $\psi(r,\theta,t)=\sum_{n=0}^{\infty}\psi_n(r,t)P_n(\cos\theta)$, the nonvanishing displacement component u_φ and the stress components $\sigma_{r\varphi}$ and $\sigma_{\theta\varphi}$ can be expressed as

$$u_\varphi=\sum_{n=1}^{\infty}u_\varphi^{(n)}\frac{dP_n(\cos\theta)}{d\theta}=\sum_{n=1}^{\infty}\psi_n\frac{dP_n(\cos\theta)}{d\theta}, \tag{1}$$

$$\sigma_{r\varphi}=\frac{\mu}{a}\sum_{n=1}^{\infty}\sigma_{r\varphi}^{(n)}\frac{dP_n(\cos\theta)}{d\theta}$$

$$=\frac{\mu}{a}\sum_{n=1}^{\infty}\frac{1}{r}\left(\psi_n-r\frac{\partial\psi_n}{\partial r}\right)\frac{dP_n(\cos\theta)}{d\theta}, \tag{2}$$

$$\sigma_{\theta\varphi}=\frac{\mu}{a}\sum_{n=1}^{\infty}\frac{\psi_n}{r}\left(\frac{d^2P_n(\cos\theta)}{d(\cos\theta)^2}-\frac{dP_n(\cos\theta)}{d(\cos\theta)}\right)\cos\theta. \tag{3}$$

The modal displacement potentials ψ_n must satisfy a transverse wave equation for each mode n of the complete solution of the form

$$\frac{\partial^2\psi_n}{\partial r^2}+\frac{2}{r}\frac{\partial\psi_n}{\partial r}-\frac{n(n+1)}{r^2}\psi_n=\alpha^2\frac{\partial^2\psi_n}{\partial t^2}, \tag{4}$$

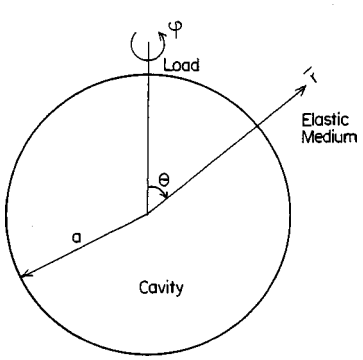


Fig. 1 Geometry of the problem

where $\alpha = c_1/c_2$ and c_2 is the transverse wave speed. An application of the residual variable method reduces by one the order of the derivative with respect to r in Eq. (4). The method is summarized below for the sake of self-sufficiency. Taking the Laplace transform of Eq. (4) we obtain the following modal equation in the Laplace domain:

$$r^2\frac{d^2\tilde{\psi}_n^2}{dr^2}+2r\frac{d\tilde{\psi}_n}{dr}+[i^2\alpha^2r^2s^2-n(n+1)]\tilde{\psi}_n=0, \tag{5}$$

in which $i=\sqrt{-1}$, s is the Laplace parameter, and $\sim$ denotes quantities in the Laplace domain. For $\tilde{\psi}_n$ to be finite at infinity the solution must be the spherical Bessel function of the third kind $h_n^{(1)}(z)$ ([18]). Thus

$$\tilde{\psi}_n=h_n^{(1)}(z)=\frac{1}{z}i^{-n-1}e^{iz}\sum_{k=0}^{n}\Gamma_k^n\left(\frac{i}{z}\right)^k, \tag{6}$$

where $z=i\alpha rs$ and

$$\Gamma_k^n=\frac{(n+k)!}{2^k k!(n-k)!}. \tag{7}$$

An integration constant is not included in Eq. (6) since it is found to be irrelevant. We note that

$$\frac{d\tilde{\psi}_n}{dr}=\frac{dh_n^{(1)}(z)}{dz}\frac{dz}{dr}=i\alpha s\frac{dh_n^{(1)}(z)}{dz}\frac{\tilde{\psi}_n}{h_n^{(1)}}. \tag{8}$$

The series expansions of $h_n^{(1)}(z)$ and its derivative with respect to z can now be substituted into Eq. (8) to yield the modal equation in the modified form

$$r\frac{d\tilde{\psi}_n}{dr}+\alpha rs\tilde{\psi}_n+\tilde{\psi}_n=-\tilde{\psi}_{Rn} \tag{9}$$

in which the residual variable $\tilde{\psi}_{Rn}$ is given by

$$\tilde{\psi}_{Rn}=\left[\frac{\sum_{k=0}^{n}k\Gamma_k^n(\alpha rs)^{n-k}}{\sum_{k=0}^{n}\Gamma_k^n(\alpha rs)^{n-k}}\right]\tilde{\psi}_n. \tag{10}$$

Inversion of Eq. (9) into the time domain gives

$$r\frac{\partial\psi_n}{\partial r}+\alpha r\frac{\partial\psi_n}{\partial t}+(n+1)\psi_n=f_n\left(\frac{t}{\alpha r}\right)*\partial\frac{\psi_n}{\partial t}, \tag{11}$$

with $f_0(t)=0$, $f_n(t)=\Sigma_{k=1}^{n}\exp(a_{nk}t)$, where a_{nk} for $k=1,\ldots,n$ are the n complex roots of $\Sigma_{k=0}^{n}\Gamma_k^nS^{n-k}=0$ given by Akkas [15] and * denotes convolution. The solutions for ψ_n must satisfy the wave Eq. (4), or its modified form (11), and the boundary conditions for $\sigma_{r\varphi}^{(n)}$ defined on the surface of the spherical cavity. In the Laplace domain, the solution is obtained from Eq. (9) as

$$\tilde{\psi}_n=C_n^3(s)\frac{e^{-\alpha rs}}{r^{n+1}}\sum_{k=0}^{n}\Gamma_k^n(\alpha rs)^{n-k}, \tag{12}$$

where $C_n^3(s)$ is an integration constant to be determined from the boundary condition at $r=1$. Obviously, Eq. (12) is equivalent to Eq. (6) as it should be. Accordingly, Eqs. (9) and (11) are the equivalent forms of Eqs. (5) and (4), respectively. This is where the basis of the residual variable method lies; in modifying the classical equations. For our problem,

$$C_n^3(s)=\frac{a}{\mu}\frac{\tilde{\sigma}_{r\varphi}^{(n)}(r=1)e^{\alpha s}}{\sum_{k=0}^{n}(2+k+\alpha s)\Gamma_k^n(\alpha s)^{n-k}}. \tag{13}$$

3 Solution for a Heaviside Load Function

Consider the case in which the cavity surface is subjected to a circumferential step load that can be expanded in a modal series as

$$\varphi_{r\varphi}(r=1,\theta,t)=\sum_{n=1}^{\infty}\sigma^{(n)}\frac{dP_n(\cos\theta)}{d\theta}H[t], \tag{14}$$

in which $\sigma^{(n)}$ are constants and $H[t]$ is the Heaviside function. Having derived the results for this special loading, the results for a more general loading follow by convolution. For the case (14) the solution becomes

$$\psi_n=\frac{a\sigma^{(n)}}{\mu r^{n+1}}\sum_{k=0}^{n+1}\left[\frac{\sum_{m=0}^{n}a_m(\alpha rs_k)^{n-m}}{\sum_{m=0}^{n+1}(n-m+2)c_m(\alpha s_k)^{n-m+1}}\right]\times e^{s_k[t-\alpha(r-1)]}H[t-\alpha(r-1)], \tag{15}$$

$$\frac{\partial\psi_n}{\partial r}=-\frac{a\sigma^{(n)}}{\mu r^{n+2}}\sum_{k=0}^{n+1}\left[\frac{\sum_{m=0}^{n+1}b_m(\alpha rs_k)^{n-m+1}}{\sum_{m=0}^{n+1}(n-m+2)c_m(\alpha s_k)^{n-m+1}}\right]\times e^{s_k[t-\alpha(r-1)]}H[t-\alpha(r-1)], \tag{16}$$

where a_m, b_m, and c_m are given by

$$a_m=\Gamma_m^n, \tag{17}$$

$$b_m=m\Gamma_{m-l}^n+\Gamma_m^n \text{ for } m\leqslant n \text{ and } b_m=m\Gamma_{m-1}^n \text{ for } m=n+1, \tag{18}$$

$$c_0=\Gamma_0^n,\ c_m=(m+1)\Gamma_{m-1}^n+\Gamma_m^n \quad \text{for } 1\leqslant m\leqslant n \text{ and } c_m=(m+1)\Gamma_{m-1}^n \quad \text{for } m=n+1, \tag{19}$$

and αs_k for $k=0,\ldots,n+1$ are the roots of

$$\sum_{m=0}^{n+1}c_m(\alpha s_k)^{n-m+2}=0. \tag{20}$$

With $s_0=0$, these roots, αs_k, for $k=1,\ldots,n+1$, are presented for modes $n=2$ to 9 in Table 1.

The formulas for the modes 1 and 2 are now given explicitly for illustration. When $n=1$ the solution (15) and (16) gives

$$\psi_1=\frac{a\sigma^{(1)}}{\mu r^2}\sum_{k=0}^{2}\frac{\alpha rs_k+1}{3\alpha^2s_k^2+6\alpha s_k+3}e^{s_k[t-\alpha(r-1)]}H[t-\alpha(r-1)], \tag{21}$$

Table 1 The roots αs_k ($k=1,\ldots,n+1$) for modes $n=2$ to 9

$n=2$	-2.0	$n=3$	$-1.549239\pm 2.930104i$
	$-1.5\pm 1.936492i$		$-2.450760\pm 0.906011i$
$n=4$	-3.159404	$n=5$	$-1.652983\pm 4.875444i$
	$-1.602012\pm 3.905351i$		$-3.719141\pm 0.8844119i$
	$-2.818285\pm 1.802312i$		$-3.127875\pm 2.701174i$
$n=6$	-4.428261	$n=7$	$-1.746899\pm 6.813909i$
	$-1.701273\pm 5.844471i$		$-3.635840\pm 4.514667i$
	$-3.396844\pm 3.605214i$		$-5.022866\pm 0.876611i$
	$-4.187751\pm 1.766199i$		$-4.594395\pm 2.651149i$
$n=8$	-5.724764	$n=9$	$-1.831069\pm 8.755852i$
	$-1.790079\pm 7.784305i$		$-4.049276\pm 6.348237i$
	$-3.851759\pm 5.429182i$		$-6.336795\pm 0.873589i$
	$-5.539981\pm 1.753160i$		$-6.000272\pm 2.631762i$
	$-4.955797\pm 3.540231i$		$-5.282587\pm 4.433846i$

$$\frac{\partial \psi_1}{\partial r}=-\frac{a\sigma^{(1)}}{\mu r^3}\sum_{k=0}^{2}\frac{\alpha^2 r^2 s_k^2+2\alpha r s_k+2}{3\alpha^2 s_k^2+6\alpha s_k+3}\times e^{s_k[t-\alpha(r-1)]}H[t-\alpha(r-1)], \tag{22}$$

where $s_0=0$, $\alpha s_1=-(3-i\sqrt{3})/2$ and $\alpha s_2=-(3+i\sqrt{3})/2$, which can be rewritten as

$$\psi_1=\frac{a\sigma^{(1)}}{\mu r^2}\left\{\frac{1}{3}+\frac{1}{75}e^{-3t^*/2\alpha}\left[4(r-6)\cos\left(\frac{\sqrt{3}t^*}{2\alpha}\right)+8(9r-4)\sin\left(\frac{\sqrt{3}t^*}{2\alpha}\right)\right]\right\}H[t^*], \tag{23}$$

$$\frac{\partial \psi_1}{\partial r}=-\frac{a\sigma^{(1)}}{\mu r^3}\left[\frac{1}{3}+\frac{2}{75}e^{-3t^*/2\alpha}(33r^2+4r-24)\cos\left(\frac{\sqrt{3}t^*}{2\alpha}\right)-\frac{16}{75}e^{-3t^*/2\alpha}(7r^2-9r+4)\sin\left(\frac{\sqrt{3}t^*}{2\alpha}\right)\right]H[t^*], \tag{24}$$

where $t^*=t-\alpha(r-1)$. Similarly, when $n=2$,

$$\psi_2=\frac{a\sigma^{(2)}}{\mu r^3}\sum_{k=0}^{3}\frac{\alpha^2 r^2 s_k^2+3\alpha r s_k+3}{4\alpha^3 s_k^3+15\alpha^2 s_k^2+24\alpha s_k+12}\times e^{s_k[t-\alpha(r-1)]}H[t-\alpha(r-1)], \tag{25}$$

$$\frac{\partial \psi_2}{\partial r}=-\frac{a\sigma^{(2)}}{\mu r^4}\sum_{k=0}^{3}\frac{\alpha^3 r^3 s_k^3+4\alpha^2 r^2 s_k^2+9\alpha r s_k+9}{4\alpha^3 s_k^3+15\alpha^2 s_k^2+24\alpha s_k+12}\times e^{s_k[t-\alpha(r-1)]}H[t-\alpha(r-1)], \tag{26}$$

where $s_0=0$, $\alpha s_1=-2$, $\alpha s_2=-(3-i\sqrt{15})/2$ and $\alpha s_3=-(3+i\sqrt{15})/2$, yielding

$$\psi_2=\frac{a\sigma^{(2)}}{\mu r^3}\left[\frac{1}{8}e^{-3t^*/2\alpha}(4r^2-6r+1)\cos\left(\frac{\sqrt{15}t^*}{2\alpha}\right)+\frac{\sqrt{15}}{40}e^{-3t^*/2\alpha}(4r^2+2r-3)\sin\left(\frac{\sqrt{15}t^*}{2\alpha}\right)+\frac{1}{4}-\frac{1}{8}e^{-2t^*/\alpha}(4r^2-6r+3)\right]H[t^*], \tag{27}$$

$$\frac{\partial \psi_2}{\partial r}=-\frac{a\sigma^{(2)}}{\mu r^4}\left[\frac{3}{4}+\frac{1}{8}e^{-2t^*/\alpha}(8r^3-16r^2+18r-9)+\frac{1}{8}e^{-3t^*/2\alpha}(16r^2-18r+3)\cos\left(\frac{\sqrt{15}t^*}{2\alpha}\right)-\frac{\sqrt{15}}{40}e^{-3t^*/2\alpha}(16r^3-16r^2-6r+9)\sin\left(\frac{\sqrt{15}t^*}{2\alpha}\right)\right]H[t^*]. \tag{28}$$

4 Results and Conclusion

The results presented in this section are calculated from the exact, analytical expressions obtained in the previous sections for a modal Heaviside loading $\sigma_{r\varphi}^{(n)}(r=1,t)=\sigma^{(n)}H[t]$ applied on the cavity surface. The realistic data used are for granite rock with $\rho=2660\ \mathrm{kg/m^3}$, $\lambda=8.444\times10^9\ \mathrm{N/m^2}$, $\mu=1.2667\times10^{10}\ \mathrm{N/m^2}$, $c_1=3563$ m/s, $c_2=2182$ m/s.

The nondimensional modal stress components $\sigma_{r\varphi}^{(n)}/\sigma^{(n)}$ as functions of the nondimensional time are presented in Figs. 2, 3, and 4 for modes 1, 2, and 3, respectively, at the locations $r=1.0$, 1.05, 1.1, 1.2, 1.5, 2.0, and 4.0. The results are available for other modes also, but are not presented here. It can be seen that the shear response initially jumps at the arrival of the transverse waves when $t=(r-1)\alpha$. The initial jump at each point in the medium corresponds to the peak value of the response history of that point and is followed by a damped response which tends to a static value at later times.

Figures 5 and 6 show the stress responses at $r=1.1$ and $r=1.2$, respectively, for modes $n=1$ to 14. It is clear that the response histories are characterized by higher frequencies and damping to the static values for the higher modes. Also the static values are lower for the higher modes.

The nondimensional circumferential modal displacement $\mu u_\varphi^{(n)}/a\sigma^{(n)}$ for mode 1 is presented in Fig. 7 for a range of points in the medium. Figure 8 shows the modal circumferential displacements at the surface of the cavity for modes $n=1$ to 14. For

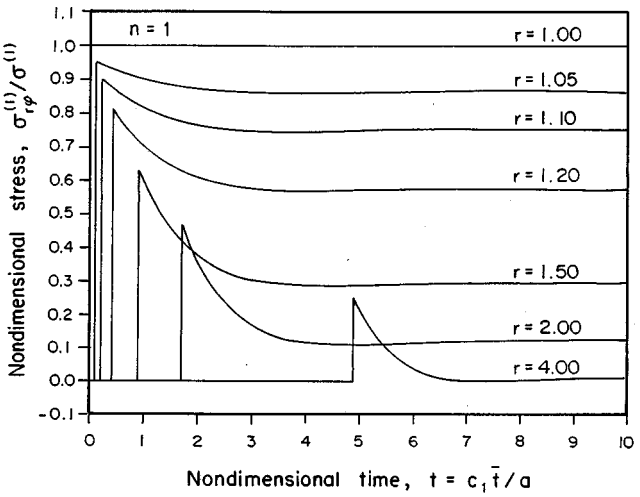


Fig. 2 Circumferential stresses for $n=1$ at various values of r

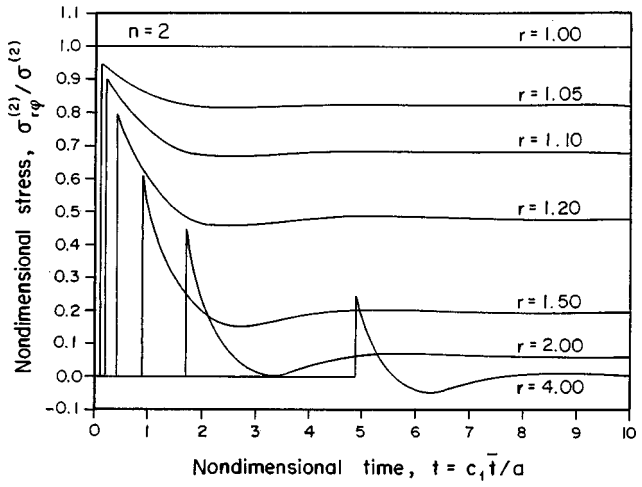


Fig. 3 Circumferential stresses for $n=2$ at various values of r

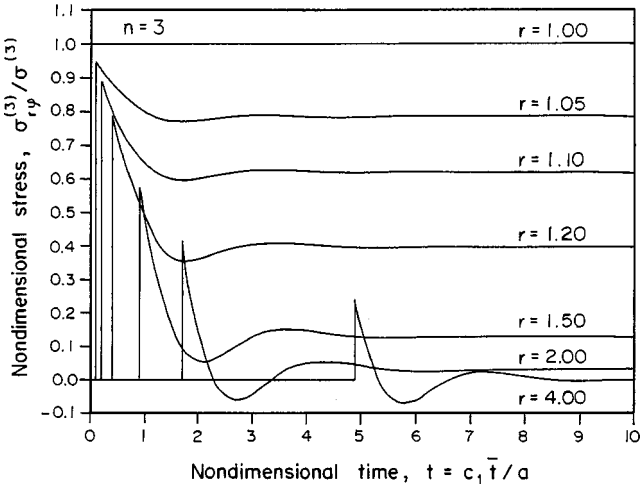


Fig. 4 Circumferential stresses for $n=3$ at various values of r

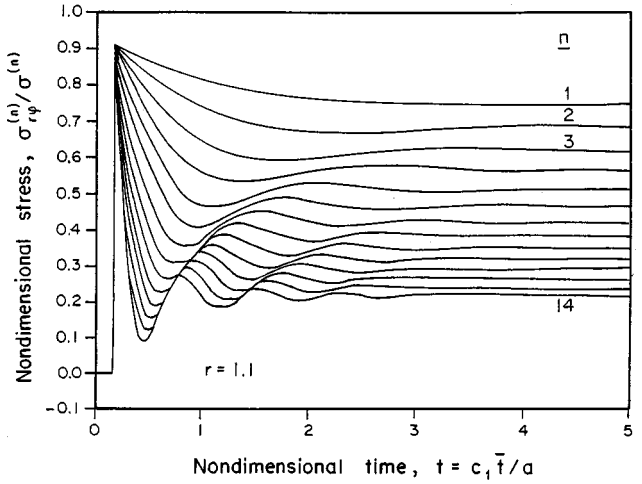


Fig. 5 Circumferential stresses for $n=1$ to 14 at $r=1.1$

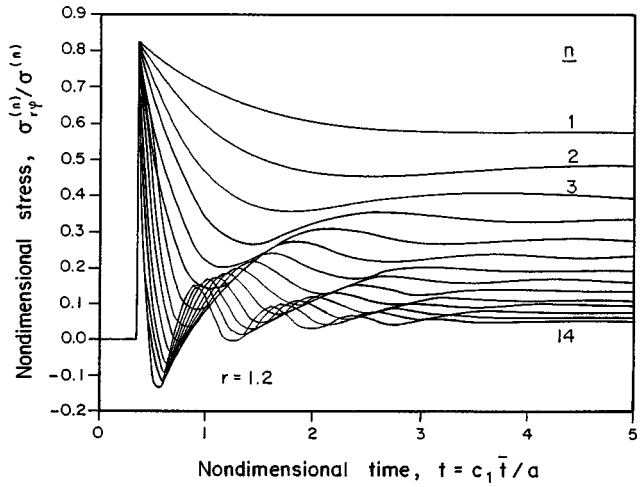


Fig. 6 Circumferential stresses for $n=1$ to 14 at $r=1.2$

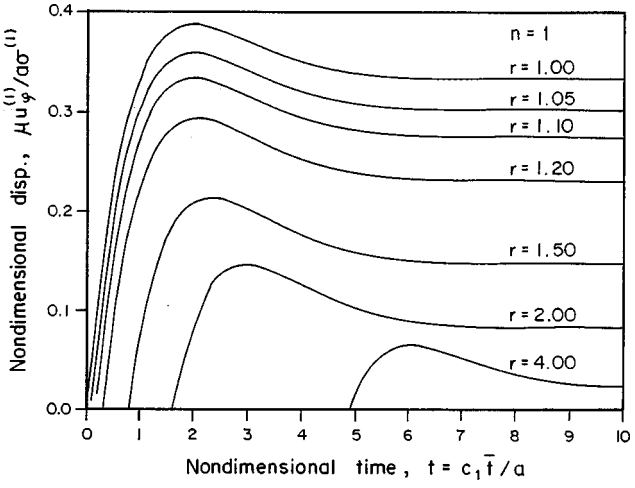


Fig. 7 Circumferential displacements for $n=1$ at various values of r

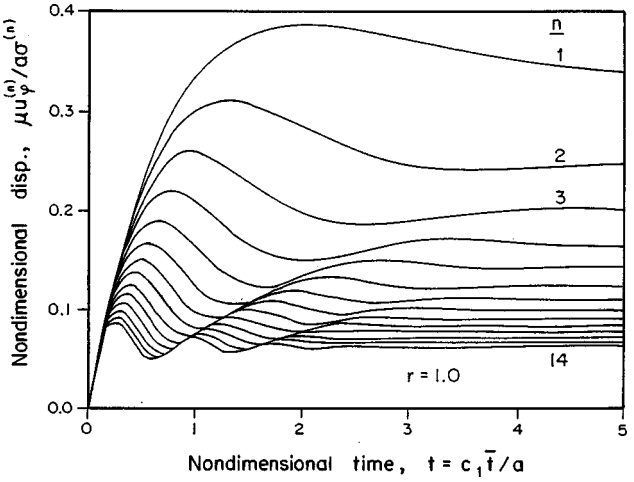


Fig. 8 Circumferential displacements for $n=1$ to 14 at $r=1.0$

each mode the displacement initially increases to its peak value followed by a damped oscillation to its static value. However, higher modes are associated with higher frequencies and larger damping and, moreover, the subsequent static values are smaller in magnitude.

The stress responses as functions of the nondimensional radial distance, r, are depicted in Fig. 9 for mode 1 at the nondimensional times $t=1,2,\ldots,10$. They are characterized by concave curves which initially gradually decrease with distance from their maximum values at the cavity surface $r=1$. Having reached their

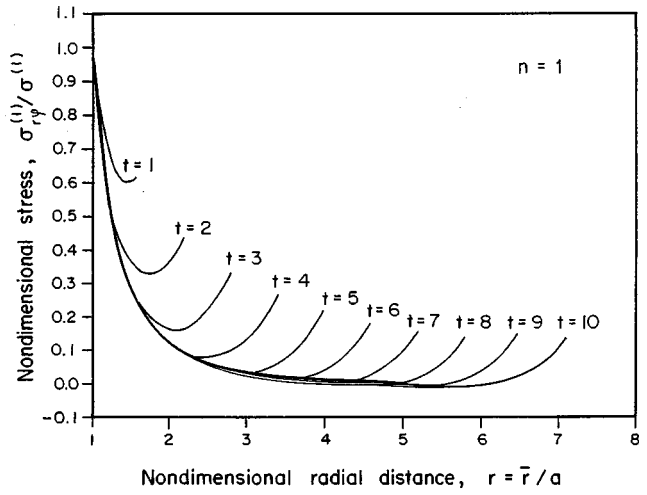


Fig. 9 Circumferential stresses for $n=1$ at various values of t

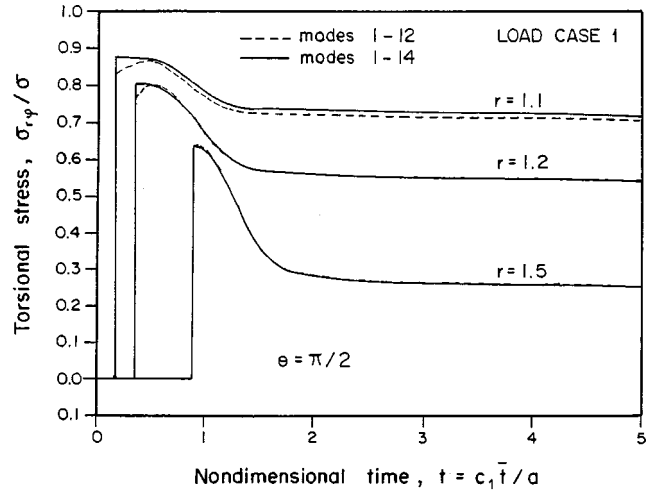


Fig. 10 Torsional stresses for $n=1$ to 12 and $n=1$ to 14 in Load Case 1 at $r=1.1$, 1.2 and 1.5 with $\theta=\pi/2$

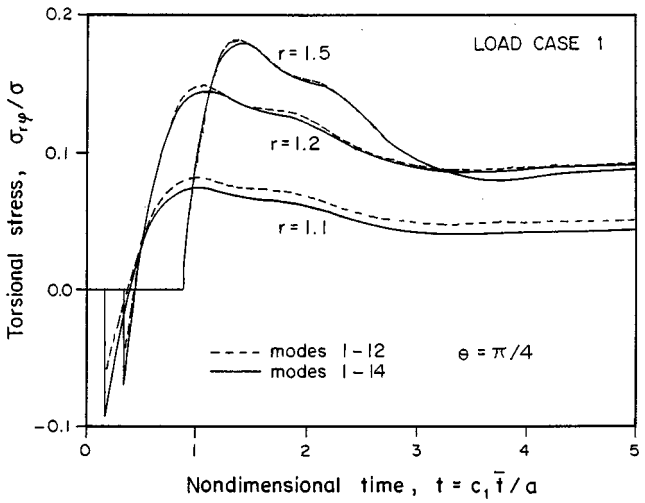


Fig. 11 Torsional stresses for $n=1$ to 12 and $n=1$ to 14 in Load Case 1 at $r=1.1$, 1.2 and 1.5 with $\theta=\pi/4$

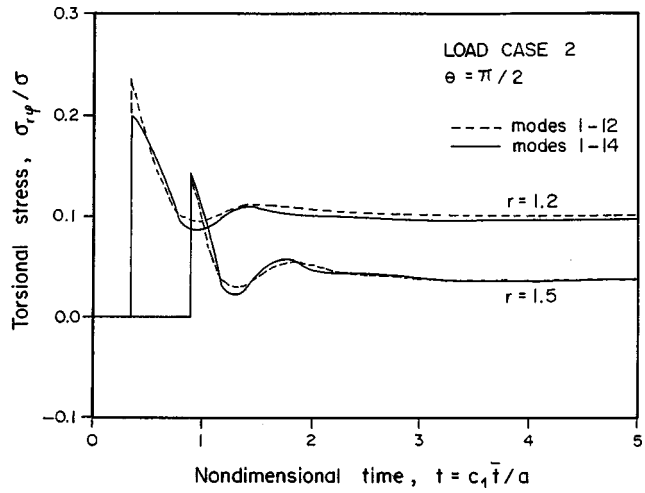


Fig. 12 Torsional stresses for $n=1$ to 12 and $n=1$ to 14 in Load Case 2 at $r=1.2$ and 1.5 with $\theta=\pi/2$

minima, they increase subsequently. The end of each curve coincides with the most distant point which can be reached by the stress wave within the medium at that particular time.

It is also interesting to analyze the response of the medium to surface ''ring loads,'' using the orthogonality properties of the Legendre polynomials and modal superposition. The results for two representative loading situations are presented here. The surface of the spherical cavity is subjected to a torsional stress σ applied first, in Case 1, at $\pi/3 \leqslant \theta \leqslant 2\pi/3$, and secondly, in Case 2, at $\pi/3 \leqslant \theta \leqslant 7\pi/18$. Figures 10, 11, 12, and 13 illustrate the values of the torsional stress responses, $\sigma_{r\phi}/\sigma$, in the medium obtained by superposing the modes 1–12 compared with those using modes 1–14.

Figures 10 and 11 show the stress response as a function of the nondimensional time t in Case 1 with $r=1.1$, 1.2, and 1.5 at $\theta=\pi/2$ and $\theta=\pi/4$, respectively. The curves in Fig. 10 are seen to initially increase to their peaks, which are achieved upon the arrival of the stress wave, and to be then gradually damped to their static values. The corresponding curves in Fig. 11 for the stress at $\theta=\pi/4$ exhibit sudden jumps followed by oscillatory damping to static values. The analogous responses for Case 2 at $r=1.2$ and 1.5 are presented in Figs. 12 and 13. Initial jumps are shown to be followed by damped oscillations.

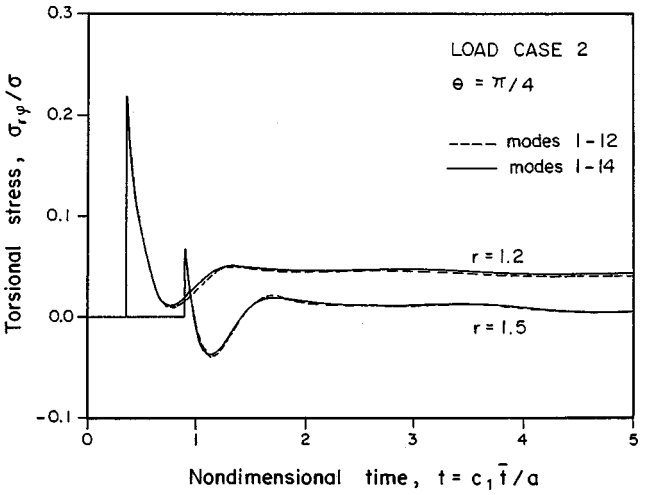


Fig. 13 Torsional stresses for $n=1$ to 12 and $n=1$ to 14 in Load Case 2 at $r=1.2$ and 1.5 with $\theta=\pi/4$

It is apparent from the results for modes 1–12 as opposed to modes 1–14 in Figs. 10–13 that the convergence of the solution at a particular point in the medium depends not only on its location but also upon the area of the cavity surface over which the load is applied. Generally, better convergence is achieved at points more distant from the cavity surface. The narrower the area over which the surface load is applied, the more modes are required to improve the convergence of the solution.

In conclusion, it can be summarized that here the residual variable method is employed to eliminate the second derivative of the displacement potential with respect to the radial coordinate from the wave equation. Integration of the resulting differential equation in the Laplace domain yields the displacement potential with an integration constant which is determined from the boundary condition. An exact closed-form expression for the potential is obtained which leads to explicit modal relations for the displacement and the stress components representing the dynamic response of the medium at any position and time. Numerical results are presented in graphical form for circumferential Heaviside loads acting on the surface of a spherical cavity within granite rock.

References

[1] Achenbach, J. D., 1972, *Wave Propagation in Elastic Solids*, North-Holland, Amsterdam.

[2] Eringen, A. C., and Suhubi, E. S., 1974–1975, *Elastodynamics*, Vols. 1 and 2, Academic Press, New York.

[3] Miklowitz, J., 1978, *The Theory of Elastic Waves and Waveguides*, North-Holland, Amsterdam.

[4] Gaunaurd, G. C., 1989, ''Elastic and Acoustic Resonance Wave Scattering,'' Appl. Mech. Rev., **42**, pp. 143–193.

[5] Sato, Y., Usami, T., and Ewing, M., 1962, ''Basic Study on the Oscillation of a Homogeneous Elastic Sphere, Part IV—Propagation of Disturbances on the Sphere,'' Geophysics, **31**, pp. 237–241.

[6] Godin, Y. A., 1995, ''An Exact Solution to a Problem of Axisymmetric Torsion of an Elastic Space With a Spherical Crack,'' Q. Appl. Math., **LIII**, pp. 679–682.

[7] Chadwick, P., and Trowbridge, E. A., 1967, ''Oscillations of a Rigid Sphere Embedded in an Infinite Elastic Solid, Part I—Torsional Oscillations,'' Proc. Cambridge Philos. Soc., **63**, pp. 1189–1205.

[8] Chadwick, P., and Trowbridge, E. A., 1967, ''Oscillations of a Rigid Sphere Embedded in an Infinite Elastic Solid, Part II—Rectilinear Oscillations,'' Proc. Cambridge Philos. Soc., **63**, pp. 1207–1227.

[9] Chadwick, P., and Trowbridge, E. A., 1967, ''Elastic Wave Fields Generated by Scalar Wave Functions,'' Proc. Cambridge Philos. Soc., **63**, pp. 1177–1187.

[10] Tupholme, G. E., 1983, ''Elastic Pulse Generation by Tractions Applied to a Spherical Cavity,'' Appl. Sci. Res., **40**, pp. 299–325.

[11] Tupholme, G. E., 1967, ''Generation of an Axisymmetrical Acoustic Pulse by a Deformable Sphere,'' Proc. Cambridge Philos. Soc., **63**, pp. 1285–1308.

[12] Chadwick, P., and Johnson, A. F., 1969, ''Torsional Oscillations of a Rigid Convex Inclusion Embedded in an Elastic Solid,'' J. Inst. Math. Appl., **5**, pp. 283–307.

[13] Geers, T. L., 1969, ''Excitation of an Elastic Cylindrical Shell by a Transient Acoustic Wave,'' ASME J. Appl. Mech., **36**, pp. 459–469.

[14] Akkas, N., 1979, ''Residual Potential Method in Spherical Coordinates and Related Approximations,'' Mech. Res. Commun., **6**, pp. 257–262.

[15] Akkas, N., 1985, ''The Residual Variable Method and its Applications,'' Acta Mech., **55**, pp. 203–217.

[16] Akkas, N., and Zakout, U., 1997, ''Transient Response of an Infinite Elastic Medium Containing a Spherical Cavity With and Without a Shell Embedment,'' Int. J. Eng. Sci., **35**, pp. 89–112.

[17] Zakout, U., and Akkas, N., 1997, ''Transient Response of a Cylindrical Cavity With and Without a Bonded Shell in an Infinite Elastic Medium,'' Int. J. Eng. Sci., **35**, pp. 1203–1220.

[18] Abramowitz, M., and Stegun, I. A., eds., 1965, *Handbook of Mathematical Functions*, Dover, New York.

Yield Functions and Flow Rules for Porous Pressure-Dependent Strain-Hardening Polymeric Materials

J. H. Lee
Mem. ASME

J. Oung
Graduate Student,

Department of Mechanical Engineering,
University of Alaska,
Fairbanks, AK 99775-5905

To characterize the response of progressively damaged glassy polymers due to the presence and evolution of voids, yield functions and flow rules were developed systematically for a pressure-dependent matrix following the modified von Mises criterion. A rigid-perfectly plastic material was first assumed. The upper bound method was used with a velocity field which has volume preserving and shape changing portions. Macroscopic yield criterion in analytical closed form was first obtained for spherical voids which is valid for all possible macroscopic strain rate fields. Macroscopic yield criteria in analytical closed form were then obtained for cylindrical voids for the special cases of axisymmetric and plane-strain modes of deformation. The upper-bound solutions were subsequently improved to better match analytical solutions for pure hydrostatic loading. Characteristics of the yield function as a function of pressure dependency and void fraction were studied in detail. Generalization of the model for spherical voids to include elasticity as well as strain hardening of the matrix was then obtained. An example for the uniaxial response of a progressively damaged material was then used to illustrate one possible application of the full set of constitutive equations. [S0021-8936(00)02902-0]

1 Introduction

One important mechanism of failure for glassy polymers is the nucleation, growth, and coalescence of voids leading to crazing, crack propagation, and subsequent failure ([1–3]). Similar mechanisms leading to ductile failure for metals have been widely studied ([4]) which was greatly facilitated by the development of the so-called Gurson's model ([5]). Gurson's model represents the plastic behavior of void-containing metals using a (porous) yield function and related flow rules. Material properties of glassy polymers are, however, very much different from those of metals. For example, unlike metals which are generally pressure independent, polymers are influenced much more by the hydrostatic pressure ([6]). The pressure dependency of yielding gives rise to different yield strengths in compression from tension, and is known as the SD (strength-differential) effect. In addition, polymers are more rate sensitive than metals and have different hardening-softening responses from metals. As a first step toward understanding the failure behavior of glassy polymers, it is the objective of this paper to develop (porous) yield criteria and related flow rules for void-containing polymers with a pressure-dependent matrix which follows the modified von Mises criterion ([7]). We focus specifically on the growth stage of the failure process. During the growth of voids for polymeric materials, the shape is generally spheroidal. Spherical voids represent one special case of the spheroidal voids whereas cylindrical ones represent another limit. Analytical yield criteria for spherical and cylindrical voids were obtained systematically using the upper bound approach ([8]) for rigid-perfectly plastic materials, which were subsequently improved such that they show an excellent match with analytical solutions of voids under pure hydrostatic pressure. Flow rules, incorporating strain hardening, were then developed for spherical voids. Lower case Greek letters indicate microscopic (matrix) properties whereas upper case Greek letters indicate macroscopic (aggregate) properties. Einstein summation convention is assumed. Most of the analytical and numerical calculations were greatly simplified by using Mathematica ([9]). More details of this work can be found in Oung [10].

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 15, 1998; final revision, June 22, 1999. Associate Technical Editor: A. K. Mal. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

2 Constitutive Equations and Upper Bound Theorem

The modified von Mises yield criterion has been found applicable for many glassy polymers (for example, [11,7,12]). The classical upper bound theorem ([8]) has been extended for pressure-dependent materials for soils ([13]) and polymers ([14]).

Modified von Mises Criterion. The modified von Mises criterion can be expressed as

$$F=\bar{\sigma}^2-CT+(C-T)\sigma_{kk}=0, \tag{1}$$

where $\bar{\sigma}=\sqrt{(3/2)\sigma'_{ij}\sigma'_{ij}}$ is the equivalent stress. σ'_{ij} is the stress deviator. $\sigma_{kk}=\delta_{ij}\sigma_{ij}$ is the hydrostatic stress where δ_{ij} is the Kronecker delta. C and T are the absolute yield strengths in compression and tension, respectively. We will assume the associated flow rule such that the plastic strain rates ($\dot{\epsilon}_{ij}$) are

$$\dot{\epsilon}_{ij}=\lambda\frac{\partial F}{\partial\sigma_{ij}}, \tag{2}$$

where λ is a proportional parameter which can be found (using Eqs. (1) and (2)) as

$$\lambda=\frac{\dot{\bar{\epsilon}}}{2\bar{\sigma}}, \tag{3}$$

where $\dot{\bar{\epsilon}}=\sqrt{(2/3)\dot{\epsilon}'_{ij}\dot{\epsilon}'_{ij}}$ is the equivalent plastic strain rate. $\dot{\epsilon}_{ij}=1/2(v_{i,j}+v_{j,i})$ where v_i is the velocity. Substituting Eq. (3) into Eq. (2) and using Eq. (1) we have

$$\dot{\epsilon}'_{ij}=\frac{3\dot{\bar{\epsilon}}}{2\bar{\sigma}}\sigma'_{ij}, \tag{4}$$

$$\dot{\epsilon}_{kk}=\frac{3}{2}\frac{\dot{\bar{\epsilon}}}{\bar{\sigma}}(C-T). \tag{5}$$

Anticipating applying the upper bound theorem, we express stresses in terms of strain rates by rearranging Eq. (5), such that

$$\bar{\sigma}=\frac{3}{2}\frac{\dot{\bar{\epsilon}}}{\dot{\epsilon}_{kk}}(C-T). \tag{6}$$

In addition, combining Eqs. (1) and (6) we have

$$\sigma_{kk}=\frac{CT}{C-T}-\frac{9}{4}\frac{\dot{\bar{\epsilon}}^2}{\dot{\epsilon}_{kk}^2}(C-T). \tag{7}$$

Using Eqs. (6) and (7), the internal plastic work density rate is

$$\dot{w}=\sigma_{ij}\dot{\epsilon}_{ij}=\frac{3}{4}\frac{\dot{\bar{\epsilon}}^2}{\dot{\epsilon}_{kk}}(C-T)+\frac{CT}{3(C-T)}\dot{\epsilon}_{kk}. \tag{8}$$

Upper Bound Theorem. The upper bound theorem for finding porous yield criteria has been discussed by Gurson [5]. The desirable velocity field is found by the minimization of the rate of plastic work done,

$$\dot{W}=\frac{1}{V}\int_V \sigma_{ij}\dot{\epsilon}_{ij}dV, \tag{9}$$

where V is the volume of a unit cell. The macroscopic stress is related to the microscopic stress and strain rate via

$$\Sigma_{ij}=\frac{\partial \dot{W}}{\partial \dot{E}_{ij}}=\frac{1}{V}\int_V \sigma_{kl}\frac{\partial \dot{\epsilon}_{kl}}{\partial \dot{E}_{ij}}dV, \tag{10}$$

where Σ_{ij} and $\dot{E}_{ij}$ are macroscopic stresses and (plastic) strain rates, respectively. It should be noted that normality of the matrix is invoked in deriving Eq. (10) such that there is a single term in the integral which simplifies the relationship. When a rigid plastic state is assumed, the microscopic velocity field at the surface (S) of the unit cell is connected to the macroscopic plastic strain rate in Cartesian coordinates as

$$v_i|_S=\dot{E}_{ik}x_k|_S. \tag{11}$$

3 The Rigid-Plastic Porous Model for Spherical Voids

For spherical voids, the unit cell is a spherical shell of inner radius a and outer radius b. The assumed velocity field is obtained following the examples of Lee [14] and Gurson [5]. Although various velocity fields can be constructed, we will seek only *simple* ones that would lead to *analytical* yield functions. The aim is relating the microscopic velocity field to the macroscopic quantities. For spherical voids, the velocity field can be conveniently separated into two parts,

$$v_i=v_i^s+v_i^v, \tag{12}$$

where v_i^s involves shape changes at constant volume and v_i^v involves volume changes at constant shape. We will first construct v_i^v. Using spherical coordinates ($0\leqslant\theta\leqslant\pi$, $0\leqslant\phi\leqslant 2\pi$), $v_\theta^v=v_\phi^v=0$. Matrix compressibility indicates

$$2\frac{v_r^v}{r}+\frac{\partial v_r^v}{\partial r}=\dot{\epsilon}_{kk}\equiv a_0 \tag{13}$$

where a_0 represents the volumetric strain rate. The simplest velocity field is obtained (cf. [14]) by assuming a_0 as constant. Using Eqs. (13) and (11), the volumetric velocity field becomes

$$v_r^v=\frac{1}{3}(\dot{E}_{kk}-a_0)\frac{b^3}{r^2}+a_0\frac{r}{3}. \tag{14}$$

When pressure dependency is not involved ($a_0=0$), the above equation reduces to that of Gurson's ([5]). The deviatoric velocity field of Gurson is used here,

$$v_i^s=\dot{E}'_{ij}x_j, \tag{15}$$

in Cartesian coordinates. Combining Eqs. (14) and (15), the strain rate field becomes

$$\dot{\epsilon}_{ij}=\dot{\epsilon}_{ij}^s+\dot{\epsilon}_{ij}^v=\dot{E}'_{ij}+\frac{1}{3}(\dot{E}_{kk}-a_0)h_{ij}+\frac{1}{3}a_0\delta_{ij}, \tag{16}$$

where $h_{rr}=-2(b/r)^3=-2h_{\theta\theta}=-2h_{\phi\phi}$, $h_{ij}|_{i\neq j}=0$. h_{ij} will be transformed into the Cartesian coordinates for subsequent calculations such that all terms in Eq. (16) refer to the same axes. The integration of any volume-averaged quantity $\langle\rangle$ takes the following form:

$$\frac{1}{V}\int_{r=a}^{r=b}\int_{\theta=0}^{\theta=\pi}\int_{\phi=0}^{\phi=2\pi}\langle\rangle r^2\sin\theta dr d\theta d\phi, \tag{17}$$

where $V=(4/3)\pi b^3$, and the void fraction f by definition is $(a/b)^3$. The yet undetermined a_0 is found by minimizing $\dot{W}$ (using Eqs. (8), (9), and (16)) with respect to a_0,

$$\frac{\partial \dot{W}}{\partial a_0}=0, \tag{18}$$

solving for a_0 and taking the positive root (for associated plasticity, a_0 is always positive), a_0 becomes

$$a_0=\left(\frac{(C-T)^2\left(\dot{E}_{kk}^2+\frac{9}{4}f\dot{\bar{E}}^2\right)}{(C-T)^2+fCT}\right)^{1/2}, \tag{19}$$

where $\dot{\bar{E}}'=(\frac{2}{3}\dot{E}'_{ij}\dot{E}'_{ij})^{1/2}$ is the macroscopic equivalent plastic strain rate and $\dot{E}_{kk}=\delta_{ij}\dot{E}_{ij}$ is the macroscopic volumetric plastic strain rate.

Using Eqs. (10), (16), and (18), the macroscopic stresses are related to the microscopic stresses as

$$\Sigma'_{ij}=\frac{1}{V}\int_V \sigma'_{ij}dV, \quad \Sigma_{nn}=\frac{1}{V}\int_V \sigma'_{kl}h_{kl}dV, \tag{20}$$

which are identical to Eq. (4.7) of Gurson [5]. This identity is a consequence of Eq. (18). In other words, as long as the minimization process is carried out, Eq. (20) is valid for the pressure-independent as well as the pressure-dependent matrices.

Yield Criterion. Using Eqs. (4), (6), (16), and (20), we find the following macroscopic stress invariants:

$$\Sigma_{eq}=\left(\frac{3}{2}\Sigma'_{ij}\Sigma'_{ij}\right)^{1/2}=\frac{3\dot{\bar{E}}(1-f)(C-T)}{2a_0}, \tag{21}$$

$$\Sigma_{nn}=\frac{2(1-f)(C-T)(\dot{E}_{kk}-a_0)}{a_0 f}. \tag{22}$$

To find the yield criterion, a_0 from Eq. (19) is substituted into Eqs. (21) and (22). Σ_{eq} and Σ_{nn} can be expressed as a function of $(\dot{E}_{kk}/\dot{\bar{E}})$ which can be eliminated. We then find the surprisingly simple yield criterion as

$$\Phi=\Sigma_{eq}^2+\frac{f}{4}\Sigma_{nn}^2+(C-T)(1-f)\Sigma_{nn}-(1-f)^2CT=0. \tag{23}$$

However, this yield criterion behaves poorly for large Σ_{nn} when compared with the analytical solution described in the next section. It was brought to our attention that in the limit of $C=T$, Eq. (23) is the same as the two-term series expansion of Gurson's model:

$$\Phi=\Sigma_{eq}^2+2fT^2\cosh\left(\frac{\Sigma_{nn}}{2T}\right)-(1+f^2)T^2=0. \tag{24}$$

We then seek an improved solution as follows. Recognizing Taylor's expansion of $\cosh x \doteq 1+\frac{1}{2}x^2$, the yield criterion can be modified as

$$\Phi=\Sigma_{eq}^2+2fCT\cosh\left(\frac{\Sigma_{nn}}{2\sqrt{CT}}\right)+(C-T)(1-f)\Sigma_{nn}-(1+f^2)CT=0. \tag{25}$$

When $C=T$, it is identical to Gurson's yield criterion for spherical voids (Eq. (24)).

Results. The yield function Φ (Eq. (25)) reduces to that of the matrix when $f=0$, and to that of Gurson's when $C=T$. The normalized yield function (Eq. (25)) for $C/T=1.1$ at various f is shown in Fig. 1. Although the yield surface of the matrix ($f=0$) is open in the direction of negative σ_{nn}, the porous yield surface is closed. As the yield surface of the matrix is not symmetric with respect to $\Sigma_{nn}=0$, the porous yield surface is also asymmetric. The porous yield surface becomes smaller as f increases and falls within the envelope of the yield surface of the matrix. Figure 2 shows the porous yield surface, at an $f=0.1$, as a function of pressure dependence (C/T).

For rigid-plastic situation, an exact solution for purely hydrostatic loading can be found by using the equilibrium equation and Eq. (1). The solutions satisfy

$$-\frac{1}{c}\left(\sqrt{b+cx}+a\ln[-a+\sqrt{b+cx}]\right)\Big|_{x=0}^{x=\Sigma_{nn}/3}=\frac{1}{3}\ln f,\quad \Sigma_{nn}\geqslant 0; \tag{26}$$

$$\frac{1}{c}\left(\sqrt{b+cx}-a\ln[a+\sqrt{b+cx}]\right)\Big|_{x=0}^{x=\Sigma_{nn}/3}=\frac{1}{3}\ln f,\quad \Sigma_{nn}<0 \tag{27}$$

where $a=C-T$, $b=C^2-CT+T^2$ and $c=3(T-C)$. Figure 3 shows the comparison of the exact solution with that of the upper bound (Eq. (23)) and the modified upper bound (Eq. (25)) solutions for f versus the hydrostatic stress. It should be noted that the

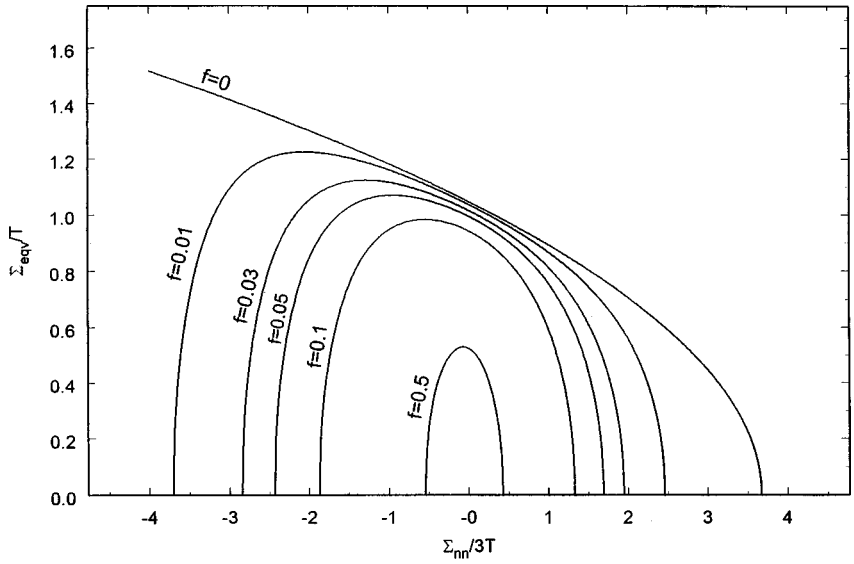


Fig. 1 Yield surfaces of spherical voids $C/T=1.1$

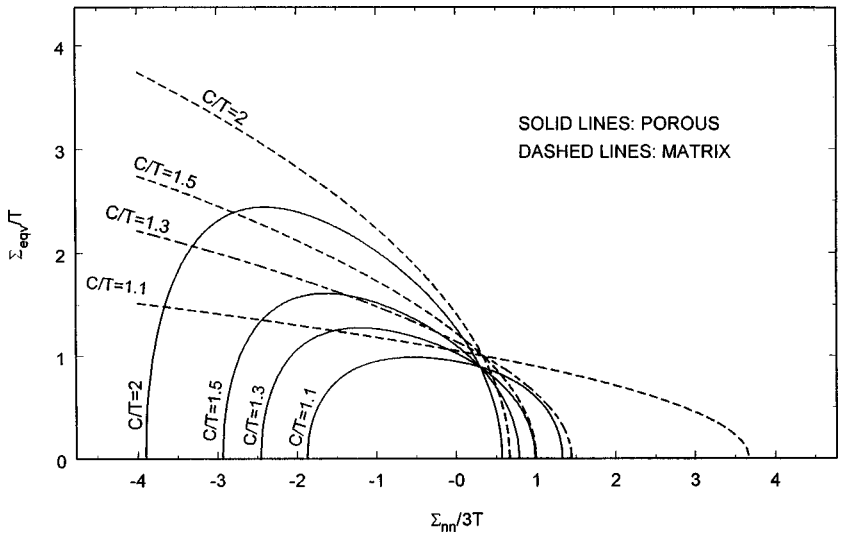


Fig. 2 Comparison of yield surfaces ($f=0.1$) with matrix yield surfaces for different C/T

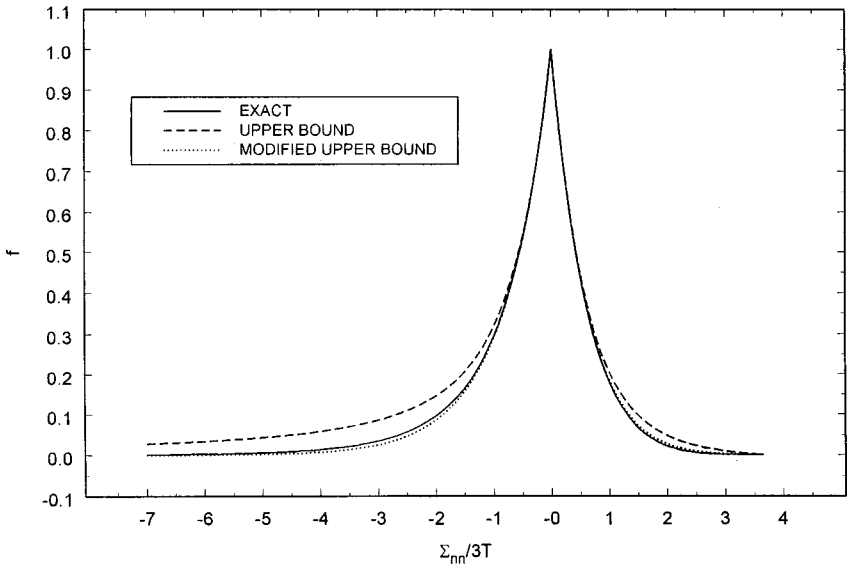


Fig. 3 Comparison of exact solutions with the upper bound and modified upper bound yield surface ($C/T=1.1$) for purely hydrostatic loading

term modified upper bound refers to a modified form of the yield function but itself is not an upper bound solution. As the exact solution falls within the envelope of that from Eq. (23), the upper bound nature of Eq. (23) is clearly seen. The modified upper bound (Eq. (25)) indicates that the result is improved significantly over the original one. It should be noted that the exact solution is valid for both associated and nonassociated flow rules.

Figure 4 shows the significant effect of pressure dependency on yielding in that the yield surfaces for a pressure-dependent matrix are generally smaller than those for a pressure-independent one. The difference becomes smaller when f reaches a high value (such as 0.5).

From our formulation, in the limit of $C=T$ (or when $a_0=0$), the velocity field is identical to that used by Gurson. However, Eq. (23) in the limit does not reduce to that of Gurson's yield criterion. A detailed examination reveals that this difference is due to the assumption of $a_0=$constant. In particular, for von Mises criterion, $\bar{\sigma}$ is a constant. In view of Eq. (6), we can also write

$$\lim_{C->T} \bar{\sigma}=T. \tag{28}$$

With the assumption of $a_0=$constant, Eq. (28) is no longer valid which contributes to the difference in Eq. (23) from Gurson's model in the limit of $C=T$. In particular, with $a_0=$constant, Eq. (5) indicates that

$$\frac{\bar{\sigma}}{\bar{\dot{\epsilon}}}=\text{constant}, \tag{29}$$

which is the equation of a viscous material.

4 Rigid-Plastic Porous Model for Cylindrical Voids

The unit cell for cylindrical voids is a cylinder of inner radius a, outer radius b, and length L. The rectangular coordinate 3 is in the same direction as z, with 1 and 2 directions in the $r-\theta$ plane. The velocity field is constructed following the same principle as that of

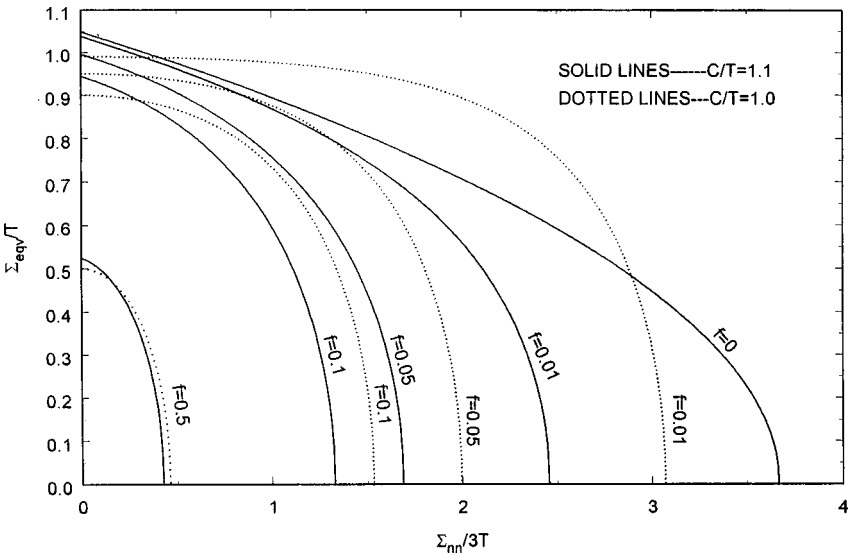


Fig. 4 Comparison of porous yield surfaces with modified von Mises and von Mises matrices

the spherical voids. However, due to the nature of transverse isotropy, the velocity field for cylindrical voids is more complicated. We will first obtain the velocity field affected by the pressure dependency of the matrix (v_i^d), i.e., the velocity field which is different from that of Gurson's. The velocity field which is common to that of Gurson's will then be added later on. Matrix compressibility requires

$$\frac{\partial v_r^d}{\partial r}+\frac{v_r^d}{r}+\frac{\partial v_z^d}{\partial z}=\dot{\epsilon}_{kk}\equiv a_0. \tag{30}$$

Assuming a_0 as constant, $v_r^d(r)$, $v_z^d(z)$, one has

$$\frac{\partial v_r^d}{\partial r}+\frac{v_r^d}{r}=c_2; \quad \frac{\partial v_z^d}{\partial z}=c_3; \quad c_2+c_3=a_0. \tag{31}$$

From Eq. (31*a*), $v_r^d=c_1/r+c_2 r/2$. Applying the boundary condition $v_z^d|_S=\dot{E}_{33}z$ leads to $c_3=\dot{E}_{33}$ such that $c_2=a_0-\dot{E}_{33}$. The boundary condition for the normal strain in the radial direction, $v_r^d|_S=\frac{1}{2}(\dot{E}_{22}+\dot{E}_{11})r|_{r=b}$, leads to $c_1=\frac{1}{2}(\dot{E}_{kk}-a_0)b^2$. One has then

$$v_r^d=\frac{b^2}{2r}(\dot{E}_{kk}-a_0)+\frac{r}{2}(a_0-\dot{E}_{33}); \quad v_z^d=\dot{E}_{33}z. \tag{32}$$

Adding the part of the velocity field which is common to both Gurson and the current model, and using conventions of Gurson [5], one has finally

$$v_r=\frac{b^2}{2r}(\dot{E}_{kk}-a_0)+\frac{r}{2}(a_0-\dot{E}_{33}) +\left(c_4r^3+c_5r+c_6\frac{1}{r}+c_7\frac{1}{r^3}\right)\cos 2\theta+V_{32}z\cos\gamma, \tag{33}$$

$$v_\theta=\left(-2c_4r^3-c_5r+c_7\frac{1}{r^3}\right)\sin 2\theta-V_{32}z\sin\gamma, \tag{34}$$

$$v_z=\dot{E}_{33}z+\left(c_8r+c_9\frac{1}{r}\right)\cos\gamma. \tag{35}$$

The velocity field on the outer surface can be expressed in terms of macroscopic quantities,

$$v_r=\dot{E}'b\cos 2\theta+\frac{1}{2}(\dot{E}_{22}+\dot{E}_{11})b+V_{32}z\cos\gamma, \tag{36}$$

$$v_z=V_{23}b\cos\gamma+\dot{E}_{33}z, \tag{37}$$

where $\dot{E}'\equiv\frac{1}{2}(\dot{E}_{22}-\dot{E}_{11})$, V_{32} is the shear velocity per unit axial length on the (3) plane and V_{23} is the normalized shear velocity parallel to the (3) axis. The condition of zero shear stress-zero shear strain rate on the void surface is

$$\dot{\epsilon}_{r\theta}|_{r=a}=\dot{\epsilon}_{rz}|_{r=a}=0. \tag{38}$$

Equations (36), (37), and (38) are four equations in the six unknowns c_4 through c_9. The remaining two equations are obtained by setting $c_4=0$ ([5]) and determining c_5 by minimizing the dissipation.

Once the velocity field is constructed, we will follow the same procedure as that of the spherical voids. The optimized a_0 is

$$a_0=(C-T)\times\sqrt{\frac{3\left[\dot{E}_{kk}^2+f\left(4C_{EP}\dot{E}'^2+\frac{4}{1+f}(\dot{E}_{13}^2+\dot{E}_{23}^2)+3\dot{E}_{33}\right)\right]}{3(C-T)^2+f(C+T)^2}}, \tag{39}$$

where

$$C_{EP}=\frac{3(1-f)^2(4+f)}{4(1+f)(3+3f-3f^2+f^3)}.$$

Yield Criterion. Using Eq. (10), the macroscopic stress invariants are

$$\Sigma_{eq}=\frac{(1-f)(C-T)}{2a_0}\times\sqrt{(a_0-3\dot{E}_{33})^2+12C_{EP}^2\dot{E}'^2+\frac{48}{1+f}(\dot{E}_{13}^2+\dot{E}_{23}^2)}, \tag{40}$$

$$\Sigma_{\gamma\gamma}=\Sigma_{11}+\Sigma_{22}=\frac{(1-f)(C-T)}{a_0 f}(\dot{E}_{kk}-a_0). \tag{41}$$

We now consider two special cases. The first is the axisymmetric case ($\dot{E}_{ij}|_{i\neq j}=0$, $\dot{E}_{11}=\dot{E}_{22}$). The procedure in the elimination of the strain rate components for spherical voids is also used for the cylindrical voids. The yield surface obtained is

$$\Phi=\Sigma_{eq}^2+\frac{3}{4}f\Sigma_{\gamma\gamma}^2+(1-f)(C-T)\left(\frac{3}{2}\Sigma_{\gamma\gamma}\mp\Sigma_{eq}\right) -(1-f)^2CT=0, \tag{42}$$

where the negative sign is used when $\dot{E}_{33}\leqslant 0$, or when $\dot{E}_{33}>0$ and $(\dot{E}_{kk}/\dot{E}_{33})^2\geqslant 9+12fCT/(C-T)$; and the positive sign is used when $\dot{E}_{33}>0$ and $(\dot{E}_{kk}/\dot{E}_{33})^2<9+12fCT/(C-T)$. Following the procedure for the spherical voids, the yield criterion can be modified as

$$\Phi=\Sigma_{eq}^2+2fCT\cosh\left(\frac{\sqrt{3}\Sigma_{\gamma\gamma}}{2\sqrt{CT}}\right)+(1-f)(C-T)\left(\frac{3}{2}\Sigma_{\gamma\gamma}\mp\Sigma_{eq}\right) -(1+f^2)CT=0. \tag{43}$$

When $C=T$, it becomes

$$\Phi=\Sigma_{eq}^2+2fT^2\cosh\left(\frac{\sqrt{3}\Sigma_{\gamma\gamma}}{2T}\right)-(1+f^2)T^2=0, \tag{44}$$

which is identical to Gurson's yield surface for axisymmetric cylindrical voids.

The second case of cylindrical void is plane strain ($\dot{E}_{ij}|_{i\neq j}=0$, $\dot{E}_{33}=0$). The yield surface becomes

$$\Phi=\frac{\Sigma_{eq}^2}{C_{EP}}+\frac{3f}{4}\Sigma_{\gamma\gamma}^2+\frac{3}{2}(1-f)(C-T)\Sigma_{\gamma\gamma} -\frac{(1-f)^2}{4}\left(\frac{(C-T)^2}{C_{EP}}+(C+T)^2\right)=0. \tag{45}$$

The modified yield criterion is

$$\Phi=\frac{\Sigma_{eq}^2}{C_{EP}}+2fCT\cosh\left(\frac{\sqrt{3}\Sigma_{\gamma\gamma}}{2\sqrt{CT}}\right)+\frac{3}{2}(1-f)(C-T)\Sigma_{\gamma\gamma} -\frac{(1-f)^2}{4}\left(\frac{(C-T)^2}{C_{EP}}+(C+T)^2\right)-2fCT=0. \tag{46}$$

When $C=T$, it is slightly different from Gurson's yield criterion and is

$$\Phi=\frac{\Sigma_{eq}^2}{C_{EP}}+2fT^2\cosh\left(\frac{\sqrt{3}\Sigma_{\gamma\gamma}}{2T}\right)-(1+f^2)T^2=0. \tag{47}$$

For comparison, we record Gurson's yield criterion below:

$$\Phi=C_{eq}\Sigma_{eq}^2+2fT^2\cosh\left(\frac{\sqrt{3}\Sigma_{\gamma\gamma}}{2T}\right)-(1+f^2)T^2=0, \tag{48}$$

where $C_{eq}=(1+3f+24f^6)^2$. It should be noted that the coefficient C_{eq} in Eq. (48) comes from numerical approximation whereas C_{EP} in Eq. (47) is in an analytical form.

Results. The yield function Φ for axisymmetric voids (Eq. (43)) reduces to that of Gurson's when $C=T$. This is not true for the plane strain case (Eq. (46), Fig. 7(*a*)), due to a different approach used in this paper in calculating the coefficient c_5. The yield surface for cylindrical voids (axisymmetric and plane strain) are shown in Fig. 5. We note that, there are two axisymmetric yield surfaces (*dashed* and *dotted*) corresponding to different modes: *dashed* curves are for $\dot{E}_{33}\leqslant 0$, as well as for $\dot{E}_{33}>0$ and $(\dot{E}_{kk}/\dot{E}_{33})^2\geqslant 9+12fCT/(C-T)$; the *dotted* curves are for $\dot{E}_{33}>0$ and $(\dot{E}_{kk}/\dot{E}_{33})^2<9+12fCT/(C-T)$.

For higher values of f (Fig. 5(*a*)) and lower values of (C/T) (Fig. 5(*b*)), the yield surface for the plane strain case falls within the envelope of that of the axisymmetric case. However, generally these two yield surfaces overlap at various locations.

For rigid-plastic situation, an approximate solution for a special case ($\dot{\epsilon}_{11}=\dot{\epsilon}_{22}\neq 0$, other strain rates are zero) can be found by using the equilibrium equation, flow rule, and Eq. (1). The solutions satisfy

$$\frac{12}{c}\left(-\sqrt{b+cx}-a\ln\left[-a+\sqrt{b+cx}\right]\right)\Big|_{x=0}^{x=\Sigma_{\gamma\gamma}/2}=\ln f,\quad \Sigma_{\gamma\gamma}\geqslant 0; \tag{49}$$

$$\frac{12}{c}\left(\sqrt{b+cx}-a\ln\left[a+\sqrt{b+cx}\right]\right)\Big|_{x=0}^{x=\Sigma_{\gamma\gamma}/2}=\ln f,\quad \Sigma_{\gamma\gamma}<0, \tag{50}$$

where $a=3(C-T)$, $b=12(C^2-CT+T^2)$, and $c=36(T-C)$. The approximate solution is valid for the associated flow rule. Figure 6 shows the comparison of the approximate solution with that of the upper bound (Eq. (42) for the axisymmetric case, Eq. (45) for plane strain) and modified upper bound (Eq. (43) for the axisymmetric case, Eq. (46) for plane strain) for f versus the transverse ''hydrostatic'' stress. Again, the upper bound nature of Eq. (42) and Eq. (45) is clearly seen. The modified upper bound yield

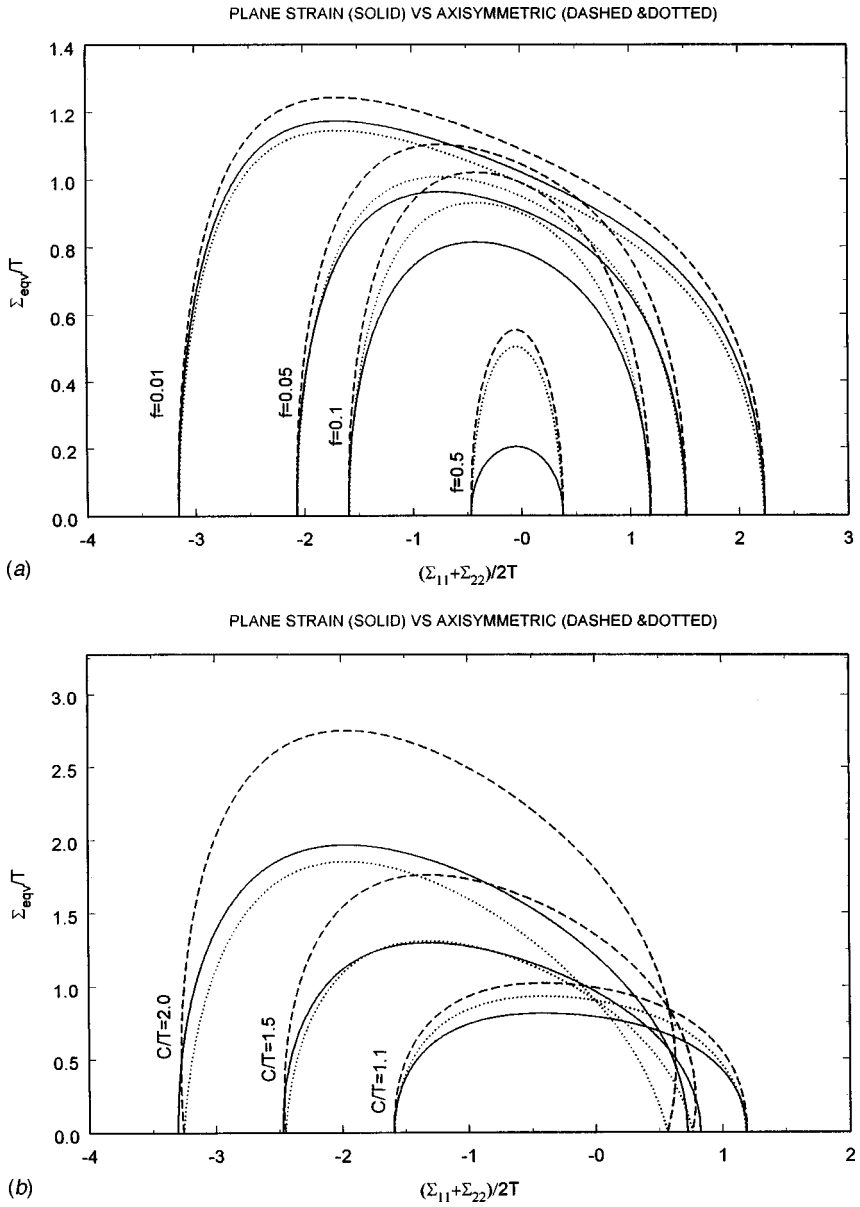


Fig. 5 Yield surfaces of cylindrical voids for axisymmetric and plane-strain cases, (*a*) $C/T=1.1$, (*b*) $f=0.1$

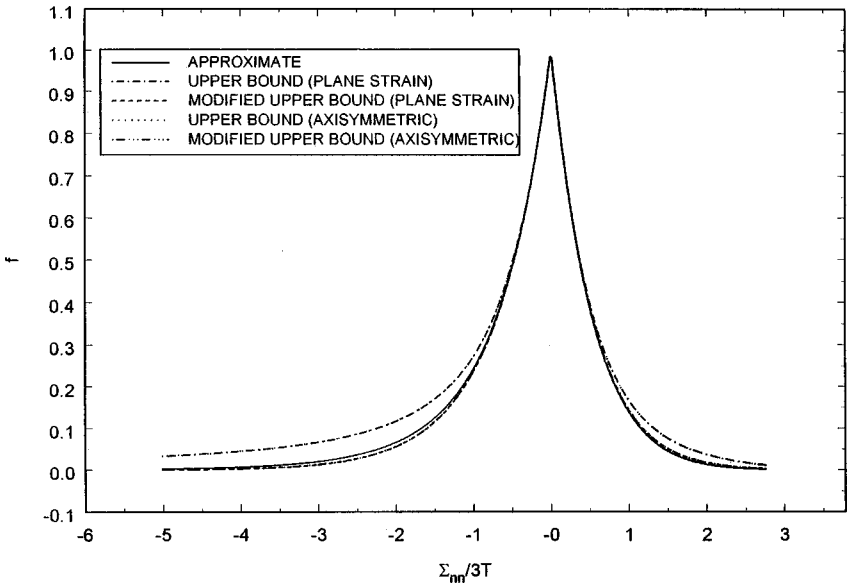


Fig. 6 Comparison of approximate solution with the upper bound and modified upper bound yield surface ($C/T=1.1$) for transverse "hydrostatic" loading

functions (Eqs. (43) and (46)) indicate that the results have been improved significantly over the original upper bound ones.

Figure 7 shows the significant effect of pressure dependency of the matrix on yielding for the plane strain (Fig. 7(*a*)) and axisymmetric (Fig. 7(*b*)) cases. The general trend is that the yield surfaces of pressure-dependent matrix ($C/T=1.1$) are smaller than those of the pressure-independent ones ($C/T=1$).

5 Generalization to Elastic-Plastic Strain-Hardening for Spherical Voids

To facilitate solutions of boundary value problems, we now formulate constitutive equations to include elasticity and strain hardening. For the hardening response of glassy polymers we will use the simple model developed by Lee [15]. As polymers generally also soften, we will use a strain-space approach ([16–19]). In generalizing the rigid perfectly plastic case to include elastic-plastic strain-hardening, one has the flexibility in choosing suitable flow rules for the matrix different from the associated one. Next we define the macroscopic plastic strain rate by using the normality rule (cf. [20,19])

$$\dot{E}^p_{ij}=\lambda\frac{\partial\Phi}{\partial\Sigma_{ij}}. \tag{51}$$

We will now use a simplified version of the hardening model in [15] assuming no plastic volume change of the matrix (nonassociated flow rule) as a first approximation. References [14] and [18] further discuss these issues in more details. The evolution of T and C is assumed to be

$$\dot{T}=H_T\dot{\bar{\epsilon}}^p,\quad \dot{C}=H_C\dot{\bar{\epsilon}}^p \tag{52}$$

where

$$H_T=\frac{\partial T}{\partial\bar{\epsilon}^p},\quad H_C=\frac{\partial C}{\partial\bar{\epsilon}^p}, \tag{53}$$

and $\bar{\epsilon}^p=\int\dot{\bar{\epsilon}}^p dt$. Next we need to find the evolution of the microscopic equivalent plastic strain rate $\dot{\bar{\epsilon}}^p$ with respect to macroscopic quantities. Following Tvergaad [4], we have the plastic work equivalence

$$\dot{\bar{\epsilon}}^p=\frac{\Sigma'_{ij}\dot{E}^p_{ij}}{(1-f)\bar{\sigma}}, \tag{54}$$

which essentially assumes that $\bar{\sigma}$ and $\dot{\bar{\epsilon}}^p$ do not vary within the unit cell. For the current situation, we substitute Eqs. (6) and (19) into Eq. (54) and solve for $\dot{\bar{\epsilon}}^p$ such that

$$\dot{\bar{\epsilon}}^p=\left(\frac{2\Sigma'_{ij}\dot{E}^p_{ij}}{3(1-f)}\sqrt{\frac{\dot{E}^2_{kk}+\frac{9}{4}f\dot{\bar{E}}^2}{(C-T)^2+fCT}}\right)^{1/2}. \tag{55}$$

The evolution of the void fraction is assumed to be

$$\dot{f}=(1-f)\dot{E}^p_{kk}. \tag{56}$$

A yield function (g) is found by substituting the following equation

$$\Sigma'_{ij}=2\mu(E'_{ij}-E'^p_{ij}),\ \ \Sigma_{nn}=3\kappa(E_{kk}-E^p_{kk}) \tag{57}$$

into Eq. (25) where μ is the shear modulus and κ is the bulk modulus. Using the consistency equation and Eqs. (52)–(56), we have

$$\dot{E}^p_{ij}=\frac{\hat{g}}{Z+\Delta}\frac{\partial\Phi}{\partial\Sigma_{ij}}, \tag{58}$$

where

$$\hat{g}=\frac{\partial g}{\partial E_{ij}}\dot{E}_{ij} \tag{59}$$

and

$$Z=\frac{\partial g}{\partial E_{ij}}\frac{\partial\Phi}{\partial\Sigma_{ij}}, \tag{60}$$

$$\Delta=-\left(\frac{\partial g}{\partial C}H_C+\frac{\partial g}{\partial T}H_T\right)Y-\frac{\partial g}{\partial f}(1-f)\frac{\partial\Phi}{\partial\Sigma_{nn}}, \tag{61}$$

where

$$Y=\left(\frac{2\Sigma'_{ij}\frac{\partial\Phi}{\partial\Sigma_{ij}}}{3(1-f)}\sqrt{\frac{\left(\frac{\partial\Phi}{\partial\Sigma_{nn}}\right)^2+9f\Sigma^2_{eq}}{(C-T)^2+fCT}}\right)^{1/2}. \tag{62}$$

During loading, $g=0$, the response of the material is hardening if $\hat{g}>0$, softening if $\hat{g}<0$ and neutral loading if $\hat{g}=0$. The stress rate versus strain rate relations can be written as

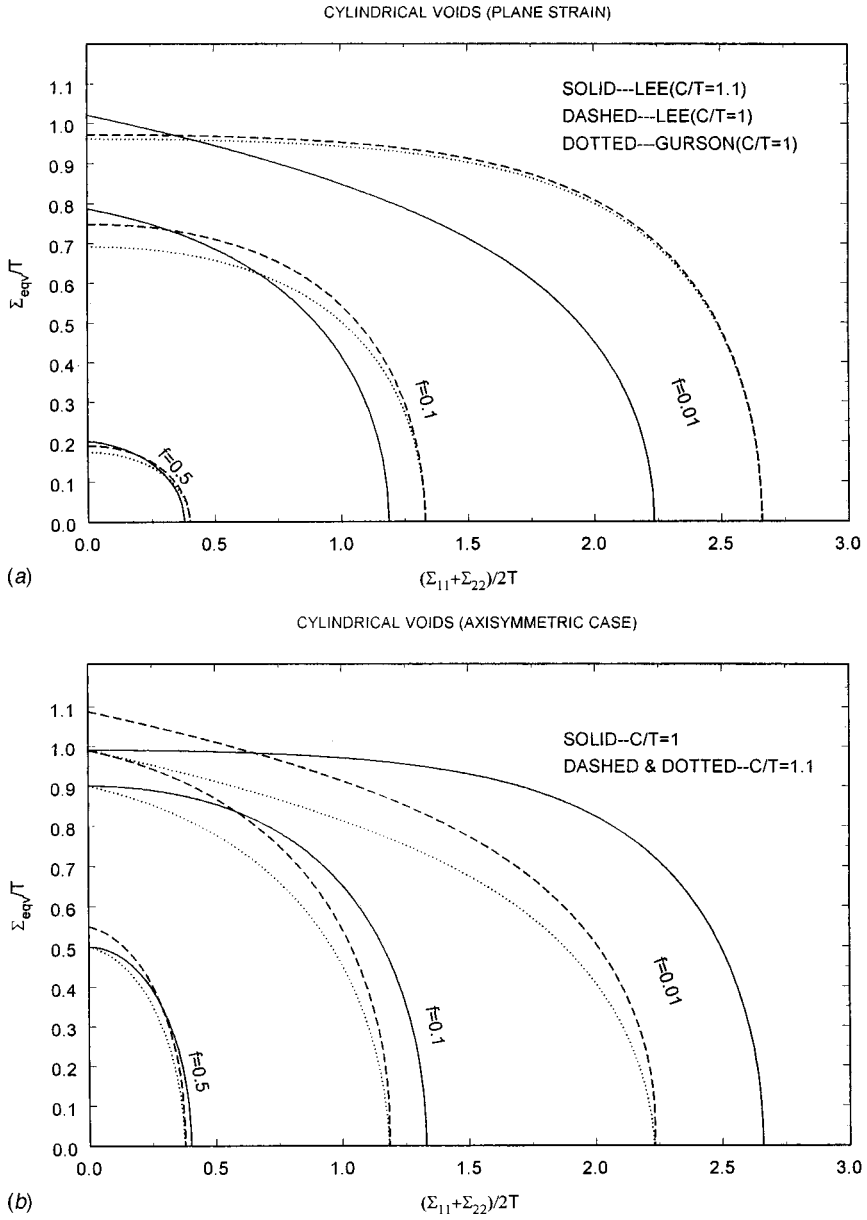


Fig. 7 Effect of pressure dependence on yield surfaces for cylindrical voids, (*a*) plane strain, (*b*) axisymmetric case

$$\dot{\Sigma}_{ij}=(M^e_{ijkl}-M^p_{ijkl})\dot{E}_{kl}=M^e_{ijkl}(\dot{E}_{kl}-\dot{E}^p_{kl}), \tag{63}$$

where the elasticity matrix is

$$M^e_{ijkl}=\mu(\delta_{il}\delta_{kj}+\delta_{ik}\delta_{lj})+\left(\kappa-\frac{2}{3}\mu\right)\delta_{ij}\delta_{kl}. \tag{64}$$

So the plastic moduli M^p_{ijkl}, using Eqs. (58), (59), become

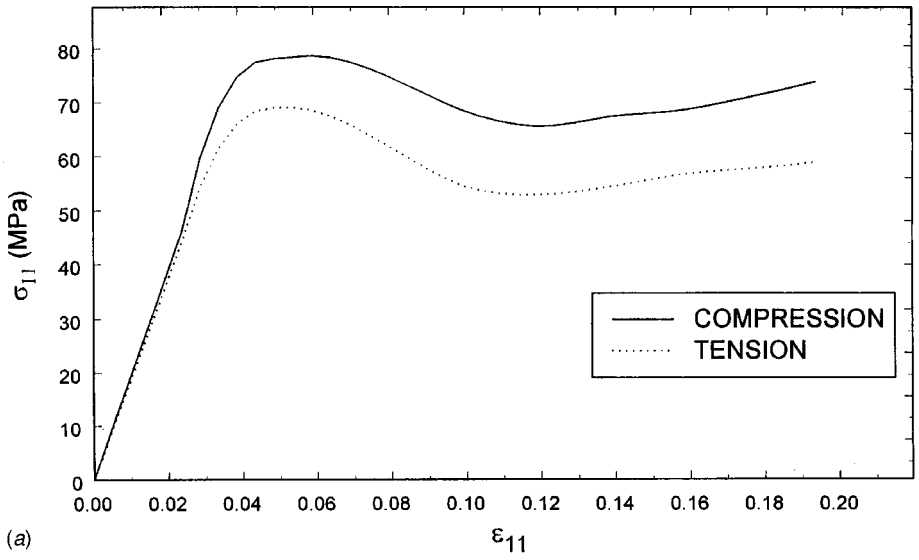
$$M^e_{ijmn}\frac{\partial\Phi}{\partial\Sigma_{mn}}\frac{\partial g}{\partial E_{kl}}/(Z+\Delta).$$

As a simple application of the constitutive equations derived, we find the response of polycarbonate under uniaxial stretching with an initial void fraction $f_0=0.001$ (which could be the result of the nucleation of a craze). The material properties are from the experimental data of [21] which have been analyzed by Lee [15]. The Young's modulus used is 2400 MPa, Poisson's ratio is 0.38. The stress-strain curves of the matrix, shown in Fig. 8(*a*), are interpolated using Mathematica which allows analytical evaluation of the hardening parameters H_T and H_C. The progressively damaged stress-strain response of the material is shown in Fig. 8(*b*) which follows similar trend of the hardening-softening of matrix but the stress is lower than that of *T*. The evolution of the void fraction is shown in Fig. 8(*c*). At this moment, there is no systematic and comprehensive experimental data to rigorously assess the accuracy of the present model.

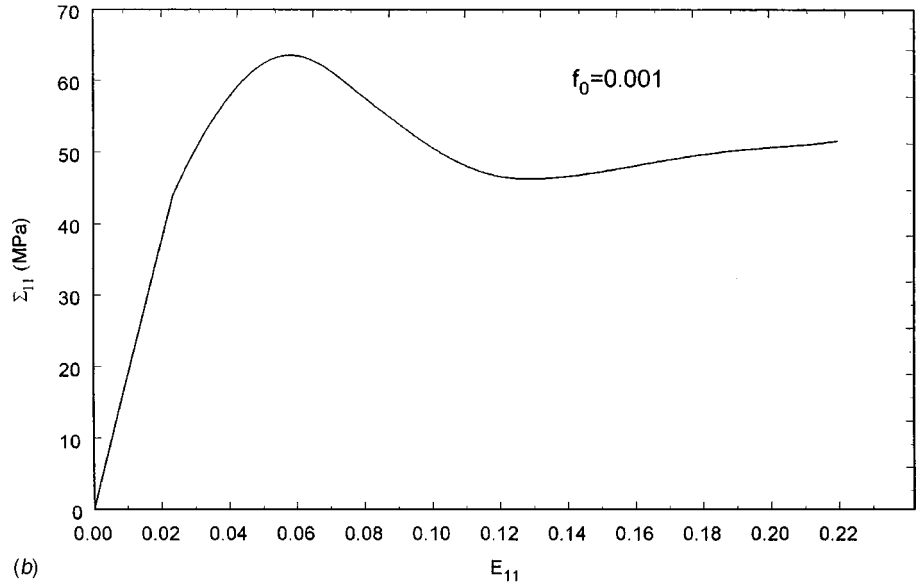
6 Discussions and Conclusions

The porous yield surface, unlike the matrix, is closed in the direction of positive and negative Σ_{nn}. The upper bound nature was evident when comparison with analytical solutions for the purely hydrostatic or the transverse ''hydrostatic'' stress loading was made. After modification, the yield functions have been improved significantly over the upper bound ones.

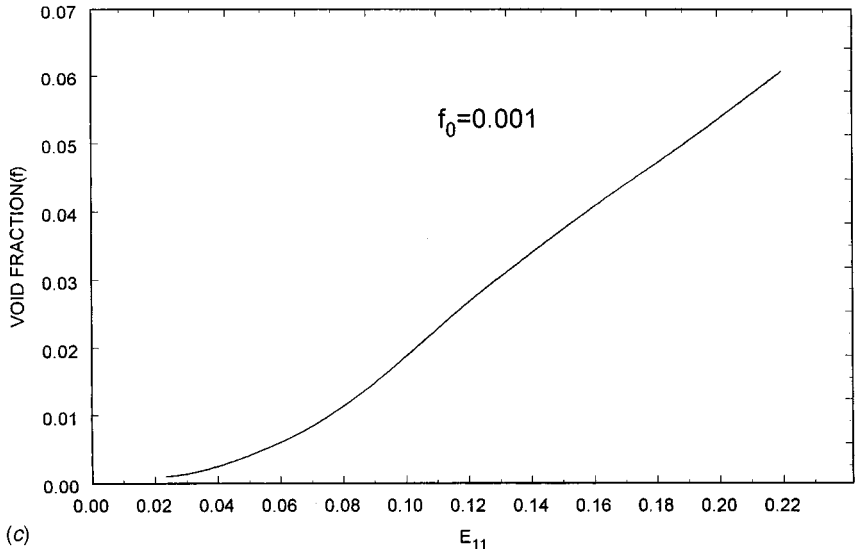
The pressure dependency of the matrix makes a significant impact on the response of the material.



(*a*)



(*b*)



(*c*)

Fig. 8 Uniaxial stretching of polycarbonate, (*a*) stress-strain curves of the matrix *C* and *T*, (*b*) stress-strain curve of the porous material, (*c*) evolution of void fraction *f*

For hardening-softening materials such as glassy polymers, the conditions for the onset of instability as well as post-localization are of interest. These conditions are very sensitive to the material models used. Detailed analysis on instability, including effects of rate sensitivity and different hardening models (such as the orientation hardening model ([22–23]) and their influence on crazing is beyond the scope of this paper and will be reported in the near future.

References

[1] Ishikawa, M., and Takahashi, H., 1991, ''Crazing Mechanism Based on Plastic Instability,'' J. Mater. Sci., **26**, pp. 1295–1300.

[2] Kramer, E. J., 1983, ''Microscopic and Molecular Fundamentals of Crazing,'' Adv. Polym. Sci., **2**, No. 53, pp. 1–56.

[3] Argon, A. S., and Salama, M. M., 1977, ''Growth of Crazes in Glassy Polymers,'' Philos. Mag., **36**, pp. 1195–1216.

[4] Tvergaard, V., 1990, ''Material Failure by Void Growth to Coalescence,'' Adv. Appl. Mech., **27**, pp. 83–151.

[5] Gurson, A. L., 1977, ''Continuum Theory of Ductile Rupture by Void Nucleation and Growth: Part I—Yield Criteria and Flow Rules for Porous Ductile Media,'' ASME J. Eng. Mater. Technol., **99**, pp. 2–15.

[6] Drucker, D. C., 1973, ''Plasticity Theory, Strength-Differential (S-D) Phenomenon, and Volume Expansion in Metals and Plastics.'' Metall. Trans., **4**, pp. 667–673.

[7] Raghava, R. S., Caddell, R. M., and Yeh, G. S. Y., 1973, ''The Macroscopic Yield Behavior of Polymers,'' J. Mater. Sci., **9**, pp. 225–232.

[8] Avitzur, B., 1980, *Metal Forming: The Application of Limit Analysis*, Marcel Dekker, New York.

[9] Wolfram, S., 1991, *Mathematica: A System for Doing Mathematics by Computer*, 2nd Ed., Addison-Wesley, Reading, MA.

[10] Oung, J., 1999, ''Yield Functions and Flow Rules for Porous Pressure-Dependent Strain-Hardening Polymeric Materials'', M.Sc. thesis, University of Alaska, Fairbanks, AK.

[11] Raghava, R. S., and Caddell, R. M., 1973, ''A Macroscopic Yield Criterion for Crystalline Polymers,'' Int. J. Mech. Sci., **15**, pp. 967–974.

[12] Freire, J. L. F., and Riley, W. F., 1980, ''Yield Behavior of Photoplastic Materials,'' Exp. Mech., **20**, pp. 118–125.

[13] Chen, W. F., 1975, *Limit Analysis and Soil Plasticity*, Elsevier, New York.

[14] Lee, J. H., 1988, ''Upper Bound Analysis of the Upsetting of Pressure-Sensitive Polymeric Rings,'' Int. J. Mech. Sci., **30**, pp. 601–612.

[15] Lee, J. H., 1989, ''Characterization of Strain Hardening for a Simple Pressure-Sensitive Plasticity Model,'' Acta Mech., **77**, pp. 133–151.

[16] Naghdi, P. M., and Trapp, J. A., 1975, ''The Significance of Formulating Plasticity Theory With Reference to Loading Surfaces in Strain Space,'' Int. J. Eng. Sci., **13**, pp. 785–797.

[17] Casey, J., and Naghdi, P. M., 1981, ''On the Characterization of Strain-Hardening Plasticity With Reference to Loading Surfaces in Strain Space,'' ASME J. Eng. Mat. Tech., **49**, pp. 285–296.

[18] Lee, J. H., 1988, ''Some Exact and Approximate Solutions for the Modified von Mises Criterion,'' ASME J. Appl. Mech., **55**, pp. 160–166.

[19] Lee, J. H., and Zhang, Y., 1994, ''A Finite-Element Work-Hardening Plasticity Model of the Uniaxial Compression and Subsequent Failure of Porous Cylinders Including Effects of Void Nucleation and Growth—Part I: Plastic Flow and Damage,'' ASME J. Eng. Mater. Technol., **116**, pp. 69–79.

[20] Lee, J. H., 1990, ''On the Asymmetry of Elastic-Plastic Moduli for a Class of Pressure-Dependent Models in Strain Space,'' Acta Mech., **84**, pp. 217–220.

[21] Spitzig, W. A., and Richmond, O., 1979, ''Effect of Hydrostatic Pressure on the Deformation Behavior of Polyethylene and Polycarbonate in Tension and Compressions,'' Polym. Eng. Sci., **19**, pp. 1129–1139.

[22] Haward, R. N., and Thackray, G., 1968, ''The Use of a Mathematical Model to Describe Isothermal Stress-Strain Curves in Glassy thermoplastics,'' Proc. R. Soc. London, Ser. A, **302**, p. 453.

[23] Boyce, M. C., Parks, D. M., and Argon, A. S., 1988, ''Large Inelastic Deformation of Glassy Polymers. Part I: Rate Dependent Constitutive Model,'' Mech. Mater., **7**, pp. 15–33.

The Influence of "Shell Behavior" on Load Distribution for Thin-Walled Conical Joints

L. Bruschelli
Doctor of Mechanical Engineering,
Viale Madonna 2,
22063 Cantu (Como), Italy

V. Latorrata
Doctor of Mechanical Engineering,
Via Resistenza 22,
22072 Cermenate (Como), Italy

This article presents a new analytical method with a numerical solution to calculate load distribution in threaded connections. Our departure model was that suggested by D. G. Sopwith who has proposed the most recent and most tested theory. Our research consists in the introduction of conicity and, above all, in the development of the influence of boundary geometry (i.e. the nonthreaded section) on load distribution. Pipe joints are analyzed in special detail, supplying us with useful finite element method comparative results. [S0021-8936(00)02002-X]

Introduction

The experimental approach in the research into threaded connections is that which has been followed by the majority of authors, especially prior to the establishment of numeric methods, such as finite elements.

In recent times, computer-based methods have attracted all the attention in giving the opportunity of obtaining complete results in terms of local stress and no longer only in load distribution.

J. N. Goodier [1] was the first in 1940 to use an analytical approach consisting in an approximate formula between displacement and axial load. In 1948 D. G. Sopwith [2] suggested the first accurate theory to obtain load distribution, having been inspired by Goodier's work. E. E. Stoeckly and H. J. Macke [3] broadened Sopwith's formula in 1952 with the introduction of the taper.

Thus, the analytical bibliography available basically consists of Sopwith's theory, which is weak in that the following are not taken into consideration:

1 threading conicity,
2 element bending behavior due to the variable load absorbed by the thread, and
3 the effect of the boundary substructures on the threaded part.

A new analytical approach is presented to improve the Sopwith model which is not suitable for thin joints with complex structural schemes.

Geometry

The calculation of load distribution is done by expressing the displacements of the contact points by means of the contact forces, by writing a congruence equation to impose the contact (without separations or penetrations of teeth flanks), and by solving the differential problem obtained.

The factors forming the relative total separation between the flanks are three:

1 δ_I, relative axial displacement due to thread bending,
2 δ_{II}, relative axial recession due to radial displacement of bolt and nut sections (according to Goodier's notation, *recession* is the axial separation of two threads consequent on a relative radial displacement between them), and
3 δ_{III}, relative axial displacement due to the different axial bolt and nut strain.

We have considered (Fig. 1) the threaded part of a conic joint (intended as the part between two planes perpendicular to the axis and crossing the first and last points of engagement between nut and bolt) as being subjected to boundary actions by the remaining parts.

The situation examined, which will be indicated as "*forcing condition,*" is that characterized by the placing of the axial make-up load $-P$ on the lower nut surface and P on the lower bolt surface (Fig. 1).

The basic geometrical measurements of the threaded part (Fig. 1) are the length L, the thread start diameter D_{0fil}, internal and external diameters D_0 and D_3, and the conicity ψ.

The characteristic abscissas (Fig. 2) are x, the axial abscissa and s, the curvilinear abscissa along the thread helix, both measured from the loaded surface.

There is a link between them, which can be obtained for the constant pitch (a) helix with conicity ψ of Fig. 2:

$$x(s)=\frac{D_{0fil}}{2\tan\psi}-\sqrt{\frac{D_{0fil}^2}{4\tan^2\psi}-\frac{sa}{\pi\tan\psi}}. \tag{1}$$

The diameters corresponding to the half-thickness surfaces of the nut and bolt, represented with dotted lines in Fig. 1, are $D_n(s)=2R_n(s)$ and $D_b(s)=2R_b(s)$; $D_{fil}(s)=2R_{fil}(s)$ is the generic engagement diameter (Fig. 1).

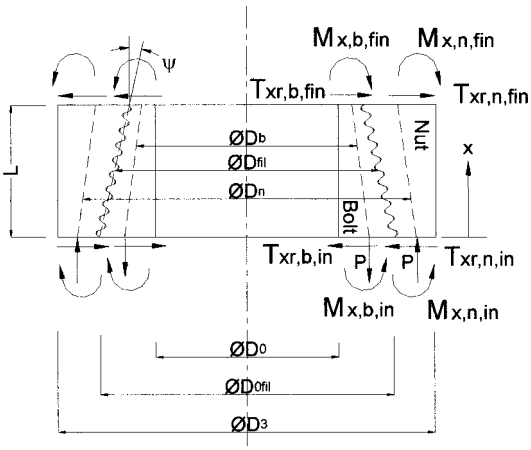


Fig. 1 Threaded part of a generic joint

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 23, 1998; final revision, April 12, 1999. Associate Technical Editor: W. K. Liu. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

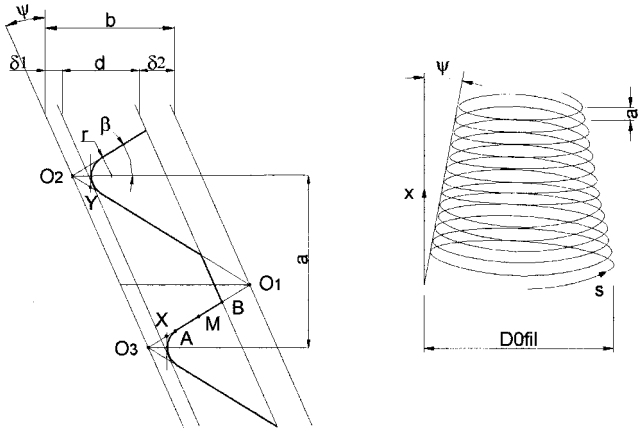


Fig. 2 Threading geometry

Factor δ_I

We assume that the thread (Fig. 2) is obtained by turning with a radial (and not perpendicular to the external conic surface) penetration of the tool.

The independent dimensions are: a, pitch; β, thread half-angle; ψ, thread conicity; δ_2, upper cut off value; and r, notch radius.

The dependent dimensions are

$$b=\frac{a}{2}\left(\frac{1}{\tan\beta}-\tan\beta*\tan^2\psi\right) \tag{2}$$

$$\delta_1\cong r\left(\frac{1}{\sin\beta}-1\right) \tag{3}$$

$$d=b-\delta_1-\delta_2\cong b-\delta_2-r\left(\frac{1}{\sin\beta}-1\right) \tag{4}$$

$$L_f=AB=\frac{a}{2\sin\beta}-\frac{a*\tan\psi}{2\cos\beta}-r*\tan\left(\frac{\pi}{2}-\beta\right)-\frac{\delta_2}{\cos\beta+\sin\beta*\tan\psi} \tag{5}$$

$$R_y=O_1Y=\frac{a}{2}\left(\frac{1}{\sin\beta}+\frac{\tan\psi}{\cos\beta}\right)-\frac{\delta_1}{\cos\beta} \tag{6}$$

$$R_x=O_1X=\frac{a}{2}\left(\frac{1}{\sin\beta}-\frac{\tan\psi}{\cos\beta}\right)-\frac{\delta_1}{\cos\beta}. \tag{7}$$

The factor δ_I is the one due to the bending, caused by contact forces, of both thread flanks in contact (Fig. 3).

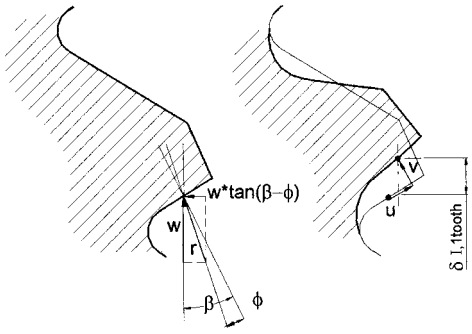


Fig. 3 Contact forces and displacements

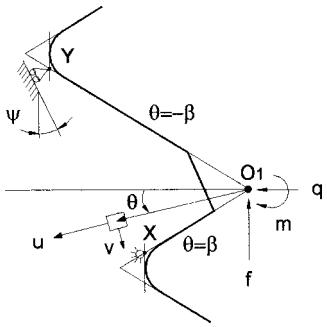


Fig. 4 Polar coordinates in the thread, equivalent loads and boundary conditions

The vertical factor w is the unknown element, the global force r is equal to $w\sec(\beta-\phi)$, and the horizontal factor is $w\tan(\beta-\phi)$ where $\phi=\text{arctg}(f)$ is the static friction angle.

The mutual forces are per unit thread length and act at the resulting force application point, which may not necessarily be the point M of middle contact area (Fig. 2).

So, we express the w lever arm referring to point O_1 as $b/2*c_f$ (Fig. 2), introducing a position coefficient c_f.

Under nominal conditions, with the resulting force r acting on M (the middle point of L_f), we have

$$c_{f,\text{nom}}=\frac{\delta_2}{b+b*\tan\beta*\tan\psi}+\frac{a}{2b*\tan\beta}-\frac{a*\tan\psi}{2b}-\frac{r}{b}\tan\left(\frac{\pi}{2}-\beta\right)\cos\beta. \tag{8}$$

After having moved the two forces to point O_1 it is possible to solve the elastic wedge problem in a plain-strain condition in the polar coordinates shown in Fig. 4, binding point X with a hinge and Y with a roller sloped at a ψ angle to the axis.

The three forces shown in Fig. 4 are $f=w$, $q=w\tan(\beta-\phi)$, $m=w[1+\tan(\beta)\tan(\beta-\phi)]*b*c_f/2$.

The stresses are ([4])

$$\begin{cases}\sigma_\vartheta=0 & 9(a)\\ \sigma_\rho=-\dfrac{2A}{\rho^2}\sin(2\vartheta)+\dfrac{2B}{\rho}\sin(\vartheta)-\dfrac{2C}{\rho}\cos(\vartheta) & (9b)\\ \tau_{\rho\vartheta}=\dfrac{A}{\rho^2}(\cos(2\vartheta)-\cos(2\beta)). & (9c)\end{cases}$$

The constants A, B, C are

$$A=\frac{m}{\sin(2\beta)-2\beta\cos(2\beta)}, \tag{10a}$$

$$B=\frac{f}{2\beta-\sin(2\beta)}, \tag{10b}$$

$$C=\frac{q}{2\beta+\sin(2\beta)}. \tag{10c}$$

If $u(\rho,\vartheta)$ and $v(\rho,\vartheta)$ are the radial and tangential displacements, we obtain from the displacement strains and strain stresses links in polar coordinates

$$\begin{cases} \dfrac{E}{1-\nu^2}\dfrac{\partial u}{\partial \rho}=\sigma_\rho-\dfrac{\nu}{1-\nu}\sigma_\vartheta & (11a)\\ \dfrac{E}{1-\nu^2}\dfrac{1}{\rho}\left(u+\dfrac{\partial v}{\partial \vartheta}\right)=\sigma_\vartheta-\dfrac{\nu}{1-\nu}\sigma_\rho & (11b)\\ \dfrac{E}{1-\nu^2}\dfrac{1}{\rho}\left(\dfrac{\partial u}{\partial \vartheta}+\rho\dfrac{\partial v}{\partial \rho}-\nu\right)=2\left(1+\dfrac{\nu}{1-\nu}\right)\tau_{\rho\vartheta}. & (11c)\end{cases}$$

We obtain the displacements solving the equation system (11):

$$\begin{aligned}\frac{E}{1-\nu^2}u(\rho,\vartheta)=&\frac{2A\sin(2\vartheta)}{\rho}+2B\sin(\vartheta)\log(\rho)\\&-2C\cos(\vartheta)\log(\rho)+C_1\cos(\vartheta)\\&+C_2\sin(\vartheta)-C\left(1-\frac{\nu}{1-\nu}\right)\vartheta\sin(\vartheta)\\&-B\left(1-\frac{\nu}{1-\nu}\right)\vartheta\cos(\vartheta). \qquad (12)\end{aligned}$$

$$\begin{aligned}\frac{E}{1-\nu^2}v(\rho,\vartheta)=&-\frac{A\nu\cos(2\vartheta)}{\rho(1-\nu)}+2B\frac{\nu}{1-\nu}\cos(\vartheta)\\&+2C\frac{\nu}{1-\nu}\sin(\vartheta)+\frac{A}{\rho}\cos(2\vartheta)\\&+2B\cos(\vartheta)\log(\rho)+2C\sin(\vartheta)\log(\rho)\\&+C_3\rho+\frac{A\cos(2\beta)}{\rho(1-\nu)}-C_1\sin(\vartheta)+C_2\cos(\vartheta)\\&+C\left(1-\frac{\nu}{1-\nu}\right)(\sin(\vartheta)-\vartheta\cos\vartheta)\\&+B\left(1-\frac{\nu}{1-\nu}\right)(\cos(\vartheta)+\vartheta\sin\vartheta). \qquad (13)\end{aligned}$$

The three unknown constants C_1, C_2, C_3 can be obtained by introducing the boundary conditions

$$\begin{cases} u(\rho=R_x,\vartheta=\beta)=0 & (14a)\\ v(\rho=R_x,\vartheta=\beta)=0 & (14b)\\ u(\rho=R_y,\vartheta=-\beta)*\cos(\beta+\psi)\\ +v(\rho=R_y,\vartheta=-\beta)*\sin(\beta+\psi)=0. & (14c)\end{cases}$$

By developing the boundary condition expressions (14) we obtain a linear equation system in the unknown C_1, C_2, C_3:

$$\begin{bmatrix} A_{11} & A_{12} & A_{13}\\ A_{21} & A_{22} & A_{23}\\ A_{31} & A_{32} & A_{33}\end{bmatrix}\begin{Bmatrix} C_1\\ C_2\\ C_3\end{Bmatrix}=\begin{Bmatrix} B_1\\ B_2\\ B_3\end{Bmatrix}, \qquad (15)$$

the coefficients

$$\begin{aligned}A_{11}&=\cos\beta,\quad A_{12}=\sin\beta,\quad A_{13}=0,\\A_{21}&=-\sin\beta,\quad A_{22}=\cos\beta,\quad A_{23}=R_x,\\A_{31}&=\cos\beta\cos(\beta+\psi)+\sin\beta\sin(\beta+\psi),\\A_{32}&=\cos\beta\sin(\beta+\psi)-\sin\beta\cos(\beta+\psi),\\A_{33}&=R_y\sin(\beta+\psi),\end{aligned}$$

and the known terms

$$\begin{aligned}B_1=&-\frac{2A\sin(2\beta)}{R_x}+2\log(R_x)(C\cos\beta-B\sin\beta)\\&+\beta\left(1-\frac{\nu}{1-\nu}\right)(C\sin\beta+B\cos\beta)\end{aligned}$$

$$\begin{aligned}B_2=&-\frac{2A\cos(2\beta)}{R_x}-2\log(R_x)(C\sin\beta+B\cos\beta)\\&+\beta\left(1-\frac{\nu}{1-\nu}\right)(C\cos\beta-B\sin\beta)-\frac{B}{1-\nu}\cos\beta\\&-\frac{C}{1-\nu}\sin\beta\end{aligned}$$

$$\begin{aligned}B_3=&\frac{2A}{R_y}[\sin(2\beta)\cos(\beta+\psi)-\cos(2\beta)\sin(\beta+\psi)]\\&-B\cos\beta\sin(\beta+\psi)\left[\frac{1}{1-\nu}+2\log(R_y)\right]+2B\sin\beta\cos(\beta\\&+\psi)\log(R_y)-B\left(1-\frac{\nu}{1-\nu}\right)\beta\cos\beta\cos(\beta+\psi)\\&-B\left(1-\frac{\nu}{1-\nu}\right)\beta\sin\beta\sin(\beta+\psi)\\&-C\left(1-\frac{\nu}{1-\nu}\right)\beta\cos\beta\sin(\beta+\psi)\\&+C\left(1-\frac{\nu}{1-\nu}\right)\beta\sin\beta\cos(\beta+\psi)+2C\log(R_y)\\&\times\cos\beta\cos(\beta+\psi)+\frac{C}{1-\nu}\beta\sin\beta\sin(\beta+\psi)\\&+2C\log(R_y)\beta\sin\beta\sin(\beta+\psi).\end{aligned}$$

The system (15) can be solved numerically.

We evaluate the displacements (especially v) at the point where the mutual forces act, having $\overline{\rho}=b/2\cos\beta c_f, \overline{\vartheta}=\beta$; so we express

$$\delta_I=2\delta_{I,1\text{tooth}}=\frac{-2v(\overline{\rho},\overline{\vartheta})}{\cos\beta}. \qquad (16)$$

We can introduce factor h (axial recession compliance factor):

$$\delta_I=\frac{2h}{E}w \qquad (17)$$

So

$$h=E\frac{-v(\overline{\rho},\overline{\vartheta})}{w\cos\beta}. \qquad (18)$$

Conical Shells With Shear Behavior

We need to develop a theory which allows us to deal with conical shells of variable thickness.

The best approach is Mindlin [5] and Reissner [6] for cylindrical axisymmetrical shells, which includes the shear strain in the kinematic model.

The extension in the case of the half-thickness conical surface is obtained by modifying the equilibrium equations keeping the

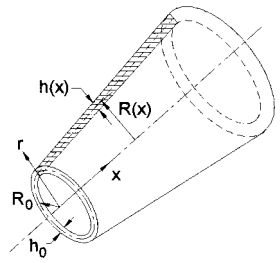


Fig. 5 Shell coordinates

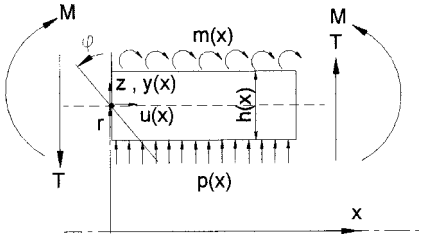


Fig. 6 Shell displacements and load conventions

kinematic model unchanged. The criteria used are shown in Fig. 5 and Fig. 6.

Displacements of a point outside the half-thickness shell surface are

$$\begin{cases} s_x = u(x) - z\varphi(x) \\ s_r = y(x) \\ s_\vartheta = 0 \end{cases} . \tag{19}$$

Strains are

$$\begin{cases} \varepsilon_x = \dfrac{ds_x}{dx} = \dfrac{du(x)}{dx} - z\dfrac{d\varphi(x)}{dx} \\ \varepsilon_r = \dfrac{ds_r}{dr} = 0 \\ \varepsilon_\vartheta = \dfrac{s_r}{r} + \dfrac{ds_\vartheta}{rdr} = \dfrac{y(x)}{R(x)} \\ \gamma_{\vartheta x} = \dfrac{ds_x}{rd\vartheta} + \dfrac{ds_\vartheta}{dx} = 0 \\ \gamma_{r\vartheta} = \dfrac{ds_r}{rd\vartheta} + \dfrac{ds_\vartheta}{dr} - \dfrac{s_\vartheta}{r} = 0 \\ \gamma_{rx} = \dfrac{ds_x}{dr} + \dfrac{ds_r}{dx} = \dfrac{dy(x)}{dx} - \varphi(x) \end{cases} \tag{20}$$

and stresses are

$$\begin{cases} \sigma_x = \dfrac{E}{(1-\nu^2)}(\varepsilon_x + \nu\varepsilon_\vartheta + \nu\varepsilon_r) = \dfrac{E}{(1-\nu^2)}\dfrac{du(x)}{dx} - \dfrac{E}{(1-\nu^2)}\dfrac{zd\varphi(x)}{dx} + \dfrac{\nu E}{(1-\nu^2)}\dfrac{y(x)}{R(x)} \\ \sigma_\vartheta = \dfrac{E}{(1-\nu^2)}(\varepsilon_\vartheta + \nu\varepsilon_x + \nu\varepsilon_r) = \dfrac{E}{(1-\nu^2)}\dfrac{y(x)}{R(x)} + \dfrac{\nu E}{(1-\nu^2)}\dfrac{du(x)}{dx} - \dfrac{\nu E}{(1-\nu^2)}\dfrac{zd\varphi(x)}{dx} \\ \tau_{rx} = G\gamma_{rx} = G\left(\dfrac{dy(x)}{dx} - \varphi(x)\right) \end{cases} \tag{21}$$

Integrating the stresses we obtain the internal forces per unit length; we introduce the equivalent thickness $h^*(x) = 5/6h(x)$ so as not to overlook the variation of τ_{xr} on the same thickness (Timoshenko [7] and Corradi Dell'acqua [8]).

The internal forces are

$$M_x = \frac{Eh^3(x)}{12(1-\nu^2)}\frac{d\varphi(x)}{dx}, \quad M_\vartheta = \frac{\nu Eh^3(x)}{12(1-\nu^2)}\frac{d\varphi(x)}{dx}$$

$$N_x = \frac{Eh(x)}{(1-\nu^2)}\frac{du(x)}{dx} + \frac{\nu Eh(x)}{(1-\nu^2)}\frac{y(x)}{R(x)},$$

$$N_\vartheta = \frac{\nu Eh(x)}{(1-\nu^2)}\frac{du(x)}{dx} + \frac{Eh(x)}{(1-\nu^2)}\frac{y(x)}{R(x)} \tag{22}$$

$$T_{xr} = Gh^*(x)\left(\frac{dy(x)}{dx} - \varphi(x)\right).$$

The following links can be observed:

$$M_\vartheta = \nu M_x$$

$$N_\vartheta = \frac{Eh(x)}{R(x)}y(x) + \nu N_x . \tag{23}$$

We write the equilibrium equations including the conicity effect of the half thickness surface (Fig. 7); ignoring the superior infinitesimal quantities and simplifying the expressions we obtain

$$\begin{cases} \dfrac{dN_x}{dx} = -\dfrac{dR}{dx} \\ \dfrac{dT_{xr}}{dx} + \dfrac{T_{xr}}{R}\dfrac{dR}{dx} = \dfrac{N_\vartheta}{R} - p \\ \dfrac{M_x}{R}\dfrac{dR}{dx} + \dfrac{dM_x}{dx} - N_x\dfrac{dR}{dx} = m - T_{xr} . \end{cases} \tag{24}$$

Substituting in the last two Eqs. (24) the expressions (22), (23) we get the following equation system:

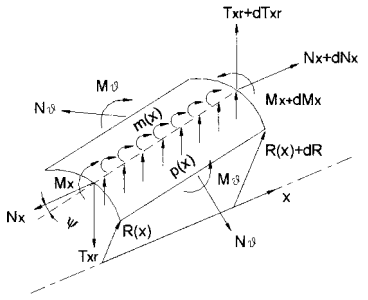


Fig. 7 Internal and applied loads

$$\begin{cases} D_2(x)\left(\dfrac{d^2y(x)}{dx^2}-\dfrac{d\varphi(x)}{dx}\right)+\left(\dfrac{D_2(x)R'(x)}{R(x)}+D_2'(x)\right)\left(\dfrac{dy(x)}{dx}-\varphi(x)\right)-D_3(x)y(x)=-p(x)+\dfrac{\nu N_x(x)}{R(x)} \\ D_1(x)\dfrac{d^2\varphi(x)}{dx^2}+\left(\dfrac{D_1(x)R'(x)}{R(x)}+D_1'(x)\right)\dfrac{d\varphi(x)}{dx}+D_2(x)\left(\dfrac{dy(x)}{dx}-\varphi(x)\right)=m(x)+N_x(x)R'(x) \end{cases} \tag{25}$$

$$D_1(x)=\frac{Eh^3(x)}{12(1-\nu^2)},\quad D_2(x)=Gh^*(x),\quad D_3(x)=\frac{Eh(x)}{R^2(x)},\quad R'(x)=\frac{dR(x)}{dx}.$$

The boundary conditions are the following:

$$D_1(x=0)\left.\frac{d\varphi}{dx}\right|_{x=0}=M_{x,in}$$

$$D_2(x=0)\left(\left.\frac{dy}{dx}\right|_{x=0}-\varphi(x=0)\right)=T_{xr,in}$$

$$D_1(x=L)\left.\frac{d\varphi}{dx}\right|_{x=L}=M_{x,fin}$$

$$D_2,(x=L)\left(\left.\frac{dy}{dx}\right|_{x=L}-\varphi(x=L)\right)=T_{xr,fin}. \tag{26}$$

Factor δ_{II}

We consider the nut and bolt as two smooth conical shells, touching along the surface defined by the points of contact on the thread.

If we ignore the radial displacement variation between the half-thickness surface and the interface we can express

$$\delta_{II}=(y_n(s)-y_b(s))\tan\beta, \tag{27}$$

coupling the shell equations to the congruence relation.

The load system of the two shells consists in the axial load $N_{x,n}=(\int_0^s wds-P)/(2\pi R_n)$ and $N_{x,b}=(P-\int_0^s wds)/(2\pi R_b)$ and in the interface pressure (made axisymmetrical by giving the same value to all the points on each parallel) $p=w/a\,\tan(\beta-\phi)$.

Since in the shell Eqs. (25) the loads refer to the half-thickness surface, it is necessary to move there the axial distributed load w, acting on the interface at a distance of $D_{fil}/2$ from the axis, thus introducing the distributed moment $m_n=w/a(R_n-D_{fil}/2)$ and $m_b=w/a(D_{fil}/2-R_b)$ (Fig. 8).

Factor δ_{III}

Factor δ_{III} is associated with the relative axial displacement in the contact area due to axial strains $\epsilon_{x,n}$ and $\epsilon_{x,b}$.

Ignoring the contribution of N_ϑ, M_x and M_ϑ into $\epsilon_{x,n}$ and $\epsilon_{x,b}$, due to the strengthening effect of the tooth, and keeping in mind the axial load expressions $P_n=-(P-\int_0^s wds)$ and $P_b=P-\int_0^s wds$ for nut and bolt, we maintain the equation already proposed by Sopwith:

$$d\delta_{III}=\frac{dx}{E}\left(\frac{P_n}{A_n}-\frac{P_b}{A_b}\right)=\frac{dx}{E}\left\{\int_0^s wds\left[\frac{1}{A_b}+\frac{1}{A_n}\right]-P\left[\frac{1}{A_b}+\frac{1}{A_n}\right]\right\}. \tag{28}$$

Differential Equation System

The congruence equation is $d\delta_{III}=d(\delta_{II}+\delta_I)$; substituting the explicit expressions (17), (27), (28) we obtain

$$-\left(\frac{1}{A_b}+\frac{1}{A_n}\right)\frac{dx}{ds}\int_0^s wds+2h\frac{dw}{ds}+E\tan\beta\left(\frac{dy_n}{dx}-\frac{dy_b}{dx}\right)\frac{dx}{ds}=-\frac{dx}{ds}P\left(\frac{1}{A_b}+\frac{1}{A_n}\right). \tag{29}$$

The associated boundary condition is $\int_0^{s_1} wds=P$; s_1 is the overall length of the thread helix.

It is necessary to emphasize how the differentiation of the term $\delta_{II}=(y_n(x(s))-y_b(x(s)))\tan\beta$ takes place by means of composite derivatives, in such a way as to obtain the derivatives of y_n and y_b as regards the axial coordinate x, the most simple to solve the shell equations.

In the case of the congruence equation it is easier to use the s variable, both for the presence of the integral term $\int_0^s wds$, and for the more immediate physical interpretation of w.

Thus in the system both variables s and x are found; however, the system will depend on a single unknown, since we have the link (1).

Keeping in mind the shell equations, it is possible to express the overall differential system of Eqs. (30):

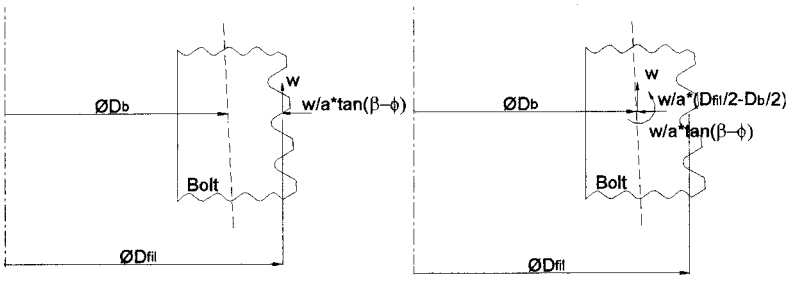


Fig. 8 Distributed moment *m* due to *w* movement

$$
\begin{cases}
-\left(\frac{1}{A_b}+\frac{1}{A_n}\right)\frac{dx}{ds}\int_0^s w\,ds+2h\frac{dw}{ds}+E\tan\beta\left(\frac{dy_n}{dx}-\frac{dy_b}{dx}\right)\frac{dx}{ds}=-\frac{dx}{ds}P\left(\frac{1}{A_b}+\frac{1}{A_n}\right)\\
D_{2,n}\left(\frac{d^2y_n}{dx^2}-\frac{d\varphi_n}{dx}\right)+\left(\frac{D_{2,n}R_n'}{R_n}+D_{2,n}'\right)\left(\frac{dy_n}{dx}-\varphi_n\right)-D_{3,n}y_n=-\frac{w}{a}\tan(\beta-\phi)+\frac{\nu}{R_n}\left(\frac{\int_0^s w\,ds-P}{2\pi R_n}\right)\\
D_{1,n}\frac{d^2\varphi_n}{dx^2}+\left(\frac{D_{1,n}R_n'}{R_n}+D_{1,n}'\right)\frac{d\varphi_n}{dx}+D_{2,n}\left(\frac{dy_n}{dx}-\varphi_n\right)=\frac{w}{a}\left(R_n-\frac{D_{fil}}{2}\right)+\frac{\int_0^s w\,ds-P}{2\pi R_n}R_n'\\
D_{2,b}\left(\frac{d^2y_b}{dx^2}-\frac{d\varphi_b}{dx}\right)+\left(\frac{D_{2,b}R_b'}{R_b}+D_{2,b}'\right)\left(\frac{dy_b}{dx}-\varphi_b\right)-D_{3,b}y_b=\frac{w}{a}\tan(\beta-\phi)+\frac{\nu}{R_b}\left(\frac{P-\int_0^s w\,ds}{2\pi R_b}\right)\\
D_{1,b}\frac{d^2\varphi_b}{dx^2}+\left(\frac{D_{1,b}R_b'}{R_b}+D_{1,b}'\right)\frac{d\varphi_b}{dx}+D_{2,b}\left(\frac{dy_b}{dx}-\varphi_b\right)=\frac{w}{a}\left(\frac{D_{fil}}{2}-R_b\right)+\frac{P-\int_0^s w\,ds}{2\pi R_b}R_b'\\
\int_0^{s_1} w\,ds=P\\
D_{1,b}(x=0)\left.\frac{d\varphi_b}{dx}\right|_{x=0}=M_{x,b,in}\\
D_{2,b}(x=0)\left(\left.\frac{dy_b}{dx}\right|_{x=0}-\varphi_b(x=0)\right)=T_{xr,b,in}\\
D_{1,b}(x=L)\left.\frac{d\varphi_b}{dx}\right|_{x=L}=M_{x,b,fin}\\
D_{2,b}(x=L)\left(\left.\frac{dy_b}{dx}\right|_{x=L}-\varphi_b(x=L)\right)=T_{xr,b,fin}\\
D_{1,n}(x=0)\left.\frac{d\varphi_n}{dx}\right|_{x=0}=M_{x,n,in}\\
D_{2,n}(x=0)\left(\left.\frac{dy_n}{dx}\right|_{x=0}-\varphi_n(x=0)\right)=T_{xr,n,in}\\
D_{1,n}(x=L)\left.\frac{d\varphi_n}{dx}\right|_{x=L}=M_{x,n,fin}\\
D_{2,n}(x=L)\left(\left.\frac{dy_n}{dx}\right|_{x=L}-\varphi_n(x=L)\right)=T_{xr,n,fin}.
\end{cases}
\tag{30}
$$

The Boundary

In order to calculate the mutual boundary forces, which depend on the overall structural geometry, it is not possible to formulate a general theory; each type of joint is an individual case and requires a specific solution.

We obtain the solution for the pipe joints, assuming that no radial interference is found between the two elements during makeup.

The geometrical dimensions of the structure and the subdivision into six elements (nut, bolt, shell A, shell B, shell C and shell D) are shown in Fig. 9.

The evaluation of the boundary actions for nut and bolt takes place considering the adjacent parts as shells, in order to evaluate numerically the border coefficients and write the kinematic congruence solving equations.

In Fig. 10 the subdivision of the joint in its substructures and mutual actions in forcing condition is shown.

It is possible to recognize the actions already introduced: $M_{x,n,in}$ and $T_{xr,n,in}$ between nut and shell A, $M_{x,n,fin}$ and $T_{xr,n,fin}$ between nut and shell D, $M_{x,b,in}$ and $T_{xr,b,in}$ between bolt and shell C and finally $M_{x,b,fin}$ and $T_{xr,b,fin}$ between bolt and shell D.

The interaction between shell A and shell C on the shoulder is characterized by the axial make-up load P (which loads by compression the lower surface of shell A and generates traction on the section of shell C between the shoulder and the start of the bolt) and tangential load $T_{xr,A,in}=fP/(\pi D_{med,sp})$ due to friction ($D_{med,sp}$ is the medium thickness shoulder diameter).

To be exact a mutual contact moment should be present between shell A and shell C (this is not shown in Fig. 10); said moment cannot be expressed in terms of P, and as such would be a further unknown source of mathematical complexity. So from now on we will ignore it, introducing an approximation which can be verified numerically.

The annotation used is the following: $y_{T_M,\text{shell},fin}^{in}$ ($\varphi_{T_M,\text{shell},fin}^{in}$) for the displacement (rotation) coefficients due to shear/moment

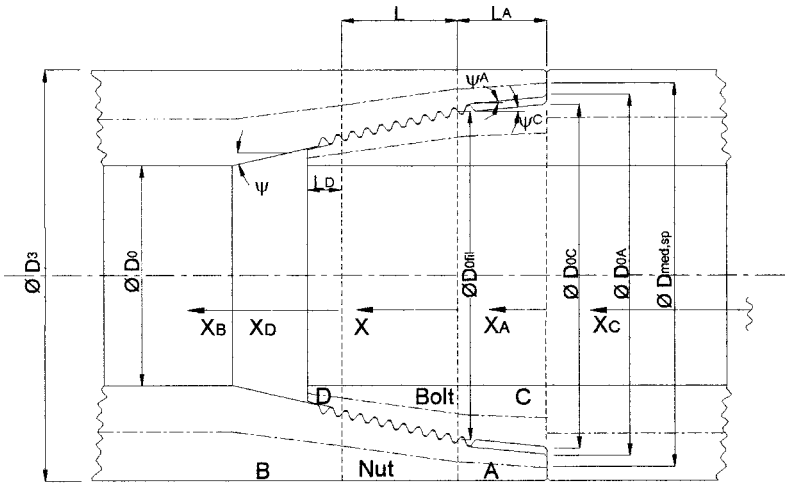


Fig. 9 Pipe joint dimensions, coordinates and substructures

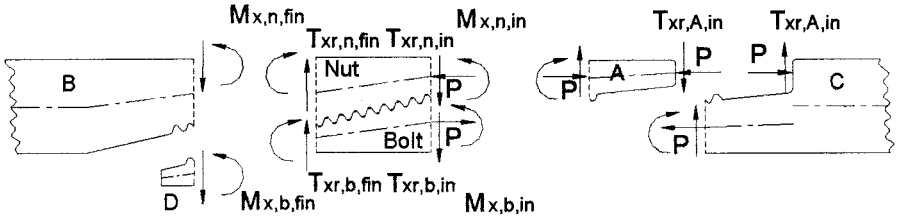


Fig. 10 Mutual forces in a pipe joint's forcing condition

on the initial/final part of ''shell''; $y_{\text{shell},\,fin}^{\,in}$ ($\varphi_{\text{shell},\,fin}^{\,in}$) for the ''border'' displacement (rotation) due to shear/moment on the initial/final part of ''shell.''

The border coefficients and the ''border'' displacements are assumed to be positive, using, in writing the congruence equations, the plus/minus criteria of the shell theory. The congruence equations at the beginning and the end of the nut and bolt are, respectively,

$$\begin{cases} -T_{xr,n,in}{}^{*}y_{T,n,in}+M_{x,n,in}{}^{*}y_{M,n,in}+y_{n,in}=T_{xr,n,in}{}^{*}y_{T,A,fin}+M_{x,n,in}{}^{*}y_{M,A,fin}+y_{A,fin} \\ T_{xr,n,in}{}^{*}\varphi_{T,n,in}-M_{x,n,in}{}^{*}\varphi_{M,n,in}+\varphi_{n,in}=T_{xr,n,in}{}^{*}\varphi_{T,A,fin}+M_{x,n,in}{}^{*}\varphi_{M,A,fin}+\varphi_{A,fin} \end{cases} \tag{31}$$

$$\begin{cases} T_{xr,n,fin}{}^{*}y_{T,n,fin}+M_{x,n,fin}{}^{*}y_{M,n,fin}+y_{n,fin}=-T_{xr,n,fin}{}^{*}y_{T,B,in}+M_{x,n,fin}{}^{*}y_{M,B,in}+y_{B,in} \\ T_{xr,n,fin}{}^{*}\varphi_{T,n,fin}+M_{x,n,fin}{}^{*}\varphi_{M,n,fin}+\varphi_{n,fin}=T_{xr,n,fin}{}^{*}\varphi_{T,B,in}-M_{x,n,fin}{}^{*}\varphi_{M,B,in}+\varphi_{B,in} \end{cases} \tag{32}$$

$$\begin{cases} -T_{xr,b,in}{}^{*}y_{T,b,in}+M_{x,b,in}{}^{*}y_{M,b,in}+y_{b,in}=T_{xr,b,in}{}^{*}y_{T,C,fin}+M_{x,b,in}{}^{*}y_{M,C,fin}+y_{C,fin} \\ T_{xr,b,in}{}^{*}\varphi_{T,b,in}-M_{x,b,in}{}^{*}\varphi_{M,b,in}+\varphi_{b,in}=T_{xr,b,in}{}^{*}\varphi_{T,C,fin}+M_{x,b,in}{}^{*}\varphi_{M,C,fin}+\varphi_{C,fin} \end{cases} \tag{33}$$

$$\begin{cases} T_{xr,b,fin}{}^{*}y_{T,b,fin}+M_{x,b,fin}{}^{*}y_{M,b,fin}+y_{b,fin}=-T_{xr,b,fin}{}^{*}y_{T,D,in}+M_{x,b,fin}{}^{*}y_{M,D,in} \\ T_{xr,b,fin}{}^{*}\varphi_{T,b,fin}+M_{x,b,fin}{}^{*}\varphi_{M,b,fin}+\varphi_{b,fin}=T_{xr,b,fin}{}^{*}\varphi_{T,D,in}-M_{x,b,fin}{}^{*}\varphi_{M,D,in}. \end{cases} \tag{34}$$

The systems (31), (32), (33) and (34) cannot be solved by $T_{xr,b,n}$, $M_{x,b,in}$, $T_{xr,b,fin}$, $M_{x,b,fin}$, $T_{xr,n,in}$, $M_{x,n,in}$, $T_{xr,n,fin}$, and $M_{x,\nu,fin}$; in fact, all the ''border'' displacements of nut and bolt depend on the unknown w which influences the load systems and is in its turn a function of all mutual actions:

$$w=w(M_{x,n,in},T_{xr,n,in},M_{x,n,fin},T_{xr,n,fin},M_{x,b,in},\times T_{xr,b,in},M_{x,b,fin},T_{xr,b,fin}).$$

The problem is thus implicit and requires an iterative solution.

The detailed functional expressions of the ''border'' displacements for nut and bolt are the following:

$$\begin{aligned} y_{n,in}&=y_{n,in}(p,m_n,N_{x,n},M_{x,n,fin},T_{xr,n,fin}) \\ \varphi_{n,in}&=\varphi_{n,in}(p,m_n,N_{x,n},M_{x,n,fin},T_{xr,n,fin}) \\ y_{n,fin}&=y_{n,fin}(p,m_n,N_{x,n},M_{x,n,in},T_{xr,n,in}) \\ \varphi_{n,fin}&=\varphi_{n,fin}(p,m_n,N_{x,n},M_{x,n,in},T_{xr,n,in}) \end{aligned} \tag{35}$$

$$y_{b,in}=y_{b,in}(p,m_b,N_{x,b},M_{x,b,fin},T_{xr,b,fin})$$
$$\varphi_{b,in}=\varphi_{b,in}(p,m_b,N_{x,b},M_{x,b,fin},T_{xr,b,fin})$$
$$y_{b,fin}=y_{b,fin}(p,m_b,N_{x,b},M_{x,b,in},T_{xr,b,in})$$
$$\varphi_{b,fin}=\varphi_{b,fin}(p,m_b,N_{x,b},M_{x,b,in},T_{xr,b,in}) \quad (36)$$

where

$$p=\frac{w}{a}\tan(\beta-\phi)$$

$$m_n=\frac{w}{a}\left(R_n-\frac{D_{fil}}{2}\right), \quad N_{x,n}=\frac{-P+\int_0^s wds}{2\pi R_n},$$

$$m_b=\frac{w}{a}\left(\frac{D_{fil}}{2}-R_b\right), \quad N_{x,b}=\frac{P-\int_0^s wds}{2\pi R_b}.$$

The shell D ''border'' displacements are always void in that on shell D external loads do not act; for shell A, shell B, and shell C ''border'' displacements are directly obtainable as the external actions are known functions of load P:

$$y_{A,fin}=y_{A,fin}(N_{x,A},T_{xr,A,in})$$
$$\varphi_{A,fin}=\varphi_{A,fin}(N_{x,A},T_{xr,A,in})$$
$$y_{B,in}=y_{B,in}(N_{x,B})$$
$$\varphi_{B,in}=\varphi_{B,in}(N_{x,B})$$
$$y_{C,fin}=y_{C,fin}(N_{x,C},T_{xr,A,in})$$
$$\varphi_{C,fin}=\varphi_{C,fin}(N_{x,C},T_{xr,A,in}) \quad (37)$$

where

$$N_{x,A}=\frac{-P}{2\pi R_A}, \quad T_{xr,A,in}=\frac{fP}{\pi D_{med,sp}} \quad N_{x,B}=0,$$

$$N_{x,C}=\frac{P}{2\pi R_C} \quad \text{where } L_C-L_A\leq x_C\leq L_C;$$

$$0 \quad \text{where } x_C<L_C-L_A.$$

The iterative solution used consists in the following steps:

1 Calculate all the border coefficients involved.
2 Calculate the ''border'' displacements of shell A, shell B and shell C.
3 Calculate w with all the void nut and bolt border actions:
$$M_{x,b,in},T_{xr,b,in},M_{x,b,fin},T_{xr,b,fin},M_{x,n,in},T_{xr,n,in},M_{x,n,fin},T_{xr,n,fin}=0.$$
4 Construct the nut and bolt load systems and calculate the associated ''border'' displacements.
5 Solve the congruence systems to find the boundary actions.
6 Calculate w with the actions of Step 5.
7 Iterate Steps 4, 5, and 6 until convergence.

The calculation of border coefficients and displacements takes place solving each shell by a numeric approach based on the finite differences method; also the overall differential system, due to its complexity, is treated in the same manner.

Results

The results given are related to pipe joint API NC38 [9] for which we have developed a finite element method model.

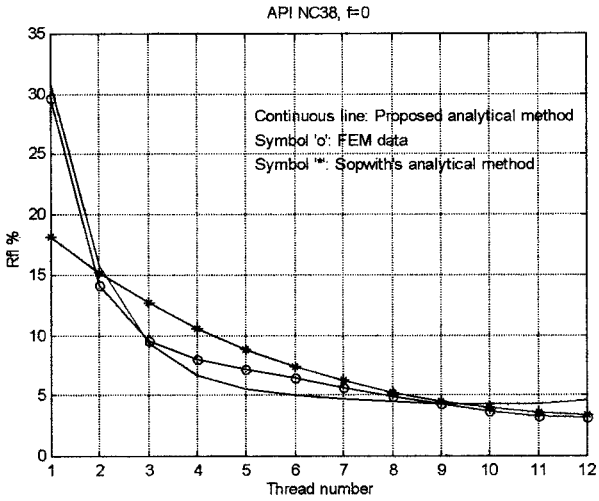


Fig. 11 API NC38 joint *Rfl* in forcing condition with coefficient of friction $f=0$ and axial load $P=543$ kN

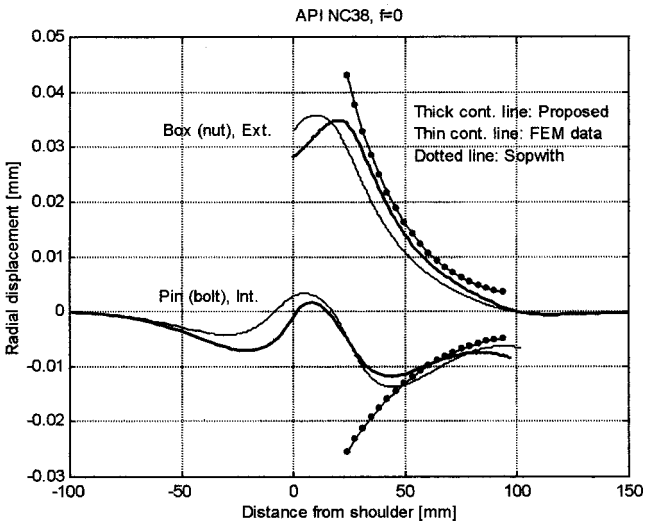


Fig. 12 API NC38 joint radial displacement with coefficient of friction $f=0$ and axial load $P=543$ kN

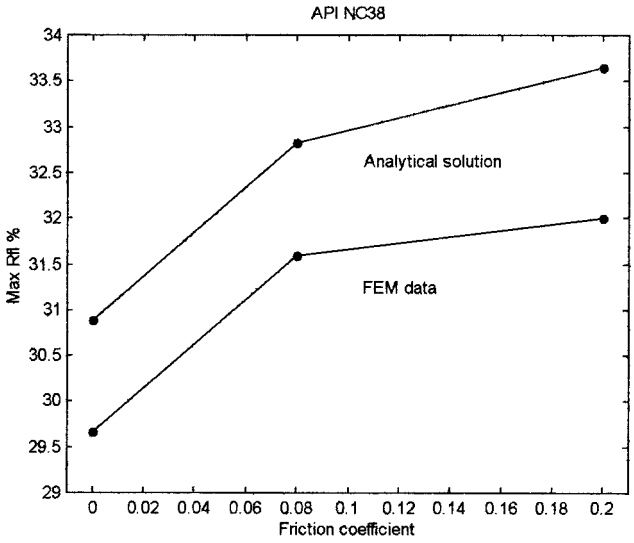


Fig. 13 API NC38 joint maximum *Rfl* value in forcing condition, with three friction values

The comparative results are *Rfl* (ratio of flank load, the percentage of axial load adsorbed by each thread pitch) and the radial displacement outside the box (''nut'' element) and inside the pin (''bolt'' element).

In order to get *Rfl* from the $w(s)$ function we only need to integrate on each thread pitch:

$$Rfl_j = \frac{100}{P} \int_{s_{\text{pitch } j}}^{s_{\text{pitch } j+1}} w(s)ds.$$

The results obtained (Fig. 11, Fig. 12) show an affinity, both in *Rfl* terms and in displacement terms, between the theoretical and finite element method values.

It is interesting to analyze the dependence, shown in Fig. 13, between the maximum *Rfl* value and the friction coefficient.

Discussion

The results shown by this new approach clearly indicate how important the correct reproduction of the threaded joint bending behavior is, especially for those with thinner characteristics and complex boundary substructures, such as the pipe joint API NC38 examined.

For this type of joint Sopwith's theory (generally accepted for common nuts and bolts) is no longer valid as it underestimates by approximately 40 percent the load absorbed by the first pitch, mainly because of the inexact deformed shape produced by the use of the thick cylinder formulas for the δ_{II} factor, which is now calculated with suitable equations for conical shells.

One of the most valuable characteristics of the theoretical approach is the speed with which the solution is obtained by the variation of the geometric parameters of the structure, which allows us to carry out the joint optimization process, otherwise difficult by the finite element method.

The optimization process involves only the load distribution without providing information on the local stress, which cannot be obtained by elaborating the internal element forces due to stress concentration induced by the notch radius.

However, the stress concentration can be limited using suitable constraints on the notch radius r, on the flank contact length L_f, and on the minimal resistance area $A_n(s=0)$.

Therefore, this proposed new approach is a very quick and efficient tool in the design and optimization process of threaded connections, which can significantly reduce the number of numerical analyses required.

References

[1] Goodier, J. N., 1940, ''Distribution of Load in Threads of Screws,'' Trans. ASME, **62**, pp. A10–A16.

[2] Sopwith, D. G., 1948, ''The Distribution of Load in Screw Threads,'' Inst. Mech. Eng. Appl. Mech. Proc., **159**, pp. 373–383.

[3] Stoeckly, E. E., and Macke, H. J., 1952, ''Effect of Taper on Screw Thread Load Distribution,'' Trans. ASME, **74**, pp. 103–112.

[4] Timoshenko, S. P., and Goodier, J. N., 1982, *Theory of Elasticity*, 3rd Ed., Mc.Graw-Hill, New York.

[5] Mindlin, R. D., 1951, ''Influence of Rotatory Inertia and Shear on Flexural Motion of Isotropic Elastic Plates,'' ASME J. Appl. Mech., **18**, p. 31.

[6] Reissner, E., 1985, ''Reflection on the Theory of Elastic Plates,'' Appl. Mech. Rev., **38**, p. 1451.

[7] Timoshenko, S. P., and Woinowsky-Krieger, S., 1959, *Theory of Plates and Shells*, McGraw-Hill, New York.

[8] Corradi Dell'acqua, L., 1992, ''Meccanica Delle Strutture, Vol. 1—II Comportamento dei Mezzi Continui,'' McGraw-Hill, Milano.

[9] American Petroleum Institute, 1994, ''Specification for Rotary Drill Stem Elements—Specification 7,'' 38th Ed.—1/4/1994.

Explicit Modal Analysis of an Axially Loaded Timoshenko Beam With Bending-Torsion Coupling

J. R. Banerjee
Department of Mechanical Engineering
and Aeronautics,
City University,
Northampton Square,
London EC1V OHB, UK

Exact analytical expressions for the natural frequencies and mode shapes of a uniform bending-torsion coupled Timoshenko beam are presented. The beam is taken to be axially loaded, and for which cantilever end conditions apply. A symbolic computing package (REDUCE) has been used, which simplifies the analysis greatly. Results are presented for an illustrative example, which confirms the accuracy of the method, and provides a convenient benchmark for the validation of the finite element or other alternative approximate methods. The approach offers the prospect of aeroelastic development, and is computationally efficient, thus holding out the promise of eventual optimization.
[S0021-8936(00)01002-3]

1 Introduction

Bernoulli-Euler and Timoshenko beams in flexure, have been studied for many years, and there are excellent texts on the subject ([1,2]). Explicit frequency equation and mode shapes for Bernoulli-Euler beams have also been available ([3]) for many years, but the corresponding results for Timoshenko beams and axially loaded Timoshenko beams have only recently appeared ([4–6]) The comparatively more difficult problem of the free vibration of a bending-torsion coupled beam with or without the effects of shear deformation, rotatory inertia, and axial load has been studied for some time ([7–14]) without the explicit development of the associated natural frequency and mode shape formulas. These are presented here for the first time and relate to axially loaded bending-torsion coupled Timoshenko beams. The degenerate cases which converge back to established results are shown to be readily obtained by suitable choice of parameters.

In the derivation of frequency equation and mode shapes, the introduction of the bending-torsion coupling increases the analytical complexity of the problem to the point at which it is scarcely feasible to proceed without the use of a symbolic algebraic package. The use of such a package—REDUCE ([15])—has enabled the bending-torsion coupled problem to become manageable.

In the subsequent text, the governing differential equations of a bending-torsion coupled Timoshenko beam are solved and boundary conditions imposed, so that the frequency equation and mode shapes are formulated. Without the use of symbolic computation, numerical methods would now be essential for further development. But, REDUCE ([15]) has been used here to obtain a complete algebraic reduction, thus replacing the numerical problem to an analytical one. An illustrative example is provided.

2. Theory

A straight uniform beam element of length L and of a tee cross section is shown in Fig. 1, with the mass axis and the elastic axis which are, respectively, the locus of the mass centers and the shear centers of the cross section, being separated by a distance x_α. The theory which follows is applicable to any singly symmetric cross section with noncoincident mass and shear center, but the tee cross section is shown as an illustrative example only for convenience. In the right-hand coordinate system of Fig. 1, the elastic axis, which is assumed to coincide with the Y-axis, is permitted flexural translation $h(y,t)$ in the Z-direction and torsional rotation $\psi(y,t)$ about the Y-axis, where y and t denote distance from the origin and time, respectively. A constant tensile axial load P is assumed to act through the centroid (mass center) of the cross section, and P can be negative, so that compression is included in the theory. Note that the axial displacement (in the Y-direction) and bending displacement in the XY-plane are not considered because they are uncoupled (with torsion), and therefore, can be dealt with separately using existing theories ([3–5]).

The governing partial differential equations of motion for the coupled bending-torsional free natural vibration of the axially loaded Timoshenko beam shown in Fig. 1 are ([13])

$$EI\theta'' + kAG(h' - \theta) - \rho I\ddot{\theta} = 0 \tag{1}$$

$$kAG(h'' - \theta') + P(h'' - x_\alpha\psi'') - m(\ddot{h} - x_\alpha\ddot{\psi}) = 0 \tag{2}$$

$$GJ\psi'' + P\{(I_\alpha/m)\psi'' - x_\alpha h''\} - I_\alpha\ddot{\psi} + mx_\alpha\ddot{h} = 0 \tag{3}$$

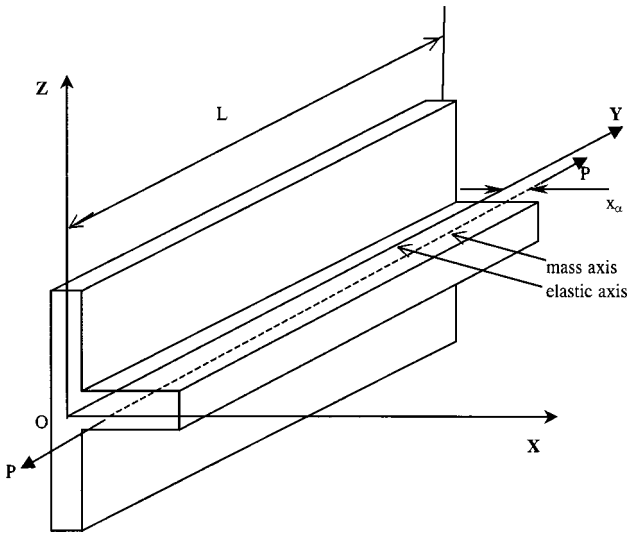


Fig. 1 Coordinate system and notation for an axially loaded bending-torsion coupled Timoshenko beam

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 3, 1998; final revision, November 8, 1999. Associate Technical Editor: V. K. Kinra. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

where E is the Young's modulus, G is the shear modulus, and ρ is the density of the material; EI, GJ and kAG are, respectively, the bending, torsional, and shear rigidities of the beam; I is the second moment of area of the beam cross section about the X-axis; k is the section shape factor; A is the cross-sectional area; $m\ (=\rho A)$ is the mass per unit length; I_α is the polar mass moment of inertia per unit length about the Y-axis (i.e., an axis through the shear center); θ is the angle of rotation in radians of the cross section due to bending alone (so that the total slope h' equals the sum of slopes due to bending and due to shear deformation) and primes and dots denote differentiation with respect to position y and time t, respectively.

Equations (1)–(3) completely define the coupled bending-torsional free vibration of an axially loaded uniform Timoshenko beam.

If a sinusoidal variation of h, θ, and ψ, with circular frequency ω, is assumed, then

$$\left.\begin{aligned} h(y,t) &= H(y)\sin\omega t \\ \theta(y,t) &= \Theta(y)\sin\omega t \\ \psi(y,t) &= \Psi(y)\sin\omega t \end{aligned}\right\} \tag{4}$$

where $H(y)$, $\theta(y)$, and $\Psi(y)$ are the amplitudes of the sinusoidally varying vertical displacement, bending rotation, and twist, respectively.

Substituting Eqs. (4) into Eqs. (1)–(3) gives

$$EI\Theta'' + kAG(H' - \Theta) + \rho I\omega^2\Theta = 0 \tag{5}$$

$$kAG(H'' - \Theta') + P(H'' - x_\alpha\Psi'') + m\omega^2 H - m\omega^2 x_\alpha\Psi = 0 \tag{6}$$

$$GJ\Psi'' + P\{(I_\alpha/m)\Psi'' - x_\alpha H''\} + I_\alpha\omega^2\Psi - m\omega^2 x_\alpha H = 0. \tag{7}$$

Introducing the nondimensional length ξ, and the differential operator D where

$$\xi = y/L$$

$$D = d/d\xi. \tag{8}$$

Eqs. (5)–(7) can be written in matrix form as follows:

$$\begin{bmatrix} kAGLD & EID^2 - kAGL^2 + \rho I\omega^2 L^2 & 0 \\ kAGD^2 + PD^2 + m\omega^2 L^2 & -kAGLD & -Px_\alpha D^2 - m\omega^2 x_\alpha L^2 \\ -Px_\alpha D^2 - m\omega^2 x_\alpha L^2 & 0 & GJD^2 + (PI_\alpha/m)D^2 + I_\alpha\omega^2 L^2 \end{bmatrix} \begin{bmatrix} H \\ \Theta \\ \Psi \end{bmatrix} = 0. \tag{9}$$

The coefficients of the differential operators D and D^2 are constants in Eq. (9), and one can eliminate any two of the three variables H, θ, and Ψ to obtain the differential equation in terms of only one variable. This can be accomplished either by using Gauss elimination or by expanding the determinant of the above 3×3 matrix and applying rules of linear operators to give

$$(D^6 + \bar{a}D^4 - \bar{b}D^2 - \bar{c})W = 0 \tag{10}$$

where

$$W = H,\ \Theta \text{ or } \Psi \tag{11}$$

and

$$\left.\begin{aligned} \bar{a} &= b^2r^2 + \frac{\{a^2b^2(1+2c^2p^2s^2) - a^2c^2p^4 - b^2(p^2 - b^2s^2)\}}{\{b^2(1+p^2s^2) + a^2p^2(1+c^2p^2s^2)\}} \\ \bar{b} &= \frac{\{b^4(1-b^2r^2s^2) - a^2b^4c^2s^2(1+2p^2r^2) + a^2b^2(2c^2p^2 - b^2r^2)\}}{\{b^2(1+p^2s^2) + a^2p^2(1+c^2p^2s^2)\}} \\ \bar{c} &= \frac{a^2b^4c^2(1-b^2r^2s^2)}{\{b^2(1+p^2s^2) + a^2p^2(1+c^2p^2s^2)\}} \end{aligned}\right\} \tag{12}$$

with

$$\left.\begin{aligned} a^2 &= I_\alpha\omega^2L^2/GJ \\ b^2 &= m\omega^2L^4/EI \\ c^2 &= 1 - mx_\alpha^2/I_\alpha = I_G/I_\alpha \\ p^2 &= PL^2/EI \\ r^2 &= I/AL^2 \\ s^2 &= EI/kAGL^2 \end{aligned}\right\} \tag{13}$$

where I_G is the polar mass moment of inertia per unit length about an axis through the centroid.

A detailed numerical study carried out by the author has shown that for a wide range of physical (engineering) problems, the coefficients $\bar{a}$, $\bar{b}$, and $\bar{c}$ are always positive so that the solution of the differential Eq. (10) is ([10])

$$W(\xi) = C_1^*\cosh\alpha\xi + C_2^*\sinh\alpha\xi + C_3^*\cos\beta\xi + C_4^*\sin\beta\xi + C_5^*\cos\gamma\xi + C_6^*\sin\gamma\xi \tag{14}$$

where $C_1^* - C_6^*$ are constants and

$$\left.\begin{aligned} \alpha &= [2(q/3)^{1/2}\cos(\phi/3) - \bar{a}/3]^{1/2} \\ \beta &= [2(q/3)^{1/2}\cos\{(\pi - \phi)/3\} + \bar{a}/3]^{1/2} \\ \gamma &= [2(q/3)^{1/2}\cos\{(\pi + \phi)/3\} + \bar{a}/3]^{1/2} \end{aligned}\right\} \tag{15}$$

with

$$\left.\begin{aligned} q &= \bar{b} + \bar{a}^2/3 \\ \phi &= \cos^{-1}[(27\bar{c} - 9\bar{a}\bar{b} - 2\bar{a}^3)/\{2(\bar{a}^2 + 3\bar{b})^{3/2}\}]. \end{aligned}\right\} \tag{16}$$

Equation (14) represents the solution for the bending displacement $H(\xi)$, bending rotation $\Theta(\xi)$, and torsional rotation $\Psi(\xi)$. Thus

$$H(\xi) = A_1\cosh\alpha\xi + A_2\sinh\alpha\xi + A_3\cos\beta\xi + A_4\sin\beta\xi + A_5\cos\gamma\xi + A_6\sin\gamma\xi \tag{17}$$

$$\Theta(\xi) = B_1\sinh\alpha\xi + B_2\cosh\alpha\xi + B_3\sin\beta\xi + B_4\cos\beta\xi + B_5\sin\gamma\xi + B_6\cos\gamma\xi \tag{18}$$

$$\Psi(\xi)=C_1\cosh\alpha\xi+C_2\sinh\alpha\xi+C_3\cos\beta\xi+C_4\sin\beta\xi+C_5\cos\gamma\xi+C_6\sin\gamma\xi \tag{19}$$

where A_1-A_6, B_1-B_6, and C_1-C_6 are the three different sets of constants.

Substituting Eqs. (17) and (18) into Eq. (5) shows that

$$\left.\begin{aligned} &B_1=(\overline{\alpha}/L)A_1; \quad B_3=-(\overline{\beta}/L)A_3; \quad B_5=-(\overline{\gamma}/L)A_5\\ &B_2=(\overline{\alpha}/L)A_2; \quad B_4=(\overline{\beta}/L)A_4; \quad B_6=(\overline{\gamma}/L)A_6 \end{aligned}\right\} \tag{20}$$

where

$$\left.\begin{aligned} \overline{\alpha}&=\alpha/(1-b^2r^2s^2-\alpha^2s^2)\\ \overline{\beta}&=\beta/(1-b^2r^2s^2+\beta^2s^2)\\ \overline{\gamma}&=\gamma/(1-b^2r^2s^2+\gamma^2s^2) \end{aligned}\right\} \tag{21}$$

Then substituting Eqs. (17) and (19) into Eq. (7) gives

$$\left.\begin{aligned} &C_1=(k_\alpha/x_\alpha)A_1; \quad C_3=(k_\beta/x_\alpha)A_3; \quad C_5=(k_\gamma/x_\alpha)A_5\\ &C_2=(k_\alpha/x_\alpha)A_2; \quad C_4=(k_\beta/x_\alpha)A_4; \quad C_6=(k_\gamma/x_\alpha)A_6 \end{aligned}\right\} \tag{22}$$

where

$$\left.\begin{aligned} k_\alpha&=a^2(1-c^2)(b^2+p^2\alpha^2)/\{a^2(b^2+p^2\alpha^2)+b^2\alpha^2\}\\ k_\beta&=a^2(1-c^2)(b^2-p^2\beta^2)/\{a^2(b^2-p^2\beta^2)-b^2\beta^2\}\\ k_\gamma&=a^2(1-c^2)(b^2-p^2\gamma^2)/\{a^2(b^2-p^2\gamma^2)-b^2\gamma^2\}. \end{aligned}\right\} \tag{23}$$

The expressions for the bending moment $M(\xi)$, the shear force $S(\xi)$, and the torque $T(\xi)$ are obtained from Eqs. (17)–(19), after some simplification, as ([13])

$$\begin{aligned} M(\xi)&=-(EI/L)\frac{d\Theta}{d\xi}\\ &=-(EI/L^2)\{A_1\alpha\overline{\alpha}\cosh\alpha\xi+A_2\alpha\overline{\alpha}\sinh\alpha\xi\\ &\quad -A_3\beta\overline{\beta}\cos\beta\xi-A_4\beta\overline{\beta}\sin\beta\xi-A_5\gamma\overline{\gamma}\cos\gamma\xi\\ &\quad -A_6\gamma\overline{\gamma}\sin\gamma\xi\} \end{aligned} \tag{24}$$

$$\begin{aligned} S(\xi)&=(EI/L^3)\left[L\frac{d^2\Theta}{d\xi^2}-p^2\left(\frac{dH}{d\xi}-x_\alpha\frac{d\Psi}{d\xi}\right)+b^2r^2\Theta L\right]\\ &=(EI/L^3)\{A_1\overline{\alpha}g_\alpha\sinh\alpha\xi+A_2\overline{\alpha}g_\alpha\cosh\alpha\xi+A_3\overline{\beta}g_\beta\sin\beta\xi\\ &\quad -A_4\overline{\beta}g_\beta\cos\beta\xi+A_5\overline{\gamma}g_\gamma\sin\gamma\xi-A_6\overline{\gamma}g_\gamma\cos\gamma\xi\} \end{aligned} \tag{25}$$

and

$$\begin{aligned} T(\xi)&=(GJ/L)\left[(1+p^2a^2/b^2)\frac{d\Psi}{d\xi}-\{p^2a^2(1-c^2)/(x_\alpha b^2)\}\frac{dH}{d\xi}\right]\\ &=(GJ/L)\{A_1(\alpha e_\alpha/x_\alpha)\sinh\alpha\zeta+A_2(\alpha e_\alpha/x_\alpha)\cosh\alpha\xi\\ &\quad -A_3(\beta e_\beta/x_\alpha)\sin\beta\xi+A_4(\beta e_\beta/x_\alpha)\cos\beta\xi\\ &\quad -A_5(\gamma e_\gamma/x_\alpha)\sin\gamma\xi+A_6(\gamma e_\gamma/x_\alpha)\cos\gamma\xi\} \end{aligned} \tag{26}$$

where

$$\left.\begin{aligned} g_\alpha&=\alpha^2+b^2r^2-p^2(1-k_\alpha)\alpha/\overline{\alpha}\\ g_\beta&=\beta^2-b^2r^2+p^2(1-k_\beta)\beta/\overline{\beta}\\ g_\gamma&=\gamma^2-b^2r^2+p^2(1-k_\gamma)\gamma/\overline{\gamma} \end{aligned}\right\} \tag{27}$$

and

$$\left.\begin{aligned} e_\alpha&=(1+a^2p^2/b^2)k_\alpha-a^2p^2(1-c^2)/b^2\\ e_\beta&=(1+a^2p^2/b^2)k_\beta-a^2p^2(1-c^2)/b^2\\ e_\gamma&=(1+a^2p^2/b^2)k_\gamma-a^2p^2(1-c^2)/b^2. \end{aligned}\right\} \tag{28}$$

2.1 Frequency Equation. The end conditions for the cantilever beam are as follows:

$$\text{at the built-in end } (\xi=0): \quad H=0, \ \Theta=0, \text{ and } \Psi=0 \tag{29}$$

$$\text{at the free end } (\xi=1): \quad S=0, \ M=0, \text{ and } T=0. \tag{30}$$

Substituting Eq. (29) in Eqs. (17)–(19), and (30) in Eqs. (24)–(26) gives

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0\\ 0 & \overline{\alpha} & 0 & \overline{\beta} & 0 & \overline{\gamma}\\ k_\alpha & 0 & k_\beta & 0 & k_\gamma & 0\\ -\overline{\alpha}g_\alpha S_{h\alpha} & -\overline{\alpha}g_\alpha C_{h\alpha} & -\overline{\beta}g_\beta S_\beta & \overline{\beta}g_\beta C_\beta & -\overline{\gamma}g_\gamma S_\gamma & \overline{\gamma}g_\gamma C_\gamma\\ \alpha\overline{\alpha}C_{h\alpha} & \alpha\overline{\alpha}S_{h\alpha} & -\beta\overline{\beta}C_\beta & -\beta\overline{\beta}S_\beta & -\gamma\overline{\gamma}C_\gamma & -\gamma\overline{\gamma}S_\gamma\\ \alpha e_\alpha S_{h\alpha} & \alpha e_\alpha C_{h\alpha} & -\beta e_\beta S_\beta & \beta e_\beta C_\beta & -\gamma e_\gamma S_\gamma & \gamma e_\gamma C_\gamma \end{bmatrix}\begin{bmatrix} A_1\\ A_2\\ A_3\\ A_4\\ A_5\\ A_6 \end{bmatrix}=0 \tag{31}$$

where

$$\begin{aligned} &C_{h\alpha}=\cosh\alpha; \quad C_\beta=\cos\beta; \quad C_\gamma=\cos\gamma\\ &S_{h\alpha}=\sinh\alpha; \quad S_\beta=\sin\beta; \quad S_\gamma=\sin\gamma. \end{aligned} \tag{32}$$

Equation (31) may be written in matrix form as

$$\mathbf{BA}=0. \tag{33}$$

The necessary and sufficient condition for nonzero elements in the column vector $\mathbf{A}$ of Eq. (31) is that $\Delta=|\mathbf{B}|$ shall be zero, and the vanishing of Δ determines the natural frequencies of the system in the usual way. Thus the frequency equation for the cantilever can be obtained for the nontrivial solution as

$$\Delta=|\mathbf{B}|=0. \tag{34}$$

Expanding the 6×6 determinant Δ of $\mathbf{B}$ algebraically is quite a formidable task, but became feasible with the recent advances in symbolic computing. Thus most of the work reported here, was carried out using the software REDUCE ([15]) in expanding the determinant $|\mathbf{B}|$, and more importantly in simplifying the expression for Δ. The final expression obtained for Δ is given below which is not necessarily in the shortest possible form, but is surprisingly concise.

$$\begin{aligned} \Delta&=\lambda C_\beta C_\gamma C_{h\alpha}-\xi_1C_\gamma(\zeta_1+\eta_1S_\beta S_{h\alpha})-\xi_2C_{h\alpha}(\zeta_2+\eta_2S_\beta S_\gamma)\\ &\quad -\xi_3C_\beta(\zeta_3+\eta_3S_\gamma S_{h\alpha}) \end{aligned} \tag{35}$$

where

$$\xi_1=\overline{\alpha}\overline{\beta}; \quad \xi_2=\overline{\beta}\overline{\gamma}; \quad \xi_3=\overline{\gamma\alpha} \tag{36}$$

$$\mu_1=k_\alpha-k_\beta; \quad \mu_2=k_\beta-k_\gamma; \quad \mu_3=k_\gamma-k_\alpha \tag{37}$$

$$\nu_1=\alpha\overline{\beta}e_\alpha g_\beta+\beta\overline{\alpha}e_\beta g_\alpha; \quad \nu_2=\beta\overline{\gamma}e_\beta g_\gamma-\gamma\overline{\beta}e_\gamma g_\beta; \quad \nu_3=\gamma\overline{\alpha}e_\gamma g_\alpha+\alpha\overline{\gamma}e_\alpha g_\gamma \tag{38}$$

Table 1 Material properties and other data used in the free vibration analysis of an axially loaded bending-torsion coupled Timoshenko beam ([8,13])

Beam Parameter	Numerical Value
E (N/m^2)	68.9×10^9
G (N/m^2)	26.5×10^9
ρ (kg/m^3)	2711
A (m^2)	3.08×10^{-4}
I (m^4)	9.26×10^{-8}
k	0.5
EI (Nm^2)	6380
GJ (Nm^2)	43.46
kAG (N)	4.081×10^6
m (kg/m)	0.835
I_α (kgm)	0.501×10^{-3}
x_α (m)	0.155
L (m)	0.82
P (N)	1790
p^2	0.1886
r^2	0.000447
s^2	0.00233

$$\eta_1=\alpha\mu_3\nu_2-\beta\mu_2\nu_3;\quad \eta_2=\beta\mu_1\nu_3-\gamma\mu_3\nu_1;\quad \eta_3=\gamma\mu_2\nu_1+\alpha\mu_1\nu_2 \tag{39}$$

$$\zeta_1=\alpha\mu_2\nu_3+\beta\mu_3\nu_2;\quad \zeta_2=\beta\mu_3\nu_1-\gamma\mu_1\nu_3;\quad \zeta_3=\gamma\mu_1\nu_2-\alpha\mu_2\nu_1 \tag{40}$$

and

$$\lambda=\alpha\bar{\alpha}^2\mu_2\nu_2+\beta\bar{\beta}^2\mu_3\nu_3-\gamma\bar{\gamma}^2\mu_1\nu_1 \tag{41}$$

with α, β, γ; $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$; k_α, k_β, k_γ; g_α, g_β, g_γ; e_α, e_β, e_γ; and $C_{h\alpha}$, C_β, C_γ, $S_{h\alpha}$, S_β, S_γ already defined in Eqs. (15), (21), (23), (27), (28), and (32), respectively. Note that it can be readily verified with the help of Eqs. (13)–(16) and (23) that the value of the determinant $\Delta=|\mathbf{B}|$ is zero when the frequency (ω) is zero. This known value of $\Delta=|\mathbf{B}|=0$ at $\omega=0$ (which corresponds to a beam with no inertial loading, i.e., at rest), can be always used to avoid any numerical problem of overflow at zero frequency when computing the value of Δ. Thus for any other (nontrivial) values of ω, the expression for Δ given by Eq. (35) can be used in locating the natural frequencies by successively tracking the changes of its sign.

2.2 Mode Shapes. Once the natural frequencies ω_n are found from Eq. (35), the modal vector $\mathbf{A}$ (in which one element may be fixed arbitrarily) is found in the usual way, namely by deleting one row of the 6th order matrix of Eq. (31) and solving for the five remaining constants in terms of the arbitrarily chosen one.

Thus, if A_1 is chosen to be the one in terms of which the remaining constants A_2-A_6 are to be expressed, as in the present case, the matrix Eq. (31), will take the following reduced order form. (Note that terms relating to A_1 are taken to the right-hand side.)

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ \bar{\alpha} & 0 & \bar{\beta} & 0 & \bar{\gamma} \\ 0 & k_\beta & 0 & k_\gamma & 0 \\ -\bar{\alpha}g_\alpha C_{h\alpha} & -\bar{\beta}g_\beta S_\beta & \bar{\beta}g_\beta C_\beta & -\bar{\gamma}g_\gamma S_\gamma & \bar{\gamma}g_\gamma C_\gamma \\ \alpha\bar{\alpha}S_{h\alpha} & -\beta\bar{\beta}C_\beta & -\beta\bar{\beta}S_\beta & -\gamma\bar{\gamma}C_\gamma & -\gamma\bar{\gamma}S_\gamma \end{bmatrix} \times \begin{bmatrix} A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ -k_\alpha \\ \bar{\alpha}g_\alpha S_{h\alpha} \\ -\alpha\bar{\alpha}C_{h\alpha} \end{bmatrix} A_1 \tag{42}$$

The symbolic computing package REDUCE ([15]) was further used to solve the above system of equations giving the following mode shape coefficients in terms of A_1

$$\left.\begin{aligned} A_2&=A_1[\bar{\beta}\bar{\gamma}(\kappa_2-\alpha\bar{\alpha}\mu_2\varepsilon_2C_{h\alpha}+\bar{\alpha}\mu_2\delta_2g_\alpha S_{h\alpha}-\sigma_2C_\beta C_\gamma+\tau_2S_\beta S_\gamma)/\chi]\\ A_3&=A_1[\mu_3/\mu_2]\\ A_4&=A_1[\bar{\gamma}\bar{\alpha}(\kappa_3+\beta\bar{\beta}\mu_3\varepsilon_3C_\beta+\bar{\beta}\mu_3\delta_3g_\beta S_\beta+\sigma_3C_\gamma C_{h\alpha}+\tau_3S_\gamma S_{h\alpha})/\chi]\\ A_5&=A_1[\mu_1/\mu_2]\\ A_6&=A_1[\bar{\alpha}\bar{\beta}(\kappa_1-\gamma\bar{\gamma}\mu_1\varepsilon_1C_\gamma-\bar{\gamma}\mu_1\delta_1g_\gamma S_\gamma+\sigma_1C_\beta C_{h\alpha}-\tau_1S_\beta S_{h\alpha})/\chi]\end{aligned}\right\} \tag{43}$$

where μ_1, μ_2, and μ_3 have already been defined in Eqs. (37) and the following further variables are introduced to compute the parameters within the square brackets.

$$\varepsilon_1=g_\alpha C_{h\alpha}+g_\beta C_\beta;\quad \varepsilon_2=g_\beta C_\beta-g_\gamma C_\gamma;\quad \varepsilon_3=g_\gamma C_\gamma+g_\alpha C_{h\alpha} \tag{44}$$

$$\delta_1=\alpha S_{h\alpha}+\beta S_\beta;\quad \delta_2=\beta S_\beta-\gamma S_\gamma;\quad \delta_3=\gamma S_\gamma+\alpha S_{h\alpha} \tag{45}$$

$$\tau_1=\alpha\bar{\beta}\mu_3g_\beta+\beta\bar{\alpha}g_\alpha\mu_2;\quad \tau_2=\beta\bar{\gamma}\mu_1g_\gamma-\gamma\bar{\beta}g_\beta\mu_3;\quad \tau_3=\gamma\bar{\alpha}\mu_2g_\alpha+\bar{\gamma}\alpha\mu_1g_\gamma \tag{46}$$

Table 2 Natural frequencies of an axially bending-torsion coupled Timoshenko beam ([8,13]) with cantilever end condition.

Frequency Number	Natural Frequency (rad/s) $P=0\,(p^2=0)$		$P=1790\ N\ (p^2=0.1886)$		$P=-1790\ N\ (p^2=-0.1886)$	
	$r^2=s^2=0$	$r^2\neq0, s^2\neq0$	$r^2=s^2=0$	$r^2\neq0, s^2\neq0$	$r^2=s^2=0$	$r^2\neq0, s^2\neq0$
1	393.4	391.7	407.4	405.8	378.5	376.8
2	818.0	816.0	828.6	826.7	806.9	805.1
3	1641	1629	1661	1649	1621	1609
4	2648	2632	2634	2668	2611	2595

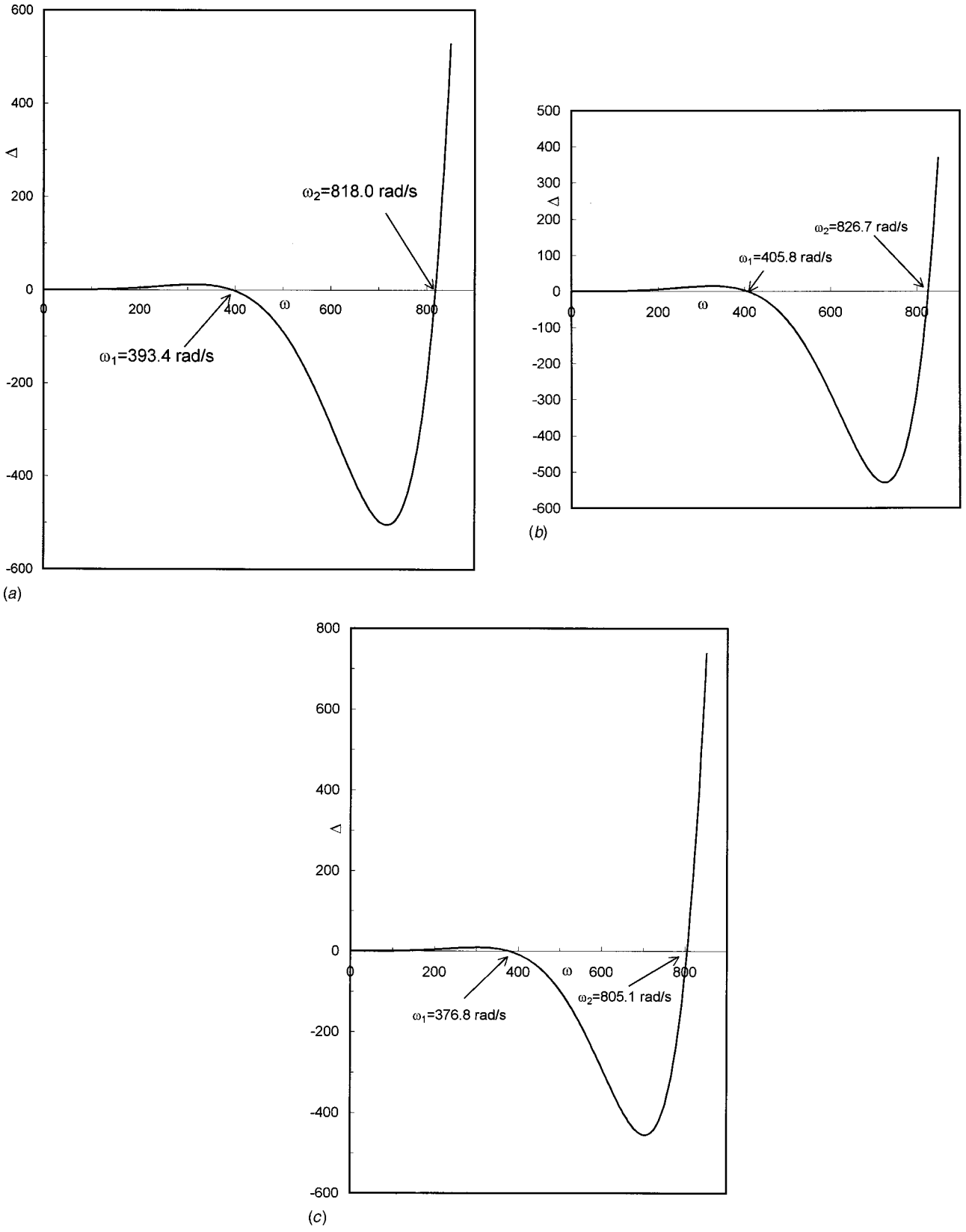


Fig. 2 (*a*) **Variation of Δ with frequency (*ω*) for the case when $p^2=r^2=s^2=0$; (*b*) variation of Δ with frequency (*ω*) for the case when $p^2=0.1886$, $r^2=0.000447$, and $s^2=0.00233$; (*c*) variation of Δ with frequency (*ω*) for the case when $p^2=-0.1886$, $r^2=0.000447$, and $s^2=0.00233$**

$$\sigma_1=\alpha\bar{\alpha}\mu_2 g_\beta-\beta\bar{\beta}\mu_3 g_\alpha;\quad \sigma_2=\beta\bar{\beta}\mu_3 g_\gamma-\gamma\bar{\gamma}\mu_1 g_\beta;$$
$$\sigma_3=\gamma\bar{\gamma}\mu_1 g_\alpha-\alpha\bar{\alpha}\mu_2 g_\gamma \tag{47}$$

$$\kappa_1=\alpha\bar{\alpha}\mu_2 g_\alpha-\beta\bar{\beta}\mu_3 g_\beta;\quad \kappa_2=\beta\bar{\beta}\mu_3 g_\beta-\gamma\bar{\gamma}\mu_1 g_\gamma;$$
$$\kappa_3=\gamma\bar{\gamma}\mu_1 g_\gamma-\alpha\bar{\alpha}\mu_2 g_\alpha \tag{48}$$

and

$$\chi=\bar{\alpha}\bar{\beta}\bar{\gamma}\mu_2(\alpha\varepsilon_2 S_{h\alpha}-\beta\varepsilon_3 S_\beta+\gamma\varepsilon_1 S_\gamma). \tag{49}$$

Note that α, β, γ, $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$; k_α, k_β, k_γ; g_α, g_β, g_γ; $C_{h\alpha}$, $S_{h\alpha}$, C_β, S_β, C_γ, and S_γ appearing in Eqs. (37), and (43)–(49) must be calculated for the particular natural frequency ω_n at which the mode shape is required.

Thus the mode shape of the bending-torsion coupled beam with cantilever end condition is given in explicit form by rewriting Eqs. (17)–(19) with the help of Eqs. (20)–(23) and (43) in the form

$$H(\xi)=A_1(\cosh\alpha\xi+R_1\sinh\alpha\xi+R_2\cos\beta\xi+R_3\sin\beta\xi+R_4\cos\gamma\xi+R_5\sin\gamma\xi) \tag{50}$$

$$\Theta(\xi)=A_1(\bar{\alpha}\sinh\alpha\xi+\bar{\alpha}R_1\cosh\alpha\xi-\bar{\beta}R_2\sin\beta\xi+\bar{\beta}R_3\cos\beta\xi-\bar{\gamma}R_4\sin\gamma\xi+\bar{\gamma}R_5\cos\gamma\xi)/L \tag{51}$$

$$\Psi(\xi)=A_1(k_\alpha\cosh\alpha\xi+R_1k_\alpha\sinh\alpha\xi+R_2k_\beta\cos\beta\xi+R_3k_\beta\sin\beta\xi+R_4k_\gamma\cos\gamma\xi+R_5k_\gamma\sin\gamma\xi)/x_\alpha \tag{52}$$

where the ratios R_1, R_2, R_3, R_4, and R_5 are, respectively, A_2/A_1, A_3/A_1, A_4/A_1, A_5/A_1, and A_6/A_1, and follow from Eqs. (43).

In spite of the apparent complexity of the frequency and mode shape equations given above, results for the degenerate case of the bending-torsion coupled Timoshenko Beam and Bernoulli-Euler beam can be obtained by substituting in the data $p^2=0$ and $p^2=r^2=s^2=0$, respectively. Note that any one or more of the terms p^2, r^2, s^2 which uniquely describe the effect of *axial load*, *rotatory inertia*, and *shear deformation*, respectively, can be set to zero (either individually or in any combination) to obtain the degenerate cases.

3 Discussion of Results

An illustrative example on the application of the frequency equation and mode shape expressions derived above is chosen from the existing literature ([8,13]). It is a cantilever beam with thin-walled semi-circular cross section, and for which substantial coupling between the bending and torsional modes of deformation is prevalent. The material properties and other data used in the analysis are listed in Table 1.

The determinant Δ of the matrix **B** of Eq. (31) was computed both numerically and using the analytical expression of Eq. (35), for a range of frequencies. The two sets of results were found to agree up to machine accuracy. The first four natural frequencies of the cantilever beam obtained from $\Delta=0$, with and without the inclusion of the effects of shear deformation and rotatory inertia, are given in Table 2 for three different loading cases, namely when the axial load is zero ($P=0$), tensile ($P=1790$) and compressive ($P=-1790$), respectively. Note that for $P=\pm1790$ N, the value of the nondimensional parameter p^2 (see Eq. (13)) is close to ±0.1886. The frequencies given in Table 2 agreed completely with published results, i.e., with those obtained using the exact dynamic stiffness theories of Friberg [8] and Banerjee and Williams [13]. The plot of Δ against frequency (ω) which identifies the first two natural frequencies of the beam is illustrated in Figs. 2(*a*), 2(*b*) and 2(*c*), respectively, for three representative cases when (*a*) $p^2=r^2=s^2=0$, (*b*) $p^2=0.1886$, $r^2=0.000447$, $s^2=0.00233$ and (*c*) $p^2=-0.1886$, $r^2=0.000447$, $s^2=0.00233$. The corresponding mode shapes were computed by using the analytical expressions of Eqs. (50)–(52). These were checked to machine accuracy by solving the system of Eqs. in (42) numerically, i.e., by using the computational steps of matrix inversion and multiplication. Further checks on the mode shapes were also performed by applying the exact dynamic stiffness theories of Friberg [8] and Banerjee and Williams [13] and again complete agreement was found in all cases. For illustrative purposes the first four modes of the beam are shown in Fig. 3 for the case when $p^2=0.1886$, $r^2=0.000447$, and $s^2=0.00233$. In order to be consistent with the unit used for the bending displacement (H) and also to make the results more meaningful, the torsional rotation (Ψ) was multiplied by x_α (i.e., the distance between the mass center and shear center) when plotting the modes. (Note that $x_\alpha\Psi$ represents the vertical displacement of the mass center relative to the shear center as a result of the twisting action.) Figure 3 shows that there is substantial coupling between the bending displacement and torsional rotation in all the four modes.

In order to demonstrate the substantial computational advantage of the proposed method, the determinant Δ was computed both numerically and analytically (when $p^2\neq0$, $s^2\neq0$, and $r^2\neq0$) for a

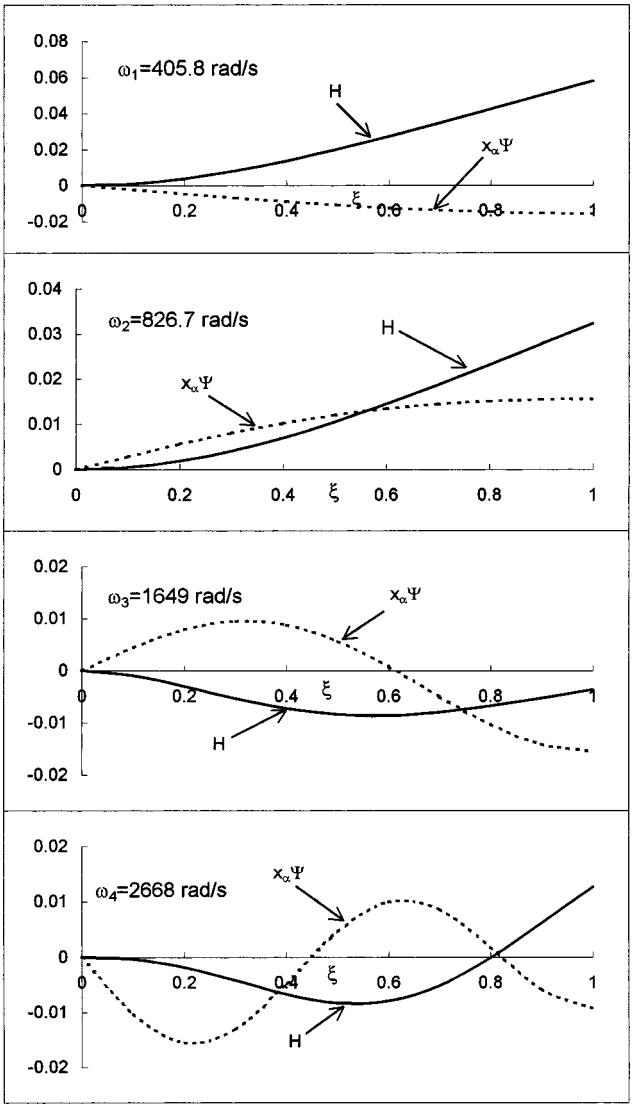


Fig. 3 The first four natural frequencies and mode shapes of an axially loaded bending-torsion coupled Timoshenko beam ([8,13]) with cantilever end condition for the case when $p^2=0.1886$, $r^2=0.000447$, and $s^2=0.00233$

Table 3 C.P.U. time on a SUN (Ultra-2) computer using Fortran

Number of Iterations (Number of Frequencies)	C.P.U. Time (s) Numerical Method	Explicit Expression
500	0.048	0.012
1000	0.092	0.023
2500	0.234	0.058
5000	0.474	0.116

large number of iterations, each performed at a different frequency. The recorded elapsed c.p.u. time on a SUN (Ultra-2) work station is shown in Table 3. It is clearly evident that programming the explicit expression for Δ has more than a fourfold advantage over the numerical method.

4 Conclusions

Exact frequency equation and mode shape expressions for an axially loaded bending-torsion coupled Timoshenko beam with cantilever end condition have been derived using the symbolic computing package REDUCE. The correctness of the expressions has been checked by numerical results which agree completely with exact published results. The expressions developed can be used to solve bench mark problems as an aid in validating the finite element and other approximate methods. Applications of the theory include further developments in aeroelastic and optimization studies. Programming the explicit expressions has a substantial advantage in c.p.u. time over numerical methods, and a typical gain of fourfold computational efficiency is demonstrated.

References

[1] Timoshenko, S. P., Young, D. H., and Weaver, W., 1974, *Vibration Problems in Engineering*, 4th Ed., Wiley, New York.

[2] Thomson, W. T., 1983, *Theory of Vibration and Application*, 2nd Ed., George Allen and Unwin, London.

[3] Tse, F. S., Morse, I. E., and Hinkle, R. T., 1978, *Mechanical Vibrations: Theory and Applications*, 2nd Ed., Allyn and Bacon Inc., London.

[4] Horr, A. M., and Schmidt, L. C., 1995, ''Closed-Form Solution for the Timoshenko Beam Theory using a Computer-Based Mathematical Package,'' Comput. Struct., **55**, No. 3, pp. 405–412.

[5] White, M. W. D., and Heppler, G. R., 1995, ''Vibration Modes and Frequencies of Timoshenko Beams With Attached Rigid Bodies,'' ASME J. Appl. Mech., **62**, No. 1, pp. 193–199.

[6] Farchaly, S. H., and Shebl, M. G., 1995, ''Exact Frequency and Mode Shape Formulae, for Studying Vibration and Stability of Timoshenko Beam System,'' J. Sound Vib., **180**, No. 2, pp. 205–227.

[7] Friberg, P. O., 1983, ''Coupled Vibration of Beams—An Exact Dynamic Element Stiffness Matrix,'' Int. J. Numer. Methods Eng., **19**, No. 4, pp. 479–493.

[8] Friberg, P. O., 1985, ''Beam Element Matrices Derived from Vlasov's Theory of Open Thin-Walled Elastic Beams,'' Int. J. Numer. Methods Eng., **21**, No. 7, pp. 1205–1228.

[9] Dokumaci, E., 1987, ''An exact Solution for Coupled Bending and Torsion Vibrations of Uniform Beams Having Single Cross-Sectional Symmetry,'' J. Sound Vib., **119**, No. 3, pp. 443–449.

[10] Banerjee, J. R., 1989, ''Coupled Bending-Torsional Dynamic Stiffness Matrix for Beam Elements,'' Int. J. Numer. Methods Eng., **28**, No. 6, pp. 1283–1298.

[11] Banerjee, J. R., and Fisher, S., 1992, ''Coupled Bending-Torsional Dynamic Stiffness Matrix for Axially Loaded Beam Elements,'' Int. J. Numer. Methods Eng., **33**, No. 6, pp. 739–751.

[12] Banerjee, J. R., and Williams, F. W., 1992, ''Coupled Bending-Torsional Dynamic Stiffness Matrix for Timoshenko Beam Elements,'' Comput. Struct., **42**, No. 3, pp. 301–310.

[13] Banerjee, J. R., and Williams, F. W., 1994, ''Coupled Bending-Torsional Dynamic Stiffness Matrix of an Axially Loaded Timoshenko Beam Element,'' Int. J. Solids Struct., **31**, No. 6, pp. 749–762.

[14] Bercin, A. N., and Tanaka, M., 1997, ''Coupled Flexural-Torsional Vibrations of Timoshenko Beams,'' J. Sound Vib., **207**, No. 1, pp. 47–59.

[15] Rayna, G., 1986, *REDUCE Software for Algebraic Computation*, Springer-Verlag, New York.

K. M. Liew[1]
Centre for Advanced Numerical Engineering Simulations, School of Mechanical and Production Engineering, Nanyang Technological University, Singapore 639798
Mem. ASME

Z. C. Feng[2]
Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139
Mem. ASME

Vibration Characteristics of Conical Shell Panels With Three-Dimensional Flexibility

A first known investigation on the three-dimensional vibration characteristics of conical shell panels is reported. A linear frequency equation is derived based on an exact three-dimensional, small-strain, linearly elastic theory. Sets of one and two-dimensional polynomial series are employed to approximate the spatial displacements of the conical shell panels in three dimension. The perturbation of frequency responses due to the variations of relative thickness L/h, slanted length L/S, vertex angle γ_v, and subtended angle γ_o is investigated. First known frequency parameters and three-dimensional deformed mode shapes of the conical shell panels are presented in vivid graphical forms. The new results may serve as benchmark references for validating the new refined shell theories and new computational techniques. [S0021-8936(00)02302-3]

Introduction

Conical shell panels are customarily used structural components in aerospace, nuclear, mechanical, and marine engineering applications. A proper understanding of the vibration characteristics of these shell panels is important before optimum design procedures can be planned. Despite the practical importance of these structural elements, researches carried out on this topic are relatively scarce. This subject is discussed at length in a monograph by Leissa [1] and Librescu [2] and follows by a review article by Liew et al. [3].

Since Love [4] introduced his shell theory in the last century, researchers have not exhausted searching for more reliable two-dimensional theories through application of reduction methods applied to three-dimensional elasticity ([5,6]). Although these two-dimensional theories are able to provide solutions for most shell problems, they may not be able to bring out the physical characteristics of shells. Models based on three-dimensional theory extract a full vibration spectrum without missing modes which are due to surface parallel vibratory motions. Three-dimensional elasticity solutions thus provide a real basis for assessing the solutions of two-dimensional theories. Some of the recent literatures for three-dimensional vibration analysis of shells are due to Hutchinson [7,8] Leissa and So [9,10], Liew and Hung [11], Liew et al. [12,13], and So and Leissa [14].

The above studies examined only the closed shell problems. This paper is a complement to the earlier works by providing a comprehensive study on the free vibration characteristics of conical shell panels (open shells) from three-dimensional analysis. The solution to this problem is made possible by using the Ritz method with sets of two-dimensional surface functions and one-dimensional thickness functions assumed in orthogonal polynomials. In this study, vibration behaviors of conical shell panels due to the variations of relative thickness L/h, slanted length L/S, vertex angle γ_v, and subtended angle γ_o are investigated.

Problem Definition

Figure 1 shows the geometric configuration of a homogeneous, isotropic, truncated conical shell panel with slanted length L, thickness h, vertex angle γ_v, and subtended angle γ_o. S denotes the slanted length of the cone from the vertex to the base. The displacement components of the conical shell panel are defined in a cylindrical coordinate system (r,θ,z). For the conical shell, there is no curvature along the spanwise direction. In the chordwise direction, however, the radius at the midsurface $R_m(z)$ is a function varying with respect to the z-direction. The displacement components at a generic point are u_1, u_2, and u_3 in the radial, circumferential, and vertical directions with respect to the polar cylindrical reference frame. In this study, the conical shell panels treated are confined to cantilevered (CFFF) and fully clamped (CCCC) boundary conditions. For the CFFF shell panel, the clamped edge is located at $z=0$ which corresponds to the base of the cone. The vibration characteristics of the CFFF and CCCC shell panels are to be investigated.

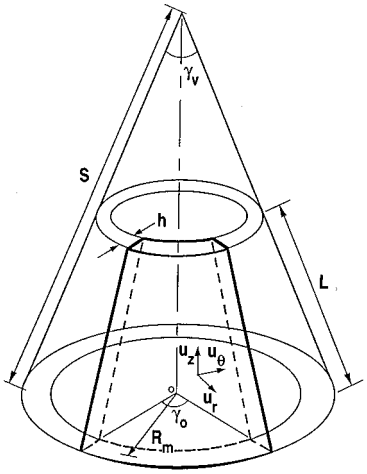


Fig. 1 Geometry and dimensions of a truncated conical shell panel

[1]This work was carried out during the author's sabbatical at the Department of Mechanical Engineering, Massachusetts Institute of Technology.

[2]Present address: Department of Mechanical and Aerospace Engineering, University of Missouri—Columbia, Columbia, MO 65211.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, April 4, 1998; final revision, October 21, 1999. Associate Technical Editor: A. Ferri. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Mathematical Formulations

The linear elastic strain energy component $\hat{V}$ for a shell panel in cylindrical coordinates can be written in an integral form as

$$\hat{V}=\hat{E}\int_{\text{Vol}}[(1-\nu)\hat{A}_1+2\nu\hat{A}_2+(1-2\nu)\hat{A}_3]r\,dr\,d\theta\,dz \quad (1)$$

where

$$\hat{E}=\frac{E}{2(1+\nu)(1-2\nu)} \quad (2)$$

$$\hat{A}_1=\varepsilon_{rr}^2+\varepsilon_{\theta\theta}^2+\varepsilon_{zz}^2 \quad (3)$$

$$\hat{A}_2=\varepsilon_{rr}\varepsilon_{zz}+\varepsilon_{rr}\varepsilon_{\theta\theta}+\varepsilon_{\theta\theta}\varepsilon_{zz} \quad (4)$$

$$\hat{A}_3=\varepsilon_{r\theta}^2+\varepsilon_{rz}^2+\varepsilon_{\theta z}^2 \quad (5)$$

and Vol is the volume, E is Young's modulus, ν is Poisson's ratio, and the strain components in cylindrical polar coordinate for small deformation are given as

$$\varepsilon_{rr}=\frac{\partial u_r}{\partial r} \quad (6)$$

$$\varepsilon_{\theta\theta}=\frac{u_r}{r}+\frac{\partial u_\theta}{r\partial\theta} \quad (7)$$

$$\varepsilon_{zz}=\frac{\partial u_z}{\partial z} \quad (8)$$

$$\varepsilon_{r\theta}=\frac{\partial u_r}{r\partial\theta}+\frac{\partial u_\theta}{\partial r}-\frac{u_\theta}{r} \quad (9)$$

$$\varepsilon_{rz}=\frac{\partial u_r}{\partial z}+\frac{\partial u_z}{\partial r} \quad (10)$$

$$\varepsilon_{\theta z}=\frac{\partial u_\theta}{\partial z}+\frac{\partial u_z}{r\partial\theta}. \quad (11)$$

For free vibration analysis, the kinetic energy $\hat{T}$ can be expressed as

$$\hat{T}=\hat{\rho}\int_{\text{Vol}}[\dot{u}_r^2+\dot{u}_\theta^2+\dot{u}_z^2]r\,dr\,d\theta\,dz \quad (12)$$

$$\dot{u}_r=\frac{\partial u_r}{\partial t};\quad \dot{u}_\theta=\frac{\partial u_\theta}{\partial t};\quad \dot{u}_z=\frac{\partial u_z}{\partial t};\quad \hat{\rho}=\frac{\rho}{2} \quad (13)$$

where ρ is the mass density per unit volume, and t is time.

For linear, small-strain, simple harmonic motion, the displacement components assume the following forms:

$$u_\alpha(r,\theta,z,t)=U_\alpha(r,\theta,z)e^{i\omega t};\quad \alpha=1,2,3 \quad (14)$$

where ω denotes the frequency of vibration.

For generality and simplicity in the subsequent derivations, the cylindrical coordinates (r,θ,z) are transformed into a set of nondimensional parameters $(\overline{x_1},\overline{x_2},\overline{x_3})$ by the following relations:

$$r=\frac{1}{2}\{[r_o(\overline{x_3})+r_i(\overline{x_3})]-[r_o(\overline{x_3})-r_i(\overline{x_3})]\overline{x_1}\} \quad (15)$$

$$\theta=\frac{1}{2}\gamma_o\overline{x_2} \quad (16)$$

$$z=\overline{x_3}L\cos\hat{\gamma}_v \quad (17)$$

in which $\hat{\gamma}_v=\gamma_v/2$.

The conical shell panel may be considered of consisting of multiple layers of sectorial slices of infinitesimal thickness. In this study, the volume integrals in the strain and kinetic energies are therefore treated as a summation of the energy contribution of each sectorial slice. At an arbitrary distance $\overline{x_3}$, the midsurface radius of curvature $R_m(\overline{x_3})$ of a sectorial slice can be approximated by

$$R_m(\overline{x_3})=L\sin\hat{\gamma}_v(\hat{S}-\overline{x_3}) \quad (18)$$

in which $\hat{S}=S/L$. The corresponding inner and outer radii can be written as

$$r_i(\overline{x_3})=R_m(\overline{x_3})-\hat{h} \quad (19)$$

$$r_o(\overline{x_3})=R_m(\overline{x_3})+\hat{h} \quad (20)$$

in which $\hat{h}=h/2$.

The nondimensional displacement trial functions are assumed as

$$U_\alpha(\overline{x_1},\overline{x_2},\overline{x_3})=\sum_{m=1}^{M}\sum_{n=1}^{N}C_{mn}^{\alpha}\phi_m^{\alpha}(\overline{x_1},\overline{x_2})\psi_n^{\alpha}(\overline{x_3}) \quad (21)$$

in which $\langle\alpha=1,2,3\rangle$, C_{mn}^{α} are the unknown coefficients, and $\phi_m^{\alpha}(\overline{x_1},\overline{x_2})$ and $\Psi_n^{\alpha}(\overline{x_3})$ are the one and two-dimensional orthogonal polynomial functions.

Let $\hat{\Pi}$ be the energy functional given by

$$\hat{\Pi}=\tilde{V}-\tilde{T} \quad (22)$$

where $\tilde{V}$ and $\tilde{T}$ are the maximum strain and kinetic energies which can be derived by substituting Eq. (21) into the respective energy expressions in Eqs. (1) and (12) with the periodic component eliminated.

The minimization of $\hat{\Pi}$ with respect to the coefficients

$$\frac{\partial\hat{\Pi}}{\partial C_{mn}^{\alpha}}=0;\quad \alpha=1,2,3 \quad (23)$$

leads to the governing eigenvalue equation of the form

$$(\hat{\mathbf{K}}-\hat{\lambda}^2\hat{\mathbf{M}})\hat{\mathbf{C}}=\mathbf{0} \quad (24)$$

where

$$\hat{\mathbf{K}}=\begin{bmatrix}\hat{\mathbf{k}}^{11} & \hat{\mathbf{k}}^{12} & \hat{\mathbf{k}}^{12}\\ & \hat{\mathbf{k}}^{22} & \hat{\mathbf{k}}^{23}\\ \text{Sym} & & \hat{\mathbf{k}}^{33}\end{bmatrix} \quad (25)$$

$$\hat{\mathbf{M}}=\begin{bmatrix}\hat{\mathbf{m}}^{11} & 0 & 0\\ & \hat{\mathbf{m}}^{22} & 0\\ \text{Sym} & & \hat{\mathbf{m}}^{33}\end{bmatrix} \quad (26)$$

and

$$\hat{\mathbf{C}}=\{\mathbf{C}^1\quad\mathbf{C}^2\quad\mathbf{C}^3\}^T. \quad (27)$$

The explicit forms of the respective elements in the stiffness submatrices $\hat{\mathbf{k}}^{\alpha\beta}$ are given by

$$\hat{k}_{mjnk}^{11}=\frac{(1-\nu)}{\tilde{\Lambda}_1}\left[\frac{4}{h^2}(\hat{\mathbf{I}}_{mjnk}^{101000;1})_{11}+(\hat{\mathbf{I}}_{mjnk}^{000000;-1})_{11}\right]+\frac{2\nu}{\tilde{\Lambda}_1 h}[(\hat{\mathbf{I}}_{mjnk}^{001000;0})_{11}+(\hat{\mathbf{I}}_{mjnk}^{100000;0})_{11}]+\frac{4}{\tilde{\Lambda}_2}\left[\frac{1}{h^2}(\hat{\mathbf{I}}_{mjnk}^{010100;1})_{11}+\frac{1}{\gamma_o^2}(\hat{\mathbf{I}}_{mjnk}^{010100;-1})_{11}\right] \quad (28)$$

$$\hat{k}^{12}_{mjnk}=\frac{2(1-\nu)}{\tilde{\Lambda}_1\gamma_o}(\hat{\mathbf{I}}^{000100;-1}_{mjnk})_{12}+\frac{4\nu}{\tilde{\Lambda}_1\gamma_o h}(\hat{\mathbf{I}}^{100100;0}_{mjnk})_{12}+\frac{2}{\tilde{\Lambda}_2\gamma_o}\left[\frac{2}{h}(\hat{\mathbf{I}}^{010010;0}_{mjnk})_{12}-(\hat{\mathbf{I}}^{010000;-1}_{mjnk})_{12}\right] \quad (29)$$

$$\hat{k}^{13}_{mjnk}=\frac{\nu}{\tilde{\Lambda}_1\hat{b}}\left[\frac{2}{h}(\hat{\mathbf{I}}^{100001;1}_{mjnk})_{13}+(\hat{\mathbf{I}}^{000001;0}_{mjnk})_{13}\right]+\frac{2}{\tilde{\Lambda}_2\hat{b}h}(\hat{\mathbf{I}}^{001010;1}_{mjnk})_{13} \quad (30)$$

$$\hat{k}^{22}_{mjnk}=\frac{4(1-\nu)}{\tilde{\Lambda}_1\gamma_o^2}(\hat{\mathbf{I}}^{010100;-1}_{mjnk})_{22}+\frac{1}{\tilde{\Lambda}_2}\left\{\frac{4}{h^2}[(\hat{\mathbf{I}}^{100010;1}_{mjnk})_{22}+(\hat{\mathbf{I}}^{001001;1}_{mjnk})_{22}]\right\}+\frac{2}{h}[(\hat{\mathbf{I}}^{000010;0}_{mjnk})_{22}+(\hat{\mathbf{I}}^{100000;0}_{mjnk})_{22}] \quad (31)$$

$$\hat{k}^{23}_{mjnk}=\frac{2\nu}{\tilde{\Lambda}_1\gamma_o\hat{b}}(\hat{\mathbf{I}}^{010001;0}_{mjnk})_{23}+\frac{2}{\tilde{\Lambda}_2\gamma_o\hat{b}}(\hat{\mathbf{I}}^{001100;0}_{mjnk})_{23} \quad (32)$$

$$\hat{k}^{33}_{mjnk}=\frac{(1-\nu)}{\tilde{\Lambda}_1\hat{b}^2}(\hat{\mathbf{I}}^{001001;1}_{mjnk})_{33}+\frac{4}{\tilde{\Lambda}_2}\left[\frac{1}{h^2}(\hat{\mathbf{I}}^{100010;1}_{mjnk})_{33}+\frac{1}{\gamma_o^2}(\hat{\mathbf{I}}^{010100;-1}_{mjnk})_{33}\right] \quad (33)$$

and the elements in the mass submatrix $\hat{\mathbf{m}}^{\alpha\beta}$ are given by

$$\hat{m}^{11}_{mjnk}=\int_{-1}^{1}\int_{-1}^{1}\int_{0}^{1}[\hat{\vartheta}(\overline{x}_1,\overline{x}_3)\hat{\Phi}^1(\overline{x}_1,\overline{x}_2,\overline{x}_3)\times\hat{\Psi}^1(\overline{x}_1,\overline{x}_2,\overline{x}_3)]d\overline{x}_1d\overline{x}_2d\overline{x}_3 \quad (34)$$

$$\hat{m}^{22}_{mjnk}=\int_{-1}^{1}\int_{-1}^{1}\int_{0}^{1}[\hat{\vartheta}(\overline{x}_1,\overline{x}_3)\hat{\Phi}^2(\overline{x}_1,\overline{x}_2,\overline{x}_3)\times\hat{\Psi}^2(\overline{x}_1,\overline{x}_2,\overline{x}_3)]d\overline{x}_1d\overline{x}_2d\overline{x}_3 \quad (35)$$

$$\hat{m}^{33}_{mjnk}=\int_{-1}^{1}\int_{-1}^{1}\int_{0}^{1}[\hat{\vartheta}(\overline{x}_1,\overline{x}_3)\hat{\Phi}^3(\overline{x}_1,\overline{x}_2,\overline{x}_3)\times\hat{\Psi}^3(\overline{x}_1,\overline{x}_2,\overline{x}_3)]d\overline{x}_1d\overline{x}_2d\overline{x}_3 \quad (36)$$

where

$$\tilde{\Lambda}_1=\frac{1}{2}(1-2\nu)\tilde{\Lambda}_2 \quad (37)$$

$$\tilde{\Lambda}_2=2(1+\nu) \quad (38)$$

$$\hat{\vartheta}(\overline{x}_1,\overline{x}_3)=\frac{1}{2}\{(r_o(\overline{x}_3)+r_i(\overline{x}_3))-(r_o(\overline{x}_3)-r_i(\overline{x}_3))\overline{x}_1\} \quad (39)$$

$$\hat{b}=L\cos\hat{\gamma}_v \quad (40)$$

and

$$(\hat{\mathbf{I}}^{defgrs;Z}_{mj})_{\alpha\beta}=\int_{-1}^{1}\int_{-1}^{1}\int_{0}^{1}\hat{\Phi}(\overline{x}_1,\overline{x}_2,\overline{x}_3)\hat{\Psi}(\overline{x}_1,\overline{x}_2,\overline{x}_3)\times\hat{\vartheta}(\overline{x}_1,\overline{x}_3)r_o(\overline{x}_3)d\overline{x}_1d\overline{x}_2d\overline{x}_3 \quad (41)$$

$$\hat{\Phi}(\overline{x}_1,\overline{x}_2,\overline{x}_3)=\frac{\partial^{d+e+r}\phi^\alpha_m(\overline{x}_1,\overline{x}_2)\psi^\alpha_j(\overline{x}_3)}{\partial\overline{x}_1^d\partial\overline{x}_2^e\partial\overline{x}_3^r} \quad (42)$$

$$\hat{\Psi}(\overline{x}_1,\overline{x}_2,\overline{x}_3)=\frac{\partial^{f+g+s}\phi^\alpha_n(\overline{x}_1,\overline{x}_2)\psi^\alpha_k(\overline{x}_3)}{\partial\overline{x}_1^f\partial\overline{x}_2^g\partial\overline{x}_3^s} \quad (43)$$

in which $\langle\alpha=1,2,3\rangle$ and $\langle\beta=1,2,3\rangle$.

The eigenvalues obtained from Eq. (24) are defined in terms of nondimensionalized frequency parameter λ of the form

$$\lambda=\omega L\sqrt{\frac{\rho}{E}}. \quad (44)$$

Admissible Displacement Functions

The admissible functions adopted in Eq. (21) are sets of one $\psi^\alpha_n(\overline{x}_3)$ and two-dimensional $\phi^\alpha_m(\overline{x}_1,\overline{x}_2)$ orthogonal polynomial functions. Details of these orthogonal polynomial functions can be found in Liew et al. [15,16]. To be able to use the procedures ([15,16]), we first need to derive the basic functions that are governed by the boundary conditions of the conical shell panels.

The two-dimensional functions $\phi^\alpha_m(\overline{x}_1,\overline{x}_2)$ approximate the displacement variations of the conical shell panel in the radial and circumferential directions. The general form of the functions may be written as

$$\phi^\alpha_1(\overline{x}_1,\overline{x}_2)=(\overline{x}_2)^{\hat{\Theta}^\alpha_1}(\overline{x}_2-1)^{\hat{\Theta}^\alpha_2} \quad (45)$$

and

$$\hat{\Theta}^\alpha_i=\begin{cases}0 & \text{if the } i\text{th edge is free } (F)\\ 1 & \text{if the } i\text{th edge is clamped } (C)\end{cases} \quad (46)$$

in which $\langle\alpha=1,2,3\rangle$.

The one-dimensional functions $\psi^\alpha_n(\overline{x}_3)$ approximate the variations of the conical shell panel in the spanwise direction. The general form of the basic functions may be defined as

$$\psi^\alpha_1(\overline{x}_3)=(\overline{x}_3-1)^{\hat{\Theta}^\alpha_1}(\overline{x}_3+1)^{\hat{\Theta}^\alpha_2} \quad (47)$$

and $\hat{\Theta}^\alpha_i$ follows the definition given in Eq. (46).

Results and Discussion

The above procedures are applied to extract the vibration frequencies and mode shapes of conical shell panels with CFFF and CCCC boundary conditions subject to variation of geometric parameters. For both cases, the vibration modes can be conveniently divided into the symmetry (S) and antisymmetry (A) classes with respect to the $\overline{x}_1\overline{x}_3$ plane at $\overline{x}_2=0$.

Table 1 shows the convergence of conical shell panel of different configurations by varying the number of terms used in the displacement fields. The number of terms N used in the one-dimensional polynomial functions, and the order of polynomial P of the two-dimensional functions have been increased steadily. The shell chosen has a vertex angle $\gamma_v=30$ deg and chordwise-subtended angle $\gamma_o=15$ deg. The slanted lengths L/S of the shell are fixed at 0.10 and the relative thickness L/h is fixed at 8.0, which corresponds to a moderately thick shell. In general, satisfactory converged solutions are obtained when $P=5$ and $N=8$ for both shells.

The present formulation is validated by direct comparison with the existing solutions. A clamped conical thin shell studied by Srinivasan and Krishnan [17] and Cheung et al. [18] from two different approaches is selected for this comparison purpose. Frequency parameters for the clamped conical shell with vertex angle $\gamma_v=60$ deg, chordwise subtended angle $\gamma_o=60$ deg, slanted length $L/S=0.60$, and thickness ratio $L/h=100$ are obtained from the present method. Table 2 shows the comparison of the authors' solutions with the classical thin shell solutions of Srinivasan and Krishnan [17], and Cheung et al. [18]. The frequency parameters computed are expressed in terms of the nondimensional form [18] $\tilde{\lambda}$ that is related to λ of Eq. (44) by

Table 1 Convergence of λ for conical shell panels with CFFF and CCCC boundary conditions (γ_v=30 deg, γ_o=15 deg, L/S=0.10 and L/h=8)

Boundary Condition	Terms (P,N)	Mode Sequences					
		S-1	S-2	S-3	A-1	A-2	A-3
CFFF	(4, 3)	0.14833	0.94792	2.0672	0.44400	0.57134	1.7945
	(5, 4)	0.14655	0.83845	1.6806	0.44185	0.56875	1.4018
	(6, 4)	0.14654	0.83843	1.6803	0.44184	0.56874	1.4018
	(5, 5)	0.14485	0.82264	1.6774	0.43887	0.56736	1.3867
	(5, 6)	0.14457	0.81392	1.6764	0.43802	0.56688	1.3776
	(5, 7)	0.14421	0.81303	1.6757	0.43715	0.56650	1.3763
	(5, 8)	0.14417	0.81298	1.6755	0.43712	0.56647	1.3760
CCCC	(4, 3)	2.0045	3.4669	4.5107	4.0300	5.0389	5.4092
	(5, 4)	1.9994	2.9070	4.4921	4.0193	4.7324	5.3974
	(6, 4)	1.9994	2.9069	4.4921	4.0192	4.7322	5.3973
	(5, 5)	1.9960	2.8868	4.2874	4.0168	4.7224	5.3949
	(5, 6)	1.9948	2.8755	4.2031	4.0159	4.7173	5.3944
	(5, 7)	1.9941	2.8736	4.1866	4.0154	4.7157	5.3939
	(5, 8)	1.9940	2.8734	4.1862	4.0153	4.7155	5.3939

$$\tilde{\lambda} = \frac{\lambda L}{h}\sqrt{12(1-\nu^2)}. \tag{48}$$

From Table 2, it is evident that the present predictions are in good agreement with their results. Both thin shell solutions of Cheung et al. [18] and Srinivasan and Krishnan [17] are formulated based on Donnell's shell theory. This comparison shows that the present formulation is capable of producing comparable frequency solutions to conical shell problems even with a highly degenerated thickness. Table 3 shows a comparison study of the present three-dimensional results with the finite element solutions obtained using the eight-node three-dimensional elements of MSC/NASTRAN commercial package. It is again shown that a good agreement has been achieved between the present results and the finite element solutions.

Table 2 Comparison of $\tilde{\lambda}$ for a CCCC shell panel (L/S=0.60, γ_v=60 deg, γ_o=60 deg, and L/h=100)

Source of Results	Mode Sequences			
	1 (A-1)	2 (S-1)	3 (A-2)	4 (S-2)
Srinivasan and Krishnan [17]	202.7	260.1	305.6	355.0
Cheung et al. [18]	213.4	262.5	314.7	358.6
Present Three-Dimensional Solutions	210.9	258.5	307.5	349.3

Frequency parameters λ for conical shell panels with cantilevered (CFFF) and fully clamped (CCCC) boundary conditions are presented in Tables 4–7. The effects of relative thickness L/h, slanted length L/S, vertex angle γ_v, and subtended angle γ_o upon the variations of λ are investigated. The frequency parameter λ is independent of shell thickness h; hence through λ, the physical vibration frequencies for different conical shell panels can be directly compared. The results for CFFF shell panels in Tables 4 and 5 possess subtended angle γ_o of 15 deg, 30 deg, and 45 deg, and vertex angle γ_v of 30 deg and 60 deg. The relative thickness

Table 3 Comparison of λ for a thick CFFF shell panel (L/S=0.10, γ_v=30 deg, γ_o=15 deg, and L/h=8)

Source of Results	Mode Sequences			
	1 (S-1)	2 (A-1)	3 (A-2)	4 (S-2)
Present Three-Dimensional Solutions	0.14417	0.43712	0.56647	0.81298
Three-Dimensional MSC/NASTRAN	0.14496	0.43809	0.56654	0.82073

Table 4 Frequency parameters λ for CFFF conical shell panels with γ_v=30 deg

γ_o	$\frac{L}{S}$	$\frac{L}{h}$	Mode Sequences					
			S-1	S-2	S-3	A-1	A-2	A-3
15 deg	0.10	8	0.14417	0.81298	1.6755	0.43712	0.56647	1.3760
		16	0.075791	0.44332	1.0808	0.23528	0.56600	0.76893
		32	0.043111	0.24420	0.56295	0.12125	0.40160	0.56586
	0.20	8	0.14604	0.81251	1.7120	0.33080	0.73009	1.4755
		16	0.074748	0.43694	1.1717	0.33124	0.43269	1.2877
		32	0.038971	0.22974	0.62587	0.23001	0.33183	0.69303
30 deg	0.10	8	0.16369	0.66258	0.85706	0.27641	0.78452	1.0027
		16	0.10755	0.35358	0.49651	0.14579	0.55279	0.78094
		32	0.084792	0.18665	0.31022	0.077593	0.29958	0.40528
	0.20	8	0.15287	0.83321	1.7148	0.46869	0.56462	1.4409
		16	0.086977	0.47869	1.2003	0.25422	0.56388	0.81076
		32	0.058916	0.30194	0.65110	0.13211	0.42789	0.56373
45 deg	0.10	8	0.19938	0.41197	0.86616	0.22264	0.74607	0.86991
		16	0.14608	0.23538	0.51406	0.12389	0.39808	0.52468
		32	0.093309	0.17571	0.30617	0.077493	0.20871	0.30483
	0.20	8	0.17379	0.86662	1.1736	0.35847	0.70568	1.1707
		16	0.12075	0.53669	0.64588	0.19170	0.65500	0.70555
		32	0.10140	0.31472	0.40459	0.10245	0.36038	0.70454

Table 5 Frequency parameters λ for CFFF conical shell panels with $\gamma_v = 60$ deg

γ_o	$\frac{L}{S}$	$\frac{L}{h}$	Mode Sequences					
			S-1	S-2	S-3	A-1	A-2	A-3
15 deg	0.10	8	0.18458	0.74001	1.0167	0.32454	0.89563	1.1858
		16	0.10209	0.39462	0.56021	0.17024	0.65458	0.85850
		32	0.065172	0.20321	0.30628	0.087596	0.34213	0.44235
	0.20	8	0.18481	1.0023	1.9133	0.54611	0.65935	1.6751
		16	0.096831	0.55151	1.3446	0.29734	0.65870	0.95271
		32	0.054134	0.30084	0.70554	0.15382	0.50084	0.65852
30 deg	0.10	8	0.20873	0.36020	0.91205	0.23255	0.56694	1.0229
		16	0.13457	0.20147	0.49271	0.12705	0.30087	0.58651
		32	0.083869	0.14231	0.25432	0.076394	0.15873	0.31792
	0.20	8	0.20552	0.82890	1.0522	0.35093	0.90528	1.2322
		16	0.13130	0.44633	0.60897	0.18608	0.68779	0.90484
		32	0.10130	0.23491	0.37433	0.098633	0.37115	0.51056
45 deg	0.10	8	0.22211	0.28689	0.53644	0.21779	0.37232	0.77255
		16	0.13262	0.19748	0.28690	0.12929	0.20582	0.41418
		32	0.079596	0.15372	0.17280	0.081014	0.13315	0.21951
	0.20	8	0.24410	0.51860	1.0591	0.28292	0.93294	1.0012
		16	0.17713	0.29238	0.62794	0.15635	0.50325	0.64845
		32	0.11729	0.20708	0.37573	0.095525	0.26344	0.37246

Table 6 Frequency parameters λ for CCCC conical shell panels with $\gamma_v = 30$ deg

γ_o	$\frac{L}{S}$	$\frac{L}{h}$	Mode Sequences					
			S-1	S-2	S-3	A-1	A-2	A-3
15 deg	0.10	8	1.9940	2.8734	4.1862	4.0153	4.7155	5.3939
		16	1.2094	1.7789	2.7015	2.6280	3.1349	3.9365
		32	0.70598	1.0032	1.5195	1.4600	1.7681	2.2601
	0.20	8	5.1331	5.7240	6.5523	9.8164	10.516	10.648
		16	3.7132	4.1808	4.7901	7.9391	8.6151	9.2405
		32	2.2615	2.5642	2.9457	5.2011	5.7173	6.1903
30 deg	0.10	8	1.0933	2.2181	2.5648	1.6526	2.6927	3.1988
		16	0.67133	1.3399	4.5272	0.95293	1.6328	2.2753
		32	0.46059	0.77532	0.82508	0.52668	0.90587	1.2356
	0.20	8	2.2234	3.0562	4.3206	4.2298	4.9800	5.6786
		16	1.4512	1.9743	2.8469	2.8194	3.3825	4.1763
		32	0.99124	1.2419	1.7008	1.5894	1.9464	2.4461
45 deg	0.10	8	0.96696	1.5655	2.1212	1.1610	2.1180	2.3122
		16	0.61017	0.89692	1.2819	0.67642	1.2274	1.3921
		32	0.42665	0.51575	0.75052	0.41370	0.65895	0.79031
	0.20	8	1.5211	2.5049	3.8964	2.4829	3.4046	4.0836
		16	1.0517	1.6133	2.5400	1.5164	2.1602	3.0837
		32	0.82370	1.0745	1.4638	0.85521	1.2543	1.7945

Table 7 Frequency parameters λ for CCCC conical shell panels with $\gamma_v = 60$ deg

γ_o	$\frac{L}{S}$	$\frac{L}{h}$	Mode Sequences					
			S-1	S-2	S-3	A-1	A-2	A-3
15 deg	0.10	8	1.2225	2.5873	2.7774	1.8342	3.0758	3.4132
		16	0.69804	1.5684	1.6673	1.0572	1.8902	2.4639
		32	0.39506	0.85194	0.89125	0.56114	1.0256	1.3417
	0.20	8	2.3161	3.3666	4.8886	4.5450	5.4218	5.8944
		16	1.4281	2.1343	3.2464	3.0497	3.7090	4.6666
		32	0.82853	1.2087	1.8500	1.7195	2.1289	2.7336
30 deg	0.10	8	1.0527	1.3802	2.2953	1.1641	1.7004	2.5650
		16	0.60394	0.77995	1.3673	0.65967	0.96685	1.5526
		32	0.35374	0.41956	0.73156	0.36683	0.51133	0.84224
	0.20	8	1.2999	2.6317	2.9703	1.9434	3.1710	3.5285
		16	0.79576	1.6226	1.8069	1.1394	1.9690	2.6627
		32	0.53035	0.93284	0.98255	0.62874	1.0973	1.4638
45 deg	0.10	8	1.2225	2.5873	2.7774	1.8342	3.0758	3.4133
		16	0.69803	1.5684	1.6673	1.0572	1.8902	2.4639
		32	0.39505	0.85196	0.89124	0.56115	1.0256	1.3417
	0.20	8	2.3161	3.3666	4.8886	4.5450	5.4219	5.8944
		16	1.4281	2.1343	3.2464	3.0497	3.7090	4.6666
		32	0.82852	1.2088	1.8500	1.7195	2.1290	2.7336

L/h of shell varies from 8, 16, and 32 (which corresponds to thick, moderately thick, and thin shells). The slanted lengths of shell L/S are fixed at 0.10 and 0.20 in the present study. The frequency parameters for the fully clamped shell panels are given in Tables 6 and 7.

For the CFFF shell panels, it is noted that the fundamental mode of vibration is dominated by the first symmetric mode (S-1) followed by the first antisymmetric mode (A-1). It is observed that when the shell with a constant vertex angle γ_v and base subtended angle γ_o, λ decreases as thickness ratio L/h increases. Since the

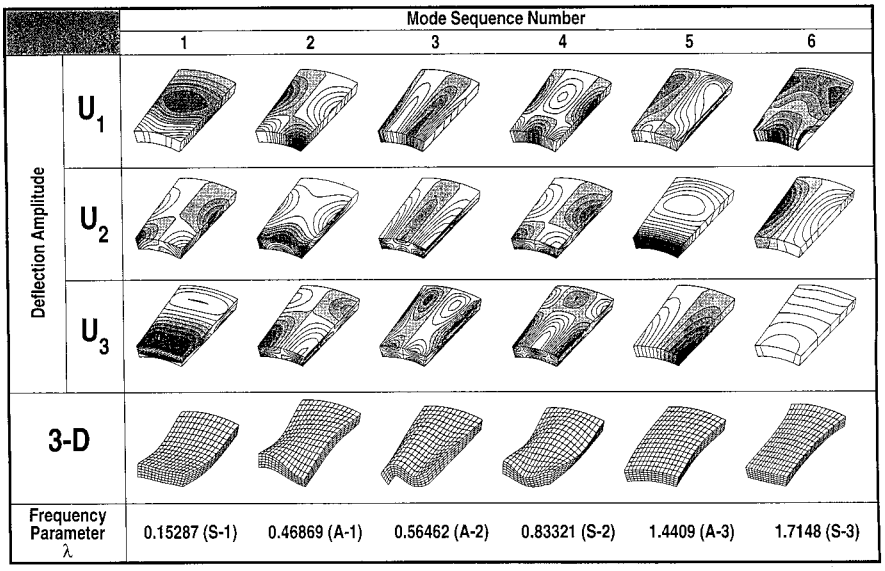


Fig. 2 Mode shapes and frequency parameters of a cantilevered conical shell panel (γ_v=30 deg, γ_o=30 deg, L/S=0.20 and L/h=8)

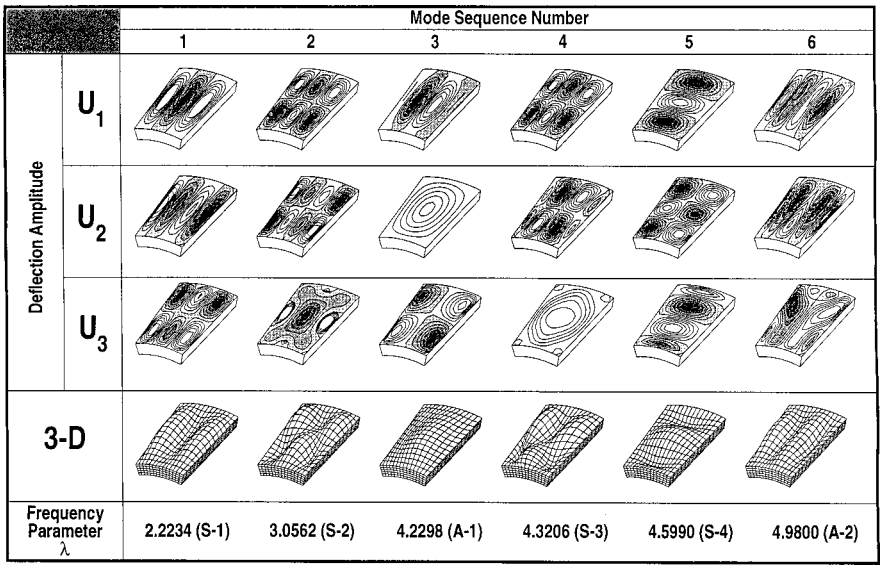


Fig. 3 Mode shapes and frequency parameters of a fully clamped conical shell panel (γ_v=30 deg, γ_o=30 deg, L/S=0.20 and L/h=8)

L/h ratio is inversely proportional to the physical thickness of shell, a higher L/h value corresponds to a thinner shell. As thinner shells possess smaller bending and torsional stiffnesses, the predicted frequency parameters inevitably reduce proportionally. The effects of vertex angle γ_v upon λ of the shell panels can be observed directly by comparing Table 4 to Table 5. It is noticed that for a shell panel of the same subtended angle γ_o, slanted length L/S and thickness ratio L/h, λ increases with γ_v of the truncated cone. The variation of λ with respect to the slanted length L/S for the fully clamped shell panel is also examined (Tables 6 and 7). For this shell panel with γ_o=30 deg and 60 deg, it is evident that λ increases as L/S increases. It is also noted that λ increases as the thickness L/h increases.

The displacement contour plots and three-dimensional deformed geometries of the vibration mode shapes for both cantilevered and fully clamped conical shell panels are depicted in Figs. 2 and 3. The conical shells considered have vertex angle γ_v =30 deg, subtended angle γ_o=30 deg, slanted length L/S =0.20, and relative thickness ratio L/h=8. For the CFFF shell panels, the first symmetric mode (S-1) shows prominent spanwise bending motion. Torsional motion is found at the first antisymmetric mode (A-1). The in-plane dominant vibration motions occur at the fifth and sixth modes. For the CCCC shell panels illustrated in Fig. 3, most of the modes presented exhibit out-of-plane transverse bending motions. The only in-plane vibration mode for this case occurs at the third mode (λ=4.2298). This in-plane mode is dominated by the in-plane shearing motion in the circumferential direction of the shell.

Concluding Remarks

The first known three-dimensional elasticity solutions for free vibrations of conical shell panels with cantilevered and clamped boundary conditions were presented. A systematic formulation of the integral expressions for strain and kinetic energies in a cylin-

drical polar coordinate system was detailed. The three-dimensional elasticity theory was adopted in this derivation that allows the extracting of full vibration spectrum for the conical shell panels. The solutions to the problem were obtained by the Ritz method with the use of one and two-dimensional orthogonal polynomial functions for the displacement fields. Convergence and comparison studies were carried out to validate the accuracy of the present formulation. Detailed parametric studies showing the influence of thickness ratio, vertex angle, subtended angle, and slanted length on the vibration responses were presented. It was noted that the frequency parameters of transverse bending modes of conical shells, whether cantilevered or clamped, always increase as the thickness increases. For the vibration modes which are dominated by the in-plane motion, the changes in shell thickness are found to have minor influence on the vibration frequencies.

References

[1] Leissa, A. W., 1973, ''Vibration of Shells,'' NASA SP-288, U.S. Government Printing Office, Washington DC, reprinted by the Acoustical Society of America, 1993.

[2] Librescu, L., 1975, *Elastostatics and Kinematics of Anisotropic and Hetrogenerus Shell-Types Structures*, Noordhoff, Leyden.

[3] Liew, K. M., Lim, C. W., and Kitipornchai, S., 1997, ''Vibration of Shallow Shells: A Review With Bibliography,'' ASME Appl. Mech. Rev., **50**, pp. 431–444.

[4] Love, A. E. H., 1888, ''Small Free Vibrations and Deformations of Thin Elastic Shells,'' Philos. Trans. R. Soc. London, Ser. A, **179**, pp. 491–549.

[5] Reddy, J. N., and Liu, C. F., 1985, ''A Higher-Order Shear Deformation Theory of Laminated Elastic Shells,'' Int. J. Eng. Sci., **23**, pp. 319–330.

[6] Liew, K. M., and Lim, C. W., 1996, ''A Higher-Order Theory for Vibration of Doubly Curved Shallow Shells,'' ASME J. Appl. Mech., **63**, pp. 587–593.

[7] Hutchinson, J. R., 1971, ''Axisymmetric Vibration of a Free Finite-Length Rod,'' J. Acoust. Soc. Am., **51**, pp. 233–240.

[8] Hutchinson, J. R., 1980, ''Vibrations of Solid Cylinders,'' ASME J. Appl. Mech., **47**, pp. 901–907.

[9] Leissa, A. W., and So, J., 1995a, ''Accurate Vibration Frequencies of Circular Cylinders From Three-Dimensional Analysis,'' J. Acoust. Soc. Am., **98**, pp. 2136–2141.

[10] Leissa, A. W., and So, J., 1995b, ''Three-Dimensional Vibrations of Truncated Hollow Cones,'' J. Vib. Cont., **1**, pp. 145–158.

[11] Liew, K. M., and Hung, K. C., 1995, ''Three-Dimensional Vibratory Characteristics of Solid Cylinders and Some Remarks on Simplified Beam Theories,'' Int. J. Solids Struct., **32**, pp. 3499–3513.

[12] Liew, K. M., Hung, K. C., and Lim, M. K., 1995, ''Free Vibration Studies on Stress-Free Three-Dimensional Elastic Solids,'' ASME J. Appl. Mech., **62**, pp. 159–165.

[13] Liew, K. M., Hung, K. C., and Lim, M. K., 1995, ''Vibration of Stress-Free Hollow Cylinders of Arbitrary Cross Section,'' ASME J. Appl. Mech., **62**, pp. 714–724.

[14] So, J., and Leissa, A. W., 1997, ''Free Vibrations of Thick Hollow Circular Cylinders from Three-Dimensional Analysis,'' ASME J. Vibr. Acoust., **119**, pp. 89–95.

[15] Liew, K. M., Hung, K. C., and Lim, M. K., 1995, ''Vibratory Characteristics of Simply Supported Thick Skew Plates in Three-Dimensional Setting,'' ASME J. Appl. Mech., **62**, pp. 880–886.

[16] Liew, K. M., Hung, K. C., and Lim, M. K., 1998, ''Vibration of Thick Prismatic Structures With Three-Dimensional Flexibility,'' ASME J. Appl. Mech., accepted for publication.

[17] Srinivasan, R. S., and Krishnan, P. A., 1987, ''Free Vibration of Conical Shell Panels,'' J. Sound Vib., **117**, pp. 153–160.

[18] Cheung, Y. K., Li, W. Y., and Tham, L. G., 1989, ''Free Vibration Analysis of Singly Curved Shell by Spline Finite Strip Method,'' J. Sound Vib., **128**, pp. 411–422.

D. W. Gulick
Graduate Researcher

O. M. O'Reilly[1]
Associate Professor,
Assoc. Mem. ASME,
e-mail: oreilly@me.berkeley.edu

Department of Mechanical Engineering,
University of California,
Berkeley, CA 94720-1740

On the Dynamics of the Dynabee

The Dynabee *is a gyroscopic device that is marketed as a wrist exerciser. In this paper, a model for the dynamics of this device is presented. With some additional work, we find that the dynamics are governed by a single ordinary differential equation. The solution of this equation also provides the moment required to operate the device. Specifically, we find that this moment is proportional to the square of the rotor's spin rate. We also show why it is necessary to give the device a large initial spin rate for its successful operation.* [S0021-8936(00)02602-7]

1 Introduction

A unique hand-held gyroscope invented by Archie Mishler [1] exhibits the intriguing phenomenon of rotor spin-up when an individual applies the appropriate torques to the supporting frame of the gyroscope. As described by Mishler, this ''new gyroscopic device in which the rotor cannot only rotate about its spin axis but can also rotate about a second axis at right angles to the spin axis, and in which the rotor can be made to increase in speed by applying a torque about a third axis,'' became popular in the early 1990s due to the marketing of a wrist and arm exerciser called the *Dynabee*™. Sold as both a therapeutic and strengthening device, the *Dynabee* is held in the palm of one hand, the rotor is manually given an initial spin, and then a coordinated motion of the wrist and hand can increase the spin rate of the rotor to speeds near 4000 rpm.

This paper is concerned with explaining the manner in which the spin rate approaches these high speeds. In addition, the paper discusses the applied moment that is required to achieve this motion. We will show that the magnitude of the moment is proportional to the spin rate squared, a feature which allows the *Dynabee* to be used as an wrist exerciser.

Our model is based on rigid-body dynamics. It can be considered as an example of a nonholonomically constrained dynamical system. Reviews of, and further details on, works in this area can be found in Neimark and Fufaev [2], Karapetyan and Rumyantsev [3], and Zenkov, Bloch, and Marsden [4].

2 Kinematics

In this section, we discuss the kinematics of the proposed mechanical model of the *Dynabee*. This model contains the important physical characteristics necessary for spin-up to occur. Specifically, bases vectors and Euler angles are introduced to describe the rotations of two bodies: a circular track and a rotor. For the relevant background on parameterizations of rotation tensors, the reader is referred to the review article by Shuster [5].

The mechanical model is comprised of two rigid bodies, a track and a rotor, as illustrated in Fig. 1. The track is a circular race or groove that constrains the motion of the rotor, which is an axisymmetric body with a cylindrical axle along its axis of symmetry. With a semi-length of R_t and a radius of R_a, both ends of the axle are constrained to remain within the track. The rotor's moment of inertia about a vector parallel to the axle, referred to as the rotor's spin-axis, is λ_1. Due to a geometric symmetry, the moment of inertia about any vector perpendicular to the spin axis is λ_2.

We prescribe the rotational motion of the track such that a body-fixed axis normal to the track's plane precesses with a constant nutation angle θ and a precessional rate $\dot{\psi}$. To describe the rotation, we introduce a right-handed orthonormal basis $\boldsymbol{\varepsilon}$ of an inertial coordinate system comprised of the vectors $\mathbf{E}_1$, $\mathbf{E}_2$ and $\mathbf{E}_3$. Starting from a reference configuration where the normal to the track's plane is aligned with $\mathbf{E}_3$, the rotation of the track may be specified by a 3-1-3 (ψ, θ, φ) set of Euler angles, as shown in Fig. 2. Since we desire that the outer shell have a purely precessional motion that does not involve the body revolving about $\mathbf{E}_3$, the final angle of rotation is $\varphi = -\psi$. In the resulting motion, a material point of the track returns to the same location after one full precession *without* revolving around $\mathbf{E}_3$.

We define a set of body-fixed vectors $\{\mathbf{e}_{t1}, \mathbf{e}_{t2}, \mathbf{e}_{t3}\}$ that form a right-handed orthonormal basis $\boldsymbol{\varepsilon}_t$ which corotates with the track. Here, $\mathbf{e}_{t1}$ and $\mathbf{e}_{t2}$ lie in the plane that contains the track, while $\mathbf{e}_{t3}$ is normal to the track's plane. Restricting the nutation angle θ to be constant, the angular velocity vector of the circular track is

$$\boldsymbol{\omega}_t = \dot{\psi}\mathbf{E}_3 - \dot{\psi}\mathbf{e}_{t3} \tag{1}$$

where a dot over a variable denotes differentiation with respect to time.

Only two independent Euler angles are necessary to specify the current orientation of the rotor relative to the track, as viewed in Fig. 3. The first rotation, through an angle α, is about $\mathbf{e}_{t3}$, and causes the axle to rotate around the track. Following this rotation, it is convenient to define a right-handed orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ such that $\mathbf{e}_1$ is parallel to the axle and $\mathbf{e}_3$ is parallel to $\mathbf{e}_{t3}$. Thus, the vector $\mathbf{e}_2$ lies in the plane spanned by $\mathbf{e}_{t1}$ and $\mathbf{e}_{t2}$. Note that this basis is neither corotational with the track nor corotational with the rotor. For the second rotation, the rotor spins about $\mathbf{e}_1$ through an angle γ. Corotational with the rotor is the right-handed orthonormal basis $\boldsymbol{\varepsilon}_r = \{\mathbf{e}_{r1}, \mathbf{e}_{r2}, \mathbf{e}_{r3}\}$. This basis is oriented such that $\mathbf{e}_{r1}$ is parallel to the rotor's axle. Given the relative rotation of the rotor, we use the methods of Casey and Lam [6] to calculate the angular velocity vector of the rotor relative to the track $\hat{\boldsymbol{\omega}}_{r,t}$:

$$\hat{\boldsymbol{\omega}}_{r,t} = \boldsymbol{\omega}_r - \boldsymbol{\omega}_t = \dot{\gamma}\mathbf{e}_1 + \dot{\alpha}\mathbf{e}_3 \tag{2}$$

where $\boldsymbol{\omega}_r$ is the absolute angular velocity vector of the rotor.

In postulating constraints on the motion of the rotor, several conditions could exist at the contact point between the track and the rotor's axle. These include frictionless sliding, sliding with friction, and rolling without sliding. We shall only consider the last case because it is the only condition that presents a mechanism for spin-up of the rotor. Figure 4 shows the proposed type of contact between the rotor and the track. The axle contacts the track at a point P on the track's lower surface and at a point Q on the upper surface. In this configuration, the center of mass of the rotor remains coincident with the center of the track as the rotor rolls at both points P and Q.

[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 1, 1999; final revision, Oct. 25, 1999. Associate Technical Editor: N. C. Perkins. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

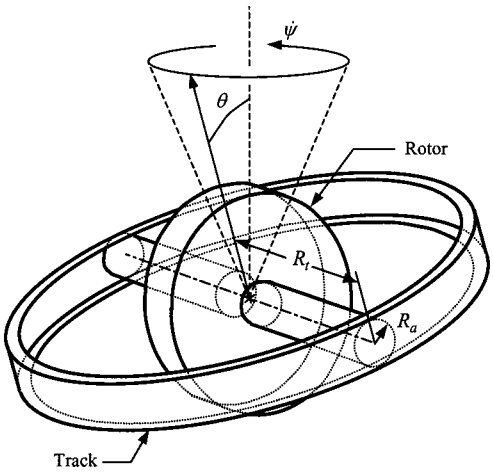


Fig. 1 Schematic of the *Dynabee*. The precessional motion of a vector normal to the circumferential track is also shown.

Introducing a right-handed orthonormal basis $\{\mathbf{e}_{\pi 1}, \mathbf{e}_{\pi 2}, \mathbf{e}_{\pi 3}\}$, as illustrated in Fig. 4, facilitates the analysis. Here, $\mathbf{e}_{\pi 1}$ points from the center of the rotor (point O) toward point P such that the angle β between $\mathbf{e}_{\pi 1}$ and $\mathbf{e}_1$ is

$$\beta = \tan^{-1}\left(\frac{R_a}{R_t}\right). \tag{3}$$

Furthermore, $\mathbf{e}_{\pi 2}$ is parallel to $\mathbf{e}_2$. Rolling at P and Q introduces the following constraint equations:

$$\begin{aligned}(\mathbf{v}_r + \boldsymbol{\omega}_r \times \boldsymbol{\pi}_P) - (\mathbf{v}_t + \boldsymbol{\omega}_t \times \boldsymbol{\pi}_P) &= \mathbf{0},\\ (\mathbf{v}_r + \boldsymbol{\omega}_r \times \boldsymbol{\pi}_Q) - (\mathbf{v}_t + \boldsymbol{\omega}_t \times \boldsymbol{\pi}_Q) &= \mathbf{0},\end{aligned} \tag{4}$$

where $\mathbf{v}_r$ is the velocity vector of the rotor's center of mass, $\mathbf{v}_t$ is the velocity vector of the center of mass of the track, and $\boldsymbol{\pi}_P$ and $\boldsymbol{\pi}_Q$ are the position vectors of P and Q relative to the rotor's center of mass, respectively. Clearly,

$$\boldsymbol{\pi}_P = -\boldsymbol{\pi}_Q. \tag{5}$$

It follows that we can separate the constraint equations into a single vector equation for translation,

$$\mathbf{v}_r - \mathbf{v}_t = \mathbf{0}, \tag{6}$$

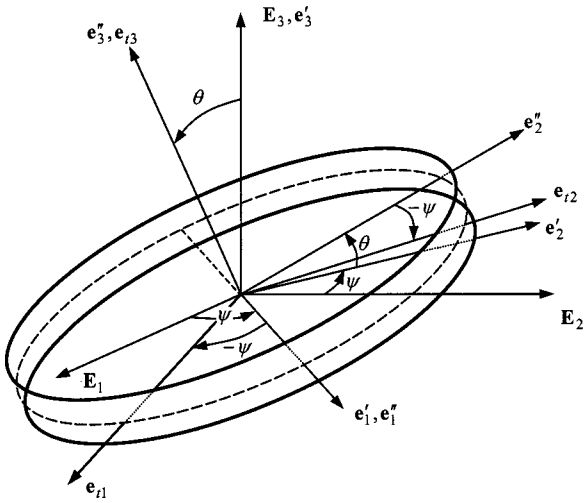


Fig. 2 The Euler angles and reference frames used to parameterize the precessional motion of the track

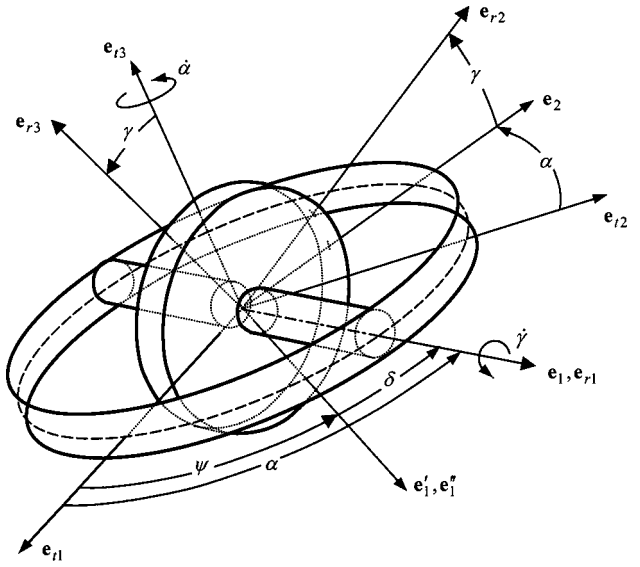


Fig. 3 The Euler angles and reference frames used to parameterize the rotor's rotation relative to the track

and another vector equation for rotation,

$$(\boldsymbol{\omega}_r - \boldsymbol{\omega}_t) \times \boldsymbol{\pi}_P = \mathbf{0}. \tag{7}$$

Equation (6) provides three scalar constraints that fully specify the translational motion of the rotor. Throughout the remainder of the paper, we shall assume that the centers of mass of the track and rotor are coincident and stationary. Taking the dot products of (7) with $\mathbf{e}_{\pi i}$ $(i = 1,2,3)$ yields two scalar constraint equations

$$\begin{aligned}\Psi_1 &= (\boldsymbol{\omega}_r - \boldsymbol{\omega}_t) \cdot \mathbf{e}_{\pi 2} = 0,\\ \Psi_2 &= (\boldsymbol{\omega}_r - \boldsymbol{\omega}_t) \cdot \mathbf{e}_{\pi 3} = 0.\end{aligned} \tag{8}$$

The second constraint may be transformed into an algebraic relationship between the various Euler angles by substituting for the relative angular velocity vector (from (2)) and then integrating over time. Allowing the initial values of the Euler angles to be zero yields the algebraic relationship

$$\gamma = -\zeta\alpha, \tag{9}$$

where $\zeta = R_t / R_a$. Thus the rotor's motion has been reduced to a problem with only one independent variable. In addition to reducing the number of independent variables, the rotational constraints will be used to postulate appropriate constraint moments on the rotor.

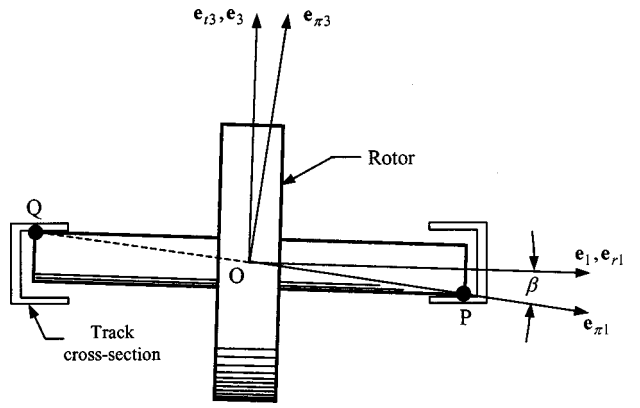


Fig. 4 Rolling contact between the rotor and track as viewed from the negative $\mathbf{e}_2$ direction

3 Kinetics

In our model, we assume that the net moment applied to the rotor, $\mathbf{M}_r$, is the sum of two moments: a constraint moment due to the track-rotor contact, denoted as $\mathbf{M}_{r_c}$, and a dissipative moment $\mathbf{M}_{r_d}$. Hence

$$\mathbf{M}_r = \mathbf{M}_{r_c} + \mathbf{M}_{r_d}. \tag{10}$$

Using a normality prescription ([7,8]), the constraint moment on the rotor (relative to the rotor's center of mass) is

$$\mathbf{M}_{r_c} = \kappa_2 \mathbf{e}_{\pi 2} + \kappa_3 \mathbf{e}_{\pi 3} \tag{11}$$

where κ_2 and κ_3 are indeterminate. A physical reason that the constraint moment has no component in the $\mathbf{e}_{\pi 1}$ direction is that this moment is generated by reaction forces $\mathbf{F}_P$ and $\mathbf{F}_Q$ at P and Q, respectively. It is easy to see that $\mathbf{M}_{r_c} = \boldsymbol{\pi}_P \times \mathbf{F}_P + \boldsymbol{\pi}_Q \times \mathbf{F}_Q$, and thus it is impossible to generate a component of the constraint moment along a vector connecting the center of the rotor and either of these points.

The constraint moment must satisfy two conditions for rolling to occur at P and Q. First, for contact at these points to be maintained

$$\kappa_2 < 0. \tag{12}$$

Additionally, the conditions for static Coulomb friction must be satisfied to permit rolling without sliding:

$$|\kappa_3 / \kappa_2| \leqslant \mu_s \sqrt{1 + \frac{1}{\zeta_2}} \cong \mu_s \tag{13}$$

where μ_s is the coefficient of static friction, and the approximation results if one employs the physically plausible assumption that $\zeta \gg 1$.

It is reasonable to believe that there is some energy loss within the system and that this loss can be modeled as a moment that opposes the rotation of the rotor. The dissipation could be caused either by frictional loss at the axle-track contact point or aerodynamic drag. A simple viscous model for the dissipative moment is

$$\mathbf{M}_{r_d} = -\sigma \omega_1 \mathbf{e}_1, \tag{14}$$

where σ is a positive constant and $\omega_1 = \boldsymbol{\omega}_r \cdot \mathbf{e}_1$. It is easy to see that $\mathbf{M}_{r_d} \cdot \boldsymbol{\omega}_r \leqslant 0$.

To calculate the angular momentum of the rotor, we note that its inertia tensor $\mathbf{J}$ is

$$\mathbf{J} = \lambda_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + \lambda_2 (\mathbf{e}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3) \tag{15}$$

in the current configuration.[2] Here, λ_1 is the principal moment of inertia about the rotor's axis of symmetry, and λ_2 is the other distinct principal moment of inertia. By combining (1) and (2) and applying the geometric relationship $\gamma = -\zeta\alpha$ (cf. (9)), the components of the rotor's angular velocity $\boldsymbol{\omega}_r$ are obtained:

$$\begin{aligned} \omega_1 &= \boldsymbol{\omega}_r \cdot \mathbf{e}_1 = -\zeta \dot{\alpha} + \dot{\psi} \sin(\alpha - \psi) \sin\theta, \\ \omega_2 &= \boldsymbol{\omega}_r \cdot \mathbf{e}_2 = \dot{\psi} \cos(\alpha - \psi) \sin\theta, \\ \omega_3 &= \boldsymbol{\omega}_r \cdot \mathbf{e}_3 = \dot{\alpha} - \dot{\psi}(1 - \cos\theta). \end{aligned} \tag{16}$$

Using (15) and (16), we find that the angular momentum $\mathbf{H}$ relative to the center of mass of the rotor is

$$\mathbf{H} = \mathbf{J}\boldsymbol{\omega}_r = \lambda_1 \omega_1 \mathbf{e}_1 + \lambda_2 (\omega_2 \mathbf{e}_2 + \omega_3 \mathbf{e}_3). \tag{17}$$

The balance of angular momentum for the rotor is

$$\dot{\mathbf{H}} = \mathbf{M}_r. \tag{18}$$

Since the constraint moment only has components along $\mathbf{e}_{\pi 2}$ and $\mathbf{e}_{\pi 3}$, the components of the balance of angular momentum in the $\mathbf{e}_{\pi i}$ $(i = 1,2,3)$ directions involve one uncoupled differential equation for the final independent Euler angle α and two equations for the constraint moment. Evaluating $\mathbf{e}_{\pi i} \cdot \dot{\mathbf{H}} = \mathbf{e}_{\pi i} \cdot \mathbf{M}_r$ yields

$$\begin{aligned} &\zeta(\eta \dot{\omega}_1) - [\dot{\omega}_3 + \zeta \dot{\alpha} \omega_2 + (1 - \eta)\omega_1 \omega_2] + \frac{\sigma \zeta}{\lambda_2} \omega_1 = 0, \\ &\dot{\omega}_2 - \zeta \dot{\alpha} \omega_3 + (\eta - 1)\omega_1 \omega_3 = \frac{\kappa_2}{\lambda_2}, \\ &\eta \dot{\omega}_1 + \zeta[\dot{\omega}_3 + \zeta \dot{\alpha} \omega_2 + (1 - \eta)\omega_1 \omega_2] + \frac{\sigma}{\lambda_2} \omega_1 = \frac{\kappa_3}{\lambda_2 \sin\beta}. \end{aligned} \tag{19}$$

The derivation of these scalar equations uses the relationships in (3) and (9), as well as a dimensionless parameter η:

$$\eta = \frac{\lambda_1}{\lambda_2}. \tag{20}$$

Using (16), $(19)_1$ becomes an uncoupled, nonlinear ordinary differential for $\alpha(t)$. We simplify this governing equation by introducing a new angle (see Fig. 3): the phase angle

$$\delta(t) = \alpha(t) - \psi(t). \tag{21}$$

Additionally, we nondimensionalize by introducing

$$\tau = \dot{\psi}_o t \tag{22}$$

where $\dot{\psi}_o$ is the initial condition for the precession rate ($\dot{\psi}_o = \dot{\psi}(0)$). Substitution into $(19)_1$, yields

$$\begin{aligned} &\delta'' + a\delta' + b\psi'^2 \cos\delta - (b\psi'' + c\psi')\sin\delta - d\psi'^2 \sin 2\delta \\ &\quad = e\psi'' - a\psi'. \end{aligned} \tag{23}$$

Here, each apostrophe denotes differentiation with respect to the dimensionless variable τ (e.g., $\delta' = d\delta/d\tau$ and $\delta'' = d^2\delta/d\tau^2$) and the constant coefficients are

$$\begin{aligned} a &= \frac{\zeta^2 \nu}{1 + \zeta^2 \eta}, \quad b = \frac{\zeta \eta \sin\theta}{1 + \zeta^2 \eta}, \quad c = \frac{\zeta \nu \sin\theta}{1 + \zeta^2 \eta}, \\ d &= \frac{1}{2} \frac{(\eta - 1)\sin^2\theta}{1 + \zeta^2 \eta}, \quad e = \frac{\zeta^2 \eta + \cos\theta}{1 + \zeta^2 \eta}. \end{aligned} \tag{24}$$

A new dimensionless parameter ν, that reflects the importance of the dissipative moment, appears in (24):

$$\nu = \frac{\sigma}{\lambda_2 \dot{\psi}_o}. \tag{25}$$

Once we find a solution to the single equation of motion for the rotor, (23), we can then calculate $\mathbf{M}_{r_c}$ and verify that the criteria for rolling are satisfied. Knowing $\delta(t)$, we can then determine

$$\begin{aligned} \mathbf{M}_{r_c} &= \lambda_2[\dot{\omega}_2 - \zeta \dot{\alpha} \omega_3 + (\eta - 1)\omega_1 \omega_3]\mathbf{e}_{\pi 2} \\ &\quad + \lambda_2 \sqrt{1 + \zeta^2}(\eta \dot{\omega}_1 + \nu \dot{\psi}_o \omega_1)\mathbf{e}_{\pi 3}, \end{aligned} \tag{26}$$

which follows from (19). Finally, the criteria (12) and (13) determine whether the rolling motion of the rotor is physically possible.

4 Constant Track Precession Rate

When the track's precession rate $\dot{\psi}$ is constant, (23) simplifies greatly. It also becomes tractable to identify steady-state motions and to study their stability. Furthermore, we can identify a conserved quantity when the only moment acting on the rotor is the constraint moment (i.e., no dissipative moment is present). We now assume that the track rotates with a constant precession rate of $\dot{\psi}(t) = \dot{\psi}_o$. Hence $\psi'(\tau) = 1$, $\psi''(\tau) = 0$, and (23) becomes

$$\delta'' + a\delta' + b\cos\delta - c\sin\delta - d\sin 2\delta = -a. \tag{27}$$

[2] See Beatty [9] and Casey [7] for detailed discussions on the use of tensors in rigid-body dynamics.

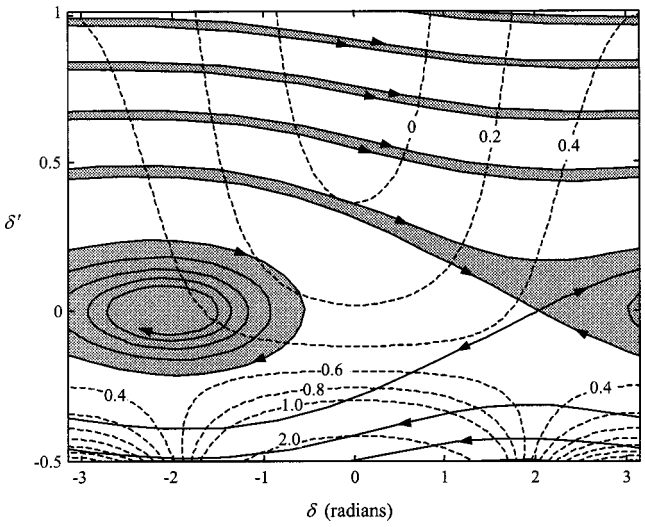


Fig. 5 Phase plane for the rotor response given a constant precession rate of the track. The solid lines are trajectories for several initial conditions, the shaded area is the basin of attraction of the stable equilibrium, and the dashed lines are contours for the ratio $|\kappa_3/\kappa_2|$. The phase plane is of (27) with $\zeta=29$, $\eta=1.7$, $\nu=0.018$, and $\theta=\pi/4$.

Figure 5 shows a phase portrait of (27) with parameters characteristic of the *Dynabee*. In this figure, the solid lines are solution trajectories for several initial conditions chosen specifically so that the solutions either originate at the unstable equilibrium point or approach it. The shaded region is the basin of attraction of the stable equilibrium. Also shown in Fig. 5 are the contour lines for the ratio $|\kappa_3/\kappa_2|$, which determines whether or not the rolling criterion in (13) is satisfied. If a trajectory passes through a portion of the phase plane where this ratio exceeds the coefficient of static friction μ_s, then sliding will occur and the equation of motion (27) is no longer applicable. The criterion for maintaining contact between the rotor's axle and the track, $\kappa_2<0$ (cf. (12)), is satisfied everywhere within the region of states shown in Fig. 5; however, this criterion is not satisfied near $\delta'=-1$. We remark that trajectories lying outside the basin of attraction, modulo a set of measure zero, will eventually result in decreasing values of δ'. As a result, these trajectories will then correspond to motions of the *Dynabee* where slipping of the rotor relative to the track occurs and (27) becomes invalid.

The solutions of (27) when $\delta'(\tau)=\delta''(\tau)=0$ determine the existence and locations of the system equilibria. For values of $\zeta \gg 1$, there exist either two equilibria (one stable and one unstable) or no equilibria, depending on the relative values of the dimensionless parameters ζ, η, and ν and the nutation angle θ. When equilibria exist, the rotor can attain a steady state motion at the stable equilibrium point where the axle revolves around the track at the same rate that the track is precessing; the resulting spin rate is $\dot{\gamma}(t)=-\zeta\dot{\psi}_o$, a result which explains the magnification of the spin rate as a function of the precession rate.

In the case that the dissipative moment is negligible ($\sigma=0$), the equation of motion (27) reduces to

$$\delta''+b\cos\delta-d\sin 2\delta=0 \tag{28}$$

and it is straightforward to show that the response conserves a quantity analogous to the total energy of the system:

$$\frac{d}{d\tau}\left(\frac{1}{2}\delta'^2+b\sin\delta+\frac{1}{2}d\cos 2\delta\right)=0. \tag{29}$$

Furthermore, for the conservative system, a stable equilibrium point *always* exists at $\delta=-\pi/2$ and an unstable equilibrium point exists at $\delta=\pi/2$. At both of these equilibria, the contact forces between the rotor's axle and the track are purely normal (i.e., there is no frictional force so $\kappa_3=0$), as is expected since the magnitude of the rotor's angular momentum is constant. The constraint moment evaluated at the stable equilibrium is

$$\mathbf{M}_{r_c}=\left(\frac{\dot{\gamma}}{\zeta}\right)^2\cos\theta[\lambda_1\zeta+(\lambda_1-\lambda_2)\sin\theta]\mathbf{e}_{\pi 2}. \tag{30}$$

In writing (30), we have used the relation $\dot{\gamma}(t)=-\zeta\dot{\psi}_o$. Ignoring the inertia of the track and its casing, this is equal to the moment that the holder of the *Dynabee* is forced to exert. It should be noted that the applied moment is a quadratic function of $\dot{\gamma}$.

5 Increasing the Track Precession Rate

We now prescribe the precessional motion of the track to begin at a given precessional rate $\dot{\psi}_o$ ($\dot{\psi}_o>0$) and increase slowly thereafter according to

$$\dot{\psi}(t)=\dot{\psi}_o(1+\varepsilon\dot{\psi}_o t), \tag{31}$$

where ε is a constant such that $0<\varepsilon\ll 1$. The dimensionless equivalent to (31) is $\psi'(\tau)=1+\varepsilon\tau$. Substituting this expression into the equation of motion for the rotor, (23), we arrive at

$$\delta''+a\delta'+b(1+\varepsilon\tau)^2\cos\delta-[b\varepsilon+c(1+\varepsilon\tau)]\times\sin\delta-d(1+\varepsilon\tau)^2\sin 2\delta=-a(1+\varepsilon\tau)-e\varepsilon \tag{32}$$

where a, b, c, d, and e were defined by (24).

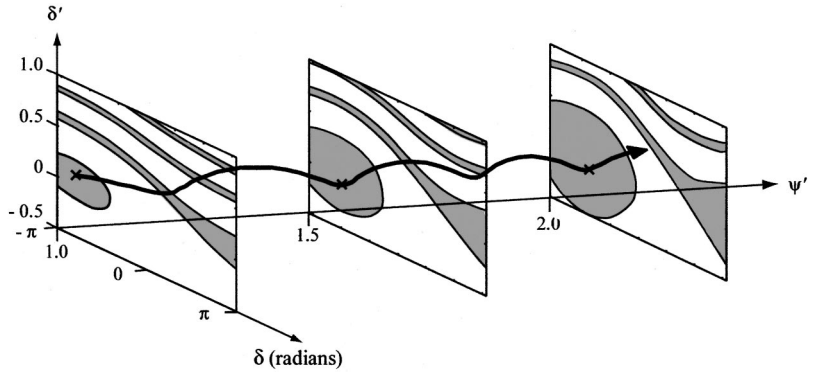


Fig. 6 Cross sections of the basin of attraction of the stable equilibrium for (32) when the track's precession rate is increasing according to (31) with $\zeta=29$, $\eta=1.7$, $\nu=0.018$, $\theta=\pi/4$, and $\varepsilon=0.01$. A sample trajectory is also shown whose initial conditions are $\psi'(0)=1$, $\delta(0)=-2.5$, and $\delta'(0)=0.1$.

As in the constant precession case, solutions to the equation of motion were numerically calculated. For a given set of parameters $\{\zeta, \eta, \nu, \theta, \varepsilon\}$, any trajectory lying in the basin of attraction of the stable equilibrium of (32) also satisfies

$$\lim_{\tau\to\infty} \gamma'(\tau) = -\zeta\psi'(\tau) = -\zeta(1+\varepsilon\tau), \tag{33}$$

and is thus a response that exhibits spin-up. Figure 6 shows cross sections of the basin of attraction at $\psi' = 1.0$, 1.5, and 2.0 for a choice of parameters characteristic of the *Dynabee* and with the spin-up coefficient ε equal to 0.01.

The trajectory of the initial conditions $\psi'(0) = 1$, $\delta(0) = -2.5$ and $\delta'(0) = 0.1$ is also included in Fig. 6 as an example of a trajectory that lies within the basin of attraction. The response extending from these initial conditions approaches the following steady-state conditions as time increases:

$$\lim_{\tau\to\infty} \delta(\tau) = -\pi/2, \qquad \lim_{\tau\to\infty} \delta'(\tau) = 0. \tag{34}$$

The maximum value of $|\kappa_3/\kappa_2| = 0.99$ is attained when $\psi' = 1.24$, and as time increases and ψ' becomes larger, this ratio approaches zero as $|\kappa_3/\kappa_2| \propto 1/\psi'$. Simultaneously, $\mathbf{M}_{r_c}$ grows proportionally to ψ'^2. Thus, when the precession rate is large, the response resembles the equilibrium conditions for the conservative system described in Section 4.

As evidenced in Fig. 6, the basin of attraction of the stable equilibrium changes as ψ' becomes larger.[3] It is therefore possible for initial conditions (δ, δ') which were originally outside this basin to become trapped inside it. This phenomenon is often known as resonance capture. It has been studied in a variety of mechanical systems including satellites and celestial bodies (cf., e.g., [10–13]).

[3]We are grateful to an anonymous reviewer for bringing this matter to our attention.

In summary, we find that as the rotor spins up, the coefficient of static Coulomb friction needed to sustain rolling decreases. This explains the observed phenomenon that it is difficult to spin up a *Dynabee* if the rotor is not given a sufficiently large initial spin rate. It would clearly be of interest to examine the case where the rotor slides on the track, but we leave this matter for future work.

Acknowledgment

The work of David Gulick was supported by a NSF Graduate Fellowship.

References

[1] Mishler, A. L., 1973, ''Gyroscopic Device,'' U.S. Patent 3726146.
[2] Neimark, J. I., and Fufaev, N. A., 1972, *Dynamics of Nonholonomic Systems*, English translation by J. R. Barbour, American Mathematics Society, Providence, RI.
[3] Karapetyan, A. V., and Rumyantsev, V. V., 1990, ''Stability of Conservative and Dissipative Systems,'' Appl. Mech. Sov. Rev., **1**, pp. 3–144.
[4] Zenkov, D. V., Bloch, A. M., and Marsden, J. E., 1998, ''The Energy-Momentum Method for the Stability of Non-Holonomic Systems,'' Dyn. Stabil. Syst., **13**, pp. 123–165.
[5] Shuster, M. D., 1993, ''A Survey of Attitude Representations,'' J. Astronaut. Sci., **41**, No. 4, pp. 439–517.
[6] Casey, J., and Lam, V. C., 1986, ''On the Relative Angular Velocity Tensor,'' ASME J. Mech., Transm. Autom. Des., **108**, pp. 399–400.
[7] Casey, J., 1995, ''On the Advantages of a Geometric Viewpoint in the Derivation of Lagrange's Equations for a Rigid Continuum,'' J. Appl. Math. Phys., **46**, pp. S805–S847.
[8] O'Reilly, O. M., and Srinivasa, A. R., 1999, ''On Constraints and Potential Energies in Systems of Rigid Bodies and Particles,'' submitted for publication.
[9] Beatty, M. F., Jr., 1986, *Principles of Engineering Mechanics*, Vol. 1, Plenum Press, New York.
[10] Haberman, R., Rand, R., and Yuster, T., 1999, ''Resonant Capture and Separatrix Crossing in Dual-Spin Spacecraft,'' Nonlinear Dyn., **18**, pp. 159–184.
[11] Hall, C. D., and Rand, R. H., 1994, ''Spinup Dynamics of Axial Dual-Spin Spacecraft,'' AIAA J. Guid. Cont. Dyn., **17**, pp. 30–37.
[12] Henrard, J., 1982, ''Capture into Resonance: An Extension of the Use of Adiabatic Invariants,'' Celest. Mech., **27**, pp. 3–22.
[13] Lochak, P., and Meunier, C., 1988, *Multiphase Averaging for Classical Systems With Application to Adiabatic Theorems*, Springer-Verlag, New York.

Generalized Hellinger-Reissner Principle

J.-H. He
Shanghai Institute of Applied Mathematics and Mechanics,
Shanghai University,
149 Yanchang Road,
Shanghai 200072, P. R. China

By the semi-inverse method of establishing variational principles, the Hellinger-Reissner principle can be obtained straightforwardly from energy trial-functionals without using Lagrange multipliers, and a family of generalized Hellinger-Reissner principles with an arbitrary constant are also obtained, some of which are unknown to us at the present time. The present theory provides a straightforward tool to search for various variational principles directly from governing equations and boundary conditions.
[S0021-8936(00)00702-9]

1 Introduction

It is well known that the Hellinger-Reissner principle can be deduced from the minimum complementary energy principle via the Lagrange multiplier method by eliminating the constraints of equilibrium equations. In this paper, we will rederive the well-established principle directly from its governing equations and boundary conditions without Lagrange multipliers via the semi-inverse method previously proposed by this author. In addition, a family of generalized Hellinger-Reissner principles will also be obtained.

2 Mathematical Formulation of Small Displacement Problems in Elasticity ([1,2])

The basic equations governing the elastic body subjected to the action of a distributed body force can be written as follows:

1 Equilibrium conditions:

$$\sigma_{ij,j}+f_i=0 \quad (\text{in } \tau) \tag{1}$$

in which σ_{ij} are stresses, $\sigma_{ij,j}=\partial\sigma_{ij}/\partial x_j$, f_i represent body forces, and τ is the volume of an elastic body.

2 Stress-strain relations:

For linear elasticity, we have

$$\sigma_{ij}=a_{ijkl}e_{kl} \quad (\text{in } \tau) \tag{2a}$$

or

$$e_{ij}=b_{ijkl}\sigma_{kl} \quad (\text{in } \tau) \tag{2b}$$

in which e_{ij} are strains, and a_{ijkl}, b_{ijkl} represent elastic and compliance constants, respectively.

Let us now introduce the strain energy density A and complementary B. They are defined in general by

$$A=\int_0^e \sigma_{ij}\mathbf{d}e_{ij} \quad \text{or} \quad \frac{\partial A}{\partial e_{ij}}=\sigma_{ij} \tag{2c}$$

$$B=\int_0^\sigma e_{ij}\mathbf{d}\sigma_{ij} \quad \text{or} \quad \frac{\partial B}{\partial \sigma_{ij}}=e_{ij} \tag{2d}$$

and satisfy the following energy identity:

$$A+B=e_{ij}\sigma_{ij}. \tag{2e}$$

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 17, 1997; final revision, January 16, 1998. Associate Technical Editor: W. K. Liu. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

It should be specially pointed out that Eqs. (2*a*)–(2*d*) are equivalent to each other.

3 Strain-displacement relations:

$$e_{ij}=\frac{1}{2}(u_{i,j}+u_{j,i}) \quad (\text{in } \tau). \tag{3}$$

4 Boundary conditions for given surface displacement:

$$u_i=\overline{u_i} \quad (\text{in } \Gamma_u) \tag{4}$$

5 Boundary conditions for given external force on boundary surface:

$$\sigma_{ij}n_j=\overline{p_i} \quad (\text{in } \Gamma_\sigma) \tag{5}$$

where $\Gamma_u+\Gamma_\sigma=\Gamma$ covers the total boundary surface.

The known variational principles are recapitulated below:

1 The Minimum Potential Energy Principle.

$$J^P(u_i)=\int\int\int (A-f_iu_i)\mathbf{d}\tau-\int_{\Gamma_\sigma}\int \overline{p_i}u_i\mathbf{d}S \tag{6}$$

where u_i is one kind of independent variable, Eqs. (2) of stress-strain relations, Eq. (3) of strain-displacement relations, and the boundary conditions (4) for given surface displacement are its variational constraints.

2 Hellinger-Reissner Variational Principle.

$$J_{HR}(\sigma_{ij},u_i)=\int\int\int (B+\sigma_{ij,j}u_i+f_iu_i)\mathbf{d}\tau-\int_{\Gamma_u}\int \sigma_{ij}n_j\overline{u_i}\mathbf{d}S$$
$$-\int_{\Gamma_\sigma}\int u_i(\sigma_{ij}n_j-\overline{p_i})\mathbf{d}S \tag{7}$$

where σ_{ij}, u_j are two kinds of independent variables, and Eq. (2) of stress-strain relations are its variational constraints.

It is interesting to note that if Eq. (3) of strain-displacement relations is treated as a constraint then its stationary conditions satisfy Eqs. (1) and (2). Here a simple proof is given.

Proof: Making the above functional (7) stationary, in view of the independent of σ_{ij} and u_i, we can obtain (1), (4), and (5) and the following equation

$$\frac{\partial B}{\partial \sigma_{ij}}-\frac{1}{2}(u_{i,j}+u_{j,i})=0 \tag{8}$$

as a Euler equation

Using the constraints of Eq. (3), the above Euler equation reduces to Eq. (2*d*), the stress-strain relation. The interesting result lies in the fact that by involutory transformation ([3,4]) the form

of the functional (7) remains unchanged, while its stationary conditions are changed. Therefore we can use Eqs. (2) or (3) as constraints when using the finite element method based on the Hellinger-Reissner principle or other direct variational methods to solve a problem.

3 Semi-Inverse Method and Trial-Functional

In this paper, we will rederive the well-known Hellinger-Reissner principle without using the Lagrange multipliers. As a first step towards this end, we will first introduce the concepts of the semi-inverse method and trial-functional ([5–8]).

Using Lagrange multipliers to remove the constraints of strain-displacement relations (3) and stress-strain relations (2) in the minimum potential energy principle (7), one obtains

$$J(u_i,e_{ij}\sigma_{ij},\lambda_{ij},\beta_{ij})=J_P(u_i)+\int\int\int \lambda_{ij}\left(e_{ij}-\frac{1}{2}u_{i,j}-\frac{1}{2}u_{j,i}\right)\mathbf{d}\tau+\int\int\int \beta_{ij}\left(\sigma_{ij}-\frac{\partial A}{\partial e_{ij}}\right)\mathbf{d}\tau \tag{9}$$

where λ_{ij} and β_{ij} are Lagrange multipliers to be determined.

In the process of identification of Lagrange multipliers, the multipliers are treated as independent variations *a priori*, but after identification, they become the functions of the other original variables. That is to say the Lagrange multipliers are not independent at all, the contradiction existing in the Lagrange multiplier method may lead to a variational crisis ([9]). To overcome the contradiction mentioned above, we can pre-assume that the multipliers in functional (9) can be expressed as follows ([10]):

$$\lambda_{ij}=\lambda_{ij}(u_i,e_{ij},\sigma_{ij},u_{i,j},e_{ij,j},\sigma_{ij,j}) \tag{10}$$

$$\beta_{ij}=\beta_{ij}(u_i,e_{ij},\sigma_{ij},u_{i,j},e_{ij,j},\sigma_{ij,j}). \tag{11}$$

To simplify the identification of the Lagrange multipliers, the author introduces an unknown function F to replace the terms involving the multipliers ([10]):

$$F=\lambda_{ij}\left(e_{ij}-\frac{1}{2}u_{i,j}-\frac{1}{2}u_{j,i}\right)+\beta_{ij}\left(\sigma_{ij}-\frac{\partial A}{\partial e_{ij}}\right). \tag{12}$$

Accordingly, the functional (9) can be rewritten as follows:

$$J(u_i,e_{ij},\sigma_{ij})=\int\int\int (A-f_iu_i+F)\mathbf{d}\tau+\int\int_{\Gamma_u} G\mathbf{d}S+\int\int_{\Gamma_\sigma} H\mathbf{d}S \tag{13}$$

where F, G, and H are unknowns to be determined.

The above functional with unknowns F, G, and H is called the *trial-functional*, or *energy trial-functional*, the identification of the unknown function is very similar to that of the Lagrange multipliers.

The energy trial-functional has several ways to be constructed, and details have been discussed in [5,6]; for example, the following three integrals with an energy form can be used as trial-functionals.

$$J=\int\int\int\left\{\sigma_{ij}\left(e_{ij}-\frac{1}{2}u_{i,j}-\frac{1}{2}u_{j,i}\right)+F\right\}\mathbf{d}\tau+\int\int_{\Gamma_u} G\mathbf{d}S+\int\int_{\Gamma_\sigma} H\mathbf{d}S \tag{14}$$

$$J=\int\int\int (\sigma_{ij}e_{ij}+F)\mathbf{d}\tau+\int\int_{\Gamma_u} G\mathbf{d}S+\int\int_{\Gamma_\sigma} H\mathbf{d}S \tag{15}$$

$$J=\int\int\int \{u_i(\sigma_{ij,j}+f_i)+F\}\mathbf{d}\tau+\int\int_{\Gamma_u} G\mathbf{d}S+\int\int_{\Gamma_\sigma} H\mathbf{d}S. \tag{16}$$

Generally speaking, all integrals with an energy form can be used as trial-functional, leading to the needed variational principles.

4 Hellinger-Reissner Principle

In this section we will use the semi-inverse method to rederive the well-established Hellinger-Reissner principle from different trial-functionals. If we want to establish a functional with two kinds of independent variations (u_i,σ_{ij}) under constraints of stress-strain relations (2), an energy trial-functional can be written down directly from its partial differential Eq. (1).

$$J_{HR1}(u_i,\sigma_{ij})=\int\int\int \{u_i(\sigma_{ij,j}+f_i)+F\}\mathbf{d}\tau+\int\int_{\Gamma_u} G\mathbf{d}S+\int\int_{\Gamma_\sigma} H\mathbf{d}S. \tag{16a}$$

Taking variations with respect to the two kinds of independent variations (u_i and σ_{ij}), we obtain

$$\delta_u J_{HR1}=\int\int\int \{(\sigma_{ij,j}+f_i)\,\delta u_i+\delta_u F\}\mathbf{d}\tau+\int\int_{\Gamma_u}\delta_u G\mathbf{d}S+\int\int_{\Gamma_\sigma}\delta_u H\mathbf{d}S \tag{17}$$

$$\delta_\sigma J_{HR1}=\int\int\int\left\{-\frac{1}{2}(u_{i,j}+u_{j,i})\,\delta\sigma_{ij}+\delta_\sigma F\right\}\mathbf{d}\tau+\oint\!\!\oint u_in_j\delta\sigma_{ij}\mathbf{d}S+\int\int_{\Gamma_u}\delta_\sigma G\mathbf{d}S+\int\int_{\Gamma_\sigma}\delta_\sigma H\mathbf{d}S \tag{18}$$

for all variations of δu_i and $\delta\sigma_{ij}$. We have

$$\delta u_i:\ \ \sigma_{ij,j}+f_i+F_u=0 \tag{19}$$

$$\delta\sigma_{ij}:\ -\frac{1}{2}(u_{i,j}+u_{j,i})+F_\sigma=0 \tag{20}$$

where $\delta_u F=F_u=\partial F/\partial u_i-(\partial F/\partial u_{i,j})_{,j}$ is called the functional derivative after Goldstein [11]. If F is expressed inexplicitly with $u_{i,j}$, then the functional derivative becomes partial derivative, i.e., $F_u=\partial F/\partial u_i$.

Equations (19) and (20) with unknown F are called the trial-Euler equations. We search for such an unknown F, so that the trial-Euler equations satisfy Eqs. (1) and (3). Accordingly we have

$$F_u=0 \tag{21}$$

$$F_\sigma=e_{ij}. \tag{22}$$

In view of the constraints of (2), we can identify the unknown as follows:

$$F=B(\sigma_{ij}). \tag{23}$$

The trial-functional (16*a*), therefore, can be rewritten as follows:

$$J_{HR1}(u_i,\sigma_{ij})=\int\int\int\{u_i(\sigma_{ij,j}+f_i)+B\}\mathbf{d}\tau+\int_{\Gamma_u}\int G\mathbf{d}S+\int_{\Gamma_\sigma}\int H\mathbf{d}S. \tag{24}$$

The unknowns G and H can be identified in the same way. From Eqs. (17) and (18) we obtain following trial-Euler equations.

At the boundary Γ_u

$$\delta u_i:\ G_u=0 \tag{25}$$

$$\delta\sigma_{ij}:\ u_i n_j+G_\sigma=0 \tag{26}$$

and at the boundary Γ_σ

$$\delta u_i:\ H_u=0 \tag{27}$$

$$\delta\sigma_{ij}:\ u_i n_j+H_\sigma=0. \tag{28}$$

The trial-Euler equations (25) and (26) should satisfy boundary conditions (4) or identity, accordingly we have

$$G=-\sigma_{ij}n_j\overline{u_i}. \tag{29}$$

The trial-Euler equations (27) and (28) should satisfy boundary conditions (5) or identity, we set

$$H_u=-(\sigma_{ij}n_j-\overline{p_i}) \tag{30}$$

$$H_\sigma=-u_i n_j. \tag{31}$$

We therefore obtain

$$H=-u_i(\sigma_{ij}n_j-\overline{p_i}). \tag{32}$$

Substituting (29) and (32) into (24) yields Hellinger-Reissner principle.

The trial-functional can be constructed in an arbitrary energy form, for example

$$J_{HR2}(u_i,\sigma_{ij})=\int\int\int(\sigma_{ij}u_{i,j}+F)\mathbf{d}\tau+\int_{\Gamma_u}\int G\mathbf{d}S+\int_{\Gamma_\sigma}\int H\mathbf{d}S. \tag{33}$$

The identification of unknown is the same as before. Making the above trial-functional stationary, we obtain the following trial-Euler equation.

$$\delta u_i:\ -\sigma_{ij,j}+F_u=0 \tag{34}$$

$$\delta\sigma_{ij}:\ \frac{1}{2}(u_{i,j}+u_{j,i})+F_\sigma=0 \tag{35}$$

The above trial-Euler equations should satisfy Eqs. (1) and (3) respectively, we therefore obtain

$$F_u=-f_i \tag{36}$$

$$F_\sigma=-e_{ij}. \tag{37}$$

In view of constraints of (2), the unknown F can be determined as follows

$$F=-B(\sigma_{ij})-f_i u_i. \tag{38}$$

The trial-functional (33), therefore, can be rewritten down as follows

$$J_{HR2}(u_i,\sigma_{ij})=\int\int\int(\sigma_{ij}u_{i,j}-B-f_i u_i)\mathbf{d}\tau+\int_{\Gamma_u}\int G\mathbf{d}S+\int_{\Gamma_\sigma}\int H\mathbf{d}S. \tag{39}$$

The trial-Euler equations at the boundary read
At the boundary Γ_u

$$\delta u_i:\ \sigma_{ij}n_j+G_u=0 \tag{40}$$

$$\delta\sigma_{ij}:\ G_\sigma=0 \tag{41}$$

and at the boundary Γ_σ

$$\delta u_i:\ \sigma_{ij}n_j+H_u=0 \tag{42}$$

$$\delta\sigma_{ij}:\ H_\sigma=0. \tag{43}$$

By the same manipulation as above we can identify the unknowns as follows

$$G=-\sigma_{ij}n_j(u_i-\overline{u_i}),\ \ H=-\overline{p_i}u_i. \tag{44}$$

We therefore obtain following functional

$$J_{HR2}(u_i\sigma_{ij})=\int\int\int(\sigma_{ij}u_{i,j}-B-f_i u_i)\mathbf{d}\tau-\int_{\Gamma_u}\int\sigma_{ij}n_j(u_i-\overline{u_i})\mathbf{d}S-\int_{\Gamma_\sigma}\int\overline{p_i}u_i\mathbf{d}S. \tag{45}$$

Integrating by parts yields Hellinger-Reissner principle.

We establish another trial-functional to approach Hellinger-Reissner principle, which reads

$$J_{HR3}(u_i,\sigma_{ij})=\int\int\int\{B+F\}\mathbf{d}\tau+\int_{\Gamma_u}\int G\mathbf{d}S+\int_{\Gamma_\sigma}\int H\mathbf{d}S. \tag{46}$$

One of its trial-Euler equations can be written down as follows

$$\delta\sigma_{ij}:\ \frac{\partial B}{\partial\sigma_{ij}}+F_\sigma=0 \tag{47}$$

which should satisfy Eq. (3), and by virtue of constraints (2), we have

$$F_\sigma=-e_{ij}=-\frac{1}{2}(u_{i,j}+u_{j,i}). \tag{48}$$

Note that $\sigma_{ij}u_{i,j}=\sigma_{ij}u_{j,i}$, we therefore can identify the unknown as follows

$$F=-\sigma_{ij}u_{i,j}+F_1 \tag{49a}$$

or

$$F=\sigma_{ij,j}u_i+F_1 \tag{49b}$$

where F_1 is newly introduced unknown function of u_i

$$J_{HR3}(u_i,\sigma_{ij})=\int\int\int\{B+\sigma_{ij,j}u_i+F_1\}\mathbf{d}\tau+\int_{\Gamma_u}\int G\mathbf{d}S+\int_{\Gamma_\sigma}\int H\mathbf{d}S. \tag{50}$$

We can obtain the other set of trial-Euler equations

$$\delta u_i:\ \sigma_{ij,j}+\frac{\delta F_1}{\delta u_i}=0 \tag{51}$$

which should satisfy Eq. (1), accordingly we have

$$\frac{\delta F_1}{\delta u_i}=f_i\,. \tag{52}$$

We therefore obtain

$$F_1=f_iu_i\,. \tag{53}$$

Identifying the unknowns G and H yields the Hellinger-Reissner principle.

Other energy trial-functionals, such as

$$J_{HR4}(u_i,\sigma_{ij})=\int\int\int\{u_i(\sigma_{ij,j}+f_i)+\alpha B+F\}\mathbf{d}\tau+\int\int_{\Gamma_u}G\mathbf{d}S+\int\int_{\Gamma_\sigma}H\mathbf{d}S. \tag{54}$$

$$J_{HR5}(u_i,\sigma_{ij})=\int\int\int\{u_i(\sigma_{ij,j}+\alpha f_i)+F\}\mathbf{d}\tau+\int\int_{\Gamma_u}G\mathbf{d}S+\int\int_{\Gamma_\sigma}H\mathbf{d}S \tag{55}$$

can also lead to Hellinger-Reissner Principle, where α is a constant.

5 Generalized Hellinger-Reissner Principle

In this section, we will use the semi-inverse method to establish a family of generalized Hellinger-Reissner principles with constant parameters which have been studied by Felippa. The presence of the free parameter offers an opportunity for the systematic derivation of energy-balanced finite elements that combine displacement and stress assumptions, details can be found in Felippa's Refs. [12–14].

As a first step to this goal, we construct the following trial-functional

$$J_{GHR1}(u_i,\sigma_{ij})=\int\int\int\{u_i(\sigma_{ij,j}+f_i)+\alpha A+F\}\mathbf{d}\tau+\int\int_{\Gamma_u}G\mathbf{d}S+\int\int_{\Gamma_\sigma}H\mathbf{d}S \tag{56}$$

where α is a constant.

Its trial-Euler equations can be written as follows:

$$\delta u_i:\ \sigma_{ij,j}+f_i-\alpha\left(\frac{\partial A}{\partial e_{ij}}\right)_{,j}+F_u=0\quad(\alpha\neq1) \tag{57}$$

$$\delta\sigma_{ij}:\ -\frac{1}{2}(u_{i,j}+u_{j,i})+F_\sigma=0. \tag{58}$$

By virtue of constraints (2), we can identify the unknown F as follows:

$$F=-\alpha f_iu_i+B. \tag{59}$$

The trial-functional (56), therefore, can be renewed as follows:

$$J_{GHR1}(u_i,\sigma_{ij})=\int\int\int\{u_i(\sigma_{ij,j}+f_i)+B+\alpha(A-f_iu_i)\}\mathbf{d}\tau+\int\int_{\Gamma_u}G\mathbf{d}S+\int\int_{\Gamma_\sigma}H\mathbf{d}S. \tag{60}$$

The trial-Euler equations at the boundary read as follows. At the boundary Γ_u,

$$\delta u_i:\ \alpha\frac{\partial A}{\partial e_{ij}}n_j+G_u=0 \tag{61}$$

$$\delta\sigma_{ij}:\ u_in_j+G_\sigma=0, \tag{62}$$

and at the boundary Γ_σ,

$$\delta u_i:\ \alpha\frac{\partial A}{\partial e_{ij}}n_j+H_u=0 \tag{63}$$

$$\delta\sigma_{ij}:\ u_in_j+H_\sigma=0. \tag{64}$$

By the same manipulation as above we can identify the unknowns as follows:

$$G=-\sigma_{ij}n_j\overline{u_i}-\alpha\sigma_{ij}n_j(u_i-\overline{u_i}) \tag{65}$$

$$H=-\alpha u_i\overline{p_i}-u_i(\sigma_{ij}n_j-\overline{p_i}). \tag{66}$$

We therefore obtain the following generalized Hellinger-Reissner principle:

$$J_{GHR1}(u_i,\sigma_{ij})=\int\int\int\{u_i(\sigma_{ij,j}+f_i)+B+\alpha(A-f_iu_i)\}\mathbf{d}\tau-\int\int_{\Gamma_u}[\sigma_{ij}n_j\overline{u_i}+\alpha\sigma_{ij}n_j(u_i-\overline{u_i})]\mathbf{d}S-\int\int_{\Gamma_\sigma}[\alpha u_i\overline{p_i}+u_i(\sigma_{ij}n_j-\overline{p_i})]\mathbf{d}S \tag{67}$$

which is unknown to us at the present time. For $\alpha=0$ we obtain the Hellinger-Reissner principle (7).

The trial-functional for the generalized Hellinger-Reissner principle under constraints of (3) can also be constructed as follows:

$$J_{GHR2}(u_i,\sigma_{ij})=\int\int\int\{u_i\sigma_{ij,j}+F\}\mathbf{d}\tau+\int\int_{\Gamma_u}G\mathbf{d}S+\int\int_{\Gamma_\sigma}H\mathbf{d}S. \tag{68a}$$

The term $u_i\sigma_{ij,j}$ can be expressed as follows:

$$u_i\sigma_{ij,j}=\alpha u_i\sigma_{ij,j}+(1-\alpha)u_i\sigma_{ij,j}\,.$$

Substituting it into (68*a*), then integrating by parts yields

$$J_{GHR2}(u_i,\sigma_{ij})=\int\int\int\{-\alpha\sigma_{ij,j}e_{ij}+(1-\alpha)u_i\sigma_{ij,j}+F\}\mathbf{d}\tau+\int\int_{\Gamma_u}G\mathbf{d}S+\int\int_{\Gamma_\sigma}H\mathbf{d}S. \tag{68b}$$

The trial-Euler equations of the above trial-functional (68*b*) can be expressed as follows:

$$\delta u_i:\ \alpha\sigma_{ij,j}+(1-\alpha)\sigma_{ij,j}+F_u=0 \tag{69}$$

$$\delta\sigma_{ij}:\ -\alpha e_{ij}-\frac{1}{2}(1-\alpha)(u_{i,j}+u_{j,i})+F_\sigma=0 \tag{70}$$

which should satisfy Eqs. (1) and (2). In virtue of constraints (3), we can identify the unknown F as follows:

$$F = B + f_i u_i . \tag{71}$$

The unknowns G and H can be identified in the same way, we therefore obtain the following generalized Hellinger-Reissner principle:

$$J_{GHR2}(u_i,\sigma_{ij}) = \int\int\int \{-\alpha\sigma_{ij,j}e_{ij} + (1-\alpha)u_i\sigma_{ij,j} + f_i u_i + B\}\mathbf{d}\tau + \int\int_{\Gamma_u} \sigma_{ij}n_j(du_i - \overline{u_i})\mathbf{d}S + \int\int_{\Gamma_\sigma} u_i[\overline{p_i} - (1-\alpha)\sigma_{ij}n_j]\mathbf{d}S \tag{72a}$$

$$= \int\int\int \{-\alpha(\sigma_{ij}e_{ij} + u_i\sigma_{ij,j}) + u_i(\sigma_{ij,j} + f_i) + B\}\mathbf{d}\tau + \int\int_{\Gamma_u} \sigma_{ij}n_j(du_i - \overline{u_i})\mathbf{d}S + \int\int_{\Gamma_\sigma} u_i[\overline{p_i} - (1-\alpha)\sigma_{ij}n_j]\mathbf{d}S. \tag{72b}$$

We will illustrate another way to arrive at a generalized Hellinger-Reissner principle; the trial-functional can be constructed as follows:

$$J_{GHR3}(u_i,\sigma_{ij}) = \int\int\int (A - f_i u_i + F)\mathbf{d}\tau + \int\int_{\Gamma_u} G\mathbf{d}S + \int\int_{\Gamma_\sigma} H\mathbf{d}S. \tag{73}$$

In this trial-functional, we will treat Eq. (3) as a variational constraint; that means we try to find such unknowns, so that the stationary conditions satisfy Eqs. (1), (2), (4), and (5). Its trial-Euler equations read

$$\delta u_i : \; -\left(\frac{\partial A}{\partial e_{ij}}\right)_{,j} - f_i + F_u = 0 \tag{74}$$

$$\delta\sigma_{ij} : \; F_\sigma = 0. \tag{75}$$

The trial-Euler Eq. (74) should satisfy Eq. (1), thus we have

$$\frac{\delta F}{\delta u_i} = 0. \tag{76}$$

By virtue of constraint (3), the unknown function F, which must also have the form of energy, can be written in a more general form:

$$F = \alpha(\sigma_{ij,j}u_i + A) + F_1(\sigma_{ij},\sigma_{ij,j}) \tag{77}$$

where α is a nonzero constant, and F_1 is the new introduced unknown function to be further determined.

Substituting (77) into (75) yields

$$-\alpha\frac{1}{2}(u_{i,j} + u_{j,i}) + \frac{\delta F_1}{\delta\sigma_{ij}} = 0 \tag{78}$$

which should satisfy the other set of its stationary conditions, Eq. (2), and in view of the constraint (3), we have

$$\frac{\delta F_1}{\delta\sigma_{ij}} = \alpha e_{ij} = \alpha\frac{\partial B}{\delta\sigma_{ij}}. \tag{79}$$

It is easy to identify the unknown function F_1 as follows:

$$F_1 = \alpha B. \tag{80}$$

So we obtain the following renewed trial-functional

$$J_{GHR3}(u_i,\sigma_{ij}) = \int\int\int \{A - f_i u_i + \alpha(\sigma_{ij,j}u_i + A + B)\}\mathbf{d}\tau + \int\int_{\Gamma_u} G\mathbf{d}S + \int\int_{\Gamma_\sigma} H\mathbf{d}S. \tag{81}$$

The trial-Euler equations at boundary conditions read

$$\delta u_i : (1+\alpha)\sigma_{ij}n_j + G_u = 0 \tag{82}$$

$$\delta\sigma_{ij} : \alpha u_i n_j + G_\sigma = 0 \tag{83}$$

$$\delta\sigma_{ij} : \alpha u_i n_j + H_\sigma = 0 \tag{84}$$

$$\delta u_i : (1+\alpha)\sigma_{ij}n_j + H_u = 0. \tag{85}$$

We therefore obtain

$$G = -(1+\alpha)\sigma_{ij}n_j u_i + \sigma_{ij}n_j\overline{u_i} \tag{86}$$

$$H = -\alpha\sigma_{ij}u_i n_j - u_i\overline{p_i}. \tag{87}$$

Substituting the identified G and H into the trial-functional (81), we can obtain

$$J_{GHR3}(u_i,\sigma_{ij}) = \int\int\int \{A - f_i u_i + \alpha(\sigma_{ij,j}u_i + A + B)\}\mathbf{d}\tau + \int\int_{\Gamma_u} \sigma_{ij}n_j[\overline{u_i} - (1+\alpha)u_i]\mathbf{d}S - \int\int_{\Gamma_\sigma} u_i(\overline{p_i} + \alpha\sigma_{ij}n_j)\mathbf{d}S \tag{88}$$

which is also a generalized Hellinger-Reissner principle and can also be found in Felippa's reference [12], and the Hellinger-Reissner principle can be obtained as its special case ($\alpha = -1$)

Integrating by part, and using the constraint (3), the above functional can be converted into the following form without changing its stationary conditions:

$$J^*_{GHR3}(u_i,\sigma_{ij}) = \int\int\int \{A - f_i u_i + \alpha(A + B - \sigma_{ij}e_{ij})\}\mathbf{d}\tau + \int\int_{\Gamma_u} \sigma_{ij}n_j(\overline{u_i} - u)\mathbf{d}S + \int\int_{\Gamma_\sigma} u_i\overline{p_i}\mathbf{d}S. \tag{89}$$

6 Generalized Variational Principles

Now we try to establish a generalized variational principle from the Hellinger-Reissner principle via the Lagrange multiplier method. We have

$$J_{GVP}(\sigma_{ij},e_{ij},u_i,\lambda_{ij}) = J_{HR}(\sigma_{ij},u_i) + \int\int\int \lambda_{ij}\left(e_{ij} - \frac{\partial B}{\partial\sigma_{ij}}\right)d\tau \tag{90}$$

where λ_{ij} are Lagrange multipliers to be determined.

Taking variations with respect to the independent variations e_{ij}, we obtain

$$\delta_e J_{GVP} = \int\int\int \lambda_{ij}\delta e_{ij}\mathbf{d}\tau = 0. \tag{91}$$

We therefore obtain

$$\lambda_{ij} = 0. \tag{92}$$

The phenomena is call the variational crisis after Chien.

It is quite convenient to remove the variational crisis via the semi-inverse method. The trial-functional for the problem reads

$$J_{GVP}(\sigma_{ij},e_{ij},u_i)=\int\int\int (B+\sigma_{ij,j}u_i+f_iu_i+F)\mathbf{d}\tau - \int_{\Gamma_u}\int \sigma_{ij}n_j\overline{u_i}\mathbf{d}S-\int_{\Gamma_\sigma}\int u_i(\sigma_{ij}n_j-\overline{p_i})\mathbf{d}S. \tag{93}$$

Its trial-Euler equations can be expressed as follows:

$$\delta u_i:\sigma_{ij,j}+f_i+Fu=0 \tag{94}$$

$$\delta\sigma_{ij}:\frac{\partial B}{\partial\sigma_{ij}}-\frac{1}{2}(u_{i,j}+u_{j,i})+F_\sigma=0 \tag{95}$$

$$\delta e_{ij}:F_e=0. \tag{96}$$

The above three kinds of trial-Euler equations should satisfy Eqs. (1)–(3), remember that the unknown F has the form of energy, we set

$$F_u=0 \tag{97}$$

$$F_\sigma=\alpha\left(e_{ij}-\frac{\partial B}{\partial\sigma_{ij}}\right) \tag{98}$$

$$F_e=\alpha\left(\sigma_{ij}-\frac{\partial A}{\partial e_{ij}}\right). \tag{99}$$

We therefore can identify the unknown F as follows:

$$F=\alpha(A+B-\sigma_{ij}e_{ij}). \tag{100}$$

Thus we obtain following generalized variational principle without any constraints.

$$J_{GVP}(u_i,e_{ij},\sigma_{ij})=J_{HR}(u_i,\sigma_{ij})+\alpha\int\int\int (A+B-\sigma_{ij}e_{ij})\mathbf{d}\tau \tag{101}$$

which is actually Chien's generalized variational principle [9].

7 Conclusion

In the paper, we have proved that the well-known Hellinger-Reissner principle can also be regarded as a principle under constraints of strain-displacement, the Hellinger-Reissner principle, and a family of generalized Hellinger-Reissner principles can be readily obtained via the semi-inverse method. By this method we can also obtain the Hu-Washizu principle, where details will be discussed in another paper.

Acknowledgments

The author thanks Prof. W. K. Liu and an anonymous referee for their valuable comments and suggestions and providing Refs. [11–14]. The work is supported by National Key Basic Research Special Fund (No. G1998020318).

References

[1] Chien, W. Z., 1984, ''Generalized Variational Principle in Elasticity,'' Eng. Mech. Civil Eng., **24**, pp. 93–153.

[2] Washizu, K., 1982, *Variational Methods in Elasticity and Plasticity*, Pergamon Press, Oxford.

[3] He, J. H., 1997, Involutory Transformation Without Using Lagrange Multipliers and Its Applications to Establishing Variational Principles With Multi-Variables in Thin Plate Bending Problems, *Modern Mechanics and Advances in Science and Technology*, F. G. Zhuang, ed., Qinghua University Press, Beijing, pp. 1417–1418 (in Chinese).

[4] Liu, G. L., 1990, ''A System Approach to the Search and Transformation for Variational Principles in Fluid Mechanics With Emphasis on Inverse and Hybrid Problems,'' *Proc. 1st Int. Symp. Aerotheromdynamics of Internal Flow*, pp. 128–135.

[5] He, J. H., 1997, ''Semi-inverse Method of Establishing Generalized Variational Principles for Fluid Mechanics With Emphasis on Turbomachinery Aerodynamics,'' Int. J. Turbo Jet Eng., **14**, pp. 23–28.

[6] He, J. H., 1997, ''Semi-Inverse Method: A New Approach to Establishing Variational Principles in Fluid Mechanics,'' J. Eng. Thermophys., **18**, No. 4, pp. 440–444 (in Chinese).

[7] He, J. H. 1997, ''Equivalent Theorem of Hellinger-Reissner and Hu-Washizu Principles,'' *J. Shanghai University*, **1**, No. 1, pp. 36–41 (English edition).

[8] He, J. H., 1997, ''On C. C. Lin's Constraints,'' *Modern Mechanics and Advances in Science and Technology*, F. G. Zhuang, ed., Qinghua University Press, Beijing, pp. 603–604.

[9] Chien, W. Z., 1983, ''Method of High-Order Lagrange Multiplier and Generalized Variational Principles of Elasticity With More General Forms of Functionals,'' Appl. Math. Mech. (in Chinese), **4**, No. 2, pp. 137–150.

[10] He, J. H., 1997, ''Modified Lagrange Multiplier Method and Generalized Variational Principles in Fluid Mechanics,'' *J. Shanghai University*, 1, No. 2 (English edition).

[11] Goldstein H., 1981, *Classical Mechanics*, 2nd Ed; Addision-Wesley, Reading PA.

[12] Felippa, C., 1989, ''Parametrized Multifield Variational Principles in Elasticity,'' Commun. Appl. Num. Eng., **5**, pp. 79–99.

[13] Felippa, C., 1994, ''A Survey of Parametrized Variational Analysis and Applications to Computational Mechanics,'' Comput. Methods Appl. Mech. Eng., **113**, pp. 109–140.

[14] Felippa, C., 1996, ''Recent Developments in Parametrized Variational Principles for Mechanics,'' Comput. Mech., **18**, pp. 159–174.

Behavior of a Rubber Spring Pendulum

R. Bhattacharyya
Associate Professor,
Department of Mechanical Engineering,
Indian Institute of Technology,
Kharagpur 721302, India
Mem. ASME

The stability of motion of a nonlinear neo-Hookean rubber spring pendulum under a special type of support oscillation is studied. The small swing motion is described by a Mathieu-Hill equation, corresponding stability curves for which are generated in a relevant parametric plane with a stability criterion obtained earlier. Autoparametric resonance in the special case of linearized motions is found to occur, as usual.
[S0021-8936(00)00801-1]

1 Introduction

It is known that a linear elastic pendulum, with certain parametric adjustments, exhibits autoparametric resonance. Although a simple mechanical system, study of the mathematical model of such a pendulum can explain rather complex behavior of systems in various fields of physics and engineering.

It has been shown earlier by Minorsky [1] and later by Olsson [2] that the autoparametric resonance occurs when the linearized natural frequency of the axial mode becomes equal to twice that of the pendulum mode. Physically, in this situation a periodic energy exchange takes place as a result of nonlinear resonant coupling between the two modes of oscillation. Several other papers [3–5] are devoted to the study of such a system.

Ryland and Meirovitch [6] with a novel solution method for Hill's equation, discussed the behavior of a spring pendulum with oscillating support. Studies on chaotic motion of this system have also appeared in the papers by Nunez-Yepez et al. [7] and Cuerno et al. [8]. In all these expositions the spring of the elastic pendulum is considered to be linear and the nonlinearities considered therein are purely geometrical.

In many engineering applications, structural components made of rubber-like materials are used to sustain uniaxial, torsional, and shearing vibrations. For example, these are employed as suspension springs in automobiles, pumps, machines, etc. Reinforced rubber-like materials are used as drive shafts in compact vehicles. Thus, it may be of no surprise to expect the possibility of autoparametric resonance in such cases in a way similar to the case of linear spring pendulum. In this paper therefore, we shall consider a nonlinear spring pendulum in which a rubber rod made of neo-Hookean material possessing material nonlinearity supports the pendulum bob and analyze the stability of swinging motion.

2 Formulation

The schematic of the nonlinear elastic pendulum is shown in Fig. 1 (inset). The pendulum is considered to be a particle of mass m and suspended by a neo-Hookean rubber rod of undeformed length l_0 and cross-sectional area A_0. The mass of the rod is considered to be much smaller than that of the pendulum bob. It is assumed that the motion of this nonlinear elastic pendulum is entirely confined to the plane of Fig. 1. The displacement of the support point is denoted by $Y(t)$. Let the uniform axial stretch λ and static equilibrium stretch λ_s be defined in terms of the static displacement x_{st}, l_0, and the instantaneous radial displacement x by

$$\lambda=\frac{l_0+x_{st}+x}{l_0}>0 \quad \text{and} \quad \lambda_s=\frac{l_0+x_{st}}{l_0}>0.$$

In a neo-Hookean rubber rod, the uniform axial force $T(\lambda)$ required to produce the stretch λ is determined by

$$T(\lambda)=A_0G(\lambda-\lambda^{-2}),$$

in which G is the shear modulus at the natural undeformed state.

The equations of motion for the pendulum bob may be obtained with the help of the axial force expression given above in the following form:

$$\ddot{\lambda}-\lambda\dot{\theta}^2+H(\lambda-1/\lambda^2)=(\omega_p^2+\ddot{\bar{Y}})\cos\theta, \tag{2.1}$$

and

$$\lambda\ddot{\theta}+2\dot{\lambda}\dot{\theta}+(\omega_p^2+\ddot{\bar{Y}})\sin\theta=0, \tag{2.2}$$

where

$$H=A_0G/ml_0, \quad \omega_p^2=g/l_0, \quad \text{and} \quad \bar{Y}=Y/l_0.$$

Equations (2.1) and (2.2) are nonlinear and coupled, exact solutions of which are difficult to obtain for arbitrary support motion. However, for small lateral motion of the pendulum, the axial response that is proportional to the forcing function can be found analytically. Thus, in (2.1) and (2.2) we assume

$$\ddot{\bar{Y}}=BH\lambda, \tag{2.3}$$

where B is a real constant. The equations of motion (2.1) and (2.2) may now be rewritten for this type of support motion as

$$\ddot{\lambda}-\lambda\dot{\theta}^2+H\left[(1-B\cos\theta)\lambda-\frac{1}{\lambda^2}\right]=\omega_p^2\cos\theta, \tag{2.4}$$

and

$$\lambda\ddot{\theta}+2\dot{\lambda}\dot{\theta}+(\omega_p^2+BH\lambda)\sin\theta=0. \tag{2.5}$$

The static equilibrium axial stretch and angular deflection can be obtained from (2.4) and (2.5) as the time-independent solutions. Thus, from (2.5) $\theta=\theta_s=0$ (hanging position), π (inverted position). Consequently, from (2.4) the axial static stretch turns out to be the solution of

$$(1-B)\lambda_s^3-(-1)^c\frac{\lambda_s^2}{2\kappa}-1=0. \tag{2.6}$$

Note that $c=0$ when $\theta_s=0$ and $c=1$ when $\theta_s=\pi$. The nondimensional stiffness ratio, $\kappa=A_0G/2$ mg. We shall consider here only the hanging pendulum case ($c=0$).

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 26, 1997; final revision, Apr. 7, 1998. Associate Technical Editor: M. M. Carroll.

Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

For $-\infty<B<1$, only one positive real root of (2.6) exists and in this range of B, as $\kappa\to\infty$, $\lambda_s\to 1/(1-B)^{1/3}$ $(=\lambda_s^*$, say). Evidently, for $B<0$, $\lambda_s^*<1$. In fact, for $\lambda=\lambda_s^*$, Eq. (2.6) shows that there exists a $\kappa^*=-1/2B$ such that for all $\kappa\geqslant\kappa^*$, $\lambda_s\leqslant 1$. This shows that for $B<0$, in a certain range of values of κ we can expect an inverted pendulum-like behavior. Figure 1 shows the variation of static equilibrium stretch with the stiffness ratio for various values of B. For $B<0$, static equilibrium values less than 1 may be noted. Also, as B becomes more negative κ^* reduces. For $B\geqslant 1$, it may be argued that (2.6) does not yield a real positive value of the static stretch. Thus, in this discussion only the values of $B<1$ are considered.

We shall now resume our discussion on the solution of the equations of motion. The following two special cases are considered for the closed-form solutions; (i) small axial and swinging motions and (ii) finite axial and small swinging motions. The case of finite swinging motion superimposed on static stretch may be analyzed in a way similar to that given in Nayfeh and Mook [9].

3 Small Oscillations Superimposed on a Finite Static Stretch

In this case, with $\lambda=\lambda_s+z$, $|z|=|x/l_0|\ll 1$, and for small θ and $\dot\theta$, Eqs. (2.4) and (2.5) reduce to

$$\ddot z+H\left(1-B+\frac{2}{\lambda_s^3}\right)z=0, \tag{3.1}$$

and

$$\lambda_s\ddot\theta+(\omega_p^2+BH\lambda_s)\theta=0. \tag{3.2}$$

Equations (3.1) and (3.2) show two completely decoupled linear harmonic oscillators with axial and swing mode nondimensional natural frequencies $\omega_a=\sqrt{H}(1-B+2/\lambda_s^3)^{1/2}$, and $\omega_s=(\omega_p^2/\lambda_s+BH)^{1/2}$, respectively. The static equilibrium Eq. (2.6) may now be used to remove the parameter $H(=2\kappa\omega_p^2)$ from these frequency expressions yielding *universal frequency formulas independent of the shear modulus, G of the neo-Hookean rod*. Thus,

$$\frac{\omega_a}{\omega_p}=\frac{1}{\sqrt{\lambda_s}}\left[\frac{(1-B)\lambda_s^3+2}{(1-B)\lambda_s^3-1}\right]^{1/2}\quad\text{and}$$

$$\frac{\omega_s}{\omega_p}=\frac{1}{\sqrt{\lambda_s}}\left[\frac{\lambda_s^3-1}{(1-B)\lambda_s^3-1}\right]^{1/2}. \tag{3.3}$$

The result $(3.3)_1$ shows that ω_a/ω_p grows indefinitely large as $\lambda_s\to 1/(1-B)^{1/3}$ $(=\lambda_s^*)$ and it approaches zero irrespective of the values of B as $\lambda_s\to\infty$. It may be mentioned here that for $B=0$, $(3.3)_1$ yields the result obtained and presented graphically by Beatty ([10], Fig. 1).

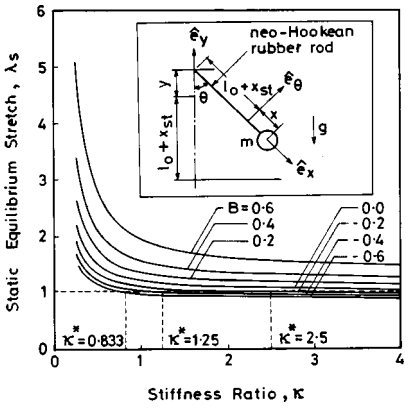


Fig. 1 Static equilibrium stretch versus stiffness ratio for various values of *B* in the suspended pendulum case. Inset shows system schema.

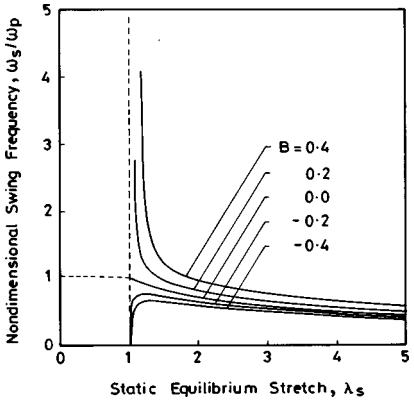


Fig. 2 Linearized swing mode natural frequency versus static equilibrium stretch for various values of *B*

The universal formula $(3.3)_2$ gives a new result, however. The graph of small swing frequency is shown in Fig. 2 for various values of B. It is clear from $(3.3)_2$ that $\omega_s/\omega_p\to 0$ as $\lambda_s\to\infty$ for all values of $B<1$. For $B>0$, as $\lambda_s\to\lambda_s^*$, $\omega_s/\omega_p\to\infty$. However, for $B<0$ the curves originate from $\omega_s/\omega_p=0$ at $\lambda_s=1$ and increase monotonically to their respective maximum values, approaching zero as $\lambda_s\to\infty$, as shown in Fig. 2. Notice that for $\lambda_s<1$, no real values of the frequency exist. Hence the asymptotic behavior will not be present since $\lambda_s^*<1$. Also for a fixed value of λ_s the swing frequency increases with the increase of B from negative to positive values. Finally, it follows from $(3.3)_2$ that for $B=0$, at $\lambda_s=1$, $\omega_s/\omega_p=1$.

4 Small Swinging Motion and Finite Amplitude Axial Vibration

We consider the case where θ and $\dot\theta$ are small so that the equations of motion (2.4) and (2.5) take the following form:

$$\ddot\lambda+H\left[(1-B)\lambda-\frac{1}{\lambda^2}\right]=\omega_p^2, \tag{4.1}$$

$$\lambda\ddot\theta+2\dot\lambda\dot\theta+(\omega_p^2+BH\lambda)\theta=0. \tag{4.2}$$

Notice here that the swing mode is influenced by the stretching oscillation.

The closed-form analytical solution of (4.1) may now be obtained by following Beatty [10] and Beatty and Bhattacharyya [11]. We shall briefly describe the procedure here.

It may be shown from the first integral of (4.1) that the time taken by the pendulum to move axially from the initial state $(\lambda_0,\dot\lambda_0)$ at time $t=0$ to a state $(\lambda,\dot\lambda)$ at time $t=t$ is

$$t=\pm\frac{1}{[H(1-B)]^{1/2}}\int_{\lambda_0}^{\lambda}\frac{\lambda\,d\lambda}{[-\lambda Q(\lambda)]^{1/2}}, \tag{4.3}$$

where

$$Q(\lambda)=\lambda^3-\left[\frac{\lambda^2}{\kappa}+\frac{E\lambda}{H}-2\right]\frac{1}{(1-B)}=(\lambda-\alpha)(\lambda-\beta)(\lambda+\gamma),\qquad B<1. \tag{4.4}$$

In (4.4),

$$E(\lambda_0,\dot\lambda_0)=\dot\lambda_0^2+H(1-B)\lambda_0^2+2H/\lambda_0-2\omega_p^2\lambda_0$$

is the energy constant of the uncoupled axial mode only. The real roots of $Q(\lambda)$ in (4.4) are denoted by α, β, and $-\gamma$ of which the

positive real roots α *and* β are the two extreme stretches during the motion such that $0<\alpha\leqslant\lambda(t)\leqslant\beta$ and $-\gamma$ is the negative real root of no physical significance.

The sign in (4.3) must be chosen appropriately consistent with the initial conditions. From (4.4), the following relationships hold:

$$\alpha+\beta-\gamma=\frac{1}{\kappa(1-B)},\quad \alpha\beta\gamma=\frac{2}{(1-B)},$$

which may be utilized to express β and γ in terms of α as

$$\beta=\frac{-\eta+[\eta^2+8(1-B)/\alpha]^{1/2}}{2(1-B)}\quad\text{and}\quad\gamma=\frac{2}{(1-B)\alpha\beta},\tag{4.5}$$

with $\eta=(1-B)\alpha-\kappa^{-1}$. Use of the substitution

$$\gamma=\frac{\alpha}{1+n\sin^2\phi},\quad n=\frac{\alpha-\beta}{\beta},\quad 0\leqslant\phi\leqslant\pi/2,\tag{4.6}$$

and the energy conservation equation $E(\lambda_0,\dot{\lambda}_0)=E(\beta,0)=E(\alpha,0)$ along with (4.4) and (4.5) in (4.3) yields the travel time t as

$$t=\pm\frac{2J_0}{\sqrt{H}}\Pi(\phi;n,k)=\pm\frac{\pi}{\sqrt{H}}\Lambda(\phi;n,k),\tag{4.7}$$

where $\Lambda(\phi;n,k)$ is the Heuman Lambda function, and $\Pi(\phi;n,k)$ is the incomplete elliptic integral of third kind given by

$$\Pi(\phi;n,k)=\int_0^{\phi}\frac{d\chi}{(1+n\sin^2\chi)(1-k^2\sin^2\chi)^{1/2}}.\tag{4.8}$$

Various parameters used above are defined as

$$J_0=\frac{\alpha}{[(1-B)\beta(\alpha+\gamma)]^{1/2}},\quad k=\left[\frac{-\gamma n}{\alpha+\gamma}\right]^{1/2}.\tag{4.9}$$

Since $B<1$, clearly $0<k<1$, and $0<k^2<-n<1$.

The periodic time for finite amplitude uniaxial oscillation is obtained from (4.7) as

$$T_p=\frac{2\pi}{\sqrt{H}}\Lambda(\pi/2;n,k)=\frac{2\pi}{\sqrt{H}}\Lambda_0(\xi;k)$$
$$\text{with }\xi=\sin^{-1}\left[\frac{1+k^2/n}{1-k^2}\right]^{1/2},\tag{4.10}$$

wherein $\Lambda_0(\xi;k)$ is the tabulated Heuman's Λ_0 function. This completes the solution of (4.1).

Next we shall discuss the stability of the solutions of (4.2). To this end, using (4.7) and (4.8) to change the independent variable from t to ϕ we find that (4.2) may be rewritten as

$$\frac{d^2\theta}{d\phi^2}+2P(\phi)\frac{d\theta}{d\phi}+R(\phi)\theta=0,\tag{4.11}$$

wherein,

$$P(\phi)=\frac{1}{\lambda}\left[\frac{d\Lambda}{d\phi}-\frac{\lambda d^2\Lambda/d\phi^2}{2d\Lambda/d\phi}\right]=\frac{(nk^2\sin^2\phi-2n-k^2)\sin 2\phi}{4(1+n\sin^2\phi)(1-k^2\sin^2\phi)},\tag{4.12}$$

and

$$R(\phi)=\frac{\pi^2}{\lambda}\left[\frac{d\Lambda}{d\phi}\right]^2\left[\frac{1}{2\kappa}+B\lambda\right]$$
$$=\frac{2J_0^2(1+n\sin^2\phi+2B\kappa\alpha)}{\kappa\alpha(1+n\sin^2\phi)^2(1-k^2\sin^2\phi)}.\tag{4.13}$$

Finally, with

$$\theta=y\exp\left(-\int^{\phi}P(\chi)d\chi\right)$$

(4.11) becomes

$$\frac{d^2y}{d\phi^2}+F(\phi)y=0,\tag{4.14}$$

in which

$$F(\phi)=R(\phi)-P^2(\phi)-\frac{dP(\phi)}{d\phi}.\tag{4.15}$$

By inspection of (4.12), (4.13), and (4.15), it may be shown that $F(\phi)$ is an even, π-periodic function of ϕ. Thus, (4.14) turns out to be the Mathieu-Hill equation, in regard to which we shall discuss the stability of the solutions.

The Mathieu-Hill equation has been studied in great detail (see [12]) earlier. However, the stability criterion obtained by Bhattacharyya [13] for this equation will be applied here to study the stability of small swinging oscillation superimposed on arbitrary uniaxial motion of the pendulum. Other examples that use this criterion may be found in Zhou [14] and Beatty and Bhattacharyya [11]. The statement of the criterion will now be provided.

Let $\Delta(0)\neq0$ and $f_0\neq0$ be the absolutely convergent Hill's infinite determinant for zero Floquet exponent ($\zeta=0$) and the constant term of the absolutely convergent even cosine Fourier series expansion of $F(\phi)$ in (4.14), respectively. Also, let $\Delta(0)$, ϕ, and f_0 be real-valued so that $f_0\neq4m^2$, $m=0,1,2,3,\ldots$. Then, we have the following stability criterion:

> if $\Delta(0)$ and f_0 have opposite sign, then the Floquet exponent, ζ is a nonzero real-valued quantity; and hence the solution of (4.14) is unstable. If $\Delta(0)$ and f_0 have the same sign and are real-valued the solution is stable if and only if
>
> $$|\sin(i\pi\zeta/2)|=|\sqrt{\Delta(0)}\sin(\pi\sqrt{f_0}/2)|\leqslant1.\tag{4.16}$$

In view of the developments presented earlier in this paper, it is clear that the values of both $\Delta(0)$ and f_0 are determined completely by three parameters—the stiffness ratio κ, the parameter B, and the smaller extreme stretch α. Thus, (4.16) may be used to generate a stability map in $\kappa\alpha$-plane for a given value of B. A brief discussion on the method of stability analysis will now follow.

4.1 Stability Analysis. The numerical method adopted here for obtaining the required stability maps in the parametric plane is very similar to that described in [11]. The requirements for the success of this method are that the Fourier series for $F(\phi)$ and the Hill's infinite determinant for all combinations of the system parameters under consideration must converge absolutely. In order to guarantee this, the Fourier coefficients and the infinite determinants are calculated every time with an accuracy of three significant digits of decimal. In most of the cases, this level of accuracy is obtained within first 9 to 23 Fourier coefficients and determinant values computed using (9×9) to (23×23) central rows and columns of $\Delta(0)$. For a given value B, and for each value of κ, numerically the critical value of $\alpha=\alpha_c$ is found for which the equality in (4.16) holds. This critical value of the extreme stretch denotes the boundary between the stable and unstable regions in the $\kappa\alpha_c$-plane for a particular value of B, the boundary points in this case being the stable states of impending instability. Note that since $E(\alpha,0)=E(\beta,0)$, in the presentation of the results we have used the greater extreme stretch (β_c) instead of the smaller one (α_c).

For $B>0$, however, f_0 and $\Delta(0)$ turn out to have opposite signs in certain range of values of κ and β. In such cases, the above process is bypassed to conclude readily that the small pendulum motion is unstable. Of course, for a given value of B and κ, to achieve sufficient accuracy in the result, few simple iterations have been performed to find the bounding values of the range of the extreme stretch within which the small swing becomes unstable due to the opposite sign, as suggested by the stability cri-

terion. These boundary points, being themselves unstable, are then plotted in the $\kappa\beta_c$-plane to identify the stable and unstable regions for each value of B.

In the presentation of stability diagrams in $\kappa\beta_c$-plane, different regions are identified as I and II to indicate occurrence of same and opposite signs of $\Delta(0)$ and f_0, respectively. Notice that region II is always unstable, whereas, I is stable if and only if (4.16) is satisfied.

4.2 Results for Fixed Support. It is clear in this case that $B=0$. Results obtained from the numerical method are plotted for this case to obtain the stability curve in the $\kappa\beta_c$-plane, as shown in Fig. 3 which divides the entire plane into unstable and stable regions. Also the static equilibrium stretch, λ_s is plotted in the same graph using the positive real root of (2.6). The shaded portion indicates the instability of infinitesimal swinging motion. In the stable region, the small amplitude swing remains bounded for all $\beta\in(\lambda_s,\beta_c]$ for a given value of κ.

It may be noticed that as κ increases from 0.794 to 3.5, the critical value of β denoted by β_c increases indicating a wider stable region. Thus, in this range of values of κ, even for a sufficiently large value of $\beta(\lambda_s<\beta<\beta_c)$ the small amplitude swinging oscillation of the pendulum remains bounded. Of course, the pendulum motion will not be simple harmonic. It is simple harmonic if β is sufficiently close to λ_s. To verify the foregoing conclusions, the time response simulations of Eqs. (4.1) and (4.2) have been performed for various parameter combinations falling in the stable and unstable regions of the stability curves. Figure 4 shows such simulation results for the stable $(\kappa,\beta=2,1.4)$ and unstable $(\kappa,\beta=0.8,1.4)$ swing motions, respectively. These figures clearly support the conclusions drawn from Fig. 3.

Results for free fall motion of the bob also appear in the same figure. In this case, for $\alpha=1.0$ the relationship between β and κ is plotted using $(4.5)_1$ shown by the dashed line. By inspection, it is clear that the free fall motion is predicted to be unstable within the approximate range $0.5<\kappa<1.0$.

Beyond $\kappa=3.5$ no finite values of β_c are found numerically. Thus, the small swing is stable for values of $\kappa>3.5$. But simulations show chaos-like behavior for finite values of β in this region. We shall, however, exclude this part from the present treatment.

Similarly, for $\kappa<0.794$ the region bounded by the stability curve and the static equilibrium curve guarantees stable swinging motion of the pendulum. The peculiarity of the stability curve at $\kappa=0.794$ will be discussed in the context of support motion analysis presented below.

It may be observed from Eq. (4.1) that the special support motion treated here amounts to the case of free oscillation of an equivalent nonlinear spring and a mass system. Hence this case can be treated for the determination of stability curves in the required parametric space in a similar manner.

4.3 Results for Oscillating Support. For $B\neq0$, results are obtained for various values of $B<1$ following the same method described in Subsection 4.1. Some of the results for $B\in[-0.6,0.8]$ are presented showing the distinctive features due to the effect of material nonlinearity.

Figures 5–9 show the stability maps in $\kappa\beta_c$-plane for various values of B within the given range along with the corresponding

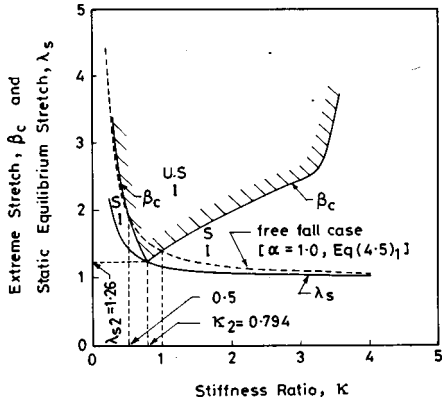


Fig. 3 Stability curves in $\kappa\beta_c$–plane for fixed support ($B=0$), obtained numerically from (4.16)

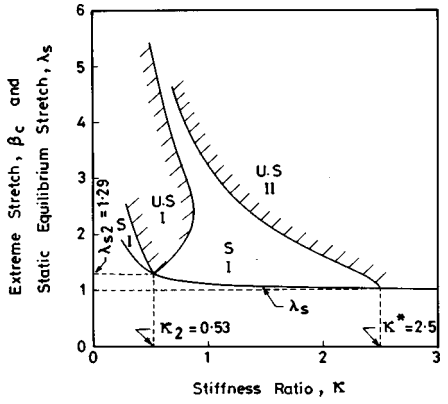


Fig. 5 Stability curves for the special type of support motion with $B=-0.2$

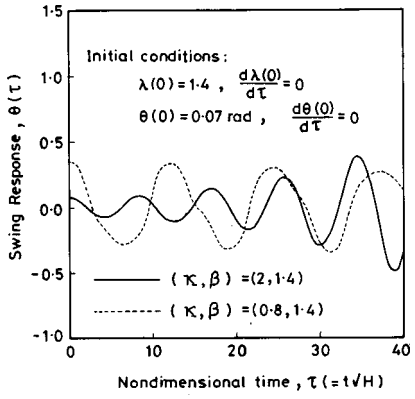


Fig. 4 Swing response curves obtained from Eqs. (4.1) and (4.2) for fixed support case ($B=0$)

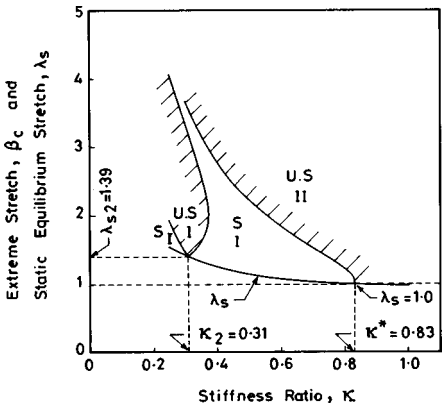


Fig. 6 Stability curves for the special type of support motion with $B=-0.6$

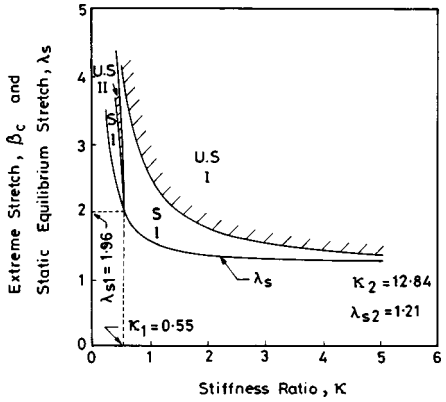


Fig. 7 Stability curves for the special type of support motion with $B=0.4$

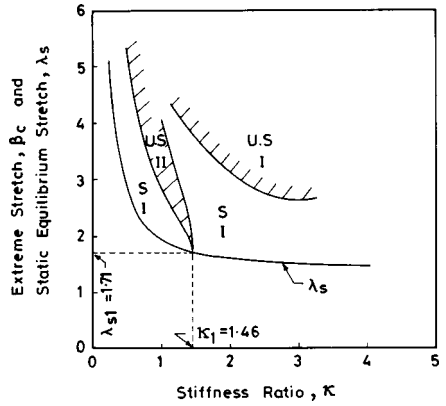


Fig. 8 Stability curves for the special type of support motion with $B=0.6$

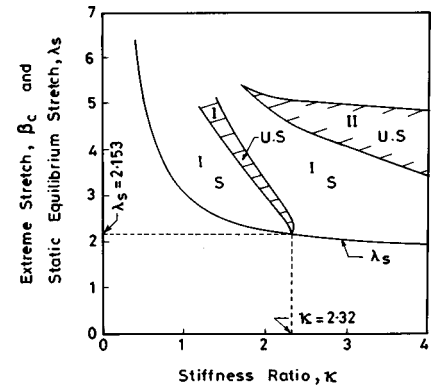


Fig. 9 Stability curves for the special type of support motion with $B=0.8$

static equilibrium curves. In all these cases, the instability of small swing is depicted by the shaded region. Notice that the stability curves for $B<0$ are to some extent different from those for $B>0$. In Figs. 3 and 5–7, the boundary of the unstable region I touches the corresponding static equilibrium curve at the point denoted by (κ_2,λ_{s2}). It is observed that with the increase of B from -0.6 to 0.4 through zero, κ_2 increases from 0.31 to 12.84. In addition, for $B>0$, Figs. 7 and 8 show similar touching points denoted by (κ_1,λ_{s1}) for the unstable region II. It is also found from stability analysis and shown by figures above that as B increases through positive values to less than 0.75, κ_1 also increases. We shall henceforth identify these two type of points as the ''touch-down'' points for future reference.

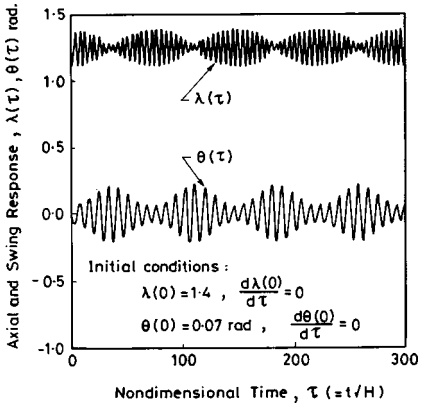


Fig. 10 Axial and swing response curves obtained from the solution of Eqs. (2.4) and (2.5) for fixed support case ($B=0$) with $(\kappa,\beta)=(0.8,1.4)$. This point lies in the unstable region in Fig. 3.

Examination of all the stability curves including that for $B=0$ shows that a point in $\kappa\beta_c$-plane lying in the stable (unstable) region for one particular value of B may become unstable (stable) for another. For example, comparison of Figs. 6 and 8 shows that for $B=0.6$ excepting a small region about $\kappa=1.46$, small swing that is stable in the range $\kappa\geqslant 0.83$ becomes unstable for $B=-0.6$ in the same range.

Recalling the case of small oscillations superimposed on a finite static stretch we see from (3.1) and (3.2) that

$$\frac{\omega_a^2}{\omega_s^2}=\frac{2\kappa\lambda_s(1-B+2/\lambda_s^3)}{1+2B\kappa\lambda_s}\equiv j^2, \tag{4.17}$$

where $j>0$ is a rational number. Solving for λ_s and κ from (4.17) and the equilibrium Eq. (2.6) we obtain the following roots:

$$\lambda_{sj}=\left[\frac{j^2+2}{j^2+B-1}\right]^{1/3} \quad \text{and} \quad \kappa_j=\frac{(j^2+2)^{2/3}(j^2+B-1)^{1/3}}{2(3-3B-Bj^2)}. \tag{4.18}$$

Since both λ_{sj} and κ_j are positive, it follows from (4.18) that the inequality

$$(1-j^2)\leqslant B<3/(3+j^2) \tag{4.19}$$

must hold. For $j=2$, these roots are found to be $\kappa_2=1.651(3+B)^{1/3}/(3-7B)$, and $\lambda_{s2}=[6/(3+B)]^{1/3}$ which for $B=0$ yields 0.794 and 1.26, respectively. In Fig. 3 this point appears as the touch-down point of the boundary of the unstable region I. For arbitrary values of the larger extreme stretch, $\beta>\lambda_{s2}(=\beta_c=1.26)$ at $\kappa_2(=0.794)$, as suggested by the stability curve, the infinitesimal swing is unstable. In particular, even if β is sufficiently close to $\lambda_{s2}(=1.26)$ for infinitesimal axial motion, the swing still remains unstable. This observation is by no means surprising, as shown by Olsson [2].

For nonzero values of B, (λ_{s2},κ_2) can be found similarly and compared with the touch-down points in Figs. 5–7. Thus, the present method of analysis convincingly shows that for a linearized spring pendulum, autoparametric resonance occurs when the linearized axial natural frequency becomes twice that of the swinging motion, i.e., when in (4.17) $j=2$. It is known that under such situation, physically a strong energy exchange between the two modes of oscillation will be noticed. The inequality (4.19) for $j=2$ gives $-3\leqslant B<0.429$. It is thus clear why the touch-down point for region I does not occur for $B=0.6$ in Fig. 8.

It also follows from Fig. 3 and Figs. 5–7 that for the nonlinear case, in general, such strong energy exchange will be noticeable when a representative point lies inside the unstable region I. Of course, to observe this in simulation of time response one has to

solve numerically the original Eqs. (2.4) and (2.5). One example is shown in Fig. 10 for the point $(B,\kappa,\beta)=(0,0.8,1.4)$ which lies in the unstable region of Fig. 3.

To explain the existence of the other type of touch-down points, substitute $j=1$ in (4.18) to obtain $\kappa_1=1.04B^{1/3}/(3-4B)$, and $\lambda_{s1}=(3/B)^{1/3}$. In this case, (4.19) translates into $0\leqslant B<0.75$. For the values of B satisfying this inequality, numerical values of κ_1 and λ_{s1} calculated from the expressions given above for $j=1$ are found to match with the touch-down points shown in Figs. 7 and 8 for $B=0.4$ and 0.6, respectively. It is also clear that the touch-down points corresponding to $j=1$ case will not occur for $B<0$ case (see Figs. 5 and 6). In passing, we note that theoretically this type of touch-down point also occurs for $B=0$ at $(\kappa_1,\lambda_{s1}=0,\infty)$; hence could not be shown in Fig. 3. Consequently, we arrive at an interesting result that autoparametric resonance also exists for infinitesimally small motions when the linearized natural frequencies of the axial and the swing modes are equal to each other. Previous investigations on the linear spring pendulum indicate the existence of such parametric instability. However, as concluded [5], in our case also this instability is of much less significance compared with the $j=2$ case for the linearized axial and swing motions. In general, however, for points lying within region II, small swing superimposed on finite axial motion is unstable denoting a strong energy exchange behavior between the modes.

The stability curves in Fig. 9 for $B=0.8$ again shows a touch-down point for the unstable region I at $\kappa=2.32$, $\lambda_s=2.153$. It turns out that use of $j=0.445$ in (4.18) yields, approximately, the values of κ and λ_s given above. Also, (4.19) yields $0.56\leqslant B<0.871$. As found earlier for the $j=1$ case, the autoparametric resonance is also of little importance for $j=0.445$. For other values of B in the range $0.75\leqslant B<1$ similar results are obtained. However, as B gets very close to 1, the range of values of λ_s becomes much larger than the value 2.5. It is known that for such large values of the stretch the tension-stretch relation corresponding to the neo-Hookean material is no longer valid. In fact, in all the results obtained so far, values of stretch greater than 2.5 hardly have any physical relevance. Hence the case $B\rightarrow 1$ is excluded from this analysis.

In the stability maps for each $B<0$, shown in Figs. 5 and 6, it is observed that the boundary for unstable region II ends abruptly at the point $(\kappa,\lambda_s)=(\kappa^*,1.0)$ identified in Fig. 1. Thus, for a given value of $B<0$ and for $\kappa\geqslant\kappa^*$, it turns out from these stability maps that infinitesimal swing is unstable for arbitrary values of the initial extreme stretch, β; it is also unstable even for infinitesimal uniaxial motion superimposed on a finite static stretch. It is easy to show from (3.2) that for infinitesimally small motions, the coefficient of θ is zero at $\kappa=\kappa^*$. Hence for $\kappa>\kappa^*$, this coefficient is negative denoting unstable small swinging motion, as expected for an inverted pendulum.

For the values of $B<-3$ the touch-down points corresponding to $j=2$ will not exist. As B becomes more negative, κ^* reduces. Only the unstable region of type II above the stability curve passing through $(\kappa,\lambda_s)=(\kappa^*,1.0)$ is found to exist for all such values of B; it becomes wider as B becomes more negative. Conclusions drawn are similar to those appear in the paragraph above. At this point it may be mentioned that a rigid simple inverted pendulum subjected to simple harmonic support motion can become stable for certain parameter combinations [15]. In our case more work is necessary to investigate similar behavior for the inverted pendulum.

4.4 Role of Material Nonlinearity. It is shown by Olsson [2] that unless the second-order terms that originate from the geometric nonlinearity are included in the equations of motion for the linear spring pendulum, parametric resonance cannot be captured. The presence of such higher-order terms yields the Mathieu equation for one of the independent coordinates. In the present case, inclusion of material nonlinearity results in the Hill's equation. The touch-down points in the stability curves automatically correspond to the results obtained by Olsson [2] and others. Owing to the presence of material nonlinearity two stability curves branch out from these touch-down points. Also, additional unstable regions exist for the values of B used here. Hence, the results of this investigation show that unlike the effect of geometric nonlinearity, material nonlinearity may give rise to unstable behavior at various values of the stiffness ratio, κ and the extreme stretch, β for a given value of B. It is also found by numerical solution of Eqs. (4.1) and (4.2) that the growth rate of unstable pendulum mode increases when a representative point in the unstable region lies away from the touch-down points inside the unstable regions.

5 Concluding Remarks

This paper deals with the dynamic behavior of a nonlinear spring pendulum with a support motion that is proportional to the axial response of the pendulum. Results are obtained for two special cases of motion. The first case deals with two uncoupled linear simple harmonic oscillators for infinitesimal motion of both the modes. The second case discusses the motion of the pendulum bob for small swing superimposed on finite axial motion. The stability for this special case helps one to identify the parameter combinations for which the original nonlinear system (2.4) and (2.5) would show a strong energy exchange behavior between the modes.

Results show that autoparametric resonance between the two modes of motion exists when the linearized axial mode natural frequency is twice that of the swing mode. Besides, for $B>0$, autoparametric resonance, although considerably weak in nature, also takes place for other values of j including 1. Inside the unstable region of either type, strong energy exchange behavior is observed for finite axial motion. Finally, this study brings out the usefulness of the stability criterion for the Mathieu-Hill equation used here.

Acknowledgments

The author would like to thank H. Das for his help with the graphics.

References

[1] Minorsky, N., 1962, *Nonlinear Oscillations*, Van Nostrand Reinhold, New York, p. 506.
[2] Olsson, M. G., 1976, ''Why Does a Mass on a Spring Sometimes Misbehave?'' Am. J. Phys., **44**, No. 12, pp. 1211–1212.
[3] Kane, T. R., and Kahn, M. E., 1968, ''On a class of Two Degrees of Freedom Oscillations,'' ASME J. Appl. Mech., Series E, **35**, pp. 547–552.
[4] Lai, H. M., 1984, ''On the Recurrence Phenomenon of a Resonant Spring Pendulum,'' Am. J. Phys., **52**, No. 3, pp. 219–223.
[5] Anicin, B. A., Davidovic, D. M., and Babovic, V. M., 1993, ''On the Linear Theory of the Elastic Pendulum,'' Eur. J. Phys., **14**, pp. 132–135.
[6] Ryland II, H. G., and Meirovitch, L., 1977, ''Stability Boundaries of a Swinging Spring With Oscillating Support,'' J. Sound Vib., **51**, No. 4, pp. 547–560.
[7] Nunez-Yepez, N., Salas-Brito, A. L., Vargas, C. A., and Vincente, L., 1990, ''Onset of Chaos in an Extensible Pendulum,'' Phys. Lett. A, **145**, pp. 101–105.
[8] Cuerno, R., Ranada, A. F., and Ruiz-Lorenzo, J. J., 1992, ''Deterministic Chaos in the Elastic Pendulum: A simple Laboratory for Nonlinear Dynamics,'' Am. J. Phys., **60**, No. 1, pp. 73–79.
[9] Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillation*, John Wiley and Sons, New York.
[10] Beatty, M. F., 1983, ''Finite Amplitude Oscillations of a Simple Rubber Support System,'' Arch. Ration. Mech. Anal., **83**, No. 3, pp. 195–219.
[11] Beatty, M. F., and Bhattacharyya, R., 1990, ''Poynting Oscillations of a Rigid Disk Supported by a Neo-Hookean Rubber Shaft,'' J. Elast., **24**, pp. 135–186.
[12] Bellman, R., 1969, *Stability Theory of Differential Equations*, Dover, New York.
[13] Bhattacharyya, R., 1995, ''A Stability Theorem for Hill's Equation for Engineering Applications,'' ASME J. Vibr. Acoust., **117**, pp. 380–381.
[14] Zhou, Z., 1993, ''Coupled Shear-Torsional Motion of a Rubber Support System,'' J. Elast., **30**, pp. 123–189.
[15] Cunningham, W. J., 1958, *Introduction to Nonlinear Analysis*, McGraw-Hill, New York.

T. H. Lin

K. K. F. Wong

Department of Civil and Environmental Engineering,
University of California,
Los Angeles, CA 90095-1593

N. J. Teng
Universal Analytics, Inc.,
Torrance, CA 90503

Micromechanics of Hysteresis Loops of Fatigue in a Single Crystal

Grain boundaries are susceptible to cause boundary corrosion, cracking, and creep deformation. Single crystals are presently used in turbine engines. A micromechanic analysis is shown to explain the occurrence of highly localized plastic strain in the slip band known as the shear band in metals under a monotonic loading. Based on the prior analyses of fatigue bands in polycrystals, a micromechanic analysis of a single crystal under plane deformation is developed. The Bauschinger effect and hysteresis loops of these single crystals were calculated and shown. The calculated results agree generally with experimental observations. [S0021-8936(00)02202-9]

1 Introduction

Single-crystal nickel-based superalloys have been developed to eliminate the grain boundaries which are susceptible to cause grain boundary corrosion, cracking, and creep deformation [1]. These single crystals are presently used in turbine engine parts. The prediction of fatigue life of these single crystals is of both scientific interest and practical need. This paper gives a method to analyze single crystals under cyclic loading in plane deformation. This analysis is based on the extension of the micromechanic high-cycle fatigue analysis of a face-centered-cubic (f.c.c.) polycrystal. This analysis is first shown to explain the formation of shear band under a monotonic loading, then to explain the growth of fatigue band in a polycrystal under a cyclic loading, and finally gives the analysis of fatigue band in high-cycle fatigue of a single crystal under plane deformation.

2 A Micromechanic Theory of Fatigue Crack Initiation

Initial defects always exist in metals and cause an initial stress field τ^i. During loading, when the resolved shear stress in some region reaches the critical shear stress τ^c, slip occurs. After unloading, this slip remains and induces a residual shear stress τ^r. Denoting the resolve shear stress caused by loading by τ^a, the total resolved shear stress is then

$$\tau = \tau^i + \tau^a + \tau^r. \tag{1}$$

The governing condition to initiate or continue sliding is to have the resolved shear stress equal to the critical shear stress, i.e.,

$$\tau = \tau^c \quad , \quad \text{slip occurs} \tag{2a}$$

$$\tau < \tau^c \quad , \quad \text{no slip.} \tag{2b}$$

(i) Role of Microstress Field on the Formation of Fatigue Band. When a piece of metal is uniformly loaded, slip lines appear on the surface. These slip lines are the results of highly localized plastic deformation. This raises the question of why the plastic strain is so heterogeneous. To explain this highly heterogeneous plastic deformation, the micromechanic shear stress field due to a uniform plastic strain e''_{12} in a thin slice (see Fig. 1) in an isotropic infinite medium is analyzed. This analysis gives the residual shear stress [2] along the x_1-axis as

$$\tau^r = \frac{4\mu e''_{12} w d}{\pi(1-\upsilon)} \frac{x_1^2 - d^2 - w^2}{(w^2+(x_1+d)^2)(w^2+(x_1-d)^2)}. \tag{3}$$

This τ^r is positive outside the slice. Hence the width $2d$ of this slice tends to increase. This explains why, in general, a slip band rapidly widens, covering the whole crystal. Along the x_2-axis, this analysis gives a residual shear stress as

$$\tau^r = \frac{4\mu e''_{12} w d}{\pi(1-\upsilon)} \left[1 - \left(\frac{x_2}{d}\right)^2\right]. \tag{4}$$

The thickness $2w$ is generally very small and approaches zero. The plastic strain e''_{12} required to yield a finite value of τ^r has to be very large. This explains why the highly localized plastic strain occurs in thin slip bands in single-phase metal under uniform loading. Hence, the combination of Eqs. (3) and (4) explains the formation of shear band under a monotonic loading. Equation (4) not only shows the large plastic strain in the shear band, but also gives the negligible variation of the residual shear stress across the thickness. This is referred to as the continuity of the resolved shear stress field. This is important in explaining the ratchet mechanism in fatigue band, which will be explained later.

(ii) Fatigue Band Model. Under cyclic loadings, slip lines appear on the surface (see Fig. 2). After removing these slip lines by electropolishing and recycling, original lines reappear. These slip lines are known as persistent slip bands (PSBs), which are the favorable sites of crack initiation. In the 1950s, thin ribbons protruding out of fatigue specimen surface were discovered [3]. These ribbons are known as extrusions. Negative extrusion, called intrusions, were also observed (see Fig. 3). Based on the hints supplied by these observations, a physical model was developed for high-cycle fatigue crack initiation.

The physical model of fatigue crack initiation is shown in Fig. 4(*a*). The extrusion or intrusion is represented by *R* in this figure.

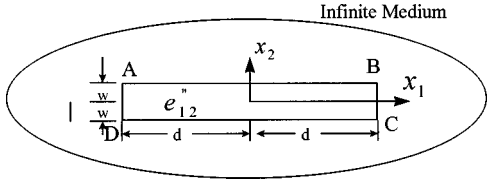


Fig. 1 Shear band model

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 17, 1998; final revision, October 1, 1999. Associate Technical Editor: J. L. Bassani. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

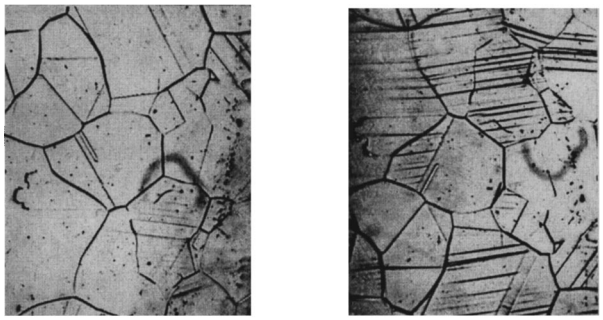


Fig. 2 Slip lines in polycrystalline nickel during two stages of cyclic loading ([17])

The formation of an extrusion requires a positive shear strain $e''_{\alpha\beta}$ in P on the top and a negative shear strain in Q on the bottom of the extrusion R. This can be caused by a positive initial shear stress τ^i in P and a negative initial shear stress $-\tau^i$ in Q. This system of initial resolved shear stress in a segment can be caused by a change of compressive stress in R, as shown in Fig. 4(b).

Consider the segment $ABCD$ in Fig. 4(b) having an initial tensile strain $e^i_{\alpha\alpha}$; i.e., this segment has an initial length longer than the slot. Imagine that this segment is cut out and compresses to the same length as the slot, and then is welded back to the slot under this imaginary compression. Since there is no such compression, this compression must be relieved by applying an equal and opposite force. This produces a compressive stress $\tau_{\alpha\alpha}$ on the segment at the free surface, which pushes the segment out of the free surface, creating extrusion. Hence a set of positive shear stress in P and negative in Q is produced by an initial tensile strain in R. This initial tensile strain can be provided by a row of interstitial dipoles [2,4].

(iii) Ratchet Mechanism. With an initial tensile strain $e^i_{\alpha\alpha}$ in R (see Fig. 4), the initial shear stress in P, τ^i_P, is positive and that of Q, τ^i_Q, is negative. The shear stress due to the applied load σ_{22} is the same in the whole crystal, thus $\tau^a_P=\tau^a_Q=\tau^a$. Due to the continuity of the residual shear stress field as given by Eq. (4), $\tau^r_P=\tau^r_Q=\tau^r$. Consider the following sequence of loadings:

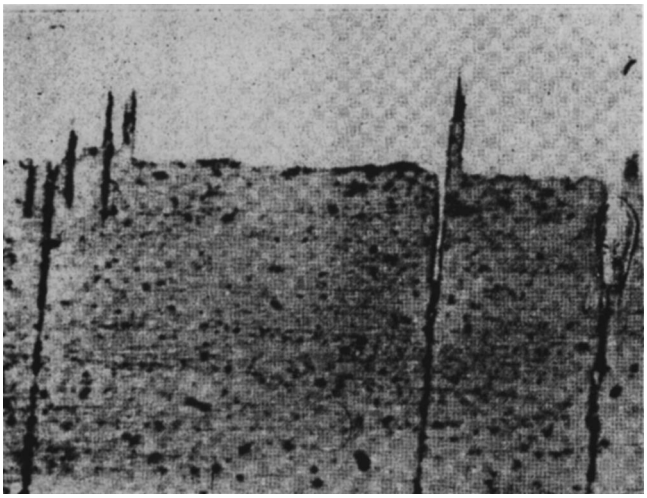


Fig. 3 Instrusions and extrusions in copper during fatigue ([17])

1 First Cycle Forward Loading ($\tau^a>0$): P slides, $\tau^r_{1f}<0$, where the subscript ''$1f$'' denotes the first forward loading. Therefore,

$$\tau_P=\tau^i_P+\tau^a+\tau^r_{1f}=\tau^c \tag{5a}$$

$$\tau_Q=\tau^i_Q+\tau^a+\tau^r_{1f}>-\tau^c. \tag{5b}$$

2 First Cycle Reversed Loading ($\tau^a<0$): Q slides, $\tau^r_{1r}>0$. Therefore,

$$\tau_P=\tau^i_P+\tau^a+\tau^r_{1f}+\tau^r_{1r}<\tau^c \tag{6a}$$

$$\tau_Q=\tau^i_Q+\tau^a+\tau^r_{1f}+\tau^r_{1r}=-\tau^c. \tag{6b}$$

3 Second Cycle Forward Loading ($\tau^a>0$): P slides, $\tau^r_{2f}<0$. Therefore,

$$\tau_P=\tau^i_P+\tau^a+\tau^r_{1f}+\tau^r_{1r}+\tau^r_{2f}=\tau^c \tag{7a}$$

$$\tau_Q=\tau^i_Q+\tau^a+\tau^r_{1f}+\tau^r_{1r}+\tau^r_{2f}>-\tau^c. \tag{7b}$$

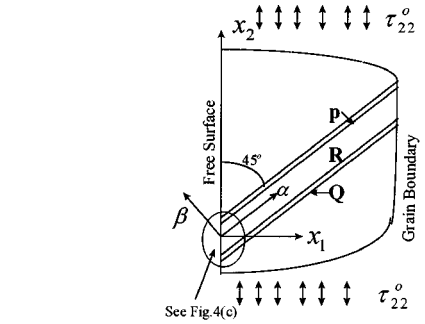
4 Second Cycle Reversed Loading ($\tau^a<0$): Q slides, $\tau^r_{2r}>0$. Therefore,

$$\tau_P=\tau^i_P+\tau^a+\tau^r_{1f}+\tau^r_{1r}+\tau^r_{2f}+\tau^r_{2r}<\tau^c \tag{8a}$$

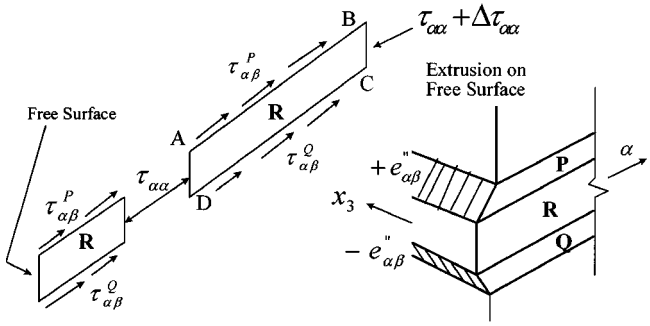
$$\tau_Q=\tau^i_Q+\tau^a+\tau^r_{1f}+\tau^r_{1r}+\tau^r_{2f}+\tau^r_{2r}=-\tau^c. \tag{8b}$$

This process is repeated. A typical numerical result of the plastic strain distributions in P at different cycles of loading of the surface crystal is shown in Fig. 5. It is seen that P always slides in the positive direction and Q in the negative direction. The magnitudes of these slips and hence the extrusions are monotonically increasing.

(iv) Secondary Slip. The buildup of the slip strain $e''_{\alpha\beta}$ in P and Q is caused by $e^i_{\alpha\alpha}$ in R. If R were cut out, the free length of



(a) Most Favorable Oriented Crystal



(b) Compressive Stress in R to Cause $+\tau^P_{\alpha\beta}$ and $-\tau^Q_{\alpha\beta}$

(c) Three-Dimensional View of Extrusion

Fig. 4 Fatigue band model

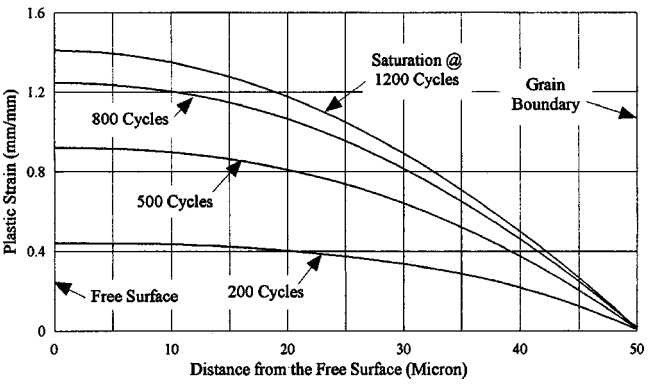


Fig. 5 Typical plastic strain distribution under cyclic loadings of aluminum

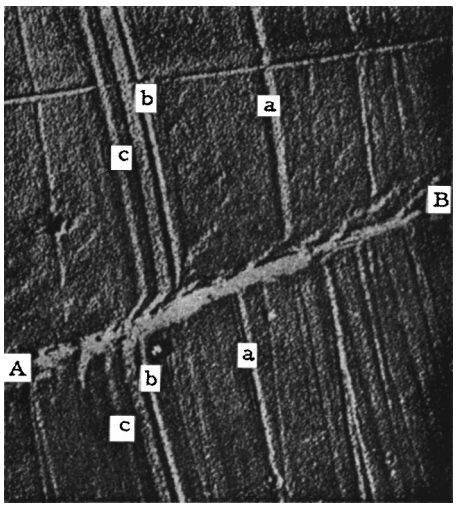


Fig. 6 Initially straight scratches a, b, c are displaced unidirectionally by static slip band AB. (Reproduced from Trans. Metal Soc. AIME, 1962, courtesy of AIME.)

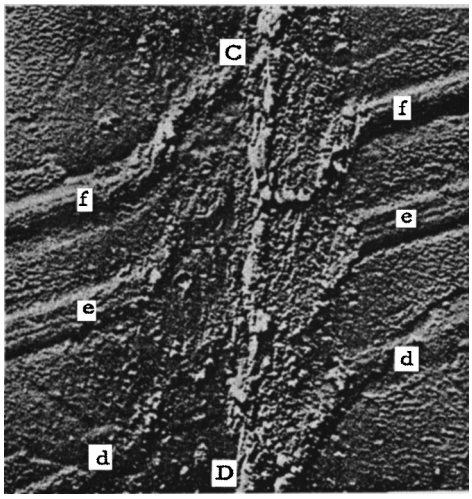


Fig. 7 Cyclic slip band CD produces no overall displacement of scratches d, e, f. Within the slip band; the scratches are displaced equally backward and forward. (The same as Fig. 6, courtesy of AIME.)

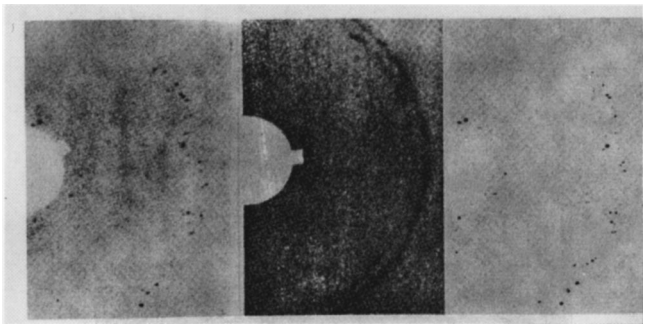


Fig. 8 X-ray reflection patterns: (*a*) Sharp X-ray annealed α-brass. (*b*) From same specimen as (*a*) after a unidirectional strain 150×50 deg twist. (*c*) From same specimen as (*a*) after 1500 reversals of plastic strain 1.5-deg twist and showing same reflections as (*a*). (Reproduced from the book *Fracture*, 1959, courtesy of Technological Press, MIT.)

R would be longer than the slot by an amount known as the static extrusion [5]. The $e^i_{\alpha\alpha}$ causes an initial compression in R, which in turn causes positive $\tau^i_{\alpha\beta}$ in P and negative $\tau^i_{\alpha\beta}$ in Q. Under cyclic loading, the extrusion grows and R increases in length. This elongation causes the compression in R to decrease. There are 12 slip systems in a f.c.c. crystal. The change of direct stress $\tau_{\alpha\alpha}$ in R causes changes in resolved shear stress in all slip systems. When the decrease in compression in R becomes large, the applied stress can cause a second slip system to have shear stress reaching the critical value and slide. The plastic strain $e''_{\xi\eta}$ caused by slip in this secondary slip system has a *tensor* component $e''_{\alpha\alpha}$, just like the initial tensile strain $e^i_{\alpha\alpha}$ in causing the positive and negative $\tau^i_{\alpha\beta}$ in P and Q, respectively [6]. Hence with secondary slip, the extrusion can grow considerably beyond the static extrusion. The occurrence of the secondary slip system was recently clearly observed [7].

3 Experimental Verifications

This model has extensive metallurgical supports [2], and two of these supports are shown as follows:

(i) Slip Band Formation. An informative experiment on slip band formation was made by Wood and Bender [8]. They tested copper circular rod specimens subject to torsion. The specimens were electropolished and then scratched as markers with a pad carrying 0.5μ diamond dust. Some specimens were subject to alternate torsion and some subject to single twist through large angles. The deformation in a typical slip band AB of a specimen subject to single twist is shown in Figs. 6 a, b, c are typical scratches which were initially straight and continuous. It is seen that the single twist causes the scratches above AB to displace relative to those below. Figure 7 shows the deformation under cyclic torsion with scratches d, e, f and a typical fatigue band DC. It is seen that the cyclic deformation caused no relative displacement of the scratches left and right of the fatigue band, but within the band the scratches have displaced equally up and down producing a zig-zag. A severely slid line with positive shear as P is sandwiched in two less severely slid lines with negative shear such as Q. This clearly verifies with the theory proposed.

(ii) Lattice Straining. X-ray reflection patterns of monotonically and cyclically loaded specimens are very different [9]. The later retain the discrete spots like that of annealed metals while the former do not (Fig. 8). This shows that slip occurrence in alternate loadings does not cause lattice straining in the bulk of the metal. Under cyclic loading, the positive shear slip lines (like P) are closely located with the negative one (like Q). At some distance from the slip lines, the stress field caused by positive slip in P is balanced by that caused by negative slip in Q. Hence the stress field and the lattice strain is small in the bulk of the metal.

Under monotonic loadings, the slip in all slip lines tends to be all of the same sign and causes a significant average plastic strain, which causes an appreciable stress field and a lattice strain in the bulk of the metal. The above theory accounts for the different X-ray reflection patterns of the monotonically deformed and cyclically deformed metals.

4 Single Crystals

The single-crystal nickel-based superalloys have been developed to eliminate the grain boundaries, which are susceptible to grain boundary corrosion, cracking, and creep deformation [1]. These single crystals are presently used in turbine engines. The prediction of high-cycle fatigue life of these single crystals is of practical need. The following shows the analysis of fatigue bands of single crystals.

(i) Method of Analysis. A crystal embedded at the free surface of the polycrystal under alternated tension and compression is first analyzed, as shown in Fig. 4(*a*). This solution gives surface tractions on the grain boundary. In a single crystal, the surface tractions are zero and hence must be removed by applying equal and opposite tractions on the boundary. The stress field caused by this equal and opposite tractions is analyzed by the finite element method.

Plastic strain is taken to occur only in the fatigue band. The band is divided into a number of grids. The plastic strain in the left half of the crystal is denoted by $e_{ij}''^{(L)}$ and in the right half by $e_{ij}''^{(R)}$. The solution of the stress field in Fig. 9(*a*) is the sum of the solutions of Figs. 9(*b*), (*c*), and (*d*). Figure 9(*d*) gives uniform stress. The stress field due to $e_{ij}''^{(L)}$ is solved by the semi-infinite solid solution with the free surface at the left, as shown in Fig. 10. This solution yields the condition of zero traction at the free surface and gives surface tractions on the right, top, and bottom planes (see Fig. 10(*b*)). These tractions are removed by applying equal and opposite tractions, as shown in Fig. 10(*c*). The plastic strain in the grid is replaced by the equivalent forces [10] and the stress field caused by these equivalent forces is solved by the finite element method. This equivalent force due to $e_{ij}''^{(L)}$ is relatively far from the considered crystal boundaries and hence the variation of the surface traction on the boundary is small, and thus the grids of the finite element method do not need to be very fine. This facilitates the finite element method solution. Similarly, for the grid in the right half of the crystal, the free surface of the semi-infinite solution is taken to be at the right side.

The initial strain e_{ij}^{i} has the same effect as plastic strain e_{ij}'' in causing a stress field, so the initial strain can be analyzed in the same way as the plastic strain. This gives a method to calculate the stress influence coefficient in the mth grid caused by a unit inelastic strain (either plastic strain or initial strain) in the nth grid. The resolved shear stress τ equals the sum of the initial stress τ^{i}, the applied stress τ^{a}, and the residual stresses $\tau^{r(L)}+\tau^{r(R)}$. Equating the resolve shear stress τ to the critical shear stress τ^{c} gives the incremental plastic strain distributions at different stages of loading.

(ii) Experimental Observations of Fatigue Bands in Single Crystals. Mecke and Blochwitz [11] observed the subgrain displacement in single nickel crystal under cyclic loading. These experiments were carried out under constant plastic strain amplitudes at room temperature. It is shown that the PSBs have penetrated across the whole crystal and extruded out on both sides as shown in Fig. 11. The case with extrusion protruding on one side and intrusion on the other side was not observed. Basinski et al. [12,13] tested copper single crystals at constant plastic strain amplitudes at room temperature under cyclic loadings. In these tests, both extrusions and intrusions are observed. Zhai et al. [14,15]

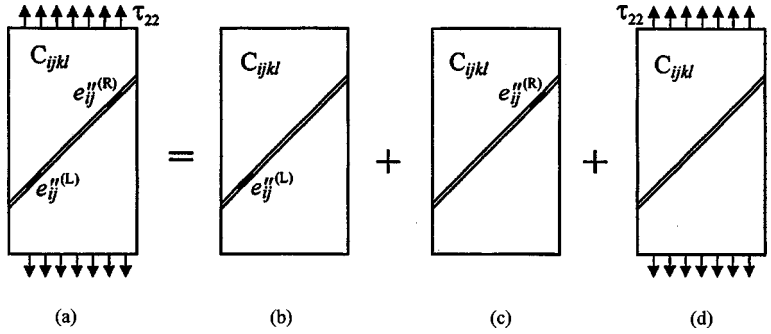


Fig. 9 Procedure for decoupling the single crystal problem

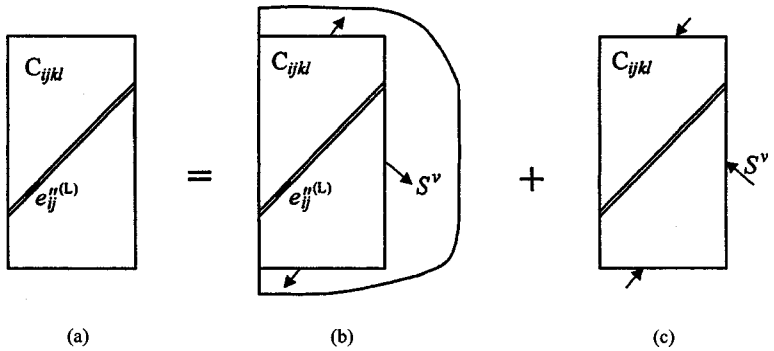


Fig. 10 Removal of boundary tractions for a single crystal

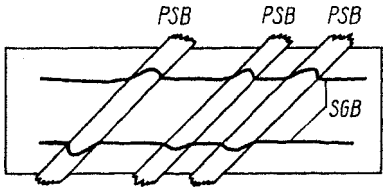


Fig. 11 Extrusions observed in single crystal [11]

performed fatigue experiments on aluminum single crystals under constant cyclic stress amplitude. Again, both extrusions and intrusions were observed on the free surfaces.

(iii) Numerical Calculations. The analytical solution developed in previous sections is here applied to analyze the single-crystal tests. To simplify the calculation, a single fatigue band in the single crystal is considered (the grain boundary shown in Fig. 4(*a*) is now a free surface). The analytical method can readily be used to analyze multiple fatigue bands in the crystal. Referring to Fig. 4(*a*), both *P* and *Q* are assumed to be 0.05 μm in thickness, and *R* to be 1.0 μm. The crystal is f.c.c. and is assumed to be elastically isotropic with shear modulus $\mu = 50$ GPa and the Poisson ratio $\upsilon = 0.3$. The critical shear stress, τ^c, is taken to be 200 MPa, and cyclic loading $\tau_{22} = 399.55$ MPa. An initial tensile strain was assumed to vary linearly from a maximum value at the center to zero at the two ends of a 1.4 mm segment in the fatigue band. This segment was divided into a number of grids, and each grid was approximated by a uniform initial tensile strain. This assumed initial tensile strain distribution was found to give an initial resolved shear stress, τ^i, quite uniform over each half of the fatigue band. Thus a uniform τ^i of 0.5 MPa was used in the present analysis. The variations of the plastic strain in *P* and *Q* along the length of the fatigue band at different cycles of loading are shown in Fig. 12.

The widening of a slip band, i.e., the increase of ''2*d*,'' has been explained in the micromechanic analysis (see Eq. (3)). This gives the spread of the fatigue band toward the two free surfaces, causing the protruding of extrusion on two sides. If the initial tensile strain in the above is replaced by an initial compression strain, intrusions instead of extrusions will occur on both sides. The present model seems to explain the observed extrusions and intrusions.

(iv) Hysteresis Loops. This analytical model is used to calculate hysteresis loops of aluminum single crystals, with shear modulus $\mu = 26.5$ GPa and Poisson ratio $\upsilon = 0.3$. The critical shear stress, τ^c, is taken to be 0.369 MPa. A calculated hysteresis loop is shown in Fig. 13. In the initial loading and unloading curve, the elastic limit in unloading occurs at a positive normal stress, i.e., at positive resolved shear stress. This indicates a large Bauschinger effect. The widths of the hysteresis loops have been found to decrease with the number of cycles, which agree with the experimental hysteresis loops of aluminum single crystals shown in Fig. 14 [16].

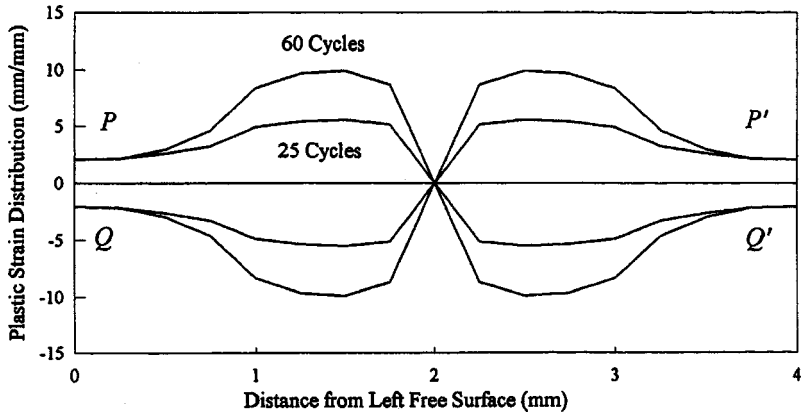


Fig. 12 Plastic strain distribution with initial strain at center. P & Q′ and Q & P′ are symmetrically located. Extrusions protruding out on both faces.

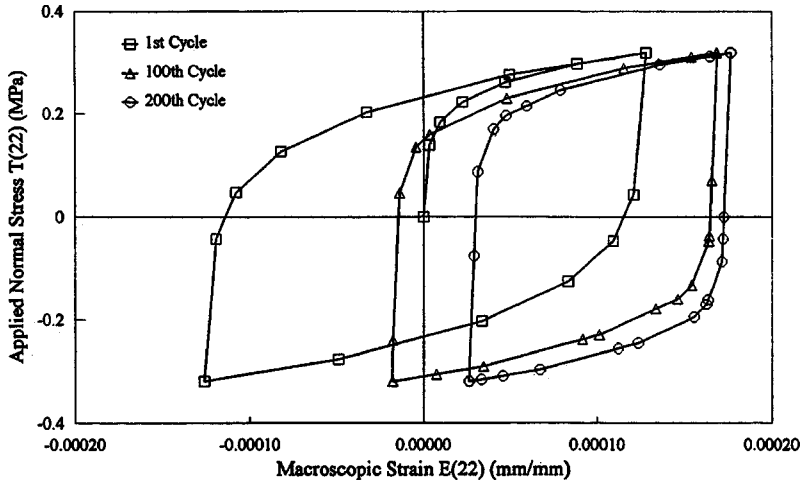


Fig. 13 Hysteresis loops of an aluminum single crystal

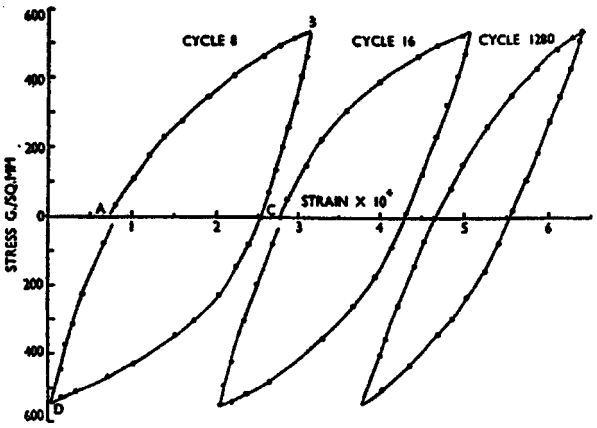


Fig. 14 Experimental observation of hysteresis loops in aluminum single crystal ([18])

5 Conclusions

For an extrusion to protrude, the shear strain in P has to be positive and that in Q has to be negative near the occurrence of extrusion. This requires positive resolved shear stress in P and negative in Q. In turn, this requires a compressive stress in R to push the extrusion out. The occurrence of extrusion on both sides of the single crystal implies compression in R on both faces. Similarly, for an intrusion to occur on the left side, the shear strain in P must be negative and that in Q must be positive. This requires a tensile stress in R to pull the intrusion in. A segment in R with an initial compressive strain tends to increase the length of the tensile stress in R under cyclic loadings. The spread of the tension in R over the length of the fatigue band will result in intrusions on both faces. Initial compression and initial tension may occur in the specimens. Hence extrusions on both faces and intrusions on both faces have been observed.

This single-crystal fatigue band analysis is for plane strain, which gives an approximate solution for the central length portion of the crystal. To remove the assumption of plane strain, a three-dimensional model is required. In the past year, an elastic-plastic boundary element method to analyze the fatigue bands and hysteresis loops in an aluminum single crystal has been developed. The calculated results by the developed method check amazingly well with the aluminum single crystal fatigue test data recently conducted by Zhai et al. in Oxford, England. These results will appear in *Philosophical Magazine A* (2000). This study is essential to improve the representation of the constitutive relation of single crystals and single crystal superalloys, which is important to the design of engine parts made of single crystals.

Acknowledgment

This work was sponsored by the Air Force Office of Scientific Research (AFOSR), USAF, under grant number F49629-96-1-350. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the AFOSR or the U.S. Government.

References

[1] Walker, K. P., and Jordan, E. M., 1989, ''Biaxial Constitutive Modeling and Testing of a Single Crystal Superalloy at Elevated Temperatures,'' *Biaxial and Multiaxial Fatigue, EGF3,''* M. W. Brown and K. J. Miller, eds., Mechanical Engineering Publication, London, pp. 145–170.

[2] Lin, T. H., 1992, ''Micromechanics of Crack Initiation in High-Cyclic Fatigue,'' Adv. Appl. Mech., **29**, pp. 1–62.

[3] Forsyth, P. J. E., and Stubbington, C. A., 1955, ''The Slip Band Extrusion Effect Observed in Some Aluminum Alloys Subjected to Cyclic Stresses,'' J. Inst. Met., **83**, p. 395.

[4] Essmann, V., Gossel, V., and Mughrabi, H., 1981, ''A Model of Extrusions and Intrusions in Fatigued Metals I—Point Defect Production and the Growth of Extrusions,'' Philos. Mag. A, **44**, pp. 405–426.

[5] Mughrabi, H., Wang, R., Differt, K., and Essmann, V., 1983, ''Fatigue Crack Initiation by Cyclic Slip Irreversibilities in High-Cycle Fatigue,'' *Fatigue Mechanism*, STM-STP-811, pp. 5–45.

[6] Lin, T. H., Lin, S. R., and Wu, X. Q., 1989, ''Micromechanics of an Extrusion in High-Cyclic Fatigue,'' Philos. Mag. A, **59**, pp. 1263–1276.

[7] Zhai, T., Briggs, G. A. D., and Matin, J. W., 1996, ''Fatigue Damage at Room Temperature in Aluminum Single Crystals IV: Secondary Slip,'' Acta Mater., **44**, pp. 3489–3496.

[8] Wood, W. A., and Bendler, A. M., 1962, ''The Fatigue Process in Copper as Studies by Electron Metallography,'' Trans. Metall. Soc. AIME, **244**, pp. 180–186.

[9] Wood, W. A., 1956, ''Mechanisms of Fatigue,'' *Fatigue in Aircraft Structure*, A. M. Freudenthal, ed., Academic Press, New York, pp. 1–19.

[10] Lin, T. H., 1968, *Theory of Inelastic Structures*, John Wiley and Sons, New York.

[11] Mecke, K., and Blockwitz, C., 1980, ''Internal Displacement of Persistent Slip Bands in Cyclically Deformed Nickle Single Crystals,'' Phys. Status Solidi A, **64**, pp. K5–K7.

[12] Basinski, Z. S., Pascual, R., and Bainski, S. J., 1983, ''Low Amplitude Fatigue of Copper Single Crystals I—The Role of the Surface in Fatigue Failure,'' Acta Metall., **31**, pp. 591–602.

[13] Basinski, Z. S., and Bainski, S. J., 1985, ''Low Amplitude Fatigue of Copper Single Crystals II—PSB Sections,'' Acta Metall., **33**, pp. 1319–1327.

[14] Zhai, T., Lin, S., and Xiao, J. M., 1990, ''Influence on Non-Geometric Effect of PSB on Crack Initiation in Aluminum Single Crystal,'' Acta Metall. Mater., **38**, pp. 1687–1692.

[15] Zhai, T., Matin, J. W., and Briggs, G. A. D., 1995, ''Fatigue Damage at Room Temperature in Aluminum Single Crystals I: On the Surface Containing the Slip Burger's Vector,'' Acta Metall. Mater., **43**, pp. 3813–3825.

[16] Thompson, N., and Wadsworth, N. J., 1958, ''Metal Fatigue,'' Adv. Phys., **7**, pp. 72–170.

[17] Kennedy, A. J., 1963, *Process of Creep and Fatigue of Metals*, John Wiley and Sons, New York.

[18] Thompson, N., Coogan, C. K., and Rider, J. R., 1955, ''Experiments on Aluminum Crystals Subjected to Slowly Alternating Stresses,'' J. Inst. Met., **84**, pp. 73–80.

Low-Gravity Sloshing in an Axisymmetrical Container Excited in the Axial Direction

M. Utsumi
Machine Element Department,
Research Institute,
Ishikawajima-Harima Heavy
Industries Company, Ltd. (IHI),
3-1-15 Toyosu, Koto-ku,
Tokyo 135-0061, Japan

The response of low-gravity propellant sloshing is analyzed for the case where an axisymmetrical container is exposed to axial excitation. Spherical coordinates are used to analytically derive the characteristic functions for an arbitrary axisymmetrical convex container, for which time-consuming and expensive numerical methods have been used in the past. Numerical results show that neglecting the surface tension results in the underestimation of the magnitude of the liquid surface oscillation. The reason for this is explained by the influences of the Bond number and the liquid filling level on the critical value of the coefficient of the excitation term in the modal equation, above which the oscillation is destabilized, and on the characteristic root of the destabilized system.
[S0021-8936(00)01502-6]

1 Introduction

The importance of low-gravity propellant sloshing in space vehicle operations is well recognized ([1]). For recent large vehicles in particular, the propellant contributes a non-negligible portion to the total mass of the vehicle, and therefore the importance of sloshing is accentuated.

The problem of low-gravity sloshing is characterized by the dominant role of the surface tension of the liquid, which curves the liquid surface strongly even in the undisturbed static case. This curved static liquid surface, which is called the meniscus, makes the problem geometrically more complicated than the sloshing problem under normal gravity, especially for an arbitrary axisymmetrical container with curved walls and top.

Several studies have been conducted for low-gravity sloshing. For a cylindrical container, we can obtain analytical expressions for the characteristic functions constituting the modal functions of the velocity potential and the liquid surface displacement ([2–6]). For an arbitrary axisymmetrical container, however, the characteristic functions were not determined analytically in previous work ([7–12]). Numerical procedures were applied instead, and the computation time and cost increased dramatically when the computational mesh was refined. To solve this problem, the author ([13,14]) developed a new analytical method for determining the characteristic functions for an arbitrary axisymmetrical container. The distinguishing feature of the method is the use of spherical coordinates whose origin is at the top of the cone that is tangent to the container at the contact line of the meniscus with the container wall. This analytical method requires little computation time and cost and also has the following advantages over numerical methods: (i) the mathematical formulation allows the liquid surface and its dynamical displacement to be expressed as a single-valued function, even when the liquid surface curves strongly due to the surface tension in a low-gravity space environment; (ii) the kinematic compatibility condition for the liquid surface displacement can be satisfied at the container wall since the liquid surface displacement at the wall can be made tangential to the container wall. The method is thus a geometrically convenient means of solving the low-gravity sloshing problem for an arbitrary axisymmetrical container.

In the previous papers ([13,14]), the meniscus shape, the eigenfrequency, and the responses of surface slosh motion and slosh force and moment to lateral excitation of a spherical container were calculated. For engineering purposes, however, the analysis must be extended to the case of axial excitation, since the axial excitation can cause critical liquid surface oscillation when the excitation frequency is twice the eigenfrequency of the fundamental mode. The purpose of the present paper is to predict the response of surface slosh motion to axial excitation and its dependency on the liquid filling level and the Bond number.

2 Computational Model

The geometry is defined as in Fig. 1, where an ellipsoidal container is drawn as a typical example of an axisymmetrical container. The container is subjected to the axial acceleration $\ddot{f}(t)$ in the z-direction. The liquid domain and the container wall are denoted by V and W, respectively. The undisturbed and disturbed liquid surfaces are represented by M and F, respectively. The low-g sloshing is characterized by the strongly curved meniscus M, which is 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 container is rigid.
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.

3 Spherical Coordinate System

As shown in Fig. 1, spherical coordinates R, θ, and φ are introduced and the liquid surface displacement ζ is considered in the R-direction. The origin O is chosen as the apex of the cone whose side wall is tangent to the container wall at the contact line of the meniscus with the container wall. The origin is above the container for $z_C > b$ (Case 1) and below otherwise (Case 2). By using the spherical coordinates, the meniscus M, the disturbed liquid surface F, and the container wall W can be expressed as

$$M: \quad R = R_M(\theta), \tag{1}$$

$$F: \quad R = R_F(\theta,\varphi,t) = R_M(\theta) + \zeta(\theta,\varphi,t), \tag{2}$$

$$W: \quad R = R_W(\theta). \tag{3}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Apr. 23, 1998; final revision, Jan. 20, 2000. Associate Technical Editor: D. A. Siginer. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

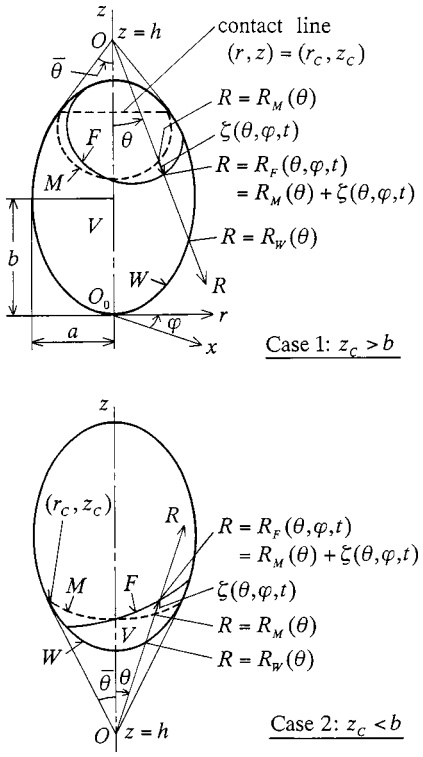


Fig. 1 Axisymmetrical container and coordinate systems

Using the spherical coordinates has the merits mentioned in the Introduction. Note that only one component of the surface displacement vector needs to be considered in order to satisfy the compatibility condition.

4 The Variational Principle

For the present problem, the variational principle can be written in the form

$$\delta\int_{t_1}^{t_2}\left\{\int\int\int_V (p_l-p_g)dV-\int\int_F \sigma dF - \int\int_{W_1}\sigma_1 dW_1-\int\int_{W_2}\sigma_2 dW_2\right\}dt=0, \tag{4}$$

where p_l is the liquid pressure; p_g is the gas pressure; and σ, σ_1, and σ_2 are the surface energy per unit area associated with the liquid-gas interface F, liquid-solid interface W_1, and gas-solid interface W_2, respectively. The Lagrangian within braces can be derived as follows. When surface energy and gas pressure are neglected, the Lagrangian density equals the liquid pressure ([15]) and therefore the Lagrangian becomes equal to $\int\int\int_V p_l dV$. From this, we must subtract the potential energy due to the gas pressure $p_g V$ and that due to the surface energy in order to estimate the Lagrangian for the case under consideration here. The gas pressure is assumed to be constant, since the gas density is much smaller than the liquid density.

5 Static Analysis

As a preliminary step to the dynamical slosh analysis, we must determine the meniscus shape $R_M(\theta)$. The variational principle (4) can be reduced to the following principle of virtual work:

$$\delta U_1+\delta U_2+\delta U_3=0, \tag{5}$$

where

$$\delta U_1=\delta\int\int\int_{V_{st}}(p_{l,st}-p_g)dV_{st}, \tag{6}$$

$$\delta U_2=-\delta\int\int_M \sigma dM, \tag{7}$$

$$\delta U_3=-\delta\left(\int\int_{W_{1,st}}\sigma_1 dW_{1,st}+\int\int_{W_{2,st}}\sigma_2 dW_{2,st}\right). \tag{8}$$

In Eq. (6), $p_{l,st}$ is given by

$$p_{l,st}=p_C+\rho_f g\varepsilon[R\cos\theta-R_M(\bar{\theta})\cos\bar{\theta}], \tag{9}$$

where ε is 1 and -1, respectively, for Cases 1 and 2 (Fig. 1). This expression is convenient for simultaneous mathematical description of the two cases. Expressing Eqs. (6)–(8) in terms of the virtual displacement of the meniscus δR_M, we obtain

$$\delta U_1=-\int\int_M (p_{l,st}-p_g)\,\delta R_M\cos\gamma_M dM, \tag{10}$$

$$\delta U_2=-\int\int_M \sigma\,\mathrm{div}\,\mathbf{N}_M\delta R_M\cos\gamma_M dM + \varepsilon\int_{C_{st}}\sigma\,\delta R_M\cos\theta_C dC_{st}, \tag{11}$$

$$\delta U_3=\varepsilon\int_{C_{st}}(\sigma_1-\sigma_2)\,\delta R_M dC_{st}, \tag{12}$$

where $\mathbf{N}_M$ is the unit normal vector of the meniscus pointing into the liquid domain, γ_M is the angle between $\mathbf{N}_M$ and the R direction, and C_{st} is the static contact line. The transformation from Eq. (6) into (10) is based on the fact that the variation in V_{st} results from the virtual displacement of the meniscus in its normal direction. The derivation of Eq. (12) from (8) is due to the fact that the variations in $W_{1,st}$ and $W_{2,st}$ are caused by the virtual displacement of the meniscus at the contact line C_{st}. The first term on the right side of Eq. (11) is the work done by the virtual displacement in the direction normal to the meniscus. This work can be evaluated by considering the variation of an infinitesimal surface element (Appendix A). On the other hand, the work done by the virtual displacement in the tangential direction of each surface element cancels out mutually over the surface M due to the interaction between the adjacent surface elements, and reduces to an integral along the contact line C_{st} (the second term on the right side of Eq. (11)), which is combined with δU_3 to yield the contact angle condition as can be seen later (Eq. (14)).

Substituting Eqs. (10)–(12) into Eq. (5), we obtain

$$p_{l,st}-p_g+\sigma\,\mathrm{div}\,\mathbf{N}_M=0 \quad \text{on } M, \tag{13}$$

$$\sigma\cos\theta_C+\sigma_1-\sigma_2=0 \quad \text{along } C_{st}. \tag{14}$$

Equation (13) represents the equilibrium condition among the liquid pressure, the gas pressure, and the surface tension of the liquid. Equation (14) is called the contact angle condition, since it requires that the contact angle between the meniscus and the container wall takes a prescribed value θ_C according to the magnitude of the surface tensions σ, σ_1, and σ_2. Using Eq. (1), these conditions can be expressed in terms of $R_M(\theta)$ as (Appendix B)

$$R_{M\theta\theta}=\frac{2R_M^2+3R_{M\theta}^2}{R_M}-\frac{R_{M\theta}}{\tan\theta}\frac{R_M^2+R_{M\theta}^2}{R_M^2}-\frac{(R_M^2+R_{M\theta}^2)^{3/2}}{R_M} \times\left\{\frac{\varepsilon(p_g-p_C)}{\sigma}-\frac{\rho_f g}{\sigma}[R_M\cos\theta-R_M(\bar{\theta})\cos\bar{\theta}]\right\}, \tag{15}$$

$$R_{M\theta}(\bar{\theta})=\frac{-\varepsilon R_M(\bar{\theta})}{\tan\theta_C}. \tag{16}$$

A backward integration of Eq. (15) is used to obtain a solution

$R_M(\theta)$ having an arbitrary prescribed value at $\theta=\overline{\theta}$ and satisfying Eq. (16). The value of unknown parameter p_g-p_C is determined such that $R_{M\theta}(0)=0$ by iterative computations.

Equation (15) is equivalent to the widely used Young-Laplace equation, as is proved in Appendix C. Thus the variational principle (4) for the sloshing problem reduces to the conventionally used equilibrium equations when the container is at rest.

6 Slosh Analysis

6.1 Governing Equations. In this section, governing equations are derived from the variational principle (4). By using the pressure equation for unsteady flow, p_l can be expressed in terms of the velocity potential ϕ describing the liquid motion relative to the moving container:

$$p_l=p_C-\rho_f\left\{\frac{\partial\phi}{\partial t}+g\varepsilon[R_M(\overline{\theta})\cos\overline{\theta}-R\cos\theta]+(h-\varepsilon R\cos\theta)\ddot{f}(t)+\frac{1}{2}(\nabla\phi)^2+\dot{G}(t)\right\}, \tag{17}$$

where $\dot{G}(t)$ is an arbitrary time function. Substituting Eq. (17) into (4) and applying the variation leads to ([14])

$$\int_{t_1}^{t_2}\left[\rho_f\int\int\int_V\nabla^2\phi\,\delta\phi dV-\rho_f\int\int_W\nabla\phi\cdot\mathbf{N}_W\delta\phi dW-\rho_f\int\int_F\left(\frac{\partial\zeta}{\partial t}\cos\gamma_F-\nabla\phi\cdot\mathbf{N}_F\right)\delta\phi dF+\int\int_F(p_g-p_l-\sigma\,\mathrm{div}\,\mathbf{N}_F)\,\delta\zeta\cos\gamma_F dF+\varepsilon\int_C(\sigma\cos\theta_C'+\sigma_1-\sigma_2)\,\delta\zeta dC-\rho_f\delta G\int\int_F\frac{\partial\zeta}{\partial t}\cos\gamma_F dF\right]dt=0. \tag{18}$$

Since the variations $\delta\phi$, $\delta\zeta$, and δG are arbitrary and independent of one another, we obtain the field equation (19), the boundary conditions (20)–(23), and the volume constant condition (24):

$$\nabla^2\phi=0 \quad \text{in } V, \tag{19}$$

$$\nabla\phi\cdot\mathbf{N}_W=0 \quad \text{on } W, \tag{20}$$

$$\frac{\partial\zeta}{\partial t}\cos\gamma_F-\nabla\phi\cdot\mathbf{N}_F=0 \quad \text{on } F, \tag{21}$$

$$p_g-p_l-\sigma\,\mathrm{div}\,\mathbf{N}_F=0 \quad \text{on } F, \tag{22}$$

$$\sigma\cos\theta_C'+\sigma_1-\sigma_2=0 \quad \text{along } C, \tag{23}$$

$$\int\int_F\frac{\partial\zeta}{\partial t}\cos\gamma_F dF=0. \tag{24}$$

The Laplace equation (19) corresponds to the condition of continuity within the liquid domain. Equation (20) means that the liquid velocity in the direction normal to the rigid wall vanishes. Equations (21) and (22) represent the dynamic and kinematic boundary conditions on the liquid surface, respectively. Equation (23) is the contact angle condition, which requires that the contact angle between liquid surface and container wall remains constant during sloshing. Equation (24) shows that the liquid volume is constant since the liquid is assumed to be incompressible. Since Eq. (24) can be derived from the other kinematic conditions (19), (20), and (21), Eqs. (19)–(23) constitute the system of basic equations which governs low-gravity sloshing. In the following analysis using the Galerkin method, the combined and integrated form (Eq. (18)) is more convenient than the separated form (Eqs. (19)–(24)).

6.2 The Variational Principle in Spherical Coordinates. To solve the sloshing problem for an axisymmetrical container, it appears most convenient to express Eq. (18) in terms of the spherical coordinates introduced in Section 3. We first express the following quantities in terms of $R_F(\theta,\varphi,t)$ and $R_w(\theta)$ and their θ and φ derivatives:

$$\mathbf{N}_F=\varepsilon(\mathbf{e}_R R_F\sin\theta-\mathbf{e}_\theta R_{F\theta}\sin\theta-\mathbf{e}_\varphi R_{F\varphi})/[(R_F^2+R_{F\theta}^2)\times\sin^2\theta+R_{F\varphi}^2]^{1/2}, \tag{25}$$

$$\mathbf{N}_W=\varepsilon(\mathbf{e}_R R_W-\mathbf{e}_\theta R_{W\theta})/(R_W^2+R_{W\theta}^2)^{1/2}, \tag{26}$$

$$dF=R_F[(R_F^2+R_{F\theta}^2)\sin^2\theta+R_{F\varphi}^2]^{1/2}d\theta d\varphi, \tag{27}$$

$$dW=R_W(R_W^2+R_{W\theta}^2)^{1/2}\sin\theta d\theta d\varphi, \tag{28}$$

$$\cos\gamma_F=\mathbf{N}_F\cdot\mathbf{e}_R=\varepsilon R_F\sin\theta/[(R_F^2+R_{F\theta}^2)\sin^2\theta+R_{F\varphi}^2]^{1/2} \tag{29}$$

$$dC=R_F|_{\theta=\overline{\theta}}\sin\overline{\theta}d\varphi, \tag{30}$$

$$\cos\theta_C'=(\mathbf{N}_F\cdot\mathbf{N}_W)_{\theta=\overline{\theta}}. \tag{31}$$

We then substitute Eqs. (25)–(31) into Eq. (18), use Eq. (2), and employ the linear approximation. We can thus obtain

$$\rho_f\int_0^{2\pi}\int_0^{\overline{\theta}}\varepsilon\int_{R_M}^{R_W}\nabla^2\phi\,\delta\phi R^2\sin\theta dR d\theta d\varphi-\rho_f\int_0^{2\pi}\int_0^{\overline{\theta}}\varepsilon\left(\left.\frac{\partial\phi}{\partial R}\right|_{R=R_W}-\frac{R_{W\theta}}{R_W^2}\left.\frac{\partial\phi}{\partial\theta}\right|_{R=R_W}\right)\delta\phi|_{R=R_W}R_W^2\sin\theta d\theta d\varphi$$
$$+\rho_f\int_0^{2\pi}\int_0^{\overline{\theta}}\varepsilon\left(\left.\frac{\partial\phi}{\partial R}\right|_{R=R_M}-\frac{R_{M\theta}}{R_M^2}\left.\frac{\partial\phi}{\partial\theta}\right|_{R=R_M}-\frac{\partial\zeta}{\partial t}\right)\delta\phi|_{R=R_M}R_M^2\sin\theta d\theta d\varphi+\int_0^{2\pi}\int_0^{\overline{\theta}}\left\{\varepsilon(p_g-p_C)\right.$$
$$+\rho_f g[R_M(\overline{\theta})\cos\overline{\theta}-R_M\cos\theta]-\sigma S_{0M}(\theta)+\varepsilon\rho_f\left.\frac{\partial\phi}{\partial t}\right|_{R=R_M}+\varepsilon\rho_f\ddot{f}(t)(h-\varepsilon R_M\cos\theta)-\rho_f\zeta\cos\theta[\ddot{f}(t)+g]$$
$$\left.-\sigma\left[S_{1M}(\theta)\zeta+S_{2M}(\theta)\frac{\partial\zeta}{\partial\theta}+S_{3M}(\theta)\frac{\partial^2\zeta}{\partial\theta^2}+S_{4M}(\theta)\frac{\partial^2\zeta}{\partial\varphi^2}\right]\right\}\delta\zeta R_M^2\sin\theta d\theta d\varphi$$
$$+\int_0^{2\pi}\left\{\sigma\cos\theta_C+\sigma_1-\sigma_2-\varepsilon\sigma\left[R_M(R_M^2+R_{M\theta}^2)^{-3/2}\left(R_M\frac{\partial\zeta}{\partial\theta}-R_{M\theta}\zeta\right)\right]_{\theta=\overline{\theta}}\right\}\delta\zeta|_{\theta=\overline{\theta}}R_M(\overline{\theta})\sin\overline{\theta}d\varphi=0, \tag{32}$$

where

$$S_{0M}(\theta)=R_M^{-1}(R_M^2+R_{M\theta}^2)^{-3/2}[2R_M^3+3R_MR_{M\theta}^2-R_M^2R_{M\theta\theta}$$
$$-R_{M\theta}(R_M^2+R_{M\theta}^2)\cot\theta],$$
$$S_{1M}(\theta)=R_M^{-2}(R_M^2+R_{M\theta}^2)^{-5/2}[-2R_M^5-5R_M^3R_{M\theta}^2+2R_M^4R_{M\theta\theta}$$
$$-R_M^2R_{M\theta}^2R_{M\theta\theta}+R_{M\theta}(2R_M^4+3R_M^2R_{M\theta}^2+R_{M\theta}^4)\cot\theta], \tag{33}$$
$$S_{2M}(\theta)=R_M^{-1}(R_M^2+R_{M\theta}^2)^{-5/2}[3R_MR_{M\theta}(R_MR_{M\theta\theta}-R_{M\theta}^2)$$
$$-R_M^2(R_M^2+R_{M\theta}^2)\cot\theta],$$
$$S_{3M}(\theta)=-R_M(R_M^2+R_{M\theta}^2)^{-3/2},$$
$$S_{4M}(\theta)=-(\sin\theta)^{-2}R_M^{-1}(R_M^2+R_{M\theta}^2)^{-1/2}.$$

The static terms found from the fourth and fifth integrands of Eq. (32)

$$\varepsilon(p_g-p_C)+\rho_f g[R_M(\overline{\theta})\cos\overline{\theta}-R_M\cos\theta]-\sigma S_{0M}(\theta)$$
$$=0 \quad \text{on} \quad M, \tag{34}$$
$$\sigma\cos\theta_C+\sigma_1-\sigma_2=0 \quad \text{along} \quad C_{st} \tag{35}$$

coincide with the formerly derived equilibrium conditions (13) (Eq. (*B*3)) and (14) used to determine the meniscus shape. In the subsequent analysis, Eq. (32) is used as the basic equation for the dynamic slosh analysis.

6.3 Dimensionless Parameters. For convenience in the subsequent analysis and numerical calculation, the following dimensionless quantities are introduced:

$$\phi=\frac{\phi^*}{(b^*)^2\omega_{ch}{}^*},\quad \zeta=\frac{\zeta^*}{b^*},\quad \ddot{f}(t)=\frac{\ddot{f}^*(t^*)}{b^*(\omega_{ch}{}^*)^2},$$
$$\{R,R_M(\theta),R_W(\theta)\}=\frac{\{R^*,R_M{}^*(\theta),R_W{}^*(\theta)\}}{b^*},\quad t=\omega_{ch}{}^*t^*,$$
$$\{S_{1M}(\theta),S_{2M}(\theta),S_{3M}(\theta),S_{4M}(\theta)\}$$
$$=(b^*)^2\{S_{1M}{}^*(\theta),S_{2M}{}^*(\theta),S_{3M}{}^*(\theta),S_{4M}{}^*(\theta)\},$$
$$\{dF,dW\}=\{dF^*,dW^*\}/(b^*)^2,$$
$$\omega=\omega^*/\omega_{ch}{}^*, \tag{36}$$
$$\text{Bo}=\frac{\rho_f{}^*g^*(b^*)^2}{\sigma^*}, \tag{37}$$

where b^* (Fig. 1) and $\omega_{ch}{}^*$ are the characteristic length and frequency, respectively. Note that dimensional quantities are distinguished from the corresponding nondimensional ones by adding an asterisk; i.e., an asterisk should be added to all the dimensional quantities, that have been used hitherto. Equation (37) defines the Bond number Bo, which is the dimensionless parameter relating the magnitude of gravity to surface tension. The characteristic frequency is defined by

$$\omega_{ch}{}^*=(g^*/b^*)^{1/2} \quad \text{for Bo}\neq 0 \quad (g^*\neq 0) \tag{38a}$$
$$\omega_{ch}{}^*=[\sigma^*/\rho_f{}^*(b^*)^3]^{1/2} \quad \text{for Bo}=0. \tag{38b}$$

6.4 Free Vibration Analysis. As a preliminary step to the forced vibration analysis, the modal functions for ϕ and ζ are determined by a free vibration analysis. The solution to the Laplace equation

$$\frac{\partial^2\phi}{\partial R^2}+\frac{2}{R}\frac{\partial\phi}{\partial R}+\frac{1}{R^2}\frac{\partial^2\phi}{\partial\theta^2}+\frac{\cot\theta}{R^2}\frac{\partial\phi}{\partial\theta}+\frac{1}{R^2\sin^2\theta}\frac{\partial^2\phi}{\partial\varphi^2}=0 \tag{39}$$

must be expressed by a linear combination of characteristic functions whose orthogonality is satisfied within the range $0\leqslant\theta\leqslant\overline{\theta}$. Since $\overline{\theta}<\pi/2$ (see Fig. 1), such orthogonality cannot be satisfied by the associated Legendre functions commonly used for spherical coordinates defined in the range $0\leqslant\theta\leqslant\pi$. Hence, the characteristic functions must be derived anew here. Assuming a solution in terms of separated variables

$$\phi(R,\theta,\varphi,t)=X(R)\Theta(\theta)\cos m\varphi e^{i\omega t} \tag{40}$$

and substituting Eq. (40) into (39) leads to

$$\frac{d^2X}{dR^2}+\frac{2}{R}\frac{dX}{dR}-\frac{\lambda}{R^2}X=0 \tag{41}$$

$$\frac{d^2\Theta}{d\theta^2}+\cot\theta\frac{d\Theta}{d\theta}+\left(\lambda-\frac{m^2}{\sin^2\theta}\right)\Theta=0 \tag{42}$$

where λ is the characteristic value to be determined later. Substituting $X=R^\alpha$ into Eq. (41) yields

$$\alpha(\alpha+1)-\lambda=0. \tag{43}$$

The solution to Eq. (42) is given by (Appendix B of [14])

$$\Theta(\theta)=(1-\xi^2)^{m/2}F(m-\alpha,\alpha+m+1,m+1,(1-\xi)/2)$$
$$=(1-\xi^2)^{m/2}$$
$$\times\left\{1+\sum_{i=1}^{\infty}\frac{(m-\alpha)(m-\alpha+1).....(m-\alpha+i-1)}{1\times2\times....\times i}\right.$$
$$\left.\times\frac{(\alpha+m+1)(\alpha+m+2).....(\alpha+m+i)}{(m+1)(m+2).....(m+i)}\left(\frac{1-\xi}{2}\right)^i\right\} \tag{44}$$

where $\xi=\cos\theta$ and F denotes the Gaussian hypergeometric series. The characteristic value $\lambda=\lambda_k$ ($k=1,2,\ldots$) and the corresponding exponents (see Eq. (43))

$$\alpha_{1k}=\frac{-1-(1+4\lambda_k)^{1/2}}{2},\quad \alpha_{2k}=\frac{-1+(1+4\lambda_k)^{1/2}}{2} \tag{45}$$

are determined using the boundary condition

$$\frac{d\Theta}{d\theta}=0 \quad \text{at} \quad \theta=\overline{\theta}. \tag{46}$$

This can be derived by considering the limit $\theta\to\overline{\theta}$ of the boundary condition on the container wall found from the second integrand of the variational principle ((32))

$$\left.\left(\frac{1}{R_{W\theta}}\frac{\partial\phi}{\partial R}-\frac{1}{R_W^2}\frac{\partial\phi}{\partial\theta}\right)\right|_{R=R_W(\theta)}=0 \quad (0\leqslant\theta\leqslant\overline{\theta}) \tag{47}$$

since $|R_{W\theta}|\to\infty$ as $\theta\to\overline{\theta}$ (see Fig. 1). Thus, the boundary condition for the characteristic function $\Theta_k(\theta)$ can be determined by the kinematic condition at the contact line $\theta=\overline{\theta}$, i.e., only (46), instead of the condition throughout the container wall, i.e., (47). Therefore, irrespective of the generatrix shape of the container, the characteristic function $\Theta_k(\theta)$ can be analytically determined by solving the boundary value problem constituted by Eqs. (42) and (46). In other words, the local satisfaction (46) of the boundary condition (47) determines the characteristic functions and allows us to employ the Galerkin method to satisfy the overall condition (47) by constituting an admissible function of the velocity potential using the characteristic functions (Eq. (48)).

Expressing $\Theta_k(\theta)$ in terms of the Gaussian hypergeometric series (Eq. (44)) is helpful for examining the convergence of the series solution. The Gaussian hypergeometric series $F(\alpha,\beta,\gamma,x)$ converges for arbitrary values of α, β, and γ, provided that $|x|<1$. Hence, solution (44) converges for $0\leqslant\theta<\pi$. In the present analysis, we have $0\leqslant\theta\leqslant\overline{\theta}$, where $\overline{\theta}<\pi/2$, as can be seen from Fig. 1. Therefore, convergence of solution (44) is ensured.

By linear combination of the characteristic functions, the velocity potential and the liquid surface displacement can be expressed as

$$\phi(R,\theta,\varphi,t)=i\omega\sum_{k=1}^{\infty}\left\{a_k\left(\frac{R}{l_a}\right)^{a_{1k}}+b_k\left(\frac{R}{l_b}\right)^{\alpha_{2k}}\right\}\Theta_k(\theta)\cos m\varphi e^{i\omega t}, \tag{48}$$

$$\zeta(\theta,\varphi,t)=\sum_{k=1}^{\infty}c_k\Theta_k(\theta)\cos m\varphi e^{i\omega t} \tag{49}$$

where a_k, b_k, and c_k are arbitrary real constants and l_a and l_b are normalization parameters introduced to improve the convergence of series (48). Namely, with the increase of k, $\lambda_k\to\infty$, i.e., $\alpha_{1k}\to-\infty$ and $\alpha_{2k}\to\infty$ (Eq. (45)), so that l_a and l_a are, respectively, the minimum and maximum of R considered, which is $R_M(\theta)$ and $R_W(\theta)$ $(0\leqslant\theta\leqslant\overline{\theta})$.

Equation (49) is obtained by imposing

$$\left.\frac{\partial\zeta}{\partial\theta}\right|_{\theta=\overline{\theta}}=0 \tag{50}$$

as a dynamical contact line condition. This is an approximate method, but it introduces only a small error even for a small static contact angle θ_C by virtue of the special way the spherical coordinates are used in the present study (Appendix C of [14]).

We can obtain the frequency equation by substituting Eqs. (48) and (49) into (32) and neglecting the excitation term. Using the Galerkin method while considering variation with respect to a_k, b_k, and c_k yields algebraic homogeneous equations for a_k, b_k, and c_k. These can be reduced to a standard eigenvalue problem for c_k:

$$(-\omega^2\mathbf{M}+\mathbf{K})\mathbf{c}=\mathbf{0} \tag{51}$$

where

$$\mathbf{c}=\{c_1,c_2,c_3,\ldots c_{\overline{k}}\}^T. \tag{52}$$

Representations of $\mathbf{M}$ and $\mathbf{K}$ in terms of the nondimensional quantities introduced in Eqs. (36) and (37) are presented in the author's previous report ([15]). The dimension of Eq. (51) is $\overline{k}$, at which the summation in Eqs. (48) and (49) is truncated. The solution to Eq. (51) gives the eigenfrequencies and the eigenmodes.

The dimension $\overline{k}$ of the eigenvalue problem (51) required for obtaining a sufficiently converged solution is low ([15]) due to the orthogonality of the characteristic function Θ_k, and therefore the present analysis requires only a small amount of computation time and cost. The analytical derivation of the characteristic function Θ_k and the rapid convergence of its constituting series (Eq. (44)) are helpful for fast computation.

6.5 Modal Equation for Axial Excitation. The purpose here is to study the parametric resonance for the case where the excitation frequency is close to twice the fundamental modal frequency with $m=1$. Hence, we may express ϕ and ζ in terms of their fundamental modal functions (48) and (49) determined in Section 6.4:

$$\phi(R,\theta,\varphi,t)=\dot{q}(t)\sum_{k=1}^{\infty}\left\{a_k\left(\frac{R}{l_a}\right)^{\alpha_{1k}}+b_k\left(\frac{R}{l_b}\right)^{\alpha_{2k}}\right\}\Theta_k(\theta)\cos\varphi, \tag{53}$$

$$\zeta(\theta,\varphi,t)=q(t)\sum_{k=1}^{\infty}c_k\Theta_k(\theta)\cos\varphi, \tag{54}$$

where $q(t)$ denotes the modal coordinate. The ratios among the constants a_k, b_k, and c_k are determined in the free vibration analysis. Substituting Eqs. (53) and (54) into the variational principle (32) and considering the variation with respect to $q(t)$ leads to the following modal equation

$$M_0\ddot{q}+[K_0+K_1\ddot{f}(t)]q=0, \tag{55}$$

where M_0, K_0, and K_1 are given in Appendix D. Introducing the modal damping ratio ζ_d modifies Eq. (55) to

$$\ddot{q}+2\zeta_d\omega\dot{q}+[\omega^2+Q\ddot{f}(t)]q=0 \tag{56}$$

where

$$\omega^2=\frac{K_0}{M_0},\quad Q=\frac{K_1}{M_0}. \tag{57}$$

6.6 Critical Value Q_{cr}. The amplitude of the liquid surface displacement can increase with time due to the parametric resonance. Let us estimate the critical value of Q, above which the response increases with time, under the condition that the excitation frequency $2\omega_f$ is close to 2ω. We express the excitation and solution as

$$\ddot{f}(t)=f_0\sin 2\omega_f t, \tag{58a}$$

$$q(t)=\alpha\cos\omega_f t+\beta\sin\omega_f t, \tag{58b}$$

$$\dot{q}(t)=-\omega_f\alpha\sin\omega_f t+\omega_f\beta\cos\omega_f t, \tag{58c}$$

$$\ddot{q}(t)=(-\omega_f^2\alpha+\omega_f\dot{\beta})\cos\omega_f t-(\omega_f^2\beta+\omega_f\dot{\alpha})\sin\omega_f t, \tag{58d}$$

where

$$\omega_f\cong\omega \tag{59}$$

and α and β are gradually varying time-dependent functions. Substituting Eq. (58) into (56) and using a harmonic balance method leads to the following system of differential equations for α and β:

$$(\omega^2-\omega_f^2)\alpha+\omega_f\dot{\beta}+\left(2\zeta_d\omega_f\omega+\frac{1}{2}Qf_0\right)\beta=0, \tag{60a}$$

$$-\omega_f\dot{\alpha}+\left(-2\zeta_d\omega_f\omega+\frac{1}{2}Qf_0\right)\alpha+(\omega^2-\omega_f^2)\beta=0. \tag{60b}$$

Substituting

$$(\alpha,\beta)=(\alpha_0,\beta_0)e^{st} \tag{61}$$

into Eq. (60) yields the following characteristic equation for s:

$$As^2+Bs+C=0, \tag{62}$$

where

$$A=\omega_f^2, \tag{63a}$$

$$B=4\zeta_d\omega\omega_f^2, \tag{63b}$$

$$C=(2\zeta_d\omega_f\omega)^2-\left(\frac{1}{2}Qf_0\right)^2+(\omega_f^2-\omega^2)^2. \tag{63c}$$

The system is unstable when either of the roots of Eq. (62) has a positive real part ($B>0$ excludes the case where both of the roots have positive real parts). Hence, the condition $C<0$ gives the following critical value:

$$Q_{cr}=\frac{4\zeta_d\omega^2}{f_0} \tag{64}$$

under the approximation (59). The unstable root can be calculated from Eq. (62) as

$$s_1=-2\zeta_d\omega+\frac{Qf_0}{2\omega}. \tag{65}$$

7 Numerical Results

Figures 2 and 3 show the meniscus shape and the eigenfrequency, respectively, for the case of a spherical container and a contact angle of five deg between meniscus and container wall.

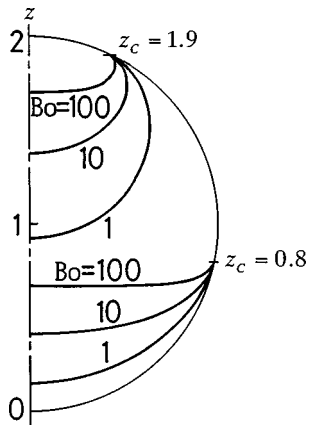


Fig. 2 Shape of meniscus for various Bond numbers and dimensionless *z*-coordinates of contact line

These results were shown and discussed in a previous paper ([14]). They are presented again here in order to facilitate the discussion of the numerical results for the liquid surface displacement given below. It can be seen from Fig. 2 that when the Bond number Bo is large, the menisci are almost flat except in the vicinity of the container wall. The menisci tend to exhibit more spherical shapes when the Bond number is decreased.

Figure 4 shows the eigenfrequency for the case Bo=0 ($g^*=0$, see Eq. (37)). Because the characteristic frequency $(g^*/b^*)^{1/2}$ used in Fig. 3 becomes zero in this case, an alternative characteristic frequency $[\sigma^*/\rho_f^*(b^*)^3]^{1/2}$ is employed.

In Appendix E, the influence of the contact angle on the eigenfrequency is examined for the cases of Bo=0 and Bo=1.

As a test for the accuracy of the present method, the numerical results obtained in the present study are compared with results in previous literature. The meniscus shape determined by this analysis is shown in a figure taken from a paper by Dodge et al. [10] (Fig. 5). Complete agreement can be confirmed. In Fig. 6, the present theoretical predictions for the eigenfrequency are compared with the previous theoretical and experimental results. Good agreement can be observed. The agreement between the theoretical results obtained by the author and by Concus et al. [8] is so good that the two curves cannot be distinguished.

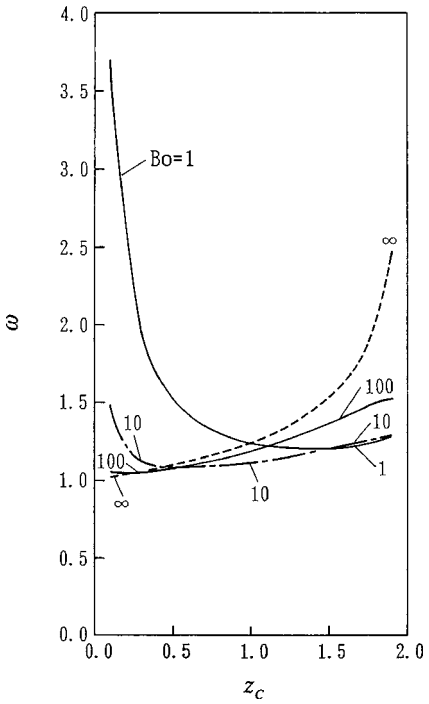


Fig. 3 Dimensionless eigenfrequency $\omega=\omega^*/\omega^*_{ch}$ (ω^* is the dimensional eigenfrequency, ω^*_{ch} is the characteristic frequency given by $\omega^*_{ch}=(g^*/b^*)^{1/2}$)

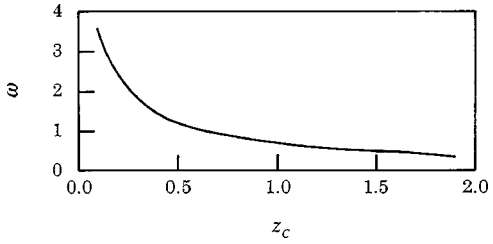


Fig. 4 Dimensionless eigenfrequency $\omega=\omega^*/\omega^*_{ch}$ for the case Bo=0 (ω^* is the dimensional eigenfrequency, ω^*_{ch} is the characteristic frequency given by $\omega^*_{ch}=[\sigma^*/\rho_f^*(b^*)^3]^{1/2}$

To consider the reason for these results, the following two parameters

$$\frac{Q}{Q_{cr}}=\frac{Qf_0}{4\zeta_d\omega^2}, \tag{66}$$

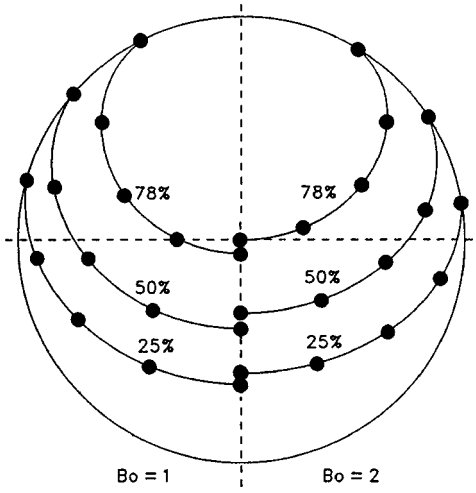


Fig. 5 Comparison of the present and the previous results for the meniscus shape (78 percent, 50 percent, and 25 percent filling levels for Bo=1 and Bo=2; θ_c=5 deg; ●, present analysis; —— analysis by Dodge et al. [10]

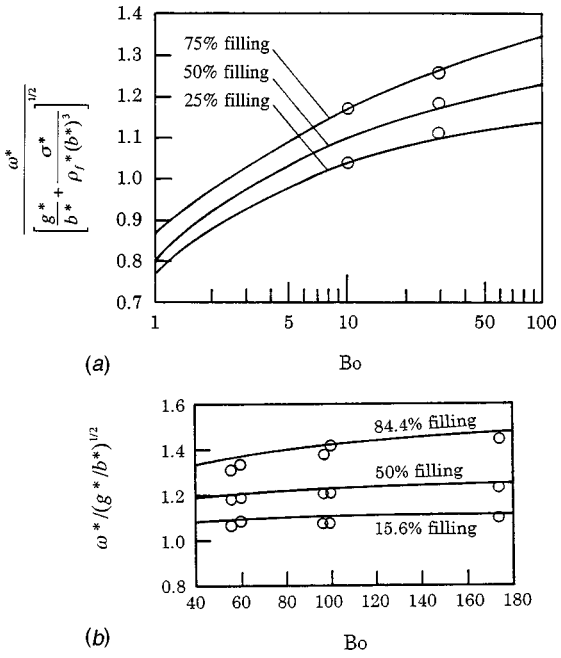


Fig. 6 Comparison of the present and the previous results for the eigenfrequency. (*a*) ——, present analysis and theoretical prediction by Concus et al. [8] both for $\theta_c=0$ deg; ●, experiment by Coney and Salzman [17]. (*b*) ——, present analysis; ●, experiment by Dodge and Garza [9].

$$\frac{s_1}{\omega}=-2\zeta_d+\frac{Qf_0}{2\omega^2} \tag{67}$$

are calculated from Eqs. (64) and (65) and are shown in Figs. 8 and 9, respectively. It can be seen that the variations of these parameters with z_C and Bo exhibit a tendency similar to the dependence of the magnitude of the response on z_C and Bo shown in Fig. 7. Independence of the parameters (66) and (67) of z_C for the case Bo$\rightarrow\infty$ can be explained as follows. For the case Bo$\rightarrow\infty$, the relation $K_0=K_1$ holds as can be seen from Eqs. (*D*2) and (*D*3), so that from Eq. (57)

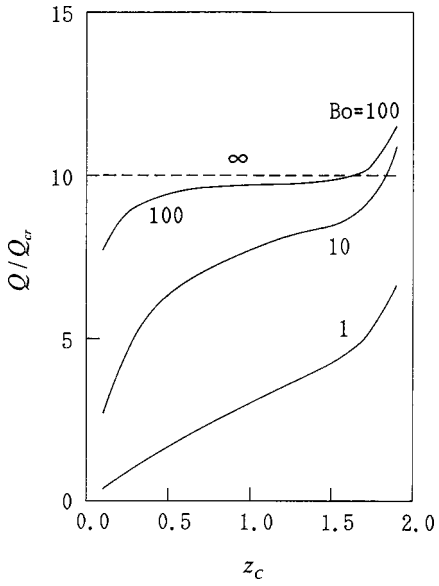


Fig. 8 Coefficient of excitation term Q in modal Eq. (56) normalized by its critical value Q_{cr}

$$Q=\omega^2 \quad \text{for Bo}\rightarrow\infty. \tag{68}$$

Therefore, the two parameters (66) and (67) do not depend on z_C. For small Bo and z_C, Q/Q_{cr} is smaller than unity (Fig. 8) and s_1 is negative (Fig. 9). Hence, the system is stable and the magnitude of the response shown in Fig. 7 is smaller than the prescribed initial value $\zeta(\bar{\theta},0,0)=0.01$.

Let us consider the physical reason why the analysis for Bo$\rightarrow\infty$, which neglects the surface tension effect, underestimates the magnitude of the dynamical displacement of the liquid surface for large z_C. We attribute the reason to the marked decrease of ω^2 in Eqs. (66) and (67) with the decrease of Bo from infinity (see Fig. 3). This decrease in ω^2 is due to the fact that the kinetic energy of the liquid increases and M_0 in Eq. (57) increases. To comprehend this intuitively, we express the kinetic energy in terms of a surface integral over the meniscus using Green's theorem and Eq. (20):

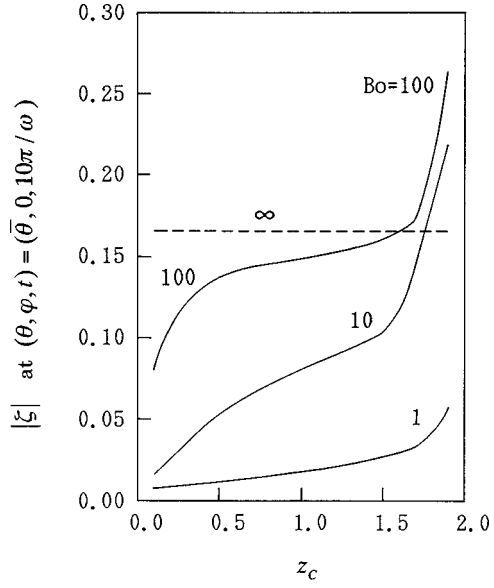


Fig. 7 Dimensionless magnitude of liquid surface displacement $|\zeta|$ at $(\theta,\varphi,t)=(\bar{\theta},0,10\pi/\omega)$

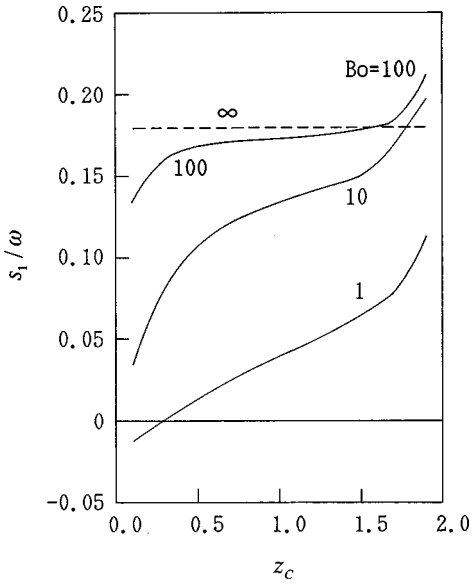


Fig. 9 Characteristic root s_1 normalized by dimensionless eigenfrequency ω

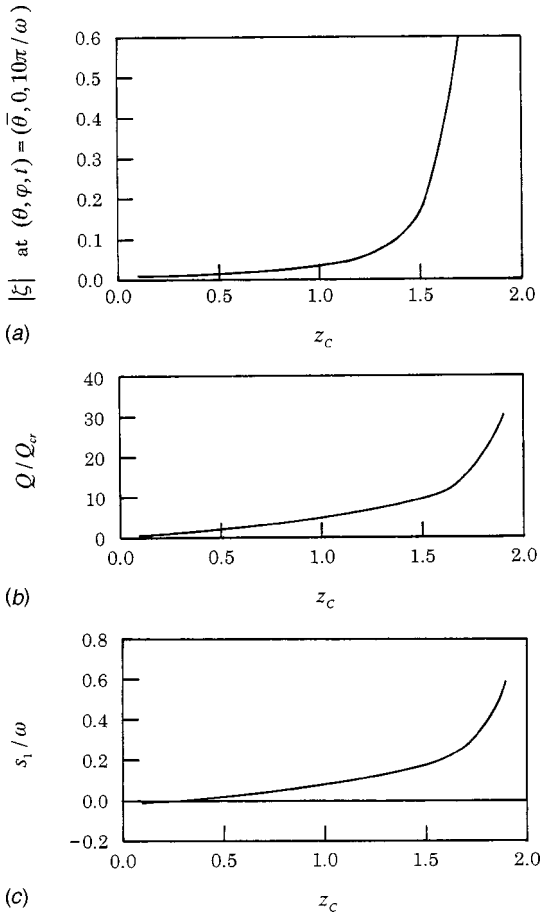


Fig. 10 Results for the case Bo=0; (*a*) dimensionless magnitude of liquid surface displacement $|\zeta|$ at $(\theta,\varphi,t)=(\bar{\theta},0,10\pi/\omega)$; (*b*) coefficient of excitation term Q in modal Eq. (56) normalized by its critical value Q_{cr}; (*c*) characteristic root s_1 normalized by dimensionless eigenfrequency ω

$$\frac{1}{2}\int\int\int_V (\nabla\phi)^2 dV=\frac{1}{2}\iint_{M+W}\phi\frac{\partial\phi}{\partial N}dS=\frac{1}{2}\int\int_M\phi\frac{\partial\phi}{\partial N}dM. \tag{69}$$

The value of the integral is greatly influenced by the area of the meniscus. The area becomes much larger for a finite Bond number than for an infinite one, for which the meniscus is reduced to the plane $z=z_C$ (see Fig. 2). It should be noted here that the large difference in the area for finite and infinite Bond numbers is due to the very small contact angle between the meniscus and the container wall ([16]) and the large tilt of the container wall from the vertical direction (i.e., parallel to the *z*-axis) near $z=2$.

Figure 10 shows the results for the case where the Bond number is exactly zero and the amplitude of the axial acceleration is $0.5b^*(\omega^*_{ch})^2=0.5\sigma^*/\rho^*_f(b^*)^2$ (see Eq. (38*b*)). Here again, it can be confirmed that the dependence of the magnitude of the liquid surface displacement on z_C (Fig. 10(*a*)) can be explained by the variations in the parameters (66) and (67) with z_C (Figs. 10(*b*) and 10(*c*)). For large z_C, the increase in these parameters with the increase in z_C is more marked for Bo=0 than for Bo$\neq$0 (Figs. 8 and 9). This is because for Bo=0, ω^2 in the denominators on the right sides of Eqs. (66) and (67) decreases monotonically with the increase in z_C (Fig. 4). Note that the effect of ω is raised to the second power in Eqs. (66) and (67). On the other hand, even for a relatively small but nonzero Bond number Bo=1, the eigenfrequency ω increases up with increasing z_C for large z_C (Fig. 3). This is due to the fact that the remaining gravity effect is accentuated compared to the surface tension effect when z_C is increased, and consequently the ratio [liquid volume]/[meniscus area] becomes large. Monotonic increase in the natural frequency with the increase in the liquid depth is a typical characteristic of gravity waves, as can be seen from the result for the case Bo$\rightarrow\infty$ (Fig. 3).

8 Summary and Conclusions

The problem of low-gravity propellant sloshing in an axisymmetrical container subjected to axial excitation is formulated by a variational principle and is solved analytically by a modal analysis method. The use of spherical coordinates enables us to analytically derive the characteristic functions for an arbitrary axisymmetrical container. Numerical results show that neglecting the surface tension leads to failure to predict the dependency of the response of the liquid surface displacement on the liquid filling level and underestimation of the magnitude of the liquid surface oscillation for high liquid filling levels. This result is explained by the influences of the Bond number and the liquid filling level on the critical value of the coefficient of the excitation term in the modal equation and the unstable characteristic root.

This paper did not consider the case in which all of the internal wall surface of the container contacts the liquid and the liquid surrounds the gas completely. The present method can be extended to such a case by altering the position of the origin of the spherical coordinates to an arbitrary position within the gas region and by using the associated Legendre polynomials, which are orthogonal within $0\leqq\theta<\pi$, instead of the characteristic functions derived in the present study. Such an extension is planned as a future work by the author.

Nomenclature

a_k, b_k = coefficients in modal functions for velocity potential (Eq. (48))
Bo = Bond number
b = half height of container (characteristic length, Fig. 1)
C = contact line
c_k = coefficients in modal functions for liquid surface displacement (Eq. (49))
$\mathbf{e}_R$, $\mathbf{e}_\theta$, $\mathbf{e}_\varphi$ = unit vectors in R, θ, and φ-directions
F = disturbed liquid surface (Fig. 1)
$\ddot{f}(t)$ = acceleration of container in *z*-direction
g = gravitational acceleration
h = *z*-coordinate of the origin O of spherical coordinates
k = characteristic value and function number
$\overline{k}$ = number of characteristic functions Θ_k taken into account
l_a, l_b = normalization parameters (Eq. (48))
M = meniscus (undisturbed liquid surface, Fig. 1)
$\mathbf{N}_F$ = 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
$p_{l,st}$ = static 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 container wall (Fig. 1)
$R_{W\theta}$ = $dR_W/d\theta$
$S_{0M}-S_{4M}$ = θ-dependent functions (Eq. (33))
V = liquid domain (Fig. 1)
W = container wall (Fig. 1)
W_1, W_2 = liquid-solid and gas-solid interfaces, respectively
z_C = z-coordinate of contact line
z_0 = z-coordinate of meniscus bottom ($\theta=0$)
α_{1k}, α_{2k} = characteristic exponents (Eqs. (45) and (48))
γ_F = angle between $\mathbf{N}_F$ and R-direction
γ_M = angle between $\mathbf{N}_M$ and R-direction
ε = 1 and -1, respectively, for Cases 1 and 2 (Fig. 1)
ζ = liquid surface displacement (Fig. 1)
Θ_k = characteristic function determined by Eqs. (42) and (46)
$\overline{\theta}$ = maximum value of θ (Fig. 1)
θ_C = contact angle between meniscus and container wall
θ'_C = contact angle between disturbed liquid surface and container wall
λ_k = characteristic value determined by Eqs. (42) and (46)
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 (motion relative to the container)
ω = eigenfrequency
ω_{ch} = characteristic frequency

Appendix A

Derivation of the First Term on the Right Side of Eq. (11). This derivation is based on the fact that the variation $\delta(dM)$ in the surface element due to the virtual displacement $\delta R_M \cos\gamma_M$ in the direction normal to the meniscus can be expressed by

$$\delta(dM)=\operatorname{div}\mathbf{N}_M\delta R_M\cos\gamma_M dM. \tag{A1}$$

This can be proved by expressing $\operatorname{div}\mathbf{N}_M$ as the limit of a surface integral, i.e.,

$$\operatorname{div}\mathbf{N}_M=\lim_{D\to 0}\left\{\oint\!\!\!\oint_A \mathbf{N}_M\cdot\mathbf{n}dA/D\right\} \tag{A2}$$

where D is an arbitrary volume including point P in which the vector $\mathbf{N}_M$ is erected, A is the closed surface bounding volume D, and $\mathbf{n}$ is the outer normal unit vector of surface A. Let volume D be the domain which the surface dM penetrates during the virtual displacement $\delta R_M\cos\gamma_M$ from dM to $dM'=dM+\delta(dM)$ (see Fig. 11). Then, the inner product $\mathbf{N}_M\cdot\mathbf{n}$ is -1 on dM, $\cos(\delta R_{M\theta}d\theta)\cong 1$ on dM', and 0 on $A-dM-dM'$ (i.e., on the entire remaining portion of A, except dM and dM'), while $D=dM\,\delta R_M\cos\gamma_M$. So, Eq. (A2) gives

$$\operatorname{div}\mathbf{N}_M=\frac{dM'-dM}{dM\,\delta R_M\cos\gamma_M}=\frac{\delta(dM)}{dM\,\delta R_M\cos\gamma_M}, \tag{A3}$$

which is identical to Eq. (A1).

Appendix B

Derivation of Eq. (15) From Eq. (13). Equation (1) can be expressed as

$$f(R,\theta)=R-R_M(\theta)=0. \tag{B1}$$

Therefore $\mathbf{N}_M$ can be determined by

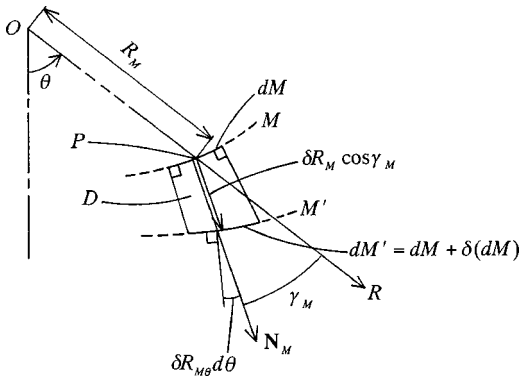


Fig. 11 Virtual displacement $\delta R_M\cos\gamma_M$ in direction normal to the meniscus M considered for derivation of Eq. (A1)

$$\mathbf{N}_M=\varepsilon\,\operatorname{grad}f/|\operatorname{grad}f|=\varepsilon(\mathbf{e}_R R_M-\mathbf{e}_\theta R_{M\theta})/(R_M^2+R_{M\theta}^2)^{1/2}. \tag{B2}$$

Substituting Eqs. (9) and (B2) into Eq. (13) gives

$$\begin{aligned}\varepsilon(p_g-p_C)+\rho_f g[R_M(\overline{\theta})\cos\overline{\theta}-R_M\cos\theta]\\ -\sigma R_M^{-1}(R_M^2+R_{M\theta}^2)^{-3/2}[2R_M^3+3R_MR_{M\theta}^2\\ -R_M^2R_{M\theta\theta}-R_{M\theta}(R_M^2+R_{M\theta}^2)\cot\theta].\end{aligned} \tag{B3}$$

Solving Eq. (B3) with respect to $R_{M\theta\theta}$ leads to Eq. (15).

Appendix C

Proof of Equivalence of Eq. (15) to Young-Laplace Equation. The Young-Laplace equation is given by

$$p_g-p_{l,st}=\sigma\left(\frac{1}{R_1}+\frac{1}{R_2}\right), \tag{C1}$$

where R_1 and R_2 are the principal radii of the meniscus. The directions of the principal curvatures are the θ and φ-directions, due to the axisymmetry of the meniscus. Hence, we have

$$\begin{aligned}\frac{1}{R_1}&=\frac{d}{ds}\left(\tan^{-1}\frac{dz}{dr}\right)=\frac{d}{d\theta}\left(\tan^{-1}\frac{dz/d\theta}{dr/d\theta}\right)\cdot\frac{d\theta}{ds}\\ &=\frac{d}{d\theta}\left(\tan^{-1}\frac{dz/d\theta}{dr/d\theta}\right)\cdot\left[\left(\frac{dr}{d\theta}\right)^2+\left(\frac{dz}{d\theta}\right)^2\right]^{-1/2},\end{aligned} \tag{C2}$$

$$\frac{1}{R_2}=\frac{1}{r}\frac{dz}{ds}=\frac{1}{r}\frac{dz}{d\theta}\frac{d\theta}{ds}=\frac{1}{r}\frac{dz}{d\theta}\left[\left(\frac{dr}{d\theta}\right)^2+\left(\frac{dz}{d\theta}\right)^2\right]^{-1/2}, \tag{C3}$$

where s denotes the arc length ($ds=[(dr)^2+(dz)^2]^{1/2}$). On the meniscus, the following relation holds:

$$r=R_M(\theta)\sin\theta,\quad z=h-\varepsilon R_M(\theta)\cos\theta. \tag{C4}$$

Using Eq. (C4) and the formula $(d/dx)\tan^{-1}x=1/(1+x^2)$, we express the principal curvatures (C2) and (C3) in terms of the function $R_M(\theta)$. Substituting the resulting equations into Eq. (C1) leads to Eq. (15).

Appendix D

Coefficients in Modal Equation (55). The coefficients in Eq. (55) are represented by

$$\begin{aligned}M_0=\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\int_0^{\overline{\theta}}\{&a_ka_l\overline{M}_{0aa}(k,l)+a_kb_l[\overline{M}_{0ab}(k,l)+\overline{M}_{0ba}(l,k)]\\ &+b_kb_l\overline{M}_{0bb}(k,l)+a_kc_l[\overline{M}_{0ac}(k,l)+\overline{M}_{0ca}(l,k)]\\ &+b_kc_l[\overline{M}_{0bc}(k,l)+\overline{M}_{0cb}(l,k)]\}d\theta,\end{aligned} \tag{D1}$$

$$K_0=\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\left\{\int_0^{\bar{\theta}} c_k c_l \bar{K}_{0cc}(k,l)\,d\theta + c_k c_l \hat{K}_{0cc}(k,l)\right\}, \quad (D2)$$

$$K_1=\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\int_0^{\bar{\theta}} c_k c_l \bar{K}_{1cc}(k,l)\,d\theta \quad (D3)$$

where

$$\bar{M}_{0\alpha\beta}(k,l)=\varepsilon f_{\alpha k}\left(\frac{R_W}{l_\alpha}\right)\Theta_k \times\sin\theta\left\{R_W^2 f'_{\beta l}\left(\frac{R_W}{l_\beta}\right)\frac{1}{l_\beta}\Theta_l - R_{W\theta} f_{\beta l}\left(\frac{R_W}{l_\beta}\right)\Theta'_l\right\} - \varepsilon f_{\alpha k}\left(\frac{R_M}{l_\alpha}\right)\Theta_k \sin\theta\left\{R_M^2 f'_{\beta l}\left(\frac{R_M}{l_\beta}\right)\frac{1}{l_\beta}\Theta_l - R_{M\theta} f_{\beta l}\left(\frac{R_M}{l_\beta}\right)\Theta'_l\right\}$$

for

$$(\alpha,\beta)=(a,a),(a,b),(b,a),(b,b); \quad (D4)$$

$$\bar{M}_{0ac}(k,l)=\varepsilon R_M^2 \sin\theta f_{ak}\left(\frac{R_M}{l_a}\right)\Theta_k\Theta_l, \quad (D5)$$

$$\bar{M}_{0bc}(k,l)=\varepsilon R_M^2 \sin\theta f_{bk}\left(\frac{R_M}{l_b}\right)\Theta_k\Theta_l, \quad (D6)$$

$$\bar{M}_{0ca}(k,l)=\bar{M}_{0ac}(l,k), \quad (D7)$$

$$\bar{M}_{0cb}(k,l)=\bar{M}_{0bc}(l,k), \quad (D8)$$

$$\bar{K}_{0cc}(k,l)=-R_M^2 \sin\theta\Theta_k[\mu_1\Theta_l\cos\theta+\mu_2(S_{1M}\Theta_l + S_{2M}\Theta'_l + S_{3M}\Theta''_l - m^2 S_{4M}\Theta_l)] \quad (D9)$$

$$\hat{K}_{0cc}(k,l)=-\mu_2[R_M^2(R_M^2+R_{M\theta}^2)^{-3/2}\Theta_k \times\sin\theta(R_M\Theta'_l - R_{M\theta}\Theta_l)]_{\theta=\bar{\theta}}, \quad (D10)$$

$$\bar{K}_{1cc}(k,l)=-R_M^2 \sin\theta\Theta_k\Theta_l\cos\theta, \quad (D11)$$

with

$$f_{ak}(x)=x^{\alpha_{1k}}, f_{bk}(x)=x^{\alpha_{2k}}, \quad (D12)$$

$$\mu_1=1 \ \text{(for Bo}\neq 0), \quad \mu_1=0 \ \text{(for Bo}=0), \quad (D13)$$

$$\mu_2=\frac{1}{\text{Bo}} \ \text{(for Bo}\neq 0), \quad \mu_2=1 \ \text{(for Bo}=0). \quad (D14)$$

Appendix E

Discussion on Eigenfrequency for Small Bond Numbers. Figure 12 examines the influence of the contact angle θ_C on the eigenfrequency for the cases of Bo=0 and Bo=1. It can be seen from Fig. 12 that for Bo=0 and $\theta_C=0$ deg, the fundamental eigenfrequency is close to zero (the eigenvalue, i.e., the square of the eigenfrequency, is smaller, but exact zero eigenvalue was not obtained due to limitation of the numerical accuracy of the computation). This is because the liquid surface can be given a rigid-body displacement that does not change the surface energy. Since such a rigid-body model is approximately valid for Bo=1, $\theta_C=0$ deg, and small values of z_C, the eigenfrequency ω^* for such cases is close to $[\sigma^*/\rho_f^*(b^*)^3]^{1/2}$. This is equal to $[g^*/\text{Bo}\,b^*]^{1/2}=[g^*/b^*]^{1/2}$, which is the eigenfrequency of the pendulum supported at $z^*=b^*$ and having a mass at $z^*=0$. These results for $\theta_C=0$ deg compare well with the results given in Concus et al. [8], which shows the values of $(\omega^*)^2/(1+\text{Bo})[\sigma^*/\rho_f^*(b^*)^3]$. For a contact angle of 5 deg, the eigenfrequency increases markedly with the decrease in the liquid filling level z_C both for Bo=0 and Bo=1. To examine the physical reason for this result, the value of $1/M$ (M is the area of the meniscus) was computed as a function of z_C and shown in Fig. 13. It can be seen from Figs. 12 and 13 that the variations of the eigenfrequency and $1/M$ with z_C exhibit a similar tendency. This means that for the case of $\theta_C=5$ deg, the liquid surface behaves as a membrane subjected to tension and bounded by the container wall. If the eigenfrequency is shown as a function of fill fraction instead of z_C (Fig. 14), the difference between the results for Bo=0 and Bo=1 becomes larger except for extremely low liquid filling levels, where the meniscus is very slender and therefore the influence of the gravity force on the eigenfrequency is very weak.

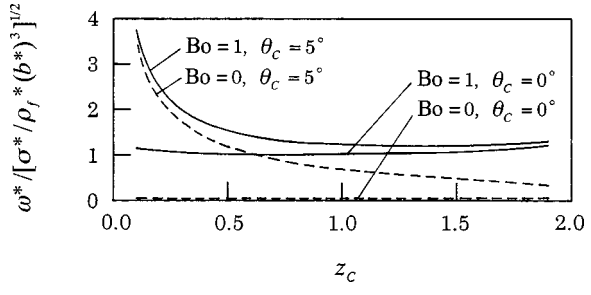


Fig. 12 Influence of contact angle on the eigenfrequency

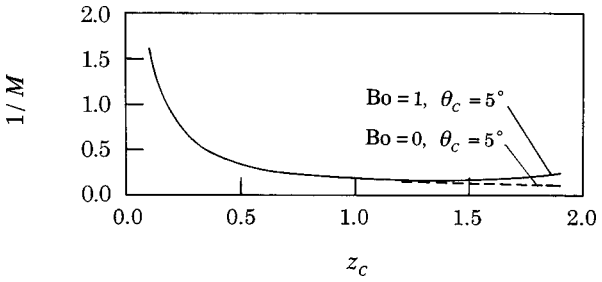


Fig. 13 Value of 1/*M*, where *M* is area of meniscus

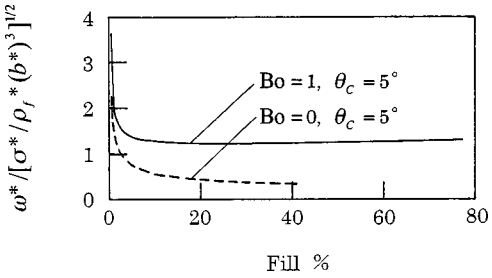


Fig. 14 Eigenfrequency shown as a function of the ratio [liquid volume]/[container volume] within the range $z_C \leqslant 1.9$

References

[1] Abramson, H. N., ed., 1966, ''The Dynamic Behavior of Liquids in Moving Containers,'' NASA SP-106.

[2] Bauer, H. F., and Siekmann, J., 1971, ''Dynamic Interaction of a Liquid with the Elastic Structure of a Circular Cylindrical Container,'' Ing. Arch., **40**, pp. 266–280.

[3] 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.

[4] 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.

[5] 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, Department of Mechanical Engineering, Stanford University, Stanford, CA.
[6] Tong, P., 1967, ''Liquid Motion in a Circular Cylindrical Container With a Flexible Bottom,'' AIAA J., **5**, pp. 1842–1848.
[7] Chu, W. H., 1970, ''Low-Gravity Fuel Sloshing in an Arbitrary Axisymmetric Rigid Tank,'' ASME J. Appl. Mech., **37**, pp. 828–837.
[8] Concus, P., Crane, G. E., and Satterlee, H. M., 1969, ''Small Amplitude Lateral Sloshing in Spheroidal Containers Under Low Gravitational Conditions,'' NASA CR-72500.
[9] 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.
[10] Dodge, F. T., Green, S. T., and Cruse, M. W., 1991, ''Analysis of Small-Amplitude Low Gravity Sloshing in Axisymmetric Tanks,'' Micrograv. Sci. Technol., **4**, pp. 228–234.
[11] 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.
[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] Utsumi, M., 1989, ''The Meniscus and Sloshing of a Liquid in an Axisymmetric Container at Low-Gravity,'' *Sloshing and Fluid Structure Vibration*, D. C. Ma, J. Tani, S. S. Chen, and W. K. Liu, eds., ASME, New York, pp. 103–113.
[14] Utsumi, M., 1998, ''Low-Gravity Propellant Slosh Analysis Using Spherical Coordinates,'' J. Fluids Struct., **12**, pp. 57–83.
[15] Seliger, R. L., and Whitham, G. B., 1968, ''Variational Principles in Continuum Mechanics,'' Proc. R. Soc. London, Ser. A, **305**, pp. 1–25.
[16] Neu, J. T., and Good, R. J., 1963, ''Equilibrium Behavior of Fluids in Containers at Zero Gravity,'' AIAA J., **1**, pp. 814–819.
[17] Coney, T. A., and Salzman, J. A., 1971, ''Lateral Sloshing in Oblate Spheroidal Tanks Under Reduced and Normal Gravity Conditions,'' NASA TN D-6250.

The Probabilistic Solutions to Nonlinear Random Vibrations of Multi-Degree-of-Freedom Systems

G.-K. Er
Assistant Professor,
Faculty of Science and Technology,
University of Macao,
P.O. Box 3001,
Macao, China
e-mail: fstgke@umac.mo

The probability density function of the responses of nonlinear random vibration of a multi-degree-of-freedom system is formulated in the defined domain as an exponential function of polynomials in state variables. The probability density function is assumed to be governed by Fokker-Planck-Kolmogorov (FPK) equation. Special measure is taken to satisfy the FPK equation in the average sense of integration with the assumed function and quadratic algebraic equations are obtained for determining the unknown probability density function. Two-degree-of-freedom systems are analyzed with the proposed method to validate the method for nonlinear multi-degree-of-freedom systems. The probability density functions obtained with the proposed method are compared with the obtainable exact and simulated ones. Numerical results showed that the probability density function solutions obtained with the presented method are much closer to the exact and simulated solutions even for highly nonlinear systems with both external and parametric excitations. [S0021-8936(00)01602-0]

1 Introduction

Practical problems are frequently formulated as multi-degree-of-freedom systems with random excitations. Therefore, the problems of nonlinear random vibrations of multi-degree-of-freedom systems were proposed in many areas of science and engineering. However, it is known that even the solution of a nonlinear single-degree-of-freedom system has challenged many researchers in the past decades. One of the key difficulties in nonlinear random vibrations lies in the determination of the probability density function of system responses because the reliability and some other statistical analyses are based on it. Even if the probability density function solution to nonlinear random vibration is governed by the Fokker-Planck-Kolmogorov (FPK) equation, it is still difficult to obtain the exact probability density function solution if the system is nonlinear or there is parametric excitation. Therefore, some methods were proposed for the approximate probability density function solution to nonlinear random vibrations of multi-degree-of-freedom systems. The most frequently used approximation method is the equivalent linearization or Gaussian closure procedure ([1–4]). However, this method is considered unsuitable when the system is highly nonlinear, or when parametric random excitations are present, because in either case the probability distribution of the system response is usually far from being Gaussian. To improve the accuracy of an approximate solution, a non-Gaussian closure method was used ([5–6]). With this method, the probability density function of the system responses is approximated with the Gram-Charlier series or Hermite polynomial. As is known, the series is not consistent with a probability theory, e.g., negative probability may result. The stochastic average method is another method for the probability density function solution of response amplitudes of nonlinear systems. It is suitable for the weakly damped systems with weak excitations ([7]). The principle of maximum entropy was attempted for an approximate probability density function solution to nonlinear random dynamic systems ([8]); however, highly nonlinear algebraic equations must be formulated and solved in the determination of the parameters with this method and it is difficult to extend for multi-degree-of-freedom problems. Another method called the multi-Gaussian closure method was proposed with which an approximate PDF is constructed with a linear superposition of some Gaussian PDFs ([9]). Such an approximate probability density function satisfies the non-negativity condition. With this method, the approximate probability density function agrees well with exactly one even in the tails of the probability density function. However, highly nonlinear algebraic equations need to be solved, which is a tedious problem. Numerical simulation is versatile ([10–12]), but the computational effort with it is usually unacceptable for estimating the probability density function solution of the system responses, specially for small probability problems. As a result, the problem of nonlinear random vibrations of multi-degree-of-freedom systems or solving the FPK equation in higher dimension have attracted much attention in the last decades.

Recently, a new method was reported for nonlinear random systems and applied to a nonlinear random single-degree-of-freedom system ([13]). In this paper, the method is extended and applied to the probability density function solutions for nonlinear random vibrations of multi-degree-of-freedom systems or the solution to the FPK equation in a higher dimension. With this method, the probability density function of the responses of nonlinear random multi-degree-of-freedom systems is assumed, in a given domain, to be an exponential function of a polynomial in state variables. Special measure is taken such that the FPK equation is satisfied in the average sense of integration with the assumed probability density function. The problem of determining the parameters in the approximate probability density function finally results in solving simultaneous quadratic algebraic equations. The nonlinear random vibrations of two highly nonlinear two-degree-of-freedom systems are analyzed with the proposed method. Numerical results are provided and compared with obtainable exact probability density function solutions to show the effectiveness of the proposed method for highly nonlinear multi-degree-of-freedom systems. In the case that no exact solution is

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 18, 1998; final revision, November 1, 1999. Associate Technical Editor: A. A. Ferri. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

obtainable for a nonlinear two-degree-of-freedom system with both external and parametric excitations, the Monte Carlo simulation is conducted to obtain the simulated probability density function solution. The presented method is further validated by the comparison of the probability density function obtained with the presented method and simulated ones.

2 Problem Formulation

We consider the following problem of nonlinear random vibrations of a multi-degree-of-freedom system:

$$\ddot{Y}_i + h_{i0}(\mathbf{Y},\dot{\mathbf{Y}}) = h_{ij}(\mathbf{Y},\dot{\mathbf{Y}})W_j(t) \quad i=1,2,...,n_y;\ j=1,2,...,m \tag{1}$$

where $\mathbf{Y}\in\Re^{n_y}$, $Y_i\in\Re$, $(i=1,2,...,n_y)$, are components of the vector process $\mathbf{Y}$, $h_{i0}:\Re^{n_y}\times\Re^{n_y}\to\Re$, $h_{ij}:\Re^{n_y}\times\Re^{n_y}\to\Re$, and $W_j(t)$ are random excitations. $h_{i0}(\mathbf{Y},\dot{\mathbf{Y}})$ and $h_{ij}(\mathbf{Y},\dot{\mathbf{Y}})$ are of the polynomial type. Setting $Y_i=X_{2i-1}$, $\dot{Y}_i=X_{2i}$, $d_{2i-1}=X_{2i}$, $d_{2i}=-h_{i0}$, $g_{2i-1,j}=0$, $g_{2i,j}=h_{ij}$, $(i=1,2,...,n_y;\ j=1,2,...,m)$, and $2n_y=n_x$, Eq. (1) can be expressed as follows:

$$\frac{d}{dt}X_i = d_i(\mathbf{X}) + g_{ij}(\mathbf{X})W_j(t) \quad i=1,2,...,n_x;\ j=1,2,...,m \tag{2}$$

in Stratonovich form, where $\mathbf{X}\in\Re^{n_x}$, X_i, $(i=1,2,...,n_x)$, are components of the vector process $\mathbf{X}$, $d_i:\Re^{n_x}\to\Re$, and $g_{ij}:\Re^{n_x}\to\Re$. Functions d_i and g_{ij} are generally nonlinear, and their functional forms are assumed to be deterministic. When the excitations $W_j(t)$ are Gaussian white noises with zero mean and cross correlation

$$E[W_j(t)W_k(t+\tau)] = S_{jk}\delta(\tau) \tag{3}$$

where $\delta(\tau)$ is the Dirac function and S_{jk} are constants, representing the cross-spectral density of W_j and W_k, Eq. (2) may also be expressed in Ito's form as

$$\frac{d}{dt}X_i = f_i(\mathbf{X}) + g_{ij}(\mathbf{X})W_j(t) \quad i=1,2,...,n_x;\ j=1,2,...,m \tag{4}$$

where

$$f_i(\mathbf{X}) = d_i(\mathbf{X}) + \frac{1}{2}\frac{\partial \mathbf{g}_{ij}(\mathbf{X})}{\partial X_k}g_{kj}(\mathbf{X}). \tag{5}$$

The system response $\mathbf{X}$ is a Markov vector and the probability density of the stationary Markov vector is governed by the following reduced FPK equation ([14]):

$$\frac{\partial}{\partial x_j}(f_j p) - \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j}(G_{ij}p) = 0 \tag{6}$$

where $\mathbf{x}$ is the state vector and $\mathbf{x}\in\Re^{n_x}$, $p=p(\mathbf{x})$ and

$$G_{ij}(\mathbf{x}) = S_{ls}g_{il}(\mathbf{x})g_{js}(\mathbf{x}). \tag{7}$$

In the following discussion it is assumed that the probability density function $p(\mathbf{x})$ of the stationary responses of random system (4) satisfies

$$\begin{cases} p(\mathbf{x}) \geqslant 0 & \mathbf{x}\in\Re^{n_x} \\ \lim_{x_i\to\infty} p(\mathbf{x}) = 0 & i=1,2,...,n_x \\ \int_{\Re^{n_x}} p(\mathbf{x})d\mathbf{x} = 1 \end{cases}. \tag{8}$$

If an approximate probability density function denoted as $\tilde{p}(\mathbf{x};\mathbf{a})$ is used, where $\mathbf{a}\in\Re^{N_p}$ and a_i, $(i=1,2,...,N_p)$ are parameters to be determined and N_p is the total number of the parameters, it is obvious that conditions (8) should also be fulfilled by the approximate probability density function. Another requirement for the approximate probability density function is that it must include as many parameters as possible so that the higher level of approximation can be reached.

It can be summarized that the following four problems exist at present and need to be solved for the probability density function solution of the nonlinear random vibrations of multi-degree-of-freedom systems:

- Formulate a versatile probability density function model so that condition (8) can be fulfilled.
- An arbitrary number of unknown parameters can be included in the probability density function model and the precision of approximate probability density function solution can be improved as the total number of the parameters increases.
- Formulate a consistent solution procedure with the probability density function model so that practical problems can be solved consistently.
- The probability density function model and solution technique is valid for multi-degree-of-freedom or multi-dimensional systems.

A possible strategy attempting to solve the above four problems will be presented in the next section.

3 Technique for Approximate Probability Density Function Solution of Multi-Degree-of-Freedom Systems

In this section, a probability density function model is formulated and then, based on the probability density function model, a solution technique is presented for the approximate solution to Eq. (6).

The probability density function $\tilde{p}(\mathbf{x};\mathbf{a})$ for the approximate solution to Eq. (6) is assumed to be of the form

$$\tilde{p}(\mathbf{x};\mathbf{a}) = \begin{cases} c\exp^{Q_n(\mathbf{x};\mathbf{a})} & \mathbf{x}\times\mathbf{a}\in D_x^{n_x}\times\Re^{N_p} \\ 0 & \mathbf{x}\notin D_x^{n_x} \end{cases} \tag{9}$$

where $\mathbf{a}^T=[a_1,a_2,\ldots,a_{N_p}]\in\Re^{N_p}$, being a_i the parameters to be determined; N_p is the total number of the parameters and $D_x^{n_x}=[m_1-\alpha_1\sigma_1,m_1+\beta_1\sigma_1]\times[m_2-\alpha_2\sigma_2,m_2+\beta_2\sigma_2]\times\ldots\times[m_i-\alpha_i\sigma_i,m_i+\beta_i\sigma_i]\times\ldots\times[m_{n_x}-\alpha_{n_x}\sigma_{n_x},m_{n_x}+\beta_{n_x}\sigma_{n_x}]\subset\Re^{n_x}$ in which m_i and σ_i denote the mean value and standard deviation of X_i, respectively. $\alpha_i>0$ and $\beta_i>0$ are defined such that $m_i-\alpha_i\sigma_i$ and $m_i+\beta_i\sigma$ are located in the tails of the probability density function of X_i and the derivatives of the probability density function of X_i with respect to x_i at $m_i-\alpha_i\sigma_i$ and $m_i+\beta_i\sigma$ equal zero; c is the normalization constant and $Q_n(\mathbf{x};\mathbf{a})$ is an n-degree polynomial in the state variables $x_1,x_2,\ldots,x_{n_x}$. In this paper, the following expression for $Q_n(\mathbf{x};\mathbf{a})$ is utilized in the presented examples:

$$Q_n(\mathbf{x};\mathbf{a}) = \sum_{i=1}^{n_x} a_i x_i + a_{n_x+1}x_1^2 + a_{n_x+2}x_1x_2 + \ldots + a_{n_x(n_x+3)/2}x_{n_x}^2 + \ldots + a_{N_p}x_{n_x}^n. \tag{10}$$

In addition to the polynomial (10), any other polynomial may also be used, for example, the polynomial obtained from the maximum entropy method.

Eq. (6) can also be written in the following form:

$$\frac{\partial f_j}{\partial x_j}p+f_j\frac{\partial p}{\partial x_j}-\frac{1}{2}\left(\frac{\partial^2 G_{ij}}{\partial x_i\partial x_j}p+\frac{\partial G_{ij}}{\partial x_j}\frac{\partial p}{\partial x_i}+\frac{\partial G_{ij}}{\partial x_i}\frac{\partial p}{\partial x_j}+G_{ij}\frac{\partial^2 p}{\partial x_i\partial x_j}\right)=0. \tag{11}$$

Generally, the reduced FPK Eq. (11) cannot be satisfied exactly with $\tilde{p}(\mathbf{x};\mathbf{a})$ because $\tilde{p}(\mathbf{x};\mathbf{a})$ is only an approximation of $p(\mathbf{x})$ and the total number N_p of the unknown parameters is always limited in practice. Substituting $\tilde{p}(\mathbf{x};\mathbf{a})$ for $p(\mathbf{x})$ in Eq. (11) yields the following residual error:

$$\Delta(\mathbf{x};\mathbf{a})=\frac{\partial f_j}{\partial x_j}\tilde{p}+f_j\frac{\partial \tilde{p}}{\partial x_j}-\frac{1}{2}\left(\frac{\partial^2 G_{ij}}{\partial x_i\partial x_j}\tilde{p}+\frac{\partial G_{ij}}{\partial x_j}\frac{\partial \tilde{p}}{\partial x_i}+\frac{\partial G_{ij}}{\partial x_i}\frac{\partial \tilde{p}}{\partial x_j}+G_{ij}\frac{\partial^2 \tilde{p}}{\partial x_i\partial x_j}\right). \tag{12}$$

Substituting Eq. (9) into Eq. (12) leads to

$$\Delta(\mathbf{x};\mathbf{a})=\delta(\mathbf{x};\mathbf{a})\tilde{p}(\mathbf{x};\mathbf{a}) \tag{13}$$

where

$$\delta(\mathbf{x};\mathbf{a})=f_j\frac{\partial Q_n}{\partial x_j}-\frac{1}{2}\left(\frac{\partial G_{ij}}{\partial x_j}\frac{\partial Q_n}{\partial x_i}+\frac{\partial G_{ij}}{\partial x_i}\frac{\partial Q_n}{\partial x_j}+G_{ij}\frac{\partial^2 Q_n}{\partial x_i\partial x_j}+G_{ij}\frac{\partial Q_n}{\partial x_i}\frac{\partial Q_n}{\partial x_j}\right)+\frac{\partial f_j}{\partial x_j}-\frac{1}{2}\frac{\partial^2 G_{ij}}{\partial x_i\partial x_j}. \tag{14}$$

Because $\tilde{p}(\mathbf{x};\mathbf{a})\neq 0$, the only possibility for $\tilde{p}(\mathbf{x};\mathbf{a})$ to satisfy Eq. (11) is $\delta(\mathbf{x};\mathbf{a})=0$. However, usually $\delta(\mathbf{x};\mathbf{a})\neq 0$ because $\tilde{p}(\mathbf{x};\mathbf{a})$ is only an approximation of $p(\mathbf{x})$. In this case, another set of mutually independent functions $H_k(\mathbf{x})$ which span the space R^{N_p} can be introduced to make the projection of $\delta(\mathbf{x};\mathbf{a})$ on R^{N_p} vanish, which yields

$$\int_{\Re^{n_x}}\delta(\mathbf{x};\mathbf{a})H_k(\mathbf{x})d\mathbf{x}=0,\quad k=1,2,...,N_p \tag{15}$$

or

$$\int_{\Re^{n_x}}\left\{f_j\frac{\partial Q_n}{\partial x_j}-\frac{1}{2}\left(\frac{\partial G_{ij}}{\partial x_j}\frac{\partial Q_n}{\partial x_i}+\frac{\partial G_{ij}}{\partial x_i}\frac{\partial Q_n}{\partial x_j}+G_{ij}\frac{\partial^2 Q_n}{\partial x_i\partial x_j}+G_{ij}\frac{\partial Q_n}{\partial x_i}\frac{\partial Q_n}{\partial x_j}\right)+\frac{\partial f_j}{\partial x_j}-\frac{1}{2}\frac{\partial^2 G_{ij}}{\partial x_i\partial x_j}\right\}H_k(\mathbf{x})d\mathbf{x}=0,\quad k=1,2,...,N_P. \tag{16}$$

Equation (16) means that the reduced FPK equation is satisfied with $\tilde{p}(\mathbf{x};\mathbf{a})$ in the average sense of integration if $\delta(\mathbf{x};\mathbf{a})H_k(\mathbf{x})$ is integrable in $\Re^{n_x}$.

The function $H_k(\mathbf{x})$ can be selected as $x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}f_N(\mathbf{x})$, being $k_1,k_2,\ldots,k_n=0,1,2,...,N_p$ and $k=k_1+k_2+\ldots+k_n$ such that $\delta(\mathbf{x};\mathbf{a})H_k(\mathbf{x})$ is integrable in $\Re^{n_x}$. In order to guarantee that $\delta(\mathbf{x};\mathbf{a})H_k(\mathbf{x})$ is integrable in $\Re^{n_x}$, the function $f_N(\mathbf{x})$ must be properly selected. Numerical experience shows that a convenient and effective choice for function $f_N(\mathbf{x})$ is the Gaussian probability density function, for instance the probability density function obtained with the Gaussian closure or an equivalent linearization procedure. Because of the particular choice of $f_N(\mathbf{x})$, there is no difficulty in formulating the algebraic equations in terms of the unknown parameters if the integration in Eq. (16) can be easily evaluated by taking into account the relationships between higher and lower order moments of Gaussian stochastic processes. Because of this simplification, this solution technique is tractable to many engineering problems for which the integration of Eq. (16) can be carried out analytically. From this solution technique it is seen that the normalization factor c drops out and therefore there is no need for a nonlinear constrain to be considered that forces the area under a probability distribution function to equal 1. Because $\delta(\mathbf{x};\mathbf{a})$ is a quadratic function of $Q_n(\mathbf{x};\mathbf{a})$, the obtained algebraic equations from Eq. (16) are quadratic algebraic equations of $\mathbf{a}$. Numerical experience showed that the solution of this algebraic equation converges very fast by utilizing some popular methods, e.g., the Newton's method.

It is also seen from the above discussion that the unknown function $\tilde{p}(\mathbf{x};\mathbf{a})$, with the presented method, is not approximated as a linear function of $\mathbf{a}$. hence the procedure is essentially different from the Galerkin procedure.

4 Examples

Example 1. The following two-degree-of-freedom system is analyzed with the presented method:

$$\ddot{Y}_1+\tfrac{1}{2}a_1(S_{11}\dot{Y}_1+2a_2S_{12}\dot{Y}_2)+2a_3Y_1+4a_4Y_1^3+6a_5Y_1^5+W_1(t) \tag{17}$$

$$\ddot{Y}_2+\tfrac{1}{2}a_1[2(1-a_2)S_{12}\dot{Y}_1+S_{22}\dot{Y}_2]+2a_6Y_2+4a_7Y_2^3+6a_8Y_2^5+W_2(t) \tag{18}$$

where $a_1,a_2,\ldots,a_8$ are some constants, $E[W_i(t+\tau)W_j(t)]=S_{ij}\delta(\tau)$, $(i,j=1,2)$. Setting $Y_1=X_1$, $\dot{Y}_1=X_2$, $Y_2=X_3$ and $\dot{Y}_2=X_4$, the system can be expressed as

$$\dot{X}_1=X_2 \tag{19}$$

$$\dot{X}_2=-\tfrac{1}{2}a_1(S_{11}X_2+2a_2S_{12}X_4)-2a_3X_1-4a_4X_1^3-6a_5X_1^5+W_1(t) \tag{20}$$

$$\dot{X}_3=X_4 \tag{21}$$

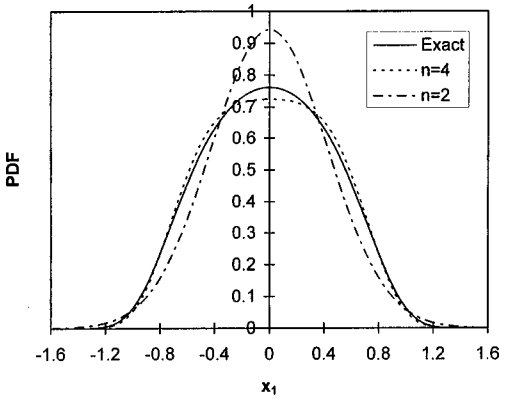


Fig. 1 The probability density functions of X_1, for Example 1

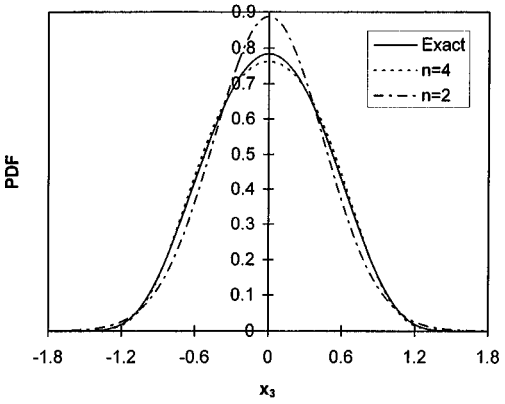


Fig. 2 The probability density functions of X_3, for Example 1

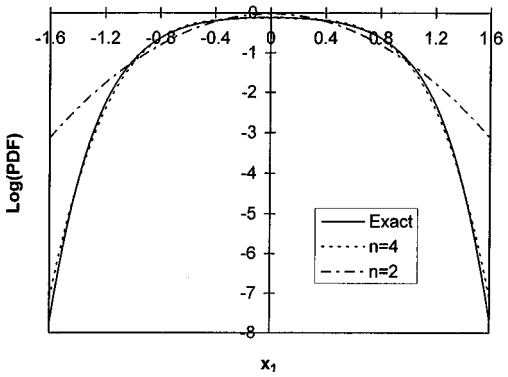


Fig. 3 The logarithmic probability density functions of X_1, for Example 1

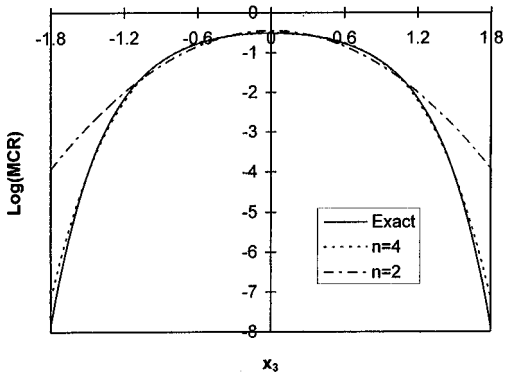


Fig. 4 The logarithmic probability density functions of X_3, for Example 1

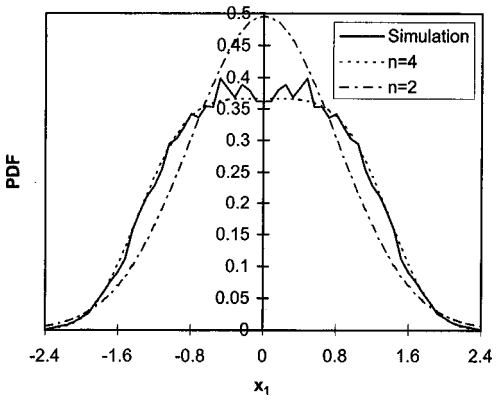


Fig. 5 The probability density functions of X_1, for Example 2

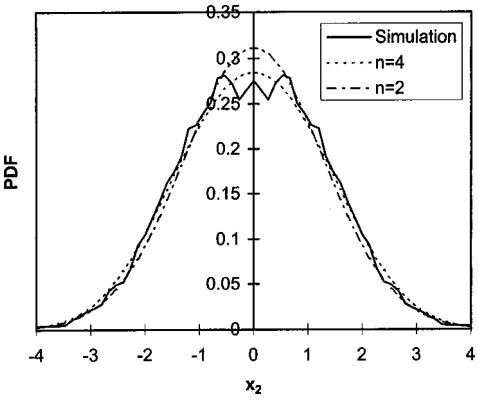


Fig. 6 The probability density functions of X_2, for Example 2

$$\dot{X}_4 = -\tfrac{1}{2}a_1[2(1-a_2)S_{12}X_2 + S_{22}X_4] - 2a_6X_3 - 4a_7X_3^3 - 6a_8X_3^5 + W_2(t). \quad (22)$$

For this system, the exact PDF solution is obtainable to be ([15])

$$p(x_1,x_2,x_3,x_4) = C\exp\{-a_1[\tfrac{1}{2}(x_2^2+x_4^2) + a_3x_1^2 + a_4x_1^4 + a_5x_1^6 + a_6x_3^2 + a_7x_3^4 + a_8x_3^6]\} \quad (23)$$

where C is the normalization constant.

For $S_{11}=S_{22}=2$, $a_1=a_3=a_4=1$, $a_5=a_7=0.5$, $a_6=1.5$, and $a_8=0.2$, as well as arbitrary values of S_{12} and a_2, the approximate probability density functions of X_1 and X_3 obtained with the presented method are compared with the exact probability density function solutions in Figs. 1 and 2. It is apparent that the approximate solutions for $n=4$ are very close to exact solutions though the system is highly nonlinear in this case. For $n=2$, the results coincide with those from the linearization or Gaussian closure procedure. The probability density function solutions for $n=4$ are much improved compared to those from the equivalent linearization procedure. In order to show the tail behavior of the probability density functions, the logarithmic probability density functions are plotted in Figs. 3 and 4. It is seen that the approximate probability density functions for $n=4$ are much closer to the exact probability density function solutions, even in the tails. Numerical analysis showed that even an exact probability density function solution can be obtained for $n=6$.

Example 2. Consider the following nonlinear two-degree-of-freedom system with both external and parametric excitations.

$$\ddot{Y}_1 + 0.2\dot{Y}_1 + Y_1[1+W_1(t)] + Y_1^3 - Y_2 = 0 \quad (24)$$

$$\ddot{Y}_2 + 0.5\dot{Y}_2 + Y_2 + 0.5Y_2^3 = W_2(t) \quad (25)$$

where $E[W_i(t+\tau)W_j(t)] = S_{ij}\delta(\tau)$, $(i,j=1,2)$, $S_{11}=0.2$, $S_{22}=1$ and $S_{12}=0$. Setting $Y_1=X_1$, $\dot{Y}_1=X_2$, $Y_2=X_3$, and $\dot{Y}_2=X_4$, the above system can also be expressed as

$$\dot{X}_1 = X_2 \quad (26)$$

$$\dot{X}_2 = -X_1[1+W_1(t)] - X_1^3 - 0.2X_2 + X_3 \quad (27)$$

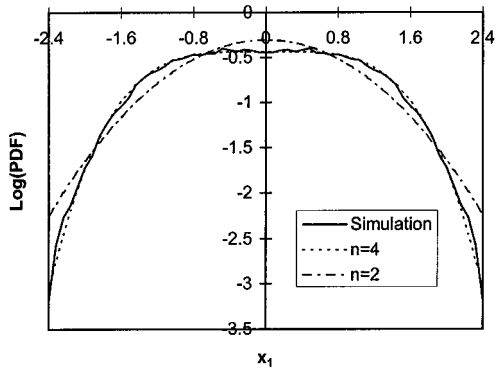


Fig. 7 The logarithmic probability density functions of X_1, for Example 2

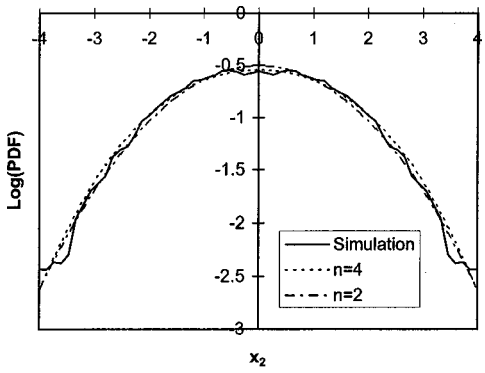


Fig. 8 The logarithmic probability density functions of X_2, for Example 2

$$\dot{X}_3 = X_4 \tag{28}$$

$$\dot{X}_4 = W_2(t) - X_3 - 0.5X_3^3 - 0.5X_4 . \tag{29}$$

There is no obtainable exact probability density function solution for this system. In order to show the effectiveness of the presented method for such systems, the Monte Carlo simulation is conducted to obtain the simulated probability density function solutions which are compared with the probability density functions for $n=4$ and 2, of X_1 and X_2 in Figs. 5 and 6. From these figures it is seen that the probability density functions for $n=4$ are much improved in comparison with those for $n=2$. The tail behaviors of the probability density functions are also compared in Figs. 7 and 8. Again excellent agreement is found between the results for $n=4$ and simulated results. It is seen that the probability density functions are also improved even in the tails as n increases from 2 to 4. The tails are truncated at the points where the simulated probability density functions become fluctuated due to the limited sample size though the selected sample size is as big as 10^7.

The above results validate the method for the nonlinear random vibrations of multi-degree-of-freedom systems with both external and parametric excitations.

5 Conclusions

The probability density function solution of nonlinear stochastic multi-degree-of-freedom systems can be reasonably approximated with $c\exp^{Q_n(\mathbf{x};\mathbf{a})}$ in the domain defined in the above discussion, where $Q_n(\mathbf{x};\mathbf{a})$ is an n-degree polynomial in $x_1, x_2, \ldots, x_{n_x}$. Such an approximate probability density function fulfills the necessary requirements for the probability density function of the stationary random variables. Special measure is taken such that the FPK equation is satisfied in the average sense of integration. The results obtained for $n=4$ are much improved in comparison with those obtained for $n=2$. From numerical results, it is also known that the results obtained with the presented method for $n>2$ are much closer to the exact or simulated probability density function solutions, even in the tails of the probability density functions, which is important for reliability analysis and some other statistical analysis. Numerical results also show that the presented method is even valid for the multi-degree-of-freedom systems with high nonlinearity and both external and parametric excitations. Because only quadratic algebraic equations are resulted with the method, which can be solved easily to determine the unknown parameters, this method provides a consistent and powerful tool for the probability density function solution of nonlinear random vibrations of the multi-degree-of-freedom systems with a polynomial type of non-linearity and both external and parametrical excitations.

Acknowledgments

The results presented in this paper were obtained in the course of research supported by the funding of the Research Committee of the University of Macao, Grant No. 6851/96/UM, 628/97/UM.

References

[1] Booton, R. C., 1954, ''Nonlinear Control Systems With Random Inputs,'' IRE Trans. Circuit Theory, CT-1, **1**, pp. 9–19.
[2] Caughey, T. K., 1959, ''Response of a Nonlinear String to Random Loading,'' ASME J. Appl. Mech., **26**, pp. 341–344.
[3] Lin, Y. K., 1967, *Probabilistic Theory of Structure Dynamics*, McGraw-Hill, New York.
[4] Iwan, W. D., 1973, ''A Generalization of the Concept of Equivalent Linearization,'' Int. J. Non-Linear Mech., **5**, pp. 279–287.
[5] Assaf, S. A., and Zirkie, L. D., 1976, ''Approximate Analysis of Nonlinear Stochastic Systems,'' Int. J. Control, **23**, pp. 477–492.
[6] Soize, C., 1988, ''Steady-State Solution of Fokker-Planck Equation in Higher Dimension,'' Prob. Eng. Mech., **3**, pp. 196–206.
[7] Stratonovich, R. L., 1963, *Topics in the Theory of Random Noise*, Vol. 1, Gordon and Breach, New York.
[8] Sobczyk, K., and Trebicki, J., 1990, ''Maximum Entropy Principle in Stochastic Dynamics,'' Prob. Eng. Mech., **5**, pp. 102–110.
[9] Er, G. K., 1998, ''Multi-Gaussian Closure Method for Randomly Excited Nonlinear Systems,'' Int. J. Non-Linear Mech., **33**, pp. 201–214.
[10] Shinozuka, M., 1972, ''Monte Carlo Simulation of Structural Dynamics,'' Comput. Struct., **2**, pp. 855–874.
[11] Harris, C. J., 1979, ''Simulation of Multivariate Nonlinear Stochastic System,'' Int. J. Numer. Methods Eng., **14**, pp. 37–50.
[12] Kloeden, P. E., and Platen, E., 1995, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin.
[13] Er, G. K., 1999, ''A Consistent Method for the PDF Solutions of Random Oscillators,'' ASCE Journal of Engineering Mechanics, **125**, pp. 443–447.
[14] Soong, T. T., 1973, *Random Differential Equations in Science and Engineering*, Academic Press, New York.
[15] Scheurkogel, A, and Elishakoff, I, 1988, ''Non-linear Random Vibration of a Two-Degree-of-Freedom System,'' *Non-Linear Stochastic Engineering Systems*, F. Ziegler and G. I. Schuëller, eds., Springer-Verlag, Berlin, pp. 285–299.

Dynamic Stability of Poroelastic Columns

G. Cederbaum
The Pearlstone Center for Aeronautical Engineering Studies,
Department of Mechanical Engineering,
Ben-Gurion University of the Negev,
Beer-Sheeva 84105, Israel

The dynamic stability of a poroelastic column subjected to a longitudinal periodic force is investigated. The column material is assumed to be transversely isotropic with respect to the column axis, and the pore fluid flow is possible in the axial direction only. The motion of the column is governed by two coupled equations, for which the stability boundaries are determined analytically by using the multiple-scales method. It is shown that due to the fluid diffusion the stability regions are expanded, relative to the elastic (drained) case. The critical (minimum) loading amplitude, for which instability occurs, is also given.
[S0021-8936(00)00902-8]

Introduction

The dynamic stability of elastic structures was investigated by Bolotin [1], where the behavior is governed by the Mathieu equation and the stability characterizations are given by the Strutt diagram (see also Timoshenko and Gere [2]). Extensive bibliographies for further results for this problem were given by Evan-Iwanovsky in a review paper ([3]) and in a monograph ([4]). Linear and nonlinear viscoelastic structures subjected to periodic forces have been investigated, e.g., in Stevens [5] and Touati and Cederbaum [6]. It was shown that due to the structural damping, resulting from the viscoelasticity of the material, the regions of stability are expanded with respect to the elastic case. In this paper we investigate the dynamic stability of a column with another time-dependent characterization, the case of a poroelastic column.

Poroelasticity is a continuum theory for porous media consisting of an elastic matrix containing interconnected fluid-filled pores. In physical terms, the theory postulates that when a porous material is subjected to stress, the resulting matrix deformation leads to volumetric changes in the pores. If the pores are fluid-filled, this results in a flow of the pore fluid between regions of higher and lower pore volume change. The phenomenological model for the behavior of fluid-saturated poroelastic materials was developed by Biot [7]. His motivation, and the main application of the theory over the years, was concerned with dynamic problems (wave propagation) in geological (massive) structures. Sub-three-dimensional structures, such as beams, plates, or shells, have received much less attention and were mainly motivated by problems in biomechanics ([8–11]). When such elements are subjected to bending the stress gradients in the transverse direction are usually much greater than those in the axial or in-plane ones, so that if the bulk material is isotropic the diffusion (the fluid flow in the pores) in the transverse direction is dominant, and the diffusion in the axial or in-plane directions is considered to be negligible. On the other hand, there are structural elements, such as plant stems, for which by virtue of the microgeometry the fluid flow in the axial direction is dominant. This case was investigated recently by Li et al. [12–14] for the quasi-static behavior, vibration, and buckling of beams, where the fluid-saturated material was taken to be transversely isotropic in the cross-sectional plane. We used the Biot's constitutive relations ([7]) and the Darcy's law in the formulation of the beams, modeled within classical beam theory. Very interesting behavior patterns were exhibited for this kind of poroelastic element, which could easily be fabricated. The same type of material is considered in the following.

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 1, 1998; final revision, July 30, 1999. Associate Technical Editor: J. T. Jenkins. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Problem Formulation

Consider the problem of the transverse vibration of a simply supported poroelastic column with a uniform cross section, subjected to a periodic longitudinal loading, as shown in Fig. 1. The cross-sectional dimensions are small compared to length, so that the classical beam theory is adopted. In Li et al. [12] it was shown that only four bulk material properties are necessary to define the behavior of the poroelastic material–E, η, λ, and K. The physical interpretation of these parameters is as follows: E is the axial Young's modulus of the solid skeleton (drained), namely, when the pore fluid pressure $p_f=0$. η is a nondimensional constant by which the axial strain must be multiplied in order to find the relative change in pore volume of the drained material when it is stressed in the axial direction. Now, let β denote the ratio of pore fluid increment to pore pressure when there is no axial strain, then $\lambda=\eta/E\beta$. Finally, K is related to axial permeability and is given by $K=k\lambda E/\mu\eta$, where k is the axial permeability of the porous skeleton (depends only on the pore geometry), and μ is the viscosity of the pore fluid.

The constitutive law and the actual Darcy's law for the quasi-static buckling problem result in the following beam equations in the form of global variables Li et al. [14]:

$$EI\frac{\partial^2 y(x,t)}{\partial x^2}-\eta M_p+M=0, \tag{1}$$

$$K\frac{\partial^2 M_p(x,t)}{\partial x^2}-\frac{\partial M_p(x,t)}{\partial t}-\lambda EI\frac{\partial^3 y(x,t)}{\partial^2 x\,\partial t}=0, \tag{2}$$

respectively, where M is the bending moment which can be determined from the equilibrium of the column, and M_p is the pore pressure moment resultant defined by

$$M_p=-\int_A p_f z\,dA \tag{3}$$

where z is distance from the neutral axis. Equations (1) and (2) are coupled through λ and η; if $\lambda=\eta=0$ (if either is zero the other must also be zero) then Eq. (1) reduces to the purely elastic case. For a simply supported column $M=Py$, where P is the end load.

In our case the load is time-dependent, given by

$$P(t)=P_0\cos\theta t. \tag{4}$$

Thus, in order to arrive at the equation for the transverse vibration of the column it is necessary to differentiate Eq. (1) twice with respect to x and to add the inertia force acting on the column. By doing so one obtains

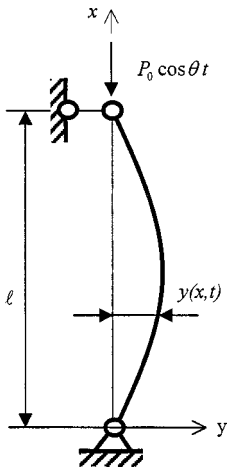


Fig. 1 Column configuration

$$EI\frac{\partial^4 y(x,t)}{\partial x^4}-\eta\frac{\partial^2 M_p}{\partial x^2}+P_0\cos\theta t\frac{\partial^2 y(x,t)}{\partial x^2}+\rho\frac{\partial^2 y(x,t)}{\partial t^2}=0 \tag{5}$$

where ρ is the mass per unit length of the column.

Using the separation of variables method, and in order to satisfy the boundary conditions of a simply supported column, the solution functions are given by the following Fourier sine series:

$$y(x,t)=\sum_{j=1}^{\infty} f_j \sin\frac{j\pi x}{l} \tag{6a}$$

$$M_p(x,t)=\sum_{j=1}^{\infty} m_j \sin\frac{j\pi x}{l} \tag{6b}$$

where l is the length of the column. In the following, the main stability region is investigated. To this end only the $j=1$ term in Eqs. (6) is considered. Let $\alpha=\pi/l$, and substituting Eqs. (6) into Eqs. (5) and (2), the equations govern the motion of the column are now in the form

$$EI\alpha^4 f+\eta\alpha^2 m-P_0\alpha^2\cos\theta t f+\rho\ddot{f}=0 \tag{7a}$$

$$-K\alpha^2 m-\dot{m}+\lambda EI\alpha^2\dot{f}=0. \tag{7b}$$

Here the dot denotes differentiation with respect to time. By solving (7*a*) for m and putting this expression together with its time derivative into (7*b*), a single differential equation involving f only is obtained:

$$\frac{\rho}{\alpha^2}\ddot{f}+[(1+\lambda\eta)EI\alpha^2-P_0\cos\theta t]\dot{f}+P_1\theta\sin\theta t f$$
$$=-K[\rho\ddot{f}+(EI\alpha^4-P_0\alpha^2\cos\theta t)f]. \tag{8}$$

Let

$$\omega^2=\frac{EI}{\rho}\alpha^4,\quad \Omega^2=(1+\lambda\eta)\omega^2,\quad P_E=(1+\lambda\eta)EI\alpha^2,$$
$$\gamma=\frac{P_0}{2P_E},\quad \delta=K\alpha^2 \tag{9}$$

by which the equation of motion is in the form

$$\ddot{f}+\Omega^2[(1-2\gamma\cos\theta t)\dot{f}+2\theta\gamma\sin\theta t f]$$
$$=-\delta[\ddot{f}+\omega^2(1-2\gamma(1+\lambda\eta)\cos\theta t)f] \tag{10}$$

where γ and δ are small and P_E is the static buckling load of the fluid-saturated column. However, since the left side of Eq. (10) is the time-derivative of the Mathieu equation, it is possible to proceed by investigating the stability of the following equation instead:

$$\ddot{f}+\Omega^2(1-2\gamma\cos\theta t)f$$
$$=-\delta\left(\dot{f}+\omega^2\int f(t)dt-\omega^2 2\gamma(1+\lambda\eta)\int\cos\theta t f(t)dt\right). \tag{11}$$

Stability Boundaries of the Equation of Motion

In this section the stability properties of the solution of Eq. (11) are investigated. To this end the multiple scale method is employed in order to obtain asymptotic solutions ([15]). This method is based on the fact that physical processes that determine the behavior of the structure take place on distinct time scales. In the system investigated here, three different time scales control its development. The first time scale serves as a reference to the others and hence is considered throughout this work to be of order 1. The second time scale, associated with the loading amplitude P_1, is given by $1/\gamma$, and is assumed to be much bigger than 1. The third time scale is associated with the diffusion characteristics and is given by $1/\delta$. As was shown in Li et al. [13] the diffusion time of a beam of length l is equal to l^2/K, and is considered in the following to be much longer than the period time of the drained beam.

We turn now to finding the instability boundaries in the (Ω, θ, γ, δ) space in the limit of γ and δ being small. It is first noted that when

$$\Omega=\frac{\theta}{2} \tag{12}$$

a regular perturbation expansion in γ and δ results in solutions that contain terms that are unbounded in time. These terms, known as secular terms, are unphysical. In order to eliminate these secular terms, the multiple scales analysis is employed by introducing two new independent variables $\tau=\gamma t$ and $\zeta=\delta t$. Within the multiscale analysis t, τ, and ζ are considered as three independent variables. As a result, f is expanded in the following form:

$$f(t,\tau,\zeta)=f_0(t,\tau,\zeta)+\gamma g(t,\tau,\zeta)+\delta h(t,\tau,\zeta)+O(\gamma^2,\delta^2,\gamma\delta) \tag{13}$$

and the second time derivative of f as

$$\frac{d^2 f}{dt^2}=\frac{\partial^2 f_0}{\partial t^2}+\gamma\left(2\frac{\partial^2 f_0}{\partial t\partial\tau}+\frac{\partial^2 g}{\partial t^2}\right)+\delta\left(2\frac{\partial^2 f_0}{\partial t\partial\zeta}+\frac{\partial^2 h}{\partial t^2}\right)$$
$$+O(\gamma^2,\delta^2,\gamma\delta). \tag{14}$$

Further, since the regions of interest are near $\Omega=\theta/2$, Ω is expanded as

$$\Omega^2=\frac{\theta^2}{4}+\gamma\Omega_1^2+O(\gamma^2) \tag{15}$$

where, to the first order of γ, Ω_1 measures how close half the loading frequency $\theta/2$ is to the natural frequency Ω. Equations (13), (14), and (15) are inserted into Eq. (11) and coefficients of equal powers of γ and δ are collected. The lowest order terms yields the following equation for f_0

$$\frac{\partial^2 f_0}{\partial t^2}+\frac{\theta^2}{4}f_0=0 \tag{16}$$

and its solution is given by

$$f_0(t,\tau,\zeta)=A(\tau,\zeta)e^{i(\theta/2)t}+\text{C.C.} \tag{17}$$

where C.C. stands for complex conjugate.

Terms of first order in γ and δ yield the following equations, respectively:

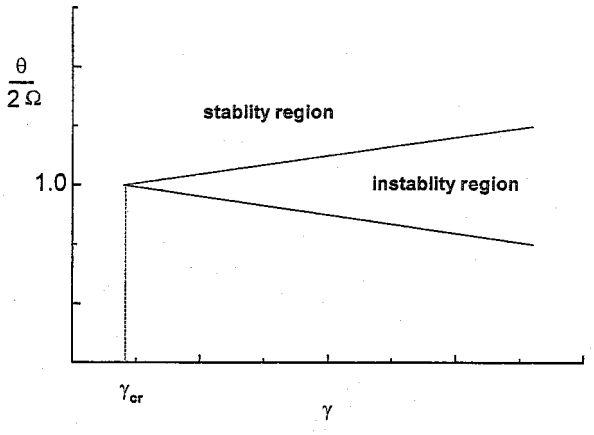


Fig. 2 Stability boundaries in a qualitative form

$$\frac{\partial^2 g}{\partial t^2}+\frac{\theta^2}{4}g=\left[-i\theta\frac{\partial A}{\partial \tau}-\Omega_1^2 A+\frac{\theta^2}{4}A^*\right]e^{i(\theta/2)t}+\text{NST}+\text{C.C.} \tag{18}$$

$$\frac{\partial^2 h}{\partial t^2}+\frac{\theta^2}{4}h=\left[-2i\frac{\theta}{2}\frac{\partial A}{\partial \zeta}-\frac{\theta^2}{4(1+\lambda\eta)}\frac{2}{i\theta}A-i\frac{\theta}{2}A\right]e^{i(\theta/2)t}+\text{NST}+\text{C.C.} \tag{19}$$

The terms on the right-hand side of Eqs. (18) and (19) that multiply $\exp(i\theta/2)$ are the secular ones since they give rise to unbounded solutions. The rest are nonsecular terms (NST). In order to eliminate the secular terms, it is required that the coefficients multiplying $\exp(i\theta t/2)$ vanish. Let

$$A(\tau,\zeta)=T(\tau)Z(\zeta) \tag{20}$$

and by introducing it into Eqs. (18) and (19) yields

$$\theta\frac{\partial T}{\partial \tau}-i\Omega_1^2 T+i\frac{\theta^2}{4}T^*=0 \tag{21}$$

$$\frac{\partial Z}{\partial \zeta}+\frac{1}{2}\frac{\lambda\eta}{1+\lambda\eta}Z=0. \tag{22}$$

These have solutions

$$T(\tau)=T_1 e^{(1/\theta)\sqrt{\theta^4/16-\Omega_1^4}\tau}+T_2 e^{-(1/\theta)\sqrt{\theta^4/16-\Omega_1^4}\tau} \tag{23}$$

and

$$Z(\zeta)=Z_0 e^{-(1/2)(\lambda\eta/(1+\lambda\eta))\zeta} \tag{24}$$

respectively, where T_1, T_2, and Z_0 are constants. Inserting Eqs. (23) and (24) into Eq. (17) yields the following form of the zeroth-order solution:

$$f_0(t,\tau,\zeta)=F_0 e^{-(1/2)(\lambda\eta/(1+\lambda\eta))\zeta}e^{\pm 1/\theta\sqrt{\theta^4/16-\Omega_1^4}\tau}e^{i(\theta/2)t} \tag{25}$$

where F_0 is a constant, obtained from T_1, T_2, and Z_0. Then, by rewriting τ and ζ in terms of t, we see that instability occurs when

$$\frac{\gamma}{\theta}\sqrt{\frac{\theta^2}{16}-\Omega_1^4}\geqslant\frac{\delta}{2}\frac{\lambda\eta}{1+\lambda\eta}. \tag{26}$$

Hence, using Eq. (15), it is concluded that for small γ and δ the stability boundaries are given by

$$\Omega^2=\frac{\theta^2}{4}\left[1\pm\sqrt{\gamma^2-\frac{4}{\theta^2}\left(\frac{\lambda\eta}{1+\lambda\eta}\right)^2\delta^2}\right]. \tag{27}$$

The critical (minimum) value of the loading parameter, for which instability occurs, is

$$\gamma_{cr}=\frac{2}{\theta}\frac{\lambda\eta}{1+\lambda\eta}\alpha^2 K. \tag{28}$$

A qualitative Strutt diagram of the above results is given in Fig. 2. It is noted that in the elastic (drained) case $\lambda=\eta=0$, and thus γ_{cr} is equal to zero, as is expected since Eqs. (2) and (5) are not coupled, and thus the column vibrates without a damping mechanism.

It is noted that while the expansion in Eq. (13) involves two independent small parameters γ and δ, the detuning in Eq. (15) only takes into account the effect of γ (and not of δ by $\delta\Omega_2^2$). If this additional detuning term is considered it will add to Eq. (19) a term of the form $-\Omega_2^2 A$ inside the brackets. This leads to an extra term in Eq. (22) in the form $-i\Omega_2^2 Z/\theta$, which, in turn, adds an extra factor to Eq. (24) of the form $e^{\pm i\Omega_2^2\zeta/\theta}$. Thus such a detuning term does not influence stability, at least to the truncation order being investigated.

Acknowledgment

The author thanks a reviewer for his constructive comments.

References

[1] Bolotin, V. V., 1964, *The Dynamic Stability of Elastic Systems*, Halden Day, San Francisco.
[2] Timoshenko, S. P., and Gere, M. G., 1961, *Theory of Elastic Stability*, McGraw-Hill Kogakusha, LTD, Tokyo.
[3] Evan-Iwanovski, R. M., 1965, ''On the Parametric Response of Structures,'' Appl. Mech. Rev. *V-I*, **18**, pp. 699–702.
[4] Evan-Iwanovski, R. M., 1976, *Resonant Oscillations in Mechanical Systems*, Elsevier, Amsterdam.
[5] Stevens, K. K., 1966, ''On the Parametric Excitation of a Viscoelastic Column,'' AIAA J., **12**, pp. 2111–2116.
[6] Touati, D., and Cederbaum, G., 1994, ''Dynamic Stability of Nonlinear Viscoelastic Plates,'' Int. J. Solids Struct., **31**, No. 17, pp. 2367–2376.
[7] Biot, M. A., 1941, ''General Theory of Three Dimensional Consolidation,'' J. Appl. Phys., **12**, pp. 155–165.
[8] Nowinski, J. L., and Davis, C. F., 1972, ''The Flexure and Torsion of Bones Viewed as Anisotropic Poroelastic Bodies,'' Int. J. Eng. Sci., **10**, pp. 1063–1079.
[9] Taber, L. A., 1992, ''A Theory for Transverse Deflection of Poroelastic Plates,'' ASME J. Appl. Mech., **59**, pp. 628–634.
[10] Yang, M., Taber, L. A., and Clark, E. B., 1994, ''A Nonlinear Poroelastic Model for the Trabecular Embryonic Heart,'' ASME J. Biomech. Eng., **116**, pp. 213–223.
[11] Zhang, D., and Cowin, S. C., 1994, ''Oscillatory Bending of a Poroelastic Beam,'' J. Mech. Phys. Solids, **42**, pp. 1575–1599.
[12] Li, L. P., Schulgasser, K., and Cederbaum, G., 1995, ''Theory of Poroelastic Beams With Axial Diffusion,'' J. Mech. Phys. Solids, **43**, No. 12, pp. 2023–2042.
[13] Li, L. P., Cederbaum, G., and Schulgasser, K., 1996, ''Vibration of Poroelastic Beams With Axial Diffusion,'' Eur. J. Mech., **15**, No. 6, pp. 1077–1094.
[14] Li, L. P., Schulgasser, K., and Cederbaum, G., 1997, ''Buckling of Poroelastic Columns With Axial Diffusion,'' Int. J. Mech. Sci., **39**, No. 4, pp. 409–415.
[15] Bender, C. M., and Orszag, S. A., 1984, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, Singapore.

A Normal Force-Displacement Model for Contacting Spheres Accounting for Plastic Deformation: Force-Driven Formulation

L. Vu-Quoc[1]
Professor,
e-mail: vu-quoc@ufl.edu
Mem. ASME

X. Zhang[2]
Graduate Research Assistant

L. Lesburg[3]
Graduate Research Assistant

Aerospace Engineering, Mechanics
and Engineering Science,
University of Florida,
Gainesville, FL 32611

In this paper, we present a simple and accurate model for the normal force-displacement (NFD) relation for contacting spherical particles, accounting for the effects of plastic deformation. This NFD model, based on the formalism of the continuum theory of elastoplasticity, is to be used in granular flow simulations involving thousands of particles; the efficiency of the model is thus a crucial property. The accuracy of the model allows for an accurate prediction of the contact force level in the plastic regime. In addition to being more accurate than previously proposed NFD models, the proposed NFD model also leads to more accurate coefficient of restitution that is a function of the approaching velocity of two particles in collision. The novelty of the present NFD model is the additive decomposition of the contact-area radius, and the correction of the curvature of the particles at the contact point due to plastic flow. The accuracy of the proposed model is validated against nonlinear finite element results involving plastic flow in both loading and unloading conditions. [S0021-8936(00)03102-0]

1 Introduction

Many industrial and agricultural processes involve particle systems. To improve the efficiency of such processes and to design the next generation of particle handling and transportation equipment, the motion behavior of particle systems must be well understood. There have been two methods applied to gain this understanding: experiment and computer simulation. The discrete element method (DEM) is an incremental method in which the equations of motion of the particles are numerically integrated in time to obtain updated positions and velocities (see Cundall and Strack [1] and Vu-Quoc, Zhang, and Walton [2]). The particle-particle interaction greatly affects the behavior of the particle system and thus plays an important role in DEM simulations. In a particle system of *dry* granular materials, the particle-particle interaction is the direct particle-particle contact, which dominates the motion behavior of such a system. Therefore, it is important to use a model that can accurately describe the contact force-displacement (FD) relationship in DEM to obtain reliable simulation results. Since DEM is computationally intensive—especially when the number of particles is large—the DEM FD model must be simple to reduce the complexity of the computation. The most important features of a successful FD model for DEM simulations are *accuracy* and *simplicity*.

Most of the existing FD models for DEM simulations are based on theories of contact mechanics. For example, the Hertz theory (see Hertz [3] and Johnson [4]) provides solutions for elastic contact between spheres subjected load; and the Mindlin and Deresiewicz [5] theory provides solutions for elastic-frictional contact between spheres subjected to a frictional contact force in the tangential direction. Shih et al. [6], provided experimental and finite element (FE) verification of the stress distributions that are predicted by Hertz theory.[4] More recently, Vu-Quoc and Lesburg [7] presents extensive FE validation of the theories of Hertz and Mindlin and Deresiewicz [5]. By showing the significant effect of plastic deformation on FD relationships, Vu-Quoc and Lesburg [7] demonstrate the severity of a common deficiency among the models that are based on elastic contact mechanics theories: These models only account for elastic deformation in the FD relationship. Applying these models to simulations of dry granular flows, in which most contact involve plastic deformation, can lead to inaccurate results.

A model based on the formalism of elasto-plasticity was proposed by Dobry et al. [8]. Even though the Dobry et al. [8] model is named as an elasto-plastic model, it applies only to elastic materials. Johnson [4] applied Hertz theory and the von Mises yield criterion to determine the normal force at which the incipient yield occurs in two spheres subjected to a normal load. This work does not provide, however, the effect of the plastic deformation on the normal force-displacement (NFD) relationship during either loading or unloading. The NFD model proposed by Walton and Braun [9] is based on finite element analysis (FEA) and accounts for plastic deformation. The Walton and Braun [9] NFD model is simple and easy to implement, but it produces a constant coefficient of restitution when simulating the collision of spheres. This behavior is not in agreement with experiments, which showed that the coefficient of restitution depended on the incoming velocity before collision (see Goldsmith [10] and Kangur and Kleis [11]).

Thornton [12] proposed a NFD model that accounts for both elastic and plastic deformation and produces a coefficient of restitution for sphere collisions that varies with the incoming velocity.[5] To produce accurate and reliable simulation results, FD models that account for both elastic and plastic deformation still

[1]To whom correspondence should be addressed.
[2]Now with Siemens Corporation, Princeton, New Jersey.
[3]Now with Parametric Technology Corporation, Boston, MA.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 6, 1998; final revision, Sept. 30, 1999. Associate Technical Editor: J. T. Jenkins. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

[4]See also Johnson [4].
[5]See also Brilliantov et al. [13] for another model. It is noted that our model (elasto-plasticity) differs markedly from the model proposed by Brilliantov et al. [13] (viscoelasticity).

need to be developed. We present here a new NFD model that satisfies these demands. Our model is based on an additive decomposition of the contact-area *radius*, a correction of the local curvature of the particles at the contact point, and a formulation inspired from the continuum theory of elastoplasticity. These cardinal features of our model, the first of its kind in the literature, came from careful nonlinear FEA and observations of the computed results.

2 Elastic Contact Between Two Spheres

The starting point of our elasto-plastic NFD model is Hertz theory, which we briefly review below.

2.1 Hertz Theory. Figure 1 depicts the contact between two spheres subjected to normal load P. Define the equivalent elastic modulus E^* and the equivalent contact curvature $1/R^*$ as

$$E^* := \left(\frac{1-{}_{(i)}\nu^2}{{}_{(i)}E} + \frac{1-{}_{(j)}\nu^2}{{}_{(j)}E} \right)^{-1}, \tag{2.1}$$

and

$$\frac{1}{R^*} := \left(\frac{1}{{}_{(i)}R} + \frac{1}{{}_{(j)}R} \right), \tag{2.2}$$

where ${}_{(i)}R$ is the radius of sphere i, ${}_{(i)}\nu$ and ${}_{(i)}E$ are the Poisson's ratio and Young's modulus of the material of sphere i, respectively. Similarly ${}_{(j)}R$, ${}_{(j)}\nu$, and ${}_{(j)}E$ are those of sphere j. The contact area is a circle with radius a. Hertz proposed that on the contact surface, the distribution of normal pressure p is axisymmetric and shaped as half of an ellipse. At a point A in the contact area, with a distance of r from the center $\mathcal{O}$ of the contact area s, the normal pressure $p(r)$ can be expressed as

$$p(r) = p_m \left[1 - \left(\frac{r}{a} \right)^2 \right]^{1/2}, \tag{2.3}$$

where p_m, the maximum normal pressure at $r=0$, is related to the normal force P and the contact-area radius a by

$$p_m = \frac{3P}{2\pi a^2}. \tag{2.4}$$

Further, we have the following expressions for the contact radius a (Johnson [4], Eq. (4.22)):

$$a = \left(\frac{3PR^*}{4E^*} \right)^{1/3}, \tag{2.5}$$

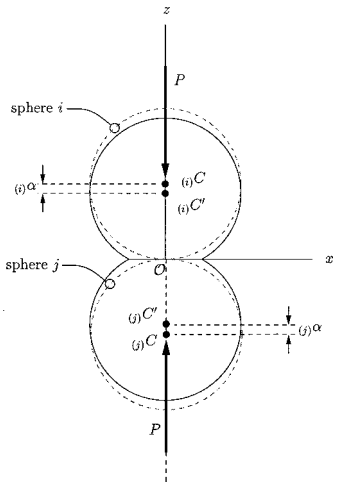


Fig. 1 Two spheres in contact, subjected to normal load P

and for the approach of distant points on the two spheres (Johnson [4], Eq. (4.23))

$${}_{(i)}\alpha + {}_{(j)}\alpha = \frac{a^2}{R^*} = \left(\frac{9P^2}{16R^*(E^*)^2} \right)^{1/3}. \tag{2.6}$$

Introducing (2.5) into (2.4), we obtain

$$p_m = \frac{3P}{2\pi a^2} = \left(\frac{6P(E^*)^2}{\pi^3 (R^*)^2} \right)^{1/3}. \tag{2.7}$$

Hertz theory assumes that the contact area is much smaller than the size of the spheres, i.e., $a \ll {}_{(i)}R$ and $a \ll {}_{(j)}R$. Therefore, the stress distribution inside the sphere can be obtained by considering concentrated forces applied to a elastic half-space. The stress along the z-axis (the axis that passes through centers of the spheres and the center of the contact area, as shown in Figure 1) thus can be expressed as (Johnson [4], Eq. (3.45ab))

$$\sigma_r = \sigma_\theta = -p_m \left\{ (1+\nu) \left[1 - \frac{z}{a} \tan^{-1} \left(\frac{a}{z} \right) \right] - \frac{1}{2} \left(1 + \frac{z^2}{a^2} \right)^{-1} \right\}, \tag{2.8}$$

and

$$\sigma_z = -p_m \left(1 + \frac{z^2}{a^2} \right)^{-1}. \tag{2.9}$$

Based on the above expressions and on the von Mises yield criterion, the normal load for incipient yield P_Y can be evaluated (see Section 3.1).

2.2 Finite Element Validation. We employed FEA to analyze the problem of two identical *elastic* spheres in contact and subjected to normal loading. By symmetry, this problem is equivalent to that of one sphere in contact with a frictionless rigid plane. In addition, by Saint-Venant's principle, only the lower half of the upper sphere is discretized with more than 2000 axisymmetric elements of ABAQUS. We discuss here the main results for the construction of our elasto-plastic NFD model (see Vu-Quoc and Lesburg [7] for more details).

As an example, we chose to study an aluminum sphere of radius $R=0.1$ m, with Young's modulus $E=7.0\times10^{10}$ N/m^2, and Poisson's ratio $\nu=0.3$. We present in this paper only a subset of our results corresponding to the following loading history: The

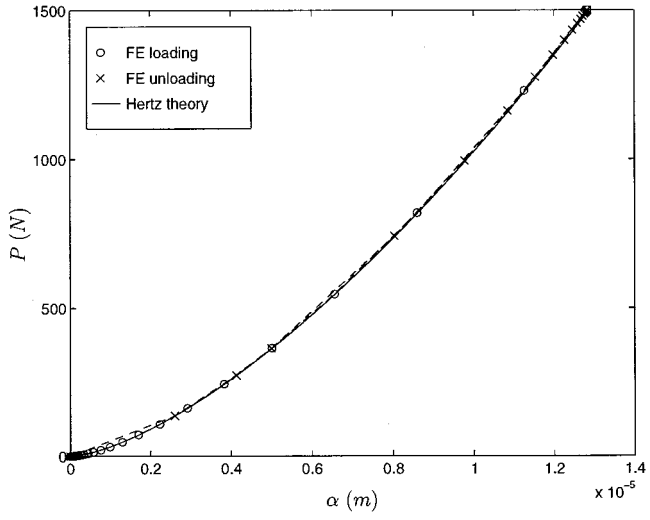


Fig. 2 Normal force P versus normal displacement α: comparison between FEA results and Hertz theory for the loading path with $P_{max}=1500$ N

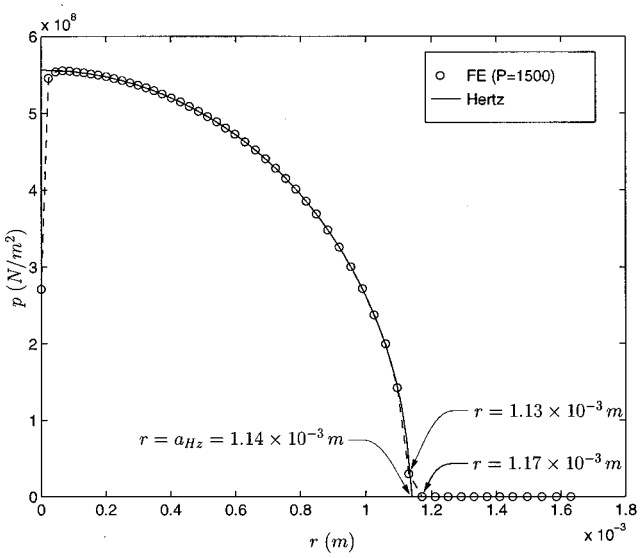


Fig. 3 Distribution of normal stress on the contact surface at maximum normal force P_{max}=1500 N: comparison between FEA results and Hertz theory

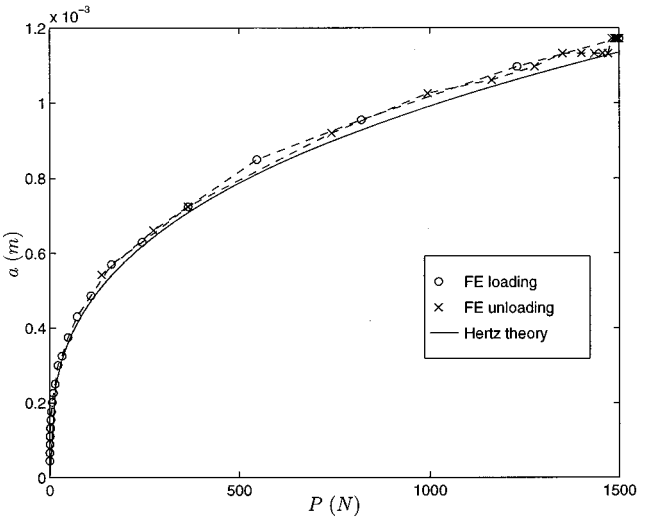


Fig. 4 Contact area radius *a* versus normal force *P*: comparison between FEA results and Hertz theory

normal contact force P is increased from 0 N to 1500 N with a constant rate, then decreased back to 0 N also with a constant rate.[6]

Figure 2 shows a good agreement between FEA results and relation (2.6) of Hertz theory for the elastic NFD relationship in both loading and unloading.

Figure 3 shows the comparison of normal stress distributions on the contact area for the normal force P at the maximal value, $P_{\max}=1500$ N. Our FEA results again agree with relation (2.3) of Hertz theory. The intersections of the normal stress curves with the horizontal axis represent the radius a of the contact area. The dashed line shows that the contact radius in the FEA results falls into the range $[1.13\times10^{-3}\,\text{m},\ 1.17\times10^{-3}\,\text{m}]$. According to (2.5) of Hertz theory, the contact radius is $a_{\text{Hz}}=1.14\times10^{-3}$ m. Thus the maximal possible difference between Hertz theory and FEA results for the contact radius is only 2.7 percent.[7]

Finally, Fig. 4 shows the variation of the contact-area radius with the normal force P. For both loading and unloading, the FEA results closely follow the prediction of (2.5).

The agreement between the FEA results and Hertz theory for elastic normal contact thus validates the FE discretization and analysis procedure. We extend the analysis of this problem to the plastic regime for the development of the present NFD model for elasto-plastic contact.

3 The Elasto-Plastic NFD Model

To construct an NFD model that accounts for both elastic and plastic deformation, we performed nonlinear FEA to observe the behavior of the sphere. We used the same FE discretization and sphere properties as those presented in Section 2.2. In addition, we assumed the sphere material to be elasto-perfectly plastic, with yield stress $\sigma_Y=1.0\times10^8$ N/m^2 (see Vu-Quoc and Lesburg [7] for more details). In this section, we present our NFD model based on our observations of FEA results.

3.1 Incipient Yield Force. Plastic deformation occurs when the normal contact force P exceeds the *incipient yield force* P_Y. According to the von Mises criterion, yield occurs at points in the material at which the second invariant of the stress deviator satisfies

$$J_2 := \frac{1}{6}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2] = \frac{(\sigma_Y)^2}{3}, \tag{3.1}$$

where σ_1, σ_2, and σ_3 are the principal stresses, and σ_Y the material yield stress under uniaxial tension. With the principles stresses along the z-axis in Figure 1 given by (2.8) and (2.9), we obtain the following expression for J_2:

$$J_2=\frac{1}{3}(p_m)^2F(\nu,u),$$

$$F(\nu,u) := \left\{ -(1+\nu)[1-u\tan^{-1}(u)]+\frac{3}{2}(1+u^2)^{-1}\right\}^2, \tag{3.2}$$

where we have introduced a new variable $u := z/a$. Since p_m does not depend on z,[8] the point where yield will occur first is where F in (3.2) is maximized with respect to u. Let u^* be the maximizer of F, and thus of J_2. The value of u^* is a function of ν, but not of p_m, and can be found by solving

$$\frac{\partial F(\nu,u^*)}{\partial u}=0. \tag{3.3}$$

For $\nu=0.3$, we obtain $u^*=0.48086$;[9] the variation of J_2 and $J_2' := \partial J_2/\partial u$ with respect to u is shown in Fig. 5.

At incipient yield, in the expression (3.2), we set $J_2=(\sigma_Y)^2/3$ as in (3.1), and p_m to $p_{m,Y}$, to obtain

$$p_{m,Y}=A_Y(\nu)\sigma_Y, \quad A_Y(\nu) := [F(\nu,u^*)]^{-1/2}, \tag{3.4}$$

[6] See Vu-Quoc and Lesburg [7] for other loading histories.

[7] Note that FEA results show a sharp drop in normal pressure at the center of the contact area ($r=0.0$). We attribute this behavior to the numerical characteristics of the finite element modeling of the axisymmetric problem using ABAQUS.

[8] See (2.4) and (2.7).

[9] Thus incipient yield occurs not at the contact surface, but at a point on the z-axis about one-half the contact radius above the contact surface. We confirmed this result by FEA (see Vu-Quoc and Lesburg [7]).

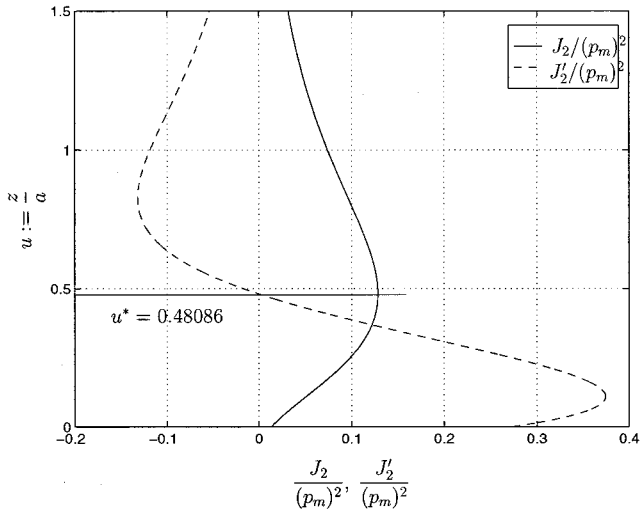


Fig. 5 Variation of J_2 and J_2' along the z-axis for ν=0.3

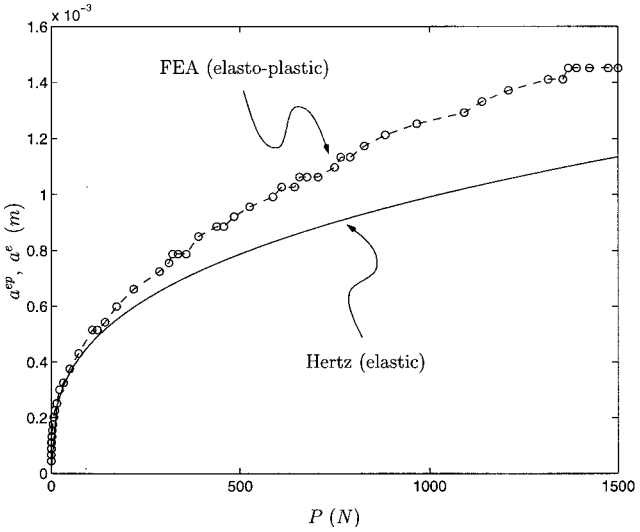


Fig. 6 Contact radius a^{ep} versus normal force P for elasto-plastic contact, with comparison to Hertz theory (elastic)

where $A_Y(\nu)$ is a function of the Poisson ratio ν only. For ν=0.3, we obtain A_Y=1.61; for ν=0.4, we obtain A_Y=1.74. Next, using (2.7) with $p_m = p_{m,Y}$ and $P = P_Y$, we obtain[10]

$$P_Y = \frac{\pi^3 R^2 (1-\nu^2)^2}{6E^2} A_Y^3 \sigma_Y^3. \tag{3.5}$$

Therefore, the *incipient yield force* P_Y for the sphere in our FEA is evaluated using (3.5) to be $P_Y \simeq 36.45$ N.

3.2 Decomposition of the Contact-Area Radius. Let a^{ep} be the radius of the contact area of an elasto-plastic contact under normal contact force P, and a^e the radius of the contact area of elastic contact under the same normal force P. It is easy to come to the following relationship:

$$\begin{cases} a^{ep} = a^e & \text{for } P \leqslant P_Y, \\ a^{ep} > a^e & \text{for } P > P_Y. \end{cases} \tag{3.6}$$

In other words, the effect of plastic deformation is to increase the size of contact-area radius. Figure 6 shows the FEA result of the elasto-plastic contact radius a^{ep} versus normal contact force P for P loading from 0 to 1500 N. The elastic contact radius a^e determined by Hertz theory (i.e., by (2.5)) is also shown in Fig. 6.

Based on the behavior described above and shown in Fig. 6 we introduce the following additive decomposition of the elasto-plastic contact radius a^{ep}:

$$a^{ep} = a^e + a^p, \tag{3.7}$$

where a^e is the elastic part determined by (2.5) of Hertz theory, and a^p the plastic correction part.[11] The next issue is how to model a^p as a function of P to reflect accurately the observed NFD behavior obtained from FEA.

Figure 7 shows the plastic contact radius a^p versus the normal force P for loading P to 1500 N and unloading. We observe from Fig. 7 that, during loading, the plastic radius a^p increases with P in an approximately linear fashion. During unloading, the plastic contact radius a^p does not obviously decrease with the decreasing P; and there is a permanent deformation left after complete unloading. In other words, the contact radius goes to a nonzero residual value (denoted later as a_{res}) as the normal force P goes to zero. Further, we note that the *elastic* part of a^{ep} is *nonlinear* with respect to P, as can be seen from (2.5). Our FEA results of some other loading histories also show similar behavior (see Vu-Quoc and Lesburg [7]). Based on these observations, the plastic contact area a^p in the proposed NFD model is approximated as

$$a^p = \begin{cases} C_a \langle P - P_Y \rangle & \text{for loading} \\ C_a \langle P_{\max} - P_Y \rangle & \text{for unloading} \end{cases}, \tag{3.8}$$

where C_a is a constant that can be determined from the properties of the spheres in contact. For example, C_a for the current elasto-plastic contact problem, the FEA results indicate $C_a = 2.33 \times 10^{-7}$ N/m. The symbol $\langle\ \rangle$ denotes the MacCauley bracket defined by

$$\langle x \rangle = \begin{cases} 0 & \text{for } x \leqslant 0 \\ x & \text{for } x > 0 \end{cases}. \tag{3.9}$$

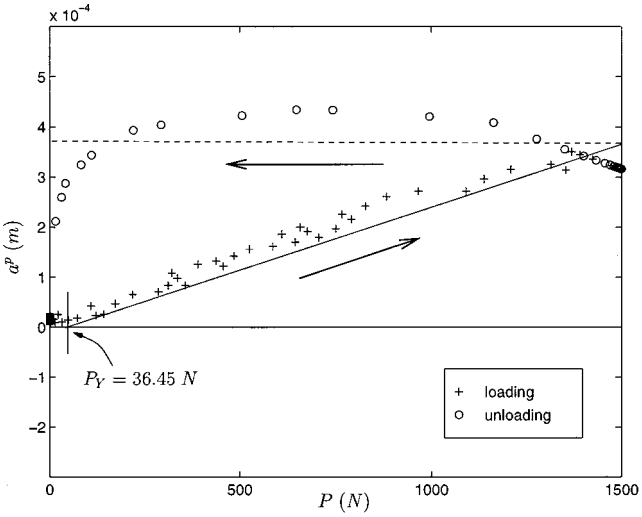


Fig. 7 Plastic contact radius a^p versus normal contact force P. Symbols (+, ○): FEA results. Solid line: model for loading. Dashed line: model for unloading.

[10]Note that (3.5) is for two identical spheres in contact, or for a sphere contacting a frictionless rigid surface. For two spheres of different materials and geometry, we obtain $P_Y = \pi^3 (R^*)^2 A_Y \sigma_Y^3 / [6(E^*)^2]$. (cf. Johnson [4], Eq. (6.10)).

[11]This decomposition is similar to the decomposition of the elastoplastic strain ϵ^{ep} into the sum of the elastic strain ϵ^e and the plastic strain ϵ^p in the theory of elasto-plasticity.

The approximation (3.8) is also consistent with the relation between plastic strain ϵ^p and stress σ in the continuum theory of elasto-plasticity with linear hardening.

This additive decomposition of elasto-plastic contact radius is not only the foundation of the proposed NFD model, but is also crucial in the elasto-plastic TFD model (see Vu-Quoc, Lesburg, and Zhang [14]) since the initial tangential contact stiffness is closely related to the radius of the contact area.

3.3 Normal Pressure Distribution. Fig. 8 shows for our current FEA the normal pressure $p_{\rm FE}$ at $P_{\max}=1500$ N. This figure shows that the normal pressure on the contact surface is approximately constant at a level of 2.3×10^8 N/m^2, or 2.3 times the material's yield stress $\sigma_Y=1.0\times10^8$ N/m^2. FEA results for maximal force levels of $P_{\max}=500$ N and $P_{\max}=1000$ N show similar results for the maximum normal pressure $p_{\rm FE}$, but different elasto-plastic contact area radii a^{ep}. That is, when the normal contact force is much greater than the incipient yield force, i.e., $P\gg P_Y$, the maximum normal pressure is always roughly twice the material yield stress. For detailed results of these other loading histories, see Vu-Quoc and Lesburg [7].

Figure 8 also shows the Hertz prediction (via (2.3)) for the distribution of normal pressure p_H for the same normal force level $P_{\max}=1500$ N. Comparing p_H with $p_{\rm FE}$, we see that the maximum normal pressure from Hertz theory is much larger than that from the FEA results. On the other hand, the radius of contact area from (elastic) Hertz theory is smaller than that from FEA results. Since both normal pressures shown in Fig. 8 arise from the same normal force level $P_{\max}$, the integrals of the normal pressures over the respective contact areas should be the same.

Remark 3.1. The elasto-plastic NFD model proposed by Thornton [12] assumes that the normal pressure on the contact surface is a constant near the center region of the contact area. The maximum normal pressure, called the *contact yield stress* by Thornton [12] is denoted here by $(\sigma_Y)_{th}$. Consistent with Davies ([15], p. 425) and with Johnson ([4], p. 155) we obtain $(\sigma_Y)_{th}\simeq1.61\,\sigma_Y$, which is less than what we obtained from FEA results.[12] To avoid a contentious situation, we refer to this model as *Model T*. The contact-area radius of Model T is therefore larger than our elasto-plastic contact radius a^{ep}. In addition, we note that the tangential stiffness of contact is closely related to the size of the contact area (see Vu-Quoc, Lesburg, and Zhang [14] for more details), an accurate TFD model cannot be developed using Model T. ■

In the present model, we approximate the distribution of the normal pressure on the elasto-plastic contact surface by an elliptic curve as shown by the dashed line in Fig. 8. This normal pressure can be expressed as

$$p_{ep}(r)=(p_m)_{ep}\left[1-\left(\frac{r}{a^{ep}}\right)^2\right]^{1/2}, \tag{3.10}$$

where the maximum normal pressure $(p_m)_{ep}$ is determined by setting the integral of p_{ep} of (3.10) over the elasto-plastic contact area equal to the normal force P. We can see that the shape of the normal pressure distribution of our model is similar to that of Hertz theory, but the distribution is over the elasto-plastic contact area. Such an approximation of normal pressure is motivated by the use of Mindlin and Deresiewicz [5] formalism in constructing the elasto-plastic tangential force-displacement (TFD) model. It is important to keep in mind that although this elliptic approximation of normal pressure is crucial for the construction of our TFD model (presented in Vu-Quoc, Lesburg, and Zhang [14] but will not affect the NFD relationship in the present NFD model.

3.4 Parabola Law: Normal Displacement Versus Contact-Area Radius. Hertz theory gives a parabolic relation between the normal displacement and the radius of the contact area. In the cases of one sphere contacting a rigid frictionless plane, the parabola law (2.6) can be simplified to

$$\alpha_H=\frac{(a_H)^2}{R}, \tag{3.11}$$

where α_H is the normal displacement between the sphere and the plane, a_H the radius of contact area. Note that α_H and a_H are both for *elastic* contact and are both determined via the relations of Hertz theory.

Figure 9 shows the FEA result for the relationship between normal displacement α and elasto-plastic contact radius a^{ep}, for monotonic loading form $P=0$ to $P=1500$ N. The α_H versus a_H relation obtained from Hertz theory (via (3.11)) is also shown. In Fig. 9, the α versus a^{ep} curve from FEA results follows a roughly parabolic relation that is similar to the curve from Hertz theory. Our analyses of other elasto-plastic contact cases reveal the similar trends (see Vu-Quoc and Lesburg [7]).

[12]The value of $(\sigma_Y)_{th}$ was actually $1.587\sigma_Y$ as given by Davies ([15], p. 425) and $1.60\sigma_Y$ as given by Johnson ([4], p. 155).

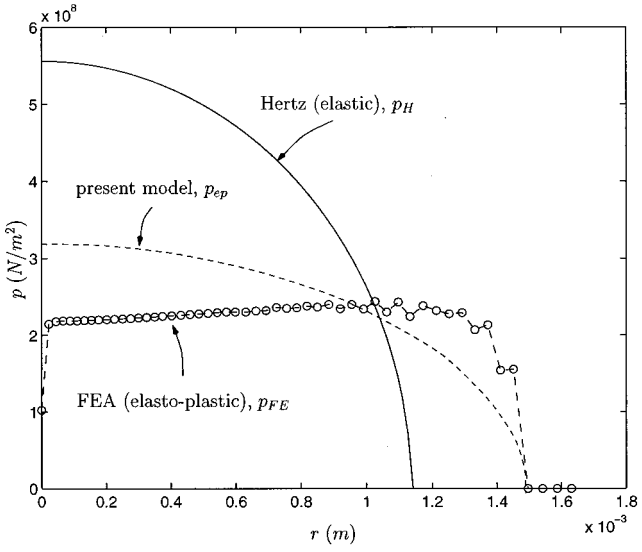


Fig. 8 Normal stress distribution on the contact surface

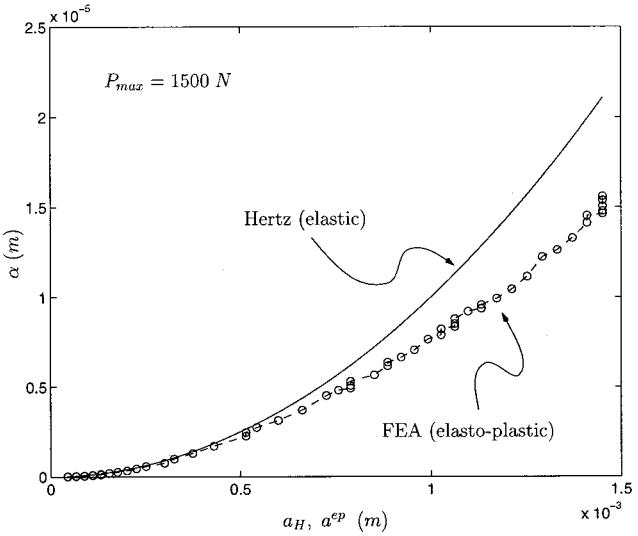


Fig. 9 Normal displacement α versus the radius of total contact area (a^{ep} for elasto-plastic contact, a_H for elastic contact)

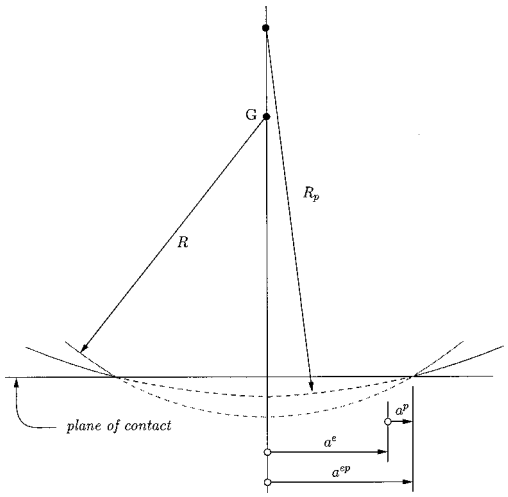


Fig. 10 Plastic deformation increases the radius of relative contact curvature

Based on the FEA results, and considering that (3.11) is from geometry (see Johnson ([4], pp. 84–89) we assume that the α versus a^{ep} relation is parabolic for the loading portion of the force history. That is, when P is increasing, we have

$$2\alpha = \frac{(a^{ep})^2}{R_p^*}, \tag{3.12}$$

where R_p^* is the equivalent radius of the local contact curvature accounting for the effect of plastic deformation. The quantity R_p^* is defined in a manner similar to (2.2) by

$$\frac{1}{R_p^*} := \left(\frac{1}{{}_{(i)}R_p} + \frac{1}{{}_{(j)}R_p} \right). \tag{3.13}$$

For the contact between two identical spheres or between a sphere and a rigid plane, we have $R_p^* = R_p/2$. Therefore, the parabola law can be written as

$$\alpha = \frac{(a^{ep})^2}{R_p}, \tag{3.14}$$

where R_p is the radius of local contact curvature accounting for plastic deformation. Using (3.15) shown shortly below this radius can be computed based on the original radius R and on the level of plastic
deformation.

Figure 10 shows such a change in the local contact curvature. Therefore, we propose to compute the radius R_p of relative curvature by the following:

$$R_p = C_R(P)R, \tag{3.15}$$

where $C_R(P)$ is the coefficient for adjusting the contact radius to account for the plastic deformation. Considering that the plastic deformation tends to flatten the contact surface and that a larger normal force P produces a larger the radius of local contact curvature, we propose to express $C_R(P)$ as

$$C_R(P) = \begin{cases} 1.0 & \text{for } P \leqslant P_Y \\ 1.0 + K_c\langle P - P_Y \rangle & \text{for } P > P_Y \end{cases}, \tag{3.16}$$

where K_c is a constant determined by the sphere properties. For the sphere used in our FEA, we obtained $K_c = 2.69 \times 10^{-4}\ \mathrm{N}^{-1}$. Therefore, when $P \leqslant P_Y$, $C_R(P) = 1.0$ leads to $R_p = R$; and when $P > P_Y$, $C_R(P) > 1.0$ leads to $R_p > R$.

For the case of normal force P unloading after loading to a maximum normal force $P_{\max}$, we assume that the relation between normal displacement and contact area still follows the parabola law. Since the plastic deformation is irreversible, there must exist a residual normal displacement α_{res} after the complete unloading of normal force P. Considering that the unloading should be an elastic process, we propose the parabola law for the P unloading stage to be

$$\alpha - \alpha_{\mathrm{res}} = \frac{(a^e)^2}{(C_R)_{P=P_{\max}} R}. \tag{3.17}$$

Let $\alpha_{\max}$ and $(a^e)_{\max}$ be the normal displacement and the elastic contact radius corresponding to the maximum normal force $P_{\max}$, respectively. Substituting $\alpha_{\max}$, $(a^e)_{\max}$, and $P_{\max}$ into (3.17), we obtain

$$\alpha_{\mathrm{res}} = \alpha_{\max} - \frac{(a^e)_{\max}^2}{(C_R)_{P=P_{\max}} R}. \tag{3.18}$$

The residual normal displacement depends on the maximum normal force $P_{\max}$.

Remark 3.2. Apart from the additive decomposition of the contact-area radius, the use of (3.15) to correct the local curvature at the contact point is another crucial component for our successful FD models. This correction also plays an important role in the TFD model in accounting for the effect of plastic deformation (see Vu-Quoc, Lesburg, and Zhang [14]) ■

Please note that the present NFD model is developed to apply to elasto-plastic contact in combination with the TFD model presented in Vu-Quoc, Lesburg, and Zhang [14]. We performed three-dimensional FEA for elasto-plastic contact subjected to both the normal and tangential forces simultaneously. The three-dimensional FEA results show that the presence of a tangential force does not significantly affect either the P versus α relationship or the plastic flow inside the sphere. In another words, the plastic deformation is mainly caused by the normal contact force P (see Vu-Quoc and Lesburg [7]). Consequently, the present NFD model does not need to be adjusted when a tangential force is present.

Remark 3.3. The Thornton [12] NFD model also proposed a coefficient to modify the radius of contact curvature when there is a plastic deformation. Such a coefficient, however, is applied only for the unloading (P decreasing) stage. In addition, the coefficient from Thornton [12] is almost a constant when P is much larger than the incipient yield force P_Y. More comparisons between the present NFD model and the Thornton [12] NFD model can be found in Vu-Quoc and Zhang [16].

4 Algorithm for the Normal Force-Displacement Model

Based on the discussion in Section 3, the algorithm for the force-driven version of present NFD model is broken into two components: loading and unloading.

During the loading stage of the normal force P, when $P \leqslant P_Y$, the normal displacement α can be determined by Hertz theory. When $P > P_Y$, the elastic contact radius a^e and the plastic contact radius a^p can be computed by (2.5) and (3.8), respectively. Then, the normal displacement α is computed using the parabola law by (3.14), with the elasto-plastic contact radius a^{ep} by (3.7), and with the radius of contact curvature modified by (3.15).

At the start of unloading (P decreasing), the maximum normal force $P_{\max}$, the maximum normal displacement $\alpha_{\max}$, and the maximum elastic contact radius $(a^e)_{\max}$ are all known from previous calculations. Consequently, the residual normal displacement can be computed by (3.18). When $P_{\max} \leqslant P_Y$, the elastic unloading follows Hertz theory. When $P_{\max} > P_Y$, the normal displacement α is computed using the parabola law for unloading by (3.17), with the elastic contact radius a^e from Hertz theory by (2.5).

Algorithm 4.1 shows the detailed pseudocode for the implementation of the force-driven version of the proposed elasto-plastic NFD model. Please note that a boolean variable $nfinc$ (which is an abbreviation for ''*n*ormal *f*orce *inc*reasing'') is used in Algorithm 4.1 for indicating the status of normal force P: When $nfinc=$**true**, P is increasing; when $nfinc=$**false**, P is decreasing.

Algorithm 4.1. Elasto-plastic NFD model: Force-driven version

1 **Data**: R, E, ν, σ_Y, C_a, K_c.
2 *Calculate* P_Y *via* (3.5).
3 **Input**: P_{n-1}, a^{ep}_{n-1}, a^{e}_{n-1}, a^{p}_{n-1}, *and* P_n.
4 **Goal**: *compute* α_n, a^{ep}_n, a^{e}_n, *and* a^{p}_n.
5 *Calculate* $\Delta P_n = P_n - P_{n-1}$.
6 ①**if** $\Delta P_n = 0$
7 *Update* $\alpha_n = \alpha_{n-1}$, $a^{ep}_n = a^{ep}_{n-1}$, $a^{e}_n = a^{e}_{n-1}$, $a^{p}_n = a^{p}_{n-1}$.
8 ①**elseif** $\Delta P_n > 0$ (*Loading*)
9 *set* $nfinc=$**true**.
10 ②**if** $P_n \leqslant P_Y$ (*elastic*)
11 *Calculate* α_n *via* (2.6).
12 *Calculate* a^{e}_n *via* (2.5).
13 $a^{p}_n = 0$ *by* (3.8).
14 $a^{ep}_n = a^{e}_n$ *by* (3.7).
15 ②**elseif** $P_n > P_Y$ (*elasto-plastic*)
16 *Calculate* a^{e}_n *via* (2.5).
17 *Calculate* a^{p}_n *via* (3.8).
18 *Calculate* a^{ep}_n *via* (3.7).
19 *Calculate* R_p *via* (3.16).
20 *Calculate* α_n *via* (3.14).
21 ②**endif**
22 ①**elseif** $\Delta P_n < 0$ (*Unloading*)
23 ③**if** $nfinc=$**true**
24 *set* $P_{max} = P_{n-1}$.
25 *set* $\alpha_{max} = \alpha_{n-1}$.
26 *set* $a^{e}_{max} = a^{e}_{n-1}$.
27 *set* $a^{p}_{max} = a^{p}_{n-1}$.
28 *Calculate* α_{res} *via* (3.18).
29 *set* $nfinc=$**false**.
30 ③**endif**
31 *set* $a^{p}_n = a^{p}_{max}$.
32 *Calculate* a^{e}_n *via* (2.5).
33 *Calculate* α_n *via* (3.17).
34 *Calculate* a^{ep}_n *via* (3.7).
35 ①**endif**

From Algorithm 4.1, we see that a normal displacement α can be determined with a given normal force P for the loading stage. In the unloading stage, with a given normal force P and the maximum normal force $P_{\max}$, the normal displacement α can also be determined directly. In other words, the proposed NFD model is ''total'' rather than incremental.

In addition to the ordinary parameters for the contact such as R, E, ν, and σ_Y, coefficients C_a and K_c, which depend on the geometric and mechanical properties of the spheres in contact, are also needed to implement this proposed NFD model. Coefficients C_a and K_c can be either extracted from FEA results as we do in this paper or extracted using optimization from data of simple experiments as we did in Zhang and Vu-Quoc [17].

5 Numerical Examples

Using the proposed elasto-plastic NFD model, we applied the loading paths ABC, ADE, and AFG shown in Fig. 11 to the problem of a sphere pressed against a frictionless rigid plane. The geometric and material properties are the same as those listed in Section 3. Figures 12–15 show the NFD curves and related coef-

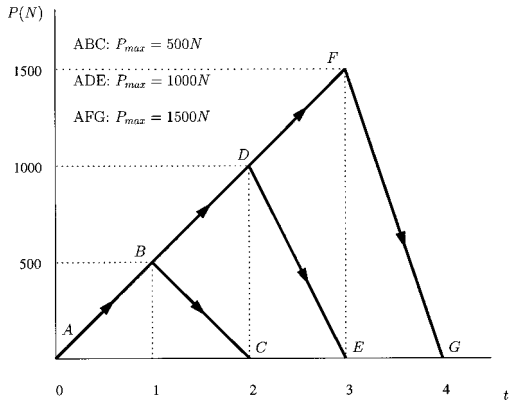


Fig. 11 Loading paths of normal force

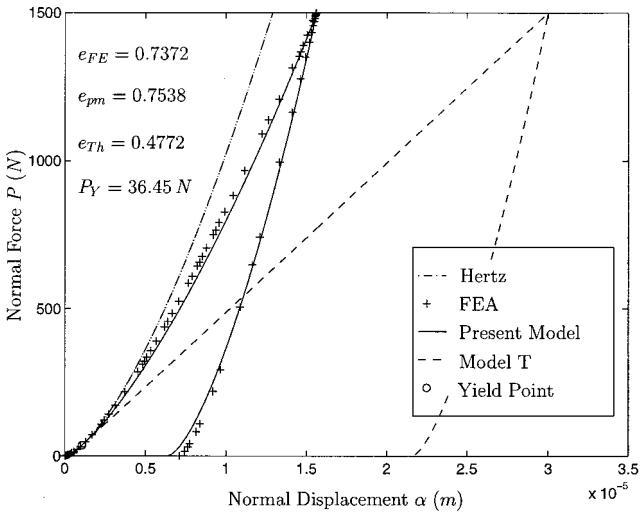


Fig. 12 Normal force P versus normal displacement α by different models for the loading path AFG in Fig. 11

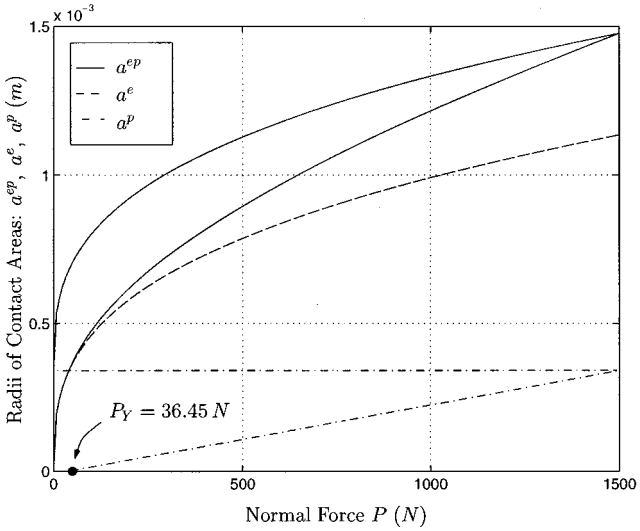


Fig. 13 Contact areas radii a^{ep}, a^{e}, a^{p} versus normal force P by the proposed elasto-plastic NFD model for the loading path: AFG in Fig. 11

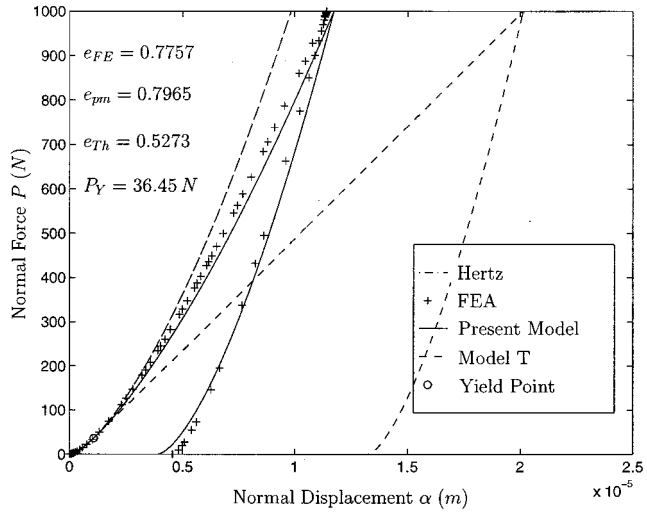


Fig. 14 Normal force *P* versus normal displacement α by different models for the loading path *ADE* in Fig. 11

ficients of restitution generated by our elasto-plastic NFD model. In each figure, we compare the results of our model with both Hertz theory and with the results from Model T (see Remark 3.1). Please note that all the NFD curves shown in these figures are produced in a force-driven procedure, i.e., the force history is input to produce the displacement history. The coefficient of restitution, representing the energy dissipation during a collision, is originally defined as the ratio of outgoing velocity to the incoming velocity when a particle collides with a static half-space. The different coefficients of restitution shown in this section are equivalently obtained by taking the square root of the ratio of the compression energy to the releasing energy during the loading process, i.e.,

$$e = \left(\frac{\text{area under the unloading curve}}{\text{area under the loading curve}} \right)^{1/2}. \qquad (5.1)$$

Figure 12 shows various P versus α curves for the loading path AFG ($P_{\max}=1500$ N) shown in Fig. 11. The coefficient of restitution from the results of the proposed elasto-plastic NFD model is $e_{pm}=0.7538$, while the coefficient of restitution from FEA results is $e_{\mathrm{FE}}=0.7372$; the difference is only 2.2 percent. It can be seen that the P versus α curve produced by the proposed elasto-plastic NFD model agrees with the P versus α curve produced by FEA. The P versus α curve produced by Model T is, however, much too soft. That is, one obtains a much larger displacement for the same force level, compared to FEA results. At the maximum normal force $P_{\max}=1500$ N, the normal displacement predicted by Model T, $(\alpha_{\max})_{th}\simeq 3.00\times 10^{-5}$ m, is about twice as the corresponding FEA result, $(\alpha_{\max})_{\mathrm{FE}}\simeq 1.56\times 10^{-5}$ m. The maximum displacement from the proposed elasto-plastic NFD model is about the same as the corresponding FEA result. The corresponding coefficient of restitution from Model T is $e_{th}=0.4772$, which is 35.3 percent lower than that predicted by FEA. This difference in the coefficient of restitution is because Model T predicts an energy dissipation (i.e., the area enclosed by the loading and unloading curves and the α axis) that is much larger than that predicted by FEA.

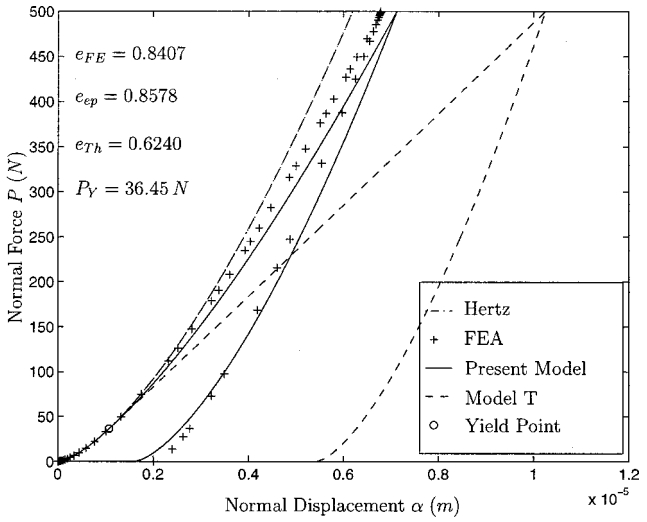


Fig. 15 Normal force *P* versus normal displacement α by different models for the loading path *ABC* in Fig. 11

Figure 13 shows the contact area radii a^{ep}, a^{e}, and a^{p}, versus the normal force P as produced by the proposed elasto-plastic NFD model for loading path AFG ($P_{\max}=1500$ N). In our elasto-plastic NFD model, the a^{e} versus P curve is based on the Hertz theory expression (2.5). The a^{p} versus P curve follows (3.8), and agrees with FEA results shown in Fig. 7. The total contact area radius a^{ep} shown in Fig. 13 is simply the sum of a^{e} and a^{p}, as implied in (3.7), and agrees with the FEA results shown in Fig. 6.

Figure 14 shows the P versus α curves for the loading path ADE ($P_{\max}=1000$ N). The results from the proposed NFD model agree closely with FEA results. The corresponding coefficient of restitution from the proposed elasto-plastic NFD model is $e_{pm}=0.7757$, while the coefficient of restitution from FEA results is $e_{\mathrm{FE}}=0.7965$; the difference between them is small (2.6 percent). As in the previous case, the results from Model T display a contact behavior that is too soft, with much larger maximum displacement and much larger energy dissipation. Quantitatively, the maximal displacement $(\alpha_{\max})_{Th}$ obtained from Model T is $(\alpha_{\max})_{Th}\simeq 2.00\times 10^{-5}$ m; from FEA, we obtain $(\alpha_{\max})_{\mathrm{FE}}\simeq 1.14\times 10^{-5}$ m; and from the proposed elasto-plastic NFD model, $(\alpha_{\max})_{pm}\simeq 1.17\times 10^{-5}$ m. The coefficient of restitution from Model T is $e_{Th}=0.5273$, which differs from e_{FE} and e_{pm} by about 33.8 percent.

Figure 15 shows the P versus α curve for the loading path ABC ($P_{\max}=500$ N) Again, it can be seen that results from the proposed elasto-plastic NFD model agree closely with FEA results. The corresponding coefficient of restitution from the proposed elasto-plastic NFD model is $e_{pm}=0.8578$, which agrees well to the coefficient of restitution from FEA results, $e_{\mathrm{FE}}=0.8407$ with a difference of 2.0 percent. Similar to the previous two loading histories, Model T yields results that are much too soft, with the maximum displacement $(\alpha_{\max})_{Th}\simeq 1.11\times 10^{-5}$ m, being much larger than the corresponding maximum displacement $(\alpha_{\max})_{\mathrm{FE}}\simeq 0.68\times 10^{-5}$ m from FEA results and the maximum displacement $(\alpha_{\max})_{pm}\simeq 0.71\times 10^{-5}$ m from the proposed elasto-plastic NFD model. Similarly, the energy dissipation in Model T is also much larger than that in FEA results. The coefficient of restitution from Model T is $e_{Th}=0.6240$, which is about 25.8 percent smaller than e_{FE}.

In summary, the proposed elasto-plastic NFD model produces not only an accurate P versus α relationship, but also a correct coefficient of restitution and energy dissipation compared with FEA results. Model T produces much softer P versus α relationship, smaller coefficient of restitution, and larger energy dissipation, for the same maximum normal force level.

Remark 5.1. The soft contact behavior of Model T is partially attributed to the fact that this model predicts that the P versus α curve is a straight line after the plastic deformation occurs during loading. The NFD curves obtained from applying the *displacement*-driven versions of the proposed NFD model and Model T to the same displacement paths show similar results, i.e., the normal force level produced by Model T from the same nor-

mal displacement is much smaller than that by the proposed NFD model as presented in Vu-Quoc and Zhang [16]. The results from the discrete element (DEM) simulation of a sphere colliding with a rigid half-space using different models also reveal similar results. We refer the readers to Vu-Quoc and Zhang [16] for more details. ■

6 Conclusion

We have presented a model for the normal force-displacement (NFD) relation of contacting spherical particles, accounting for the effects of both elastic and plastic deformation. This NFD model, based on the formalism of the continuum theory of elastoplasticity, is to be used in granular flow simulations involving thousands of particles (see also Vu-Quoc et al. [2] and Vu-Quoc and Zhang [16]); the efficiency of the model is thus a crucial property. The novelty of the present NFD model is the additive decomposition of the contact-area radius, and the correction of the radii of the particles at the contact point due to plastic flow. The presented construction and the algorithm of this model show that the implementation of this model is *simple* and that the computation using this model is efficient. (The displacement-driven version of this NFD model is also simple and efficient, see Vu-Quoc and Zhang [16].) A comparison of the FD results produced by this model with nonlinear elasto-plastic FEA results validated the accuracy of the proposed model. The same formalism for the construction of the proposed NFD model is also employed to construct a tangential force-displacement (TFD) model accounting for the effects of both elastic deformation and plastic deformation (presented in Vu-Quoc et al. [14].) Therefore, the presented NFD model and the TFD model form a set of consistent FD models for the simulation of contacting spherical particles. For additional information on the application of previous FD models to granular-flow simulations, the readers are referred to [18–21].

Acknowledgments

The support of the National Science Foundation is gratefully acknowledged.

References

[1] Cundall, P., and Strack, O., 1979, ''A Discrete Numerical Model for Granular Assemblies,'' Géotech., **29**, No. 1, pp. 47–65.
[2] Vu-Quoc, L., Zhang, X., and Walton, O. R., 2000, ''A 3-D Discrete Element Method for Dry Granular Flows of Ellipsoidal Particles,'' Comput. Methods Appl. Mech. Eng., invited paper for the special issue on Dynamics of Contact/Impact Problems, to appear.
[3] Hertz, H., 1882, ''Über die Berührung fester elastischer Körper (On the Contact of Elastic Solids),'' J. Reine Angew. Math., **92**, pp. 156–171.
[4] Johnson, K. L., 1985, *Contact Mechanics*, 2nd Ed., Cambridge University Press, New York.
[5] Mindlin, R. D., and Deresiewicz, H., 1953, ''Elastic Spheres in Contact Under Varying Oblique Forces,'' ASME J. Appl. Mech., **20**, pp. 327–344.
[6] Shih, C. W., Schlein, W. S., and Li, J. C. M., 1992, ''Photoelastic and Finite Element Analysis of Different Size Spheres in Contact,'' J. Mater. Res., **7**, No. 4, pp. 1011–1017.
[7] Vu-Quoc, L., and Lesburg, L., 2000, ''Contact Force-Displacement Relations for Spherical Particles Accounting for Plastic Deformation,'' Int. J. Solids Struct., submitted for publication.
[8] Dobry, R., Petrakis, E., and Seridi, A., 1991, ''General Model for Contact Law Between Two Rough Spheres,'' J. Eng. Mech., **117**, No. 6, pp. 1365–1381.
[9] Walton, O. R., and Braun, R. L., 1986, ''Viscosity, Granular-Temperature, and Stress Calculations for Shearing Assemblies of Inelastic, Frictional Disks,'' J. Rheol., **30**, No. 5, pp. 949–980.
[10] Goldsmith, W., 1960, *Impact: The Theory and Physical Behavior of Colliding Solids*, Edward Arnold, London.
[11] Kangur, K. F., and Kleis, I. R., 1988, ''Experimental and Theoretical Determination of the Coefficient of Velocity Restitution Upon Impact,'' Mech. Solids, **23**, No. 5, pp. 182–185.
[12] Thornton, C., 1997, ''Coefficient of Restitution for Collinear Collisions of Elastic Perfectly Plastic Spheres,'' ASME J. Appl. Mech., **64**, pp. 383–386.
[13] Brilliantov, N. V., Spahn, F., Hertzsch, J., and Poschel, T., 1996, ''Model for Collisions on Granular Gases,'' Phys. Rev. E, **53**, No. 5, pp. 5382–5392.
[14] Vu-Quoc, L., Lesburg, L., and Zhang, X., 1999, ''A Tangential Force-Displacement Model for Contacting Spheres Accounting for Plastic Deformation: Force-Driven Formulation,'' J. Mech. Phys. Solids, submitted for publication.
[15] Davies, R. M., 1949, ''The Determination of Static and Dynamic Yield Stresses Using a Steel Ball,'' Proc. R. Soc. London, Ser. A, **197**, pp. 416–432.
[16] Vu-Quoc, L., and Zhang, X., 1999, ''An Elasto-Plastic Contact Force-Displacement Model in the Normal Direction: Displacement-Driven Version,'' Proc. R. Soc. London, Ser. A, **455**, No. 1991, pp. 4013–4044.
[17] Zhang, X., and Vu-Quoc, L., 2000, ''A Method to Extract the Mechanical Properties of Particles in Collision Based on a New Elasto-Plastic Normal Force-Displacement Model,'' Int. J. Plast., submitted for publication.
[18] LoCurto, G. J., Zhang, X., Zakirov, V., Bucklin, R. A., Vu-Quoc, L., Hanes, D. M., and Walton, O. R., 1997, ''Soybean Impacts: Experiments and Dynamic Simulations,'' Trans. Am. Soc. Agr. Eng. (ASAE), **40**, No. 3, pp. 789–794.
[19] Vemuri, B. C., Chen, L., Vu-Quoc, L., Zhang, X., and Walton, O. R., 1998, ''Efficient and Accurate Collision Detection for Granular Flow Simulation,'' Graph. Models Image Process., **60**, No. 6, pp. 403–422.
[20] Vu-Quoc, L., and Zhang, X., 1999, ''An Accurate and Efficient Tangential Force-Displacment Model for Elastic-Frictional Contact in Particle-Flow Simulations,'' Mech. Mater., **31**, pp. 235–269.
[21] Zhang, X., and Vu-Quoc, L., 2000, ''Simulation of Chute Flow of Soybeans Using an Improved Tangential Force-Displacement Model,'' Mech. Mater., **32**, No. 2, pp. 115–129.

On Higher-Order Crack-Tip Fields in Creeping Solids

B. N. Nguyen
Postdoctoral Researcher

P. R. Onck
Postdoctoral Researcher

E. van der Giessen
Professor

Delft University of Technology,
Koiter Institute Delft,
Mekelweg 2, 2628 CD Delft,
The Netherlands

In view of the near-tip constraint effect imposed by the geometry and loading configuration, a creep fracture analysis based on C only is generally not sufficient. This paper presents a formulation of higher-order crack-tip fields in steady power-law creeping solids which can be derived from an asymptotic development of near-tip fields analogous to that of Sharma and Aravas and Yang et al. for elastoplastic bodies. The higher-order fields are controlled by a parameter named A_2^*, similar as in elastoplasticity, and a second loading parameter, σ_∞. By means of the scaling properties for power-law materials, it is shown that A_2^* for a flat test specimen is independent of the loading level. Finally, we carry out small-strain finite element analyses of creep in single-edge notched tension, centered crack panel under tension, and single-edge notched bending specimens in order to determine the corresponding values of A_2^* for mode I cracks under plane-strain conditions.* [S0021-8936(00)01202-2]

1 Introduction

There is growing appreciation of the role of higher-order terms in the asymptotic stress and deformation fields near cracks in nonlinear materials. In elastic materials, Williams [1] was the first to publish the second term in the series expansion of the asymptotic field; since then this term is commonly denoted as the *T*-stress. Together with the magnitude of the leading-order square-root singular term governed by the stress intensity factor *K*, this yields a two-parameter characterization of the near-tip crack fields through $\{K,T\}$. In the context of the experimental testing of fracture properties, the significance of the *T*-stress, parallel to the crack in mode I loading, lies in the dependence of *T* on the geometry of the test specimen. Later, Li and Wang [2] performed a two-term asymptotic analysis for a mode I crack under plane-strain conditions for power-law hardening materials. Their second-order parameter along with the *J*-integral then describe the near-tip fields.

At the beginning of this decade, Betegón and Hancock [3] proposed a two-parameter characterization of elastic-plastic crack-tip fields by using a modified boundary layer formulation based on the first two terms *K* and *T*. They concluded that geometries, which maintain *J*-dominance, are characterized by zero or positive values of *T*-stresses while geometries with negative *T* stresses can be described by a two-parameter characterization using *J* and *T*. O'Dowd and Shih [4] have advanced a similar two-parameter characterization of the crack-tip fields in elastic-plastic solids in terms of the value of the *J*-integral and a constraint factor *Q*, i.e., $\{J,Q\}$. Furthermore, they take *Q* as the near-tip hydrostatic stress normalized by the yield stress σ_0. Detailed comparison with finite element calculations ([4]) has confirmed that the hydrostatic stress field is constant to a fair degree of accuracy over a wedge-shaped zone in front of the crack tip. They also devised a procedure to relate *Q* to the elastic *T*-stress under small-scale yielding conditions. This two-parameter characterization has been quite successful in the classification of brittle compared to ductile failure in many metals ([5,6]).

In the same period, Sharma and Aravas [7] performed an analytical asymptotic analysis of the fields in an elastoplastic material. Their analysis has been carried out in terms of a J_2 deformation theory of plasticity, assuming the following stress (σ_{ij})–strain (ϵ_{ij}) relationship (in Cartesian coordinates):

$$\epsilon_{ij}=\frac{1+\nu}{E}s_{ij}+\frac{1-2\nu}{E}\frac{1}{3}\sigma_{kk}\delta_{ij}+\frac{3}{2}\alpha\epsilon_e\frac{s_{ij}}{\sigma_e} \quad \text{if } \sigma_e\geqslant\sigma_0 \tag{1}$$

where ϵ_e is given by the power-law hardening expression

$$\epsilon_e=\epsilon_0(\sigma_e/\sigma_0)^n \tag{2}$$

with σ_e the Mises stress, $\sigma_e=\sqrt{3/2s_{ij}s_{ij}}$ and s_{ij} the deviatoric part of the stress tensor. In (1), *E* is Young's modulus, ν Poisson's ratio, *n* is the hardening exponent, $\epsilon_0=\sigma_0/E$ is the elastic yield strain, and α is a material parameter. The series expansion of the near-tip stress field is found to be of the form ([7])

$$\frac{\sigma_{ij}(r,\theta)}{\sigma_0}=\left(\frac{J}{\alpha\epsilon_0\sigma_0 I_n r}\right)^{1/n+1}\tilde{\sigma}_{ij}^{(0)}(\theta)+Q\left(\frac{r}{J/\sigma_0}\right)^p\tilde{\sigma}_{ij}^{(1)}(\theta)+\ \ldots, \tag{3}$$

where *r* and θ are polar coordinates centered at the crack tip, and where the $\tilde{\sigma}_{ij}^{(k)}(\theta)$, $k=0,1$ are normalized nondimensional functions corresponding to the scaling factor I_n. The first term in (3) corresponds to the HRR field ([8,9]) with the nondimensional function $\tilde{\sigma}_{ij}^{(0)}$ being tabulated by Shih [10]. According to (3), the second-order term is actually not constant in (r,θ). However, the exponent *p* is found to be rather small: $p\in(-0.102,\ 0.053)$ for realistic values of *n*, $n\in(2,\ 20)$ ([7]). The dimensionless functions for $k=1$ have been tabulated by Sharma [11]. A few years later, Xia et al. [12] performed a similar asymptotic analysis up to the fifth-order term, and Yang et al. [13] carried out a complete analysis of higher-order terms for both mode I and mode II loadings. The latter authors, as well as Nikishkov et al. [14] also applied the three-term expansion to some important selected test specimens (e.g., single-edge notched bending (SENB), centered crack panel under tension (CCT)).

What we are interested in here are the higher-order near-tip fields in a steadily creeping power-law material. Riedel and Rice [15] demonstrated that the leading-order terms in the expansion can be directly obtained from those for elastic-plastic materials by invoking the Hoff analogy, i.e., replacing displacements with displacement rates and strains by strain rates in all governing equations. The hardening law (2) then transforms into the creep law $\dot{\epsilon}_e=\dot{\epsilon}_0(\sigma_e/\sigma_0)^n$, and the singular term in the near-tip stress field becomes equivalent to the leading term ($k=0$) in (3) when *J* is replaced with C^*, and $\alpha\epsilon_0$ with $\dot{\epsilon}_0$. Sharma and Aravas [16] have suggested to proceed along these lines to obtain the higher-order terms. Thus, they replace the radial dependence of the second

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 17, 1998; final revision, Oct. 5, 1999. Associate Technical Editor: I. M. Daniel. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

term in (3), $r/(J/\sigma_0)$, with $r/(C^*/\dot{\epsilon}_0\sigma_0)$ (the additional factor $\dot{\epsilon}_0$ is needed here to nondimensionalize the time dimension in C^*).

However, here we arrive at a principal difficulty which is associated with the fact that the creep parameters σ_0 and $\dot{\epsilon}_0$ are not independent. In fact, the creep law can be written in a fully equivalent manner as

$$\dot{\epsilon}_e = B\sigma_e^n, \tag{4}$$

with the single material parameter B replacing the combination $\dot{\epsilon}_0/\sigma_0^n$. Careful consideration of the leading term in the asymptotic expansion shows that, owing to the particular power $1/(n+1)$, the coefficient can indeed be expressed in terms of B (see also [15]). However, the second and higher-order terms cannot be immediately re-grouped in terms of the combination $B = \dot{\epsilon}_0/\sigma_0^n$. Henceforth, we have to conclude that the simple application of Hoff's analogy to (3) cannot deliver the correct higher-order fields. The underlying reason for this is that σ_0 in plasticity not only enters in (2), where ϵ_0/σ_0^n is the only parameter [cf. (4)], but also serves as a separate quantity to define the yield surface. The latter is absent in creep.

The purpose of this paper therefore is to reconsider the derivation of the higher-order crack-tip fields in a steadily creeping solid.

2 Dimensional Considerations

We consider a crack under tensile loading (mode I) in an elastic power-law creeping material governed by the constitutive equations

$$\dot{\epsilon}_{ij} = \frac{1+\nu}{E}\dot{s}_{ij} + \frac{1-2\nu}{E}\frac{1}{3}\dot{\sigma}_{kk}\delta_{ij} + \frac{3}{2}\dot{\epsilon}_e\frac{s_{ij}}{\sigma_e}, \tag{5}$$

with $\dot{\epsilon}_e$ according to (4). A creep zone around the crack is defined as the region in which the creep strains are larger or equal to the elastic strains ([15]). If the creep zone is small compared to the crack length (or specimen size), the crack-tip situation is controlled by the stress intensity factor K_I. We concentrate on the case of extensive creep, i.e., when the creep zone is not small compared to the crack length, for which the value of $C(t)$ is the relevant loading parameter ([15]). Under constant remote loading conditions and if the crack is stationary, steady-state conditions are attained after sufficiently long times and $C(t)$ approaches C^*, i.e.,

$$C^* = \int_\Gamma \left(\dot{w}dx_2 - \sigma_{ij}n_j\frac{\partial \dot{u}_i}{\partial x_1}ds \right) \tag{6}$$

with, for power-law creep,

$$\dot{w} = \int_0^{\dot{\epsilon}_{kl}} \sigma_{ij}d\dot{\epsilon}_{ij} = \frac{n}{n+1}\sigma_e\dot{\epsilon}_e, \tag{7}$$

where the Cartesian coordinate system x_i is centered at the crack tip and Γ is an arbitrary counterclockwise path around the crack tip parametrized by s.

When the near-tip fields have approached the steady state, i.e., $\dot{\sigma}_{ij}=0$, the constitutive Eqs. (5) only involve the material parameters n and B. It then follows purely from dimensional considerations that the stress field must scale with $(C^*/Br)^{1/(n+1)}$ when the loading is governed solely by C^*. This agrees exactly with the leading-order term that one obtains from (3) when J is replaced with C^*, and $\alpha\epsilon_0$ with $\dot{\epsilon}_0$ ([15,16]). Obviously, the higher-order terms cannot scale in the same way; but, there is no other way available with just C^*, B, and n. As the constitutive model does not involve a separate parameter with dimensions of stress, we must introduce a second loading parameter, with dimensions of stress, σ_∞. There is not a unique definition of this parameter. One possibility is to identify σ_∞ with the net average stress over the ligament in a test specimen, as is done by Riedel [17] when normalizing the C^*-integral (see also Section 5.1). Another possibility is to identify σ_∞ with the Mises stress that would be acting if there were no crack. This definition applies to test specimens as well as to components.

Hence, the steady-state stress field is a function of r, θ, B, n, C^*, and σ_∞. Dimensional considerations then show that apart from r there is only a single parameter with dimensions of length, namely

$$L = C^*/B\sigma_\infty^{n+1}. \tag{8}$$

Moreover, one finds that the near-tip stress field must be of the form

$$\hat{\sigma}_{ij} = f_{ij}(\hat{r}, \theta, n),$$

and similar forms result for the strain rate $\dot{\epsilon}_{ij}$ and the velocity $\dot{u}_i$. Here and in the sequel, the $(\hat{})$ denotes nondimensionalization on the basis of σ_∞ and L:

$$\hat{r} \equiv \frac{r}{L}, \quad \hat{\sigma}_{ij} \equiv \frac{\sigma_{ij}}{\sigma_\infty}, \quad \hat{\dot{\epsilon}}_{ij} \equiv \frac{\dot{\epsilon}_{ij}}{B\sigma_\infty^n}, \quad \hat{\dot{u}}_i \equiv \frac{\dot{u}_i}{B\sigma_\infty^n L}. \tag{9}$$

With the associated normalizations

$$\hat{x}_i \equiv \frac{x_i}{L}, \quad \hat{\dot{w}} = \frac{\dot{w}}{B\sigma_\infty^{n+1}},$$

the C^* integral (6) becomes

$$C^* = B\sigma_\infty^{n+1}L\int_\Gamma \left(\hat{\dot{w}}d\hat{x}_2 - \hat{\sigma}_{ij}n_j\frac{\partial \hat{\dot{u}}_i}{\partial \hat{x}}d\hat{s} \right).$$

It then follows from (7)–(8) that the normalized near-tip fields must be scaled so that

$$\int_\Gamma \left(\frac{n}{n+1}\hat{\sigma}_e^{n+1}d\hat{x}_2 - \hat{\sigma}_{ij}n_j\frac{\partial \hat{\dot{u}}_i}{\partial \hat{x}}d\hat{s} \right) = 1. \tag{10}$$

3 Three-Term Expansion of Near-Tip Fields

With an asymptotic development analogous to Sharma-Aravas's ([7]) but going further than the second term as Yang et al. [13] did in their elastoplastic analysis, the near-tip stress field for creeping solids under plane-strain mode I loading conditions can be written as

$$\hat{\sigma}_{ij}(r,\theta) = Q_0[\hat{r}^s\hat{\sigma}_{ij}^{(0)}(\theta) + Q_1\hat{r}^p\hat{\sigma}_{ij}^{(1)}(\theta) + Q_2\hat{r}^q\hat{\sigma}_{ij}^{(2)}(\theta) + \ldots] \quad \text{as } \hat{r}\to 0, \tag{11}$$

where $\hat{\sigma}_{ij}^{(k)}$ $(k=0,1,2)$ are to be determined later. As Yang et al. [13] pointed out, for $n \geqslant 3$, the first three terms are controlled by plastic strains and not by elastic strains. Therefore, the use of a three-term expansion for steadily creeping solids is fully justified.

The nondimensional deviatoric stress field is readily obtained as

$$\hat{s}_{ij}(r,\theta) = Q_0[\hat{r}^s\hat{s}_{ij}^{(0)}(\theta) + Q_1\hat{r}^p\hat{s}_{ij}^{(1)}(\theta) + Q_2\hat{r}^q\hat{s}_{ij}^{(2)}(\theta) + \ldots] \quad \text{as } \hat{r}\to 0. \tag{12}$$

Expressing the effective Mises stress in terms of deviatoric stress components and limiting the development to dominant terms, we obtain the following expansion:

$$\hat{\sigma}_e = Q_0[\hat{r}^s\hat{\sigma}_e^{(0)} + Q_1\hat{r}^p\hat{\sigma}_e^{(1)} + Q_2\hat{r}^q\hat{\sigma}_e^{(2)} + \ldots] \quad \text{as } \hat{r}\to 0 \tag{13}$$

with

$$\hat{\sigma}_e^{(0)} = \left(\frac{3}{2}\hat{s}_{ij}^{(0)}\hat{s}_{ij}^{(0)} \right)^{1/2}, \quad \hat{\sigma}_e^{(1)} = \frac{3}{2}\frac{\hat{s}_{ij}^{(0)}\hat{s}_{ij}^{(1)}}{\hat{\sigma}_e^{(0)}}, \quad \text{and}$$

$$\hat{\sigma}_e^{(2)} = \frac{3}{2}\frac{\hat{s}_{ij}^{(0)}\hat{s}_{ij}^{(2)}}{\hat{\sigma}_e^{(0)}}. \tag{14}$$

The strain-rate field can be immediately obtained from the constitutive Eq. (5) for steady conditions by substitution of (12) and (13). Making use of the nondimensionalization (9), and retaining again only up to the third-order terms, we find

$$\hat{\epsilon}_{ij} = Q_0^n[\hat{r}^{sn}\hat{\epsilon}_{ij}^{(0)} + Q_1\hat{r}^{s(n-1)+p}\hat{\epsilon}_{ij}^{(1)} + Q_2\hat{r}^{s(n-1)+q}\hat{\epsilon}_{ij}^{(2)} + \dots] \quad \text{as } \hat{r}\to 0 \tag{15}$$

where

$$\hat{\epsilon}_{ij}^{(0)} = \frac{3}{2}\hat{s}_{ij}^{(0)}[\hat{\sigma}_e^{(0)}]^{n-1}, \tag{16}$$

$$\hat{\epsilon}_{ij}^{(k)} = \frac{3}{2}[\hat{\sigma}_e^{(0)}]^{n-1}\left[\hat{s}_{ij}^{(k)} + (n-1)\hat{s}_{ij}^{(0)}\frac{\hat{\sigma}_e^{(k)}}{\hat{\sigma}_e^{(0)}}\right], \quad k=1,2. \tag{17}$$

The expression for the displacement rate can be obtained by integration of the strain-displacement relations in a way analogous to that of Li and Wang [2] in their elastoplastic analysis. For completeness, we list the result here in terms of the normalized quantities according to (9):

$$\hat{u}_i = Q_0^n[\hat{r}^{sn+1}\hat{u}_i^{(0)} + Q_1\hat{r}^{s(n-1)+p+1}\hat{u}_i^{(1)} + Q_2\hat{r}^{s(n-1)+q+1}\hat{u}_i^{(2)} + \dots] \quad \text{as } \hat{r}\to 0 \tag{18}$$

where

$$\hat{u}_r^{(0)} = (n+1)\hat{\epsilon}_{rr}^{(0)}, \quad \hat{u}_r^{(1)} = \frac{\hat{\epsilon}_{rr}^{(1)}}{p+\dfrac{2}{n+1}} \quad \text{and} \quad \hat{u}_r^{(2)} = \frac{\hat{\epsilon}_{rr}^{(1)}}{q+\dfrac{2}{n+1}} \tag{19}$$

$$\hat{u}_\theta^{(0)} = \frac{2\hat{\epsilon}_{r\theta}^{(0)} - \hat{u}_{r,\theta}^{(0)}}{\dfrac{-n}{n+1}}, \quad \hat{u}_\theta^{(1)} = \frac{2\hat{\epsilon}_{r\theta}^{(1)} - \hat{u}_{r,\theta}^{(1)}}{p-\dfrac{n-1}{n+1}} \quad \text{and}$$

$$\hat{u}_\theta^{(2)} = \frac{2\hat{\epsilon}_{r\theta}^{(2)} - \hat{u}_{r,\theta}^{(2)}}{q-\dfrac{n-1}{n+1}} \tag{20}$$

with $(\,)_{,\theta}$ denoting differentiation with respect to θ.

As mentioned in Section 2, the solution must satisfy the normalization condition (10). Evaluation of (10) along a circular contour around the crack tip for $\hat{r}\to 0$ yields that

$$s = \frac{-1}{n+1}, \quad Q_0 = \left(\frac{1}{I_n}\right)^{1/n+1}, \tag{21}$$

with (cf. [10])

$$I_n = \int_{-\pi}^{\pi}\left(\frac{n}{n+1}[\hat{\sigma}_e^{(0)}]^{n+1}\cos\theta - \left[\sin\theta(\hat{\sigma}_{rr}^{(0)}(\hat{u}_\theta^{(0)} - \hat{u}_{r,\theta}^{(0)}) - \hat{\sigma}_{r\theta}^{(0)}(\hat{u}_r^{(0)} + \hat{u}_{\theta,\theta}^{(0)})) + \frac{1}{n+1}\cos\theta(\hat{\sigma}_{rr}^{(0)}\hat{u}_r^{(0)} + \hat{\sigma}_{r\theta}^{(0)}\hat{u}_\theta^{(0)})\right]\right)d\theta, \tag{22}$$

provided that the dimensionless functions $\hat{\sigma}_{ij}^{(0)}(\theta)$ are normalized so that $\max \hat{\sigma}_e^{(0)}(\theta)=1$. With (21), the stress field can be rewritten in dimensional form as

$$\sigma_{ij}(r,\theta) = \left(\frac{C^*}{BI_nL}\right)^{1/n+1}\left[\left(\frac{r}{L}\right)^{-1/n+1}\hat{\sigma}_{ij}^{(0)}(\theta) + Q_1\left(\frac{r}{L}\right)^p\hat{\sigma}_{ij}^{(1)}(\theta) + Q_2\left(\frac{r}{L}\right)^q\hat{\sigma}_{ij}^{(2)}(\theta) + \dots\right] \tag{23}$$

to elucidate that the first-order term represents the HRR creep solution.

The final solution for the higher-order parameters p, q, as well as the associated functions $\hat{\sigma}_{ij}^{(k)}$, etc., $(k=1,2)$ follows from substitution of the expansion (11) into the equilibrium equation and the appropriate boundary conditions. This analysis has been carried out for elastoplastic solids by Yang et al. [13] in terms of an equivalent asymptotic expansion as used here, be it that their length scale L is not defined by (8) but regarded as an independent parameter. They derive that the third term in the expansion is related to the second term through

$$q = 2p - s = 2p + \frac{1}{n+1} \quad \text{for } n\geqslant 3, \tag{24}$$

and the final result can be written as (cf. Eq. (3))

$$\frac{\sigma_{ij}(r,\theta)}{\sigma_0} = \left(\frac{J}{\alpha\sigma_0\epsilon_0 I_n L}\right)^{1/n+1}\left[\left(\frac{r}{L}\right)^{-1/n+1}\tilde{\sigma}_{ij}^{(0)}(\theta) + A_2\left(\frac{r}{L}\right)^p\tilde{\sigma}_{ij}^{(1)}(\theta) + A_2^2\left(\frac{r}{L}\right)^q\tilde{\sigma}_{ij}^{(2)}(\theta) + \dots\right]. \tag{25}$$

Tables giving the values of the stress functions $\tilde{\sigma}_{ij}^{(k)}(\theta)$ and the higher-order exponents p and q as a function of n for mode I plane-strain conditions have been summarized by Chao and Zhang [18]. One can repeat the Yang et al. [18] derivation for the present asymptotic expansion to show that the second and third term in (23) are related in a completely similar fashion, and that $Q_2 = Q_1^2$. In fact, the angular functions $\hat{\sigma}_{ij}^{(k)}(\theta)$, etc. $(k=1,2)$ appearing in (11)–(20) are identical to the functions $\tilde{\sigma}_{ij}^{(k)}(\theta)$, etc., given by Yang et al. [13] and Chao and Zhang [18]. To emphasize the similarity with (25), we will from now on use the substitutions

$$Q_1 := A_2^*, \quad Q_2 := A_2^{*2} \tag{26}$$

and write the nondimensional stress field in a steady creeping solid as

$$\hat{\sigma}_{ij}(r,\theta) = \left(\frac{1}{I_n}\right)^{1/n+1}[\hat{r}^{-1/n+1}\hat{\sigma}_{ij}^{(0)}(\theta) + A_2^*\hat{r}^p\hat{\sigma}_{ij}^{(1)}(\theta) + A_2^{*2}\hat{r}^q\hat{\sigma}_{ij}^{(2)}(\theta) + \dots]. \tag{27}$$

The only unknown left is the second-order coefficient A_2^*. Its properties for creeping test specimens will be investigated further in the next section.

4 Scaling Properties for Test Specimens

The higher-order terms in (27) control the deviation of the crack tip fields from the HRR solution. The intensity of this deviation, in terms of the value of A_2^*, in a fracture test specimen depends on the geometry of the cracked specimen and on the loading. Before addressing some typical test specimens in the next section, we show that the scaling properties of power-law creeping solids render A_2^* independent of the loading level.

Consider a cracked specimen submitted to steady-state, i.e., extensive power-law creep conditions. Following Ilyushin [19], the response of the real cracked specimen can be obtained from that of a normalized cracked specimen in which we normalize all lengths by a characteristic specimen dimension, W; all stresses by the remote applied stress σ_∞, and all strain rates by $B\sigma_\infty^n$. The latter two are identical to the scaling used for the near-tip fields in (9), but the normalization of lengths is different; for that reason, quantities normalized by W are denoted by $\bar{(\,)}$ rather than $\hat{(\,)}$. The normalized governing equations (equilibrium, compatibility and constitutive relations) thus reduce to

$$\frac{\partial\hat{\sigma}_{ij}}{\partial\bar{x}_j} = 0, \quad \hat{\epsilon}_{ij} = \frac{1}{2}\left(\frac{\partial\bar{u}_i}{\partial\bar{x}_j} + \frac{\partial\bar{u}_j}{\partial\bar{x}_i}\right), \quad \hat{\epsilon}_{ij} = \frac{3}{2}\hat{\sigma}_e^{n-1}\hat{s}_{ij}. \tag{28}$$

This means that C^* is normalized by $B\sigma_\infty^{n+1}W$:

$$\bar{C}^* = \frac{C^*}{B\sigma_\infty^{n+1} W} = \frac{L}{W} = \bar{L} \tag{29}$$

where we recall that L is the characteristic length of the near-tip creep fields (see (8)).

Accordingly, the stress field in a specimen whose geometry and loading are symmetric with respect to the crack plane is expressed as (cf. [17])

$$\sigma_{ij}(x_k, a, W, P, n, \mathcal{S}, \mathcal{F}) = \sigma_\infty \hat{\sigma}_{ij}(\bar{x}_k, \bar{a}, n, \mathcal{S}, \mathcal{F}) \tag{30}$$

where σ_∞ is, for example, taken as the net section stress $P/(W-a)$, and a is the crack length. Variable $\mathcal{S}$ symbolizes the shape of the specimen (but not the size W), and $\mathcal{F}$ represents the loading configuration (but not the magnitude P). In view of the governing Eq. (28), the dimensionless function $\hat{\sigma}_{ij}$ depends on the specimen shape and loading configuration, but does not depend on the specimen size nor on the loading level. Also, it depends on the material only through n but not through B.

The normalized asymptotic near-tip stress field in (27) should possess the same properties. To check this, we recall that, also by scaling arguments, Riedel [17] showed that the C^*-integral for a specimen in which the stress field obeys (30) can be expressed as

$$C^* = WB\sigma_\infty^{n+1} g(\bar{a}, n, \mathcal{S}, \mathcal{F}) \quad \text{or} \quad \bar{C}^* = g(\bar{a}, n, \mathcal{S}, \mathcal{F}),$$

where the dimensionless function g for a given specimen and loading configuration only depends on $\bar{a}$ and n. It then follows from (8) that $L = Wg(\bar{a}, n, \mathcal{S}, \mathcal{F})$, so that $\hat{r} = \bar{r}/g(\bar{a}, n, \mathcal{S}, \mathcal{F})$. When this is substituted into (27), we immediately see that the asymptotic field also has the functional form $\hat{\sigma}_{ij}(\bar{x}_k, \bar{a}, n, \mathcal{S}, \mathcal{F})$ as in (30), provided that the coefficient A_2^* depends only on $\bar{a}$, n, $\mathcal{S}$, and $\mathcal{F}$. So, the higher-order coefficient for the creep fields is indeed independent of the loading level.

In general, the equivalent parameter A_2 for the elastoplastic solid in (25) does depend on the loading level too, because the near-tip fields are dependent on the size of the plastic zone. Only in the limit of a fully plastic specimen is A_2 independent of the loading level ([20]), just like the creep coefficient A_2^*. In fact, the fields in a fully plastic power-law hardening material would be similar to those in a power-law creeping solid by virtue of the Hoff analogy, if the length scales L used in (25) versus (23) were the same. Under these conditions, the second-order parameters A_2 and A_2^* would be identical. Now, Chao et al. [20] and Chao and Zhu [21] use the specimen width W as the arbitrary length parameter L in (25) to identify the value of their second-order parameter, $A_2^{(W)}$. The value A_2 of the parameter corresponding to the length scale L is related to $A_2^{(W)}$ according to

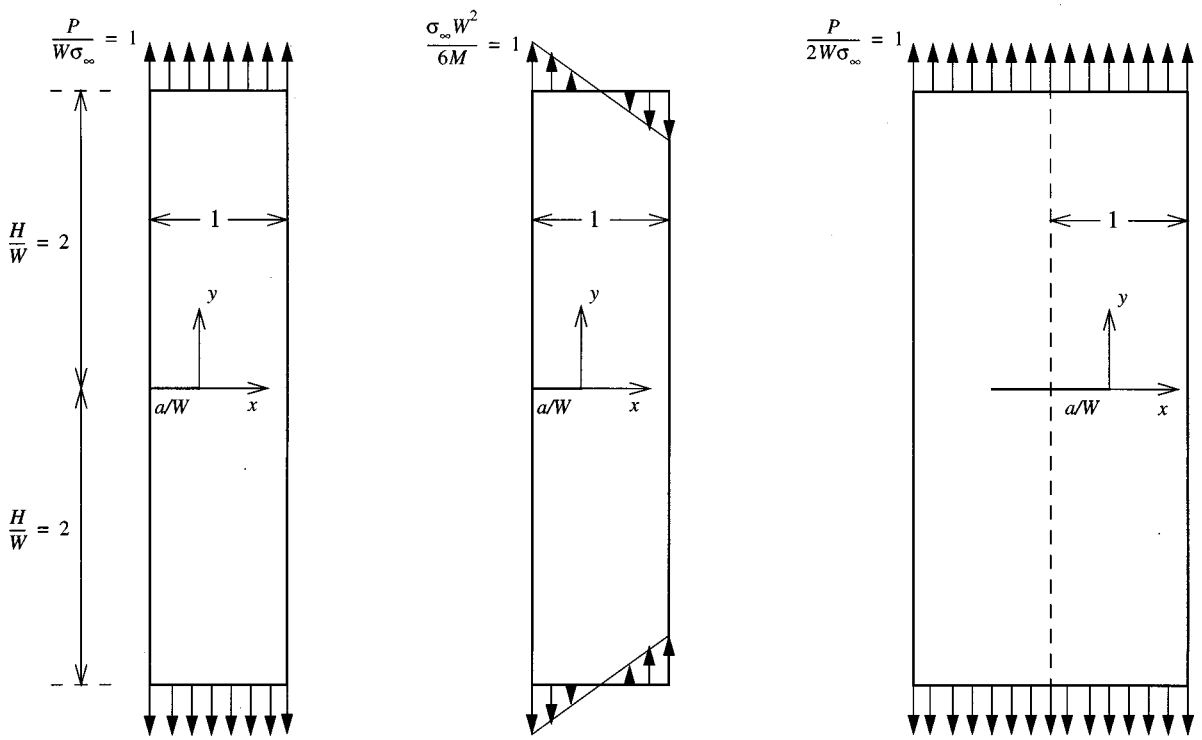
$$A_2 = A_2^{(W)} (L/W)^{p+1/n+1}. \tag{31}$$

This follows immediately from (25) by equating the second-order terms using (L, A_2) with those using $(W, A_2^{(W)})$ (with reference to (24) it is readily checked that the same conclusion is reached by equating the respective third-order terms). The correspondence (31) can be used now to convert the values of the second-order parameter $A_2^{(W)}$ reported by Chao and Zhu [21] for selected test specimens, values of a/W and n, to the appropriate value of A_2^* in the creep solution (27) for any C^* normalized according to (29) through

$$A_2^* = A_2^{(W)} (\bar{C}^*)^{p+1/n+1}. \tag{32}$$

5 Application to Test Specimens

5.1 Loading Functions. In this section, higher-order crack tip fields presented in Section 3 are applied to three test specimen configurations: SENT (single edge notched tension specimen), SENB (single edge notched bending specimen), and CCT (centered crack panel under tension) (see Fig. 1). As indicated in Fig.



(a) SENT (b) SENB (c) CCT

Fig. 1 Geometries of normalized test specimens: (*a*) SENT, (*b*) SENB, and (*c*) CCT, including normalized loading

1, the calculations are performed in a fully normalized way using the scaling presented in Section 4. For practical reasons, the dimensions of the SENT, SENB, and CCT specimens considered in this section are taken to be identical. A span of $2H=4W$ is used so that end effects on the near-tip behavior can be neglected.

In order to obtain the HRR and higher-order fields related to SENT, SENB, and CCT specimens, we need to determine first the value of $\overline{C}^*$ for these specimens. One way to compute $\overline{C}^*$ rapidly and efficiently is to refer to fully plastic solutions providing the values of the *J*-integral determined by Kumar et al. [22] as suggested by Riedel [17]. For the SENT specimen, the fully-plastic solution for power-law hardening materials is given by [22]

$$J=\alpha\sigma_0\epsilon_0 W(1-\overline{a})\overline{a}h_1\left(\frac{P}{P_0}\right)^{n+1} \tag{33}$$

where P_0 is the limit load,

$$P_0=1.4555\eta W(1-\overline{a})\sigma_0 \quad \text{with} \quad \eta=\sqrt{1+\left(\frac{\overline{a}}{1-\overline{a}}\right)^2}-\frac{\overline{a}}{1-\overline{a}},$$

and h_1 is a dimensionless function of the loading configuration, $\overline{a}$ and *n*. Equation (33) can be expressed in terms of the remote applied stress $\sigma_\infty \equiv P/W$ as

$$J=\frac{\alpha\epsilon_0}{\sigma_0^n}W(1-\overline{a})\overline{a}h_1\left(\frac{\sigma_\infty}{1.455\eta(1-\overline{a})}\right)^{n+1}. \tag{34}$$

From (34), we directly obtain the corresponding creep solution for C^* by application of Hoff's analogy, and subsequent normalization according to (29) yields

$$\overline{C}^*=(1-\overline{a})\overline{a}h_1\left(\frac{1}{1.455\eta(1-\overline{a})}\right)^{n+1} \quad \text{for SENT specimens.} \tag{35}$$

The value of h_1 in (33) can be obtained by interpolation from the tables given by Kumar et al. [22] or by an approximate expression proposed by Riedel [17]. However, refined numerical investigations by Shih and Needleman [23,24] have provided more accurate solutions of this problem, which have shown (Shih and Needleman, [23]) that, in the range $1\leqslant n\leqslant 10$, the calculated values of h_1 may differ by as much as 100 percent from those given by Kumar et al. [22]. Therefore, the coefficients h_1 for the SENT specimens used in this paper will be based on those given in [23].

SENB specimens were also investigated by Shih and Needleman [23] and they write the fully plastic solutions as

$$J=\alpha\sigma_0\epsilon_0 W(1-\overline{a})h_1\left(\frac{M}{M_0}\right)^{n+1} \tag{36}$$

where

$$M_0=0.364\sigma_0 W^2(1-\overline{a})^2 \quad \text{and} \quad M=-\int_0^W \sigma_{yy}(x,y=H)x\,dx. \tag{37}$$

Expressing the applied moment *M* in terms of the remote stress $\sigma_\infty=6M/W^2$ as indicated in Fig. 1 and applying the Hoff analogy to expression (36) with a normalization according to (29), we obtain the creep solution

$$\overline{C}^*=(1-\overline{a})h_1\left(\frac{1}{2.184(1-\overline{a})^2}\right)^{n+1} \quad \text{for SENB specimens.} \tag{38}$$

In a similar way, the fully plastic solution of Kumar et al. [22] for CCT specimens,

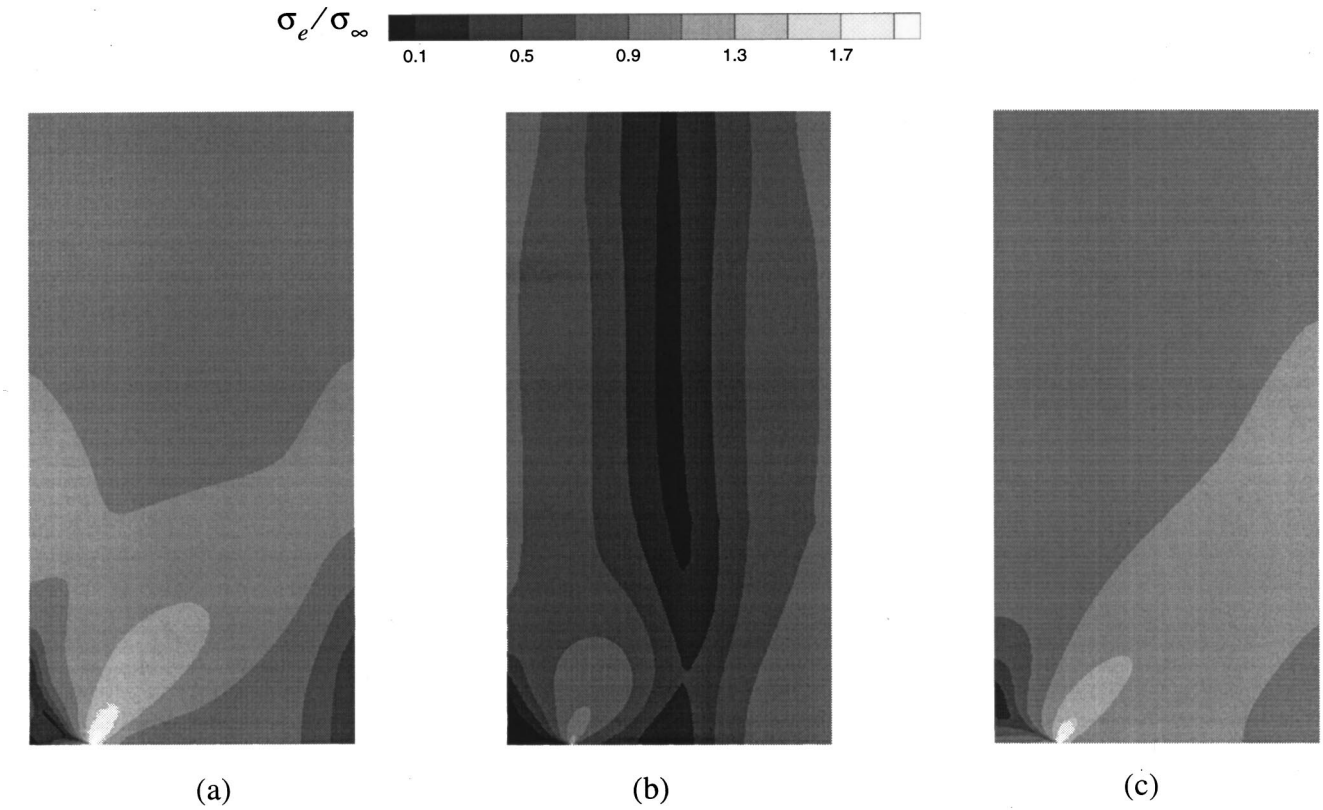


Fig. 2 Contours of Mises stress according to finite element analysis in (*a*) SENT, (*b*) SENB, and (*c*) CCT specimens (*a*/*W*=0.2)

$$J=\frac{\alpha\epsilon_0}{\sigma_0^n}a(1-\overline{a})h_1\left(\frac{\sqrt{3}\sigma_\infty}{2(1-\overline{a})}\right)^{n+1}, \tag{39}$$

allows to directly obtain

$$\overline{C}^*=\overline{a}(1-\overline{a})h_1\left(\frac{\sqrt{3}}{2(1-\overline{a})}\right)^{n+1} \quad \text{for CCT specimens.} \tag{40}$$

The values of h_1 from the accurate fully plastic solutions of Shih and Needleman [23] are available only for SENT and SENB specimens. Therefore, the values for the CCT specimens will be taken from [22].

5.2 Results and Discussion. After having computed $\overline{C}^*$ for SENT, SENB, and CCT specimens, we need to identify the parameter A_2^* for each configuration in order to completely determine the near-tip fields. For this purpose, we have carried out small-strain finite element analyses of steady-state creep in the normalized SENT, SENB, and CCT specimens. For these calculations, we take $n=5$ and consider plane-strain specimens with $\overline{a}=a/W=0.2$, 0.5, and 0.75. By virtue of symmetry, only a half of the specimen geometry needs to be discretized. The computations are carried out incrementally with the elastic-creeping constitutive law (5) with $\nu=0.3$ and $E/\sigma_\infty=10^3$, and are stopped when steady conditions have been attained.

The differences in steady-creep stress fields among the three specimens with $a/W=0.2$ are shown in Fig. 2. These results show that only quite close to the crack tip, the stress distributions are similar while they become more and more different with increas-

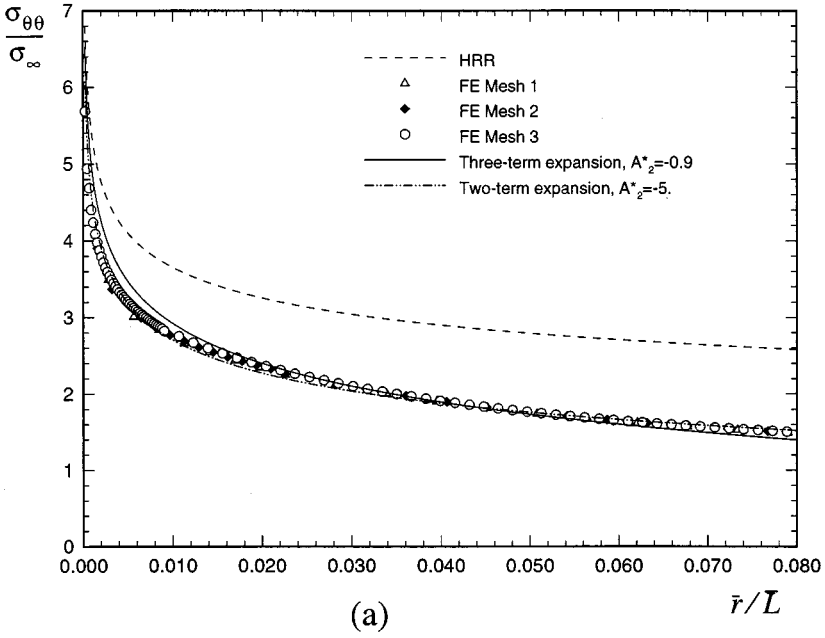


Fig. 3 Normalized opening stress $\hat{\sigma}_{\theta\theta}(\overline{r}/\overline{L},\theta=0)$ in (*a*) SENT and (*b*) SENB specimens ($a/W=0.2$)

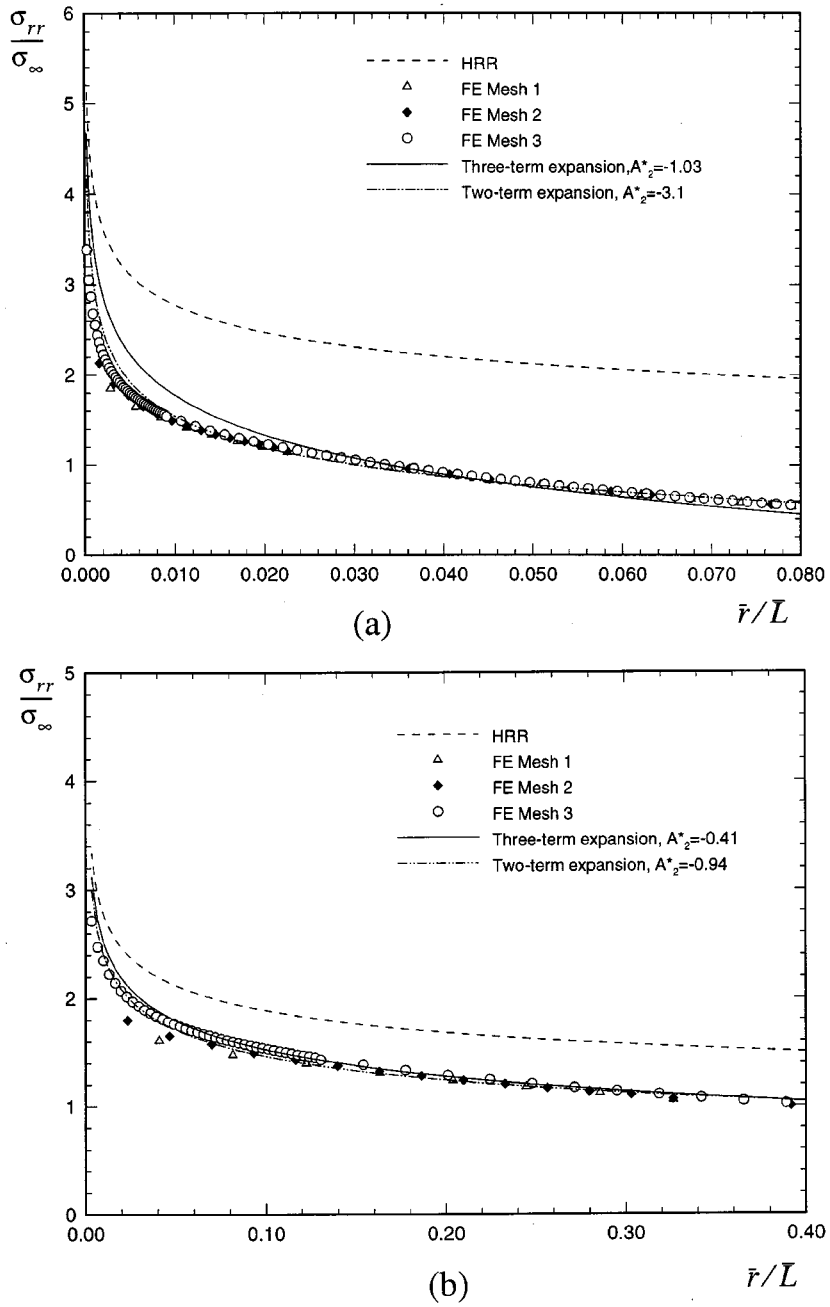


Fig. 4 Normalized radial stress $\hat{\sigma}_{rr}(\bar{r}/\bar{L}, \theta=0)$ in (*a*) SENT and (*b*) SENB specimens ($a/W=0.2$)

ing distance $\bar{r}$. This serves to emphasize that the constraint effect in creep structures is important and cannot be neglected.

A convergence study has been carried out for the SENT and SENB specimens having $a/W=0.2$ by using three meshes with different degrees of refinement around the crack tip. Figures 3 and 4 show the variations of the normalized opening stresses $\hat{\sigma}_{\theta\theta}(\bar{r}/\bar{L}, \theta=0)$ and the radial stresses $\hat{\sigma}_{rr}(\bar{r}/\bar{L}, \theta=0)$ ahead of the crack. These figures show that good convergence has been obtained. Also they show that the deviations with respect to the corresponding HRR solutions are significant and even quite large in the SENT specimen. In both cases, the HRR field is only dominant in a very small region near the crack tip.

Furthermore, Figs. 3 and 4 illustrate the fitting results to obtain the corresponding parameter A_2^*. Two-term and three-term expansions are obtained by limiting the asymptotic field (27) to the second and third terms, respectively. For $n=5$, Yang et al. [13] give $p=0.05456$ for the second term exponent while $q=0.2758$ according to (24). One way to identify A_2^* is by fitting the asymptotic fields of $\hat{\sigma}_{\theta\theta}$ or $\hat{\sigma}_{rr}$ to small-strain finite element solutions for a complete specimen. Yang et al. [13] and Chao et al. [20] have found that the second-order parameter identified in the range $0 \leqslant \theta \leqslant 45$ deg varies only slightly with the location, and generally, the identification of their parameter A_2 in this region has been found to deliver appropriate higher-order fields for a wide range of θ. Motivated by this experience, we here determine A_2^* by fitting the two-term or three-term expansions of $\hat{\sigma}_{\theta\theta}$ and $\hat{\sigma}_{rr}$ along $\theta=0$ deg to the finite element results over a certain range $0<\hat{r}<\hat{r}_c$, with $\hat{r}_c$ depending on geometry and crack depth $\bar{a}$. Chao et al. [20] have shown that in the ''near-tip'' range of

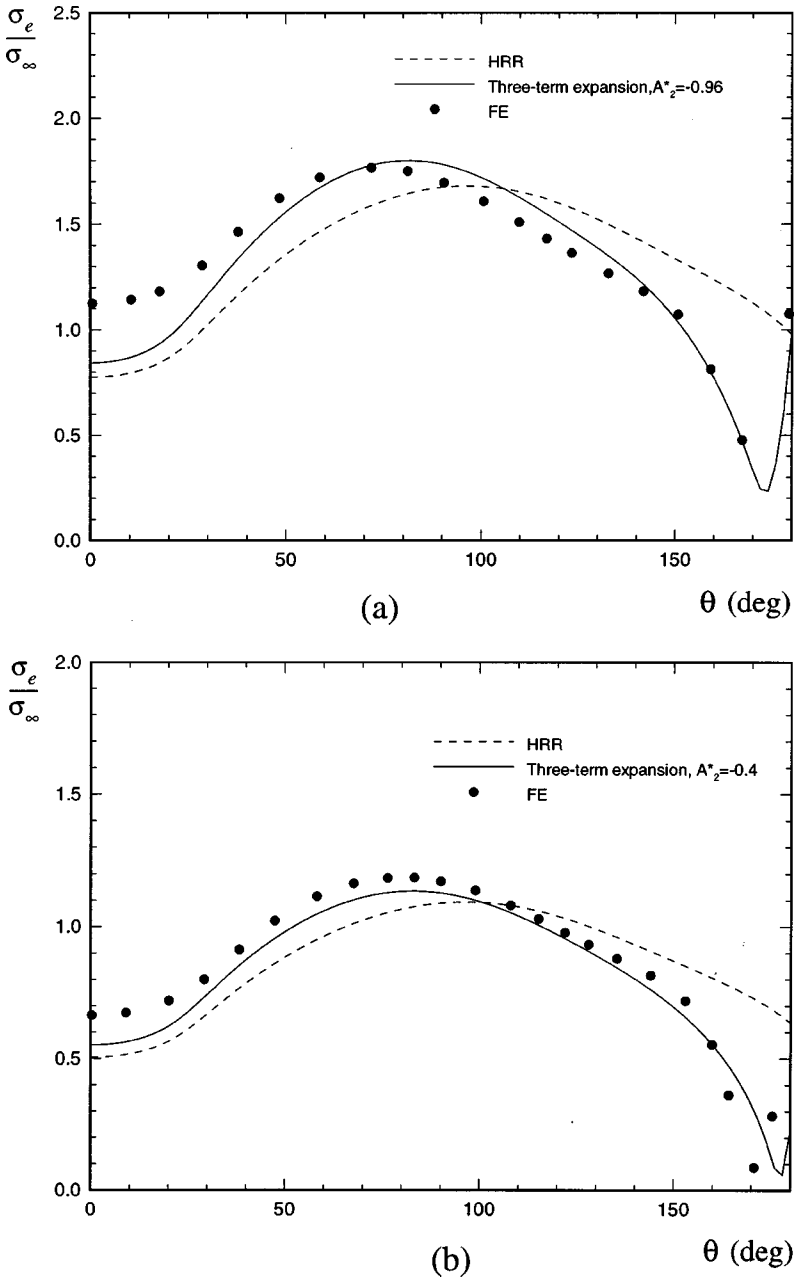


Fig. 5 Angular distributions of the effective stresses at $\bar{r}=0.0196$ in (*a*) SENT and (*b*) SENB specimens ($a/W=0.2$)

$0<\bar{r}<0.1$, their parameter identified under fully plastic conditions can describe the near-tip stress distribution. This range covers the interval $0<\bar{r}/L=\hat{r}<0.4$ for all specimens considered in this paper.

Figures 3 and 4 show that $\hat{\sigma}_{\theta\theta}$ is less sensitive to the mesh refinement than $\hat{\sigma}_{rr}$. Hence, by using a relatively fine mesh, A_2^* can be identified from $\hat{\sigma}_{\theta\theta}$ with a good degree of accuracy. By so doing for the SENT specimen with $a/W=0.2$, we obtained $A_2^*=-5$ for the two-term expansion and -0.9 for the three-term expansion (Fig. 3). On the other hand, fitting on the basis of $\hat{\sigma}_{rr}$ yields $A_2^*=-3.1$ and -1.03 for the two and three-term expansions, respectively. For the SENB specimen with $a/W=0.2$, the A_2^* values based on $\hat{\sigma}_{\theta\theta}$ and $\hat{\sigma}_{rr}$ agree very well ($A_2^*\approx-0.4$) by using the three-term expansion whereas its values for the two-term expansion are -1.2 from $\hat{\sigma}_{\theta\theta}$ and -0.94 from $\hat{\sigma}_{rr}$. Since the difference in $\hat{\sigma}_{\theta\theta}$- and $\hat{\sigma}_{rr}$-based A_2^* values is smaller for the three-term than for the two-term expansion, we can conclude that a three-term expansion indeed provides a more accurate description of the actual near-tip fields than a two-term expansion. Figures 3 and 4 indicate that the fitting can be performed either on the radial variations of $\hat{\sigma}_{\theta\theta}$ or on both $\hat{\sigma}_{\theta\theta}$ and $\hat{\sigma}_{rr}$ to obtain the corresponding A_2^*. For moderately low constraint geometries, such as the SENB specimen, the deviation of the near-tip fields from the HRR solution is not very large, so that A_2^* can be reliably obtained from $\hat{\sigma}_{\theta\theta}$. However, in the cases where the deviation with respect to the HRR solution is quite large, a compromise between the values of A_2^* from $\hat{\sigma}_{\theta\theta}$ and $\hat{\sigma}_{rr}$ may be necessary. In the remainder of this paper, the identification of A_2^* for all considered specimens will be based on $\hat{\sigma}_{\theta\theta}$.

In order to have a further appreciation of the quality of the three-term expansion, Fig. 5 shows the angular distributions of effective stress at $\bar{r}=0.0196$ according to the HRR solution, a

Table 1 Values of A_2^* for a two-term expansion of the asymptotic fields

a/W	0.2	0.5	0.75
SENB	−1.20	−2.80	−18.50
SENT	−3.10	−10.00	−69.00
CCT	−4.75	−8.00	−18.50

Table 2 Values of A_2^* for a three-term expansion of the asymptotic fields

a/W	0.2	0.5	0.75
SENB	−0.4	−1.00	−6.25
SENT	−0.90	−2.50	−20.00
CCT	−0.79	−1.45	−3.00

three-term expansion and the finite element results for the two specimens referred to above. Especially in the interval $0 \leqslant \theta \leqslant 30$ deg approximately (but also for 150 deg$\leqslant\theta\leqslant$180 deg), the deviations from the HRR solution are significant. Such differences are completely ignored in $\{J,Q\}$ approaches ([4]), which collapse all higher-order terms in an additional hydrostatic stress field; therefore, these approaches are expected to have very limited applicability in steady creep and fully plastic conditions. The three-term expansion can indeed pick up the main features of the deviation of the finite element results from the HRR solution quite well, but small quantitative differences remain. The quality of the approximation by a three-term expansion is better in the SENB specimen (Fig. 5(*b*)) than in the SENT specimen (Fig. 5(*a*)). This observation is consistent with the discussion above and is essentially due to the fact that when the deviation from the HRR solution is relatively large, the truncation of the series expansion at the third-order term has some limitation to fully represent the crack-tip stress fields.

The identification results of A_2^* using two-term and three-term expansions for different types of specimen with various ratios a/W are summarized in Tables 1 and 2, respectively. The value of A_2^* is negative for all the studied specimens, and varies substantially with the order of truncation. The general tendency is that the absolute value of A_2^* increases drastically with a/W, whereas the values of $A_2^{(W)}$ for fully plastic flow ([20,21]) do not. This can be understood by examining relation (32) and by plotting $(\bar{C}^*)^{p+1/n+1}$ according to (35), (38), and (40) as a function of a/W (see Fig. 6). For all specimens considered here, $(\bar{C}^*)^{p+1/n+1}$ increases rapidly with a/W for deep cracks ($a/W \geqslant 0.5$) and this explains the observed tendency. Moreover, relation (32) along with the values of A_2^* for the three-term expansion given in Table 2 (for $n=5$), can be used to compute the corresponding values of $A_2^{(W)}$. These are shown in Fig. 7(*a*) for SENB, SENT, and CCT specimens, and indeed vary only mildly with a/W.

To estimate the reduction in stress level at a certain distance from the crack tip (i.e., loss of constraint), we focus on (23) and (26) and substitute $L=W$. It is noted that both the second and third-order terms contribute to a reduction in σ_{rr} and $\sigma_{\theta\theta}$ for $A_2^{(W)}<0$ since $\hat{\sigma}_{rr}^{(1)}>0$ and $\hat{\sigma}_{\theta\theta}^{(1)}>0$ while $\hat{\sigma}_{rr}^{(2)}<0$ and $\hat{\sigma}_{\theta\theta}^{(2)}<0$. Therefore, for a given C^* and σ_∞, the loss of constraint is directly related to $A_2^{(W)}$. Figure 7(*a*) shows that the CCT specimens have the lowest constraint.

For comparison, Fig. 7(*b*) includes Chao and Zhu's [21] fully plastic results for SENB and CCT specimens. The observation that our creep results for $n=5$ do not order exactly with their values for $n=3$ and $n=10$ may be partly due to the fact that different finite element meshes were used, but are believed to stem mainly from differences in the way that the finite element stress fields are fitted to the asymptotic fields. With this in mind, the general tendency is good and shows that our results based on steady creep calculations and relation (32) allow to recover Chao and Zhu's [21] results obtained under large-scale yielding up to fully plastic conditions.

It is pertinent to note that the finite element computations have ignored finite strain effects leading to blunting. This is not only for a fair comparison with the asymptotic solution, which is also for infinitesimal strains, but is in fact essential since a finite strain steady creep solution does not exist. Under a constant remote loading, creep deformations will continue to accumulate, leading to blunting and continuously evolving fields.

We are not aware of any published experimental work that evidences the effects of crack-tip constraint on creep crack growth. However, Parks [25] and his co-workers observed radically different patterns of creep damage in SENB and CCT specimens. As we know that CCT specimens have a much lower constraint than SENB specimens, it is likely that the constraint effect

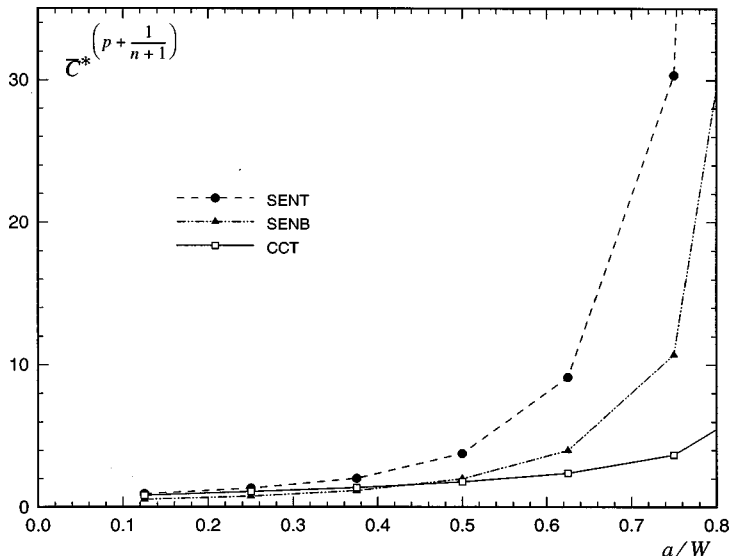


Fig. 6 Evolution of $(\bar{C}^*)^{p+1/n+1}$ with a/W in SENT, SENB, and CCT specimens for $n=5$ ($p=0.05456$)

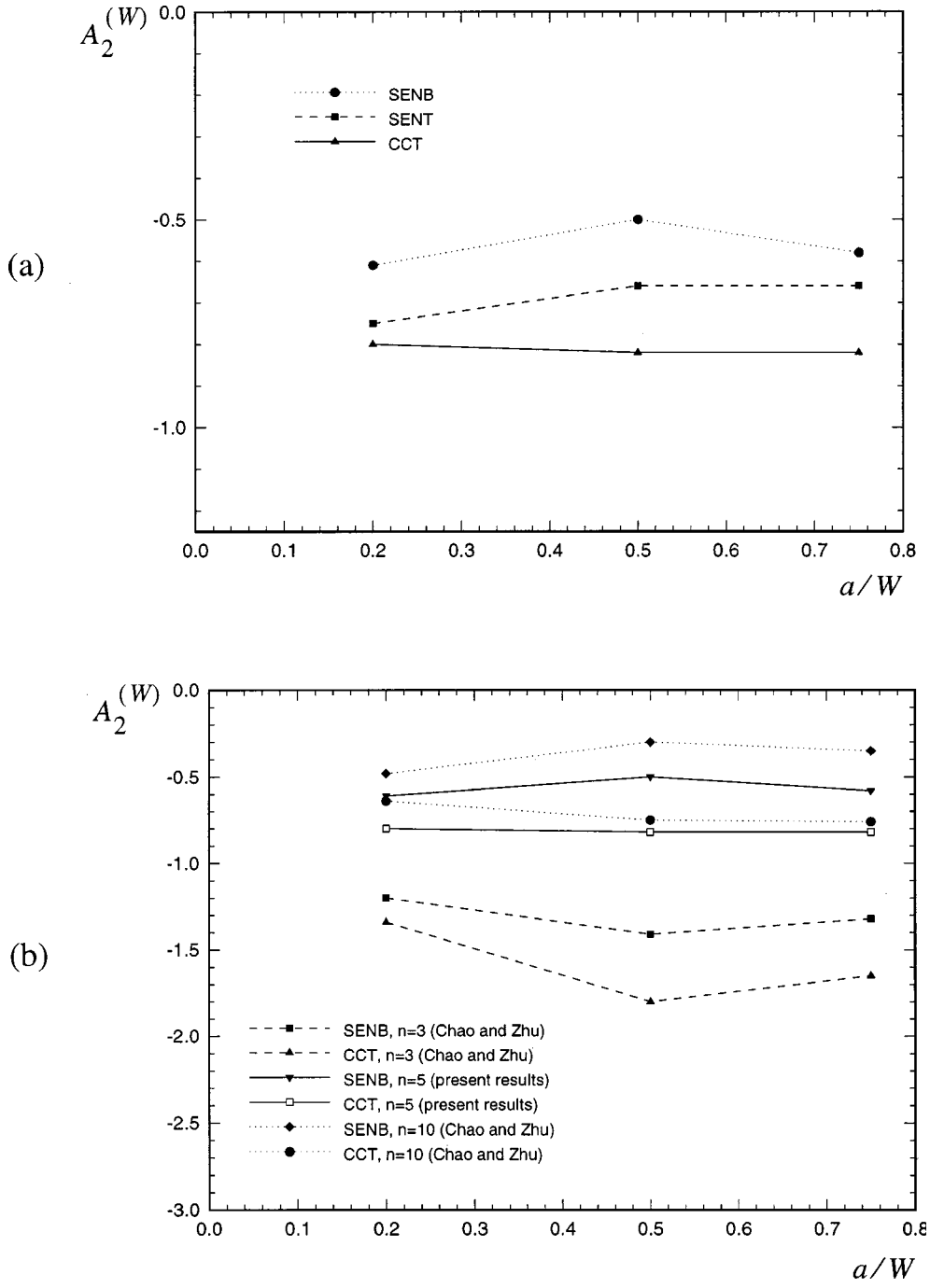


Fig. 7 Values of $A_2^{(W)}$ for (*a*) SENB, SENT, and CCT specimens as obtained from our creep calculations and converted using (32), (*b*) SENB and CCT specimens compared with the fully plastic results of [21]

has a significant influence on the damage pattern as well as the crack growth rate and direction in the near-tip region. Crack-tip constraint effects on creep fracture will be studied in a forthcoming paper.

6 Conclusion

This paper proposes higher-order fields near a sharp crack in a steadily creeping solid, limited to the third term, which are derived from an asymptotic development analogous to that of Sharma and Aravas [16] and Yang et al. [13] for elastoplastic bodies. The analysis has shown that the higher-order fields, which account for the constraint effect imposed by the specific geometry and loading configuration, are controlled by two additional parameters: A_2^* and σ_∞. The latter is a reference stress, which can be for instance defined as the effective Mises stress in the absence of the crack. Thus, the remote loading level is characterized by C^* and σ_∞. Together, they also define a creep characteristic length, $L=C^*/(B\sigma_\infty^{n+1})$, that serves as a natural scaling parameter for the radial variation of the higher-order fields. By means of the scaling properties applied to power-law materials, we have proved that A_2^* is only dependent on n, the geometry shape and loading configuration, but is independent of the loading level.

Small-strain finite element computations of three popular test

specimen geometries have been carried out to identify the value of A_2^*. The numerical results have confirmed that the near-tip stress distribution varies significantly from one type of specimen to another. Therefore, a creep fracture characterization based on $\{C^*, A_2^*, \sigma_\infty\}$ is expected to be more accurate than the current practice based on C^* only, because the latter neglects the constraint effect.

Finally, by virtue of the Hoff analogy, we have derived a relation between A_2^* and the second-order parameter $A_2^{(W)}$ in the fully plastic case. This relation has been confirmed numerically and allows to directly exchange the values of these parameters between creeping and fully plastic situations.

References

[1] Williams, M. L., 1957, ''On the Stress Distribution at the Base of a Stationary Crack,'' ASME J. Appl. Mech., **24**, pp. 111–114.

[2] Li, Y., and Wang, Z., 1986, ''Higher Order Asymptotic Field of Tensile Plane Strain Nonlinear Crack Problem,'' Sci. Sin. (Series A), **29**, pp. 941–955.

[3] Betegón, C., and Hancock, J. W., 1991, ''Two-Parameter Characterization of Elastic-Plastic Crack-Tip Fields,'' ASME J. Appl. Mech., **58**, pp. 104–110.

[4] D'Dowd, N. P., and Shih, C. F., 1991, ''Family of Crack-Tip Fields Characterized by a Triaxiality Parameter—I. Structures of Fields,'' J. Mech. Phys. Solids, **39**, pp. 989–1015.

[5] Parks, D. M., 1992, ''Advances in Characterization of Elastic-Plastic Crack-Tip Fields,'' *Topics in Fracture and Fatigue*, A. S. Argon, ed., Springer-Verlag, New York, pp. 59–98.

[6] D'Dowd, N. P., 1995, ''Application of Two-Parameter Approaches in Elastic-Plastic Fracture Mechanics,'' Eng. Fract. Mech., **52**, No. 3, pp. 445–465.

[7] Sharma, S. M., and Aravas, N., 1991, ''Determination of Higher-Order Terms in Asymptotic Elastoplastic Crack-Tip Solutions,'' J. Mech. Phys. Solids, **39**, pp. 1043–1072.

[8] Hutchinson, J. W., 1968, ''Plastic Stress and Strain Fields at a Crack Tip,'' J. Mech. Phys. Solids, **16**, pp. 337–347.

[9] Rice, J. R., and Rosengren, G. F., 1968, ''Plane Strain Deformation near a Crack Tip in a Power Law Hardening Materials,'' J. Mech. Phys. Solids, **16**, pp. 1–12.

[10] Shih, C. F., 1983, ''Tables of HRR Singular Field Quantities,'' Report MRL E-147, Materials Research Laboratory, Brown University, Providence, RI.

[11] Sharma, S. M., 1997, private communication.

[12] Xia, L., Wang, T. C., and Shih, C. F., 1993, ''Higher-Order Analysis of Crack Tip Fields in Elastic Power-Law Hardening Materials,'' J. Mech. Phys. Solids, **41**, No. 4, pp. 665–687.

[13] Yang, S., Chao, Y. J., and Sutton, M. A., 1993, ''Higher Order Asymptotic Crack Tip Fields in a Power-Law Hardening Material,'' Eng. Fract. Mech., **45**, pp. 1–20.

[14] Nikishkov, G. P., Brückner-Foit, A., and Munz, D., 1995, ''Calculation of the Second Fracture Parameter for Finite Cracked Bodies Using a Three-Term Elastic-Plastic Asymptotic Expansion,'' Eng. Fract. Mech., **52**, No. 4, pp. 685–701.

[15] Riedel, H., and Rice, J. R., 1980, ''Tensile Cracks in Creeping Solids,'' ASTM-STP-700, ASTM, Philadelphia, pp. 112–130.

[16] Sharma, S. M., and Aravas, N., 1995, ''Two-Parameter Characterization of Crack Tip Fields in Edge-Cracked Geometries: Plasticity and Creep Solutions,'' ASTM-STP-1220, ASTM Philadelphia, pp. 309–327.

[17] Riedel, H., 1987, *Fracture at High Temperatures* (Materials Research and Engineering Series), Springer-Verlag, New York.

[18] Chao, Y. J., and Zhang, L., 1997, ''Tables of Plane Strain Crack Tip Fields: HRR and Higher Order Terms,'' ME-Report 97-1, University of South Carolina.

[19] Ilyushin, A. A., 1946, Prikl. Mat. Mekh., **10**, p. 347.

[20] Chao, Y. J., Yang, S., and Sutton, M. A., 1994, ''On The Fracture of Solids Characterized by One or Two Parameters: Theory and Practice,'' J. Mech. Phys. Solids, **42**, pp. 629–647.

[21] Chao, Y. J., and Zhu, X. K., 1998, ''J-A_2 Characterization of Crack-Tip Fields: Extent of J-A_2 Dominance and Size Requirements,'' Int. J. Fract., **89**, pp. 285–307.

[22] Kumar, V., German, M. D., and Shih, C. F., 1981, ''An Engineering Approach for Elastic-Plastic Fracture Analysis,'' Report NP-1931 on Project 1237-1 for Electric Power Research Institute, Palo Alto, CA.

[23] Shih, C. F., and Needleman, A., 1984, ''Fully Plastic Crack problems, Part 1: Solutions by a Penalty Method,'' ASME J. Appl. Mech., **51**, pp. 48–56.

[24] Shih, C. F., and Needleman, A., 1984, ''Fully Plastic Crack problems, Part 2: Application of Consistency Checks,'' ASME J. Appl. Mech., **51**, pp. 57–64.

[25] Parks, D. M., 1999, private communication.

Constitutive Dynamic-Order Model for Nonlinear Contact Phenomena

D. Ingman
e-mail: qadov@tx.technion.ac.il

J. Suzdalnitsky
e-mail: isosef@tx.technion.ac.il

M. Zeifman
zmicael@tx.technion.ac.il

QAR,
Technion-Israel Institute of Technology,
Haifa 32000, Israel

A dynamic integro-differential operator of variable order is suggested for a more adequate description of processes, which involve state dependent measures of elastic and inelastic material features. For any negative constant order this operator coincides with the well-known operator of fractional integration. The suggested operator is especially effective in cases with strong dependence of the behavior of the material on its present state—i.e., with pronounced nonlinearity. Its efficiency is demonstrated for cases of viscoelastic and elastoplastic spherical indentation into such materials (aluminum, vinyl) and into an elastic material (steel) used as a reference. Peculiarities in the behavior of the order function are observed in these applications, demonstrating the ''physicality'' of this function which characterizes the material state. Mathematical generalization of the fractional-order integration-differentiation in the sense of variability of the operator order, as well as definitions and techniques, are discussed. [S0021-8936(00)02102-4]

1 Introduction

The large body of research on material modeling in general and polymer modeling in particular can be classified in two distinct branches. The first seeks the origin of the material behavior at the micro level, i.e., at that of molecular chains, microdefects, etc. The second group deals with phenomenological description at the macro parametric level of explanation. The present paper addresses the second branch only. Note that probabilistic features of the corresponding statistical ensembles are also essential for evaluation of the macro response of the material.

The most general and common approach to material behavior postulates, that local material response to an external stress at a point occupied at the time t by a particle in the deformed body, is uniquely determined by the history of motion and temperature [1]. Krajcinovic [2] assumes that if the continuum is homogeneous and isotropic on the macro scale, the constitutive description of the interconnection between stress, deformation, and temperature can be relaxed to the following form:

$$\sigma(t)=\Psi[F(s),T(s),\dot{T}(s),0<s\leqslant t], \tag{1.1}$$

where Ψ is the tensor-valued response functional, $F(s)$ the history of the deformation gradient up to time t, and T temperature. The constitutive law (1.1) implies that the stress is a functional of all preceding values of the deformation gradients, temperature, and temperature gradient. The current stress value is derived by integrating this functional over the time interval $(0,t)$. Obviously, a similar expression can be written for the case of deformation in response to a given load. The functional representation (1.1) is much too complicated and cumbersome for engineering applications, especially in its general form. It is important to note that this constitutive law is deterministic and assumes that a single stress-versus-time function can correspond to predetermined histories of temperature and deformation fields. It makes no allowance for possible changes in the continuum caused by onset of defects like microcracks and displacement of their aggregations—processes of highly probabilistic nature.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, December 21, 1998; final revision, April 9, 1999. Associate Technical Editor: A. K. Mal.

Lee and Radok [3] suggested a simplified form of the functional Ψ, which may be regarded as the well-known integro-differential operator

$$\sigma(t)=\psi[x(t)]=\int_0^t G(t-\tau)\frac{d}{d\tau}[x(\tau)]d\tau. \tag{1.2}$$

For some specific set of functions, for instance $G(t)=(t-\tau)^{\alpha-1}/\Gamma(\alpha)$, this is the operator of fractional-order integration ([4]).

The generalized time-dependent relationship between one-dimensional stress σ and strain ϵ has the form of the operator equation

$$\sigma(t)=\mathbf{D}\epsilon(t). \tag{1.3}$$

For the two extremes: (a) purely elastic deformation, $\mathbf{D}$ is the operator of multiplication by the modulus of elasticity E; and (b) purely viscous deformation, it is the differential operator $\mathbf{D}=\eta(d/dt)$, where η is the viscosity coefficient. Thus, one can assume (1.2) to be a valid description of the intermediate state ([3]). These operators were recently applied to viscoelastic problems by Bagley and Calico [5], Bagley et al. [6], and Rossikhin and Shitikova [7].

Another, perhaps the most popular in ''macro''-material modeling, presentation of a viscoelastic solid is a system of springs and dashpots (Maxwell elements). The mathematical model of this system is a linear differential equation, whose fundamental solution (Green's function) is accepted as the relaxation function or creep compliance function. Lee and Rogers [8] and Rabotnov [9] noted that a more complex behavior of material response necessitates a larger number of structural elements and thus leads to a higher order of the corresponding differential equations. Bardenhagen et al. [10] and many others tried to improve this model by introducing additional types of elements—nonlinear springs, sliding elements, etc.

Some theories of plasticity also should be mentioned. Both classical, i.e., implementing the yield surface concept ([11]), viscoplasticity (which this paper does not address), and rate-independent plasticity lead to description of the phenomena in terms of a set of differential equations ([12]). A more modern (fairly close to the present paper) philosophy lets us refer to the so-called endochronic theory of Valanis [13], originally formulated for application to the mechanical behavior of metals and

more recently applied to other materials, such as concrete ([14]) and polymers ([15]). The basis of the theory is an internal material time related to the material deformation history.

For a rate-independent one-dimensional case, this time measure ξ can be defined as some parameter proportional to the strain ([11]). The thermodynamics of the phenomenon ([13]) leads to the Lee-Radok type ((1.2) type) of stress dynamics versus deformation history. The only difference is the internal time ξ instead of the real time t in the integral.

2 Problem

From the physical point of view, the model of the Maxwell solid is merely a tool for representing the actual material behavior via solution of differential equations. It is important to emphasize that evaluation of material response within the framework of such a set of springs and dashpots does not imply any changes in their features versus material status. That is, the level of nonlinearity considered is unambiguously defined by both the system structure and the structural elements and cannot be modified throughout the evolution of the system. Most experimental studies, however, show that a proper material model should be modifiable in the sense of material evolution—for example, propagation of the plastic zone under load (see, for instance, Rajic et al., [16]) corresponds to transformation of linear springs at the propagation front into the nonlinear ones, changing the local properties of the material and its macro response.

Another type of internal inconsistency is characteristic for the case of the integro-differential operator of constant order (1.2). Consider, for instance, a material stress-strain curve under a load with a constant strain rate. By (1.2), scaling of this strain rate ($\dot{\epsilon}^* = a\dot{\epsilon}$, $a = \text{const}$) leads to the same scaling of the material response ($\sigma^*(t) = a\sigma(t)$) and magnification of any strain history causes the same magnification of the response ($\dot{\epsilon}^*(t) = K\dot{\epsilon}(t) \rightarrow \sigma^* = K\sigma(t)$). This result conflicts with most experimental studies of the stress-strain behavior. Let us demonstrate it by the next simple theoretical experiment. The direct method of reconstruction of the kernel G (in (1.2)) consists in the determination of the material response to the step load:

$$x(t) = \begin{cases} 0 & \text{for } t < t_0 \\ h & \text{for } t \geqslant t_0 \end{cases}.$$

By (1.2), the material response will be proportional to h, with the same time pattern.

$$\sigma(t) = h \cdot G(t)$$

In reality, small values of h correspond to more elastic material behavior and the stress should follow the strain time pattern while larger ones should reflect some features of inelastic behavior. Therefore, the uniqueness of G in (1.2) is inapplicable for wide ranges of stresses or strains over the loading history.

Finally, the intrinsic drawback of all the theories mentioned above is the asymmetry in the remedy for the inverse problem, that is, the strain-valued material response to the applied stress. Moreover, although some theories (like that of Valanis) use the concept of thermodynamics involving state variables, no physically sound parameter responsible for the material state has been suggested.

3 The Model

Let us consider, for simplicity, the one-dimensional isothermal problem of material response. The definition of this response depends on the parameter serving as input. In the case where a given load is considered, the response is the deformation, while increase of the deformation leads to response in the form of internal stress.

As the ''elasticity'' and ''viscosity'' effects accompany the loading process throughout and the link between them varies, satisfactory description of the phenomenon is only possible under a governing operator of dynamic order. This dynamic order reflects the material evolution within the space bounded by the purely elastic and purely viscous extremes. Thus, the correct operator $\mathbf{D}$ in (1.3) should be compatible with the continuous process of the state $S(t)$, i.e.,

$$\mathbf{D} = \mathbf{D}^{\alpha(S(t))} \tag{3.1}$$

and

$$\sigma(t) = \mathbf{D}^{\alpha(S(t))}\epsilon(t). \tag{3.1a}$$

Here, the fractional differentiation order function $\alpha(S(t))$ is some bounded function of the material state $S(t)$, i.e., $0 \leqslant \alpha(S(t)) \leqslant 1$. The limiting values of α correspond to the states of pure elasticity ($\alpha = 0$) and pure viscosity ($\alpha = 1$). For the case of material response to a load,

$$\epsilon(t) = \mathbf{D}^{\beta(S(t))}\sigma(t). \tag{3.2}$$

$\mathbf{D}^{\beta(S(t))}$ is the operator of creep compliance, with limiting β values of $-1 \leqslant \beta \leqslant 0$. The operator $\mathbf{D}$ in (3.1) depends exclusively on the state function $S(t)$. The present work simplifies the problem by dealing with the dynamic order, namely $\alpha = \alpha(\sigma(t))$ and $\beta = \beta(\epsilon(t))$ in (3.1) and (3.2), rather than with their dependence on the material state. Suitable choice of $\beta(\epsilon)$ permits correct presentation of the nonlinear plastic flow process. Note that both (3.2) and (3.1a) are functional equations with unknowns on both sides of the equation. Thus, we are introducing a new class of operators, namely, those of state evolution, whose outstanding feature is that their order depends on the history of the material response—so that they are extremely nonlinear. Unlike their constant-order counterparts, they exhibit nonlinear behavior even for an input which is a sum of other inputs. That is, if $\sigma_i(t) = D^{\alpha(\sigma_i(t))}\epsilon_i(t)$, $i = 1,2$, then $\sigma_1(t) + \sigma_2(t) \neq \mathbf{D}(\epsilon_1(t) + \epsilon_2(t))$.

The application of this type of operators is demonstrated on two examples of contact problems of viscoelastic and elastoplastic deformation. For simplicity, the one-dimensional problem of isothermal material response is examined in the present work.

The next section deals with the integro-differential operator of variable order (IDOVO), its definition and features, as well as operator reciprocal to IDOVO.

4 Integro-Differential Operator of Variable Order

Consider a linear functional set $L = \{x(t)\}$, whose elements $x(t)$ are real or complex functions of real variable t. The function $x(t)$ is supposed to be definite for $t \geqslant 0$, absolutely integrable, one-valued, and with a finite number of discontinuities in any finite interval $[0,T]$. Let $\alpha = \alpha(t)$ be a continuous function of the parameter t. Then, we define the dynamic differential operator of time-dependent order $\mathbf{D}^{\alpha(t)}$ as follows:

$$\mathbf{D}^{\alpha}x(t) = \frac{d^m}{dt^m}\int_0^t \frac{(t-\xi)^{m-1-\alpha(t-\xi)}}{\Gamma(m-\alpha(t-\xi))}x(\xi)d\xi, \tag{4.1}$$

where m is the integer part of $\alpha + 1$ for $\alpha \geqslant 0$ and $m = 0$ for $\alpha < 0$; $\alpha(t)$ is the order function of the operator $\mathbf{D}$ and Γ is the gamma function

$$\Gamma(x) = \int_0^\infty e^{-\xi}\xi^{x-1}d\xi.$$

The operator $\mathbf{D}^{\alpha}$ is linear on the set $L = \{x(t)\}$. For any negative constant $\alpha(t) = \alpha_0 < 0$, it is a Riemann-Liouville fractional integral of order α_0 or an Euler transform of the first kind ([4]). For integer $\alpha = n$, $\mathbf{D}^n$ is an ordinary operator of integration (for $n < 0$), or of differentiation (for $n > 0$). Note that for a constant α, $\mathbf{D}^{\alpha}$ coincides with (1.2). $\mathbf{D}$ is commutative: $\mathbf{D}^{\alpha}\mathbf{D}^{\beta} = \mathbf{D}^{\beta}\mathbf{D}^{\alpha}$. Limitations on α and β are the same as in ([17]). The operator $\mathbf{D}^{-\beta}$ is the reciprocal of $\mathbf{D}^{\alpha}$, if α and β satisfy the relation

$$\int_0^t \frac{(t-\xi)^{-\beta(t-\xi)}\xi^{-1+\alpha(\xi)}}{\Gamma(1-\beta(t-\xi))\Gamma(\alpha(\xi))}d\xi = U(t) \tag{4.2}$$

where $U(t)$ is the unit function (proofs are given in the Appendices). A more detailed study of the IDOVO is forthcoming.

The problem of reconstruction of $\alpha(t)$ for predetermined $x(t)$ and $y(t)=\mathbf{D}^{\alpha}x(t)$ now arises. Denoting the kernel of $\mathbf{D}^{\alpha}$ as

$$g_m(t)=\frac{t^{m-1-\alpha(t)}}{\Gamma(m-\alpha(t))} \tag{4.3}$$

we have

$$y(t)=\frac{d^m}{dt^m}\int_0^t g_m(t-\xi)x(\xi)d\xi.$$

Methods for solving this Volterra integral equation are well known. Rewriting (4.3) in the form

$$\alpha(t)=-\log(t^{1-m}\Gamma(m-\alpha(t))g_m(t))/\log t \tag{4.4}$$

permits reconstruction of $\alpha(t)$ by an iterative procedure. This technique is presented in Section 6.

5 Viscoelastic Deformation of Material by Spherical Indentor (Principles of Demonstration Model)

Although the IDOVO was originally intended for plastic materials, the examples presented below are based on indentation tests, which usually involve metals. The main reasons for this choice are: the indentation tests are strongly dependent on the viscoelastic and elastoplastic features of the tested materials, and moreover, the tests reported in literature are often supported by available experimental data.

In the articles quoted above, in which fractional differentiation was applied to viscoelastic problems, the order of the differential operator was fractional but constant. In the following, the differential operator of variable order is applied to the problem of penetration of a spherical indentor into a plane. According to the Hertz formulas for the purely elastic case, the radius of the contact circle a and the pressure $p(r)$ within this circle are given by

$$a^3(t)=\frac{3}{8}\frac{1}{2G}P(t)R \tag{5.1}$$

$$p(r)=\frac{4}{\pi R}2G(a^2-r^2)^{1/2} \tag{5.2}$$

where R is the radius of the sphere; $P(t)$ is the load on the indentor at time tT; T is a some unit of time; t dimensionless variable, $2G=E/(1+\nu)$; E,ν are the modulus of elasticity and Poisson's ratio of the compressed material; and r is the distance from the center of the contact circle $(0\leqslant r\leqslant a)$. To obviate the dimensionality problem, we replace $P(t)$ by $P_0\cdot x(t)$, where $x(t)$ is a dimensionless function. Since $1/2G$ represents the elastic properties of the material, it should be replaced in the viscoelastic situation by the IDOVO as follows:

$$\frac{1}{2G}\rightarrow\frac{1}{2G_0}\mathbf{D}^{-\alpha(S(t))}.$$

The function $\alpha(S(t))$ in this illustrative example is presented as $\alpha(t)$. Thus in the viscoelastic situation (5.1) the following change takes place through the variable-order operator of fractional integration:

$$a^3(t)=\left(\frac{3RP_0}{16G_0}\right)\mathbf{D}^{-\alpha(t)}x(t)=\left(\frac{3RP_0}{16G}\right)\int_0^t\frac{\xi^{-1+\alpha(\xi)}}{\Gamma(\alpha(\xi))}x(t-\xi)d\xi. \tag{5.3}$$

According to the earlier assumptions, $\alpha(t)$ should vary from 0 (purely elastic state) to 1 (purely viscous state). Let us demonstrate the elastic-viscous dynamics for the following model of $\alpha(t)$

$$\alpha(t)=1-2^{-t}.$$

Figure 1 represents the dependence of the radius of the contact circle a on t for $x(t)=t$.

The inverse problem can be written as follows: $P(t)=(8/3)(2G/R)a^3(t)$ or $x(t)=\mathbf{D}^{\beta(t)}(a(t)/c)^3$, where $c^3=3P_0R/16G_0$ and $\mathbf{D}^{\beta(t)}$ is the reciprocal operator to $\mathbf{D}^{-\alpha(t)}$. This direct approach, however, does not yield the pressure distribution over the contact zone. Therefore, in order to reconstruct it we have to use an operator analogous to (5.2), which should also be considered as the fractional differentiation: $p(r,t)=(8G_0/\pi R)\mathbf{D}^{\beta^*(t)}[a^2(t)-r^2]^{1/2}$. As indicated above, in the general case $\beta(t)\neq\beta^*(t)$. The notion of model verification consists on deriving $p(r,t)$ subject to $\beta(t)=\beta^*(t)$. Thus, integrating $p(r,t)$ over the entire contact zone must yield

$$P^*(t)=P(t), \tag{5.4}$$

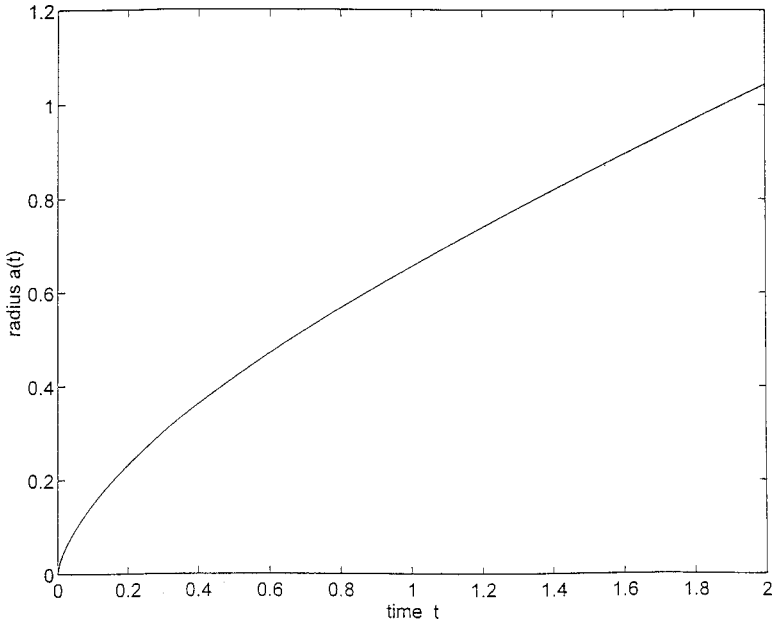


Fig. 1 Radius of contact circle versus time of load application

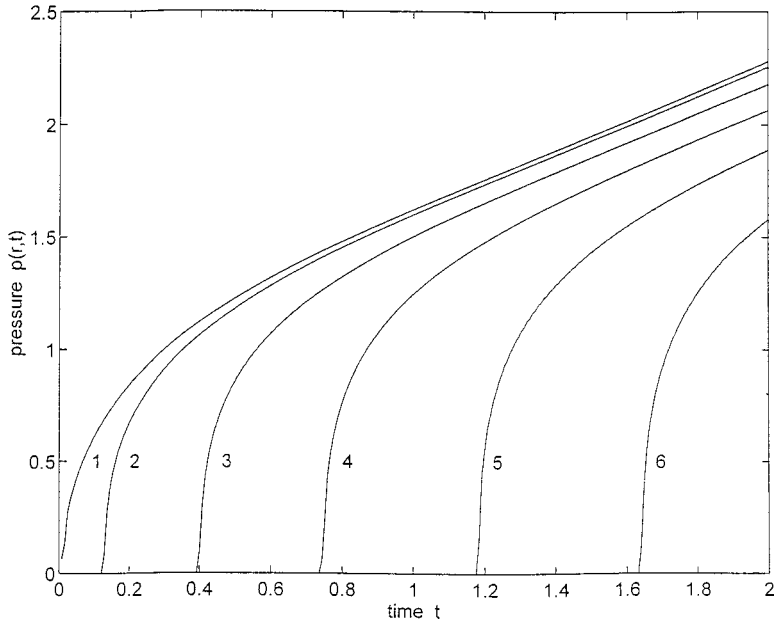


Fig. 2 Pressure versus time of load application

if the model properly describes the phenomenon.

Denoting

$$\nu(t)=\int_0^t \frac{\xi^{-1+\alpha(\xi)}}{\Gamma(\alpha(\xi))}x(t-\xi)d\xi,$$

we have

$$P(r,t)=\frac{8G_0}{\pi R}c\frac{d}{dt}\int_0^t \frac{\xi^{-\beta(\xi)}}{\Gamma(1-\beta(\xi))}[\nu^{2/3}(t-\xi)-(r/c)^2]^{1/2}d\xi. \tag{5.5}$$

The function $g_1(\xi)=\xi^{-\beta(\xi)}/\Gamma[1-\beta(\xi)]$ was determined according to (3.3) (see Appendix B). Note that the condition (5.4) is met in our case.

Figure 2 shows the dependence of the pressure p on time for $r=0,0.2,0.4,0.6,0.8,1.0$ (curves 1–6, in that order), and Fig. 3—that of p on r for $t=0.01,0.5,1.0,1.5,2.0$ (curves 1–5 in that order). The area of the contact zone and the pressure at any point in it increase with the loading time, while the pressure versus distance from the center decreases.

Johnson [18], who considered a purely viscous process of deformation assuming $2G\rightarrow d/dt$, obtained

$$p(r,t)=\frac{P_0}{2\pi a}(a^2-r^2)^{-1/2}.$$

In such a case the pressure tends to infinity at the boundaries of the contact zone. The present approach, as can be seen from Figs. 2–3, suggests that the pressure (5.4) tends to zero near the bound-

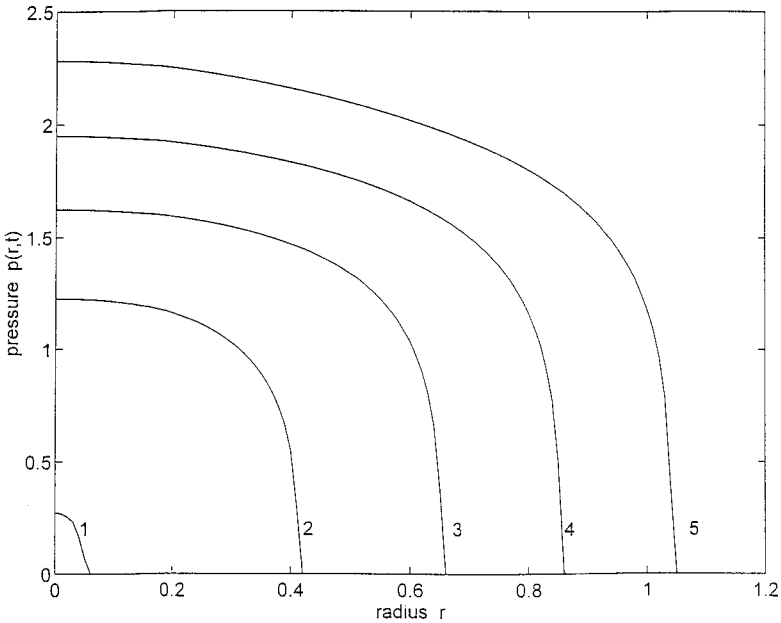


Fig. 3 Dependence of pressure on contact radius

ary, and, therefore, these results seem to be more correct from the physical point of view. Note, that the results also satisfy the criterion (5.4).

6 Elastoplastic Deformation (Experiment-Based Model)

As an actual example illustrating the reconstruction of the order of IDOVO, let us consider time-independent (nonviscous) elastoplastic deformation.

Brinell introduced the method of the material hardness determination in which a steel spherical indentor penetrates into the studied material. The method permits estimation of some material characteristics (elasticity modulus, Poisson's ratio, yield stress, and hardness). As a supplement, Sheivechman and Suzdalnitsky [19] suggested an indentational method for determination of the adhesion parameter. Francis [20] showed that spherical indentation in an elastoplastic material comprises in three distinct stages. In the elastic stage the deformation is reversible and can be described by the Hertz solution. A transitional regime sets in with formation of a plastic zone under the contact region. In the third stage the behavior of the material under the indentor remains fully plastic.

We resort to experimental indentation data obtained by Field and Swain [21,22] in order to determine the operator equation relating the stresses and strains in the form (1.3). They represent the results of action of a 10 μm indentor on an aluminum specimen. These data (op. cit., [22], Fig. 12) are shown in Fig. 4. The test values are marked by asterisks.

We consider the relationship between the load P and the penetration depth h in the form

$$h^{1.5}(p)=\frac{3P_0}{4E_0R^{0.5}}(p+D^{-\beta(p)}p)$$

$$=h_e^{1.5}(p)+c_p\int_0^p\frac{\xi^{-1+\beta(\xi)}}{\Gamma(\beta(\xi))}(p-\xi)d\xi, \qquad (6.1)$$

where $c_p=3P_0/4E_0R^{0.5}$, $h_e(p)=c_pp$—the linear (elastic) part of the deformation, p—the dimensionless compressive load.

Denoting $y(p)=c_p^{-1}h^{1.5}(p)$, $y_e(p)=c_p^{-1}h_e^{1.5}(p)$ and

$$g(\xi)=\frac{\xi^{-1+\beta(\xi)}}{\Gamma(\beta(\xi))} \qquad (6.2)$$

we obtain the Volterra equation

$$y(p)=\int_0^p g(\xi)(p-\xi)d\xi+y_e(p). \qquad (6.3)$$

Consider the polynomial approximation of the experimental data:

$$y(p)=p+\sum_{k=2}^{n}a_kp^k. \qquad (6.4)$$

Substituting (6.4) in (6.3) and solving obtained equation we easily find that the polynomial approximation $g(\xi)$ has to satisfy

$$g(\xi)=\frac{1}{a_1}\sum_{k=0}^{n-2}(k+1)(k+2)a_{k+2}\xi^k. \qquad (6.5)$$

Equating (6.2) and (6.5) and using a table of test loads p_i, $i=1, N$, we obtain a set of nonlinear algebraic equations

$$P_i^{\beta}-\Gamma(\beta)P_i\sum_{k=0}^{n-2}(k+1)(k+2)a_{k+2}P_i^k=0.$$

Further, we find β_i for each point (P_i,h_i), with the aid of a routine iterative procedure of interval halving $\beta_i^{(j+1)}=F(P_i,h_i,\beta_i^{(j)})$, $0\leq\beta_i^j<1$, $j=0,1,2\ldots$.

The behavior of the order function $\beta(P)$ is shown in Fig. 5. The operator $\mathbf{D}^{-\beta(P)}$ performs integration. The function $\beta(P)$ increases up to 0.3 and subsequently stabilizes at that level. This result leads to an alternative conclusion relative to Field and Swain [22], who held that for loads of 20 mN and higher the behavior of aluminum is fully plastic. According to our own results, the aluminum specimen loaded by an indentor of small radius 10 μm exhibits propagation of the plastic flow at the initial stages of penetration (increase of the operator order) and acquires some ''equilibrium'' semi-plastic state (saturation of the operator order below 1). As a comparative example, indentation of elastic material was also tested. Figure 3 in Field and Swain [21] presents the dependence of the depth penetration on the applied load ($P<178$ mN). Here, the loaded material is steel and the spherical indentor has a diameter of 125 μm. Plastic flow does not appear in this case, the deformation remains elastic and the calculation shows that the order of integration $\beta(P)$ is constant at zero.

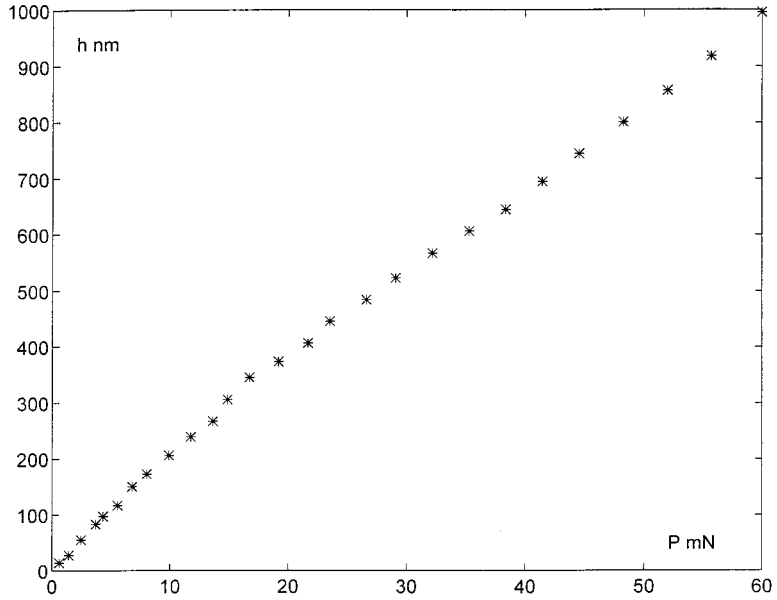


Fig. 4 Experimental results of indentation

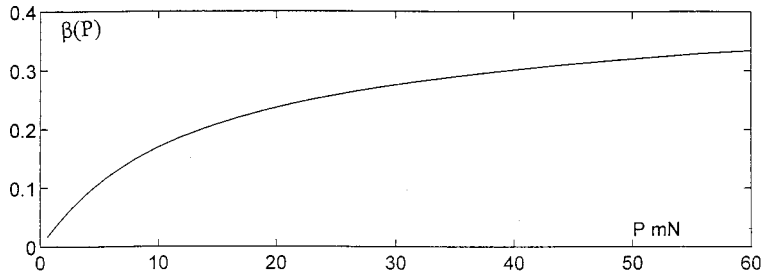


Fig. 5 Order function $\beta(P)$

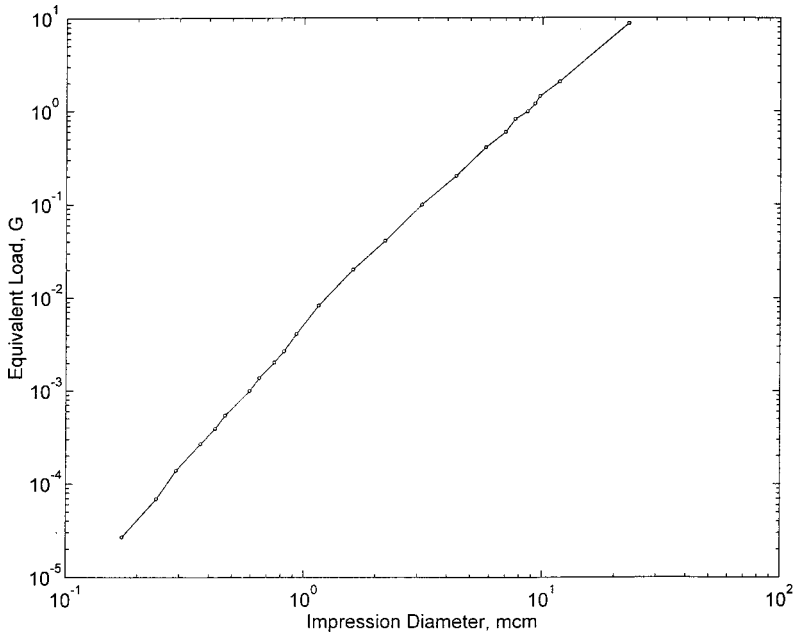


Fig. 6 Results of indentation for polymer

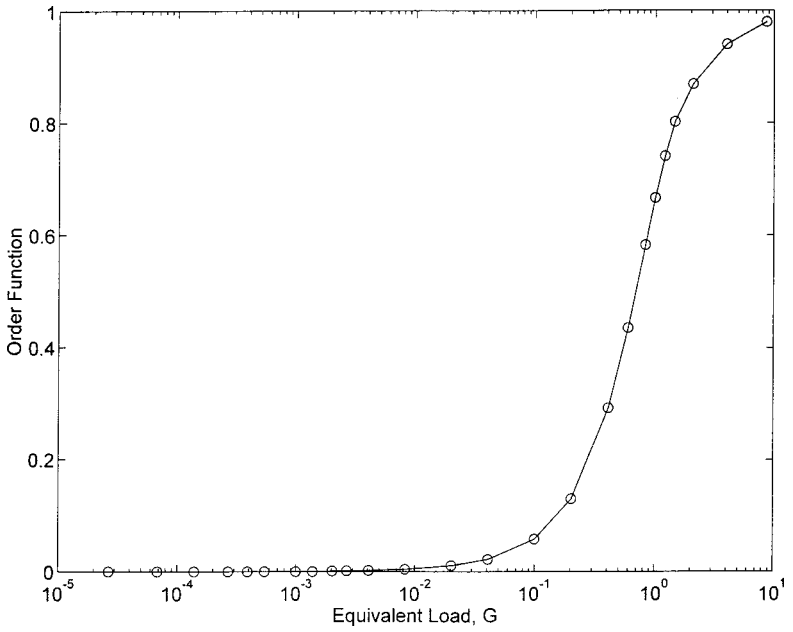


Fig. 7 Order function $\alpha(p)$ for polymer

As a final example let us consider indentation test data obtained by Barlow [23] on a vinyl specimen. Log load versus log impression diameter results are plotted in Fig. 6. Note that the test covered a total of 5.5 decades of equivalent load. In his original work, Barlow concluded as follows. (1) For equivalent loads below 3 mg, the vinyl specimen remains fully elastic (a constant slope of 3 on the log load versus log diameter curve, according to the Hertz solution (5.1)). (2) For the load interval of 3–750 mg, the material exhibits partly plastic behavior. (3) For loads above 750 mg, the behavior is fully plastic.

As in the preceding example, let us consider the relationship between the load P and the contact diameter a as

$$a^3(p)=\frac{3RP_0}{4E_0}D^{-\alpha(p)}p+a_e^3=c_a\int_0^p\frac{\xi^{-1+\alpha(\xi)}}{\Gamma(\alpha(\xi))}(p-\xi)d\xi+a_e^3 \tag{6.4}$$

where $a_e(p)$—the elastic part of the deformation, p—the dimensionless load, $c_a=3P_0R/4E_0$.

The behavior of the reconstructed order function $\alpha(p)$ is shown in Fig. 7. While Barlow's first conclusion is clearly confirmed by IDOVO ($\alpha=0$ for $P<3$ mg), the fully plastic range in terms of the order function is not really reached. Unlike its aluminum counterpart, no stable semiplastic range is observed for the vinyl specimen; rather, the order function increases continuously up to the limit value $\alpha=1$.

Conclusions

The suggested integro-differential operator with dynamic order (order function) generalizes the well-known operator of fractional integration. For an integer order value (positive or negative) this operator performs the standard operations of differentiation or integration. A rapidly convergent iteration process for calculation of the order function is suggested.

The order function as a material state function permits better understanding of the viscoelastic and elastoplastic deformation processes. In fact, the considered examples of spherical indentation show that the reconstructed order functions clearly bring out the well-known differences in the elastic-plastic features of such diverse materials as aluminum, steel, and vinyl. They support our understanding that steel is almost absolutely elastic throughout the investigated range of applied loads, that aluminum becomes partially plastic under elevated loads, and that vinyl goes through three stages—elastic under low loads, semi-elastic/semiplastic under intermediate loads, and almost absolutely plastic under elevated loads.

The authors believe that the suggested approach will have both a scientific and a practical impact on this field of science and engineering.

Acknowledgments

The authors wish to express their appreciation to the two referees for their valuable suggestions and comments and to Mr. E. Goldberg for his valuable help and advice in preparation of the manuscript.

Appendix A

Let operators D^α and D^β act on $x(t)$ successively: $\mathbf{D}^\alpha\mathbf{D}^\beta x(t)$. Let limited ourselves to the case of $0<\alpha,\beta<1$. Applying the Laplace transform to this product and using the property of the Laplace transform for a convolution, we have

$$\begin{aligned}\mathbf{L}\{\mathbf{D}^\alpha\mathbf{D}^\beta x(t)\}&=\mathbf{L}\left\{\frac{d^m}{dt^m}\int_0^t\frac{(t-\xi)^{m-1-\alpha(t-\xi)}}{\Gamma(m-\alpha(t-\xi))}\right.\\&\quad\left.\times d\xi\frac{d^n}{d\xi^n}\int_0^\xi\frac{(\xi-\eta)^{n-1+\beta(\xi-\eta)}}{\Gamma(n-\beta(\xi-\eta))}x(\eta)d\eta\right\}\\&=p^m\mathbf{L}\left\{\frac{t^{m-1-\alpha(t)}}{\Gamma(m-\alpha(t))}\right\}p^n\mathbf{L}\left\{\frac{t^{n-1-\beta(t)}}{\Gamma(n-\beta(t))}\right\}\mathbf{L}\{x(t)\}\\&=\mathbf{L}\left\{\frac{d^n}{dt^n}\int_0^t\frac{(t-\xi)^{n-1-\beta(t-\xi)}}{\Gamma(n-\beta(t-\xi))}\right.\\&\quad\left.\times d\xi\frac{d^m}{d\xi^m}\int_0^\xi\frac{(\xi-\eta)^{m-1+\alpha(\xi-\eta)}}{\Gamma(m-\alpha(\xi-\eta))}x(\eta)d\eta\right\}\\&=\mathbf{L}\{\mathbf{D}^\beta\mathbf{D}^\alpha x(t)\}.\end{aligned}$$

According to the theorem of the single-valuedness of transforms we have $\mathbf{D}^\alpha\mathbf{D}^\beta=\mathbf{D}^\beta\mathbf{D}^\alpha$. Hence, operator $\mathbf{D}$ is commutative.

Appendix B

Let

$$0<\alpha(t)<1, \tag{B1}$$

$0\leqslant\beta(t)<1$, $y=\mathbf{D}^{-\alpha}x$, $z=\mathbf{D}^\beta y$. Then

$$\begin{aligned}z(t)&=\mathbf{D}^\beta\mathbf{D}^{-\alpha}x(t)\\&=\frac{d}{dt}\int_0^t\frac{(t-\xi)^{-\beta(t-\xi)}}{\Gamma(1-\beta(t-\xi))}d\xi\int_0^\xi\frac{(\xi-\eta)^{-1+\alpha(\xi-\eta)}}{\Gamma(\alpha(\xi-\eta))}x(\eta)d\eta\\&=\frac{d}{dt}\int_0^t x(\xi)d\xi\int_0^{t-\xi}\frac{(t-\xi-\eta)^{-\beta(t-\xi-\eta)}\eta^{-1+\alpha(\eta)}}{\Gamma(1-\beta(t-\xi-\eta))\Gamma(\alpha(\eta))}d\eta.\end{aligned}$$

Hence, $z(t)=x(t)$ and $\mathbf{D}^\beta$ is the reciprocal operator to $\mathbf{D}^{-\alpha}$, if α and β are related by the dependence

$$\int_0^t\frac{(t-\xi)^{-\beta(t-\xi)}\xi^{-1+\alpha(\xi)}}{\Gamma(1-\beta(t-\xi))\Gamma(\alpha(\xi))}d\xi=U(t) \tag{B2}$$

where $U(t)$ is the unit function. Thus, we obtain (4.2).

Denoting $f(t)=t^{-1+\alpha(t)}/\Gamma(\alpha(t))$, $g(t)=t^{-\beta(t)}/\Gamma[1-\beta(t)]$, one can rewrite equation (*B*2) as follows:

$$\int_0^t f(\xi)g(t-\xi)d\xi=U(t). \tag{B3}$$

Let us assume that for $t\to0$

$$\alpha(t)=ct+o(t). \tag{B4}$$

Then $f(0)=c$. The left-hand side of (B3) is the convolution of the functions $f(t)$ and $g(t)$. Their Laplace transforms $F(p)$ and $G(p)$ according to (B3) are related by the equation $F(p)G(p)=1/p$. Further, $f(t)$ in the neighborhood of $t=0$ is a regular function, i.e., for $t\to0$ we have $f(t)=c+\Sigma_{k=1}^\infty a_kt^k$. Then, for $p\to\infty$, $F(p)=(c/p)+\Sigma_{k=1}^\infty a_{(k}(k!/p^{k+1})$, hence $G(p)=(1/c)+\Sigma_{k=1}^\infty bk(k!/p^{k+1})$, $g(t)=(1/c)\delta(t)+\varphi(t)$, where $\varphi(t)$ is a function regular in the neighborhood of $t=0$, $\delta(t)$ is the delta function. Equation (*B*3) may be rewritten as follows:

$$\int_0^f f(\xi)\left[\frac{1}{c}\delta(t-\xi)+\varphi(t-\xi)\right]d\xi=U(t)$$

or

$$\int_0^t f(\xi)\varphi(1-\xi)d\xi=U(t)-\frac{1}{c}f(t). \tag{B5}$$

After numerical integration of (B5) using the trapezoid rule, we obtain the iterative sequence

$$\varphi_0=\varphi_1=\frac{2}{ch}\frac{c-f_1}{c+f_1},$$

$$\varphi_j=\frac{2}{c^2h}(c-f_j)-\frac{1}{c}\left[\varphi_0 f_j+2\sum_{k=1}^{j-1} f_k\varphi_{j-k}\right], \quad j=1,\ldots,N,$$

where $f_j=f(t_j)$, $\varphi_j=\varphi(t_j)$, $t_j=jT/N$, $(0,T)$ is the interval of integration and N the number of its subdivisions for numerical integration.

Remark. When condition (B1) is replaced by the inequality $c_1<\alpha(t)<c_2$, (B2) undergoes changes connected with the value m. The derivation of this equation is similar to that of (B2). Condition (B4) is of interest in applications. If α is constant, $\alpha>0$, Eq. (*B*2) is satisfied for $\alpha=\beta$ in accordance with the properties of the gamma and beta functions and $D^{-\alpha}$ is the reciprocal operator to D^{α}.

References

[1] Billington, E. W., and Tate, A., 1981, *The Physics of Deformation and Flow*, McGraw-Hill, New York.

[2] Krajcinovic, D., 1996, *Damage Mechanics*. Elsevier, Amsterdam.

[3] Lee, E. N., and Radok, J. R. M., 1960, ''The Contact Problem for Viscoelastic Bodies,'' ASME J. Appl. Mech., **82**, pp. 438–444.

[4] Sneddon, J. N., 1979, *The Use of Operators of Fractional Integration in Applied Mathematics*, PWN, Warsaw.

[5] Bagley, R. L., and Calico, R. A., 1991, ''Fractional Order State Equations for the Control of Viscoelastically Damped Structures.'' J. Guidance Control Dynam. , **14**, No. 2, pp. 304–311.

[6] Bagley, R. L., Swinney, D. V., and Griffin, K. E., 1993, ''Fractional Order Calculus Model of the Generalized Theodorsen Function,'' J. Aircr., **30**, No. 6, pp. 1003–1005.

[7] Rossikhin, Yu. A., and Shitikova, M. V., 1997, ''Application of Fractional Operators to the Analysis of Damped Vibrations of Viscoelastic Single-Mass Systems,'' J. Sound Vib., **199**, No. 4, pp. 567–586.

[8] Lee, E. N., and Rogers, T. G., 1963, ''Solution of Viscoelastic Stress Analysis Problems Using Measured Creep or Relaxation Functions,'' ASME J. Appl. Mech., **30**, pp. 127–133.

[9] Rabotnov, Yu. N., 1977, *Elements of Hereditary Mechanics of Solids*, Nauka, Moscow.

[10] Bardenhagen, S. G., Stout, M. G., and Gray, G. T., 1997, ''Three-Dimensional, Finite Deformation, Viscoplastic Constitutive Models for Polymeric Materials,'' Mech. Mater., **25**, No. 4, pp. 235–253.

[11] Lubliner, J., 1990, *Plasticity Theory*, Macmillan, New York.

[12] Skrzypek, J. J., and Hetnarski, R., 1993, *Plasticity and Creep*, CRC Press, Boca Raton, FL.

[13] Valanis, K. S., 1971, ''A Theory of Viscoplasticity Without a Yield Surface I—II,'' Archiwum Mechaniki Stosowanej, **23**, No. 4, pp. 517–551.

[14] Valanis, K. C., and Read, H. E., 1986, ''An Endochronic Plasticity Theory for Concrete,'' Mech. Mater., **5**, No. 3, pp. 277–295.

[15] Valanis, K. C., and Yilmazer, U., 1978, ''An Intrinsic Time to Fracture Criterion for Amorphous Polymers,'' Eng. Fract. Mech., **10**, No. 3, pp. 659–676.

[16] Rajic, N., et al., 1997, ''Plastic Zone Size Determination by Temperature Measurement,'' ASME J. Eng. Mater. Technol., **119**, pp. 32–39.

[17] Oldham, K. B., and Spanier, J., 1974, *The Fractional Calculus*, Academic Press, New York.

[18] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.

[19] Sheivekhman, A., and Suzdalnitsky, J., 1997, ''Determination of the Elasto-plastic Properties of Engineering Materials From Spherical Indentation,'' J. Mech. Behav. Mater., **8**, No. 4, pp. 283–294.

[20] Francis, H. A., 1976, ''Phenomenological Analysis of Plastic Spherical Indentation,'' ASME J. Eng. Mater. Technol., **98**, pp. 272–281.

[21] Field, J. S., and Swain, M. V., 1993, ''A Simple Predictive Model for Spherical Indentation,'' J. Mater. Res., **8**, No. 2, pp. 297–305.

[22] Field, J. S., and Swain, M. V., 1995, ''Determining the Mechanical Properties of Small Volumes of Material From Submicrometer Spherical Indentations,'' J. Mater. Res., **10**, No. 1, pp. 101–112.

[23] Barlow, D. A., 1973, ''The Indentation and Scratch Hardness of Plastics,'' ASME J. Eng. Mater. Technol., **95**, pp. 243–251.

S. Roy
Assistant Professor,
Mem. ASME

W. X. Xu
Graduate Research Assistant

Department of Mechanical Engineering,
University of Missouri-Rolla,
1870 Miner Circle,
Rolla, MO 65401

S. J. Park
Graduate Research Assistant

K. M. Liechti
Professor

Center for the Mechanics of Solids,
Structures and Mechanics,
ASE/EM WRW 110C,
The University of Texas,
Austin, TX 78712

Anomalous Moisture Diffusion in Viscoelastic Polymers: Modeling and Testing

It is now well known that Fick's Law is frequently inadequate for describing moisture diffusion in polymers or polymer composites. Non-Fickian or anomalous diffusion typically occurs when the rates of diffusion and viscoelastic relaxation in a polymer are comparable, and the ambient temperature is below the glass transition temperature (T_g) of the polymer. As a result, it is necessary to take into account the time-dependent response of a polymer, analogous to viscoelastic relaxation of mechanical properties, in constructing such a model. In this paper, a simple yet robust methodology is proposed that would allow characterization of non-Fickian diffusion coefficients from moisture weight gain data for a polymer below its T_g. Subsequently, these diffusion coefficients are used for predicting moisture concentration profiles through the thickness of a polymer. Moisture weight gain data at different temperatures for an epoxy adhesive is employed to calibrate the model. Specimen thickness independence of the modeling parameters is established through comparison with test data. A finite element procedure that extends this methodology to more complex shapes and boundary conditions is also validated.
[S0021-8936(00)02402-8]

Introduction

It is now widely recognized that moisture plays a significant role in influencing the mechanical behavior, and therefore, the long-term durability of polymers and polymer matrix composites (PMC). Numerous diffusion models have been proposed over the years for modeling hygrothermal effects in polymers and PMC. The one most frequently used by researchers is the one-dimensional Fickian model due to its simplicity and mathematical tractability. Unfortunately, this model tends to overestimate the moisture absorption in panels for short diffusion time ([1]). Some researchers have suggested that the deviation can be explained by a two-stage Fickian process ([2,3]). Others claim that the diffusion process in a PMC is really non-Fickian ([4,5]). The applicability of Fick's law for a given material system under a specified loading cannot be guessed a priori but must be determined from moisture absorption/desorption test data.

Frisch [6] and Crank [7] were among the first researchers to recognize and attribute non-Fickian moisture transport in resins to time-dependent molecular mechanisms within a polymer. In a landmark article, Frisch [8] suggested that a polymer below its glass transition temperature (T_g) must possess history-dependent diffusion coefficients and experience time-dependent changes in surface concentrations in order to maintain sorption-equilibrium at its boundaries. These time-dependencies are intrinsically related to the relaxation times for molecular rearrangement in the polymer. To model this phenomenon, Weitsman [9] applied the basic principles of continuum mechanics and irreversible thermodynamics to derive governing equations and boundary conditions for coupled stress-assisted diffusion in elastic and viscoelastic materials. Following an approach originally proposed by Biot [10] and adapted by Schapery [11], viscoelasticity was introduced by means of scalar-valued internal state variables, also referred to as hidden coordinates. The results of the analysis allowed an insight into the mechanism that causes a time-drift toward equilibrium at the boundary of a viscoelastic material subjected to a constant chemical potential of the ambient vapor. The governing equations also indicated that the saturation levels vary quadratically with stress and that they can be expressed in terms of the dilatational and the deviatoric stress invariant. However, the mathematical expression for diffusivity obtained from the analysis was too cumbersome for experimental characterization and it was recommended that simplifications should be used in practical cases. In this vein, Cai and Weitsman [12] developed a simplified methodology within the framework of Fickian diffusion with step-wise time-varying boundary conditions. Their procedure allowed the reduction of non-Fickian moisture weight-gain data in a manner that enabled the evaluation of the diffusion coefficient and through-thickness concentration profiles. Even so, an iterative numerical procedure had to be employed for the evaluation of the time varying boundary concentrations in the form of an exponential (Prony) series. They also discussed characterization of the non-Fickian diffusion coefficient using an approach that involved the Carson transform.

Subsequently, Weitsman [13] extended the rigorous thermodynamic model to incorporate polymer ''free volume'' as an internal state variable in order to include the effect of physical aging on moisture diffusion. The resulting governing equations were not amenable to closed-form solution due to mathematical complexity, although it was observed that the diffusion equation followed a time-retardation process analogous to mechanical viscoelastic response, and that it exhibited an aging behavior characteristic of glassy polymers. Employing a similar approach, Weitsman [14] developed a model for coupled damage and moisture transport in an elastic, transversely isotropic, fiber-reinforced polymer composite. The damage entity was represented as a skew-symmetric tensor and was included in the model as an internal state variable. All of these models are mathematically complex and are not amenable to simple closed-form solutions.

For stretched polymer sheets where the diffusion-governing equations are coupled with mechanical response through volumetric strain, Roy et al. [15] presented a numerical procedure for solving coupled diffusion equations using an approach based on

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 12, 1998; final revision, Oct. 29, 1999. Associate Technical Editor: I. M. Daniel. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

the free volume theory. Sancaktar and Baechtler [16] showed that there is a substantial change in the free volume ratio in a polymer as a result of stress whitening, which in turn results in an increase in moisture uptake in the stress-whitened region. A multivalued diffusion coefficient, based on an earlier model proposed by Wong and Broutman [17,18], was employed to model this effect. More recently, Roy [19] derived governing equations for history-dependent diffusion using irreversible thermodynamics, and developed a novel numerical framework for solving the complex non-Fickian governing equations using the finite element method.

The objective of this paper is to present a new methodology that enables characterization of non-Fickian diffusion coefficients from moisture weight gain data in a polymer. The proposed approach is very simple yet robust, and does not require Carson transforms and iterative procedures. Subsequently, these diffusion coefficients are used to predict moisture concentration profiles through the thickness of a polymer. Moisture weight gain data at different temperatures for an epoxy adhesive is employed to calibrate the model. Because the specimens were fully immersed in a salt solution, the variation in the moisture boundary concentration with time is assumed to be negligible. A finite element procedure that would extend this methodology to more complex shapes and time-varying moisture boundary conditions is also discussed.

Governing Equations for Diffusion in Viscoelastic Media

Using the concept of internal degrees of freedom in a polymer molecule within the framework of irreversible thermodynamics, Frisch [8] employed the Gibbs-Duhem relation to model penetrant flux in a two component system with component 1 as the polymer and component 2, the penetrant. He showed that for a polymer below its glass transition temperature, the diffusivity and boundary concentration are not constants but vary continuously with time. Based on this theory for polymers with glass-like transition, it can be rigorously shown that, below its glass-transition temperature (T_g), the diffusion governing equation in a polymer takes the form

$$\frac{\partial C}{\partial t}=\frac{\partial}{\partial X_i}\left(\overline{D}_{ij}\frac{\partial C}{\partial X_j}\right) \tag{1}$$

where the effective diffusivity is given by

$$\overline{D}_{ij}(X_i,T,\sigma,t)=\left[\frac{L_{ij}}{(1-C)}\left(\frac{\partial \mu}{\partial C}+\sum_{r=1}^{n}\frac{\partial \mu}{\partial \ln q_r}\right)\right] \tag{2}$$

where C is moisture concentration, X_i are the spatial coordinates, T is temperature, σ represents an invariant stress measure, t is time, L_{ij} are the Onsager coefficients, μ is the chemical potential of the diffusing vapor in the polymer, and q_r are hidden coordinates that define the internal motion of individual chain segments in polymer molecules, thereby representing viscoelastic behavior in a polymer. In addition, Eq. (1) is subject to the sorption-equilibrium boundary condition that requires the chemical potential of the ambient vapor to remain unchanged with time.

For an unstressed isotropic viscoelastic polymer, the time-varying effective diffusivity can be expressed in the form of a Prony series ([8]),

$$\tilde{D}(T,t)=D_0(T)+\sum_{r=1}^{n}D_r(T)(1-e^{-t/\tau_r}) \tag{3}$$

where D_0, D_r are the unknown temperature-dependent Prony coefficients, τ_r are the corresponding retardation times, and n is the number of terms in the Prony series.

Diffusion With Time-Varying Diffusivity. The governing equation for one-dimensional diffusion in a polymer sheet of thickness h is given by

$$\frac{\partial C}{\partial t}=D(t)\frac{\partial^2 C}{\partial x^2} \tag{4}$$

where $D(t)$ is the time-varying diffusion coefficient assumed to be uniform through the thickness of the polymer.

The concentration boundary conditions are

$$C(0,t)=C(h,t)=C_0. \tag{5}$$

Defining a change in variable

$$dU=D(t)dt, \tag{6}$$

giving

$$U(t)=\int_0^t D(t')dt'. \tag{7}$$

With this change of variable, Eq. (4) reduces to

$$\frac{\partial C}{\partial U}=\frac{\partial^2 C}{\partial x^2}. \tag{8}$$

Equation (8) can now be solved using a standard solution for constant diffusivity (i.e., $D=1$) to give C as a function of x and U. The modified boundary conditions in terms of the variable U take the form

$$C(0,U)=C(h,U)=C_0 \tag{9}$$

subject to the initial condition

$$C(x,0)=C_i. \tag{10}$$

The solution to the initial boundary value problem defined by Eqs. (8), (9), and (10) is given by [7]

$$\frac{C-C_i}{C_0-C_i}=1-\frac{4}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\times\exp\{-(2n+1)^2\pi^2U/h^2\}\cos\frac{(2n+1)\pi x}{h}. \tag{11}$$

If M_i and M_t denote the total amount of diffusing substance which has entered the sheet at time 0 and t, respectively, and M_∞ is the corresponding quantity after infinite time, then integrating Eq. (11) over the thickness h gives

$$\frac{M_t-M_i}{M_\infty-M_i}=1-\frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}\exp\{-(2n+1)^2\pi^2U/h^2\}. \tag{12}$$

Assuming that the diffusivity can be expressed in the form of a Prony Series

$$D(t)=D_0+\sum_{r=1}^{R}D_r(1-e^{-t/\tau_r}). \tag{13}$$

Substituting Eq. (13) in (7) gives

$$U(t)=D_0t+\sum_{r=1}^{R}D_r[t+\tau_r(e^{-t/\tau_r}-1)]. \tag{14}$$

Substituting Eq. (14) in (11) results in an expression for moisture concentration as a function of distance and time

$$\frac{C-C_i}{C_0-C_i}=1-\frac{4}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\exp\left\{\frac{-(2n+1)^2\pi^2}{h^2}\times\left\{D_0t+\sum_{r=1}^{R}D_r[t+\tau_r(e^{-t/\tau_r}-1)]\right\}\right\}\cos\frac{(2n+1)\pi x}{h}. \tag{15}$$

Finally, substituting Eq. (14) in (12) gives the moisture weight-gain fraction as a function of time

$$\frac{M_t - M_i}{M_\infty - M_i} = 1 - \frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}\exp\left\{\frac{-(2n+1)^2\pi^2}{h^2}\times\left\{D_0 t + \sum_{r=1}^{R} D_r[t+\tau_r(e^{-t/\tau_r}-1)]\right\}\right\}. \quad (16)$$

Finite Element Formulation. In order to extend the simple one-dimensional analytical model to more complex shapes and boundary conditions, a three-dimensional finite element code (NOVA-3D) was developed. The variational (weak) form of Eq. (4) in three dimensions is given by

$$\int_{v^{(e)}}\left[u\frac{\partial C^t}{\partial t} + D^t\frac{\partial u}{\partial X_i}\frac{\partial C^t}{\partial X_i}\right]dV - \int_{A^{(e)}}\left[u\left(D^t\frac{\partial C^t}{\partial X_i}\right)n_i\right]dA = 0, \quad i=1,3 \quad (17)$$

where u is an admissible variational test function. Based on the variational statement, the diffusion boundary conditions can now be identified as

$$\left(D^t\frac{\partial C^t}{\partial X_i}\right)n_i + \hat{q} = 0 \quad \text{on } A_1^{(e)} \quad \text{(specified solvent flux)}$$

$$C = \hat{C} \quad \text{on } A_2^{(e)} \quad \text{(specified concentration)}$$

where

$$A_1^{(e)} + A_2^{(e)} = A^{(e)}$$

and n_i are the components of the unit outward normal at the boundary. Thus,

$$\int_{v^{(e)}}\left[u\frac{\partial C^t}{\partial t} + D^t\frac{\partial u}{\partial X_i}\frac{\partial C^t}{\partial X_i}\right]dV = -\int_{A_1^{(e)}} u\hat{q}\,dA. \quad (18)$$

Finite Element Approximation. A standard finite element interpolation of the concentration field over each element is given by

$$C(X_i,t) = \sum_{j=1}^{N} N_j(X_i)C_j(t) \quad (19)$$

where C_j are the nodal concentrations, N_j are the interpolation functions, and N is the number of nodes per element. Substituting Eq. (19) in (18) and employing matrix notation, Eq. (18) becomes

$$[T^{(e)}]\{\dot{C}\} + [K^{(e)}]\{C\} = \{F^{(e)}\} \quad (20)$$

where the subscript (e) is used to denote that the equations are satisfied over each element and

$$T_{jk}^e = \int_{v^{(e)}}(N_j N_k)dV$$

$$K_{jk}^e = \int_{v^{(e)}}\left\{D^t\frac{\partial N_j}{\partial X_i}\frac{\partial N_k}{\partial X_i}\right\}dV$$

$$F_j^e = -\int_{A_1^{(e)}} N_j\hat{q}\,dA \quad , \quad i=1,3 \quad \text{and} \quad j,k=1,N.$$

The time derivative $\{\dot{C}\}$ is approximated using a standard theta family of approximations, yielding for time t_n and t_{n+1},

$$[A^{(e)}]\{C\}_{n+1} + [B^{(e)}]\{C\}_n = \{p^{(e)}\}_n, \quad (21)$$

where

$$[A^{(e)}] = [T^{(e)}] + \theta\Delta t_{n+1}[K^{(e)}]$$

$$[B^{(e)}] = [T^{(e)}] + (1-\theta)\Delta t_{n+1}[K^{(e)}]$$

$$\{P^{(e)}\} = \Delta t_{n+1}(\theta\{F^{(e)}\}_{n+1} + (1-\theta)\{F^{(e)}\}_n).$$

Equation (21) is solved using a value of $\theta = 0.5$, which corresponds to the Crank-Nicholson scheme and is unconditionally stable. Note that for $n=1$, the value of the starting concentration in Eq. (21) is known from initial conditions.

Diffusion Experiments

Absorption experiments were conducted on neat epoxy adhesive specimens in a salt water solution. The epoxy specimens were cut from 0.51-mm (0.02-in.) thick sheets that were provided by Oak Ridge National Laboratory (ORNL) personnel. Special care was taken to eliminate air bubbles from the sheets. The 25.4-mm (1-in.) square specimens were dehydrated for 24 hours, weighted, dehydrated for an additional 65 hours, and weighed to ensure complete dehydration. Both steps were conducted at 70°C in order to accelerate dehydration and the length of the second step was established by a series of preliminary experiments. The specimens were immersed in a salt water solution at four different temperature levels (23, 50, 60, and 70°C). The solution was formed by mixing salt (five percent by weight) with de-ionized water. The temperatures were controlled to within ±1°C. The mass measurements were made using a Mettler balance that has a range of 160 g with a resolution of 0.1 mg. The initial mass of the specimens was typically 0.35 g, so that changes in mass of approximately 0.05 percent could be resolved. The specimens were weighted after wiping excess liquid from the surface of the specimen. The frequency of the measurements decreased with time as saturation approached. The duration of the experiments was approximately six months. Subsequently, in order to ensure that the model parameters are indeed material properties that are independent of specimen thickness, a separate set of absorption tests was performed at 23°C on epoxy specimens of thickness 1.194 mm (0.047 in.).

Results and Discussion

The salt-water weight gain data clearly indicated that the diffusion characteristic of this epoxy does not conform to Fick's law for the range of time and temperature considered in this study. As a result, it became necessary to apply the non-Fickian diffusion model described in the previous sections in order to model diffusion in this material. For this purpose, time-varying diffusivity characterization data necessary for the viscoelastic diffusion computer modeling were extracted from the salt water absorption test data. A Prony series representation of the time-varying diffusivity for the epoxy was obtained by fitting Eq. (16) to the test data employing a least-squares approach. The Prony coefficients and the corresponding retardation times obtained in this manner for the epoxy adhesive at 23°C, 50°C, 60°C, and 70°C are listed in Table 1. The same retardation times are used to describe the time-dependent diffusion in epoxy at the four temperatures. Because the specimens were fully immersed in a salt solution, the variation in the moisture boundary concentration with time was assumed to be negligible. The boundary moisture concentration used as input to the model for each temperature is listed in Table 2. The boundary concentrations were obtained by dividing the saturation moisture weight gain for each specimen by the volume of that specimen. A significant increase in boundary concentration with temperature was observed as indicated by the data in Table 2.

Figure 1 shows the comparison of closed-form solution and finite element analysis (FEA) with test data for epoxy adhesive at 23°C, assuming time-varying diffusivity. In this figure, percent moisture weight gain is plotted versus square root of time, in hours. A Fickian uptake curve with constant diffusivity is also included to underscore the inaccuracy of using Fick's law to model diffusion in this material. The parameters used for modeling Fickian uptake are tabulated in Table 3 for the four test tem-

Table 1 Diffusion Coefficient Data for Epoxy

Diffusion Coefficient (mm^2/s)	Temperature 23°C	Temperature 50°C	Temperature 60°C	Temperature 70°C	Retardation Time, τ_r (sec)
D_0	2.68998×10^{-7}	2.7849×10^{-7}	1.3172×10^{-6}	9.7293×10^{-7}	–
D_1	-2.2383×10^{-7}	-4.6998×10^{-7}	-1.6365×10^{-6}	-1.3923×10^{-6}	600
D_2	1.3841×10^{-8}	9.2648×10^{-7}	1.3012×10^{-6}	2.4826×10^{-6}	3000
D_3	5.3247×10^{-8}	-3.341×10^{-7}	-3.9708×10^{-7}	-1.597×10^{-6}	6000
D_4	2.19705×10^{-7}	-1.4022×10^{-7}	-5.4208×10^{-7}	-3.034×10^{-7}	30000
D_5	-2.5323×10^{-7}	-2.3545×10^{-7}	2.7423×10^{-9}	-6.6972×10^{-8}	60000
D_6	-7.4449×10^{-8}	-1.8449×10^{-8}	-3.9448×10^{-8}	-8.3294×10^{-8}	300000

Table 2 Boundary Concentration Data for Epoxy

Temperature (°C)	Boundary Concentration (gm/mm^3)
23	3.366×10^{-5}
50	4.035×10^{-5}
60	4.18×10^{-5}
70	4.16×10^{-5}

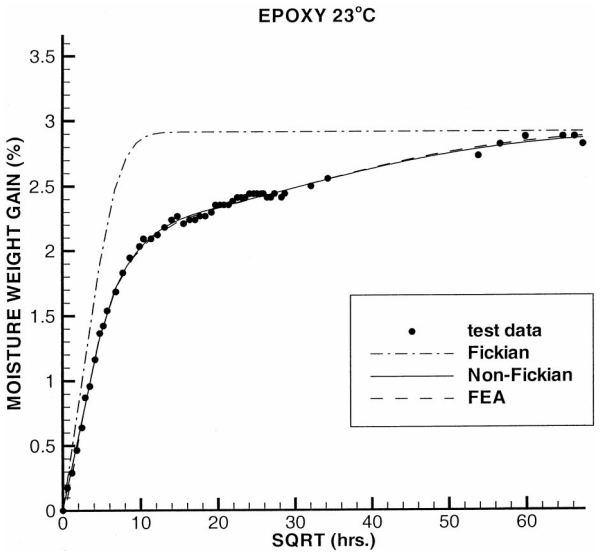


Fig. 1 Percent moisture weight gain versus square root of time for epoxy at 23°C

Table 3 Fickian Curve Fit Parameters for Epoxy at Test Temperatures

Temperature (C)	Diffusivity ($\times10^{-7}$ mm^2/s)	M_{max} (percent)
23	1.46029	2.43985
50	2.93600	2.78245
60	3.85563	2.97554
70	3.68567	3.26497

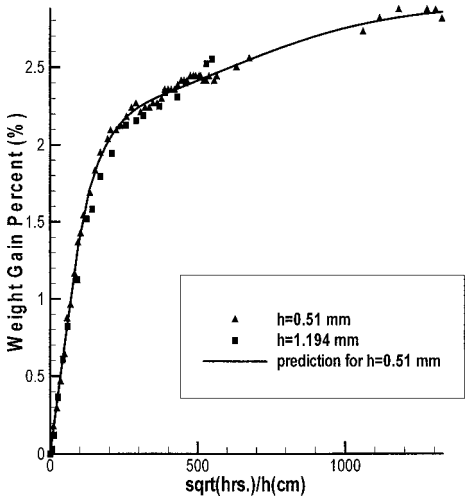


Fig. 2 Reduced absorption plot for epoxy specimens of different thickness at 23°C

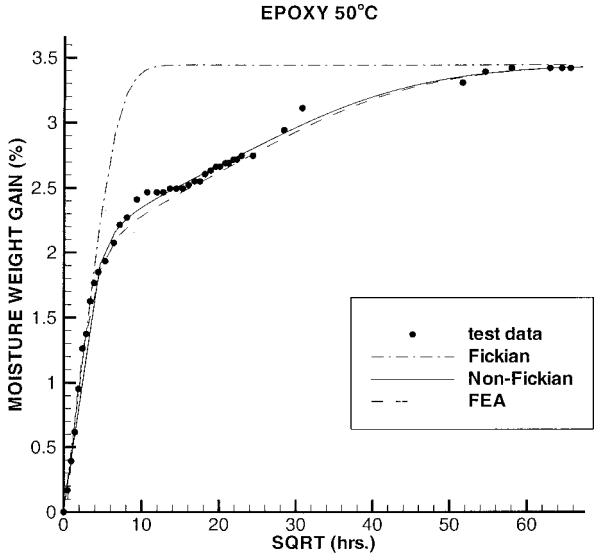


Fig. 3 Percent moisture weight gain versus square root of time for epoxy at 50°C

peratures. The agreement between the non-Fickian model prediction and test data is quite good. However, in order to ensure that the model parameters contained in Table 1 are indeed material properties that are independent of specimen thickness, a separate set of absorption tests was performed at 23°C on epoxy specimens of thickness 1.194 mm (0.047 in.). Figure 2 depicts the reduced absorption plot at a test temperature of 23°C for epoxy specimens of thickness 0.51 mm. (0.02 in.) and 1.194 mm (0.047 in.), respectively, together with the analytical model prediction. Excellent correlation is observed between test data and model prediction over the specimen thickness range for this temperature.

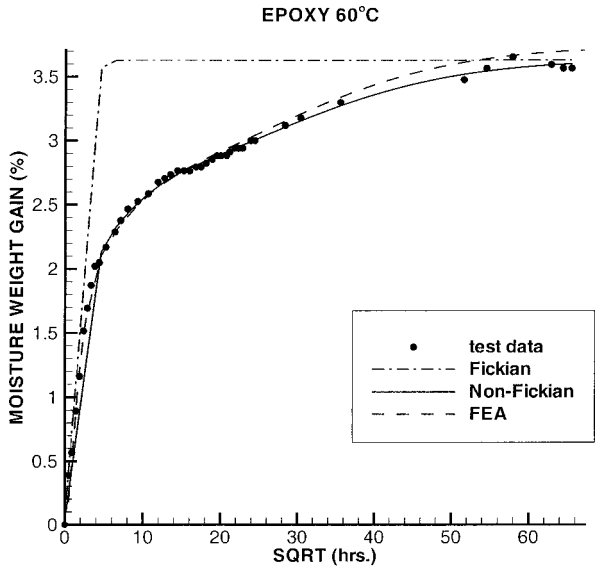


Fig. 4 Percent moisture weight gain versus square root of time for epoxy at 60°C

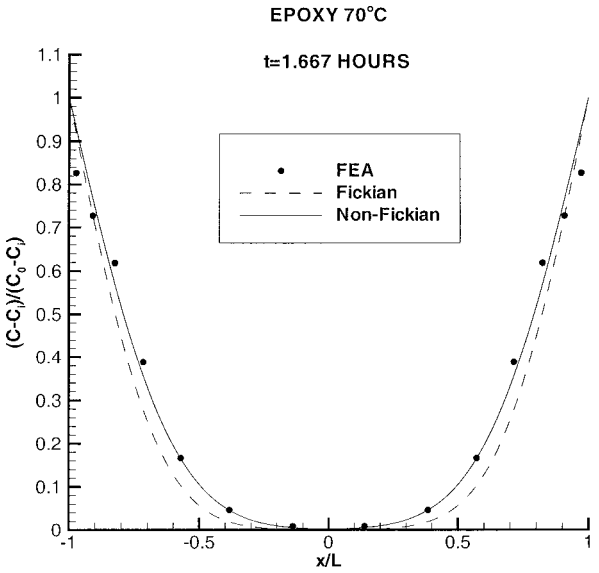


Fig. 6 Through-thickness moisture concentration profile at time=1.66 hours

Good agreement is also observed for moisture uptake in epoxy at 50°C, 60°C, and 70°C as shown in Figs. 3, 4, and 5, respectively. However, the shape of the curve in Fig. 5 seems to indicate that moisture uptake in epoxy becomes less anomalous and more Fickian as the test temperature approaches its glass transition temperature of 125°C (257°F). The knockdown in the T_g of the specimens due to moisture ingress was not measured, but is unlikely to be below 70°C. Therefore, these results seem to corroborate Frisch's [8] original hypothesis regarding the occurrence of time-varying diffusion conditions in a glassy polymer at temperatures below its glass transition.

Figure 6 shows a comparison of through-thickness normalized moisture concentration profiles at $t=1.66$ hours, predicted by the non-Fickian analytical model and the finite element analysis for epoxy at 70°C. Close agreement is observed between the two predictions. A comparison with test data was not possible due to difficulties involved with measuring moisture concentrations within a specimen. A Fickian profile is also included for comparison. The figure shows that at early stages of uptake, the differences between the Fickian and non-Fickian profiles are insignificant. However, significant differences in the two profiles become manifest at later times as shown in Figs. 7 and 8. This anomaly could result in a significant error in stress computation if Fickian concentration profiles are used in determining swelling strains and residual stresses in the epoxy.

Figure 9 shows the variation of the predicted non-Fickian diffusion coefficient with time for epoxy at the four test temperatures. The exponential decay in the diffusivity with time from an initial value toward a final threshold value is analogous to the change in the viscoelastic relaxation modulus of a polymer. Interestingly, the final threshold value of diffusivity for each case is roughly the same, indicating that the influence of test temperature on diffusivity becomes less significant as time progresses and saturation is approached.

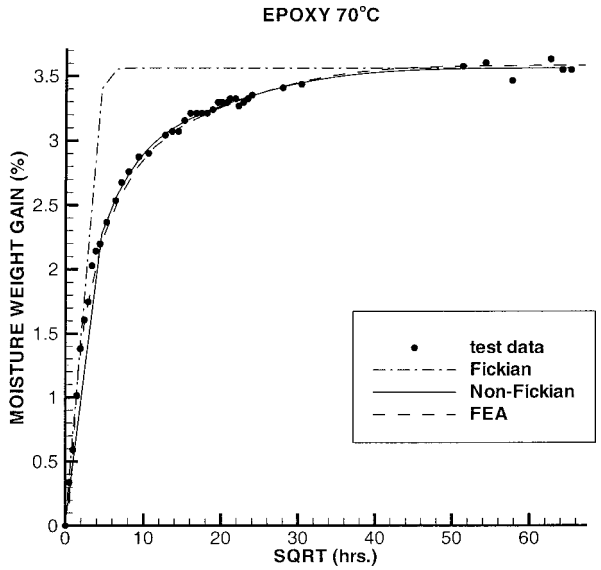


Fig. 5 Percent moisture weight gain versus square root of time for epoxy at 70°C

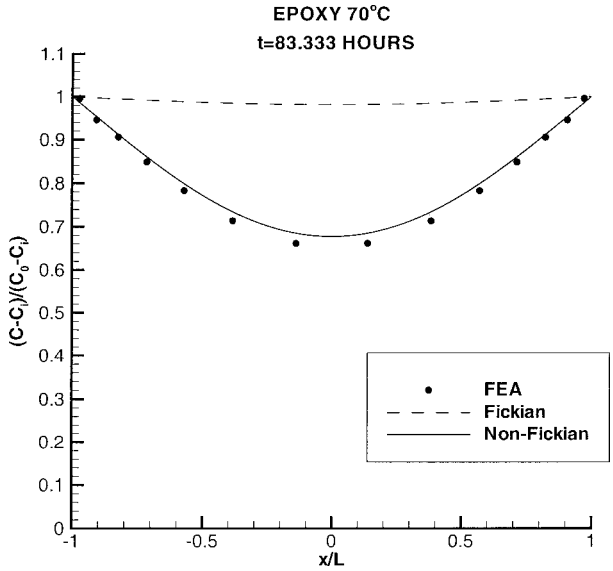


Fig. 7 Through-thickness moisture concentration profile at time=83.3 hours

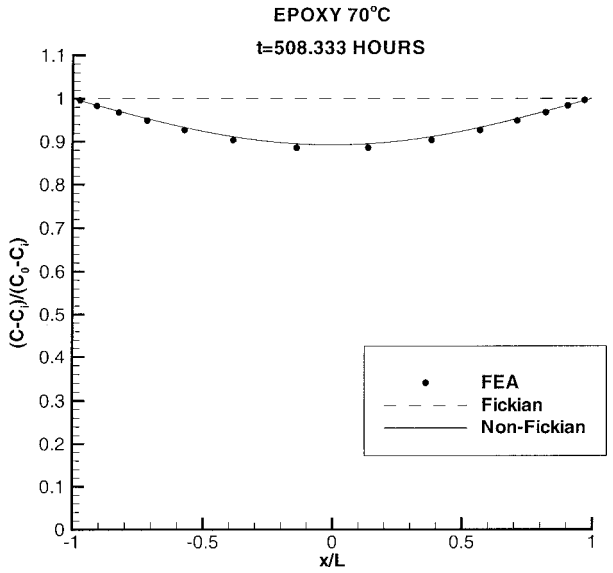


Fig. 8 Through-thickness moisture concentration profile at time=508.3 hours

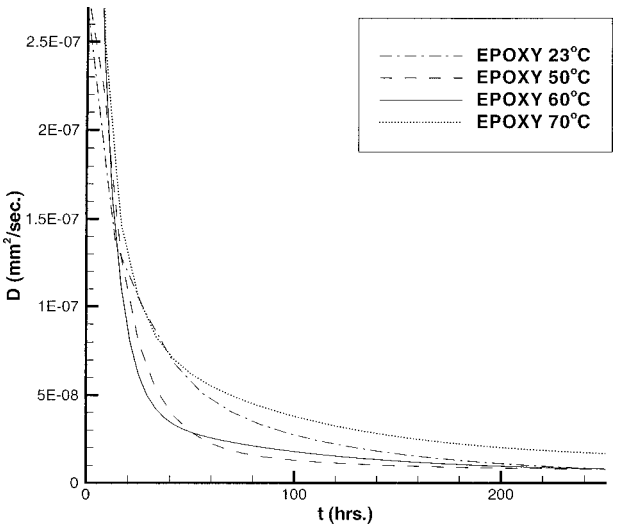


Fig. 9 Variation of diffusion coefficient with time for epoxy at test temperatures

Conclusions

A simple yet robust methodology that enables the non-Fickian diffusion coefficients from weight gain data in a polymer to be extracted without recourse to Carson transforms and iterative procedures was presented. Subsequently, these diffusion coefficients were used to predict moisture concentration profiles through the thickness of the polymer. Because the specimens were fully immersed in a salt solution, the variation in the moisture boundary concentration with time was assumed to be negligible. Salt water weight gain data at different temperatures for an epoxy adhesive was employed to calibrate the model. A finite element procedure that would extend this methodology to more complex shapes and time-varying moisture boundary conditions was also validated. Good agreement between model prediction and test data was observed at all temperatures that were considered. Specimen thickness independence of the modeling parameters was established through comparison with test data at one test temperature. It is envisioned that additional model verifications will be performed using different specimen thickness at other test temperatures.

Acknowledgments

The authors would like to thank Mr. David Warren and Dr. Ray Boeman at Oak Ridge National Laboratories for funding this research and providing material for specimens. Thanks are also due to Dr. Kenneth Morman and Dr. Jessica Schroeder of the Automotive Composites Consortium for technical discussions and providing industrial relevance to the work.

References

[1] Shen, C. H., and Springer, G. S., 1981, ''Effects of Moisture and Temperature on the Tensile Strength of Composite Materials,'' *Environmental Effects on Composite Materials*, G. S. Springer, ed., Technomic, Lancaster, PA, pp. 79–93.
[2] Gurtin, M. E. and Yatomi, C., 1979, ''On a Model for Two Phase Diffusion in Composite Materials,'' J. Compos. Mater., **13**, pp. 126–130.
[3] Carter, H. G. and Kibler, K. G., 1978, ''Langmuir-Type Model for Anomalous Diffusion in Composite Resins,'' J. Compos. Mater., **12**, pp. 118–130.
[4] Shirrell, C. D., Leisler, W. H., and Sandow, F. A., 1979, ''Moisture-Induced Surface Damage in T300/5208 Graphite/Epoxy Laminates,'' *Nondestructive Evaluation and Flaw Criticality for Composite Materials*, ASTM STP 696, R. B. Pipes, ed., American Society for Testing and Materials, Philadelphia, PA, pp. 209–222.
[5] Weitsman, Y., 1991, ''Moisture in Composites: Sorption and Damage,'' *Fatigue of Composite Materials*, K. L. Reifsnider, ed., Elsevier Science Publishers B.V., pp. 385–429.
[6] Frisch, H. L., 1966, ''Irreversible Thermodynamics of Internally Relaxing Systems in the Vicinity of the Glass Transition,'' *Non-Equilibrium Thermodynamics, Variational Techniques, and Stability*, R. J. Dennelly, R. Herman, and I. Prigogine, eds., University of Chicago Press, Chicago, IL, pp. 277–280.
[7] Crank, J., 1975, *The Mathematics of Diffusion*, 2nd Ed., Oxford University Press, Oxford, UK.
[8] Frisch, H. L., 1964, ''Isothermal Diffusion in Systems with Glasslike Transitions,'' J. Chem. Phys., **41**, No. 12, pp. 3379–3683.
[9] Weitsman, Y., 1987, ''Stress Assisted Diffusion in Elastic and Viscoelastic Materials,'' Mech. Phys. Solids, **35**, No. 1, pp. 73–93.
[10] Biot, M. A., 1956, ''Thermoelasticity and Irreversible Thermodynamics,'' J. Appl. Phys., **27**, No. 3, pp. 240–253.
[11] Schapery, R. A., 1969, ''Further Development of a Thermodynamic Constitutive Theory: Stress Formulation,'' A&S Report No. 69-2, Purdue University, West Lafayette, IN.
[12] Cai, L. W., and Weitsman, Y., 1994, ''Non-Fickian Moisture Diffusion in Polymeric Composites,'' J. Compos. Mater., **28**, No. 2, pp. 130–154.
[13] Weitsman, Y., 1990, ''A Continuum Diffusion Model for Viscoelastic Materials,'' J. Phys. Chem., **94**, No. 2, pp. 961–968.
[14] Weitsman, Y., 1987, ''Coupled Damage and Moisture Transport in Fiber-Reinforced, Polymeric Composites,'' Int. J. Solids Struct., **23**, No. 7, pp. 1003–1025.
[15] Roy, S., Lefebvre, D. R., Dillard, D. A., and Reddy, J. N., 1989, ''A Model for the Diffusion of Moisture in Adhesive Joints. Part III: Numerical Simulations,'' J. Adhes., **27**, pp. 41–62.
[16] Sancaktar, E., and Baechtle, D., 1993, ''The Effect of Stress Whitening on Moisture Diffusion in Thermosetting Polymers,'' J. Adhes., **42**, pp. 65–85.
[17] Wong, T., and Broutman, L., 1985, ''Moisture Diffusion in Epoxy Resins Part I: Non-Fickian Sorption Processes,'' Polym. Eng. Sci., **25**, No. 9, pp. 521–528.
[18] Wong, T., and Broutman, L., 1985, ''Water in Epoxy Resins Part II: Diffusion Mechanisms,'' Polym. Eng. Sci., **25**, No. 9, pp. 529–534.
[19] Roy, S., 1999, ''Modeling of Anomalous Diffusion in Polymer Matrix Composites: A Finite Element Approach,'' J. Compos. Mater., **33**, No. 14, pp. 1318–1343.

Stiffening Effects of High-Frequency Excitation: Experiments for an Axially Loaded Beam

J. S. Jensen

D. M. Tcherniak

J. J. Thomsen[1]

Department of Solid Mechanics,
Technical University of Denmark,
DK-2800 Lyngby, Denmark

According to theoretical predictions one can change the effective stiffness or natural frequency of an elastic structure by employing harmonic excitation of very high frequency. Here we examine this effect for a hinged-hinged beam subjected to longitudinal harmonic excitation. A simple analytical expression is presented, that relates the effective natural frequencies of the beam to the intensity of harmonic excitation. Experiments performed with a laboratory beam confirm the general tendency of this prediction, though there are discrepancies that cannot be explained in the framework of the linear Galerkin-discretized beam model. [S0021-8936(00)01302-7]

1 Introduction

According to certain theoretical predictions, one can change the stability of an elastic structure by employing harmonic excitation of very high frequency ([1]). We present a theory for a simple beam structure and test it experimentally.

There are a number of nontrivial effects of fast harmonic excitation. As ''nontrivial'' we consider any response to high-frequency excitation that cannot be described simply as high-frequency vibrations about the corresponding equilibrium without excitation. For example, one can stabilize a pendulum in the inverted position by rapidly shaking its support up and down, as demonstrated experimentally and explained theoretically already by Stephenson [2] and Kapitza [3]. This well-known curiosity of classical mechanics is a nontrivial effect, whereas the accompanying overlay of small but fast vibrations is considered trivial.

Blekhman [4] gives a broad overview as well as in-depth theoretical treatments of many such phenomena. Chelomei [5] and Blekhman [6], in more popular expositions, provide several illustrative examples. Among recently investigated examples we mention vibration-induced sliding and transport of fluids or solids ([7–12]), change of friction properties ([13,14]), and stabilization, stiffening, shift or drift of static equilibriums, and change of linear natural frequencies and nonlinear properties ([15,16,5,17–27]).

This paper deals with experimental validation of one particular nontrivial effect of fast excitation for a particular structure: that of changing natural frequencies for a hinged-hinged beam—or equivalently: change of beam stiffness and buckling load. V. N. Chelomei [5] describes an experiment where such a beam is buckled due to static loading, and then straightened under the action of longitudinal harmonic excitation of high frequency. In this popular exposition there is no quantitative data for the experiment. S. V. Chelomei [16] apparently uses the same beam setup for experiments along with theoretical developments, however, without giving experimental details. Jensen [20] and Tcherniak [24] recently provided more detailed analyses of this beam system paying attention also to nonlinear effects and shear deformation and rotary inertia, though still without quantitative experimental support. The present study was initiated to help remedy a seeming lack of experimental support in this area.

The change of stability due to fast harmonic excitation is, in fact, a consequence of a shift in effective or ''averaged'' stiffness—or equivalently: of linear natural frequencies. In most cases natural frequencies are more easily measured than is stiffness or buckling load, and equally easy to predict. Therefore, to test theoretical predictions of effects of fast excitation, we rely on small-amplitude natural frequencies as an adequate measure.

Specifically we measured the lowest two resonance frequencies of a hinged-hinged beam for different intensities of fast harmonic longitudinal excitation. As will be shown, theory predicts these frequencies to increase with the intensity of the excitation. The experiments reproduce this main tendency, though there are certain discrepancies that cannot readily be explained in the framework of the linear Galerkin-discretized beam model.

Section 2 describes the model system and the theoretical predictions to be tested, whereas Section 3 describes the experimental setup and procedures. Section 4 gives the results, and Section 5 the conclusions.

2 Theory

We derive here a theoretical model for predicting the effective natural frequencies when fast harmonic excitation is added to the beam. The predictions will be tested against experimental results in Section 4.

2.1 Equation of Motion. Figure 1 shows the beam model. We consider a hinged-hinged elastic, slender beam with length l. The beam is assumed to be homogeneous and inextensible with bending stiffness EI and mass per unit length ρA. The boundary conditions are hinged-hinged.

In the mobile end the beam is loaded by an axial force comprising a constant load N and a dynamic component $\Omega P \cos(\Omega t)$, where Ω is the frequency of excitation and P is denoted the *intensity* of excitation.

In the following we analyze *small transverse motion* of the beam. Consequently, axial motions are negligible if the pulsating excitation does not induce axial resonance and thus the inextensibility condition is not violated. The frequency of excitation is in the following assumed to be high, but sufficiently low so that no resonance with very high-frequency axial modes occurs.

[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, Mar. 17, 1999; final revision, Nov. 10, 1999. Associate Technical Editor: J. T. Jenkins. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

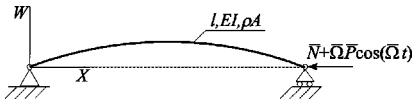


Fig. 1 A slender elastic beam with static plus time harmonic axial loading

A linear beam equation governing small transverse deflections $W(X,t)$ is obtained by considering force and moment balances for an infinitesimal beam element, giving

$$\ddot{W}+c\dot{W}+\eta\pi^{-2}\dot{W}''''+\pi^{-4}W''''+(N+\Omega P\cos(\Omega t))W''=0, \tag{1}$$

where X, W, and t have been rescaled into nondimensional variables,

$$X\rightarrow\frac{X}{l},\quad W\rightarrow\frac{W}{l},\quad t\rightarrow\omega t, \tag{2}$$

with $\omega\equiv(\pi/l)^2\sqrt{EI(\rho A)^{-1}}$ being the first natural frequency for the unloaded beam, the nondimensional parameters are

$$\Omega\equiv\frac{\overline{\Omega}}{\omega},\quad N\equiv\frac{\overline{N}}{\rho Al^2\omega^2},\quad P\equiv\frac{\overline{P}}{\rho Al^2\omega}, \tag{3}$$

and where viscous damping $c\dot{W}$ and internal damping $\eta\pi^{-2}\dot{W}''''$ has been added.

2.2 Spatial Discretization. We use ordinary Galerkin discretization, i.e., the transverse deflection is expanded as

$$W(X,t)=\sum_{i=1}^{N}q_i(t)\varphi_i(X), \tag{4}$$

where $q_i(t)$ are modal amplitudes and $\varphi_i(X)$ mode shapes for the unloaded beam, $\varphi_i(X)=\sin(i\pi X)$. Equation (4) is inserted into (1) and we employ the standard Galerkin procedure, multiplying with the mode shape functions and integrating over the beam length. Thus, we obtain the following set of discretized equations for the modal amplitudes:

$$\ddot{q}_i+(c+i^2(i\pi)^2\eta)\dot{q}_i+(\omega_i^2-\Omega P(i\pi)^2\cos(\Omega t))q_i=0, \tag{5}$$

$$\omega_i\equiv i\sqrt{i^2-\pi^2N}, \tag{6}$$

where ω_i are the natural frequencies, modified by the action of the static axial load. For $N>\pi^{-2}$ we have $\omega_1^2<0$. Hence, the load $N=\pi^{-2}$ represents the buckling load, for which the beam buckles in its fundamental mode.

The equations in (5) are uncoupled due to the boundary conditions, and each equation is recognized as a damped Mathieu equation ([28]). It is well known that the response may be resonant if $\Omega\approx2\omega_i$, i.e., in case of primary parametric resonance. In the following it will be assumed that the oscillation mode under consideration is not resonantly excited.

Full information about the effect of the fast harmonic excitation can be deduced from (5), e.g., by applying any of many integration techniques. We rely instead on an analytical approach which allows us to obtain straightforward predictions for the change in the effective natural frequencies when fast excitation is added.

An expression for the effective natural frequencies is obtained in the following.

2.3 Separation of Fast and Slow Motion. We here apply the method of direct partition of motion (DPM), ([4]), in order to obtain predictions for the effective frequency of slow oscillation modes governed by Eq. (5). The method applies a formal separation between the fast and the slow motion of the individual oscillation modes. This separation allows for obtaining equations governing the *slow* motion of the modes, where the nontrivial effect of the fast excitation is included as additional slowly varying terms.

Two time scales are introduced,

$$T_0=\Omega t,\quad T_1=t, \tag{7}$$

where T_0 is a fast time scale describing the motion at the rate of the excitation frequency, and T_1 is the slow time scale which describes the motion at the rate of the slow natural frequencies.

The motion of each beam mode is split up in a fast and a slow part,

$$q_i(T_0,T_1)=z_i(T_1)+\epsilon\phi_i(T_0,T_1), \tag{8}$$

where the fast part $\epsilon\phi_i$ is assumed to have a zero T_0-average. The small parameter ϵ, is included to emphasize that we restrict ourselves to the case where the fast motion is small compared to the slow oscillations. We choose $\epsilon\equiv\Omega^{-1}$, since Ω already represents a large parameter of the system.

Inserting (8) into (5), and using the notation $D_i^j\equiv\partial^j/\partial T_i^j$, we obtain

$$\epsilon^{-1}D_0^2\phi_i+2D_0D_1\phi_i+D_1^2z_i+\epsilon D_1^2\phi_i+(c+i^2(i\pi)^2\eta)(D_0\phi+D_1z+\epsilon D_1\phi)+(\omega_i^2-\epsilon^{-1}P(i\pi)^2\cos(T_0))(z_i+\epsilon\phi_i)=0. \tag{9}$$

Now (9) is averaged over one period of fast excitation. Using $\langle\ \rangle=1/2\pi\int_0^{2\pi}(\ )dT_0$, we obtain

$$D_1^2z_i+(c+i^2(i\pi)^2\eta)D_1z+\omega_i^2z_i-\langle\epsilon^{-1}P(i\pi)^2\cos(T_0)\epsilon\phi_i\rangle=0, \tag{10}$$

as the equation governing the slow component z_i of the ith mode.

To obtain the fast component $\epsilon\phi_i$, (10) is subtracted from (9) to yield

$$\epsilon^{-1}D_0^2\phi_i+2D_0D_1\phi_i+\epsilon D_1^2\phi_i+(c+i^2(i\pi)^2\eta)(D_0\phi+\epsilon D_1\phi)+\omega_i^2\epsilon\phi_i-\epsilon^{-1}P(i\pi)^2\cos(T_0)(z_i+\epsilon\phi_i)+\langle P(i\pi)^2\cos(T_0)\phi_i\rangle=0. \tag{11}$$

Retaining only dominant terms of order ϵ^{-1} we find

$$D_0^2\phi_i-P(i\pi)^2\cos(T_0)z_i=0, \tag{12}$$

which is readily solved for ϕ_i,

$$\phi_i=-P(i\pi)^2\cos(T_0)z_i. \tag{13}$$

We have obtained a first-order approximation for the fast vibrating components, also known as the *inertial* approximation ([4]). Equation (13) is inserted into (10),

$$D_1^2z_i+(c+i^2(i\pi)^2\eta)D_1z+\omega_i^2z_i+P^2(i\pi)^4z_i\langle\cos^2(T_0)\rangle=0, \tag{14}$$

which since $\langle\cos^2T_0\rangle=1/2$ yields the final equation for the slow components

$$D_1^2z_i+(c+i^2(i\pi)^2\eta)D_1z+\tilde{\omega}_i^2z_i=0, \tag{15}$$

where

$$\tilde{\omega}_i\equiv i\pi\sqrt{(i\pi)^2(\tfrac{1}{2}P^2+\pi^{-4})-N}. \tag{16}$$

2.4 Effective Natural Frequencies in the Presence of Fast Excitation. The modified natural frequency $\tilde{\omega}_i$ takes into account the effect of the fast excitation. As appears, the effective natural frequencies increase with added fast harmonic excitation. For very small values of the intensity P of excitation, (16) shows the effect of the excitation to be negligible, whereas for P large compared to the buckling load π^{-2}, the natural frequencies are predicted to increase asymptotically linearly with P.

3 Experiments

The experiments were conducted in order to determine the effects of fast harmonic excitation on the slow properties of an axially loaded beam. Specifically, we measured the two lowest natural frequencies with varying intensity of excitation applied.

3.1 Experimental Setup. Figure 2(*a*) shows the experimental setup, and Fig. 2(*b*) the corresponding schematic.

Two steel beams with identical material properties were used; their data are given in Table 1. Beam A was used as a test specimen, in order to estimate the beam stiffness. The experiments for beam A were carried out with free-free boundary conditions, approximately realized by suspending the beam in soft rubber bands placed near the nodal points for the first mode of vibration.

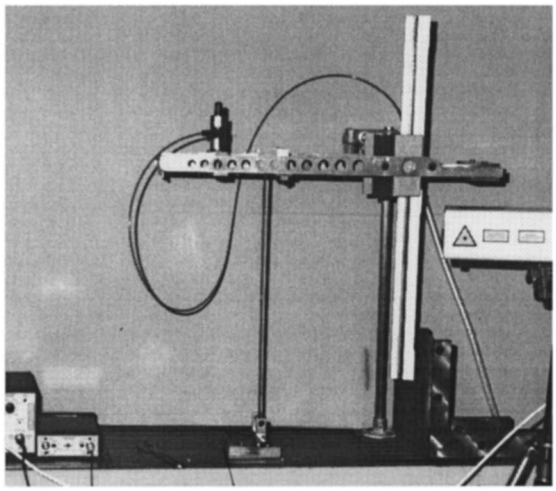
We measured the two lowest natural frequencies for Beam A. The values are given in Table 1. By using the measured frequencies we were able to estimate the bending stiffness by applying the standard formula $EI=\rho A\omega_i^2(l/\lambda_i)^4$, where $\lambda_1\approx 4.7300$ and $\lambda_2\approx 7.8532$.

Beam B was used for the actual experiments. To accomplish pinned-pinned boundary conditions small pins were soldered along the ends of the beam. The pins, about twice the width of the beam, were then fitted into roller bearings built into the top and bottom part of the setup.

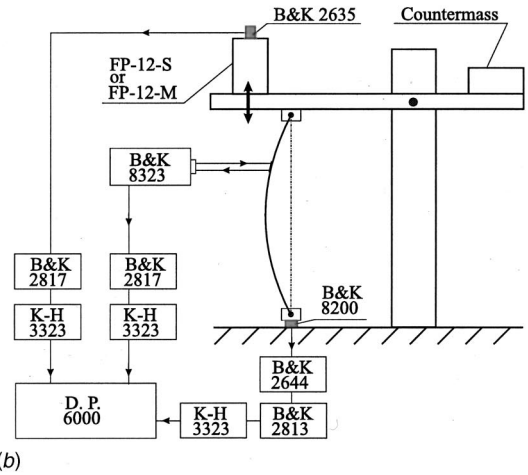
The estimated bending stiffness was used to predict the natural frequencies for the unloaded Beam B (Table 1), as these are not easily measured in the setup. As it turns out, a slight correction of the predicted frequencies is necessary to obtain a closer fit to the experimental results. The need for this adjustment is thought to be a result of imperfections in the experimental setup.

Table 1 Material data and natural frequencies for the unloaded beams

Beam	A	B
Length, l [m]	0.500	0.538
Width [m]	0.028	0.028
Thickness [m]	0.001	0.001
Mass per unit length, ρA [kg/m]	0.215	0.215
Measured ω_1 [s^{-1}]	131.	-
Measured ω_2 [s^{-1}]	361.	-
Estimated stiffness, EI [Nm^2]	0.461	0.461
Predicted ω_1 [s^{-1}]	-	49.9
Predicted ω_2 [s^{-1}]	-	200.
Corrected ω_1 [s^{-1}]	-	48.7
Corrected ω_2 [s^{-1}]	-	195.



(*a*)



(*b*)

Fig. 2 (*a*) **The experimental system,** (*b*) **schematic of the setup**

We applied fast axial excitation to the beam by using pneumatic piston vibration exciters, having internal vibrating masses. The frequency of vibrations is adjusted by regulating the pressure of the air supply. The dynamic force produced by the exciters is specified to increase linearly with the frequency of vibrations. In our setup the exciters proved capable of producing a dynamic force of above 100N in the range 60–100 Hz for the medium sized exciter (FP-12-M), and 75–120 Hz for the small exciter (FP-12-S).

3.2 Experimental Procedure. The natural frequencies of the beams were measured by applying broad-band excitation with an impact hammer (B&K 8202), and picking up the response with a laser velocity transducer (B&K 8323). The signal from the transducer was led through the power supply (B&K 2817), a low pass filter (KROHN-HITE 3323), and to the digital signal analyzer (DATA PRECISION 6000). Here the time series and a frequency spectrum computed using a standard FFT-procedure could be recorded.

The dynamic force transmitted to the beam from the exciter was measured by a force transducer (B&K 8200) built into the lower beam support. The signal was led through a line-drive amplifier (B&K 2644), a line-drive supply (B&K 2813), and through the filter to the signal analyzer. The integrated force measuring system is specified to output 4.09 mV/N. The analyzer was set up to display time series of the dynamic force, the peak-to-peak force amplitude, as well as the frequency. The frequency of excitation was also occasionally checked using the signal from an accelerometer, mounted near the exciters, and led through the charge-amplifier (B&K 2635) to the analyzer.

Four sets of experiments were carried out with each exciter by varying the static load on the beam. The static load was changed by increasing or reducing the mass of the counter load. The magnitude of the static load, N, was estimated from the natural frequencies measured with a passive exciter, i.e., with $\bar{P}=P=0$. For each experimental set, corresponding to a particular exciter and static load, 5–15 measurements were taken with varying exciter air pressure.

3.3 Experimental Observations. We here touch upon some qualitative observations made during the course of the experiments and also point out some of the problems encountered. Quantitative experimental results and comparison with theory will be presented in Section 4.

Theory predicts the natural frequencies to increase with the intensity of the fast excitation. By and large the experiments support this prediction, both for the first and for the second natural frequency. The theory assumes, however, that the excitation frequency is much larger than the natural frequencies considered. In our setup, the second natural frequency is about 30 Hz, and so

with excitation frequencies in the range 60–120 Hz this assumption is hardly fulfilled. Nevertheless, we did note an increase also in the second frequency with added excitation.

Further, when the excitation frequency was close to twice the second natural frequency we observed parametric resonance of the second vibration mode. This caused the frequency of oscillations for the second mode to be locked at half the excitation frequency, and somewhat hindered estimation of the frequency of first mode natural oscillations. Occasionally, we also observed resonant response of the third mode of oscillation, and in this case both of the lowest natural frequencies were rather difficult to measure. Nevertheless, we have included these measurements in our result section whenever we believed that the estimation of the natural frequencies was reasonably accurate.

For a wide range of excitation frequencies the response of the beam was seemingly chaotic, sometimes violently. This was an interesting observation, but consequently no useful measurements concerning natural frequencies could be obtained.

Occasionally we noticed that the measured input force was not simple time-harmonic, but included higher or lower harmonics. We believe this reflects the nonsinusoidal character of the force delivered by the pneumatic vibration exciter.

4 Results

Experimental results are here shown and compared to predictions based on (16). We present the results in nondimensional form. Consequently, we have $\tilde{\omega}_1 = 1$ and $\tilde{\omega}_2 = 4$ for $N = P = 0$.

Of a total of eight experimental sets, the sets #1, #2, #3, and #8 were performed with the small exciter (FP-12-S) with excitation frequencies in the range 60–100 Hz, and sets #4, #5, #6, #7 were performed with the medium-size exciter (FP-12-M) and 75–120 Hz.

4.1 Predictions for $\tilde{\omega}_1$ and $\tilde{\omega}_2$. Figure 3 shows results for the first natural frequency $\tilde{\omega}_1$ as the intensity P of fast harmonic excitation is varied. For each of the two vibration exciters, three different experimental sets are presented with different static loads. In the figure the diagrams are ordered corresponding to increasing static load N.

As appears, the experimental results by and large support the theoretical predictions. The experiments generally produce increasing values of $\tilde{\omega}_1$ as the intensity of excitation increases, as predicted. The order of magnitude of this increase is seen to be fairly well predicted too, even though for most experimental sets the theory predicts too high values of $\tilde{\omega}_1$ when P is large. In set #1 a decrease in $\tilde{\omega}_1$ is noted when P is small. We observed for this excitation frequency a resonant response of the third vibration mode. This indicates some form of modal interaction, which is not accounted for in our linear model.

Figure 4 shows predictions and experimental results for the second natural frequency $\tilde{\omega}_2$. As for Fig. 3 the diagrams are ordered corresponding to increasing static load.

The predictions for $\tilde{\omega}_2$ are also seen to be supported by most experimental sets. However, for sets #6 and #7 the theory predicts far too low frequencies compared to the measured values. The experimental results for these sets were however carried out with the medium size exciter, with the excitation frequency being only slightly above twice the second natural frequency. The discrepancies may originate from interaction between natural oscillations and parametric resonant oscillations. In set #4 a drop in $\tilde{\omega}_2$ is seen when P is small. These experimental values were obtained with excitation frequencies only slightly less than $2\tilde{\omega}_2$, i.e., also close to a parametric resonance.

4.2 Additional Observations. Additional interesting observations were made which are not predicted by the theoretical analysis. Often it appeared as if the fast excitation not only increased the natural frequencies, but also *decreased* the effective damping considerably. This appeared in the experiments as a significant increase in the time of decay for first mode vibrations.

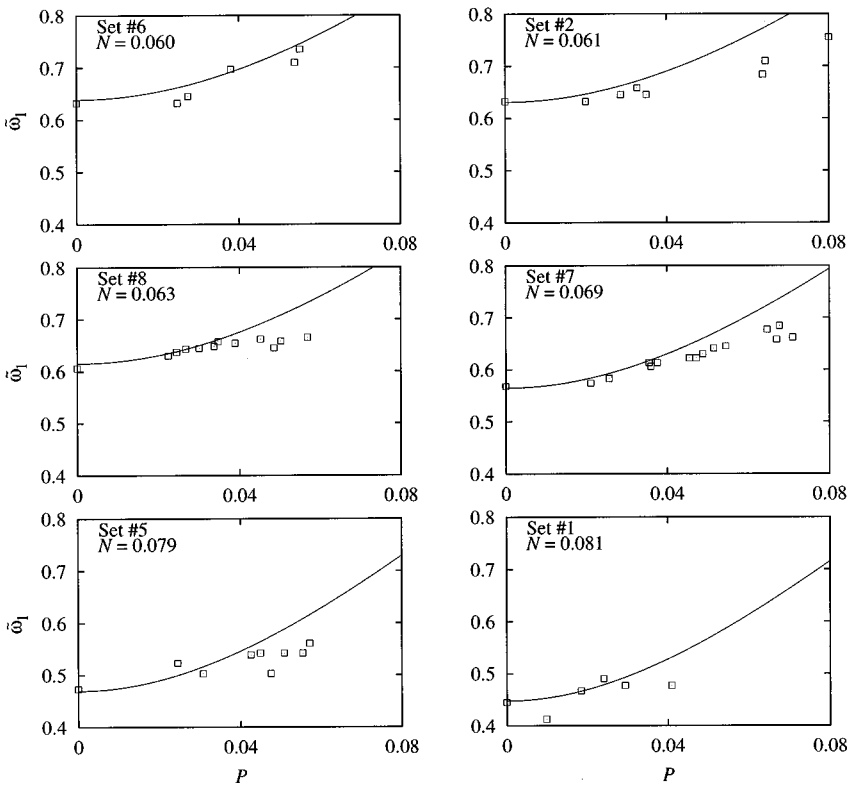


Fig. 3 The first natural frequency $\tilde{\omega}_1$ versus the intensity of excitation P for increasing values of the static axial force N

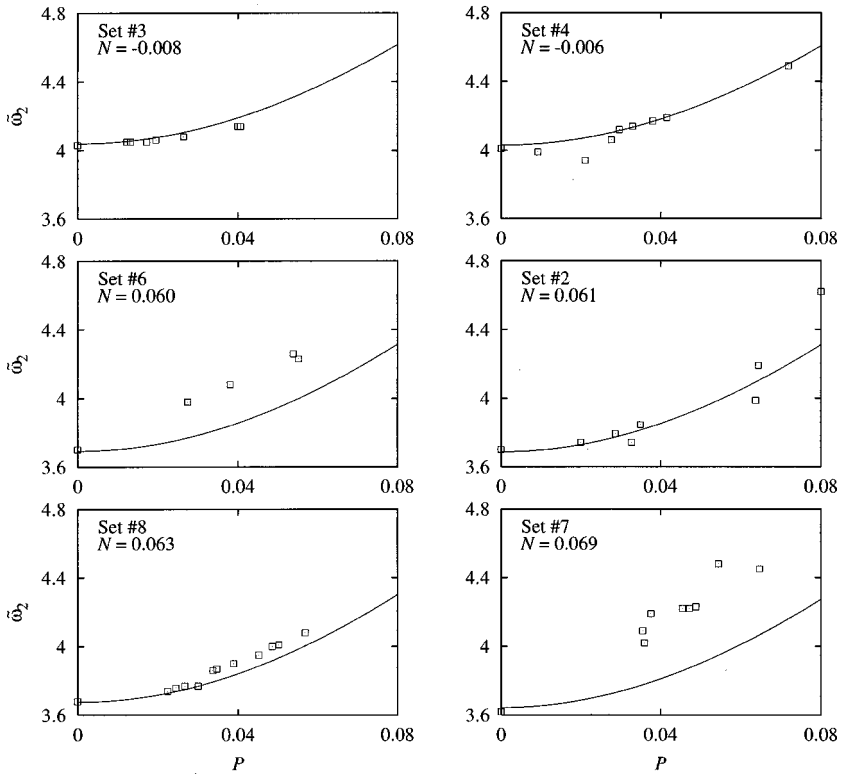


Fig. 4 The second natural frequency $\tilde{\omega}_2$ versus the intensity of excitation *P* for increasing values of the static axial force *N*

More interestingly, it was noted that for excitation of very high intensity, we observed overdamped response of the first mode. If the beam was forced away from its trivial equilibrium position and then released, it would immediately snap back to the trivial position, instead of performing damped oscillations in the first mode.

5 Conclusion

We have attempted to validate a specific non-trivial effect of applying fast harmonic excitation along the undeformed axis of a hinged-hinged beam. This effect appears as an apparent change in natural frequency, or equivalently, in beam stiffness and buckling load.

A simple analytical expression has been presented that relates effective beam natural frequencies to excitation intensity. According to this expression natural frequencies increase with excitation intensity (=force/frequency), almost linearly for higher levels of intensity.

Experiments performed with a laboratory beam showed that the measured natural frequencies (lowest two) generally do increase with excitation intensity. For some measurement sets the functional relationship between frequency and intensity is quite close to the one predicted theoretically. For other sets there are pronounced discrepancies. We believe the discrepancies can be explained by nonlinear interaction (in between beam vibration modes and between the beam and other parts of the setup), by mechanical clearances, and by the frequency of the available excitation source not being sufficiently high.

The experimental setup might be further improved on these factors, so as to match the theoretical beam model more closely. Or the theoretical model could be elaborated to include nonlinearities and the possibility of modal interaction. However, we find that the evidence given in this study is sufficient to conclude that (a) fast harmonic excitation can change the effective natural frequencies of a real physical beam; (b) for low levels of excitation the change is negligible, whereas for higher levels the natural frequencies grows almost linearly with the intensity of excitation.

References

[1] Chelomei, V. N., 1956, ''On the Possibility of Increasing the Stability of Elastic Systems by Using Vibration,'' Dokl. Akad. Nauk, **110**, pp. 345–347 (in Russian).

[2] Stephenson, A., 1908, ''On a New Type of Dynamical Stability,'' Mem. Proc. Manch. Lit. Philos. Soc., **52**, pp. 1–10.

[3] Kapitza, P. L., 1951, ''Dynamic Stability of a Pendulum With an Oscillating Point of Suspension,'' Zh. Eksp. Teor. Fiz. Pis'ma Red., **21**, pp. 588–597 (in Russian).

[4] Blekhman, I. I., 2000, *Vibrational Mechanics*, World Scientific Publishing Company, Singapore.

[5] Chelomei, V. N., 1983, ''Mechanical Paradoxes Caused by Vibrations,'' Sov. Phys. Dokl., **28**, pp. 387–390.

[6] Blekhman, I. I., 2000, ''Forming the Properties of Nonlinear Mechanical Systems by Means of Vibration,'' *Proceedings of the IUTAM/IFToMM Symposium on Synthesis of Nonlinear Dynamical Systems*, Riga, Latvia, 1998, Kluwer, Dordrecht, The Netherlands.

[7] Jensen, J. S., 1996, ''Transport of Continuous Material in Vibrating Pipes,'' *Proceedings of the EUROMECH 2nd European Nonlinear Oscillation Conference*, Vol. 1, Czech Technological University, Prague, Czech Republic, pp. 211–214.

[8] Jensen, J. S., 1997, ''Fluid Transport due to Nonlinear Fluid-Structure Interaction,'' J. Fluids Struct., **11**, pp. 327–344.

[9] Thomsen, J. J., 1996, ''Vibration Induced Sliding of Mass: Non-trivial Effects of Rotatory Inertia,'' *Proceedings of the EUROMECH 2nd European Nonlinear Oscillation Conference*, Vol. 1, Czech Technological University, Prague, Czech Republic, pp. 455–458.

[10] Thomsen, J. J., 1996, ''Vibration Suppression by Using Self-Arranging Mass: Effects of Adding Restoring Force,'' J. Sound Vib., **197**, pp. 403–425.

[11] Thomsen, J. J., 1997, *Vibrations and Stability, Order and Chaos*, McGraw-Hill, London.

[12] Miranda, E. C., and Thomsen, J. J., 1998, ''Vibration Induced Sliding: Theory and Experiments for a Beam With a Spring-Loaded Mass,'' Nonlinear Dyn., **16**, pp. 167–186.

[13] Thomsen, J. J., 1999, ''Using Fast Vibrations to Quench Friction-Induced Oscillations'' J. Sound Vib., , **228**, No. 5, pp. 1079–1102.

[14] Thomsen, J. J., 2000, ''Vibration-Induced Displacement Using High-Frequency Resonators and Friction Layers,'' *Proceedings of the IUTAM/IFToMM Symposium on Synthesis of Nonlinear Dynamical Systems*, Riga, Latvia, 1998, Kluwer, Dordrecht, The Netherlands.

[15] Blekhman, I. I., and Malakhova, O. Z., 1986, ''Quasiequilibrium Positions of the Chelomei Pendulum,'' Sov. Phys. Dokl., **31**, pp. 229–231.
[16] Chelomei, S. V., 1981, ''Dynamic Stability Upon High-Frequency Parametric Excitation,'' Sov. Phys. Dokl., **26**, pp. 390–392.
[17] Hansen, M. H., 2000, ''Non-trivial Effects of High-Frequency Excitation of Spinning Disks,'' J. Sound Vib., accepted for publication.
[18] Jensen, J. S., 1998, ''Non-linear Dynamics of the Follower-Loaded Double Pendulum With Added Support-Excitation,'' J. Sound Vib., **215**, pp. 125–142.
[19] Jensen, J. S., 1999, ''Articulated Pipes Conveying Fluid Pulsating With High Frequency,'' Nonlinear Dyn., **19**, pp. 171–191.
[20] Jensen, J. S., 1999, ''Buckling of an Elastic Beam With Added High-Frequency Excitation,'' Int. J. Non-Linear Mech., **32**, pp. 43–54.
[21] Jensen, J. S., 2000, ''Effects of High-Frequency Bi-directional Support-Excitation of the Follower-Loaded Double Pendulum,'' *Proceedings of the IUTAM/IFToMM Symposium on Synthesis of Nonlinear Dynamical Systems*, Riga, Latvia, 1998, Kluwer, Dordrecht, The Netherlands.
[22] Schmitt, J. M., and Bayly, P. V., 1998, ''Bifurcations in the Mean Angle of a Horizontally Shaken Pendulum: Analysis and Experiment,'' Nonlinear Dyn., **15**, pp. 1–14.
[23] Shapiro, B., and Zinn, B. T., 1997, ''High-Frequency Nonlinear Vibrational Control,'' IEEE Trans. Autom. Control, **42**, pp. 83–89.
[24] Tcherniak, D. M., 1999, ''The Influence of Fast Excitation on a Continuous System'' J. Sound Vib., , **227**, No. 2, pp. 343–360.
[25] Tcherniak, D. M., 2000, ''Using Fast Vibration to Change the Nonlinear Properties of Mechanical Systems,'' *Proceedings of the IUTAM/IFToMM Symposium on Synthesis of Nonlinear Dynamical Systems*, Riga, Latvia, 1998, Kluwer, Dordrecht, The Netherlands.
[26] Tcherniak, D. M., and Thomsen, J. J., 1998, ''Slow Effects of Fast Harmonic Excitation for Elastic Structures,'' Nonlinear Dyn., **17**, pp. 227–246.
[27] Weibel, S., Kaper, T. J., and Baillieul, J., 1997, ''Global Dynamics of a Rapidly Forced Cart and Pendulum,'' Nonlinear Dyn., **13**, pp. 131–170.
[28] Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillations*, John Wiley and Sons, New York.

A Variational Boundary Integral Method for the Analysis of Three-Dimensional Cracks of Arbitrary Geometry in Anisotropic Elastic Solids

G. Xu
Department of Mechanical Engineering,
University of California,
Riverside, CA 92521
Mem. ASME

A variational boundary integral method is developed for the analysis of three-dimensional cracks of arbitrary geometry in general anisotropic elastic solids. The crack is modeled as a continuous distribution of dislocation loops. The elastic energy of the solid is obtained from the known expression of the interaction energy of a pair of dislocation loops. The crack-opening displacements, which are related to the geometry of loops and their Burgers vectors, are then determined by minimizing the corresponding potential energy of the solid. In contrast to previous methods, this approach results in the symmetric system of equations with milder singularities of the type 1/R, which facilitate their numerical treatment. By employing six-noded triangular elements and displacing midside nodes to quarter-point positions, the opening profile near the front is endowed with the accurate asymptotic behavior. This enables the direct computation of stress intensity factors from the opening displacements. The performance of the method is assessed by the analysis of an elliptical crack in the transversely isotropic solid. It also illustrates that the conventional average schemes of elastic constants furnish quite inaccurate results when the material is significantly anisotropic. [S0021-8936(00)02702-1]

1 Introduction

Three-dimensional cracks have been studied extensively because of their importance in many engineering and scientific applications. To name a few, these applications, ranging from atomic to geological scale, include the analysis of dislocation nucleation from the crack tip and brittle to ductile transition in cleavage fracture [1], toughening mechanisms of particle and fiber-reinforced composites [2–4], hydraulic fracturing simulation [5], and earthquake modeling [6,7]. In most of the previous analyses of the three-dimensional crack, the problems are reduced to the solutions of integral equations defined on the faces of the crack. Most of these methods are restricted to planar crack problems ([8–10,2–4]). Some of these methods use the Somigliana identity as the means of effecting the reduction to the boundary, which results in highly singular integral equations of difficult numerical treatment ([11,12]). To overcome these geometric restrictions and numerical difficulties, Xu and Ortiz [13] have developed a variational boundary integral method for the analysis of cracks of arbitrary three-dimensional geometry in isotropic elastic solids. By representing the opening displacement of the crack as a continuous distribution of dislocation loops, they obtained the elastic strain energy of the deformed crack from the known expression of the interaction energy of a pair of dislocation loops. The geometry of the loops and their Burgers vectors are readily related to the crack-opening displacements, which are then determined by minimizing the total potential energy of the solid. Because of the variational basis of the method, the resulting system of equations is symmetric. In addition, the distributed dislocation representation gives rise to the integral equations with milder singularity of the type $1/R$. These features facilitate the numerical treatment of the integral equations. In their numerical analysis, they employed six-noded triangular elements and displaced midside nodes to quarter-point positions to ensure the accurate asymptotic behavior of the opening profile near the front. This enables the direct accurate computation of stress intensity factors from the opening displacements.

In this paper, we extend the integral equation developed by Xu and Ortiz [13] to the case that three-dimensional cracks with arbitrary geometry are in general anisotropic elastic solids. Using the same technique described in their work, this can be achieved by using the known expression for the interaction energy of a pair of dislocation loops in anisotropic elastic solids ([14]). Since the similar approach is implemented, the integral equation shares the same characteristic of the one for the isotropic case.

Modeling of cracks with dislocation segments has received considerable attention in the past (e.g., [15,16,5]). The technique implemented in this paper, although differing in using dislocation loops which appear to be more direct, is an extension of this approach to three dimensions. It is noteworthy that this technique can be applied directly to any type of the surface with displacement discontinuity. This proves to be of particular advantage in the analysis of dislocation nucleation from the crack tip because it enables the uniform numerical treatment of both the crack and the dislocation profile on the slip plane ([16,1]).

2 Integral Equation Formulation

In this section, we follow the approach developed by Xu and Ortiz [13] to obtain the integral equation for an arbitrary three-dimensional crack in a general anisotropic elastic solid. The elastic constants for the solid are denoted by c_{ijkl}. Due to the symmetry and existence of elastic strain energy, there are at most 21 distinct elastic constants. Let the surface of the crack be S and crack front be C. Define $\mathbf{u}(\mathbf{x})$, $\mathbf{x} \in S$, the displacement jump across S, with components u_i relative to a Cartesian basis $\mathbf{e}_i$. Our first

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 22, 1999; final revision, Nov. 26, 1999. Associate Technical Editor: M. Ortiz. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

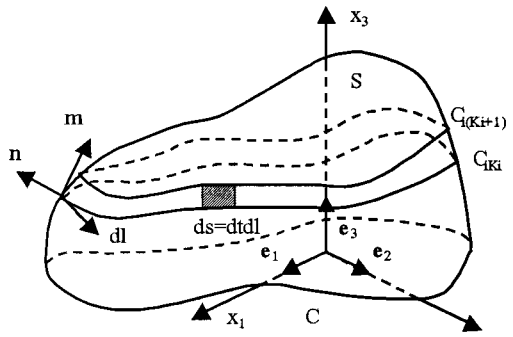


Fig. 1 Distributed dislocation loop representation of opening displacement field

aim is to compute the elastic strain energy of the cracked solids. This is accomplished by representing the opening displacements as a continuous distribution of dislocation loops, and subsequently using known expressions for dislocation loop interaction energies.

The dislocation distribution equivalent to $\mathbf{u}(\mathbf{x})$ is determined as follows. For $i=1,2,3$, let C_{iK_i} be the level contours on S corresponding to values of $u_i=K_i\Delta u_i$ (no sum), $K_i=0,1,2,\ldots N_i$. This defines three sets of level contours inscribed on the crack surface (Fig. 1). Since $u_i=0$ at the crack front C, it follows that $C_{i0}=C$. We begin by replacing the original opening displacement field by one consisting of discrete steps of magnitude Δu_i across the level contours C_{iK_i}. Evidently, the stepwise opening displacement field so defined approaches $\mathbf{u}(\mathbf{x})$ as $\Delta u_i\to 0$. Each contour C_{iK_i} may be viewed as carrying a dislocation of Burgers vector

$$\mathbf{b}_i=\Delta u_i\mathbf{e}_i \quad (\text{no sum on } i). \tag{1}$$

By definition, the Burgers vector for the loop C_{iK_i} points in the ith coordinate direction and has a constant magnitude equal to Δu_i. The elastic interaction energy between two dislocation loops C_1 and C_2 of Burgers vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ in an infinite anisotropic elastic solid is ([14]) (see Fig. 2)

$$W_{12}=\frac{1}{8\pi^2}\oint_{C_1}\oint_{C_2}\frac{1}{R}\int_0^{2\pi}\mathbf{b}_1\cdot[(d\mathbf{l}_1\times\mathbf{z},d\mathbf{l}_2\times\mathbf{z})-(d\mathbf{l}_1\times\mathbf{z},\mathbf{z})\cdot(\mathbf{z},\mathbf{z})^{-1}\cdot(\mathbf{z},d\mathbf{l}_2\times\mathbf{z})]\cdot\mathbf{b}_2 d\phi \tag{2}$$

where $d\mathbf{l}_1$ and $d\mathbf{l}_2$ are vectors of infinitesimal length, tangent to C_1 and C_2, respectively, R is the length of the relative position vector between C_1 and C_2, $\mathbf{z}$ and $\mathbf{z}_0$ are unit vectors perpendicular to $\mathbf{R}$, and ϕ is the angle between $\mathbf{z}$ and arbitrary chosen reference $\mathbf{z}_0$. Let $\mathbf{a}$ and $\mathbf{b}$ be two vectors, $(\mathbf{a},\mathbf{b})$ denotes a second-rank tensor with elements

$$(\mathbf{a},\mathbf{b})_{jk}=a_i c_{ijkl} b_l. \tag{3}$$

Invoking the principle of superposition, the elastic strain energy of the cracked solid follows as

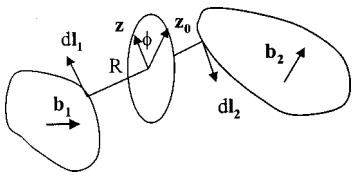


Fig. 2 Two interacting dislocation loops

$$W[\mathbf{u}]=\frac{1}{2}\lim_{\Delta u_1,\Delta u_2,\Delta u_3\to 0}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{K_i=0}^{N_i}\sum_{K_j=0}^{N_j}\left[\frac{1}{8\pi^2}\oint_{C_{iK_i}}\oint_{C_{jK_j}}\frac{1}{R}\times\int_0^{2\pi}\mathbf{b}_i\cdot[(d\mathbf{l}_{iK_i}\times\mathbf{z},d\mathbf{l}_{jK_j}\times\mathbf{z})-(d\mathbf{l}_{iK_i}\times\mathbf{z},\mathbf{z})\cdot(\mathbf{z},\mathbf{z})^{-1}\cdot(d\mathbf{l}_{jK_j}\times\mathbf{z},\mathbf{z})]\cdot\mathbf{b}_j d\phi\right]. \tag{4}$$

The factor 1/2 in this expression compensates for the fact that the double sum accounts for the interaction energy between each pair of loops twice. The computation of the elastic strain energy is completed by passing to the limit of $\Delta u_i\to 0$. To this end, let $\mathbf{m}$ be a unit surface vector (i. e., a unit vector tangent to S) normal to C_{iK_i}. Since C_{iK_i} is a level contour of u_i, it follows that

$$\mathbf{m}=\nabla u_i/|\nabla u_i|. \tag{5}$$

Let t measure the distance from the loop along $\mathbf{m}$, and Δt be the distance along $\mathbf{m}$ between the consecutive loops C_{iK_i} and $C_{i(K_i+1)}$ shown in Fig. 1. The derivative of u_i with respect to t follows as

$$\frac{du_i}{dt}=\nabla u_i\cdot\mathbf{m}=|\nabla u_i| \tag{6}$$

where use has been made of identity (5). Hence, for small Δu_i, one has

$$\Delta u_i\approx|\nabla u_i|\Delta t \tag{7}$$

and the Burgers vectors (1) become

$$\mathbf{b}_i\approx|\nabla u_i|\Delta t\mathbf{e}_i. \tag{8}$$

Furthermore, the line element along a contour may be expressed as

$$d\mathbf{l}=ds\mathbf{n}\times\mathbf{m}=ds\mathbf{n}\times\nabla u_i/|\nabla u_i| \tag{9}$$

where ds is the element of are length, and $\mathbf{n}$ is the unit normal vector to the surface. Substituting (8) and (9) into (4), and noting that in the limit of $\Delta u_i\to 0$, $\Delta t ds$ may be identified with the element of surface area dS, one finds

$$W[\mathbf{u}]=\frac{1}{16\pi^2}\int_S\int_S\frac{1}{R}\int_0^{2\pi}\mathbf{e}_i\cdot[(\mathbf{n}\times\nabla u_i)_1\times\mathbf{z},(\mathbf{n}\times\nabla u_j)_2\times\mathbf{z})-((\mathbf{n}\times\nabla u_i)_1\times\mathbf{z},\mathbf{z})\cdot(\mathbf{z},\mathbf{z})^{-1}\cdot(\mathbf{z},(\mathbf{n}\times\nabla u_j)_2\times\mathbf{z})]\cdot\mathbf{e}_j d\phi dS_1 dS_2 \tag{10}$$

where $(\cdot)_1$ and $(\cdot)_2$ denote two different points on S, and R is then the distance between these two points. Equation (10) is the sought expression for the strain energy of the cracked solid. It takes the form of a double integral over the crack surface S. An advantage of the present formulation which is immediately apparent from (10) is that the kernels of all integrals have milder singularities of the type $1/R$. Their accurate numerical treatment were given by Xu and Ortiz [13].

The potential energy of the solid follows as the sum of the strain energy (10) and the potential energy of the applied loads. Without loss of generality, we shall confine our attention to the case in which traction $\mathbf{t}$ is applied directly to the faces of the crack. Then, the potential energy of the solid is

$$\Pi[\mathbf{u}]=W[\mathbf{u}]-\int_S\mathbf{t}\cdot\mathbf{u}dS. \tag{11}$$

For given traction, the opening displacements of the crack can be obtained by minimizing the potential energy Π. The distribution of stress intensity factors over the crack front C can then be extracted from the known opening displacements. As is evident from (11), Π defines a positive definite quadratic form in $\mathbf{u}$. Consequently, Rayleigh-Ritz methods of approximation based on the

constrained minimization of Π over a finite dimensional interpolation space result in symmetric systems of equations. One such method based on a finite element discretization of the crack surface refers to Xu and Ortiz [13] in which the treatment of singularity along the crack front, the singular integration, the calculation of stress intensity factors, and the treatment of the periodic semi-infinite cracks are addressed in detail. As a final remark, it is not necessary to recast (10) in terms of the surface coordinates as previously indicated by Xu and Ortiz [13]. The use of flat six-noded triangular elements to represent the curved crack surface enables the simple linear transform of the local and global variables between two Cartesian coordinates.

3 Numerical Examples for Elliptical Cracks in Transversely Isotropic Solids

An elliptical crack in the unbounded transversely isotropic solid is selected to demonstrate the performance of the method. This configuration is selected because the numerical results can be compared to the exact analytical solutions. The elliptical crack is assumed to be situated in the plane perpendicular to the axis of elastic symmetry of a transversely isotropic elastic solid. With reference to a system of rectangular Cartesian coordinates (x,y,z), the crack bounded by $x^2/a^2+y^2/b^2=1$ is placed in the xy-plane such that the origin coincides with the center of the ellipse having major and minor semi-axes a and b in the ratio $k'=b/a$. Points on the crack front can be parameterized by the polar angle θ as $x_1=a\cos\theta$ and $x_2=b\sin\theta$ (see Fig. 3). The crack is assumed to be under mode I remote loading σ. The elastic constants in contracted notation for two transversely isotropic solid zinc and barium titanate are listed in Table 1 ([18]).

The solutions of the above elliptical crack in transversely isotropic materials are equivalent to the solutions of the same crack in isotropic materials if the appropriate isotropic elastic constants are selected ([18]). The nonzero opening displacement component u_3 is given by

$$u_3=\frac{2A\sigma b}{E(k)}\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}} \tag{12}$$

where

$$A=\Re\left\{\frac{c_{11}(c_{13}+c_{44})(n_1^{1/2}+n_2^{1/2})}{c_{44}(c_{13}+c_{11}n_1)(c_{13}+c_{11}n_1)}\right\} \tag{13}$$

and n_1 and n_2 are the roots of the quadratic in n

$$c_{11}c_{44}n^2+[c_{13}(c_{13}+2c_{44})-c_{11}c_{33}]n+c_{33}c_{44}=0. \tag{14}$$

Table 1 Elastic constants of some transversely isotropic materials (10^6 psi)

Material	c_{11}	c_{12}	c_{13}	c_{33}	c_{44}
Zinc	16.09	3.35	5.01	6.10	3.83
Barium titanate	16.80	7.80	7.10	18.90	5.64

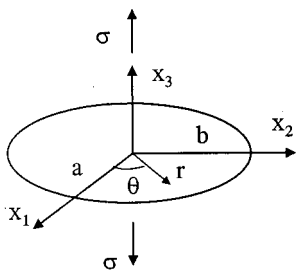


Fig. 3 An elliptical crack under mode I loading

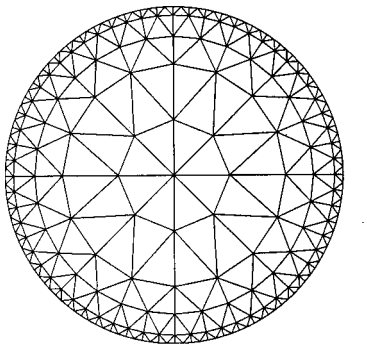


Fig. 4 Example of mesh used in the analysis

For an isotropic material with the shear modulus μ and Poisson's ratio ν,

$$A=\frac{1-\nu}{\mu}. \tag{15}$$

The mode I stress intensity factor along the crack front is given by

$$K_I=\frac{\sigma\sqrt{\pi b}}{E(k)}(\sin^2\theta+k'^2\cos^2\theta)^{1/4} \tag{16}$$

where $k^2=1-k'^2$ and

$$K(k)=\int_0^{\pi/2}\frac{d\phi}{\sqrt{1-k^2\sin^2\phi}},\quad E(k)=\int_0^{\pi/2}\sqrt{1-k^2\sin^2\phi}\,d\phi \tag{17}$$

are the elliptic integrals of the first and second kinds, respectively.

The mesh of a circular crack used in the analysis is depicted in Fig. 4. The mesh for the elliptical crack is transformed from the mesh of the circular crack. The circumferential mesh lines divide the range of the function $\sqrt{a^2-r^2}$ into equal intervals. An additional inner circle is added in the middle for the aesthetic mesh. Then, the innermost circumferential mesh line is divided into eight segments, the next circumferential mesh line into 16 segments, and so on. This defines the positions of the corner nodes of the elements. The mesh is then constructed by Delaunay triangulation, and the midside nodes are added subsequently.

The computed opening profiles for the elliptical cracks in zinc and barium titanate are shown in Figs. 5 and 6 against the exact solutions. The openings are normalized by $u_0=4A\sigma b/\pi$, which is the center opening of the penny-shaped crack with radius b. Several other approximated solutions are also plotted in the figures. The Voigt solution represents the solution of the crack in an isotropic solid with the shear modulus μ and Lamé constant λ obtained by Voigt averages ([19]), i.e.,

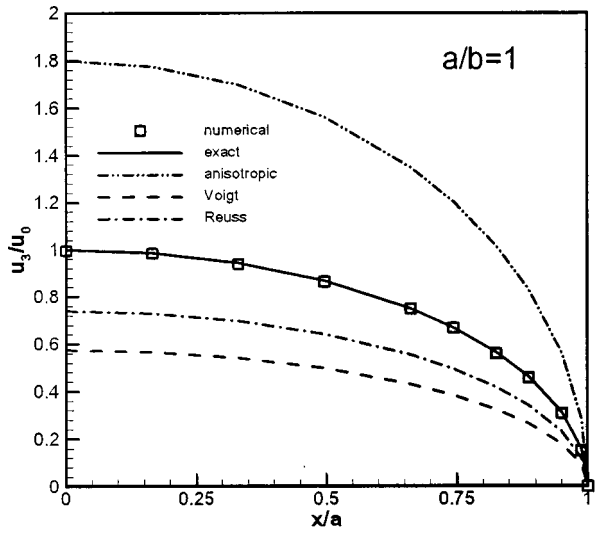
$$\mu=\frac{1}{30}(3c_{ijij}-c_{iijj}),\quad \lambda=\frac{1}{30}(2c_{iijj}-c_{ijij}). \tag{18}$$

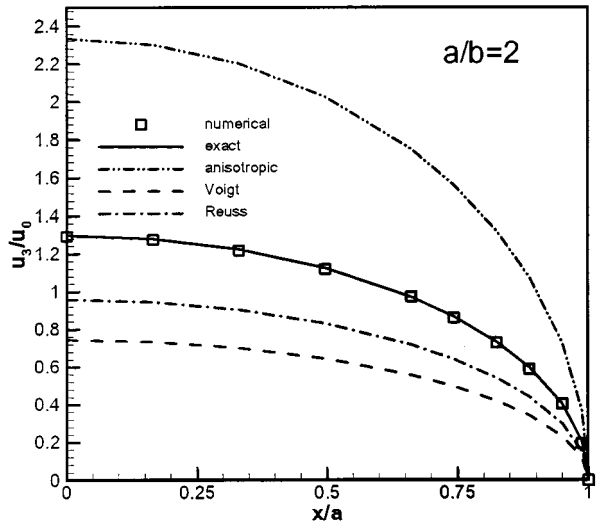
The Reuss solution represents the solution of the crack in an isotropic solid with the Young's modulus E and Poisson's ratio ν obtained by Reuss averages ([19]), i.e.,

$$\frac{1}{E}=\frac{1}{15}(2s_{ijij}+s_{iijj}),\quad \frac{\nu}{E}=\frac{1}{15}(s_{ijij}-2s_{iijj}), \tag{19}$$

where s_{ijkl} are the compliance constants. The anisotropic solution represents the solution of the crack in an isotropic solid with the Young's modulus E and Poisson's ratio ν being approximated as

$$\frac{1}{E}=s_{3333},\quad \frac{\nu}{E}=-s_{1133}. \tag{20}$$





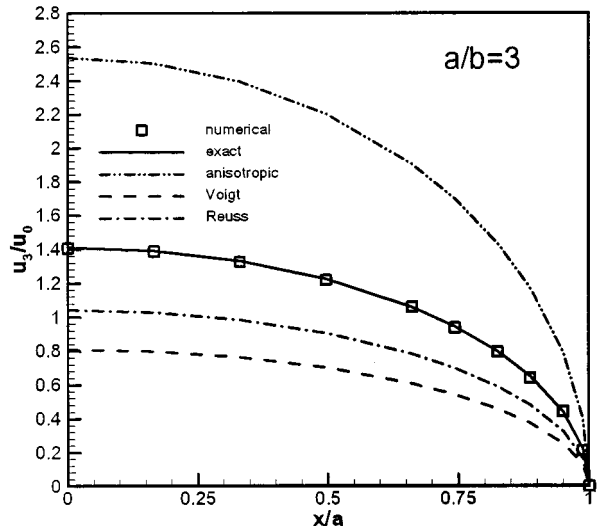
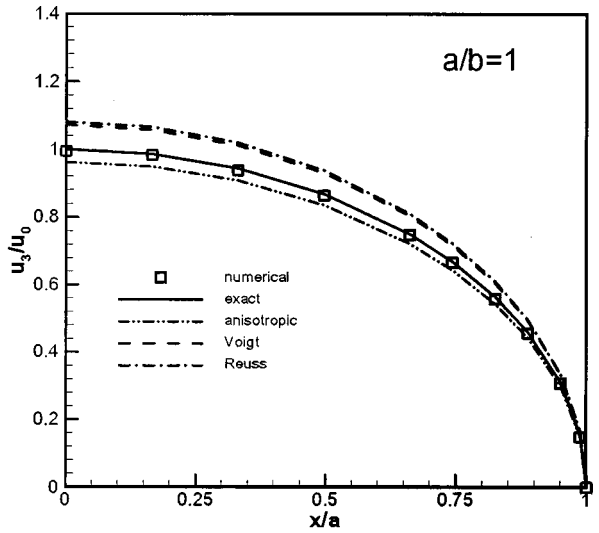
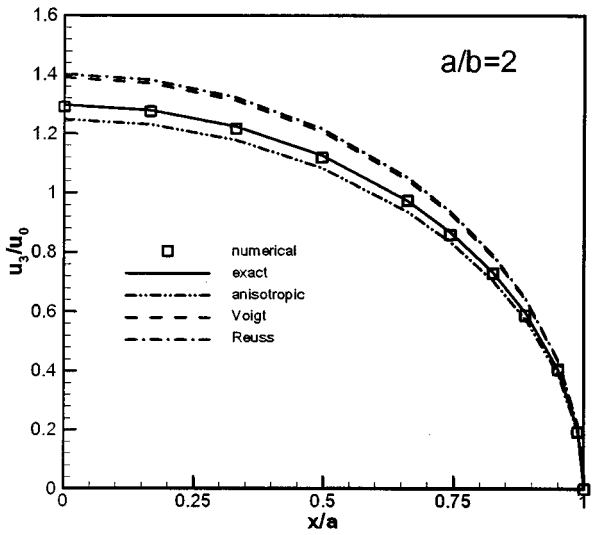


Fig. 5 Opening displacements along *x*-axes of elliptical cracks in zinc





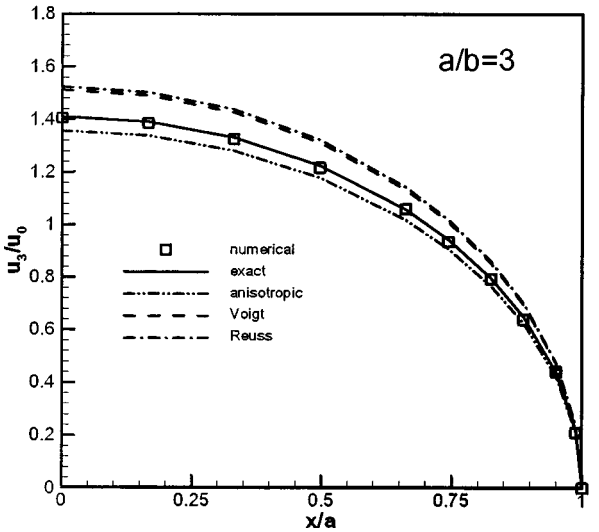


Fig. 6 Opening displacements along *x*-axes of elliptical cracks in barium titanate

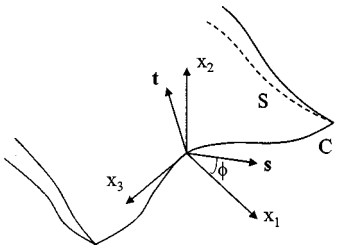


Fig. 7 Local reference frame for the calculation of stress intensity factors in anisotropic materials

These approximate solutions are commonly used in many engineering applications where general anisotropic materials are averaged as isotropic materials. As shown in the figures, these solutions are quite inaccurate when the material is significantly anisotropic such as the material zinc. The large error in the displacement particularly complicates the application such as hydraulic fracturing simulation where accurate computation of fracture openings is critical for predicting the fluid flow between crack surfaces ([5]).

A primary interest in the analysis of cracks in elastic solids is the distribution of stress intensity factors along the crack front. The accurate calculation of these quantities is of critical importance for predicting the path of a growing crack. For a three-dimensional crack of arbitrary geometry, the stress intensity factors can be defined with reference to a local Cartesian frame such that the x_1-x_3 plane is tangent to the crack surface, with the x_3-axis tangent to the crack front, as shown in Fig. 7.

The asymptotic solutions in anisotropic solids used here are obtained by taking the limit of the solution of a slit crack by Barnett and Asaro [16]. With the three stress intensity factors K_i $(i=1,2,3)$ defined as

$$\sigma_{i2}|_{x_1\to 0,x_2=0}=\frac{K_i}{\sqrt{2\pi x}}+\text{nonsingular terms}, \tag{21}$$

the opening displacements u_i are given by

$$u_i=4B_{ij}^{-1}K_j\sqrt{\frac{-x_1}{2\pi}}. \tag{22}$$

Here **B** is a 3×3 matrix that depends only on the direction of the crack front. For isotropic solids, **B** is in pure diagonal form with

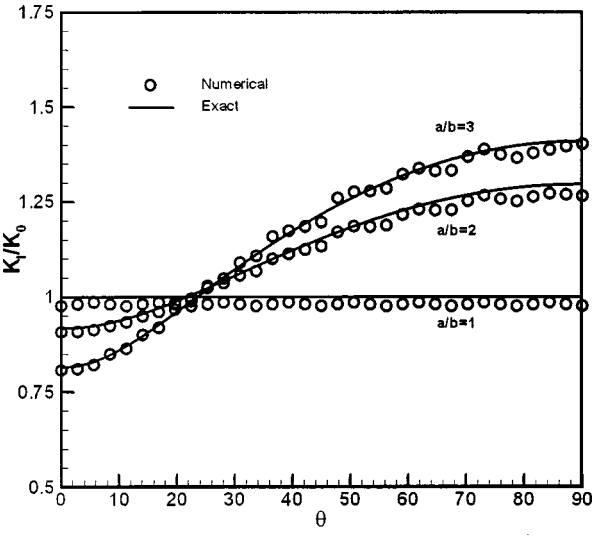


Fig. 8 Variation of K_I along the elliptical crack in zinc

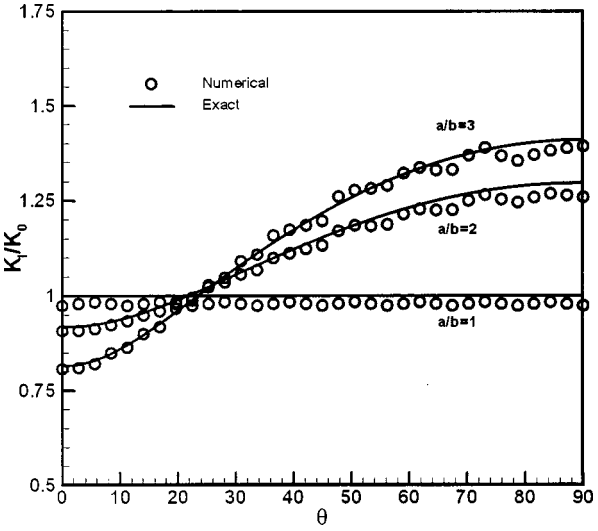


Fig. 9 Variation of K_I along the elliptical crack in barium titanate

$$B_{11}=B_{22}=\frac{\mu}{1-\nu},\quad B_{33}=\mu. \tag{23}$$

For anisotropic solids, **B** can be readily solved through the integral

$$\mathbf{B}=-\frac{1}{2\pi}\int_0^{2\pi}[(s,t)(t,t)^{-1}(t,s)-(s,s)]d\phi \tag{24}$$

where **s** and **t** are unit vectors normal to the crack front and each other as shown in Fig. 7. **B** is integrated by rotating **s** and **t** simultaneously by 2π.

Figures 8 and 9 compare the analytical and computed stress intensity factors for the cracks of aspect ratios $a/b=1$, 2, and 3. Stress intensity factors are normalized by $K_0=(2/\pi)\sqrt{\pi b}\sigma$, which is the stress intensity factor for the penny-shaped crack with radius b. In general, the calculation results in accurate stress intensity factors for predicting the crack growth. Somewhat better results can be obtained when the regular (such as isosceles and equilateral) triangular elements are placed along the crack front ([13]).

4 Summary and Conclusion

A variational boundary integral method is developed for the analysis of three-dimensional cracks of arbitrary geometry in general anisotropic elastic solids. Using the technique developed in the earlier work of Xu and Ortiz [13], the deformed crack is modeled as a continuous distribution of dislocation loops. The elastic strain energy of the deformed crack is obtained from the interaction energy of a pair of dislocation loops. The geometry of the loops and their Burgers vectors are readily related to the crack-opening displacements, which are then determined by minimizing the potential energy of the solid. Because of the variational basis of the method, the resulting system of equations is symmetric. In addition, the distributed dislocation loop representation gives rise to the integral equations with milder singularity of the type $1/R$. These features facilitate the numerical treatment of the integral equation comparing to previous numerical methods. By employing six-noded triangular elements and displacing midside nodes to quarter-point positions, the opening profile near the front is endowed with the accurate asymptotic behavior. This enables the direct accurate computation of stress intensity factors from the opening displacements. The selected numerical examples of elliptical cracks in anisotropic materials demonstrate that accurate opening displacements and stress intensity factors can be obtained

with practical meshes. The results also illustrate that several conventional average schemes of elastic constants result in quite inaccurate opening displacements in the case when the material is significantly anisotropic.

References

[1] Xu, G., Argon, A. S., and Ortiz, M., 1997, ''Critical Configurations for Dislocation From Crack Tips,'' Philos. Mag. A, **75**, p. 341.

[2] Fares, N., 1989, ''Crack Fronts Trapped by Arrays of Obstacles: Numerical Solutions Based on Surface Integral Representation,'' ASME J. Appl. Mech., **56**, p. 837.

[3] Gao, H., and Rice, J. R., 1989, ''A First Order Perturbation Analysis on Crack-Trapping by Arrays of Obstacles,'' ASME J. Appl. Mech., **56**, p. 828.

[4] Bower, A. F., and Ortiz, M., 1991, ''A Three-Dimensional Analysis of Crack Trapping and Bridging by Tough Particles,'' J. Mech. Phys. Solids, **39**, No. 6, p. 815.

[5] Clifton, R. J., 1989, ''Three-Dimensional Fracture-Propagation Models,'' J. L. Gidley, ed., *Hydraulic Fracturing*, SPE Monograph.

[6] Rice, J. R., 1993, ''Spatio-Temporal Complexity of Slip on a Fault,'' J. Geophys. Res., **98**, pp. 9885–9907.

[7] Ben-Zion, Y., and Rice, J. R., 1995, ''Slip Patterns and Earthquake Populations Along Different Classes of Faults in Elastic Solids,'' J. Geophys. Res., **100**, p. 12959.

[8] Bui, H. D., 1977, ''An Integral Equation Method for Solving the Problem of a Plane Crack of Arbitrary Shape,'' J. Mech. Phys. Solids, **25**, p. 29.

[9] Weaver, J., 1977, ''Three-Dimensional Crack Analysis,'' Int. J. Solids Struct., **13**, p. 321.

[10] Cruse, T. A., 1988, *Boundary Element Analysis in Computational Fracture Mechanics*, Kluwer, Dordrecht, The Netherlands.

[11] Sladek, V., and Sladek, J., 1983, ''Three-Dimensional Curved Crack in an Elastic Body,'' Int. J. Solids Struct., **19**, No. 5, p. 425.

[12] Murakami, Y., and Nemat-Nasser, S., 1983, ''Growth and Stability of Interacting Surface Flaws of Arbitrary Shape,'' Eng. Fract. Mech., **17**, No. 3, p. 193.

[13] Xu, G., and Ortiz, M., 1993, ''A Variational Boundary Integral Method for the Analysis of 3-D Cracks of Arbitrary Geometry Modeled as Continuous Distributions of Dislocation Loops,'' Int. J. Numer. Methods Eng., **36**, p. 3675.

[14] Lothe, J., 1982, ''Dislocations in Anisotropic Media: The Interaction Energy,'' Philos. Mag. A, **46**, p. 177.

[15] Billy, B. A., and Eshelby, J. D., 1968, *Fracture: An Advanced Treatise*, Vol. 1, H. Liebowitz, ed., Academic Press, San Diego, CA.

[16] Barnett, D. M., and Asaro, R. J., 1972, ''The Fracture Mechanics of Slit-Like Cracks in Anisotropic Elastic Media,'' J. Mech. Phys. Solids, **20**, p. 353.

[17] Xu, G., Argon, A. S., and Ortiz, M., 1995, ''Nucleation of Dislocations From Crack Tips Under Mixed Modes of Loading: Implications for Brittle Against Ductile Behaviour of Crystals,'' Philos. Mag. A, **72**, p. 415.

[18] Kassir, M. K., and Sih, G. C., 1975, *Mechanics of Fracture: Three-Dimensional Crack Problems*, Noordhoff, Groningen.

[19] Hirth, J. P., and Lothe, J., 1982, *Theory of Dislocations*, 2nd Ed., John Wiley and Sons, New York.

Strength Analysis of Spherical Indentation of Piezoelectric Materials

A. E. Giannakopoulos
Department of Materials Science and Engineering,
Massachusetts Institute of Technology,
Cambridge, MA 02139
Mem. ASME

The present paper deals with theoretical and computational analysis of quasi-static, normal indentation of a transversely isotropic, linear elastic, piezoelectric half-space by a rigid spherical indenter. The contact is axisymmetric, nonconforming, monotonically advancing with load, frictionless and adhesionless. The indenter was modeled either as perfect conductor or as perfect insulator. The mechanical and electrical fields below the surface were examined. The issues of mechanical and dielectric strength due to indentation were examined using Weibull statistics of surface imperfections. The particular cases of PZT-4, PZT-5A, $BaTiO_3$, and $(Ba_{0.917}Ca_{0.083})TiO_3$ indented by rigid punches having either zero electrical potential or zero electric charge were solved with finite element analysis. [S0021-8936(00)02502-2]

Introduction

Piezoelectric materials have been investigated extensively from the atomic and continuum point of view by many investigators. The early history and treatment of piezoelectric crystals can be found in the classic work of Cady [1]. Other well-known treaties of piezoelectricity can be found in the books of Manson [2] and Tiersten [3] and most recently in Jaffe et al. [4] and Uchino [5]. From the mechanics point of view, the general formulation of piezoelectricity was developed by Toupin [6] and later by Mindlin [7]. Important general theorems on piezoelectricity were given by Nowacki [8].

Many natural inorganic materials, such as ice, granite, bones, teeth, coral, etc., are piezoelectric. Single and polycrystalline ferroelectric ceramics is another class of piezoelectrics. In recent years, new piezoelectric materials in the form of composites have been developed, Newnham et al. [9]. Such composites are made by ferroelectric ceramics like PZT in the form of particles of different shapes (e.g., spheres, hollow spheres, fibers, and tubes) which are dispersed in polymer matrices which have low dielectric constants. Other composites use a piezoelectric ceramic skeleton, backfilled with a polymer. Certain polymers can also be piezoelectric, having the advantage of softness which facilitates their formation into various shapes.

Piezoelectric materials show linearity between components of stress and strain, as well as between electric field and electric displacements, only over limited ranges of mechanical or electrical applied fields. The limits of linear behavior depend on the coercive field used to polarize the material and on the material composition. High temperature, mechanical and electrical applied fields tend to fracture and/or depolarize piezoelectric materials. Large-scale commercialization of piezoelectric materials has produced problems of reliability and durability [5]. Most piezoelectric actuators, ultrasonic motors, and other piezoelectric components require shapes in the form of thin films, beams, and plates with sizes that are becoming ever smaller. There is a need for systematic analyses on composition dependence of mechanical properties such as strength, creep, depoling, humidity, electrode delamination, and displacement drift. Lack of easy and inexpensive means of reproducibility and quality control of mass-produced piezoelectric components add largely to their production cost.

In this work, indentation of piezoelectric materials by an axisymmetric spherical indenter is proposed as a test for assessing their mechanical and dielectric strength or to examine pre-existing critical microcrack distributions at their surface. In particular, the indentation induced stresses and electric fields were investigated. The indenter was modeled as a rigid insulating or conducting sphere. The mechanical and electrical fields at the surface were used to model indentation induced mechanical and dielectric strength using Weibull statistics.

1 Field Equations of Axisymmetric Indentation

The analysis concerns materials with transverse isotropy, which is typically the case of many piezoelectric polycrystalline ceramics and ceramic composites. The general problem is formulated in cylindrical coordinates (r,θ,z), taking Oz as the vertical axis to the surface $(z=0)$ which is also the axis of transverse isotropy (hexagonal material symmetry) and the polarization axis (Fig. 1). The substrate is taken semi-infinite, $z\geqslant 0$, with θ being the angular position. In the following, subscript 3 refers to the Oz-axis and 1 and 2 (collectively noted as radial, r-axis) refer to arbitrary chosen orthogonal axes in the plane normal to the poling direction.[1] The shear strain in the plane 12, perpendicular to the polar axis is not excited piezoelectrically. The planar isotropy of poled materials along the z-axis means that an electric field parallel to the poling axis interacts in the same way with the axial stress along any radial direction, whereas an electric field parallel to any radial direction interacts in the same way with a shear stress, σ_{rz}, in the corresponding rz-plane. The present analysis is considered within the theory of small strains and small electric displacements, which is justified in the context of shallow elastic indentations. Isothermal conditions are implied for the material constants, their deviations from the adiabatic counterparts are expected to be small (which is true for piezoelectric ceramics ([4])).

In the absence of body and inertia forces, the stress equilibrium equations are

$$\partial\sigma_{rr}/\partial r+\partial\sigma_{rz}/\partial z+(\sigma_{rr}-\sigma_{\theta\theta})/r=0,$$
$$\partial\sigma_{rz}/\partial r+\partial\sigma_{zz}/\partial z+\sigma_{rz}/r=0. \quad (1)$$

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, May 21, 1999; final revision, Oct. 13, 1999. Associate Technical Editor: M.-J. Pindera. Discussion on the paper should be addressed to the Technical Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

[1]The poling direction is the direction of the applied electric field that polarizes permanently the initially unpoled material.

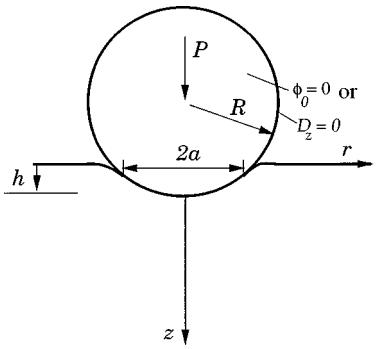


Fig. 1 Schematic of the normal indentation of piezoelectric materials by a rigid spherical indenter

In the absence of volume electric charges, the Maxwell electrostatic equation is

$$\partial D_r/\partial r + D_r/r + \partial D_z/\partial z = 0. \tag{2}$$

The small strain-displacement geometric relations give the strains in terms of the displacement vector (u_r, u_z)

$$\epsilon_{rr} = \partial u_r/\partial r, \quad \epsilon_{\theta\theta} = u_r/r, \quad \epsilon_{zz} = \partial u_z/\partial z,$$
$$\gamma_{rz} = \partial u_r/\partial z + \partial u_z/\partial r. \tag{3}$$

The electric flux vector (E_r, E_z) is given in terms of the electric potential ϕ by the Gauss equations:

$$E_r = -\partial\phi/\partial r, \quad E_z = -\partial\phi/\partial z. \tag{4}$$

In the absence of thermal strains, the constitutive equations of linear piezoelectricity are

$$\sigma_{rr} = c_{11}\epsilon_{rr} + c_{12}\epsilon_{\theta\theta} + c_{13}\epsilon_{zz} - e_{31}E_z,$$
$$\sigma_{\theta\theta} = c_{12}\epsilon_{rr} + c_{11}\epsilon_{\theta\theta} + c_{13}\epsilon_{zz} - e_{31}E_z,$$
$$\sigma_{zz} = c_{13}(\epsilon_{rr} + \epsilon_{\theta\theta}) + c_{33}\epsilon_{zz} - e_{33}E_z, \quad \sigma_{rz} = c_{44}\gamma_{rz} - e_{15}E_r, \tag{5}$$

where $c_{11}, c_{12}, c_{13}, c_{33}, c_{44}$ are the elastic constants at constant electric flux and e_{15}, e_{31}, e_{33} are the piezoelectric constants at constant stress. Initial stresses can be superposed within the context of linear elastic response.

The electric displacements (D_r, D_z) contain the coupling between strains and electric fluxes according to

$$D_r = e_{15}\gamma_{rz} + \epsilon_{11}E_r, \quad D_z = e_{31}(\epsilon_{rr} + \epsilon_{\theta\theta}) + e_{33}\epsilon_{zz} + \epsilon_{33}E_z, \tag{6}$$

where $\epsilon_{11}, \epsilon_{33}$ are the dielectric constants at constant strain.

Substituting Eqs. (3)–(6) into Eqs. (1)–(2), the problem is reduced to a system of three partial differential equations with unknowns the principal quantities (u_r, u_z, ϕ).

2 Problem Formulation and General Solution

2.1 Mechanical Boundary Conditions. A rigid, frictionless punch is indenting normally the piezoelectric half-space by an applied normal load, P (Fig. 1). The contact is nonconformal and is assumed to be monotonically advancing with load. If h is the penetration depth measured from the initially flat surface, then $dP/dh > 0$ for $h > 0$. As a result, the contact surface is a circular disk of radius a which expands with loading, $da/dh > 0$ for $h > 0$. Small strain analysis is adequate for small levels of indentation, which implies large diameter of the spherical indenter, D, compared to the contact radius, $a < 0.2D$.

The principal quantities (u_r, u_z, ϕ) are required to have continuous second derivatives with respect to the coordinates (r, z) and the following regularity condition at infinity:

$$(u_r, u_z\phi) \to o(1/\sqrt{z^2+r^2}), \quad \sqrt{z^2+r^2} \to \infty. \tag{7}$$

Inside the contact area, the mechanical boundary conditions must satisfy the applied normal displacement due to the rigid, frictionless, and adhesionless spherical punch profile. A spherical indenter of diameter D, where $a/D < 0.2$, can be approximated by a paraboloid of revolution, therefore,

$$u_z(r,0) = h - r^2/D; \quad 0 \leqslant r < a. \tag{8}$$

In all cases

$$\sigma_{rz}(r,0) = 0; \quad r \geqslant 0, \tag{9}$$

$$\sigma_{zz}(r,0) = 0; \quad r > a. \tag{10}$$

Equation (9) indicates absence of frictional or other applied shear tractions and Eq. (10) means zero normal tractions outside the contact region. Assuming smooth deformation transition at the contact perimeter $(r = a)$, the continuity condition requires $du_z(r = a-, z = 0)/dr = du_z(r = a+, z = 0)/dr$.

2.2 Electrical Boundary Conditions. The electrical boundary conditions depend on the conductivity of the punch. The following electrical boundary conditions reduce the general problem to two well-posed mixed boundary value problems which have unique solutions [10].

For the case of the indenter being a perfect electrical conductor with constant potential ϕ_0,

$$\phi(r,0) = \phi_0; \quad 0 \leqslant r < a, \tag{11}$$

$$D_z(r,0) = 0; \quad r > a. \tag{12}$$

Equation (11) indicates that the potential of the sphere will be the same as that of the indented material at the contact area, whereas Eq. (12) shows that the indentation-induced electric charge distribution outside the contact area will be zero.

For the case of the indenter being a perfect insulator with zero electric charge distribution,

$$D_z(r,0) = 0; \quad r \geqslant 0. \tag{13}$$

Equation (13) indicates that the indentation-induced electric charge distribution of the entire surface will be zero.[2]

2.3 General Features of the Solution. Hankel transformation with respect to the radial direction r, $(u_r, u_z, \phi) \to (\overline{u_r}(\xi,z), \overline{u_z}(\xi,z), \overline{\phi}(\xi,z))$, is used to transform the governing partial differential to a homogeneous system of ordinary differential equations with respect to z [11]. Hence, a general exponential form, $e^{-k\xi z}$, of the solution for the transformed principal variables is examined (ξ is the radial coordinate in the transformed space). The parameter k must satisfy the characteristic (3×3) determinant of the system of ordinary differential equations

$$\det[a_{ij}] = 0. \tag{14}$$

The same equations were derived by Ding et al.[12] using potential theory. The characteristic Eq. (14) is of sixth order and has two real roots, $k = \pm k_1$ and four complex roots $k = \pm(\delta \pm i\omega)$ $(i = \sqrt{-1})$, where it is assumed that all roots are distinct (without loss of generality, k_1, δ are taken positive definite and ω nonnegative). In order to satisfy regularity, select $k = k_1$ as the real root and then k_1, δ, ω can be found from Eq. (14) [13],

$$a_{11} = c_{44}k^2 - c_{11}, \quad a_{12} = -a_{21} = (c_{13} + c_{44})k, \quad a_{22} = c_{33}k^2 - c_{44},$$
$$a_{13} = a_{31} = (e_{31} + e_{15})k, \quad -a_{23} = a_{32} = -e_{33}k^2 + e_{15},$$

[2]It should be pointed out that Eqs. (12) and (13) indicate the contact-induced *additional* electric charge distribution of a piezoelectric. The surface of the piezoelectric is charged in its polarized state, however, this charge is ''bounded'' through the electric dipoles that are created in the material. In the present work, the problem formulation is such that the initial electric state does not enter the problem explicitly, but implicitly through the elastic, dielectric, and piezoelectric constants. Other approaches consider the poling explicitly, however, they lead to very cumbersome analysis.

$$a_{33}=\epsilon_{33}k^2-\epsilon_{11}. \tag{15}$$

Evaluating the coefficients a_{ij} for $k=k_1$, the parameters α_1, β_1, and γ_1 are defined as

$$\alpha_1=a_{12}a_{23}-a_{13}a_{22}, \quad \beta_1=-a_{11}a_{23}-a_{12}a_{13},$$

$$\gamma_1=a_{11}a_{22}+a_{12}^2. \tag{16}$$

The additional constants $\alpha_{21},\ldots,\gamma_{22}$ can be defined from δ and ω, using the complex identities (the parenthesis indicates evaluation of Eqs. (16) at $k=\delta+i\omega$):

$$\alpha_{21}+i\alpha_{22}=\alpha_1(\delta+i\omega), \quad \beta_{21}+i\beta_{22}=\beta_1(\delta+i\omega),$$

$$\gamma_{21}+i\gamma_{22}=\gamma_1(\delta+i\omega). \tag{17}$$

Inverting the Hankel transforms, the solution at the surface ($z=0$) can be represented in a general form as

$$u_z(r,0)=\int_0^\infty (M_1A_1(\xi)+M_2A_2(\xi))J_0(\xi r)d\xi, \tag{18}$$

$$\sigma_{zz}(r,0)=\int_0^\infty (M_5A_1(\xi)+M_6A_2(\xi))\xi J_0(\xi r)d\xi, \tag{19}$$

$$\phi(r,0)=\int_0^\infty (M_3A_1(\xi)+M_4A_2(\xi))J_0(\xi r)d\xi, \tag{20}$$

$$D_z(r,0)=\int_0^\infty (M_7A_1(\xi)+M_8A_2(\xi))\xi J_0(\xi r)d\xi, \tag{21}$$

where the constants M_i are defined in the Appendix and J_0, J_1 are Bessel functions of zeroth and first-order, respectively. The solution at infinity ($\sqrt{z^2+r^2}\to\infty$) tends asymptotically to the point force and point charge results. Equation (9) eliminates the $A_3(\xi)$ function. The remaining unknown functions $A_1(\xi)$, $A_2(\xi)$ can be found from the remaining surface electrical and mechanical boundary conditions.

At the surface ($z=0$), Eq. (5*d*) gives $\gamma_{rz}=e_{15}E_r/c_{44}$, indicating that the shear strain at the surface is not zero, although the shear stress is. It is clear that all mechanical and electrical constants affect the problem because they interact through the characteristic Eq. (14) in a very complex way. For the uncoupled problem ($e_{ij}=0$), the classic mechanical spherical indentation [14] and the rigid dielectric electrostatic results [15] of a transversely isotropic material are recovered separately. Note that the macroscopic response is independent of c_{12}.

3 Conducting Sphere

Equations (8)–(12) can be cast as a system of two dual integral equations, for the interior problem ($0\leqslant r<a$) and for the exterior problem ($r>a$). The solution follows from Sneddon [15] and is unique according to Walton [16].

Using the continuity contact condition $du_z(a,0)/dr=-2a/D$, the relation between the penetration depth h and the contact radius a is $h=2a^2/D$. Other continuity conditions, e.g., $\sigma_{zz}(a,0)=0$ or $D_z(a,0)=0$, are possible but could lead to jumps in the slope of the surface deformation at the contact perimeter and will not be addressed further in this work.

The surface vertical displacement outside the contact area ($r\geqslant a$) is

$$u_z(r,0)=(h/\pi)((2-(r^2/a^2))\arcsin(a/r)+(r/a)\sqrt{1-(a^2/r^2)}). \tag{22}$$

The radial displacement at the surface inside the contact area ($0\leqslant r\leqslant a$) is

$$u_r(r,0)=-\frac{8}{9\pi rD}\frac{M_9M_4-M_{10}M_3}{M_1M_4-M_2M_3}[a^3-(a^2-r^2)^{3/2}]$$
$$-\frac{8\phi_0}{(c_{11}-c_{12})\pi r}\frac{M_6M_1-M_2M_5}{M_1M_4-M_2M_3}[a-(a^2-r^2)^{1/2}], \tag{23a}$$

and outside the contact area ($r\geqslant a$) is

$$u_r(r,0)=-\frac{8a^3}{3\pi rD}\frac{M_9M_4-M_{10}M_3}{M_1M_4-M_2M_3}$$
$$-\frac{8\phi_0 a}{(c_{11}-c_{12})\pi r}\frac{M_6M_1-M_2M_5}{M_1M_4-M_2M_3}. \tag{23b}$$

The constants M_i are given in the Appendix. In order to have a nontrivial solution, $M_1M_4\neq M_2M_3$ and $M_5M_8\neq M_6M_7$, which are always satisfied for known piezoelectric ceramics. The above analytical expressions agree with those derived by Chen and Ding [17] who used the potential theory method; the present results are more straightforward and simpler.

4 Insulating Sphere

Equations (8)–(10) and (13) can be cast as a system of two dual integral equations for the interior and the exterior problem.

The continuity contact condition gives the relation between the penetration depth h and the contact radius a, $h=2a^2/D$, which is exactly as for the uncoupled mechanical indentation. The analytical expressions for the contact pressure $p(r)$, the charge distribution under the contact $q(r)$, and the resultant force P, are given in a previous work [18]. Of interest are the relations for the surface radial displacement and electric potential and were derived in this work.

The radial displacements at the surface are

$$u_r(r,0)=-\frac{8a^3}{3\pi rD}\frac{M_9M_8-M_{10}M_7}{M_1M_8-M_2M_7}\left(1-\left(1-\frac{r^2}{a^2}\right)^{3/2}\right);$$
$$0\leqslant r\leqslant a, \tag{24a}$$

$$u_r(r,0)=-\frac{8a^3}{3\pi rD}\frac{M_9M_8-M_{10}M_7}{M_1M_8-M_2M_7}; \quad r\geqslant a. \tag{24b}$$

At the surface, the electrical potential is

$$\phi(r,0)=\frac{(2a^2/D)(M_3M_8-M_4M_7)}{\pi(M_1M_8-M_2M_7)}\left(2-\frac{r^2}{a^2}\right); \quad 0\leqslant r\leqslant a, \tag{25a}$$

$$\phi(r,0)=\frac{(2a^2/D)(M_3M_8-M_4M_7)}{\pi(M_1M_8-M_2M_7)}\frac{2}{\pi}\left(\left(2-\frac{r^2}{a^2}\right)\arcsin\frac{a}{r}\right.$$
$$\left.+\frac{r}{a}\sqrt{1-\frac{a^2}{r^2}}\right); \quad r\geqslant a. \tag{25b}$$

The constants M_i are given in the Appendix. In order to have a nontrivial solution, $M_1M_8\neq M_2M_7$ and $M_5M_8\neq M_6M_7$, which are always satisfied for known piezoelectric ceramics.

5 Mechanical and Dielectric Strength

When investigating the mechanical and dielectric strength of brittle materials with spherical indenters, it is often the case that the most critical region is at the surface, close to the contact perimeter ($r=a$, $z=0$). Therefore, it seems important to examine the stresses and the electric flux in that region. From the boundary conditions and the constitutive relations, it can be shown that at the surface ($z=0$)

$$\epsilon_{zz}=E_z=0 \ \ (\text{for } r\geqslant 0), \quad \epsilon_{rr}+\epsilon_{\theta\theta}=0 \ \ (\text{for } r>a), \tag{26}$$

and therefore for $z=0$, $r>a$, Eq. (6) gives

$$\sigma_{rr} = -\sigma_{\theta\theta} = (c_{11}-c_{12})\epsilon_{rr} = (c_{11}-c_{12})\partial u_r/\partial r. \quad (27)$$

Using Eq. (23*b*) for the conducting indenter, the surface radial stress outside the contact radius ($r \geq a$) is found to be

$$\sigma_{rr}(r,0) = \frac{8(c_{11}-c_{12})}{\pi r^2}\left(\frac{M_9M_4 - M_{10}M_3}{M_1M_4 - M_2M_3}\frac{a^3}{3D} + \frac{M_6M_1 - M_2M_5}{M_1M_4 - M_2M_3}\frac{\phi_0 a}{c_{11}-c_{12}}\right), \quad (28)$$

whereas using Eq. (25*b*) for the insulating, the surface radial stress outside the contact radius ($r \geq a$) is found to be

$$\sigma_{rr}(r,0) = \frac{8a^3(c_{11}-c_{12})}{3\pi D r^2}\frac{M_9M_8 - M_{10}M_7}{M_1M_8 - M_2M_7}. \quad (29)$$

For most piezoelectric materials, the maximum tensile stress appears at the contact perimeter in the radial direction, $\max \sigma_1 = \sigma_{rr}(a,0)$.

For the perfect insulating indenter, the radial component of the electric flux at the surface is

$$E_r = \frac{2r}{a^2}\frac{(2a^2/D)(M_3M_8 - M_4M_7)}{\pi(M_1M_8 - M_2M_7)}; \quad 0 \leq r \leq a. \quad (30)$$

$$E_r = \frac{(2a^2/D)(M_3M_8 - M_4M_7)}{\pi(M_1M_8 - M_2M_7)}\frac{2}{\pi}\left(\frac{2r}{a^2}\arcsin\frac{a}{r} - 2\frac{\sqrt{r^2-a^2}}{ar}\right); \quad r \geq a. \quad (31)$$

Therefore, for piezoelectric materials indented by a perfect insulator, the maximum magnitude of electric flux is at the contact perimeter

$$\max\sqrt{E_z^2+E_r^2} = |E_r(a,0)| = \frac{(4a/D)|M_3M_8 - M_4M_7|}{\pi|M_1M_8 - M_2M_7|}. \quad (32)$$

If the piezoelectric constants are zero, then $M_3 = 0$ and $M_7 = 0$ and the problem decouples to a mechanical indentation and to a rigid dielectric problem of a transversely isotropic half space.

5.1 Statistical Aspects of Mechanical Strength. The mechanical strength of piezoelectric materials can be related to the density distribution of the surface microcracks. Spherical indentation has been used successfully to assess the distribution of microcracks on glass surfaces (e.g. Argon [19]). Such approach can be justified in the case of piezoelectrics from fracture mechanics analysis and experiments (e.g., see [20]) that indicate the mechanical strain energy release rate as the most suitable parameter for fracture criterion for non-conducting cracks (this may not be true for conducting cracks). Various surface stress combinations can be used in order to examine the statistical aspects of strength [21]. Since the radial stress, $\sigma_{rr}(r,0) = \Sigma$, is also the maximum tensile principal stress at the surface, the simplest stress criterion would involve Σ alone. If $g_c(\Sigma)$ is the critical stress distribution for semi-elliptical cracks per unit surface area, then the probability of no fracture below load P in the entire surface outside the contact circle a is

$$\Phi_c(P) = \exp\left(-2\pi\int_{a(P)}^{\infty} r\,dr\int_0^{\Sigma(r,P)} g_c(\Sigma)d\Sigma\right). \quad (33)$$

The relation between a and P is given by Giannakopoulos and Suresh [18]:

$$-P = \frac{16a^3}{3D}\frac{M_4M_5 - M_3M_6}{M_1M_4 - M_2M_3} + 4a\phi_0\frac{M_6M_1 - M_2M_5}{M_1M_4 - M_2M_3}; \quad \text{conducting indenter,} \quad (34)$$

$$-P = \frac{16a^3}{3D}\frac{M_8M_5 - M_7M_6}{M_1M_8 - M_2M_7}; \quad \text{insulating indenter.} \quad (35)$$

For the cases of an insulating indenter, or a conducting indenter with zero potential ($\phi_0 = 0$), Eqs. (28)–(29) and (34)–(35) can be expressed in a more compact form as

$$\Sigma(P) = C_cP/r^2 \quad (r \geq a), \quad (36)$$

where C_c is a constant which depends on the material properties and the electrical conditions of the indenter. Equation (33) simplifies to [19]

$$\ln\Phi_c(P) = -\pi a^2\int_0^{C_cP/a^2} g_c(\Sigma)\left(\frac{C_cP}{a^2\Sigma} - 1\right)d\Sigma. \quad (37)$$

Assume that g_c can be expanded in simple power functions of Σ [19],

$$g_c(\Sigma) = V_1(n_1-1)n_1\Sigma^{(n_1-1)} + \ldots (n_i \geq 1), \quad (38)$$

where V_i and n_i are constants that depend on the material, the surface damage condition, and the environment. Then the probability of no fracture for loads below P is

$$\Phi_c(P) = \exp(-\pi a^2(V_1(C_cP/a^2)^{n_1} + \ldots)), \quad (39)$$

which is in the classic form proposed by Weibull [22], with n_1 being Weibull modulus.

The previous analysis assumes zero residual surface stresses. However, Pohanka et al. [23,24] showed that tensile internal stresses due to polarization degrade the strength of $BaTiO_3$ and PZT-4 piezoelectric ceramics. Pohanka et al. found that machining defects at the surface act as sources of failure. Their results are in accord with the permittivity analysis of Buessem et al. [25] who found tensile internal stresses in constrained piezoelectric grains that undergo cubic-to-tetragonal transformation. Pohanka et al. assumed a relation between the microcrack size c the applied stress Σ and the internal stress Σ_R according to linear elastic fracture mechanics

$$\Sigma + \Sigma_R \approx C_A/\sqrt{c}, \quad (40)$$

where C_A is a constant that depends on the fracture energy G_c, the elastic modulus, c_{33} and, weakly, on the crack configuration $C_A \approx \sqrt{G_c c_{33}}$. Pohanka et al. found that the internal tensile stress depends on the microcrack size c and the grain size d. Their experimental results may be approximated in a simple way as

$$\Sigma_R \approx \Sigma_R^0\sqrt{d}/\sqrt{c} = C_R/\sqrt{c}, \quad (41)$$

where Σ_R^0 is the internal stress for $c \to d$. Combining (40) and (41)

$$\Sigma \approx (C_A - C_R)\sqrt{c}, \quad (42)$$

where $C_A > C_R$, i.e., $\sqrt{G_c c_{33}} > \Sigma_R^0\sqrt{d}$, in the absence of spontaneous cracking.

The failure probability analysis may be equivalently reformulated in terms of a critical distribution of microcrack sizes at the surface, using (42)

$$g_c(c) = V_1(n_1-1)n_1(C_A - C_R)^{(n_1-1)}c^{(1-n_1)/2} + \ldots. \quad (43)$$

Assuming that the essential statistic parameters are the same for the poled and unpoled specimens and that the first term in (43) dominates the problem, then the condition for similar probability of failure for both poled and unpoled cases gives the following relation between the critical indentation loads:

$$P_{\text{poled}}/P_{\text{unpoled}} = (1-(C_R/C_A))^{1-(1/n_1)}C_c^{\text{unpoled}}/C_c^{\text{poled}}. \quad (44)$$

The previous analysis is simplistic since it ignores the toughness anisotropy due to microcracking orientation with respect to the poling direction [26].

5.2 Statistical Aspects of Dielectric Strength. The dielectric strength is the resistance of the material when changing from

its dielectric to the conducting state, in the presence of high electric flux. Obviously, the dielectric strength of piezoelectric materials can be related to the density distribution of surface imperfections. However, imperfections in the bulk of the material also trigger dielectric breakdown (for PZT ceramics, Gerson and Marshall [27] reported that dielectric strength reduces logarithmically with bulk porosity). In spherical indentation, the induced radial electric flux at the surface $E_r(r,0)=E$ is also the maximum principal flux, therefore, the simplest dielectric strength criterion would involve E alone. Such an assumption implies that dielectric breakdown could start from a surface imperfection. If $g_e(E)$ is the critical electric flux distribution for such imperfections per unit surface area, then the probability of no dielectric breakdown below load P in the entire surface is

$$\Phi_e(P)=\exp\left(-2\pi\int_0^\infty r\,dr\int_0^{E(P)}g_e(E)dE\right). \tag{45}$$

The electric field intensity factor at the tips of the surface flaws is analogous to the stress intensity factor (see, for example, [20]). Thus, pre-existing surface flaws are expected to have a similar role in both mechanical failure and dielectric breakdown. Indeed, Yamashita et al. [28] and Kishimoto et al. [29] found similarity between the shapes of mechanical and dielectric strength distributions for piezoelectric ceramics of small grain size, in both ferroelectric and paraelectric state. These experimental results suggest that

$$g_e(E)=W_1(n_1-1)n_1E^{(n_1-1)}+\ \ldots, \tag{46}$$

where the Weibull modulus n_1 is the same with that of $g_c(\Sigma)$, Eq. (38), and $V_1\Sigma^{(n_1-1)}\approx W_1E^{(n_1-1)}$.

5.3 The Case of an Indenter With Nonzero Electric Potential. The case of the indenter being a perfect conductor currying nonzero potential ϕ_0 is interesting for the reason that the normal contact stress and the surface electric charge at the contact area $(0\leqslant r\leqslant a)$ have square-root radial singularity, as also noted by Chen and Ding [17]

$$\sigma_{zz}(r,0)=\frac{8}{\pi D}\frac{M_4M_5-M_3M_6}{M_1M_4-M_2M_3}\sqrt{a^2-r^2}+\frac{2}{\pi}\frac{M_6M_1-M_2M_5}{M_1M_4-M_2M_3}\frac{\phi_0}{\sqrt{a^2-r^2}}, \tag{47}$$

$$D_z(r,0)=\frac{8}{\pi D}\frac{M_7M_4-M_8M_3}{M_1M_4-M_2M_3}\times\sqrt{a^2-r^2}\,\frac{2}{\pi}\frac{M_1M_8-M_2M_7}{M_1M_4-M_2M_3}\frac{\phi_0}{\sqrt{a^2-r^2}}. \tag{48}$$

A compressive stress singularity is expected to create microdamage at the contact perimeter; the electric charge singularity is expected to depole the material at the contact perimeter. These effects attenuate with decreasing indenter's diameter D. It is then expected that due to the local damage, energy will be dissipated and the probability of macroscopic indentation-induced cracking be lowered. High positive electric potential would increase the tensile radial stress at the contact perimeter, Eq. (28), and therefore increase the probability of macroscopic cracking. On the other hand, high negative electric potential could result in lower tensile stresses at the contact perimeter, Eq. (29), and reduce the probability of macroscopic cracking.

The electric potential, however, cannot be arbitrarily negative. The stability condition $dP/dh>0$ for $a>0$ and $dP/dh=0$ if $a=0$ is satisfied if

$$\phi_0D(M_6M_1-M_2M_5)+4a^2(M_4M_5-M_3M_6)\leqslant 0, \tag{49}$$

which also guarantees that the contact stresses are compressive everywhere, $\sigma_{zz}(r,0)\geqslant 0$. High negative electric potential may violate Eq. (49) and result in singular tensile contact stresses at the contact perimeter. This situation is possible, if adhesion between the contacting surfaces is also permitted. In the absence of adhesion, the solution that violates (49) is inadmissible and the problem needs to be reformulated.

6 Finite Element Analysis

The ABAQUS [30] general purpose finite element program was used with certain modifications regarding the electric contact conditions. A mesh of four-node axisymmetric elements was used, with progressively varying element size. The final mesh had 4747 elements and 5058 nodes (Fig. 2). Full Gauss integration scheme was used. No special types of elements were used (''singular'' or ''infinite'' type of elements that have the spatial $r^{1/2}$ or r^{-1} singularity of the problem). The contact radius was resolved with 24 elements. The outer boundary was at least $20a$ away from the contact regime. With reference to Fig. 2, the outer boundary conditions were vertically constrained along the sides CD and AB. In addition, sides CD and AB were given zero electrical potential. In all cases, the uncoupled problem $(e_{ij}=0)$ was also solved. The contact stresses were found to be in agreement within five percent error when compared to analytic results.

Four piezoelectric materials were analyzed, PZT-4, PZT-5A, $BaTiO_3$, and 95 percent $BaTiO_3$–5 percent $CaTiO_3$. Their fully poled, room temperature mechanical, dielectric, and piezoelectric properties are shown in Table 1. The data were collected from Jaffe et al. [4] and Bechmann [31]. The indenting sphere was taken to be rigid and approximated by an axisymmetric paraboloid of diameter D, which was kept the same in all cases. For the coupled cases, the sphere was taken either as a perfect conductor with zero electric potential $(\phi_0=0)$, or as a perfect insulator with no surface electric charge distribution $(D_z=0)$. A constant average pressure $P/(\pi a^2)=33.84$ GPa was used in all calculations (the results for different average pressures scale according to Table 2).

The normalized applied load $P/(D^{1/2}h^{3/2})$, maximum principal tensile stress $\max\sigma_1$, and the average electric charge distribution $0.1875Q(D/a^3)$ are tabulated in Table 2. These are results that are useful for macroscopic observations of the mechanical strength due to indentation.

The contours of principal tensile stress, σ_1, are shown in Fig. 3(*a*) for the uncoupled case of PZT-4, in Fig. 3(*b*) for the coupled case of PZT-4 with conducting sphere and in Fig. 3(*c*) for the coupled case of PZT-4 with insulating sphere. The contours of principal tensile stress are shown in Fig. 4(*a*) for the uncoupled

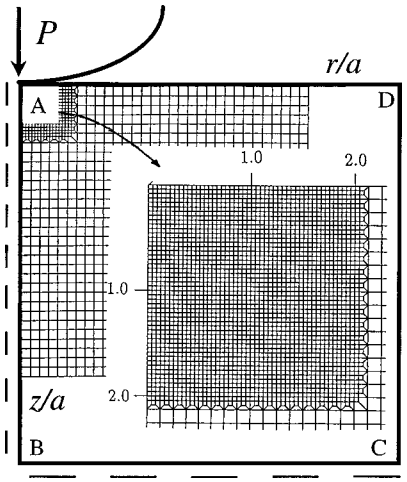


Fig. 2 Overall view of the finite element mesh used in the present calculations; details of the mesh close to and away from the contact area are included

Table 1 Piezoelectric properties

Elastic Stiffness Coefficients (GPa)	PZT-4	PZT-5A	$BaTiO_3$	$(Ba_{0.917}Ca_{0.083})TiO_3$
C_{11}	139.00	121.00	166.00	158.00
C_{33}	115.00	111.00	162.00	150.00
C_{44}	25.60	21.10	42.90	45.00
C_{12}	77.80	75.40	76.60	69.00
C_{13}	74.30	75.20	77.50	67.50
Piezoelectric Coefficients (C/m^2)	PZT-4	PZT-5A	$BaTiO_3$	$(Ba_{0.917}Ca_{0.083})TiO_3$
e_{31}	−5.200	−5.400	−4.400	−3.100
e_{33}	15.10	15.80	18.60	13.50
e_{15}	12.70	12.30	11.60	10.90
Dielectric Constants (10^{-9} F/m)	PZT-4	PZT-5A	$BaTiO_3$	$(Ba_{0.917}Ca_{0.083})TiO_3$
ϵ_{11}	6.461	8.107	11.151	8.850
ϵ_{33}	5.620	7.346	12.567	8.054

case of 95 percent $BaTiO_3$–5 percent $CaTiO_3$, in Fig. 4(*b*) for the coupled case of 95 percent $BaTiO_3$–5 percent $CaTiO_3$ with conducting sphere and in Fig. 4(*c*) for the coupled case of 95 percent $BaTiO_3$–5 percent $CaTiO_3$ with insulating sphere. Comparison of Figs. 3 and 4 reveals the influence of the strong anisotropy of the poled PZT-4, compared to the less anisotropic 95 percent $BaTiO_3$–5 percent $CaTiO_3$. In all cases, the maximum tensile stress appears at the contact perimeter and is in the radial direction. However, for the coupled (poled) cases, another strong tensile stress region appears below the contact area. The fields indicate that possible cracking locations are at the contact perimeter (Hertzian type) and to a lesser extent in the interior, along the axis of loading. In the coupled cases, the cracking, once it occurs, is likely to be unstable because, contrary to the uncoupled cases, the crack paths will be under strong tensile fields. Moreover, the toughness is weaker for a crack perpendicular to the poling direction ([26]). The magnitudes of the tensile stresses below the indentation are not very different for the cases of insulating or conducting type of indenter. The maximum tensile stress below the indentation occurs deeper for the conducting type of indenter than for the insulating type of indenter.

Figure 5(*a*) shows the magnitude of the electric flux distribution $\sqrt{E_r^2+E_z^2}$, for the coupled case of 95 percent $BaTiO_3$–5 percent $CaTiO_3$ with conducting sphere and Fig. 5(*b*) shows the magnitude of the electric flux distribution for the coupled case of 95 percent $BaTiO_3$–5 percent $CaTiO_3$ with insulating sphere. The maximum magnitude of the electric flux occurs at the contact perimeter ($r=a$) for the insulating type of the indenter and at the center of the contact area ($r=0$) for the conducting type of in-

Table 2 Finite element results for spherical indenter

Material	Indenter	$P/(D^{1/2}h^{3/2})$ (GPa) (Theory)	max $\sigma_1(\pi a^2)/P$ (Theory)	$0.1875Q(D/a^3)$ C/m^2 (Theory)
PZT-4	(uncoupled)	105.14 (89.5)	0.07181	0
	conductor	86.16 (84.7)	0.06620 (0.0578)	−12.6 (−13.1)
	insulator	75.06 (79.6)	0.07683 (0.0864)	0
PZT-5A	conductor	67.63	-	32.36
	insulator	58.95	-	0
95%$BaTiO_3$ 5%$CaTiO_3$	(uncoupled)	121.23 (112.6)	0.10370	0
	conductor	124.03 (123.1)	0.08687 (0.0432)	−6.11 (−6.35)
	insulator	122.78	0.1371	0
$BaTiO_3$	(uncoupled)	150.69 (148.9)	0.1176	0
	conductor	157.28 (153.1)	0.1031	−8.32 (−8.7)
	insulator	157.88 (154.0)	0.1359	0

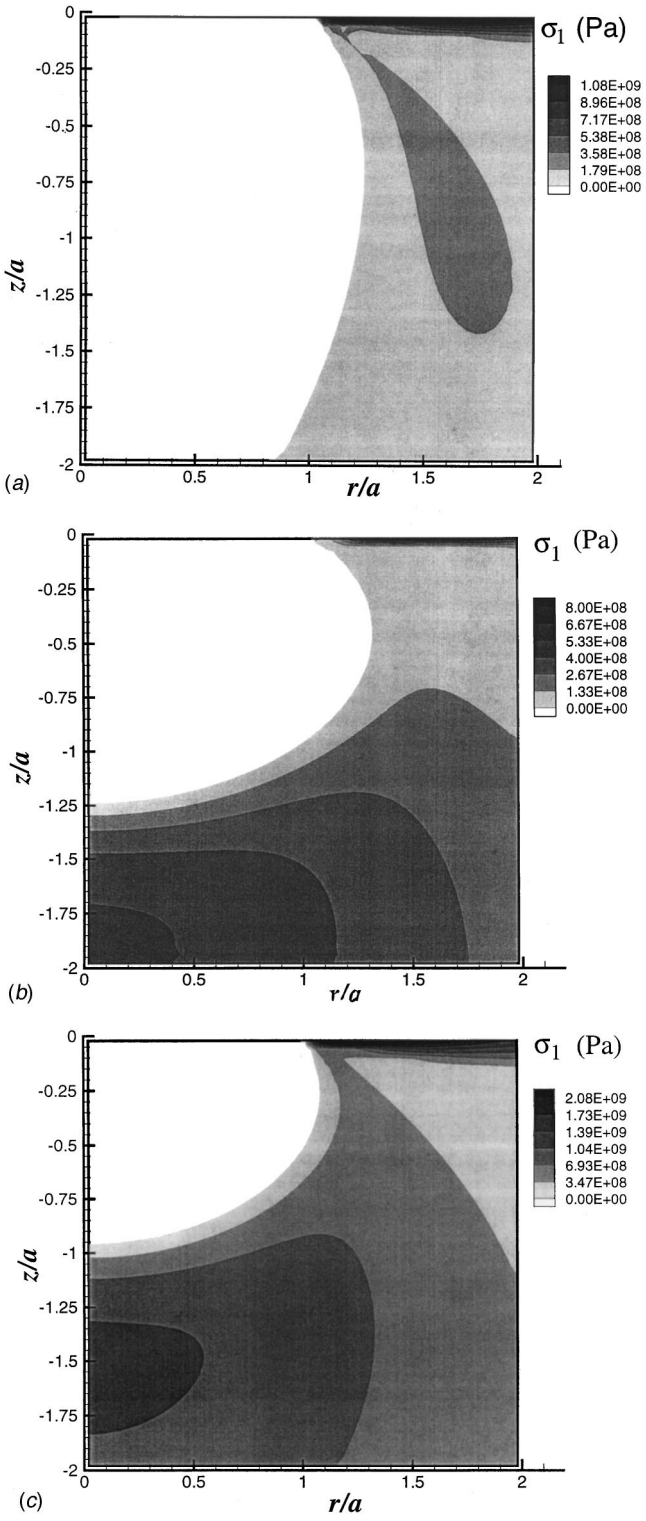


Fig. 3 Maximum tensile principal stress distribution for spherical indentation of PZT-4; (*a*) uncoupled case ($P/(\pi a^2)=33.84$ GPa), (*b*) coupled case, with indenter being a perfect conductor of zero electric potential ($P/(\pi a^2)=33.84$ GPa), (*c*) coupled case, with indenter being a perfect insulator of zero surface electric charge ($P/(\pi a^2)=33.84$ GPa)

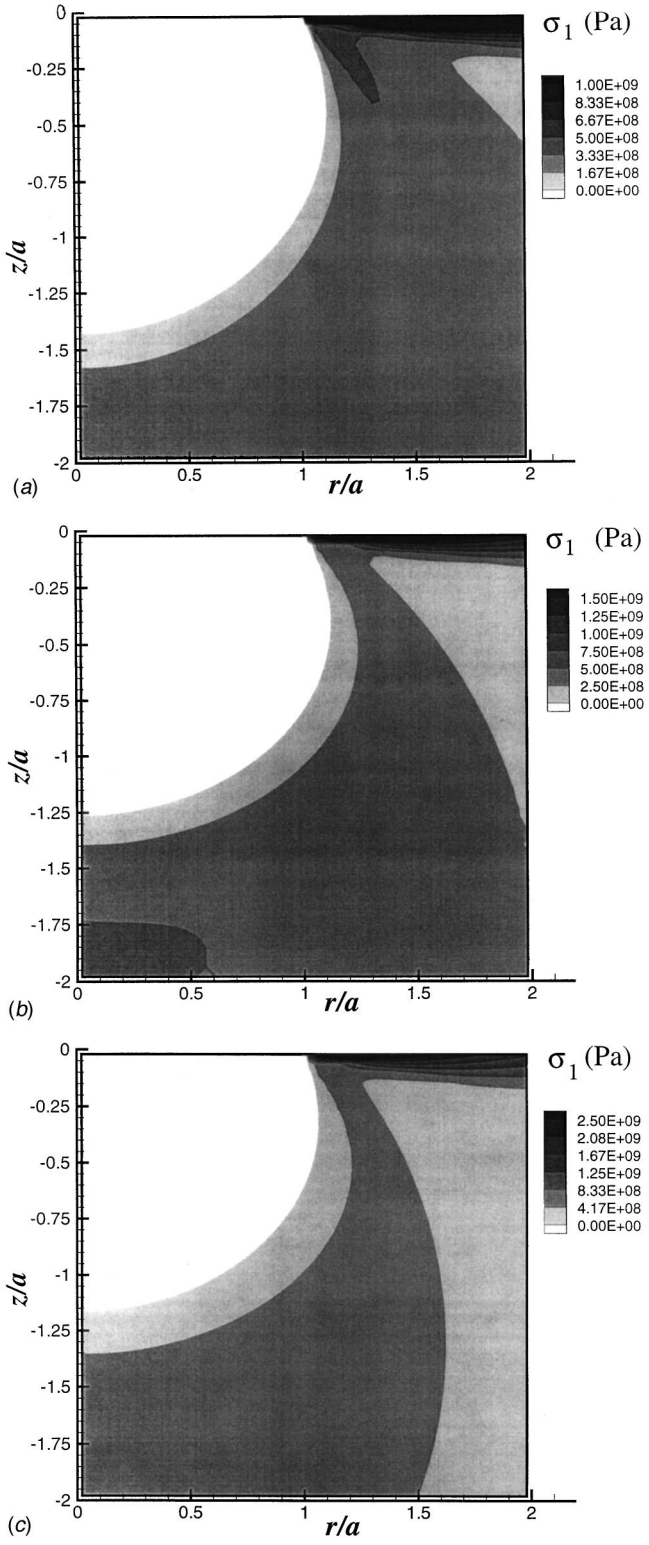


Fig. 4 Maximum tensile principal stress distribution for spherical indentation of 95 percent $BaTiO_3$ – 5 percent $CaTiO_3$; (*a*) uncoupled case ($P/(\pi a^2) = 33.84$ GPa), (*b*) coupled case, with indenter being a perfect conductor of zero electric potential ($P/(\pi a^2) = 33.84$ GPa) (*c*) coupled case, with indenter being a perfect insulator of zero surface electric charge ($P/(\pi a^2) = 33.84$ GPa)

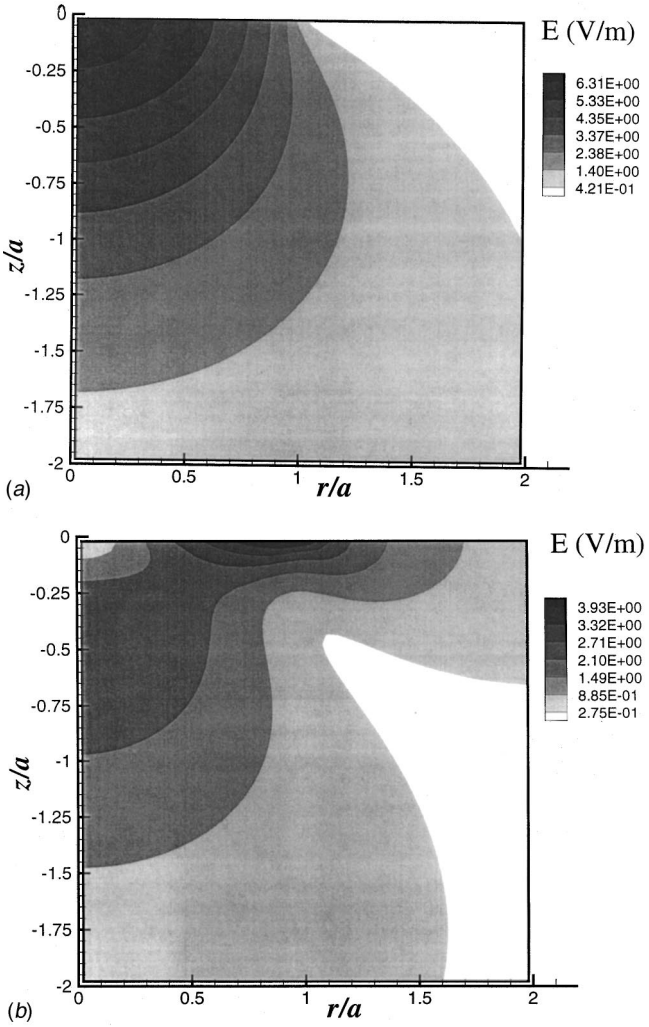


Fig. 5 Magnitude of electric flux distribution, $\sqrt{E_r^2 + E_z^2}$, for spherical indentation of 95 percent $BaTiO_3$ – 5 percent $CaTiO_3$ (coupled case), (*a*) indented with conducting sphere of zero electric potential ($P/(\pi a^2) = 33.84$ GPa), (*b*) indented with nonconducting sphere of zero surface electric charge ($P/(\pi a^2) = 33.84$ GPa)

denter. This indicates that either the contact perimeter or the contact center are the most critical regions where contact induced piezoelectric depoling may first occur. The maximum magnitude of electric flux is higher for the conducting type of indenter.

It was found earlier that if the critical stress distribution of microcracks per unit surface area, g_c, remains the same for both the unpoled and the poled material, then the probability of no fracture under spherical indentation depends on the electric contact conditions. The present analysis predicts that for PZT-4 the probability of no failure at indentation load, P, is higher for the unpoled material and lower for the poled material indented by the same sphere. Therefore, indentation induced cracking is expected for the poled PZT-4 material at lower load than the unpoled material. In addition, the probability of no failure is slightly higher for the indentation of the poled material with an insulating indenter than with a conducting indenter. Therefore, indentation induced cracking is expected for the poled PZT-4 material indented by a conducting indenter at lower load than indented by an insulating indenter of the same diameter.

Conclusions

The quasi-static spherical indentation of homogeneous, transversely isotropic, linear piezoelectric materials has been examined analytically and numerically. The surface values of stresses, as well as the electrical potential and electric charges, were found explicitly. The mechanical and electrical fields below the surface were also computed. A Weibull strength analysis was developed based on critical surface microcrack distributions. The electric conditions of the indenter were found to be important.

The analysis suggests the potential use of spherical indentation as a mechanical or dielectric strength test of piezoelectrics. In many cases, indentation can be the only method for testing small volumes of piezoelectric materials existing or processed in small volumes like thin films, layered plates, or composites. Indentation can be used to investigate unwanted time effects like loss of piezoelectricity due to aging during service or storage. Spherical indentation can model contact-induced damage in tribological applications like low-velocity impact, scratching, and wear, therefore providing design methods for protecting piezoelectric surfaces.

Appendix

Constants Used in the Analysis

$$m_1 = -k_1 e_{15}\gamma_1 - c_{44}(k_1\alpha_1 + \beta_1)$$

$$m_2 = -e_{15}(\delta\gamma_{21} - \omega\gamma_{22}) - c_{44}(\delta\alpha_{21} - \omega\alpha_{22} + \beta_{21})$$

$$m_3 = -e_{15}(\delta\gamma_{22} + \omega\gamma_{21}) - c_{44}(\delta\alpha_{22} + \omega\alpha_{21} + \beta_{22})$$

$$B_1 = c_{13}\alpha_1 - k_1\beta_1 c_{33} - k_1\gamma_1 e_{33},$$

$$B_4 = e_{31}\alpha_1 - k_1\beta_1 e_{33} + k_1\gamma_1\epsilon_{33}$$

$$B_2 = c_{13}\alpha_{21} - c_{33}(\delta\beta_{21} - \omega\beta_{22}) - e_{33}(\delta\gamma_{21} - \omega\gamma_{22})$$

$$B_3 = -c_{13}\alpha_{22} + c_{33}(\delta\beta_{22} + \omega\beta_{21}) + e_{33}(\delta\gamma_{22} + \omega\gamma_{21})$$

$$B_5 = e_{31}\alpha_{21} - e_{33}(\delta\beta_{21} - \omega\beta_{22}) + \epsilon_{33}(\delta\gamma_{21} - \omega\gamma_{22})$$

$$B_6 = -e_{31}\alpha_{22} + e_{33}(\delta\beta_{22} + \omega\beta_{21}) - \epsilon_{33}(\delta\gamma_{22} + \omega\gamma_{21})$$

$$M_1 = \beta_1 - \beta_{22}m_1/m_3, \quad M_2 = \beta_{21} - \beta_{22}m_2/m_3$$

$$M_3 = \gamma_1 - \gamma_{22}m_1/m_3, \quad M_4 = \gamma_{21} - \gamma_{22}m_2/m_3$$

$$M_5 = B_1 + B_3m_1/m_3, \quad M_6 = B_2 + B_3m_2/m_3$$

$$M_7 = -B_4 - B_6m_1/m_3, \quad M_8 = B_5 + B_6m_2/m_3$$

$$M_9 = |\alpha_1 - \alpha_{22}m_1/m_3|, \quad M_{10} = |\alpha_{21} - \alpha_{22}m_2/m_3|$$

References

[1] Cady, W. G., 1964, *Piezoelectricity*, Dover, New York.
[2] Mason, W. P., 1950, *Piezoelectric Crystals and Their Application to Ultrasonics*, Van Nostrand, New York.
[3] Tiersten, H. F., 1969, *Linear Piezoelectric Plate Vibrations*, Plenum, New York.
[4] Jaffe, B., Cook, W. R., and Jaffe, H., 1971, *Piezoelectric Ceramics*, Academic Press, San Diego, CA.
[5] Uchino, K., 1997, *Piezoelectric Actuators and Ultrasonic Motors*, Kluwer, Boston.
[6] Toupin, R. A., 1956, ''The Elastic Dielectric,'' J. Ration. Mech. Anal., **5**, pp. 849–915.
[7] Mindlin, R. D., 1972, ''Elasticity, Piezoelectricity and Crystal Lattice Dynamics,'' J. Elast., **2**, pp. 217–282.
[8] Nowacki, W., 1978, ''Some General Theorems of Thermo-Piezo-Electricity,'' J. Therm. Stresses, **1**, pp. 171–182.
[9] Newnham, R. E., Skinner, D. P., and Cross, L. E., 1978, ''Connectivity and Piezoelectric—Pyroelectric Composites,'' Mater. Res. Bull., **13**, pp. 525–536.
[10] Majorkowska-Knap, K., 1987, ''Uniqueness Theorem of Linear Thermo/Piezoelectricity,'' Bull. Pol. Acad. Sci.: Tech. Sci., **35**, pp. 163–177.
[11] Sneddon, I. N., 1980, *Special Functions of Mathematical Physics and Chemistry*, Longman, London.
[12] Ding Haojiang, Chenbuo, and Lian, Gjian, 1996, ''General Solutions for Coupled Equations for Piezoelectric Media,'' Int. J. Solids Struct., **33**, pp. 2283–2298.
[13] Matysiak, S., 1985, ''Axisymmetric Problem of Punch Pressing into a Piezoelectric Half Space,'' Bull. Pol. Acad. Sci.: Tech. Sci., **33**, pp. 25–33.
[14] Dahan, M., and Zarka, J., 1977, ''Elastic Contact Between a Sphere and a Semi-infinite Transversely Isotropic Body,'' Int. J. Solids Struct., **13**, pp. 229–238.
[15] Sneddon, I. N., 1966, *Mixed Boundary Problems in Potential Theory*, North-Holland, Amsterdam.
[16] Walton, J. R., 1975, ''A Distributional Approach to Dual Integral Equations of Titchmarch Type,'' SIAM (Soc. Ind. Appl. Math.) J. Math. Anal., **6**, pp. 628–643.
[17] Chen, W., and Ding, H., 1999, ''Indentation of a Transversely Isotropic Piezoelectric Half-Space by a Rigid Sphere,'' Acta Mech. Solid. Sin., **12**, pp. 114–120.
[18] Giannakopoulos, A. E., and Suresh, S., 1999, ''Theory of Indentation of Piezoelectric Materials,'' Acta Mater., in press.
[19] Argon, A. S., 1959, ''Distribution of Cracks on Glass Surfaces,'' Proc. R. Soc. London, Ser. A, **250**, pp. 482–492.
[20] Park, S., and Sun, C.-T., 1995, ''Fracture Criteria for Piezoelectric Ceramics,'' J. Am. Ceram. Soc., **78**, pp. 1475–1480.
[21] Batdorf, S. B., and Heinisch, H. L., 1978, ''Weakest Link Theory Reformulated for Arbitrary Fracture Criterion,'' J. Am. Ceram. Soc., **61**, pp. 355–358.
[22] Weibull, W., 1951, ''A Statistical Distribution Function of Wide Applicability,'' ASME J. Appl. Mech., **18**, pp. 293–297.
[23] Pohanka, R. C., Rice, R. W., and Walker, B. F., 1976, ''Effect of Internal Stress on the Strength of $BaTiO_3$,'' J. Am. Ceram. Soc., **59**, pp. 71–74.
[24] Pohanka, R. C., Freiman, S. W., and Rice, R. W., 1980, ''Fracture Process in Ferroic Materials,'' Ferroelectrics, **28**, pp. 337–342.
[25] Buessem, W. R., Cross, L. E., and Goswami, A. K., 1966, ''Phenomenological Theory of High Permittivity in Fine Grained Barium Titanate,'' J. Am. Ceram. Soc., **49**, pp. 33–36.
[26] Pisarenko, G. G., Chushko, V. M., and Kovalev, S. P., 1985, ''Anisotropy of Fatigue Toughness of Piezoelectric Ceramics,'' J. Am. Ceram. Soc., **68**, pp. 259–265.
[27] Gerson, R., and Marshall, T. C., 1959, ''Dielectric Breakdown of Porous Ceramics,'' J. Appl. Phys., **30**, pp. 1650–1653.
[28] Yamashita, K., Koumoto, K., and Yanagida, H., 1984, ''Analogy Between Mechanical and Dielectric Strength Distributions for $BaTiO_3$ Ceramics,'' J. Am. Ceram. Soc., **67**, pp. C31–C33.
[29] Kishimoto, A., Koumoto, K., and Yanagida, H., 1989, ''Mechanical and Dielectric Failure of $BaTiO_3$ Ceramics,'' J. Mater. Sci., **24**, pp. 698–702.
[30] ABAQUS, 1996, *Finite Element Code*, Version 5.5, Hibbit, Karlsson, and Sorensen, Pawtucket, RI.
[31] Bechmann, R., 1956, ''Elastic, Piezoelectric and Dielectric Constants of Polarized Barium Titanate Ceramics and Some Applications of the Piezoelectric Equations,'' J. Acoust. Soc. Am., **28**, pp. 347–350.

Journal of Applied Mechanics

Brief Notes

A Brief Note is a short paper that presents a specific solution of technical interest in mechanics but which does not necessarily contain new general methods or results. A Brief Note should not exceed 1500 words *or equivalent* (a typical one-column figure or table is equivalent to 250 words; a one line equation to 30 words). Brief Notes will be subject to the usual review procedures prior to publication. After approval such Notes will be published as soon as possible. The Notes should be submitted to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME, United Engineering Center, Three Park Avenue, New York, NY 10016-5990, or to the Technical Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on Brief Notes appearing in this issue will be accepted until two months after publication. Readers who need more time to prepare a Discussion should request an extension of the deadline from the Editorial Department.

Zener's Crack and the *M*-Integral

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

In a pair of bonded solids, the interface may block dislocation gliding. The pileup may cause a crack to nucleate either on the interface, or in one of the solids. The model, proposed by Zener half a century ago, has been analyzed in various forms. This note shows that the energy release rate of the crack can be calculated by an application of the M-integral. Both solids are anisotropic, and the interface is flat. The result leads to a discussion of the crack orientation. [S0021-8936(00)00701-7]

Introduction

Zener [1] proposed that a dislocation pileup concentrates stress, which may cause a crack to nucleate. Any crystallographic discontinuity (e.g., a grain boundary or a phase boundary) may act as an obstacle to block dislocation gliding. In a pair of bonded solids, when the interface blocks dislocations, the crack can nucleate either on the interface (Fig. 1(*a*)), or in one of the solids (Fig. 1(*b*, *c*)). The model has been analyzed by many authors in the last 50 years, as reviewed by Cottrell [2] and more recently by Cherepanov [3] and Fan [4]. A main result is the energy release rate of the crack, which has been obtained by solving elasticity boundary value problems for various special cases. When the two solids have dissimilar elastic properties, only the case where the crack lies on the interface (Fig. 1(*a*)) has been solved. This note calculates the energy release rate by using a path-independent integral, following the procedure of Freund [5]. He gave several examples of cracks in an isotropic and homogeneous solid. Here the two solids are anisotropic and dissimilar. The crack can be either on the interface, or in one of the solids.

Calculation

Consider an elastic solid in a state of plane-strain deformation. Knowles and Sternberg [6] introduced the following integral:

$$M=\int_C (wn_\alpha - t_\beta u_{\beta,\alpha})x_\alpha dS. \tag{1}$$

The integral is over a closed curve C in the plane. Here x_α is the rectangular coordinate, n_α the unit normal vector to C, w the strain energy density, u_α the displacement, and t_α the traction. If the solid is homogeneous along rays from the coordinate origin and if C encloses no singularity, $M=0$; that is, M is a path-independent integral. For a pair of dissimilar solids bonded on a flat interface, placing the coordinate origin at any point on the interface satisfies the homogeneity requirement.

If C does enclose a singularity, M may not vanish. For example, Rice [7] showed that the M-integral over a closed curve around a dislocation equals the pre-logarithmic factor of the dislocation energy. Consider a dislocation of the Burgers vector $\boldsymbol{b}$ lying on the

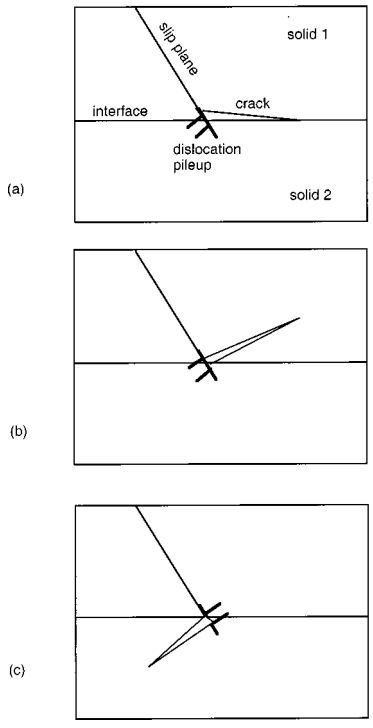


Fig. 1 Zener's model. Dislocations pile up at an obstacle, i.e., the interface between two solids. The intense stress causes a crack to nucleate either on the interface, or in one of the solids.

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 21, 1999; final revision, Oct. 12, 1999. Associate Technical Editor: K. Ravi-Chandar.

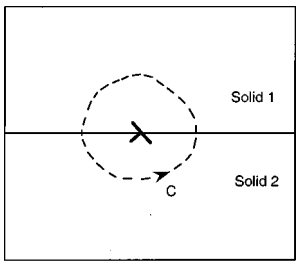


Fig. 2 A dislocation lies on an interface. An arbitrary curve *C* encloses the dislocation.

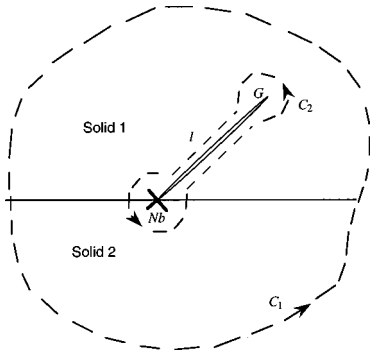


Fig. 3 A superdislocation on the interface, and a crack either on the interface or in one of the solids. The *M*-integral is evaluated over the two closed curves.

interface between a pair of solids (Fig. 2). Place the coordinate origin at the dislocation and a curve C around the dislocation. The M-integral is

$$M = \frac{1}{2\pi} \boldsymbol{b}^T \boldsymbol{H}^{-1} \boldsymbol{b}. \tag{2}$$

Here $\boldsymbol{H}$ is a positive-definite hermitian matrix, and can be calculated once the elastic constants of the two solids are given [8]. The location and the shape of the curve C do not affect the value of M. Furthermore, M is unaffected by the presence of an external boundary or other singularities in the solids, so long as C encloses no other singularities than the dislocation. This is understood by making C tightly surround the dislocation, where the singular stress field due to the dislocation prevails over the stress field due to other sources.

It has been shown that, for a pair of anisotropic solids with a fixed relative orientation, under an in-plane coordinate rotation, $\boldsymbol{H}$ transforms like a second-order tensor [9–11]. Consequently, if the interface rotates with the coordinate, but the Burgers vector $\boldsymbol{b}$ remains fixed relative to the solids, the number $\boldsymbol{b}^T\boldsymbol{H}^{-1}\boldsymbol{b}$ is invariant, and so is the value of M.

Figure 3 illustrates a flat interface between a pair of semi-infinite solids. The interface blocks N dislocations, each of the Burgers vector $\boldsymbol{b}$. A crack of length l lies either on the interface or in one of the solids. All the dislocations glide into the crack and blunt one crack tip. The other crack tip advances either on the interface or in one of the solids. We now calculate the energy release rate G at the advancing crack tip. Place the origin of the coordinates at the point where the dislocations are blocked. Because of the path-independence, the M-integral over any curve enclosing the dislocation-crack complex has the same value. First look at the curve C_1 far away from the complex. At a distance far from the dislocations, $r \to \infty$, all the dislocations behave collectively like a single superdislocation having the Burgers vector $N\boldsymbol{b}$; the stress field due to the superdislocation decays as $1/r$, but the modification due to the presence of the crack decays as $1/r^2$. Consequently, the M-integral over C_1 is the same as that over an isolated interface dislocation of the Burgers vector $N\boldsymbol{b}$. Equation (2) is applicable once $\boldsymbol{b}$ is replaced by $N\boldsymbol{b}$. Next look at the curve C_2 tightly surrounding the dislocation-crack complex. The stress field near the origin is now less singular than $1/r$, so that the small circle around the origin does not contribute to the M-integral. Nor do the traction-free crack faces. It can be shown that the M-integral over the small circle around the crack tip away from the origin equals lG [5]. Equating the M-integrals evaluated over C_1 and C_2, one obtains that

$$G = \frac{N^2}{2\pi l} \boldsymbol{b}^T \boldsymbol{H}^{-1} \boldsymbol{b}. \tag{3}$$

This is the desired result.

Discussion

For the case where the crack lies on the interface, Fan [4] solved the elasticity boundary value problem using a complex variable method, and calculated the energy release rate from the stress field. Equation (3) agrees with his result. Furthermore, this note demonstrates that the same equation is applicable to a crack inside one of the solids.

Let γ_1, γ_2, and γ_i be the surface energy per unit area of solid 1, solid 2, and the interface, respectively. When a crack advances a unit area, the surface energy increases by Γ, where $\Gamma = 2\gamma_1$ if the crack is in solid 1, or $\Gamma = \gamma_1 + \gamma_2 - \gamma_i$ if the crack is on the interface. Crack can nucleate if the energy release rate compensates the surface energy increase, $G = \Gamma$, namely,

$$\frac{N^2}{2\pi l} \boldsymbol{b}^T \boldsymbol{H}^{-1} \boldsymbol{b} = \Gamma. \tag{4}$$

The crack length l scales with N^2, everything else being fixed. It is also interesting to predict crack orientation from this model. The value of Γ depends on the crystalline orientation. As pointed out above, once the relative orientations of the two solids and the slip plane are fixed, the factor $\boldsymbol{b}^T\boldsymbol{H}^{-1}\boldsymbol{b}$ is invariant with the rotation of the interface or the crack. Consequently, when the relative orientations of the two solids and the slip plane are fixed, the crack orientation is entirely selected by the anisotropy in Γ. According to this model, anisotropy of elastic constants plays no role in selecting the crack orientation.

Acknowledgments

This note was prepared when the writer was supported by the National Science Foundation by a grant CMS-9820713.

References

[1] Zener, C., 1948, ''The Micro-Mechanism of Fracture,'' *Fracturing of Metals*, American Society of Metals, Cleveland, OH, pp. 3–31.
[2] Cottrell, A. H., 1975, ''Fracture,'' *The Physics of Metals, 2: Defects*, P. B. Hirsch, ed., Cambridge University Press, Cambridge, UK, pp. 247–280.
[3] Cherepanov, G. P., 1993, ''Interface Microcrack Nucleation,'' J. Mech. Phys. Solids, **42**, pp. 665–680.
[4] Fan, H., 1994, ''Interfacial Zener-Stroh Crack,'' ASME J. Appl. Mech., **61**, pp. 829–834.
[5] Freund, L. B., 1978, ''Stress Intensity Factor Calculations Based on a Conservative Integral,'' Int. J. Solids Struct., **14**, pp. 241–250.
[6] Knowles, J. K., and Sternberg, E., 1972, ''On a Class of Conservation Laws in Linearized and Finite Elastostatics,'' Arch. Ration. Mech. Anal., **44**, pp. 187–211.
[7] Rice, J. R., 1985, ''Conserved Integrals and Energetic Forces,'' *Fundamentals of Deformation and Fracture*, B. A. Bilby, K. J. Miller, and J. R. Willis, eds., Cambridge University Press, Cambridge, UK, pp. 33–56.
[8] Suo, Z., 1990, ''Singularities, Interfaces and Cracks in Dissimilar Anisotropic Media,'' Proc. R. Soc. London, Ser. A, **427**, pp. 331–358.
[9] Barnett, D. M., and Lothe, J., 1974, ''An Image Force Theorem for Dislocations in Anisotropic Bi-crystals,'' J. Phys. F: Met. Phys., **4**, pp. 1618–1635.
[10] Tucker, M. O., 1969, ''Plane Boundaries and Straight Dislocations in Elastically Anisotropic Materials,'' Philos. Mag., **19**, pp. 1141–1159.
[11] Ting, T. C. T., 1982, ''Effects of Change of Reference Coordinates on the Stress Analysis of Anisotropic Elastic Materials,'' Int. J. Solids Struct., **18**, pp. 139–152.

Alternative Derivation of Marguerre's Displacement Solution in Plane Isotropic Elasticity

X.-L. Gao[1]
Department of Aeronautics and Astronautics, Air Force Institute of Technology, 2950 P Street,
Wright-Patterson Air Force Base, OH 45433-7765

An alternative derivation of Marguerre's solution for displacements in plane isotropic elasticity is provided. It is shown that the present approach, which is based on Green's theorem and parallel to the Airy stress function approach, is straightforward. Also, the current derivation establishes the completeness of the Marguerre solution. [S0021-8936(00)00302-0]

1 Introduction

It is well known that there are two different formulation methods in plane isotropic elasticity, one of which is in terms of stress and the other in terms of displacement The stress formulation method using the Airy stress function has been fully developed and successfully applied to solve many problems with the first-kind boundary conditions (see, for example, Muskhelishvili [1], Teodorescu [2], and Gao [3]). However, limited attention has been paid to the displacement formulation method, whose application is always desired for problems with the second or third-kind boundary conditions.

The solution for displacements in terms of a biharmonic function was first derived by Marguerre [4] using an approach basically the same as Galerkin's three-dimensional reduction method based on the Helmholtz theorem (see, for example, Little [5], pp. 90–92). Marguerre's approach, as a two-dimensional specialization of the Galerkin method (see, for example, Barber [6], pp. 190–191) is concise, but it is not straightforward. Moreover, Marguerre's [4,7] approach appears to start with special forms and hence involves certain degrees of arbitrariness (Little [5]). This leaves the completeness of the Marguerre solution in doubt (and unaddressed).

The objective of this note is to provide an alternative derivation of Marguerre's solution using a straightforward approach. The present derivation, which makes use of Green's theorem and parallels the Airy stress function approach, also establishes the completeness of the Marguerre solution.

2 Derivation

Consider a homogeneous, isotropic elastic body undergoing infinitesimal plane-strain deformations in the usual Cartesian space $\{x_1,x_2,x_3\}$. This implies that $u_1=u_1(x_1,x_2)$, $u_2=u_2(x_1,x_2)$, $u_3=0$ and hence $\varepsilon_{31}=\varepsilon_{32}=\varepsilon_{33}=0$, $\sigma_{31}=\sigma_{32}=0$, $\sigma_{33}=\sigma_{33}(x_1,x_2)$. Then, the governing equations (in a displacement formulation) include the equilibrium equations (in the absence of body forces)

$$\frac{\partial\sigma_{11}}{\partial x_1}+\frac{\partial\sigma_{12}}{\partial x_2}=0,\quad \frac{\partial\sigma_{12}}{\partial x_1}+\frac{\partial\sigma_{22}}{\partial x_2}=0, \tag{1}$$

the constitutive equations

$$\sigma_{11}=\lambda(\varepsilon_{11}+\varepsilon_{22})+2G\varepsilon_{11},$$
$$\sigma_{22}=\lambda(\varepsilon_{11}+\varepsilon_{22})+2G\varepsilon_{22}, \tag{2}$$
$$\sigma_{12}=2G\varepsilon_{12},$$

and the geometrical equations

$$\varepsilon_{11}=\frac{\partial u_1}{\partial x_1},\quad \varepsilon_{22}=\frac{\partial u_2}{\partial x_2},\quad \varepsilon_{12}=\frac{1}{2}\left(\frac{\partial u_1}{\partial x_2}+\frac{\partial u_2}{\partial x_1}\right). \tag{3}$$

In Eqs. (1)–(3), σ_{ij}, ε_{ij} and u_i $(i,j\in\{1,2\})$ are, respectively, the in-plane stress, strain and displacement components, and λ, G are Lamé's constants defined by

$$\lambda\equiv\frac{E\nu}{(1+\nu)(1-2\nu)},\quad G\equiv\frac{E}{2(1+\nu)}, \tag{4}$$

with E and ν being Young's modulus and Poisson's ratio of the elastic material, respectively.

In a displacement formulation, $u_1=u_1(x_1,x_2)$ and $u_2=u_2(x_1,x_2)$ are regarded as the basic unknowns and are to be determined first. Using Eqs. (2) and (3) in Eq. (1) yields the Navier equations for the plane-strain case in linear elasticity (see, for example, Chou and Pagano [8], p. 73):

$$(\lambda+G)\left(\frac{\partial^2u_1}{\partial x_1^2}+\frac{\partial^2u_2}{\partial x_1\partial x_2}\right)+G\left(\frac{\partial^2u_1}{\partial x_1^2}+\frac{\partial^2u_1}{\partial x_2^2}\right)=0, \tag{5a}$$

$$(\lambda+G)\left(\frac{\partial^2u_1}{\partial x_1\partial x_2}+\frac{\partial^2u_2}{\partial x_2^2}\right)+G\left(\frac{\partial^2u_2}{\partial x_1^2}+\frac{\partial^2u_2}{\partial x_2^2}\right)=0. \tag{5b}$$

Equations (5*a*,*b*) are the basic governing equations for the homogeneous, isotropic elastic body considered in terms of displacements u_1 and u_2. This system of two partial differential equations can be solved simultaneously, which is the approach used by Marguerre [4,7]. Alternatively, the solution of this system can also be derived by solving one equation first and then enforcing the other (see, for example, Gao [9–11]). The latter approach will be adopted here to solve Eqs. (5*a*,*b*). For simplicity, it is assumed in the following derivation that the region occupied by the elastic body in the x_1x_2-plane is simply connected.

Note that Eq. (5*a*) may be rewritten as (cf. Wang [12])

$$\frac{\partial}{\partial x_1}\left[(\lambda+2G)\frac{\partial u_1}{\partial x_1}\right]=\frac{\partial}{\partial x_2}\left[-G\frac{\partial u_1}{\partial x_2}-(\lambda+G)\frac{\partial u_2}{\partial x_1}\right]. \tag{6}$$

Then, by an extended version of Green's theorem (see Appendix) there exists a function $\Lambda(x_1,x_2)$ such that

$$(\lambda+2G)\frac{\partial u_1}{\partial x_1}=\frac{\partial\Lambda}{\partial x_2},$$
$$-\left[G\frac{\partial u_1}{\partial x_2}+(\lambda+G)\frac{\partial u_2}{\partial x_1}\right]=\frac{\partial\Lambda}{\partial x_1}. \tag{7}$$

These two equations can also be written as

$$\frac{\partial}{\partial x_1}[(\lambda+2G)u_1]=\frac{\partial\Lambda}{\partial x_2}, \tag{8a}$$

$$\frac{\partial}{\partial x_1}[\Lambda+(\lambda+G)u_2]=\frac{\partial}{\partial x_2}(-Gu_1). \tag{8b}$$

Again, by the extended Green's theorem, Eq. (8*a*) implies that there is a function $\Gamma(x_1,x_2)$ such that

$$(\lambda+2G)u_1=\frac{\partial\Gamma}{\partial x_2}, \tag{9a}$$

$$\Lambda=\frac{\partial\Gamma}{\partial x_1}, \tag{9b}$$

and Eq. (8*b*) implies that there exists a function $\Omega(x_1,x_2)$ such that

[1]Current address: Department of Mechanical Engineering–Engineering Mechanics, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931-1295.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received and accepted by the ASME Applied Mechanics Division, Aug. 11, 1999; final revision, Nov. 22, 1999. Associate Technical Editor: J. R. Barber.

$$\Lambda+(\lambda+G)u_2=\frac{\partial\Omega}{\partial x_2}, \tag{10a}$$

$$-Gu_1=\frac{\partial\Omega}{\partial x_1}. \tag{10b}$$

Now solving Eq. (9*a*,*b*) and (10*a*,*b*) for u_1 and u_2 gives

$$\begin{aligned} u_1&=\frac{1}{\lambda+G}\left(\frac{\partial\Gamma}{\partial x_2}+\frac{\partial\Omega}{\partial x_1}\right),\\ u_2&=\frac{1}{\lambda+G}\left(\frac{\partial\Omega}{\partial x_2}-\frac{\partial\Gamma}{\partial x_1}\right). \end{aligned} \tag{11}$$

Next, eliminating u_1 from Eqs. (9*a*) and (10*b*) yields

$$\frac{\partial}{\partial x_1}[(\lambda+2G)\Omega]=\frac{\partial}{\partial x_2}(-G\Gamma). \tag{12}$$

This implies, once again by the extended Green's theorem, that there exists a function $F(x_1,x_2)$ such that

$$(\lambda+2G)\Omega=\frac{\partial F}{\partial x_2},\quad -GT=\frac{\partial F}{\partial x_1}. \tag{13}$$

Substituting Eq. (13) into Eq. (11) results in

$$\begin{aligned} u_1&=-\frac{1}{G(\lambda+2G)}\frac{\partial^2F}{\partial x_1\partial x_2},\\ u_2&=\frac{1}{G(\lambda+G)}\left(\frac{\partial^2F}{\partial x_1^2}+\frac{G}{\lambda+2G}\frac{\partial^2F}{\partial x_2^2}\right). \end{aligned} \tag{14}$$

These expressions determine the displacement components solely in terms of $F(x_1,x_2)$. Hence, $F(x_1,x_2)$ may be regarded as a displacement function (potential) (see, for example, Barber [6], pp. 189–197).

Clearly, since the two expressions in Eq. (14) are derived from Eq. (5*a*) alone, they must be substituted into the other basic governing equation, Eq. (5*b*), to obtain the defining equation for $F(x_1,x_2)$. Substituting Eq. (14) into Eq. (5*b*) yields

$$\frac{\partial^4F}{\partial x_1^4}+2\frac{\partial^4F}{\partial x_1^2\partial x_2^2}+\frac{\partial^4F}{\partial x_2^4}=0. \tag{15}$$

This biharmonic equation is the basic governing equation to solve for $F(x_1,x_2)$ in the plane-strain case considered. Equation (15) says that the displacement function, $F(x_1,x_2)$, is a biharmonic function. This means that the role of the displacement function here is similar to that of the Airy stress function in the standard stress formulation in plane elasticity, which solely defines the stress components and is also a biharmonic function.

To simplify the solution, define

$$\psi(x_1,x_2)\equiv\frac{1}{(\lambda+G)(\lambda+2G)}F(x_1,x_2). \tag{16}$$

Clearly, $\psi(x_1,x_2)$ still satisfies Eq. (15) and hence is a biharmonic function. Then, it follows from Eqs. (14) and (16) that

$$\begin{aligned} u_1&=-\frac{\lambda+G}{G}\frac{\partial^2\psi}{\partial x_1\partial x_2},\\ u_2&=\frac{\lambda+2G}{G}\frac{\partial^2\psi}{\partial x_1^2}+\frac{\partial^2\psi}{\partial x_2^2}. \end{aligned} \tag{17}$$

Using Eq. (4) in Eq. (17) gives

$$\begin{aligned} u_1&=-\frac{1}{1-2\nu}\frac{\partial^2\psi}{\partial x_1\partial x_2},\\ u_2&=\frac{2(1-\nu)}{1-2\nu}\frac{\partial^2\psi}{\partial x_1^2}+\frac{\partial^2\psi}{\partial x_2^2} \end{aligned} \tag{18}$$

as the displacement components in the plane-strain case in terms of the new biharmonic function $\psi(x_1,x_2)$ defined by Eq. (16), with only one material property ν being involved.

The two expressions given by Eq. (17) are identical to those of the in-plane displacement components initially derived by Marguerre [4].

For the plane-stress case, by invoking the formal equivalence between the plane-strain and plane-stress equations in plane isotropic elasticity (see, for example, Chou and Pagano [8], p. 93) it immediately follows from Eq. (18), after replacing ν by $\nu/(1+\nu)$, that

$$\begin{aligned} u_1&=-\frac{1+\nu}{1-\nu}\frac{\partial^2\psi}{\partial x_1\partial x_2},\\ u_2&=\frac{2}{1-\nu}\frac{\partial^2\psi}{\partial x_1^2}+\frac{\partial^2\psi}{\partial x_2^2} \end{aligned} \tag{19}$$

as the displacements in the case of plane stress. These are exactly the expressions of Marguerre's displacement solution in the plane-stress case derived by Little [5] using Galerkin's reduction method. This ends the plane-stress case and hence completes our derivation of Marguerre's solution for displacements in plane isotropic elasticity.

Clearly, the derivation presented above has also established that the Marguerre solution is complete (or general) in the sense that every solution of Eqs. (5*a*,*b*) may be represented in the form of Eq. (17) (see, for example, Gurtin [13]) in terms of a biharmonic (displacement) function. The reason for this is that the present approach is based on an extended Green's theorem (see Appendix), which provides the necessary and sufficient condition for the existence of the displacement function that defines the displacements in the form of Eq. (17). Therefore, the completeness of the solution given in Eq. (17) automatically follows (see, for example, Truesdell [14]).

As demonstrated, the present approach begins with solving the equilibrium equations and applies Green's theorem to introduce a displacement function. It is therefore concluded that the current approach parallels the Airy stress function approach, which also starts from the equilibrium equations and uses Green's theorem to establish the existence of Airy's stress function (see, for example, Chou and Pagano [8], p. 114). This implies that the approach used here, as an alternative of Marguerre's [4,7] approach based on the Helmholtz theorem for vector decomposition, is elementary by nature. Other approaches based on more advanced mathematical theories such as the differential operator theory may also be used to derive solutions for displacements in plane isotropic elasticity (see, for example, Raack [15] and Lurie and Vasiliev [16]). The completeness of Marguerre's displacement solution discussed above implies, however, that any solution so obtained should have the form listed in Eq. (17) (or Eq. (19)).

Appendix

An Extended Green's Theorem

Theorem: Suppose that P and Q are two functions of x_1 and x_2, with P, Q, $\partial P/\partial x_2$, $\partial Q/\partial x_1$ being continuous and single-valued at any point $(x_1,x_2)\in R$ (a simply connected region). Then, the necessary and sufficient condition for the existence of a potential function $f(x_1,x_2)$ satisfying $\partial f/\partial x_1=P$ and $\partial f/\partial x_2=Q$ is that $\partial P/\partial x_2=\partial Q/\partial x_1$.

Proof: Note that under the given conditions Green's theorem holds, i.e.,

$$\int\int_R\left(\frac{\partial Q}{\partial x_1}-\frac{\partial P}{\partial x_2}\right)dx_1dx_2\equiv\oint_\Gamma Pdx_1+Qdx_2, \tag{A1}$$

where Γ is the boundary of the simply connected region R.

By definition, the exact differential of $f(x_1,x_2)$ has the form

$$df(x_1,x_2)\equiv\frac{\partial f}{\partial x_1}dx_1+\frac{\partial f}{\partial x_2}dx_2=Pdx_1+Qdx_2. \qquad (A2)$$

Applying Green's theorem then gives, from Eqs. (*A*1) and (*A*2),

$$\int\int_R\left(\frac{\partial Q}{\partial x_1}-\frac{\partial P}{\partial x_2}\right)dx_1dx_2\equiv 0. \qquad (A3)$$

From the continuities of $\partial P/\partial x_2$ and $\partial Q/\partial x_1$, it immediately follows that

$$\frac{\partial Q}{\partial x_1}\equiv\frac{\partial P}{\partial x_2} \qquad (A4)$$

at any point $(x_1,x_2)\in R$. This proves the necessity. Next, we show that Eq. (*A*4) is also sufficient for the existence of $f(x_1,x_2)$. Again, applying Green's theorem yields, from Eqs. (*A*1) and (*A*4),

$$\oint_\Gamma Pdx_1+Qdx_2=0, \qquad (A5)$$

which implies that Pdx_1+Qdx_2 is the exact differential of a smooth function of the arguments x_1 and x_2. Label this function as $f(x_1,x_2)$. Then, it follows

$$P(x_1,x_2)dx_1+Q(x_1,x_2)dx_2\equiv df(x_1,x_2)\equiv\frac{\partial f}{\partial x_1}dx_1+\frac{\partial f}{\partial x_2}dx_2, \qquad (A6)$$

which immediately gives $\partial f/\partial x_1=P$ and $\partial f/\partial x_2=Q$. This ends the proof for the sufficiency.

The extended Green's theorem recorded and proved here was also listed in Chou and Pagano ([8], p. 114) (without proof). It was mentioned there that the theorem establishes the existence of Airy's stress function from the two equilibrium equations.

Acknowledgment

The author is very grateful to Prof. J. R. Barber and two anonymous reviewers for their insightful and constructive comments on an earlier version of this paper. The Appendix is added in response to a comment of one reviewer.

References

[1] Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff, Groningen, The Netherlands.

[2] Teodorescu, P. P., 1964, ''One Hundred Years of Investigations in the Plane Problem of the Theory of Elasticity,'' Appl. Mech. Rev., **17**, pp. 175–186.

[3] Gao, X.-L., 1996, ''A General Solution of an Infinite Elastic Plate With an Elliptic Hole Under Biaxial Loading,'' Int. J. Pressure Vessels Piping, **67**, pp. 95–104.

[4] Marguerre, V. K., 1933, ''Ebenes und achsensymmetrisches Problem der Elastizitätstheorie,'' Z. Angew. Math. Mech., **13**, pp. 437–438.

[5] Little, R. W., 1973, *Elasticity*, Prentice-Hall, Englewood Cliffs, NJ.

[6] Barber, J. R., 1992, *Elasticity*, Kluwer Academic, Dordrecht, The Netherlands.

[7] Marguerre, V. K., 1955, ''Ansätze zur Lösung der Grundgleichungen der Elastizitätstheorie,'' Z. Angew. Math. Mech., **35**, pp. 242–263.

[8] Chou, P. C., and Pagano, N. J., 1967, *Elasticity: Tensor, Dyadic, and Engineering Approaches*, Van Nostrand, Princeton, NJ.

[9] Gao, X.-L., 1998, ''A Mathematical Analysis of the Elasto-Plastic Plane Stress Problem of a Power-Law Material,'' IMA J. Appl. Math., **60**, pp. 139–149.

[10] Gao, X.-L., 1999, ''An Exact Elasto-Plastic Solution for the Plane Wedge Problem of an Elastic Linear-Hardening Material,'' Math. Mech. Solids, **4**, 289–306.

[11] Gao. X.-L., 2000, ''Two Displacement Methods for In-Plane Deformations of Orthotropic Linear Elastic Materials,'' submitted for publication.

[12] Wang, L., 1985, ''On General Solution of Problem of Elastic Shallow Thin Shells,'' Acta Mech. Sin., **17**, pp. 64–71.

[13] Gurtin, M. E., 1972, ''The Linear Theory of Elasticity,'' *Handbuch der Physik*, Vol VIa/2, C. Truesdell, ed., Springer-Verlag, Berlin, pp. 1–295.

[14] Truesdell, C., 1959, ''Invariant and Complete Stress Functions for General Continua,'' Arch. Ration. Mech. Anal., **4**, pp. 1–29.

[15] Raack, W., 1989, *Ebene Flächentragwerke, Scheiben*, Schriftenreihe Ebene Flächentragwerke, Band 2, Technische Universität Berlin Press, Berlin, Germany.

[16] Lurie, S. A., and Vasiliev, V. V., 1995, *The Biharmonic Problem in the Theory of Elasticity*, Gordon and Breach, Luxembourg.

The Carothers Paradox in the Case of a Nonclassical Couple

M. Paukshto

Professor, Institute for Mechanical Engineering Problems, RAS, V. O. Bolshoy Prospect 61, Saint Petersburg 199178, Russia

A. Pitkin

Graduate Student, Department of Mathematics, State University of Marine Technology, Lotsmanskaya Street 3, Saint Petersburg 190008, Russia

The Carothers solution for a wedge loaded by a concentrated couple at its vertex is known to be valid for the wedge angles $2\alpha<2\alpha^*\approx 257$ deg *only. Moreover, for* $\pi<2\alpha<2\alpha^*$ *it exists for antisymmetric loading only. The more realistic model of the concentrated couple of the arbitrary orientation is examined by the approach of Dundurs-Markenscoff. It is shown that the Carothers type solution holds for the edge angles* $2\alpha<\pi$. [S0021-8936(00)00402-5]

The Carothers Paradox in the Case of Nonclassical Moment

The study deals with the state of plane strain in the elastic wedge with free flanks loaded by a concentrated couple at its vertex. It is simulated by two concentrated forces. The model of such a nonrotation-invariant couple was used by Ya. S. Uflyand [1]. He has studied equilibrium of the infinite strip fixed on the flanks at the action of such couple in a center. There are three approaches to introduce a concentrated couple at the wedge vertex ([2]): (a) placing a compensating loading on an arc of radius a with a center in the vertex (the approach of Carothers), (b) placing a compensating loading on the wedge flanks on a segment of the length a from the vertex (the approach of Sternberg-Koiter) or (c) using a concentrated couple applied at an interior point placed on a distance a (the approach of Dundurs-Markenscoff). Then the parameter a tends to zero. Here the approach of Dundurs-Markenscoff will be applied.

The solution of the main problem is based on the solution of the subsidiary problem, when wedge flanks are rigidly fixed. X. Markenscoff and M. Paukshto [3] solved such problem for a classical couple. According to the approach [4] the expression for stresses in the problem for a wedge with free flanks is obtained from the solution of the problem with the rigidly fixed flanks by the limiting process as Poisson's ratio tends to unit.

A concentrated couple is defined in the following way: two parallel opposite directed forces of a magnitude P are applied to a body. The distance between them is equal to $2e$. Tending e to zero so that the product Pe was constant we have a concentrated couple $M=\lim_{e\to 0}2Pe$ [1] in the limit.

The corresponding polar displacements under moment action in an unbounded plane are as follows:

$$u_r=-\frac{M}{4\pi \mathrm{Gr}}[\sin 2\theta\cos 2\varphi-0.5\cos 2\theta\sin 2\varphi],$$

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 10, 1997; final revision, December 6, 1999. Associate Technical Editor: R. Abeyaratne.

$$u_\theta=\frac{M}{2\pi G(\chi+1)r}[(\cos^2\theta+\chi\sin^2\theta)\cos^2\varphi - 0.5(\chi-1)\sin 2\theta\sin 2\varphi+(\sin^2\theta+\chi\cos^2\theta)\sin^2\varphi], \quad (1)$$

where φ is the angle between the x-axis and the moment axis. Here G is the shear modulus and v is the Poisson's ratio.

Now if the components depending on the angle θ are discarded, a representation will be obtained for the classical ''invariant couple'' usually used at consideration of the Carothers problem:

$$u_r=0,$$

$$u_\theta=\frac{M}{4\pi \mathrm{Gr}}. \quad (2)$$

These expressions can be obtained also by consideration of the limit of four concentrated forces: $M=\lim_{e\to 0}4Pe$. The simulation of a concentrated couple by two concentrated forces complicates the solving process, but it corresponds more to the mechanic sense of a couple.

The following stresses take place in the wedge with the free flanks under the action of a concentrated couple in its vertex:

1 $0<\alpha<0.5\pi$

$$\sigma_r=\frac{4M\sin 2\theta}{(\sin 2\alpha-2\alpha\cos 2\alpha)r^2}[\cos^2\alpha\cos^2\varphi+\sin^2\alpha\sin^2\varphi],$$

$$\sigma_\theta=0,$$

$$\tau_{r\theta}=\frac{4M\sin 2\theta}{(\sin 2\alpha-2\alpha\cos 2\alpha)r^2}[\cos^2\alpha\sin^2\theta\cos^2\varphi + \sin^2\alpha\sin^2\theta\sin^2\varphi]. \quad (3)$$

2 $\alpha=0.5\pi$

At $\varphi=0$ and $\varphi=0.5\pi$, all components of stresses are defined by expression (3). For other values φ:

$$\sigma_r\to\infty,$$

$$\sigma_\theta=\frac{2M\cos^2\theta}{\pi r^2}\sin 2\varphi,$$

$$\tau_{r\theta}\to\infty.$$

3 $0.5\pi<\alpha<\alpha^*$

At $\varphi=0$ and $\varphi=0.5\pi$, all components of stresses are defined by expression (3) but at other values φ they are infinite.

4 $\alpha=\alpha^*$

At $\varphi=0$ and $\varphi=0.5\pi$:

$$\sigma_r\to\infty,$$

$$\sigma_\theta=\frac{M}{2\alpha^2\sin 2\alpha r^2}[\cos^2\alpha\cos^2\varphi(2\theta-\sin 2\theta) - \sin^2\alpha\sin^2\varphi(2\theta-\sin 2\theta)],$$

$$\tau_{r\theta}\to\infty.$$

At other values φ they are infinite.

5 $0.5\pi<\alpha<\alpha^*$

All components of stresses are infinite.

The obtained solution is compared with those for the classical couple [5,3]. The sum of solutions for noninvariant couples at $\varphi=\varphi_o$ and $\varphi=\varphi_o+0.5\pi$ (φ is arbitrary) gives the solution for the classical couple. Sternberg and Koiter showed that in the case of the classical couple the critical angle separated cases of finite and infinitely increased stresses $2\alpha=2\alpha^*$ (for antisymmetric loading only). Later S. M. Belonosov [6] noted that for $\pi\leqslant 2\alpha<2\alpha^*$ the Carothers solution ''has no physical sense'' for antisymmetric loading also because it is impossible to realize an antisymmetric loading accurately. For the couple considered in the study at its arbitrary orientation the critical angle is $2\alpha=\pi$ with the exception of two special cases at $\varphi=0$ and at $\varphi=0.5\pi$. Thus, at the consideration of a more realistic model of a concentrated couple, the solution of the Carothers problem is correct for $2\alpha<\pi$ only.

Acknowledgment

The work was supported by a grant provided by the Government of St. Petersburg and the Ministry of the Common and Professional Education of Russia (M97-2. 1K-550).

References

[1] Uflyand, Y. S., 1967, *The Integral Transforms in the Problems of Theory of Elasticity*, Gostechizdat, Leningrad (in Russian).

[2] Markenscoff, X., 1994, ''Some Remarks on the Wedge Paradox and Saint-Venant's Principle,'' ASME J. Appl. Mech., **61**, pp. 519–523.

[3] Markenscoff, X., and Paukshto, M., 1998, ''The Wedge Paradox and a Correspondence Principle,'' Proc. R. Soc. London, Ser. A, **454**, pp. 147–154.

[4] Markenscoff, X., and Paukshto, M., 1995, ''The Correspondence Between Cavities and Rigid Inclusions in Three-Dimensional Elasticity and the Cosserat Spectrum,'' Int. J. Solids Struct., **32**, pp. 431–438.

[5] Sternberg, E., and Koiter, V., 1958, ''The Wedge Under a Concentrated Couple: A Paradox in the Two-Dimensional Theory of Elasticity,'' ASME J. Appl. Mech., **25**, pp. 575–581.

[6] Belonosov, S. M., 1962, *The Basic Plane Static Problems of Theory of Elasticity for Simply Connected and Double Connected Domains*, Siberia Department AS USSR., Novosibirsk, Russia (in Russian).

On Eigenfrequencies of an Anisotropic Sphere

W. Q. Chen,[1] J. B. Cai, G. R. Ye, and H. J. Ding

Department of Civil Engineering, Zhenjiang University, Hangzhou 310027, P. R. China

This note presents exact frequency equations of two independent classes of vibrations of a spherically isotropic solid sphere with fixed boundary conditions. Numerical calculations are performed and comparison between two different materials is made. Some useful observations are obtained. [S0021-8936(00)00102-1]

Introduction

Spherically isotropic material is of particular importance because of its application in aerospace and nuclear technology ([1]). In addition, the latest geophysical results revealed that the Earth, in fact, should be modeled as a spherically isotropic inhomogeneous sphere with liquid nucleus ([2]). Chen [3] recently recalled the research history of spherically isotropic bodies. Although great achievements have been made to general solutions ([4,5]) and vibration theories ([6,7]), just as it was recently noticed by Schafbuch et al. [8] for an isotropic sphere, there is no work on the fundamental case of fixed displacement boundary conditions. The purpose of this brief note is to present the exact three-dimen-

[1]Corresponding Author. e-mail: caijb@ccea.zju.edu.cn

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 17, 1998; final revision, February 17, 2000. Associate Technical Editor: M. M. Carroll.

sional frequency equations of a spherically isotropic sphere with fixed boundary conditions, from which the ones for an isotropic sphere can be readily derived. As mentioned by Schafbuch et al. [8] the results can provide valuable and necessary information for integral equation representations of exterior domain elastodynamic problems and can be a benchmark to check numerical methods for elastic continuum modal analysis.

Basic Formulations

For a spherically isotropic medium, in the spherical coordinates (r,θ,ϕ), the linear constitutive relations can be expressed as follows ([4,6]):

$$\begin{cases} \sigma_{\theta\theta}=c_{11}s_{\theta\theta}+c_{12}s_{\phi\phi}+c_{13}s_{rr}, & \sigma_{r\theta}=2c_{44}s_{r\theta}, \\ \sigma_{\phi\phi}=c_{12}s_{\theta\theta}+c_{11}s_{\phi\phi}+c_{13}s_{rr}, & \sigma_{r\phi}=2c_{44}s_{r\phi}, \\ \sigma_{rr}=c_{13}s_{\theta\theta}+c_{13}s_{\phi\phi}+c_{33}s_{rr}, & \sigma_{\theta\phi}=(c_{11}-c_{12})s_{\theta\phi}, \end{cases} \tag{1}$$

where σ_{ij} and s_{ij} are the stress and strain tensors, respectively, and c_{ij} are the elastic stiffness constants. The equations of motion and the geometric relationships can be found in any textbook. To simplify the basic equations, Ding and Chen [7] employed three displacement functions to rewrite the displacement components as follows:

$$u_\theta=-\frac{1}{\sin\theta}\frac{\partial\psi}{\partial\phi}-\frac{\partial G}{\partial\theta},\quad u_\phi=\frac{\partial\psi}{\partial\theta}-\frac{1}{\sin\theta}\frac{\partial G}{\partial\phi},\quad u_r=w. \tag{2}$$

It is obvious that the present use of three displacement functions is much superior to the Helmholtz formulas wherein an additional equation for the vector potential function is required ([8]). It is also simpler than that of two potential functions employed by other authors ([4,6]). In fact, it seems natural that three displacement components are represented by three displacement functions so that it is easier to be understood. It has been shown in Chen [3] and Ding and Chen [7] that the substitution of Eq. (2) into the basic equations leads to an uncoupled partial differential equation and a coupled system of three such ones. On assuming

$$\psi=R\sum_{n=1}^{\infty}U_n(\xi)S_n^m(\theta,\phi)e^{i\omega t},\quad w=R\sum_{n=0}^{\infty}W_n(\xi)S_n^m(\theta,\phi)e^{i\omega t},$$

$$G=R\sum_{n=1}^{\infty}V_n(\xi)S_n^m(\theta,\phi)e^{i\omega t}, \tag{3}$$

where $\xi=r/R$ is the nondimensional radial variable and R is the radius of the sphere; $S_n^m(\theta,\phi)$ is the spherical harmonics; n and m are integers; and ω is the circular frequency, one can further transfer the partial differential equations to the ordinary ones with their solutions listed below ([3,7]):

$$U_n(\xi)=B_{n1}\xi^{-1/2}J_\eta(\Omega\xi),\quad (n\geqslant 1), \tag{4}$$

$$W_0(\xi)=C_{01}\xi^{-1/2}J_\nu(\Omega\xi f_4^{-1/2}),\quad (n=0), \tag{5}$$

$$W_n(\xi)=\sum_{j=1}^{2}C_{nj}W_{nj}(\xi),\quad V_n(\xi)=\sum_{j=1}^{2}C_{nj}V_{nj}(\xi),\quad (n\geqslant 1), \tag{6}$$

where J is Bessel function of the first kind; B_{n1} and C_{nj} are arbitrary constants; $\Omega^2=\rho\omega^2R^2/c_{44}$ is the nondimensional frequency; and W_{nj} and V_{nj} are convergent, infinite series in the variable ξ, which can be obtained by the matrix Frobenius method ([9]) and

$$\begin{cases} \eta^2=\dfrac{1}{4}[9+2(n^2+n-2)(f_1-f_2)]>0, \\ \nu^2=\dfrac{1}{4}+2(f_1+f_2-f_3)/f_4>0, \\ f_1=c_{11}/c_{44},\ f_2=c_{12}/c_{44},\ f_3=c_{13}/c_{44},\ f_4=c_{33}/c_{44}. \end{cases} \tag{7}$$

It is noted here that the limited value condition at $r=0$ has been satisfied by the solutions (4)–(6).

Frequency Equations

From Eqs. (2)–(6), one can derive the expressions of displacement without difficulty. Considering free vibration of the sphere with fixed boundary conditions, i.e., $u_r=u_\theta=u_\phi=0$ at $r=R$, one finds that the vibration can be divided into two classes, just as the case for isotropic sphere ([8]). Details are omitted for simplicity and the frequency equations for the both classes are given in the following:

The First Class

$$\tan(\Omega)=\Omega\quad \text{for}\quad n=1,$$

and

$$J_\eta(\Omega)=0\quad \text{for}\quad n>1. \tag{8}$$

The Second Class

$$J_\nu(\Omega f_4^{-1/2})=0\quad \text{for}\quad n=0,$$

and

$$W_{n1}(1)V_{n2}(1)-W_{n2}(1)V_{n1}(1)=0\quad \text{for}\quad n>0. \tag{9}$$

It is also noted here that the corresponding frequency equations for isotropic materials can be easily derived upon the following substitution:

$$c_{11}=c_{33}=\lambda+2\mu,\quad c_{12}=c_{13}=\lambda,\quad c_{44}=\mu \tag{10}$$

where λ and μ are the Lamé constants.

Numerical Results and Discussion

Because the forms of frequency equations of the first class ($n=1$ corresponds to the torsional mode) and of the second class when $n=0$ (the breathing mode) are very simple, the following numerical calculation will be performed only for the second class when $n>0$ (nonbreathing modes). Two materials will be considered: Material I is nearly isotropic like magnesium, while Material II is a hypothetical one exhibiting substantial anisotropy. The elastic constants of the two materials are available in Cohen and Shah [6] and the corresponding dimensionless values of f_i are given in Table 1.

Table 1 gives values of the nondimensional frequency Ω of the nonbreathing mode (the second class when $n>0$) for two materials mentioned above. It is also noted here that the results of Schafbuch et al. [8] for isotropic materials are reproduced and identical agreement is obtained. As in Schafbuch et al. [8], only modes with Ω less than 20 are presented in Table 1. It can be seen that for Material I, there are 43 nonbreathing modes, while for Material II, only 34 modes. As the isotropic case, the eigenfrequencies of each harmonic (n) and mode are interlaced, but there is an orderly increase in the eigenfrequency for the fundamental mode of each subsequent harmonic. In fact, the eigenfrequency of any ith mode of each harmonic increases with the increase of harmonic.

It should be mentioned that the integer m, which appears in the spherical harmonics and represents the nonaxisymmetric motion ($m\neq 0$) of the sphere is not included in the frequency equations. The explanation has been given by Silbiger [10] for a thin isotro-

Table 1 Nondimensional natural frequencies (Ω) of non-breathing modes

	Material I:		Material II:	
n	$f_1=3.64$ $f_3=1.32$	$f_2=1.60$ $f_4=3.76$	$f_1=20$ $f_3=2$	$f_2=12$ $f_4=2$
1	4.314791	6.503848	3.488849	6.881477
	9.663102	11.31541	10.09261	13.12653
	12.88867	15.99893	15.59743	16.52056
	17.67265	19.17840	19.42807	
2	6.211372	8.186220	6.564605	10.33649
	11.21155	13.71557	13.75748	16.79022
	14.57968	17.66929	19.25504	
3	7.786947	9.963098	8.068259	12.09205
	12.71925	15.77926	15.80303	19.17390
	16.45163	19.30718		
4	9.177424	11.75927	9.319468	13.50838
	14.21575	17.42921	17.40240	
	18.60310			
5	10.47039	13.50285	10.49346	14.80725
	15.73208	18.95204	18.81884	
6	11.70884	15.15068	11.63776	16.05675
	17.29190			
7	12.91345	16.68883	12.76782	17.28085
	18.89945			
8	14.09500	18.12956	13.88946	18.48923
9	15.25970	19.49673	15.00503	19.68633
10	16.41139		16.11584	
11	17.55267		17.22251	
12	18.68535		18.32548	
13	19.81076		19.42509	

pic spherical shell. His conclusion is still valid here for a spherically isotropic sphere with fixed boundary conditions.

By comparing our results for anisotropic materials with those of Schafbuch et al. [8] for isotropic ones, one can find the differences in the values of frequency and mode numbers (for example, with Ω less than 20) between them. That is really an important point that will affect the traditional design of sphere using isotropic metallic materials in various engineering applications.

Acknowledgments

This paper was supported by the National Natural Science Foundation of China and the Zhejiang Provincial Natural Science Foundation. The author (CWQ) would like to acknowledge the financial support from the Japanese Committee of Culture, Sports, Education and Science. He would like to express sincere thanks to Prof. Tadashi Shioya for his kind assistance.

References

[1] Maiti, M., 1975, ''Stress in Anisotropic Nonhomogeneous Sphere,'' J. Eng. Mech., **101**, pp. 101–108.
[2] Montagner, J. P., and Anderson, D. L., 1989, ''Constrained Reference Mantle Model,'' Phys. Earth Planet. Inter., **54**, pp. 205–227.
[3] Chen, W. Q., 1996, *Coupled Free Vibrations of Spherically Isotropic Hollow Spheres*, Ph.D. dissertation, Zhejiang University, Hangzhou, China (in Chinese).
[4] Hu, H. C., 1954, ''On the General Theory of Elasticity for a Spherically Isotropic Medium,'' Acta Sci. Sin., **3**, pp. 247–260.
[5] Chen, W. T., 1966, ''On Some Problems in Spherically Isotropic Elastic Materials,'' ASME J. Appl. Mech., **33**, pp. 539–546.
[6] Cohen, H., Shah, A. H., and Ramakrishnan, C. V., 1972, ''Free Vibrations of a Spherically Isotropic Hollow Sphere,'' Acustica, **26**, pp. 329–333.
[7] Ding, H. J., and Chen, W. Q., 1996, ''Nonaxisymmetric Free Vibrations of a Spherically Isotropic Spherical Shell Embedded in an Elastic Medium,'' Int. J. Solids Struct., **33**, pp. 2575–2590.
[8] Schafbuch, P. J., Rizzo, F. J., and Thompson, R. B., 1992, ''Eigenfrequencies of an Elastic Sphere With Fixed Boundary Conditions,'' ASME J. Appl. Mech., **59**, pp. 458–459.
[9] Ding, H. J., Chen, W. Q., and Liu, Z., 1995, ''Solutions to Equations of Vibrations of Spherical and Cylindrical Shells,'' Appl. Math. Mech., **16**, pp. 1–15.
[10] Silbiger, A., 1962, ''Non-axisymmetric Modes of Vibrations of Thin Spherical Shell,'' J. Acoust. Soc. Am., **34**, p. 862.

Torsion of a Viscoelastic Cylinder

R. C. Batra
Fellow ASME

J. H. Yu
Department of Engineering Science and Mechanics, M/C 0219, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

Finite torsional deformations of an incompressible viscoelastic circular cylinder are studied with its material modeled by two constitutive relations. One of these is a linear relation between the determinate part of the second Piola-Kirchhoff stress tensor and the time history of the Green-St. Venant strain tensor, and the other a linear relation between the deviatoric Cauchy stress tensor and the left Cauchy-Green tensor, its inverse, and the time history of the relative Green-St. Venant strain tensor. It is shown that the response predicted by the latter constitutive relation is in better agreement with the test data, and this constitutive relation is used to compute energy dissipated during torsional oscillations of the cylinder. [S0021-8936(00)00502-X]

Batra and Yu [1] recently studied the stress relaxation in an isotropic, incompressible, and homogeneous viscoelastic body deformed either in finite simple shear or finite simple extension. The material response was modeled by two constitutive relations, one linear in the history of the Green-St. Venant strain tensor $\mathbf{E}$ (e.g., see Christensen [2]) and the other linear in the history of the relative Green-St. Venant strain tensor $\mathbf{E}_t$ (e.g., see Bernstein, Kearsley, and Zapas [3] and Fosdick and Yu [4]). For each one of the two deformations studied, the former constitutive relation predicted that the tangent modulus (i.e., the slope of the stress-strain curve) is an increasing function of the strain but according to the latter constitutive relation, the tangent modulus is a nonincreasing function of the strain which agrees with the behavior observed experimentally for most materials (e.g., see Bell [5]). A similar result had been obtained by Batra [6] for two linear constitutive relations in isotropic finite elasticity. We note that both simple shearing and simple extension are homogeneous deformations and are universal in the sense that they can be produced by surface tractions alone in every elastic or viscoelastic body. Batra [7] has recently compared the response predicted by four linear constitutive relations for finite shearing, finite extension, biaxial loading, and triaxial loading of an isotropic elastic body.

Here we study finite torsional deformations of an incompressible, homogeneous, and isotropic viscoelastic circular cylinder. Even though these deformations are inhomogeneous, Ericksen [8] and Carroll [9] have shown that they are universal for elastic and viscoelastic bodies, respectively. In cylindrical coordinates, torsion of a circular cylinder is described by $r=R$, $\theta=\Theta+\kappa Z$, $z=Z$, where (r,θ,z) denote cylindrical coordinates of a point in the present configuration that occupied the place (R,Θ,Z) in the stress-free reference configuration, and κ is the angle of twist per unit length of the cylinder. Relative to an orthonormal set of bases, the physical components of the deformation gradient $\mathbf{F}$, the left Cauchy-Green tensor $\mathbf{B}$, and tensors $\mathbf{E}$ and $\mathbf{E}_t$ are given by

$$\mathbf{F}=\begin{bmatrix}1&0&0\\0&1&\kappa r\\0&0&1\end{bmatrix},\quad \mathbf{B}=\begin{bmatrix}1&0&0\\0&1+\kappa^2r^2&\kappa r\\0&\kappa r&1\end{bmatrix}, \tag{1}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Mar. 9, 1999; final revision, Feb. 1, 2000. Associate Technical Editor: K. T. Ramesh.

$$\mathbf{E}=\frac{1}{2}\begin{bmatrix}0&0&0\\0&0&\kappa r\\0&\kappa r&\kappa^2 r^2\end{bmatrix},$$

$$\mathbf{E}_t(\tau)=\frac{1}{2}\begin{bmatrix}0&0&0\\0&0&(\kappa(\tau)-\kappa(t))r\\0&(\kappa(\tau)-\kappa(t))r&(\kappa(\tau)-\kappa(t))^2r^2\end{bmatrix}. \quad (2)$$

We model the material by the following two constitutive relations (e.g., see Christensen [10], Bernstein et al. [3], and Fosdick and Yu [4])

$$\overline{\mathbf{T}}=-p\mathbf{1}+\rho\mathbf{F}\dot{\overline{\psi}}_{,\dot{\mathbf{E}}}\mathbf{F}^T, \quad (3a)$$

$$\mathbf{T}=-p\mathbf{1}+\rho\mathbf{F}(\psi_{,\mathbf{F}})^T, \quad (3b)$$

where a superimposed dot indicates the material time-derivative $\mathbf{C}=\mathbf{F}^T\mathbf{F}$, $\mathbf{B}=\mathbf{F}\mathbf{F}^T$, $2\mathbf{E}=(\mathbf{C}-\mathbf{1})$, $\overline{\psi}_{,\mathbf{E}}\equiv\partial\overline{\psi}/\partial\mathbf{E}$, $\psi_{,\mathbf{F}}\equiv\partial\psi/\partial\mathbf{F}$,

$$\rho\overline{\psi}=\int_{-\infty}^{t}\int_{-\infty}^{t}G_1(t-\tau,t-\eta)\frac{\partial tr\mathbf{E}(\tau)}{\partial\tau}\frac{\partial tr\mathbf{E}(\eta)}{\partial\eta}d\tau d\eta + \frac{1}{2}\int_{-\infty}^{t}\int_{-\infty}^{t}G_2(t-\tau,t-\eta)tr\left(\frac{\partial\mathbf{E}(\tau)}{\partial\tau}\frac{\partial\mathbf{E}(\eta)}{\partial\eta}\right)d\tau d\eta, \quad (4)$$

$$\rho\psi=\frac{1}{2}\beta_1 tr\mathbf{B}+\frac{1}{2}\beta_{-1}tr\mathbf{B}^{-1}+\int_{-\infty}^{t}g(t-\tau)\frac{\partial tr\mathbf{E}_t(\tau)}{\partial\tau}d\tau.$$

Here, $\mathbf{T}$ is the Cauchy stress tensor; p the hydrostatic pressure not determined by the deformation; ρ the mass density; ψ and ψ are specific (per unit mass) strain energy functionals; and $g(\cdot)$, $G_1(\cdot,\cdot)$, and $G_2(\cdot,\cdot)$ are material relaxation functions which are smooth, positive, and monotonically decreasing functions of time t. $G_1(\cdot,\cdot)$ and $G_2(\cdot,\cdot)$ satisfy $G_1(x,y)=G_1(y,x)$. The constants β_1 and β_{-1} satisfy $\beta_1>0$, $\beta_{-1}<0$. Substituting (4) into (3) yields, in physical components,

$$\overline{T}_{ij}=-p\delta_{ij}+\left(\delta_{KL}\int_{-\infty}^{t}2G_1(t-\tau,0)\frac{\partial E_{MM}(\tau)}{\partial\tau}d\tau + \int_{-\infty}^{t}G_2(t-\tau,0)\frac{\partial E_{KL}(\tau)}{\partial\tau}d\tau\right)F_{iK}F_{jL}, \quad (5a)$$

$$T_{ij}=-p\delta_{ij}+\beta_1 B_{ij}+\beta_{-1}B_{ij}^{-1}+\int_{-\infty}^{t}g(t-\tau)\frac{\partial E_{tij}(\tau)}{\partial\tau}d\tau, \quad (5b)$$

where δ_{ij} is the Kronecker delta. Here and below, quantities for the constitutive relation (3a) are indicated by a superposed bar. Constitutive relations (5a) and (5b) are more general than those studied by Batra and Yu [1].

Christensen [10] has analyzed the torsional deformations of a homogeneous viscoelastic cylinder made of material (5a). Following the same procedure or that given by Truesdell and Noll [11] for the torsion of an isotropic elastic cylinder, we determine the hydrostatic pressure and the components of the Cauchy stress tensor that satisfy the balance of linear momentum without body and inertia forces, and the boundary condition of null tractions on the mantle of the cylinder.

The stress components, T_{zz} and $T_{\theta z}$, have the expressions

$$\overline{T}_{zz}(t)=-\left(\frac{1}{4}\kappa^2(t)(a^4-r^4)(2F_1(t)+F_2(t)) + \frac{1}{2}\kappa(t)(a^2-r^2)F_3(t)\right)+r^2F_2(t), \quad (6a)$$

$$T_{zz}(t)=\left(\frac{\beta_1}{2}+\beta_{-1}\right)r^2\kappa^2(t)+\frac{1}{2}\int_{-\infty}^{t}g(t-\tau)r^2\frac{\partial}{\partial\tau}(\kappa(\tau) - \kappa(t))^2d\tau\beta_1\frac{a^2}{2}\kappa^2, \quad (6b)$$

$$\overline{T}_{\theta z}(t)=(2F_1(t)+F_2(t))\kappa(t)r^3+\frac{1}{2}rF_3(t), \quad (7a)$$

$$T_{\theta z}(t)=\kappa(t)r(\beta_1-\beta_{-1})+\frac{1}{2}\int_{-\infty}^{t}g(t-\tau)r\frac{\partial}{\partial\tau}(\kappa(\tau)-\kappa(t))d\tau, \quad (7b)$$

where a is the radius of the cylinder, and

$$F_\xi(t)=\frac{1}{2}\int_{-\infty}^{t}G_\xi(t-\tau,0)\frac{d\kappa^2(\tau)}{d\tau}d\tau,\quad \xi=1,2;$$

$$F_3(t)=\int_{-\infty}^{t}G_2(t-\tau,0)\frac{d\kappa(\tau)}{d\tau}d\tau. \quad (8)$$

We now consider a stress-relaxation test with $\kappa(t)=\kappa_0 h(t)$; $h(t)$ being the Heaviside unit step function. Noting that $2F_\xi(t)=G_\xi(t,0)\kappa_0^2$, and $F_3(t)=G_2(t,0)\kappa_0$ (e.g., see Christensen [10]), we obtain the following expressions for the resultant normal force, $N_z(t)$, and the resultant torque, $M_z(t)$, acting on a cross section of the cylinder.

$$\overline{N}_z(t)=-\frac{\pi\kappa_0^4a^6}{6}\left[G_1(t,0)+\frac{1}{2}G_2(t,0)\right],$$

$$N_z(t)=-\frac{\pi\kappa_0^2a^4}{4}[-2\beta_{-1}+\beta_1+g(t)], \quad (9)$$

$$\overline{M}_z(t)=\frac{\pi}{2}\kappa_0a^4\left[\frac{\kappa_0^2a^2}{3}(2G_1(t,0)+G_2(t,0))+\frac{1}{2}G_2(t,0)\right], \quad (10a)$$

$$M_z(t)=\frac{\pi}{2}\kappa_0a^4\left\{(\beta_1-\beta_{-1})+\frac{1}{2}g(t)\right\}. \quad (10b)$$

Recalling that g, G_1, G_2, and β_1 are positive and β_{-1} is negative, each constitutive relation predicts that a compressive axial force must be applied to the end faces of the cylinder in order to maintain its length. The average axial stress is proportional to $\kappa_0^4a^4$ and $\kappa_0^2a^2$ for the constitutive relations (5a) and (5b), respectively. Whereas M_z is a linear function of κ_0 for the constitutive relation (5b), it also depends upon κ_0^3 for the constitutive relation (5a).

We now compare average shear stress versus shear strain curves as predicted from these two constitutive relations without the experimental data of Lenoe et al. ([12], Fig. 3), and set $\kappa(t)=\dot{\kappa}t$, where $\dot{\kappa}$ is the torsional rate. Lenoe et al. assume that $G(t)=\Sigma_{i=0}^{3}\Psi_i e^{-\gamma_i t}$, where Ψ_i is the relaxation modulus and γ_i equals the reciprocal of the relaxation time. For the polyurethane rubber studied, they found that $\Psi_0=2.896$ MPa, $\Psi_1=0.387$ MPa, $\Psi_2=0.152$ MPa, $\Psi_3=0.689$ MPa, and $\gamma_0=0$ s^{-1}, $\gamma_1=0.001316$ s^{-1}, $\gamma_2=0.0050$ s^{-1}, $\gamma_3=0.002631$ s^{-1} provided a good fit to the test data. Recall that the average shear stress, $T_{\theta z}^m=\int_0^a 2rT_{\theta z}dr/a^2$. We assign following values to various material parameters:

$$G_1(t)=\frac{2\Psi_0(1+\nu)}{3(1-2\nu)},\quad \nu=0.49;\quad G_2(t)=\sum_{i=0}^{3}\Psi_i e^{-\gamma_i t};$$

$$\beta_1-\beta_{-1}=\Psi_0;\quad g(t)=\sum_{i=1}^{3}\Psi_i e^{-\gamma_i t}. \quad (11)$$

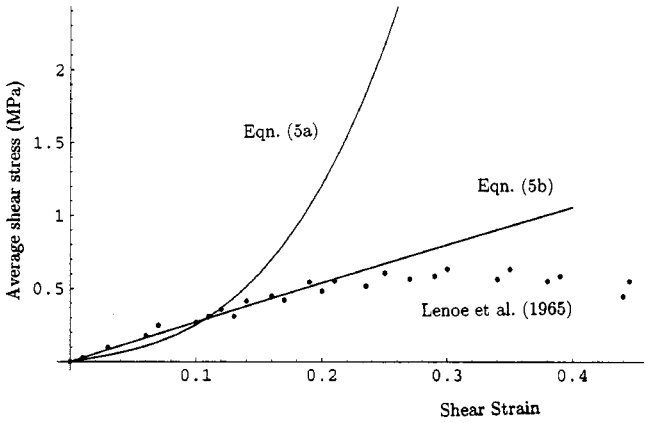


Fig. 1 Average shear stress versus shear strain curves computed from constitutive relations (5*a*) and (5*b*) and the test data of Lenoe et al. [12]. The test data is indicated as dots.

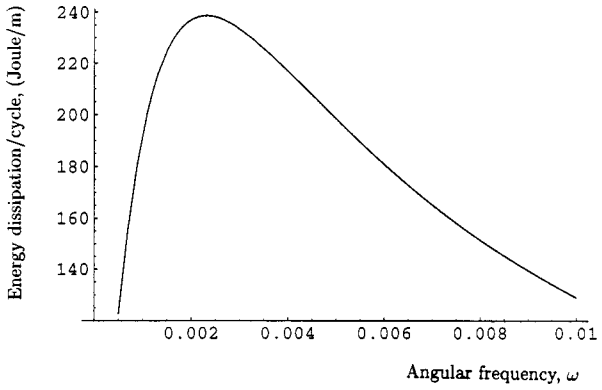


Fig. 3 Energy loss/cycle per unit length of the cylinder as a function of the forcing frequency for the polyurethane rubber tested by Lenoe et al. [12]

Figure 1 exhibits the average shear stress versus the shear strain curves for $\dot{\kappa}=0.0036\ \ s^{-1}$ as computed from constitutive relations (5*a*) and (5*b*), and also the experimental data of Lenoe et al. For shear strains up to 0.1, the three curves are close to each other. However, for large shear strains, the experimental curve is concave downwards but the ones obtained from constitutive relations (5*a*) and (5*b*) are concave upwards and nearly linear, respectively. For large shear strains, predictions from the constitutive relation (5*a*) are not even in qualitative agreement with the test data.

Henceforth we only use constitutive relation (5*b*) and analyze the damping of vibrations. We consider steady-state oscillations with $\kappa(t)=\kappa_0 \sin \omega t$, where ω, the frequency of oscillations, is such that inertia effects can be neglected (e.g., see Christensen [10]). The energy loss per cycle is given by $\Delta = \int_0^{2\pi/\omega} M(t)\dot{\kappa}(t)dt$ since there is no work done by N_z because of null axial elongation of the cylinder. For $g(t)=g_0 e^{-\gamma t}$ we obtain

$$\Delta(\gamma,\omega)=\frac{\pi g_0 \overline{\kappa}_0^2 a^4}{2}\cdot\frac{\pi\gamma\omega(\gamma^2+\omega^2)+\omega^2\gamma^2(e^{-2\pi\gamma/\omega}-1)}{(\gamma^2+\omega^2)^2}. \tag{12}$$

Whenever the term $e^{-2\pi\gamma/\omega}$ can be neglected, the energy loss will be a symmetric function of γ and ω. Figure 2 depicts the normalized energy loss $\Delta_n=4\Delta(\gamma,\omega)/\pi g_0\overline{\kappa}_0^2 a^4$ as a function of γ and ω. For $e^{-2\pi\gamma/\omega}\ll\gamma$ or ω, we see that the energy dissipation per cycle is maximum when $\gamma=\omega$. One possible explanation is that when the material relaxes faster than the frequency of the applied torque, i.e., $\gamma>\omega$, or when the material relaxes very slowly, i.e., $\gamma<\omega$, there will be a larger component of M_z in phase with κ during a part of the loading cycle which will decrease Δ. For $\gamma=0$, the material takes forever to relax, and there is no energy dissipation.

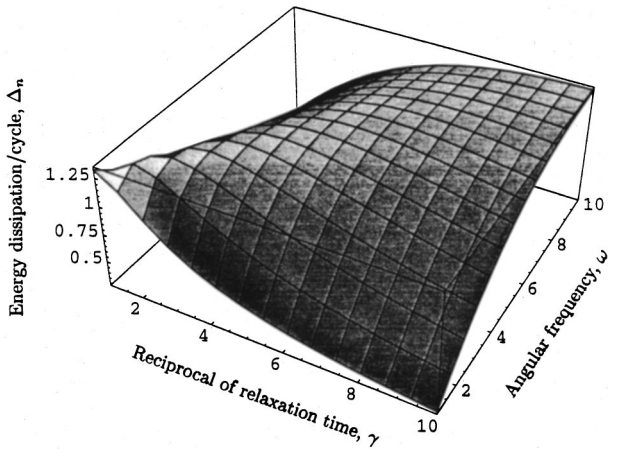


Fig. 2 The normalized energy loss/cycle per unit length of the cylinder as a function of the reciprocal of the relaxation time and the angular frequency

For a viscoelastic material with $g(t)=\Sigma_{i=1}^{3}\Psi_i e^{-\gamma_i t}$, the energy loss per cycle is

$$\Delta(\omega)=\frac{\pi\overline{\kappa}_0^2 a^4}{4}\sum_{i=1}^{3}\Psi_i\cdot\frac{\pi\gamma_i\omega(\gamma_i^2+\omega^2)+\omega^2\gamma_i^2(e^{-2\pi\gamma_i/\omega}-1)}{(\gamma_i^2+\omega^2)^2}, \tag{13}$$

For the aforestated values of material parameters and $\kappa_0=1$, $a=0.1$ m, the energy loss is plotted in Fig. 3. The energy loss is high for $0.002\leqslant\omega\leqslant0.004$. One can similarly find the optimum frequency range for other materials.

In conclusion, we note that the predictions from the constitutive relation (5*b*) are in better qualitative agreement with the test observations than those from the constitutive relation (5*a*). A real test of a constitutive relation is its ability to predict results in agreement with test data for configurations other than those used to find the values of material parameters. This arduous task has not been pursued here.

Acknowledgment

This work was supported by the ARO grant DAAG 55-98-1-0030 to Virginia Polytechnic Institute and State University.

References

[1] Batra, R. C., and Yu, J. H., 1999, ''Linear Constitutive Relations in Isotropic Finite Viscoelasticity,'' J. Elast., **55**, No. 1, pp. 73–77.
[2] Christensen, R. M., 1980, ''A Nonlinear Theory of Viscoelasticity for Applications to Elastomers,'' ASME J. Appl. Mech., **47**, pp. 762–768.
[3] Bernstein, B., Kearsley, E. A., and Zapas, L. J., 1963, ''A Study of Stress Relaxation With Finite Strain,'' Trans. Soc. Rheol., **7**, pp. 391–410.
[4] Fosdick, R. L., and Yu, J. H., 1998, ''Thermodynamics, Stability and Nonlinear Oscillations of Viscoelastic Solids—II. History Type Solids,'' Int. J. Non-Linear Mech., **33**, No. 1, pp. 165–188.
[5] Bell, J. F., 1973, ''The Experimental Foundation of Solid Mechanics,'' *Handbuch der Physik*, VIa/1, C. Truesdell, ed., Springer-Verlag, Berlin, pp. 1–813.
[6] Batra, R. C., 1998, ''Linear Constitutive Relations in Isotropic Finite Elasticity,'' J. Elast., **51**, pp. 243–245.
[7] Batra, R. C., 2000, ''Comparison of Results From Four Linear Constitutive Relations in Isotropic Finite Elasticity,'' Int. J. Nonlinear Mech., accepted for publication.
[8] Ericksen, J. L., 1954, ''Deformations Possible in Every Isotropic Incompressible Perfectly Elastic Body,'' Z. Angew. Math. Phys., **5**, pp. 466–486.
[9] Carroll, M. M., 1967, ''Controllable Deformations of Incompressible Simple Materials,'' Int. J. Eng. Sci., **5**, pp. 515–525.
[10] Christensen, R. M., 1968, ''On Obtaining Solutions in Nonlinear Viscoelasticity,'' ASME J. Appl. Mech., **35**, pp. 129–133.

[11] Truesdell, C., and Noll, W., 1965, *The Non-Linear Field Theories of Mechanics, Handbook of Physics*, Vol. III/3, Springer, New York.
[12] Lenoe, E. M., Heller, R. A., and Fruedenthal, A. M., 1965, ''Viscoelastic Behavior of a Filled Elastomer in the Linear and Nonlinear Range,'' Trans. Soc. Rheol., **9**, No. 2, pp. 77–102.

A Strip Element Method for Analyzing Wave Scattering by a Crack in an Axisymmetric Cross-Ply Laminated Composite Cylinder

Z. C. Xi

G. R. Liu

K. Y. Lam
Mem. ASME

H. M. Shang
Mem. ASME

Department of Mechanical and Production Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 11926

A strip element method is presented for analyzing waves scattered by a crack in an axisymmetric cross-ply laminated composite cylinder. The cylinder is at the outset discretized as axisymmetric strip elements through the radial direction. The application of the Hamilton variational principle develops a set of governing ordinary differential equations. The particular solutions to the resulting equations are found using a modal analysis approach in conjunction with the Fourier transform technique. The complementary solutions are formulated by the superposition of eigenvectors, the unknown coefficients of which are determined from axial stress boundary conditions at the tips of the crack. The summation of the particular and complementary solutions gives the general solutions. Numerical examples are given for cross-ply laminated composite cylinders with radial cracks. The results show that the present method is effective and efficient.
[S0021-8936(00)00202-6]

Introduction

Wave propagation in anisotropic media is one of the most fundamental and important subjects in the practice of engineering. Relevant literature is vast. Mal [1] and Nayfeh [2] reviewed it well. Because of the inherent complexities involved in material itself, an analysis of wave propagation in layered composite cylinders needs to resort to numerical techniques. Dealing with propagating waves and edge vibration in anisotropic composite cylinders, Huang and Dong [3] proposed an efficient numerical-analytical method in which a composite cylinder was modeled by finite elements, triangular functions, and wave function expansions in the radial, circumferential, and axial directions, respectively. The salient features of the method are to be capable of reducing the spatial dimensions of a problem by one and to omit tedious pre-processors occupying a substantial part of a finite element method. Rattanwangcharoen et al. [4] utilized the numerical-analytical method to solve the reflection problem of waves at the free edge of a laminated circular cylinder. Recently, Rattanwangcharoen et al. [5] combined the numerical-analytical method and the finite element method to analyze scattering of axisymmetric guided waves by a weldment between two laminated cylinders. In their treatise, the numerical-analytical method was employed to model the cylinders and the finite element method was used to model the weldment. This combinatory procedure was applied to axisymmetric guided wave scattering by cracks in welded steel pipes by Zhuang et al. [6]. The advantage of the combinatory procedure is to be able to treat complex local domains of a cylinder, such as weldment, hole, and imperfection. The disadvantage is to reduce the efficiency of the numerical-analytical method. Therefore, it is interesting to develop a numerical-analytical method for analyzing waves in a composite cylinder containing a crack.

In this paper, a strip element method is formulated for analyzing wave scattering by a crack in an axisymmetric cross-ply laminated composite cylinder, subjected to a harmonic excitation of a line source along the circumferential direction. The method is based on a strip element method proposed by Liu and Achenbach [7,8] for a cracked laminated composite plate as well as the numerical-analytical method proposed by Huang and Dong [3] for a perfect laminated composite cylinder. The cylinder is first modeled using axisymmetric strip elements in the radial direction. Then the Hamilton variational principle is used to derive a system of governing ordinary differential equations for the cylinder in a frequency domain. A particular solution for the resulting equations is found using a modal analysis approach and inverse Fourier transform techniques. A general solution is obtained with axial stress boundary conditions. Lastly, numerical examples are presented for multilayered cylinders with outer surface-breaking and radial interior cracks.

Formulation

Consider an infinitely long cracked cross-ply laminated composite cylinder made of an arbitrary number of linearly elastic cylinder-like laminae. The bonding between plies is perfect except in the region of the crack. Deformations of the cylinder are assumed small under a harmonic excitation. A radial line load of $q=q_0\exp(i\omega t)$ uniformly distributed along the circumferential direction is applied on the outer surface of the cylinder.

Because the geometry of the cylinder and the load are independent of the circumferential direction, the problem is axisymmetric. Let z and r denote, respectively, the axial and radial coordinates, then the strain-displacement relations are given by

$$\boldsymbol{\varepsilon}=\mathbf{L}\mathbf{u} \tag{1}$$

where $\boldsymbol{\varepsilon}[\varepsilon_z\ \varepsilon_\theta\ \varepsilon_r\ \gamma_{rz}]^T$ is the vector of strains and $\mathbf{u}=[u\ w]^T$ is the vector of displacements. Here u and w are the displacement components in the axial and radial directions, respectively. The operator matrix $\mathbf{L}$ is given by

$$\mathbf{L}=\begin{bmatrix} \dfrac{\partial}{\partial z} & 0 & 0 & \dfrac{\partial}{\partial r} \\ 0 & \dfrac{1}{r} & \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial z} \end{bmatrix}^T=\mathbf{L}_1\frac{\partial}{\partial z}+\mathbf{L}_2\frac{\partial}{\partial r}+\mathbf{L}_3\frac{1}{r} \tag{2}$$

where the matrices $\mathbf{L}_1$, $\mathbf{L}_2$, and $\mathbf{L}_3$ can be obtained by inspection of Eq. (2).

A lamina under consideration is transversely isotropic, so the stresses are related to strains by

$$\boldsymbol{\sigma}=\overline{\mathbf{Q}}\boldsymbol{\varepsilon} \tag{3}$$

where $\boldsymbol{\sigma}=[\sigma_z\ \sigma_\theta\ \sigma_r\ \tau_{rz}]^T$ is the vector of stresses and

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received and accepted by the ASME Applied Mechanics Division, May 3, 1999; final revision, December 7, 1999. Associate Technical Editor: A. K. Mal.

$$\overline{\mathbf{Q}}=\begin{bmatrix} \overline{Q}_{11} & \overline{Q}_{12} & \overline{Q}_{13} & 0 \\ \overline{Q}_{12} & \overline{Q}_{22} & \overline{Q}_{23} & 0 \\ \overline{Q}_{13} & \overline{Q}_{23} & \overline{Q}_{33} & 0 \\ 0 & 0 & 0 & \overline{Q}_{55} \end{bmatrix} \tag{4}$$

is the matrix of the off-principal-axis stiffness coefficients of the lamina whose expressions, in terms of engineering constants, are given by Vinson and Sierakowski [9].

A governing equation for the cylinder follows from the Hamilton variational principle which takes the form

$$\int_{t_0}^{t_1}\delta(V-T)dt=0 \tag{5}$$

where t_0 and t_1 are arbitrary time instants and V and T are the potential energy and kinetic energy of the cylinder, respectively. The potential energy of the cylinder in the absence of body forces is given by

$$V=\pi\int_{R_i}^{R_o}\varepsilon^T\sigma r dr-2\pi R_o q\overline{w} \tag{6}$$

where R_i and R_o are, respectively, the inner and outer radii of the cylinder, and $\overline{w}$ is the radial displacement at the loaded position which is expressible as $\overline{w}=\mathbf{nu}$. Here $\mathbf{n}=[0\ 1]$.

Substitution of Eqs. (1) to (3) into Eq. (6) gives

$$\begin{aligned} V=\pi\int_{R_i}^{R_o}\Bigg(&\frac{\partial\mathbf{u}^T}{\partial z}\mathbf{D}_{11}\frac{\partial\mathbf{u}}{\partial z}+\frac{\partial\mathbf{u}^T}{\partial z}\mathbf{D}_{12}\frac{\partial\mathbf{u}}{\partial r}+\frac{1}{r}\frac{\partial\mathbf{u}^T}{\partial z}\mathbf{D}_{13}\mathbf{u}\\ &+\frac{\partial\mathbf{u}^T}{\partial r}\mathbf{D}_{12}^T\frac{\partial\mathbf{u}}{\partial z}+\frac{\partial\mathbf{u}^T}{\partial r}\mathbf{D}_{22}\frac{\partial\mathbf{u}}{\partial r}+\frac{1}{r}\frac{\partial\mathbf{u}^T}{\partial r}\mathbf{D}_{23}\mathbf{u}+\frac{1}{r}\mathbf{u}^T\mathbf{D}_{13}^T\frac{\partial\mathbf{u}}{\partial z}\\ &+\frac{1}{r}\mathbf{u}^T\mathbf{D}_{23}^T\frac{\partial\mathbf{u}}{\partial r}+\frac{1}{r^2}\mathbf{u}^T\mathbf{D}_{33}\mathbf{u}\Bigg)rdr-2\pi\mathbf{u}^T\mathbf{n}^TR_oq \end{aligned} \tag{7}$$

where

$$\begin{aligned} &\mathbf{D}_{11}=\mathbf{L}_1^T\overline{\mathbf{Q}}\mathbf{L}_1\quad \mathbf{D}_{12}=\mathbf{L}_1^T\overline{\mathbf{Q}}\mathbf{L}_2\quad \mathbf{D}_{13}=\mathbf{L}_1^T\overline{\mathbf{Q}}\mathbf{L}_3\\ &\mathbf{D}_{22}=\mathbf{L}_2^T\overline{\mathbf{Q}}\mathbf{L}_2\quad \mathbf{D}_{23}=\mathbf{L}_2^T\overline{\mathbf{Q}}\mathbf{L}_3\quad \mathbf{D}_{33}=\mathbf{L}_3^T\overline{\mathbf{Q}}\mathbf{L}_3. \end{aligned} \tag{8}$$

The kinetic energy of the system is expressed in terms of the displacement vector as

$$T=\pi\int_{R_i}^{R_o}\frac{\partial\mathbf{u}^T}{\partial t}\frac{\partial\mathbf{u}}{\partial t}\rho r dr \tag{9}$$

where ρ is the mass density of the material.

Assume that the cylinder is divided into N axisymmetric strip elements in the radial direction and let R_{m-1} and R_m be, respectively, the inner and outer radii of any element m. Then the displacements within an element are approximated as

$$\mathbf{u}=\mathbf{N}(r)\mathbf{U}(z)\exp(i\omega t) \tag{10}$$

where $\mathbf{N}(r)$ is the shape function matrix of the element and $\mathbf{U}(z)$ is the vector of unknown displacement amplitudes.

Substituting Eqs. (7) and (9) into Eq. (5), in view of Eq. (10), and then taking variation with respect to $\mathbf{U}$ leads to the following governing ordinary differential equations for the cylinder

$$-\mathbf{A}_2\frac{d^2\mathbf{U}}{dz^2}+\mathbf{A}_1\frac{d\mathbf{U}}{dz}+(\mathbf{A}_0-\omega^2\mathbf{M})\mathbf{U}=\mathbf{q} \tag{11}$$

where

$$\begin{aligned} \mathbf{A}_0=\sum_{m=1}^{N}\int_{R_{m-1}}^{R_m}\Bigg(&\frac{\partial\mathbf{N}^T}{\partial r}\mathbf{D}_{22}\frac{\partial\mathbf{N}}{\partial r}+\frac{1}{r}\frac{\partial\mathbf{N}^T}{\partial r}\mathbf{D}_{23}\mathbf{N}\\ &+\frac{1}{r}\mathbf{N}^T\mathbf{D}_{23}^T\frac{\partial\mathbf{N}}{\partial r}+\frac{1}{r^2}\mathbf{N}^T\mathbf{D}_{33}\mathbf{N}\Bigg)rdr \end{aligned} \tag{12}$$

$$\begin{aligned} \mathbf{A}_1=\sum_{m=1}^{N}\int_{R_{m-1}}^{R_m}\Bigg(&-\mathbf{N}^T\mathbf{D}_{12}\frac{\partial\mathbf{N}^T}{\partial r}-\frac{1}{r}\mathbf{N}^T\mathbf{D}_{13}\mathbf{N}\\ &+\frac{\partial\mathbf{N}^T}{\partial r}\mathbf{D}_{12}^T\mathbf{N}+\frac{1}{r}\mathbf{N}^T\mathbf{D}_{13}^T\mathbf{N}\Bigg)rdr \end{aligned} \tag{13}$$

$$\mathbf{A}_2=\sum_{m=1}^{N}\int_{R_{m-1}}^{R_m}\mathbf{N}^T\mathbf{D}_{11}\mathbf{N}rdr$$

$$\mathbf{M}=\sum_{m=1}^{N}\int_{R_{m-1}}^{R_m}\mathbf{N}^T\mathbf{N}\rho rdr\quad \mathbf{q}=\mathbf{N}^T\mathbf{n}^TR_oq_0. \tag{14}$$

As can be seen from Eq. (11), the original two-dimensional problem has simplified to a one-dimensional one through the strip element idealization in the radial direction. Tedious preprocessors in finite element methods are omitted here. Consequently, computational labor can be greatly reduced. In addition, the strip element method requires a nodal line numbering along the r-axis only. This yields the minimum matrix bandwidth and therefore the strip element method requires much less computer memory and time compared with finite element methods.

Equation (11) is a set of nonhomogeneous ordinary differential equations in the frequency domain. A combination of modal analysis and inverse Fourier transform techniques gives its particular solution:

$$\mathbf{U}_p=\begin{cases} -i\displaystyle\sum_{m=1}^{M}\frac{\varphi_{m2}^{+L}\mathbf{q}_0R_o\varphi_{m1}^{+R}}{B_m^+}e^{-ik_m^+z}, & \text{for } z\geqslant 0\\ i\displaystyle\sum_{m=1}^{M}\frac{\varphi_{m2}^{-L}\mathbf{q}_0R_o\varphi_{m1}^{-R}}{B_m^-}e^{-ik_m^-z}, & \text{for } z<0 \end{cases} \tag{15}$$

where $B_m=\varphi_m^L\mathbf{B}\varphi_m^R$, k_m, φ_{m1}^L, φ_{m2}^L, φ_{m1}^R, and φ_{m2}^R are, respectively, the eigenvalues—the left and right eigenvectors obtained from the following characteristic equations:

$$\left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \omega^2\mathbf{M}-\mathbf{A}_0 & -i\mathbf{A}_1 \end{bmatrix}-k_m\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}\right)\begin{Bmatrix} \varphi_{m1}^R \\ \varphi_{m2}^R \end{Bmatrix}=\mathbf{0} \tag{16}$$

$$\begin{Bmatrix} \varphi_{m1}^L \\ \varphi_{m2}^L \end{Bmatrix}^T\left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \omega^2\mathbf{M}-\mathbf{A}_0 & -i\mathbf{A}_1 \end{bmatrix}-k_m\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}\right)=\mathbf{0} \tag{17}$$

The complementary solution of the associated homogeneous equation of Eq. (11) can be expressed by superposition of the right eigenvectors φ_m^R:

$$\mathbf{U}_c=\sum_{m=1}^{2M}C_m\varphi_m^R\exp(ik_mz)=\mathbf{G}(z)\mathbf{C} \tag{18}$$

where the subscript c denotes the complementary solution and the coefficient vector $\mathbf{C}$ is to be specified. The addition of the particular and complementary solutions yields the general solution of Eq. (11) in the form

$$\mathbf{U}=\mathbf{U}_c+\mathbf{U}_p=\mathbf{G}(z)\mathbf{C}+\mathbf{U}_p. \tag{19}$$

Thus the coefficient vector $\mathbf{C}$ can be expressed in terms of particular and general solutions at radial boundaries:

$$\mathbf{C}=\mathbf{G}_b^{-1}(\mathbf{U}_b-\mathbf{U}_{pb}) \tag{20}$$

where the subscript b denotes boundaries. Substitution of Eq. (20) into Eq. (19) gives

$$\mathbf{U}=\mathbf{G}(z)\mathbf{G}_b^{-1}(\mathbf{U}_b-\mathbf{U}_{pb})+\mathbf{U}_p\,. \tag{21}$$

The stress boundary conditions at the crack tips are given by

$$\mathbf{R}_b=\mathbf{K}\mathbf{U}_b+\mathbf{S}_p \tag{22}$$

where

$$\mathbf{R}_b=\begin{Bmatrix}\mathbf{R}_b^L\\ \mathbf{R}_b^R\end{Bmatrix}\qquad \mathbf{U}_b=\begin{Bmatrix}\mathbf{U}_b^L\\ \mathbf{U}_b^R\end{Bmatrix} \tag{23}$$

$$\mathbf{K}=\begin{bmatrix}\mathbf{R}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_1\end{bmatrix}+\begin{bmatrix}\mathbf{R}_2\dfrac{\partial\mathbf{G}^L}{\partial z}\mathbf{G}_b^{-1}\\ \mathbf{R}_2\dfrac{\partial\mathbf{G}^R}{\partial z}\mathbf{G}_b^{-1}\end{bmatrix} \tag{24}$$

$$\mathbf{S}_p=\begin{bmatrix}\mathbf{R}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_2\end{bmatrix}\begin{Bmatrix}\dfrac{\partial\mathbf{U}_p^L}{\partial\mathbf{z}}\\ \dfrac{\partial\mathbf{U}_p^R}{\partial z}\end{Bmatrix}-\begin{bmatrix}\mathbf{R}_2\dfrac{\partial\mathbf{G}^L}{\partial z}\mathbf{G}_b^{-1}\\ \mathbf{R}_2\dfrac{\partial\mathbf{G}^R}{\partial z}\mathbf{G}_b^{-1}\end{bmatrix}\begin{Bmatrix}\mathbf{U}_p^L\\ \mathbf{U}_p^R\end{Bmatrix} \tag{25}$$

are the external traction and displacement vectors on the radial boundaries, the stiffness matrix and the equivalent external force acting on the radial boundaries, respectively. The superscripts L and R represent the left and right sides of the crack, respectively.

Numerical Results and Discussion

In this section, numerical examples are given for $(C90/G0)_s$ and $(C0/G90)_s$ laminated composite cylinders. In the laminate codes, a lamina numbering increases from the inner to outer surface. The letters C and G represent carbon/epoxy and glass/epoxy, respectively. The number following the letters indicates the azimuthal angle of the fiber orientation with respect to the axis z. The subscript s denotes that the multilayered cylinders are symmetrically stacked. The material properties of the cylinders are taken from Takahashi and Chou [10]. Reference properties Q_{44} and ρ are the material constant and mass density of C0.

Figure 1 shows the distributions of the radial displacement on the outer surface of a $(C90/G0)_s$ composite cylinder with an outer surface-breaking crack. For comparison, the results for the corresponding uncracked case are also plotted in the same figure (dotted lines). It can be seen from the figure that the presence of the crack causes singularity of the displacement at the position of the crack. This phenomenon is of practical importance, from which the position of the crack can be clearly identified. It should be noted that the displacements at the left and right tips of the crack are discontinuous although this is not visible in the figure. Because of superposition of the incident and scattered wave fields, the absolute value of the displacement between the loaded point and the crack becomes irregular.

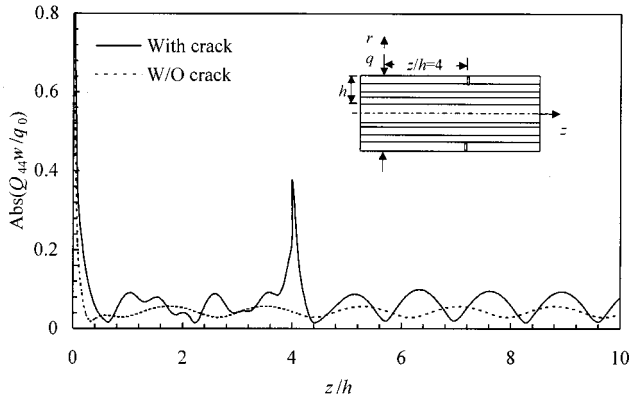


Fig. 1 Distribution of the radial displacement on the outer surface of a $(C90/G0)_s$ cylinder with an outer surface-breaking crack ($\omega h\sqrt{\rho/Q_{44}}=3.14$, $R_i/h=1$)

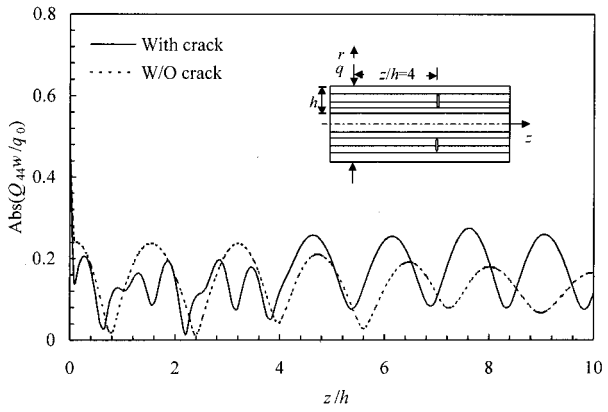


Fig. 2 Distribution of the radial displacement on the outer surface of a $(C0/G90)_s$ cylinder with a radial interior crack ($\omega h\sqrt{\rho/Q_{44}}=6.28$, $R_i/h=1$)

The present method may also be employed to detect a radial interior crack in a composite cylinder. Figure 2 shows the distributions of the radial displacement on the outer surface of a $(C0/G90)_s$ composite cylinder with a radial interior crack. It is apparent that when a wave strikes the radial interior crack in the cylinder, it generates scattering and causes an irregular oscillation of the absolute value of the displacement between the load point of application and crack. From the different patterns of wave fields at the left and right sides of the crack, the position of the interior crack can be easily identified although the crack is at a distance from the outer surface.

References

[1] Mal, A. K., 1988, ''Wave Propagation in Layered Composite Laminates Under Periodic Surface Loads,'' Wave Motion, **10**, pp. 257–266.

[2] Nayfeh, A. H., 1995, *Wave Propagation in Layered Anisotropic Media With Applications to Composites*, North-Holland, Amsterdam.

[3] Huang, K. H., and Dong, S. B., 1984, ''Propagating Waves and Edge Vibrations in Anisotropic Composite Cylinders,'' J. Sound Vib., **96**, pp. 363–379.

[4] Rattanwangcharoen, N., Shah, A., and Datta, S. K., 1994, ''Reflection of Waves at the Free Edge of a Laminated Circular Cylinder,'' ASME J. Appl. Mech., **61**, pp. 323–329.

[5] Rattanwangcharoen, N., Zhuang, W., Shah, A., and Datta, S. K., 1997, ''Axisymmetric Guided Waves in Jointed Laminated Cylinders,'' J. Eng. Mech., ASCE, **123**, pp. 1020–1026.

[6] Zhuang, W., Shah, A., and Datta, S. K., 1997, ''Axisymmetric Guided Wave Scattering by Cracks in Welded Steel Pipes,'' ASME J. Pressure Vessel Technol., **119**, pp. 401–406.

[7] Liu, G. R., and Achenbach, J. D., 1994, ''A Strip Element Method for Stress Analysis of Anisotropic Linearly Elastic Solids,'' ASME J. Appl. Mech., **61**, pp. 270–277.

[8] Liu, G. R., and Achenbach, J. D., 1995, ''Strip Element Method to Analyze Wave Scattering by Cracks in Anisotropic Laminated Plates,'' ASME J. Appl. Mech., **62**, pp. 607–613.

[9] Vinson, J. R., and Sierakowski, R. L., 1987, *The Behavior of Structures Composed of Composite Materials*, Martinus Nijhoff, Dordrecht, The Netherlands.

[10] Takahashi, K., and Chou, Tsu-Wei, 1987, ''Non-linear Deformation and Failure Behavior of Carbon/Glass Hybrid Laminates,'' J. Compos. Mater., **21**, May, pp. 396–407.

Journal of Applied Mechanics

Book Review

Thermal Stresses V, edited by R. B. Hetnarski. Lastran Corporation, Rochester, NY, 1999. 542 pages. Price: 135.00.

REVIEWED BY D. H. ALLEN[1]

This is the fifth volume in the series edited by R. B. Hetnarski on thermal stresses. While the first four volumes dealt with a variety of diverse issues, the current volume deals primarily with the issue of thermomechanically induced stresses and deformations in composite solids. The text is comprised of four chapters, each written by different authors, and dealing with somewhat different issues in composites. The first two chapters are quite closely related in scope, whereas the final two chapters deal with subjects that are related to the other two chapters, but may be treated separately.

The first chapter is written by C. T. Herakovich and J. Aboudi, both of whom are well known for their past research in composites, especially those that contain at least one inelastic phase. The thrust of this chapter is aimed primarily at stress analysis of metal matrix composites, with special emphasis on three aspects of this subject: micromechanics; lamination theory; and composite structures. Each of these subjects is treated in sufficient detail to be read without referring to additional materials. However, both the first and last emphasize only those methods that the authors were instrumental in formulating over the years.

In the case of micromechanics, the approach described by the authors is the method of cells. Particularly noteworthy in this chapter is the study of yielding of metal matrix composites under multiaxial stress states, a subject in which the authors are at the forefront of research.

The authors then review lamination theory for the case of transient temperatures. In the case of structures, the authors discuss primarily a methodology for modeling composite tubes.

This chapter can be considered to be an introduction to the subject of stress analysis of composites under transient temperature conditions, with emphasis on analytic methods of problem solution.

The second chapter is authored by K. K. Tamma and A. F. Avila. It covers the same issue as chapter one, namely, the prediction of thermally induced stresses in (primarily metal matrix) composites. However, the emphasis in this chapter is on computational techniques of stress analysis. The chapter opens with a lengthy and exhaustive review of computational methods for stress analysis in composites. This section is quite detailed and informative, especially for those who are new to the field. This is followed by two equally useful discussions of various micromechanics models, and several viscoplasticity models for metals. The section on computational methods of structural analysis introduces the finite element method for stress and deformation analysis of solids undergoing thermal transients. The chapter ends with several example calculations for both aluminum and titanium matrix composite structures such as blades in hot gas turbine engines. It is significant to note that such calculations were not possible as recently as twenty years ago. Today, such algorithms are available in numerous commercially available codes, and these calculations are performed routinely as a part of the design process. As such, this chapter presents a concise review of these advances in the last score of years.

The third chapter is written by R. Wojnar, S. Bytner, and A. Galka, all from the Polish Academy of Sciences in Warsaw. This chapter treats an entirely different subject from the first two chapters, dealing with the prediction of effective properties of heterogeneous media. While the emphasis in the current chapter is on homogenization techniques for thermally related properties, other properties such as mechanical, electric, and diffusive properties are also treated.

The scope of this chapter is much narrower than the previous two chapters. Whereas the opening chapters tend to give overviews of a broad field of research, this chapter considers the subject of homogenization theory in great detail. Indeed, to this reviewer's knowledge this is one of the two or three most detailed studies of this important issue in composites. Particularly noteworthy is the historical review that opens the chapter. The authors have taken great care to detail the important events on this subject dating back to the mid 19th century.

The interested reader will find a rigorous review of homogenization theory, and while the treatment is quite mathematical in nature, it is in a notation that is familiar to those who follow this area of research.

The final chapter in the text is written by N. Rajic, of the Aeronautical and Maritime Research Laboratory in Melbourne, Australia. This chapter deals with an important but quite dissimilar topic from the first three. The issue herein centers on the conversion of mechanical energy to heat during plastic deformations and/or crack growth in solids. While this subject was touched upon in previous volumes in this series, the current treatment is a welcome entry to the current text, and due to the emphasis on metal matrix composites in chapters one and two, is well placed in the current volume.

The author reviews the phenomenological description of the thermodynamics of plastic dissipation, and the model is utilized to predict the temperature rise due to a plastic zone near a circular hole in a metallic plate. This discussion is followed by a short section on the prediction of the temperature change induced in a plate by the energy dissipated when a crack runs in a ductile solid. While this chapter is somewhat shorter and less detailed than the others, it is nevertheless inciteful in its treatment of the subject of dissipation.

Despite the fact that there are differing authors, this volume does contain a common scope, which is often not the case in volumes of this type. Furthermore, the assemblage of all of this information into a single volume constitutes at the very least one of the most voluminous treatments of this subject heretofore seen by this reviewer. Thus, the volume would make a quite useful addition to both the reference and educational collections of scientists working in this field.

To order contact: Lastran Corporation, 78 Partridge Hill, Honeoye Falls, NY, 14472, or visit the page: www.lastran.com

[1]Department of Aerospace Engineering, Texas A&M University, College Station, TX 77843. Fellow ASME.

