Transactions of the ASME®

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published bimonthly (Jan., Mar., May, July, Sept., Nov.) The American Society of Mechanical Engineers, Three Park Avenue, New York, NY 10016. Periodicals postage paid at New York, NY and additional mailing office. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Applied Mechanics, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300. **CHANGES OF ADDRESS** must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address. **STATEMENT from By-Laws.** The Society shall not be responsible for statements or opinions advanced in papers or . . . printed in its publications (B7.1, Para. 3).
 INDEXED by Applied Mechanics Reviews and Engineering Information, Inc. Canadian Goods & Services Tax Registration #126148048.

Journal of Applied Mechanics

Published Bimonthly by The American Society of Mechanical Engineers

VOLUME 68 • NUMBER 3 • MAY 2001

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 68, Number 3 MAY 2001

BRIEF NOTES

ANNOUNCEMENTS AND SPECIAL NOTES

On the Behavior of Folded Tape-Springs

K. A. Seffen
Department of Mechanical Engineering,
University of Manchester Institute
of Science and Technology,
Sackville Street,
Manchester M60 1QD, UK
e-mail: keith.seffen@umist.ac.uk

Tape-springs are thin-walled beams with a curved cross section that can be elastically deformed to yield a flexible region of high curvature known as a fold. This feature is exploited in the folding and self-deployment of a number of recently proposed deployable structures. This study characterizes the quasi-static response of a folded tape-spring under a prescribed rotation and separation between its support points. It is shown that the corresponding end loads and fold shape are accurately predicted by a variational technique, and are confirmed by a finite element analysis. This information may then be used in further design of tape-spring hinge systems. [DOI: 10.1115/1.1365153]

1 Introduction and Background

This study deals with a specialized structural form of the everyday carpenter's tape known as a ''tape-spring.'' A tape-spring is a thin-walled, open cylindrical structure with a natural transverse curvature. It can be elastically deformed within its longitudinal plane to yield a well-defined elastic crease or *fold*. In order to properly characterize such behavior, moment-rotation relationships for symmetric bending are usually performed and are highly nonlinear and direction dependent ([1]). For example, if the direction of bending is in the *opposite sense* to the original transverse curvature, the spring exhibits a snap-through buckling to suddenly form a fold. During *equal-sense* bending—in the same direction as the transverse curvature—formation of a fold is more gradual. In both cases, the fold is connected on either side to relatively undeformed straight parts and resembles a continuous hinge.

When released, a folded tape-spring quickly returns to the straight configuration. This ability to self-actuate and the simplicity of form are exploited in the folding and self-deployment of several structures recently proposed for spacecraft applications, see Seffen [2] for an extensive review. Two examples are a rigid panel connected at its base to a spacecraft through rows of short, parallel tape-springs ([3,4]) and a membrane antenna with radially mounted tape-springs that are either discretely folded or wrapped around a central hub ([5–7]). At the end of deployment, the tape-springs ''lock'' into their undeformed shape to impart some stiffness to the foldable structure.

In dynamic laboratory tests, a folded tape-spring structure is restrained by a minimum of forces and simple release mechanisms ensure that the deployment is free from interference of the holding parts. In practice, extra loads may be applied to the structure *before* deployment, as given by the large inertia forces during launch or by over-constraining the packaged state. If measures are not taken to prevent transmission of these loads through the tape-spring ends to the more flexible fold regions, the application may prematurely unfold or some tape-springs may become damaged.

It would be difficult to ascertain the degree of distortion or damage without doing a detailed study of the overall structure; an appropriate starting point is to consider a single tape-spring. Moreover, the post-buckled response under symmetric end couples is well understood ([1,2,7]). Therefore, it is natural to extend study to a folded tape-spring also subject to end forces. To further simplify matters, only deformations within the longitudinal plane of symmetry are to be considered; the effect of lateral loads is not pursued. Thus, the aim of this paper is to formulate compact models of behavior, which may be used in subsequent design for stowed tape-spring structures, and the layout is as follows.

Section 2 describes some intuitive observations associated with end-loaded tape-springs and reviews previous work. In Section 3, a finite element analysis provides more accurate insight, and simplifying assumptions on the behavior are stated. An expression for the radius of curvature of fold in a tape-spring under end forces and end couples is derived by a variational method in Section 4. The performance is then compared to simulations from finite element analysis in Section 5. The study concludes in Section 6.

2 Fold Behavior

The geometry of an undeformed tape-spring is shown in Fig. 1(*a*). The transverse radius of curvature is R, the overall length is L, and the cross section has uniform thickness t and subtends angle α.

Under increasing equal and opposite end couples, a narrow region of the tape deforms into a fold connected on either side by transition or *ploy regions* to straight parts, and is shown schematically in Fig. 1(*b*). The formation process is analogous to the behavior in systems with *propagating instabilities* ([8]) and is discussed at length by Seffen [2] with reference to tape-springs.

Upon formation, the fold has a constant longitudinal radius of curvature, R^*, and zero transverse curvature irrespective of the relative rotation, θ, between the ends. A number of approximate, yet insightful studies ([9–11]) have suggested that R^* is equal to R. However, extensive finite element simulations for a range of tape-spring geometries ([2]) have shown that there is a marginal difference between these values and is assumed to be the case here. Note that θ is approximately the fold angle of embrace if the ploy regions are assumed to have negligibly small longitudinal curvature, Fig. 1(*b*).

The change in longitudinal curvature is denoted as $\kappa_l = 1/R^*$ and the transverse curvature change is κ_t equal to $0-(-1/R) = +1/R$ for opposite-sense bending and to $-1/R$ for equal-sense bending. Following Calladine [12], multiplying the flexural rigidity D by $\kappa_l + \nu\kappa_t$ and by the transverse arc-length $R\alpha$ provides an estimate of the bending moment, $M = M^*$, at *any* point in the fold as

$$M = DR\alpha[\kappa_l + \nu\kappa_t], \tag{1a}$$

$$\Rightarrow M^* = DR\alpha\left[\frac{1}{R^*} \pm \frac{\nu}{R}\right], \tag{1b}$$

where $D = Et^3/12(1-\nu^2)$, E is Young's modulus and ν is Poisson's ratio. Note that Eq. 1(*b*) is independent of the arc-length of fold, $R\theta$. Thus, as θ is increased or decreased under end couples $C = M^*$, Fig. 1(*b*), the fold grows or shrinks in arc-length. More

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 18, 2000; final revision, Dec. 6, 2000. Associate Editor: R. C. Benson. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

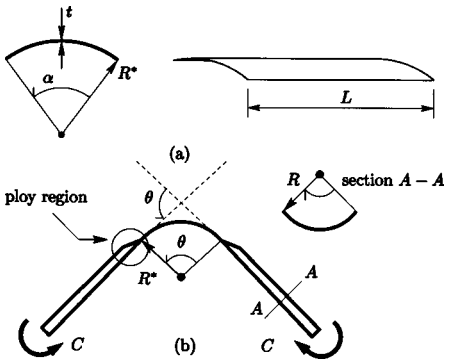


Fig. 1 Tape spring definitions: (*a*) undeformed geometry; (*b*) deformation with a single fold under end-wise couples *C*. The section view is end-on and the direction of bending is in the opposite sense to the transverse curvature.

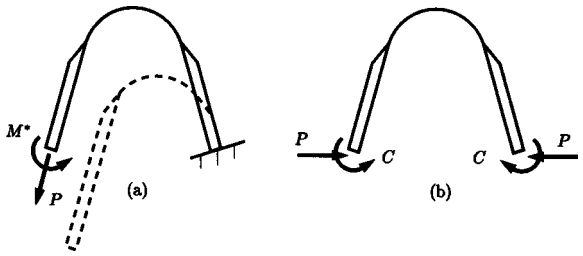


Fig. 2 Performance of fold under end loads: (*a*) end force applied to built-in tape-spring and the fold freely moves; (*b*) symmetric end loading and a folded tape-spring and the fold does not move

generally, the position of the fold along the tape and θ are not uniquely determined, provided θ is not so small that the tape snaps back to the straight configuration.

A relatively simple experiment that illustrates this indeterminacy has one end of a folded tape-spring clamped and a force P applied parallel to the other end, see Fig. 2(*a*). When θ is held fixed, the fold is observed to move towards the base as the tip translates under virtually zero force. Near to the support, the rigid base shape interacts with and arrests the motion of the fold. Subsequently, points within the fold are constrained to unnaturally deform with highest curvature on the base side, and P is no longer small.

A detailed study and confirmation of the above behavior is provided in Seffen and Pellegrino [1]. Importantly, if like loads persist at tape-spring ends in a more complex tape-spring structure, then it cannot be guaranteed that the packaged state remains intact.

When the fold is not near to a support, only loads that attempt to induce symmetric deformation about the fold midpoint may be carried in equilibrium. For an initially symmetric layout, the end forces are therefore co-linear, as shown in Fig. 2(*b*). In this case, the end couples are not necessarily equal to M^* nor is P small. This may not be entirely obvious, but is easily demonstrated again with a bent length of carpenter's tape. If the distance between the tape-spring ends is decreased by application of inwardly pointing forces, the relative rotation between the ends also increases if the end couple is held at M^*. However, if the end couple in the direction of rotation is reduced, so does the rotation. In either case, the bending moment at the fold center is greater than M^*, leading to an increased apex curvature above $1/R^*$. Conversely, if θ is fixed and the end forces repel one another, i.e., point away from each other, the midpoint bending moment is a minimum and may be less than M^*, causing the fold curvature to decrease. Under large forces, the fold can cleave apart into two separate folds, which then freely run along the tape in opposite directions to either end, and can be readily demonstrated.

In a folded tape-spring application, end loads that compress the folds in tapes may give rise to large changes of curvature and, possibly, plastic deformation; if the ends are being pulled apart, a fold may separate, thereby upsetting the packaged configuration. Therefore, it is imperative to quantify such behavior, in order to ensure safe stowage of the application during transportation as well as preservation of the proper shape required for accurate deployment. However, the problem is uniquely complicated by the interdependency of the large displacement deformation and the stress-resultants at any section within the tape-spring, and needs to be included in any analytical procedure. Previous studies on tape-springs referenced to thus far point to the success of finite element analyses in reproducing experimental behavior and is the route chosen here to provide, first, more detailed insight into the end-loaded behavior. This will permit simplifying assumptions to be made for a subsequent theoretical study.

3 Finite Element Analysis

The commercially available software package ABAQUS ([13]) is used to model the tape-spring as a mesh of S4R5 shell elements. Each element is quadrilateral with four corner nodes, has five degrees-of-freedom, and the displacement field within the element is bi-linear. Due to symmetry, only half the mesh in the transverse direction needs to be modeled. Typically, the total number of elements is 240, with 48 along the length of the tape, and the straight mesh is shown in Fig. 3(*a*). In addition, two reference nodes are defined at the centroid of cross section at each end and are connected to nodes in the same end of tape-spring by means of rigid, massless bars. The corresponding force and kinematic quantities at the reference nodes define the overall equilibrium response of the tape-spring. The material properties are linearly elastic and a small-strain, large-displacement behavior is assumed.

The tape-spring is deformed in two stages. First, a fold is formed under equal and opposite end couples. One of the reference nodes is fixed in position, but is permitted rotation about the axis of bending while the other node is free to translate and to

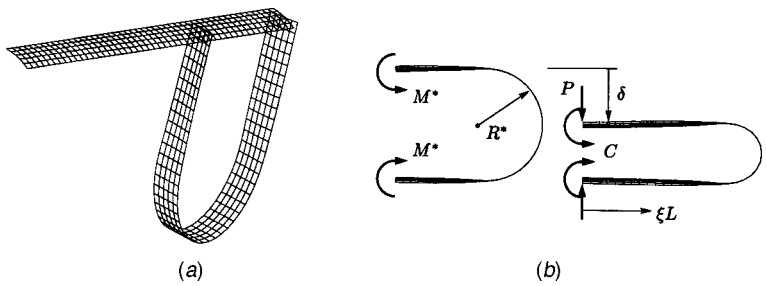


Fig. 3 Finite element model and deformation response: (*a*) original mesh and initial bending to form an opposite-sense fold; (*b*) tape-spring before and after compression by end forces with fixed end rotation

Table 1 Geometry and material properties of tape-spring, and initially folded geometry

R [mm]	α [rad]	t [mm]	L [$\times R\alpha$]	E [GPa]	ν [-]	R^* [mm]	θ [rad]
20.3	0.94	0.16	8	630	0.3	23.6	π

rotate. To capture the nonlinear snap-through phase, the solution procedure empowers an arc-length method to extract the equilibrium path from a generalized load-displacement space of the type discussed by Seffen and Pellegrino [1]. The final relative rotation between the ends is arbitrarily specified, and in Fig. 3(*a*) it is π radians.

At the start of the second stage, both reference nodes are rotationally constrained. A concentrated load is specified on the translating reference node in a direction towards or away from the opposite node, and is incrementally increased from zero using a standard static solution procedure from the ABAQUS library. Figure 3(*b*) indicates the deformation due to end forces that compress the fold with δ as the relative displacement between ends. The longitudinal radius of curvature has clearly decreased throughout the fold and, even for δ of the order of R^*, the straight parts are undeformed but increase in length and the fold region subtends the same angle.

The tape-spring geometry and material properties are listed in Table 1 and are typical of an everyday mild-steel carpenter tape, except for E, the Young's modulus, which is three times larger than normal. As δ increases, the end couples reverse direction leading to compression in the free edges of tape near to the ends. For values of δ greater than those reported here, local buckling took place at the ends for a lower value of E; by artificially increasing E, this is avoided and the deformation is confined to the folded region.

Also tabulated are θ and R^*, the latter value being computed at the end of the first deformation stage. Note that R^* differs from R by approximately 15 percent, thereby confirming that $R^* \neq R$ with this more accurate approach.

More insight is provided in Fig. 4(*a*), which indicates the variation in the ratio of total bending strain energy, U_B, to total stretching strain energy, U_S, with δ. The ratio is large suggesting that stretching effects are small in comparison to bending effects. The former can be attributed to the ploy regions where there is a small Gaussian curvature ([11]); as δ increases, the ploy regions move further away from the ends and their deformation remains unchanged but, overall, bending effects increase, as does U_B.

This is readily seen in Fig. 4(*b*), which plots the longitudinal curvature change along the tape when C is equal to M^* and at some stage during compression; a dimensionless intrinsic length parameter is ξ measured from the fixed end, see Fig. 3(*b*). The fold radius is everywhere constant under M^* but the change in curvature increases towards the midpoint during compression. This also confirms that, except for small changes in the ploy regions, the parts on either side of the fold have negligible curvature.

It is possible to report in detail on many other aspects of deformation but, for the sake of brevity, the assumptions they lead to are now summarized. Presuming to characterize behavior under fixed θ due to, first, M^* and to, second, P and C when δ is varied, then: the fold undergoes additional bending, symmetric about its midpoint and the transverse curvature is zero throughout the fold; the original straight parts remain straight, the fold angle of embrace and the relative end rotation are the same, and stretching effects can be neglected along with twist. Although these assumptions have been extracted from the behavior under end-wise compression, it can be expected that they also apply for the case of repelling end forces.

4 Theoretical Analysis

A theoretical model for the quasi-static response of a tape-spring to symmetric end-wise loading is now proposed. This is formulated in terms of potential energy expressions for the work done by the end loads and the strain energy stored in the folded region according to the *current* geometry of deformation; a simple, but effective variational approach is used to admit a closed-form expression for the longitudinal radius of curvature.

Consider Fig. 5. Initially, P and δ are equal to zero, M^* is the couple at both ends, and the fold has radius R^*. In the deformed configuration, P and δ are nonzero and the corresponding end couple for fixed θ is C. Point A is the midpoint of fold. The angle subtended by a general point, S, in the fold from the line of symmetry through A is β and $r=r(\beta)$ is its unknown longitudinal radius of curvature; β extends to values of $\theta/2$ on either side.

In the directions shown, P and C are defined as positive. Their current potential energy is simply $-P\delta$, since C does no work if the end rotation does not change. An expression for δ is now determined. Observe first that the original separation between the ends is X defined by the relationship

$$X=[L-R^*\theta]\cos\frac{\theta}{2}+2R^*\sin\frac{\theta}{2}. \tag{2}$$

The current separation is the distance x equal to

$$2\lambda L\cos\frac{\theta}{2}+\int_{-\theta/2}^{\theta/2} r\cos\beta d\beta, \tag{3}$$

where a dimensionless length for the straight parts including the ploy regions, λ, is calculated by subtracting the current arc-length of fold from the overall length, that is

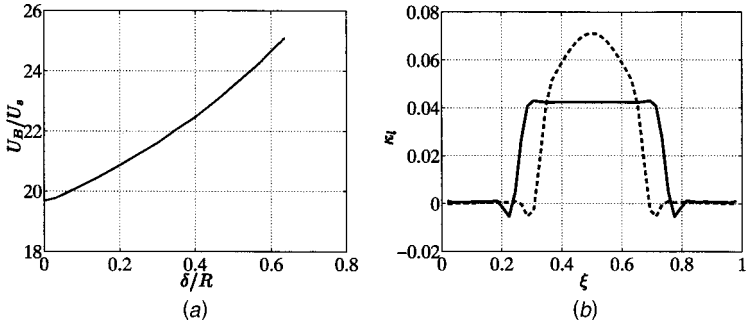


Fig. 4 Comparison of behavior for end-wise compression: (*a*) ratio of the bending strain energy to stretching strain energy in tape-spring during end-wise compression; (*b*) variation in longitudinal curvature along a compressed tape-spring: solid, constant bending moment M^*; dashed, compression under P

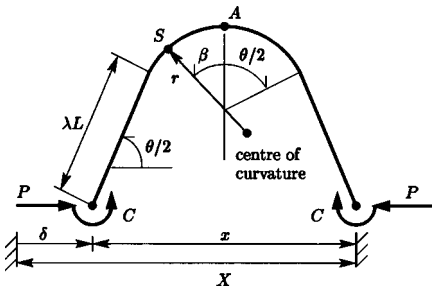


Fig. 5 Definition of parameters associated with deformation of a tape-spring

$$\lambda L = \frac{1}{2}\left[L - \int_{\beta=-\theta/2}^{\beta=\theta/2} r\, d\beta \right]. \tag{4}$$

Substituting Eq. (4) into Eq. (3) noting that $\delta = X - x$, subtracting the resulting expression from Eq. (2) and tidying up yields

$$\delta = R^*\left[2\sin\frac{\theta}{2} - \theta\cos\frac{\theta}{2}\right] + \int_{-\theta/2}^{\theta/2} r\left[\cos\frac{\theta}{2} - \cos\beta\right] d\beta. \tag{5}$$

For the fold, an expression for its potential strain energy of bending per unit surface area without twist is ([12])

$$\frac{D}{2}[\kappa_l^2 + \kappa_t^2 + 2\nu\kappa_l\kappa_t], \tag{6}$$

where the longitudinal curvature change is now equal to $1/r$; recall $\kappa_t = \pm 1/R$ for opposite and equal-sense bending, respectively. Integrating Eq. (6) over the surface area of fold yields for opposite-sense bending

$$\frac{D}{2}\int_{-\theta/2}^{\theta/2}\left[\frac{1}{r^2} + \frac{1}{R^2} + \frac{2\nu}{rR}\right] R\alpha r\, d\beta. \tag{7}$$

Multiplying $-P$ by Eq. (5) and adding to Eq. (7) results in the total potential energy expression for the deformed tape-spring as

$$\int_{-\theta/2}^{\theta/2}\left[\frac{DR\alpha}{2}\left(\frac{1}{r} + \frac{r}{R^2} + \frac{2\nu}{R}\right) - Pr\left(\cos\frac{\theta}{2} - \cos\beta\right)\right] d\beta + PR^*\left[2\sin\frac{\theta}{2} - \theta\cos\frac{\theta}{2}\right]. \tag{8}$$

In order to admit a solution for $r(\beta)$, Eq. (8) needs to be minimized subject to the initial conditions $P = 0$ when $\delta = 0$ and $r = R^*$. Since δ is also a function of r in Eq. (5), the required approach follows a procedure found in Hestenes [14] for minimizing simple integrals with *integral side conditions*. Note that although C does not appear explicitly in Eq. (8) it reenters implicitly by fixing θ; an expression for C is derived later.

The total potential energy is, first, re-expressed as a function J_0 of r in terms of the constant and integrand expressions as follows:

$$J_0(r) = c_0 + \int_{-\theta/2}^{\theta/2} \Omega_0(r(\beta))\, d\beta. \tag{9}$$

Comparing to Eq. (8) then

$$c_0 = PR^*\left[2\sin\frac{\theta}{2} - \theta\cos\frac{\theta}{2}\right], \tag{10a}$$

$$\Omega_0 = \frac{DR\alpha}{2}\left[\frac{1}{r} + \frac{r}{R^2} + \frac{2\nu}{R}\right] - Pr\left[\cos\frac{\theta}{2} - \cos\beta\right]. \tag{10b}$$

Note that Ω_0 must be a continuous function and obviates the need for $r(\beta)$ to be smooth, as would be expected for a fold with no sharp changes in geometry.

Similarly, to account for the initial conditions, it is necessary to define another function J_1 of r using Eq. (5), that is

$$J_1(r) = c_1 + \int_{-\theta/2}^{\theta/2} \Omega_1(r(\beta))\, d\beta = 0, \tag{11}$$

with

$$c_1 = R^*\left[2\sin\frac{\theta}{2} - \theta\cos\frac{\theta}{2}\right], \tag{12a}$$

$$\Omega_1 = r\left[\cos\frac{\theta}{2} - \cos\beta\right]. \tag{12b}$$

Since J_0 and J_1 are both functions of $r(\beta)$, their integrand parts can be combined into a single composite function, $F(r)$: the constants c_0 and c_1 do not have any bearing on the variational procedure which follows.

This composite function needs to be strictly defined by

$$F = \gamma_0\Omega_0 + \gamma_1\Omega_1, \tag{13}$$

and the variation with r is implied. γ_0 and γ_1 are nonzero, positive multipliers that are either constants or functions determined from minimizing F on the initial conditions and on nonzero P. Using Eqs. 10(b) and 12(b), Eq. (13) is fully written as

$$F = \gamma_0\left[\frac{DR\alpha}{2}\left(\frac{1}{r} + \frac{r}{R^2} + \frac{2\nu}{R}\right) - Pr\left(\cos\frac{\theta}{2} - \cos\beta\right)\right] + \gamma_1 r\left[\cos\frac{\theta}{2} - \cos\beta\right]. \tag{14}$$

Before either multiplier can be determined, some sanity checks are essential. If γ_0 is equal to zero, then

$$\frac{\partial F}{\partial r} = \gamma_1\left[\cos\frac{\theta}{2} - \cos\beta\right] = 0 \tag{15}$$

can never have a minimum for $\gamma_1 > 0$. Likewise, if $\gamma_1 = 0$ then

$$\frac{\partial F}{\partial r} = \gamma_0\left[\frac{DR\alpha}{2}\left(-\frac{1}{r^2} + \frac{1}{R^2}\right) - P\left(\cos\frac{\theta}{2} - \cos\beta\right)\right] = 0 \tag{16}$$

admits for zero end force $r = R$ and is invalid, and γ_1 cannot be zero.

The simplest function for either multiplier is a nonzero constant. Choosing a value of γ_0 equal to 1, substituting into Eq. (14), differentiating the resulting expression with respect to r, and setting equal to zero leads to

$$\frac{\partial F}{\partial r} = \frac{DR\alpha}{2}\left[-\frac{1}{r^2} + \frac{1}{R^2}\right] - P\left[\cos\frac{\theta}{2} - \cos\beta\right] + \gamma_1\left[\cos\frac{\theta}{2} - \cos\beta\right] = 0. \tag{17}$$

Substituting for the initial conditions, $P = 0$ and $r = R^*$, and solving for γ_1 explicitly gives

$$\gamma_1 = \left[\frac{DR\alpha}{2}\left(-\frac{1}{R^2} + \frac{1}{R^{*2}}\right)\right]\cdot\left[\cos\frac{\theta}{2} - \cos\beta\right]^{-1}. \tag{18}$$

An expression for r is obtained by replacing γ_1 in Eq. (17) with Eq. (18) and rearranging to give

$$r(\beta) = R^*\cdot\left[1 + \frac{2PR^{*2}}{DR\alpha}\left(\cos\beta - \cos\frac{\theta}{2}\right)\right]^{-1/2}. \tag{19}$$

The above expression can be shown to be the same for equal-sense bending and $r = R^*$ when P is equal to zero.

Finally, satisfaction of moment equilibrium about point A in Fig. 5 yields for the end couple

$$C = M_A - P\left(\lambda L \cos\frac{\theta}{2} + \int_0^{\theta/2} r \sin\beta d\beta\right). \tag{20}$$

The bending moment M_A at A is derived from Eq. 1(a) with $\kappa_l = 1/r$ for $\beta = 0$ and λ is determined from Eq. (4). Note that Eq. (20), upon substitution of r with Eq. (19), cannot be integrated to yield a closed-form expression for C; the process needs to be tackled by numerical means, as is now performed in comparison to finite element results.

5 Results and Discussion

The initial geometry and material properties are the same as in Table 1, and the tape has been folded with opposite-sense curvatures. The results from two tests are now presented for, first, P increasingly positive and, second, P decreasingly negative. In so doing, note that the quotient term inside parentheses in Eq. (19) is dimensionless and it is convenient to normalize force (and moment) in a similar way. However, since R^* depends on finite element analysis, it is replaced with the natural parameter R, leading to the following ''barred'' dimensionless groups:

$$\overline{P} = \frac{PR}{D\alpha}, \quad \overline{C} = \frac{C}{D\alpha}. \tag{21}$$

Likewise, δ is normalized by dividing by R, i.e., $\overline{\delta} = \delta/R$ and a measure of dimensionless longitudinal curvature compares the current value to the initial fold curvature by defining $\overline{\kappa_l} = R^*/r$.

(a) End-Wise Compression. Figure 6(a) indicates the (P,δ) behavior. The theoretical curve was computed by, first, specifying a piece-wise variation in P up to the maximum value from finite element analysis. For each value of P, the constant terms in r, Eq. (19), are known and the resulting expression is substituted into Eq. (5), which is numerically integrated over the range $(-\pi,\pi)$ via the software package MATLAB [15] to yield the corresponding value of δ.

The correlation between theory and the exact solution is initially close but diverges as δ increases. The performance is nonetheless encouraging, even at large end displacements, given the assumptions of the theoretical model.

A number of snap-shots of the change in longitudinal curvature are presented in Fig. 6(b) according to the indicated values of P. The lengths of the straight parts, as given by Eq. (4), were first calculated to define the regions either side of the fold with assumed zero curvature change; the end points of these lines then discontinuously connect to the variation due to $1/r$ from Eq. (19). As can be clearly seen, the difference between theory and finite element results is virtually negligible in most of the fold.

The variation of P with end-couple C is plotted in Fig. 7(a). The numerical procedure for calculating C in Eq. (20) was similar to that employed for δ. This figure shows that as P increases, in order to maintain a fixed relative rotation between the ends, C decreases and at $\overline{P} = 0.487$ the end couple is zero. The corresponding end displacement from Fig. 6(a) is approximately $\delta = 0.36R$. Since the angle of fold arc-length is π, an ''average'' radius of fold is $(2R^* - 0.36R)/2$, which is less then R^*. Thus, it is incorrect to presume that R^* can be simply found in this way, for example, when holding the ends of tape between the tips of fingers and measuring the separation distance. The subsequent value of M^*, Eq. 1(b), may also be inaccurate. This may seem trivial to highlight, however, dynamic models of the deployment of tape-spring systems ([1]) equate M^* to the torque applied by the tape-spring fold to the rest of the structure and, therefore, needs to be accurately measured. At larger values of P, the end couple acts in the opposite direction, as discussed in Section 3.

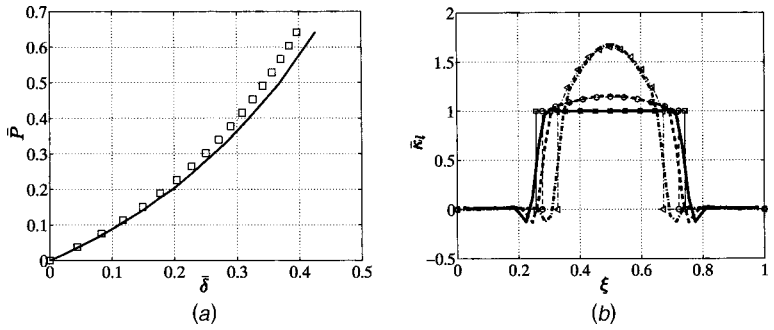


Fig. 6 Comparison of behavior for end-wise compression: (*a*) force-displacement response of a bent tape-spring with geometry in Table 1. The continuous line is the finite element analysis result and the squares denote the theoretical prediction. (*b*) Variation in longitudinal curvature change, $\kappa_l = 1/r$, within the fold for $P = 0N$ (solid), $P = 1.266N$ (dashed), and $P = 7.020N$ (dashed-dot). FEA given by thick lines, theory as symbols and thin lines.

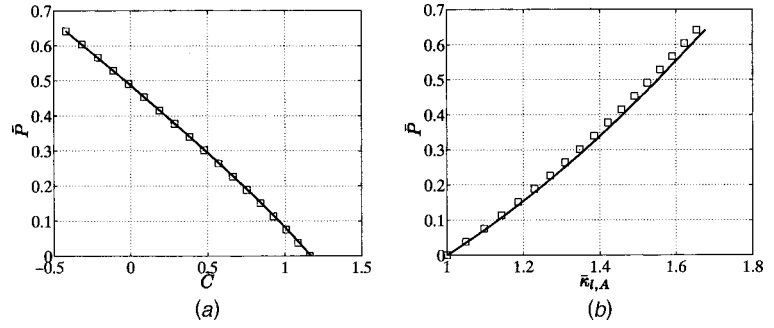


Fig. 7 Comparison of behavior between end loads and maximum longitudinal radius curvature: (*a*) Variation in end couple with applied force: FEA (solid); theory (squares). Geometry as in Table 1. (*b*) Variation in longitudinal curvature at the fold mid-point, $\kappa_{l,A}$ with end force P. Line-style as in part (*a*).

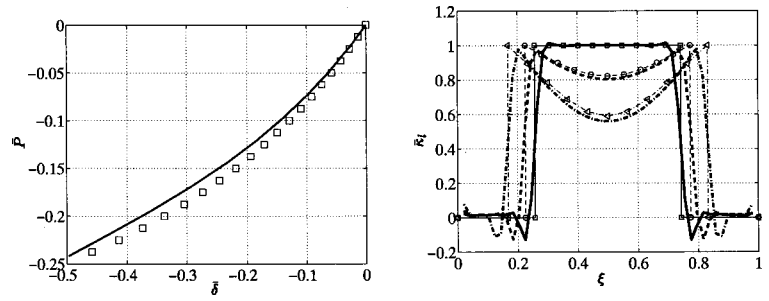


Fig. 8 Comparison of behavior of repelling end-wise forces: (*a*) Variation in force *P* with end displacement $\bar{\delta}$: FEA given by continuous line; theoretical prediction is squares. Geometry as in Table 1. (*b*) Behavior of longitudinal radius of curvature for $P=0N$ (solid), $P=-1.318N$ (dashed), and $P=-2.644N$ (dashed-dot). FEA is thick lines, theory is thin lines and symbols.

In addition, Fig. 7(*b*) indicates that the maximum longitudinal curvature (at the fold midpoint) for zero end couple is some 50 percent greater than $1/R^*$ for the tape-spring of this study. Hence, the maximum strain in a fold supported by end forces alone is much greater than in the case of end couples M^*, and may lead to unforeseen and permanent deformation in the material. Again, the discrepancies between theory and computation in both plots of Fig. 7 are very small.

(b) End-Wise Tension. The force-displacement plot is shown in Fig. 8(*a*). The procedure for obtaining the theoretical curve is exactly the same as in Fig. 6(*a*), but with negative values for *P*. The correlation between results is, again, very good at low values of δ, and marginally diverges at higher values. Figure 8(*b*) indicates that the fold curvature rapidly decreases towards the midpoint: as the magnitude of *P* increases, the distance between points in the fold and the tape-spring end decreases, and the moment-carrying capabilities are reduced. This manifests as a steadily reducing gradient in the (P, δ) response in the previous figure.

Finally, the maximum magnitude for *P* is given by noting that the curvature at the midpoint cannot be less than zero. Inverting Eq. (19) and setting equal to zero with $\beta=0$ defines the inequality constraint

$$\frac{2PR^{*2}}{DR\alpha}\left[1-\cos\frac{\theta}{2}\right]>-1, \quad \Rightarrow \bar{P}>-\frac{1}{2}\left(\frac{R}{R^*}\right)^2\frac{1}{1-\cos[\theta/2]}, \tag{22}$$

making use of Eq. (21). For $\theta=\pi$, Eq. (22) yields $\bar{P}>-0.370$, but there is no guarantee that the largest value can be carried in practice without the singly folded tape-spring now forming two distinct folds. The finite element procedure was unable to capture this separation so that the results of Eq. (22) could be verified; nonetheless, the above calculation provides a simple upper-bound estimate on the maximum load.

6 Summary

If unsymmetric loads are applied to a folded tape-spring, the fold may freely roll along the tape; if the tape-spring forms part of an application, then the unbalanced loads may alter the packaged arrangement. The only combination of end loads that may be carried without movement of a fold far away from any constraint has co-linear end forces in the simplest case of a symmetric layout. A theoretical model for the symmetric deformation of a fold under tape-spring end forces and end couples has been derived using a simple, but expeditious variational technique. This points to a closed-form solution for the variation in longitudinal radius of curvature in the fold as a function of the prescribed end force, material properties, and geometry of the folded tape-spring. The large-displacement response has been shown to be accurately correlated by a finite element analysis. Thus, for folded tape-spring structures, if tape-spring end conditions are known in advance of launch, the foregoing analysis may be used to predict the degree of distortion in the softer folds and, hence, the implications for safe stowage and correct deployment.

Nomenclature

c_0, c_1 = constant potential energy terms
C = tape-spring end couple
D = flexural rigidity$=Et^3/12(1-\nu^2)$
E = Young's modulus
F = composite function
J_0, J_1 = variational functions
L = tape-spring length
M = bending moment
M^* = fold moment under end couples
P = tape-spring end force
r = longitudinal radius of curvature
R = transverse radius of curvature
R^* = fold radius under end couples
t = tape-spring thickness
x, X = tape-spring end separation

Symbols

α = subtended angle of cross section
β = fold angular coordinate
γ_0, γ_1 = variational multipliers
δ = linear displacement between ends
ξ = intrinsic tape coordinate
κ_l, κ_t = longitudinal and transverse curvature changes
λ = dimensionless length of straight parts
θ = fold angle of embrace; relative end rotation
ν = Poisson's ratio
Ω_0, Ω_1 = potential energy integrands

References

[1] Seffen, K. A., and Pellegrino, S., 1999, ''Deployment Dynamics of Tape-Springs,'' Proc. R. Soc. London, Ser. A, **455**, pp. 1003–1048.
[2] Seffen, K. A., 1997, *Analysis of Structures Deployed by Tape-Springs*, Ph.D. thesis, University of Cambridge, Cambridge, UK.
[3] Seffen, K. A. and Pellegrino, S., 1997, ''Deployment of a Rigid Panel by Tape-Springs,'' Technical report, University of Cambridge, Department of Engineering, CUED/D-STRUCT/TR 168.
[4] Seffen, K. A., Pellegrino, S., and Parks, G. T., 1998, ''Deployment of a Panel by Tape-Spring Hinges,'' *Proceedings of IUTAM-IASS Symposium on Deployable Structures: Theory and Applications*, S. Pellegrino and S. D. Guest, eds., Kluwer, The Netherlands, pp. 355–364.
[5] Cambridge Consultants Ltd., 1989, ''Design Study for a Mars Sailcraft,'' Technical Report, Cambridge Consultants, Ltd., Q7844/JPA/Issue 1.
[6] Rits, W. J., 1996, ''A Multipurpose Deployable Membrane Reflector,'' ESA Bull., **88**, pp. 66–71.
[7] Seffen, K. A., You, Z., and Pellegrino, S., 2000, ''Folding and Deployment of Curved Springs,'' Int. J. Mech. Sci., **42**, pp. 2055–2073.
[8] Kyriakides, S., 1994, ''Propagating Instabilities in Structures,'' *Advances in Applied Mechanics*, J. W. Hutchinson and T. Y. Wu, eds. Academic Press, San Diego, CA, pp 67–189.
[9] Wuest, W., 1954, ''Einige Anvendungen der Theorie der Zylinderschale,'' Z. Angew. Math. Mech., **34**, pp. 444–454.

[10] Rimrott, F. P. J., 1970, ''Querschnittsverformung bei Torsion Offnerer Profile,'' Z. Angew. Math. Mech., **50**, pp. 775–778.
[11] Calladine, C. R., 1988, ''The Theory of Thin Shell Structures, 1888–1988,'' Proc. Inst. Mech. Eng., **202**, pp. 1–9.
[12] Calladine, C. R., 1983, *Theory of Shell Structures*, Cambridge University Press, Cambridge, UK.
[13] Hibbitt, Karlsson, and Sorenson, 1997, *ABAQUS Version 5.7*, Hibbitt, Karlsson, and Sorenson, Pawtucket.
[14] Hestenes, M. R., 1980, *Calculus of Variations and Optimal Control Theory*, Robert E. Krieger Publishing, New York.
[15] Mathworks, 1997, *MATLAB Version 5.1*. The Mathworks Inc., Natick.

On Saint-Venant's Problem for an Inhomogeneous, Anisotropic Cylinder—Part I: Methodology for Saint-Venant Solutions

S. B. Dong
Civil and Environmental Engineering Department, University of California, Los Angeles, CA 90095-1593
Mem. ASME

J. B. Kosmatka
Department of Applied Mechanics and Engineering Science, University of California, San Diego, CA 92093-0085
Mem ASME

H. C. Lin
Civil and Environmental Engineering Department, University of California, Los Angeles, CA 90095-1593

In this paper, the first in a series of three, a procedure based on semi-analytical finite elements is presented for constructing Saint-Venant solutions for extension, bending, torsion, and flexure of a prismatic cylinder with inhomogeneous, anisotropic cross-sectional properties. Extension-bending-torsion involve stress fields independent of the axial coordinate and their displacements may be decomposed into two distinct parts which are called the primal field and the cross-sectional warpages herein. The primal field embodies the essence of the kinematic hypotheses of elementary bar and beam theories and that for unrestrained torsion. The cross-sectional warpages are independent of the axial coordinate and they are determined by testing the variationally derived finite element displacement equations of equilibrium with the primal field. For flexure, a restricted three-dimensional stress field is in effect where the stress can vary at most linearly along the axis. Integrating the displacement field based for extension-bending-torsion gives that for the flexure problem. The cross-sectional warpages for flexure are determined by testing the displacement equations of equilibrium with this displacement field. In the next paper, the cross-sectional properties such as the weighted-average centroid, center of twist and shear center are defined based on the Saint-Venant solutions established in the present paper and numerical examples are given. In the third paper, end effects or the quantification of Saint-Venant's principle for the inhomogeneous, anisotropic cylinder is considered. [DOI: 10.1115/1.1363598]

Introduction

This paper is the first of three devoted to the equilibrium analysis of a finite length prismatic elastic cylinder whose cross section may be arbitrary in its geometry and may be composed of any number of distinct materials. Each constituent material may have linear elastic mechanical properties exhibiting the most general form of rectilinear anisotropy. The materials are perfectly bonded so that full intersurface kinematic and traction continuity is assured. The cylinder's lateral surface is traction-free. The two ends of the cylinder are acted upon by tractions which are presumed to be prescribed on a pointwise basis. These traction states lead to an axial force, bending moment, torque, and flexure (transverse shear) force, and they must occurred in such a way that overall equilibrium of the cylinder is maintained.

Specializing the above problem description to a homogeneous, isotropic cylinder gives the celebrated Saint-Venant's problem. The Saint-Venant's solutions or that to the relaxed formulation of Saint-Venant's problem ([1,2]) are solutions in which the pointwise specification of tractions on the ends was replaced by integrals representing the axial force, bending moment, torque, and flexure force. Saint-Venant asserted that differences between traction states according to his solutions and any other equipollent traction state were confined to regions at the ends of the cylinder, i.e., Saint-Venant's principle. The Saint-Venant solutions for homogeneous, isotropic cylinders occupy a very important place in structural engineering. They validate the kinematic hypotheses of elementary mechanics of material theories for the extension, bending, torsion, and flexure of slender members as well as delineate their ranges of validity.

With integral end conditions in the relaxed formulation, many competing solutions are possible as there are limitless end traction states capable of producing identical force and moment resultants. But the Saint-Venant solutions are distinguished by certain characteristics. Clebsch [3] indicated that the resultant traction vector on any material plane normal to the cross section of the cylinder is parallel to its generator. For extension, bending, and torsion, Voigt [4] indicated that the stress and strain fields are independent of the axial coordinate, and for flexure, these states can vary at most linearly along the cylinder's axis. Sternberg and Knowles [5] showed that the Saint-Venant extension-bending-torsion solutions for a homogeneous, isotropic cylinder produce absolute minimum strain energy states. For the flexure problem, Sternberg and Knowles provided a proof of a minimum strain energy state for the special case of Poisson's ratio equal to zero.

Herein, Saint-Venant solutions, i.e., solutions to the relaxed formulation of the Saint-Venant problem, for an inhomogeneous, anisotropic cylinder are constructed. For this task, two sequential problems are addressed according to the decomposition of these end cross section force and moment resultants into the following two sets, (1) extension-bending-torsion and (2) flexure. Variationally derived semi-analytical finite element displacement equations of equilibrium are used in the analysis. The solution methodology calls for the displacement fields to be set forth at the outset. Each of the fields for the two sequential problems can be put into two essential parts, (1) a primal field and (2) cross-sectional warpages. Ieşan's rational scheme ([6]) is used to generate these fields, because it is systematic and does not require any other a priori assumptions, such as those used in the semi-inverse method. By his scheme, the primal field for extension-bending-torsion is obtained by integrating the most general form of rigid-body dis-

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. 7, 1999; final revision, July 21, 2000. Associate Editor: J. W. Ju. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

placement with respect to the axial coordinate. That for flexure is obtained by integrating that for extension-bending-torsion. These displacement fields are expressed in terms of unknown amplitude coefficients. It is mentioned that these fields may also be derived from Voigt's characterization of the Saint-Venant problem, i.e., that for extension-bending-torsion is obtained by integrating the strain-displacement relations for a strain state that is independent of the axial coordinate and that for flexure is obtained by integrating the extension, bending, torsion field once with respect to the axial coordinate. Testing the governing equilibrium equations with the extension-bending-torsion and flexure displacement fields results in systems of equations for the cross-sectional warpages that may be seen to be driven by the primal field. Once the warpages are found, the stress components can be determined in terms of the displacement amplitude coefficients. By integrating the appropriate stress components over the cross section, cross-sectional stiffness relations are formed that enable the resultant forces and moments to be related to the unknown displacement coefficients.

From the Saint-Venant solutions, cross-sectional properties such as the centroid (or stiffness weighted centroid), the principal bending axes, the center of twist and the shear center can be determined. These issues are addressed in Kosmatka, Lin, and Dong [7], the second paper in this three-part series. End effects are considered in Lin, Dong, and Kosmatka [8], the third of the three-part series, where a method for the quantitative analysis of Saint-Venant's principle is given. Any pointwise specification of end tractions can be taken as the sum a Saint-Venant field and a self-equilibrated stress state. This self-equilibrated stress state can be represented by the eigendata of the spectral representation of the matrix operator in the governing equilibrium equation. The specific case of end effects in restrained torsion was given by Kazic and Dong [9].

The present procedure for Saint-Venant solutions differs significantly from that used by Kosmatka and Dong [10] for a homogeneous, anisotropic cylinder even though all theoretical considerations are the same. Comments on the differences are given in the Concluding Remarks section.

Semi-Analytical Finite Element Equations

Consider a prismatic cantilevered cylinder of length L with an arbitrarily shaped cross section of perfectly bonded linearly elastic anisotropic materials as shown in Fig. 1. Let R denote the open region occupied by the cylinder with B as its lateral surface. Let Π denote a generic cross section of the cylinder whose lateral boundary curve is S_π. The surfaces at the tip and root ends are identified as Π_1 and Π_2, respectively. Establish right-hand Cartesian coordinates (x, y, z) and take the origin at some point at the tip end Π_1. The mechanical variables of the problem are stress, strain, and displacement, $\boldsymbol{\sigma}(x,y,z), \boldsymbol{\epsilon}(x,y,z), \mathbf{u}(x,y,z)$, with components $\boldsymbol{\sigma}=[\sigma_{xx},\sigma_{yy},\sigma_{zz},\sigma_{yz},\sigma_{xz},\sigma_{xy}]^T$, $\boldsymbol{\epsilon}=[\epsilon_{xx},\epsilon_{yy},\epsilon_{zz},\gamma_{yz},\gamma_{xz},\gamma_{xy}]^T$, $\mathbf{u}=[u,v,w]^T$. The constitutive equation for a given anisotropic material in the cross section has the form $\boldsymbol{\sigma}=\mathbf{C}\boldsymbol{\epsilon}$, where the symmetric (6×6) matrix $\mathbf{C}$ contains the 21 independent elastic moduli.

In the present version of the semi-analytical finite element method, the cross section of the cylinder undergoes discretization. In each element, the displacement field $\mathbf{u}$ is taken in product form of interpolation functions $\mathbf{n}(x,y)$ over the cross section and nodal variables $(\mathbf{u}_u,\mathbf{u}_v,\mathbf{u}_w)$ of unspecified functions of the axial coordinate z. In the computer code for the examples in the two subsequent papers, both six-node triangular and eight-node quadrilateral elements are used, so that the cross-sectional interpolations are complete second-order polynomials. The displacement field $\mathbf{u}(x,y,z)$ has the form

$$\mathbf{u}(x,y,z)=\mathbf{n}_e(x,y)\mathbf{u}_e(z) \tag{1a}$$

or

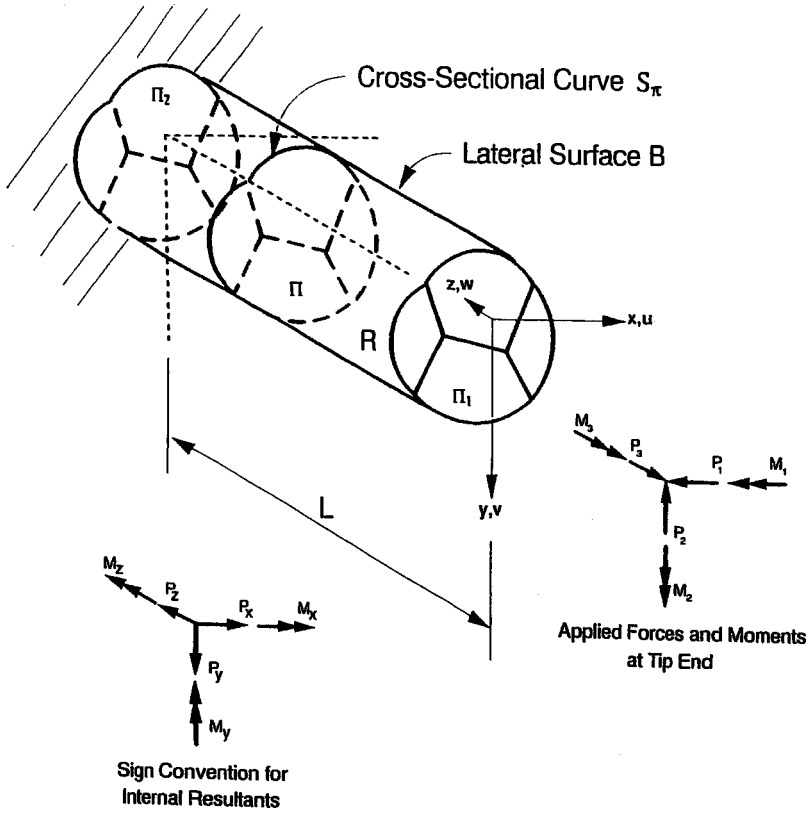


Fig. 1 Coordinate system for anisotropic cylinder

$$\begin{Bmatrix} u(x,y,z) \\ v(x,y,z) \\ w(x,y,z) \end{Bmatrix} = \begin{bmatrix} \mathbf{n}(x,y) & \cdot & \cdot \\ \cdot & \mathbf{n}(x,y) & \cdot \\ \cdot & \cdot & \mathbf{n}(x,y) \end{bmatrix} \begin{Bmatrix} \mathbf{u}_u(z) \\ \mathbf{u}_v(z) \\ \mathbf{u}_w(z) \end{Bmatrix}. \tag{1b}$$

The separation of the dependent variables with one part stated in terms of an assumed field occupies in an intermediate position between an exact solution and a Ritz/Galerkin technique, a procedure due to Kantorovich and Krylov [11].

With this partitioned dependence of displacement field (1), the differential strain-displacement operators can be separated into two parts.

$$\boldsymbol{\epsilon} = \mathbf{L}_{xy}\mathbf{u} + \mathbf{L}_z\mathbf{u} \tag{2}$$

where $\mathbf{L}_{xy}$ and $\mathbf{L}_z$ are matrices of linear differential operators given by

$$\mathbf{L}_{xy} = \begin{bmatrix} \frac{\partial}{\partial x} & \cdot & \cdot \\ \cdot & \frac{\partial}{\partial y} & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \frac{\partial}{\partial y} \\ \cdot & \cdot & \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & \cdot \end{bmatrix}; \quad \mathbf{L}_z = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \frac{\partial}{\partial z} \\ \cdot & \frac{\partial}{\partial z} & \cdot \\ \frac{\partial}{\partial z} & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}. \tag{3}$$

Substitution of Eq. (1) into Eq. (2) gives the strain-transformation equations as

$$\boldsymbol{\epsilon} = \mathbf{b}_1\mathbf{u}_e + \mathbf{b}_2\mathbf{u}_{e,z} \tag{4}$$

where

$$\mathbf{b}_1 = \begin{bmatrix} \mathbf{n}_{,x} & \cdot & \cdot \\ \cdot & \mathbf{n}_{,y} & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \mathbf{n}_{,y} \\ \cdot & \cdot & \mathbf{n}_{,x} \\ \mathbf{n}_{e,y} & \mathbf{n}_{,x} & \cdot \end{bmatrix}; \quad \mathbf{b}_2 = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \mathbf{n} \\ \cdot & \mathbf{n} & \cdot \\ \mathbf{n} & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}. \tag{5}$$

The governing equilibrium equations are variationally derived from the theorem of minimum potential energy in the form of

$$\delta\left(\frac{1}{2}\int_0^L \left\{ \int\int_\Pi \boldsymbol{\epsilon}^T\mathbf{C}\boldsymbol{\epsilon}\,dx\,dy \right\} dz + V_E\right) = 0 \tag{6}$$

where V_E represents potential energy of end tractions on Π_1. As the variational process on V_E gives boundary data only, its appropriate form for Saint-Venant solutions is deferred until the discussion of boundary conditions. Inserting strains $\boldsymbol{\epsilon}$ for all elements into Eq. (6) and carrying out the variation leads to the following governing equilibrium equations for the anisotropic cylinder:

$$\mathbf{K}_1\mathbf{U}_{,zz} + \mathbf{K}_2\mathbf{U}_{,z} - \mathbf{K}_3\mathbf{U} = 0 \tag{7}$$

where $\mathbf{U} = [\mathbf{U}_u, \mathbf{U}_v, \mathbf{U}_w]^T$ denotes the assembled ordered nodal displacement components and system stiffness matrices $\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_3$ are given by

$$[\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_3] = \sum_{n=1}^{N} \int\int [(\mathbf{b}_2^T\mathbf{C}\mathbf{b}_2), \; (\mathbf{b}_2^T\mathbf{C}\mathbf{b}_1 - \mathbf{b}_1^T\mathbf{C}\mathbf{b}_2), \; (\mathbf{b}_1^T\mathbf{C}\mathbf{b}_1)]\,dx\,dy. \tag{8}$$

Note that $\mathbf{K}_1$ and $\mathbf{K}_3$ are symmetric, while $\mathbf{K}_2$ is antisymmetric. The specific roles of these stiffness matrices can be inferred from their dependence on $\mathbf{b}_1$ and $\mathbf{b}_2$. As $\mathbf{K}_3$ is constituted from $\mathbf{b}_1$, it governs generalized plane strain, while $\mathbf{K}_1$ formed from $\mathbf{b}_2$ relates to behavior complement to generalized plane strain. Stiffness matrix $\mathbf{K}_2$ is the agent that couples these two behaviors. These operators are the same as those in the governing equations of Kazic and Dong [9] for the analysis of restrained torsion and Taweel, Dong, and Kazic [12] for reflection of monochromatic waves at the free end of a cylinder. The derivation follows standard isoparametric finite element methodology, and further details may be found in the aforementioned papers.

In the relaxed formulation, force and moment resultants are employed rather than the tractions on a cross section. On a generic cross section Π at some arbitrary axial coordinate z, these force and moment resultants are given by

$$\int\int_\Pi \sigma_{zx}\,dx\,dy = P_x(z) \qquad \int\int_\Pi \sigma_{zz}y\,dx\,dy = M_x(z)$$

$$\int\int_\Pi \sigma_{zy}\,dx\,dy = P_y(z); \qquad \int\int_\Pi \sigma_{zz}x\,dx\,dy = -M_y(z) \tag{9}$$

$$\int\int_\Pi \sigma_{zz}\,dx\,dy = P_z(z) \qquad \int\int_\Pi (\sigma_{zy}x - \sigma_{zx}y)\,dx\,dy = M_z(z)$$

This equation can be recast in compact form as

$$\int\int_\Pi \mathbf{h}^T\boldsymbol{\sigma}\,dx\,dy = \mathbf{F}(z) \tag{10}$$

where $\mathbf{F}(z)$ and $\mathbf{h}(x,y)$ are given by

$$\mathbf{F}^T(z) = [P_x, P_y, P_z, M_x, -M_y, M_z] \tag{11}$$

$$\mathbf{h}(x,y) = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & y & x & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & x \\ 1 & \cdot & \cdot & \cdot & \cdot & -y \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}. \tag{12}$$

In replacing the pointwise specification of the end tractions by their force and moment resultants, the form of V_E in Eq. (6) is written as

$$V_E = -\overline{\mathbf{F}}_I^T(0)\mathbf{a} = -\{\overline{P}_1a_1 + \overline{P}_2a_2 + \overline{P}_3a_3 + \overline{M}_1a_4 - \overline{M}_2a_5 + \overline{M}_3a_6\}_{\Pi_1} \tag{13}$$

where a_i are displacement amplitudes of the Saint-Venant solutions to be presented. They are associated with the cross-sectional deformational measures corresponding to these force and moment resultants. At the root end Π_2, a fully restrained condition is assumed so that kinematic boundary conditions are met in the variational process.

Rigid-Body Displacements

There are six distinct rigid-body displacement modes for the cylinder, which satisfy governing Eq. (7) identically. Furthermore, they lead to zero strains when substituted into strain-transformation Eq. (4). These two sets of relations are useful identities in the solution procedure.

Let M be the total number of nodes in a given finite element model. The matrix form of the rigid-body displacements can be set forth with the help of the following six $3M$ column vectors $\mathbf{R}_i$.

$$\mathbf{R}_1=\begin{Bmatrix}\mathbf{I}_1\\ \mathbf{0}\\ \mathbf{0}\end{Bmatrix};\quad \mathbf{R}_2=\begin{Bmatrix}\mathbf{0}\\ \mathbf{I}_1\\ \mathbf{0}\end{Bmatrix};\quad \mathbf{R}_3=\begin{Bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{I}_1\end{Bmatrix}$$
$$\mathbf{R}_4=\begin{Bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{y}\end{Bmatrix};\quad \mathbf{R}_5=\begin{Bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{x}\end{Bmatrix};\quad \mathbf{R}_6=\begin{Bmatrix}-\mathbf{y}\\ \mathbf{x}\\ \mathbf{0}\end{Bmatrix} \tag{14}$$

where $(\mathbf{I}_1,\mathbf{x},\mathbf{y})$ are column vectors of length M with unit entries, the x and the y-coordinates of the M nodes, respectively. Then, rigid-body displacement vector $\mathbf{U}_{RB}(z)$ has the form

$$\mathbf{U}_{RB}(z)=\mathbf{\Phi}_{RB}\mathbf{a}_{RB}=[-z\mathbf{N}_1+\mathbf{N}_2]\mathbf{a}_{RB} \tag{15}$$

where

$$\mathbf{N}_1=[\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{R}_2,\mathbf{R}_1,\mathbf{0}]^T;\quad \mathbf{N}_2=[\mathbf{R}_1,\mathbf{R}_2,\mathbf{R}_3,\mathbf{R}_4,\mathbf{R}_5,\mathbf{R}_6]^T \tag{16}$$

and $\mathbf{a}_{RB}=[u_o,v_o,w_o,\omega_1,\omega_2,\omega_3]^T$ contains the six translational and rotational amplitudes. Substituting Eq. (15) into governing Eq. (7) and strain-transformation Eq. (4) gives

$$\mathbf{K}_3\mathbf{R}_i=0\quad (i=1,2,3,6);\quad \mathbf{b}_1\mathbf{r}_i=0\quad (i=1,2,3,6) \tag{17a}$$

$$\begin{aligned}&\mathbf{K}_3\mathbf{R}_4=-\mathbf{K}_2\mathbf{R}_2;\qquad &&\mathbf{b}_1\mathbf{r}_4=\mathbf{b}_2\mathbf{r}_2\\ &\mathbf{K}_3\mathbf{R}_5=-\mathbf{K}_2\mathbf{R}_1;\qquad &&\mathbf{b}_1\mathbf{r}_5=\mathbf{b}_2\mathbf{r}_1\end{aligned} \tag{17b}$$

where $\mathbf{r}_i$s are the counterparts of $\mathbf{R}_i$s on the element level.

Problem I—Extension-Bending-Torsion

Problem I refers to extension-bending-torsion by applied tractions on the end cross section Π_I. The displacement field for Saint-Venant's solutions can be written as

$$\begin{aligned}u(x,y,z)&=a_{I1}z-a_{I5}\frac{z^2}{2}-a_{I6}yz\\ &\quad+\sum_{i=1}^{6}a_{Ii}\psi_{Iiu}(x,y)-\omega_3 y-\omega_2 z+u_o\\ v(x,y,z)&=a_{I2}z-a_{I4}\frac{z^2}{2}+a_{I6}xz\\ &\quad+\sum_{i=1}^{6}a_{Ii}\psi_{Iiv}(x,y)+\omega_3 x-\omega_1 z+v_o\\ w(x,y,z)&=(a_{I3}+a_{I5}x+a_{I4}y)z\\ &\quad+\sum_{i=1}^{6}a_{Ii}\psi_{Iiw}(x,y)+\omega_1 y+\omega_2 x+w_o\end{aligned} \tag{18a}$$

where a_{Ii} ($i=1$ to 6) are displacement amplitudes and $(\psi_{Iiu},\psi_{Iiv},\psi_{Iiw})$ the components of the cross-sectional warpages. The coefficients a_{I1} and a_{I2} are associated with longitudinal shear, a_{I3} with extension, a_{I4} and a_{I5} with bending, and a_{I6} with torsion. That part of the field devoid of warpage functions and rigid-body displacement is called the *primal* field. The rigid-body displacements $(u_o,v_o,w_o,\omega_1,\omega_2,\omega_3)$ were included for completeness sake.

Displacement field (18*a*) can be derived by integrating the strain-displacement equations for a strain field independent of z. It can also be obtained by integrating the rigid-body displacement with respect to z as shown by Ieşan [6].

The roles of the two longitudinal shears and their amplitude coefficients a_{I1} and a_{I2} need clarification. These longitudinal shears appear in Eq. (18*a*) as a direct consequence of integrating the most general form of the rigid-body displacement. They produce shear tractions on the lateral surface B, which violate the side conditions of the Saint-Venant problem. In the present solution procedure, carrying these two modes along produce null results. They are completely uncoupled with the extension-bending-torsion problem. It may be possible that these deformation modes will participate in the solution of Almansi [13] and Michell [14] problems where tractions occur on the lateral surface.

Recasting displacement field (18*a*) in matrix form gives

$$\mathbf{U}(z)=[\mathbf{\Phi}_I(z)+\mathbf{\Psi}_I]\mathbf{a}_I+\mathbf{\Phi}_{RB}(z)\mathbf{a}_{RB} \tag{18b}$$

where $\mathbf{\Phi}_I$ and $\mathbf{\Psi}_I$ are $(3M\times 6)$ matrices given by

$$\mathbf{\Phi}_I(z)=-\frac{z^2}{2}\mathbf{N}_1+z\mathbf{N}_2 \tag{19}$$

and

$$\begin{aligned}\mathbf{\Psi}_I&=[\psi_{I1},\psi_{I2},\psi_{I3},\psi_{I4},\psi_{I5},\psi_{I6}]^T\\ \mathbf{a}_I&=[a_{I1},a_{I2},a_{I3},a_{I4},a_{I5},a_{I6}]^T\end{aligned} \tag{20}$$

Observe that $\mathbf{\Phi}_I=\int\mathbf{\Phi}_{RB}dz$, which is the essence of Ieşan's scheme. An inspection of the primal field terms associated with a_{I3} to a_{I6} in Eq. (18*a*) or its matrix components in Eq. (19) shows that they express the kinematic hypotheses of elementary structural theories of extension, pure bending, and unrestrained torsion.

The cross-sectional warpages ψ_{Ii}s can be established by substituting Eq. (18*b*) into Eq. (7) and equating all terms multiplied by deformation coefficients a_{Ii} to zero.

$$\begin{aligned}&\mathbf{K}_3\psi_{Ii}=\mathbf{K}_2\mathbf{R}_i;\quad (i=1,2,3,6)\\ &\mathbf{K}_3\psi_{I4}=\mathbf{K}_2\mathbf{R}_4-\mathbf{K}_1\mathbf{R}_2\\ &\mathbf{K}_3\psi_{I5}=\mathbf{K}_2\mathbf{R}_5-\mathbf{K}_1\mathbf{R}_1\end{aligned} \tag{21}$$

Observe that the right-hand sides of Eq. (21) involve the primal field $\mathbf{\Phi}_I$, so that the cross-sectional warpages may be said to be driven by the displacement field embodying the kinematic hypotheses of the elementary structural theories. The warpages are elastic responses due to cross-elasticity (Poisson ratio) effects and the longitudinal shear warpages of free torsion. Note also that ψ_{I1} and ψ_{I2} in Eq. (21) satisfy the rigid body identities (17a) so that

$$\psi_{I1}=-\mathbf{R}_5;\quad \psi_{I2}=-\mathbf{R}_4. \tag{22}$$

This result shows that the two longitudinal shear fields, γ_{xz} and γ_{yz}, which are associated with a_{I1} and a_{I2}, vanish identically and are not involved in Problem I for extension-bending-torsion.

The solution to Eq. (21) requires the inverse of $\mathbf{K}_3$. Because $\mathbf{K}_3$ is singular due to the presence of rigid-body motion, it cannot be factorized without administering kinematic constraints. Four rigid-body modes, the three translations along the coordinate directions and a rotation about the z-axis must be suppressed from $\mathbf{K}_3$ prior to its inverse.

Once the nodal warpages $\mathbf{\Psi}_{Ii}$s are found, the functional dependence of displacement field (18*b*) is completely defined. Using strain-transformation Eq. (4), identities (17*b*), and the anisotropic stress-strain relation, the strain and stress components in an element can be written as

$$\boldsymbol{\epsilon}=[\mathbf{b}_2\mathbf{n}_2+\mathbf{b}_1\mathbf{\Psi}_{Ie}]\mathbf{a}_I=[\mathbf{h}+\mathbf{b}_1\mathbf{\Psi}_{Ie}]\mathbf{a}_I=\boldsymbol{\epsilon}_0\mathbf{a}_I \tag{23a}$$

$$\boldsymbol{\sigma}=\mathbf{C}\boldsymbol{\epsilon}_0\mathbf{a}_I=\boldsymbol{\sigma}_0\mathbf{a}_I \tag{23b}$$

where $\mathbf{n}_2$ and $\mathbf{\Psi}_{Ie}$ are the counterparts of $\mathbf{N}_2$ and nodal warpages $\mathbf{\Psi}_I$ for the given element. In Eqs. (23*a*), (23b), $\mathbf{b}_2\mathbf{n}_2$ was replaced by $\mathbf{h}$, which is possible by using identity (17*b*) and recalling the criteria that preserve rigid-body displacements and constant strain modes in isoparametric finite element mappings, i.e.,

$$\sum n_i=1;\quad \sum n_i x_i\equiv\mathbf{nx}=x;\quad \sum n_i y_i\equiv\mathbf{ny}=y. \tag{24}$$

The resultants on the end cross section Π_I for Problem I in terms of the vector $\mathbf{F}$ in Eq. (13) takes the form

$$\mathbf{F}_I=[0,0,P_3,M_1,-M_2,M_3]^T. \tag{25}$$

These resultants can be found by integrating the stress components $(\sigma_{zz},\sigma_{xz},\sigma_{yz})$ of Eq. (23*b*) over each element's cross-sectional area as indicated by Eq. (10) and summing the contributions of the N elements of the total cross section.

$$\sum_{i=1}^{N}\int\int_{\Pi_1}\mathbf{h}^T\boldsymbol{\sigma}dxdy=\sum_{i=1}^{N}\left[\int\int_{\Pi_1}\mathbf{h}^T\mathbf{C}[\mathbf{h}+\mathbf{b}_1\boldsymbol{\Psi}_{Ie}]dxdy\right]\mathbf{a}_I=\mathbf{F}_I \tag{26}$$

or in the form of a cross-sectional stiffness relation

$$\underset{(6\times6)(6\times1)}{\boldsymbol{\kappa}_I\ \mathbf{a}_I}=\underset{(6\times1)}{\mathbf{F}_I}\rightarrow\begin{bmatrix}\underset{(2\times2)}{\mathbf{0}} & \underset{(2\times4)}{\mathbf{0}}\\ \underset{(4\times2)}{\mathbf{0}} & \underset{(4\times1)}{\boldsymbol{\kappa}_{Ibb}}\end{bmatrix}\begin{Bmatrix}\underset{(2\times1)}{\mathbf{a}_{Ia}}\\ \underset{(4\times1)}{\mathbf{a}_{Ib}}\end{Bmatrix}=\begin{Bmatrix}\underset{(2\times1)}{\mathbf{0}}\\ \underset{(4\times1)}{\mathbf{F}_{Ib}}\end{Bmatrix} \tag{27a}$$

where the expanded form of its nontrivial portion $\boldsymbol{\kappa}_{Ibb}$ is

$$\begin{bmatrix}\kappa_{I33} & \kappa_{I34} & \kappa_{I35} & \kappa_{I36}\\ \kappa_{I34} & \kappa_{I44} & \kappa_{I45} & \kappa_{I46}\\ \kappa_{I35} & \kappa_{I45} & \kappa_{I55} & \kappa_{I56}\\ \kappa_{I36} & \kappa_{I46} & \kappa_{I56} & \kappa_{I66}\end{bmatrix}\begin{Bmatrix}a_{I3}\\ a_{I4}\\ a_{I5}\\ a_{I6}\end{Bmatrix}=\begin{Bmatrix}P_3\\ M_1\\ -M_2\\ M_3\end{Bmatrix}. \tag{27b}$$

Note that $\boldsymbol{\kappa}_{Ibb}$ is a symmetric matrix. The absence of terms in the first two rows and columns of the $(6\times6)\boldsymbol{\kappa}_I$ matrix further substantiates that the longitudinal shear components play no part in extension-bending-torsion. Solution of Eq. (27*b*) gives

$$\mathbf{a}_{Ib}=\boldsymbol{\kappa}_{Ibb}^{-1}\mathbf{F}_{Ib}. \tag{28}$$

The coefficients of $\boldsymbol{\kappa}_{Ibb}$ and its inverse contain a number of properties related to the cross section, such as the weighted-average centroid, principal bending axes, center of twist, etc., and these properties will be considered in the next paper of this series of three papers.

Problem II—Flexure

Saint-Venant flexure involves a stress field that is at most linear with z. The appropriate displacement field according to Ieşan's scheme [6] is obtained by integrating the displacement field of Problem I once with respect to z to give

$$\mathbf{U}(x,y,z)=[\boldsymbol{\Phi}_{II}(z)+z\boldsymbol{\Psi}_I+\boldsymbol{\Psi}_{II}]\mathbf{a}_{II}+[\boldsymbol{\Phi}_I(z)+\boldsymbol{\Psi}_I(z)+\boldsymbol{\Psi}_I]\mathbf{b}_{II}+\boldsymbol{\Phi}_{RB}(z)a_{\mathbf{RB}} \tag{29}$$

where $\boldsymbol{\Phi}_{II}=\int\boldsymbol{\Phi}_I dz$. There are six new warpage functions in $\boldsymbol{\Psi}_{II}$ and two new sets of displacement coefficient $\mathbf{a}_{II}$ and $\mathbf{b}_{II}$.

$$\begin{aligned}\boldsymbol{\Psi}_{II}&=[\psi_{II1},\psi_{II2},\cdots,\psi_{II6}]^T\\ \mathbf{a}_{II}&=[a_{II1},a_{II2},\cdots,a_{II6}]\\ \mathbf{b}_{II}&=[b_{II1},b_{II2},\cdots,b_{II6}]\end{aligned} \tag{30}$$

To determine the warpage functions, substitute Eq. (29) into Eq. (7) and set all terms multiplied by coefficients $\mathbf{a}_{II}$ and $\mathbf{b}_{II}$ to zero. Two sets of equations arise. One set is identical to Eq. (21), which verifies that the warpage $\boldsymbol{\Psi}_I$ is the same as that of Problem I. The other set $\boldsymbol{\Psi}_{II}$ is found by solving the following equation:

$$\mathbf{K}_3\psi_{IIi}=\mathbf{K}_1\mathbf{R}_i+\mathbf{K}_2\psi_{Ii};\ (i=1\ \text{to}\ 6). \tag{31}$$

This equation shows that the flexural warpages are driven by the primal field and the warpages of Problem I.

Substituting displacement field (29) into strain-transformation Eq. (3) and the anisotropic constitutive equations gives the element's strain and stress fields as

$$\boldsymbol{\epsilon}=[z\boldsymbol{\epsilon}_0+\boldsymbol{\epsilon}_1]\mathbf{a}_{II}+\boldsymbol{\epsilon}_0\mathbf{b}_{II} \tag{32a}$$

$$\boldsymbol{\sigma}=[z\boldsymbol{\sigma}_0+\boldsymbol{\sigma}_1]\mathbf{a}_{II}+\boldsymbol{\sigma}_0\mathbf{b}_{II} \tag{32b}$$

where $\boldsymbol{\epsilon}_0$ and $\boldsymbol{\sigma}_0$ were defined in Eqs. (23*a*) and (23*b*), respectively, and $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\sigma}_1$ are given by

$$\boldsymbol{\epsilon}_1=\mathbf{b}_2\boldsymbol{\Psi}_{Ie}+\mathbf{b}_1\boldsymbol{\Psi}_{IIe} \tag{33a}$$

$$\boldsymbol{\sigma}_1=\mathbf{C}\boldsymbol{\epsilon}_1=\mathbf{C}\mathbf{b}_2\boldsymbol{\Psi}_{Ie}+\mathbf{C}\mathbf{b}_1\boldsymbol{\Psi}_{IIe}. \tag{33b}$$

Integrating the stresses over an element's cross section according to Eq. (9) and summing the contributions of N elements define the array of resultant forces and moments on a generic cross section at any location z along the cylinder's axis as

$$\mathbf{F}(z)=\left[\int\int_{\Pi}\mathbf{h}^T[z\boldsymbol{\sigma}_0+\boldsymbol{\sigma}_1]dx\,dy\right]\mathbf{a}_{II}+\left[\int\int_{\Pi}\mathbf{h}^T\boldsymbol{\sigma}_0dx\,dy\right]\mathbf{b}_{II}=z(\boldsymbol{\kappa}_I\mathbf{a}_{II})+(\boldsymbol{\kappa}_{II}\mathbf{a}_{II}+\boldsymbol{\kappa}_I\mathbf{b}_{II}) \tag{34}$$

where $\boldsymbol{\kappa}_I$ was given by Eq. (27*a*) and the expanded form of $\boldsymbol{\kappa}_{II}$ is shown in Eq. (40*a*).

The coefficients in $\mathbf{a}_{II}$ are determined from global equilibrium considerations. Differentiating Eq. (34) and noting that

$$\frac{\partial M_x}{\partial z}=P_2;\ \ \frac{\partial M_y}{\partial z}=-P_1 \tag{35}$$

shows that

$$\frac{\partial\mathbf{F}(z)}{\partial z}\equiv\mathbf{F}'_{II}=[0,0,0,P_2,P_1,0]^T \tag{36}$$

so that

$$\boldsymbol{\kappa}_I\mathbf{a}_{II}=\mathbf{F}'_{II}\leftrightarrow\begin{bmatrix}\underset{(2\times2)}{\mathbf{0}} & \underset{(2\times4)}{\mathbf{0}}\\ \underset{(4\times2)}{\mathbf{0}} & \underset{(4\times4)}{\boldsymbol{\kappa}_{Ibb}}\end{bmatrix}\begin{Bmatrix}\underset{(2\times1)}{\mathbf{a}_{IIa}}\\ \underset{(4\times1)}{\mathbf{a}_{IIb}}\end{Bmatrix}=\begin{Bmatrix}\underset{(2\times1)}{0}\\ \underset{(4\times1)}{\mathbf{F}'_{IIb}}\end{Bmatrix}. \tag{37}$$

The solution for amplitudes a_{II3} to a_{II6} in $\mathbf{a}_{IIb}$ is given by extracting the second equation of the partitioned matrix in Eq. (37) and solving it.

$$\boldsymbol{\kappa}_{Ibb}\mathbf{a}_{IIb}=\mathbf{F}'_{IIb} \tag{38}$$

The coefficients a_{II1} and a_{II2} have no role in the flexure problem. With the coefficients in $\mathbf{a}_{II}$ known, the coefficients in $\mathbf{b}_{II}$ can be determined from Eq. (34) by invoking the following conditions at $z=0$,

$$\mathbf{F}(z)\big|_{z=0}\equiv\mathbf{F}_{II}=[P_1,P_2,0,0,0,P_2e_x+P_1e_y]^T, \tag{39}$$

where the term $P_2e_x+P_1e_y$ represents the torque of P_1 and P_2 because they may be acting at distances e_x and e_y off of the shear center. At $z=0$, Eq. (34) in expanded form appears as

$$\begin{bmatrix}\underset{(2\times2)}{\mathbf{0}} & \underset{(2\times4)}{\mathbf{0}}\\ \underset{(4x\times2)}{\mathbf{0}} & \underset{(4x\times4)}{\boldsymbol{\kappa}_{Ibb}}\end{bmatrix}\begin{Bmatrix}\underset{(2\times1)}{\mathbf{b}_{IIa}}\\ \underset{(4\times1)}{\mathbf{b}_{IIb}}\end{Bmatrix}+\begin{bmatrix}\underset{(2\times2)}{\mathbf{0}} & \underset{(2\times4)}{\boldsymbol{\kappa}_{IIab}}\\ \underset{(4x\times2)}{\boldsymbol{\kappa}_{IIba}} & \underset{(4x\times4)}{\boldsymbol{\kappa}_{IIbb}}\end{bmatrix}\begin{Bmatrix}\underset{(2\times1)}{\mathbf{0}}\\ \underset{(4\times1)}{\mathbf{a}_{IIb}}\end{Bmatrix}=\begin{Bmatrix}\underset{(2\times1)}{\mathbf{F}_{IIa}}\\ \underset{(4\times1)}{\mathbf{F}_{\mathbf{IIb}}}\end{Bmatrix} \tag{40a}$$

or

$$\boldsymbol{\kappa}_{IIab}\mathbf{a}_{IIb}=\mathbf{F}_{IIa};\ \ \boldsymbol{\kappa}_{Ibb}\mathbf{b}_{IIb}+\kappa_{IIbb}\mathbf{a}_{IIb}=\mathbf{F}_{IIb}. \tag{40b}$$

Although it is not obvious in the form given, the first of Eq. (40*b*) is identically satisfied. Upon its expansion, the results repeat that which are contained in Eq. (38). The second of Eq. (40*b*) enables the solution of $\mathbf{b}_{IIb}$.

$$\mathbf{b}_{IIb}=-\boldsymbol{\kappa}_{Ibb}^{-1}\boldsymbol{\kappa}_{IIbb}\mathbf{a}_{IIb}-\boldsymbol{\kappa}_{Ibb}^{-1}\mathbf{F}_{IIb} \tag{41}$$

The coefficients b_{II1} and b_{II2} do not have a role in the flexure problem.

It should be emphasized that the form of Eq. (39) implies that the flexure forces pass through the origin of coordinate system at the tip end, which may not coincide with the shear center. For flexure forces applied elsewhere, a torsional moment M_z must be included which is reflected in the sixth component in Eq. (39). A more general discussion of the shear center is given by the next paper.

Concluding Remarks

A procedure based on displacement equations of equilibrium and finite element modeling of the cross section has been presented for solution of the relaxed Saint-Venant problems for extension-bending-torsion and flexure for a prismatic inhomogeneous, anisotropic cylinder. The displacement field for extension-bending-torsion is based on a strain field independent of the axial coordinate and that for flexure is obtained by one integration of the extension-bending-torsion field with respect to z. This systematic method of setting forth the Saint-Venant displacement fields is due to Ieşan [6]. Many characteristics in Saint-Venant's solutions for homogeneous, isotropic cylinders were also apparent in the case of a cylinder with arbitrary cross-sectional properties. Axially independent stress field in extension-bending-torsion was exhibited. By virtue of variationally derived equilibrium equations and application of Saint-Venant kinematics, the Saint-Venant solutions for an inhomogeneous, anisotropic cross section do in fact manifest absolute minimum strain energy states among competing solutions whose end tractions may lead to the same integral end conditions.

Kosmatka and Dong's analysis ([10]) of a homogeneous, anisotropic cylinder differ from the present formulation, in that the former sets forth the complete displacement field for all six end resultants at the outset. Therefore, it was necessary to distinguish the warpages due to torsion from that due to flexure, which was possible through the use of an additional degree-of-freedom. The analysis enables all of the deformation coefficients to be identified with their corresponding resultant forces and moments. The present formulation, aside from its capability for treating inhomogeneous cross-sections, establishes the Saint-Venant solutions in a more systematic manner.

In the next paper, the cross-sectional properties of an inhomogeneous, anisotropic properties are discussed. In the third companion paper, quantitative analysis of Saint-Venant's principle is considered.

Acknowledgments

The authors thank Rokuro Muki for his critical reading of the manuscript and his many helpful suggestions.

References

[1] de Saint-Venant, A. J. C. B., 1856, ''Memoire sur la Torsion des Prismes,'' Mem. Savants Etrangers, **14**, pp. 233–560.
[2] de Saint-Venant, A. J. C. B., 1856, ''Memoire sur la Flexion des Prismes,'' J. Math. de Liouville, Ser. II, **1**, pp. 89–189.
[3] Clebsch, 1862,
[4] Voigt, W., 1887, ''Theoretische Studienüber die Elasticitätsverhältnisse der Krystalle,'' Gött. Abhandl., **34**, pp. 53–153
[5] Sternberg, E., and Knowles, J. K., 1966, ''Minimum Energy Characterizations of Saint-Venant's Solution to the Relaxed Saint-Venant Problem,'' Arch. Ration. Mech. Anal., **21**, No. 2, pp. 89–107.
[6] Ieşan, D., 1986, ''On Saint-Venant's Problem,'' Arch. Ration. Mech. Anal., **91**, pp. 363–373.
[7] Kosmatka, J. B., Lin, H. C., and Dong, S. B., 2001, ''On Saint-Venant's Problem for an Inhomogeneous, Anisotropic Cylinder, Part II: Cross-Sectional Properties,'' ASME J. Appl. Mech., 68
[8] Lin, H. C., Dong, S. B., and Kosmatka, J. B., 2001, ''On Saint-Venant's Problem for an Inhomogeneous, Anisotropic Cylinder—Part III: End Effects,'' ASME J. Appl. Mech., 68
[9] Kazic, M., and Dong, S. B., 1990, ''Analysis of Restrained Torsion,'' J. Eng. Mech., **116**, No. 4, pp. 870–891.
[10] Kosmatka, J. B., and Dong, S. B., 1991, ''Saint-Venant Solutions for Prismatic Anisotropic Beams,'' Int. J. Solids Struct., **28**, No. 7, pp. 917–938.
[11] Kantorovich, L. V., and Krylov, V., 1958, *Approximate Methods for Higher Analysis*, Noordhoff, Groningen, The Netherlands.
[12] Taweel, H., Dong, S. B., and Kazic, M., 2000, ''Wave Reflection from the Free End of a Cylinder with an Arbitrary Cross-Section,'' Int. J. Solids Struct., **37**, pp. 1701–1726.
[13] Almansi, E., 1901, ''Sopra la Deformazione dei Cilindri Solecitati Lateralmente,'' Atti Accad. Naz. Lincei, Cl. Sci. Fis., Mat. Nat., Rend., **10**, pp. I:333–338; II:400–408.
[14] Michell, J. H., 1901, ''The Theory of Uniformly Loaded Beams,'' Q. J. Math., **32**, pp. 28–42.

On Saint-Venant's Problem for an Inhomogeneous, Anisotropic Cylinder—Part II: Cross-Sectional Properties

J. B. Kosmatka
Department of Applied Mechanics and Engineering Science,
University of California,
San Diego, CA 92093-0085
Mem. ASME

H. C. Lin

S. B. Dong
Mem. ASME

Civil and Environmental Engineering Department,
University of California,
Los Angeles, CA 90095-1593

Cross-sectional properties of a prismatic inhomogeneous, anisotropic cylinder are determined from Saint-Venant solutions for extension-bending-torsion and flexure, whose method of construction was presented in a previous paper. The coupling of extensional, bending, and twisting deformations due to anisotropy and inhomogeneity leads to some very interesting features. Herein, it is shown that for an inhomogeneous, anisotropic cylinder whose cross-sectional plane is not a material symmetry plane, distinct modulus-weighted and compliance-weighted centroids and distinct principal bending axes are possible. A line of extension-bending centers is given on which an axial force causes extension and bending only but no twist. Two shear centers are given, one using the Griffith-Taylor definition that ignores cross-sectional warpages and the other by stipulating a zero mean rotation over the cross section. The center of twist is discussed, and this property depends on root end fixity conditions that are prescribed in terms of their mean values based on integrals over the cross section rather than by a pointwise specification. While these shear center and center of twist definitions have some rational bases, it is recognized that other definitions are possible, for example those based on modulus or compliance-weighted integrals. Two examples, an angle and a channel, both composed of a two-layer ±30 deg angle-ply composite material, illustrate the procedures for determining these cross-sectional properties. [DOI: 10.1115/1.1365152]

Introduction

Saint-Venant's problem is concerned with equilibrium of an elastic cylinder subjected to traction on its cross sections at the two ends, and Saint-Venant's solutions refer to those based on a relaxed formulation where the pointwise traction specification is replaced by integrals representing the end forces and moments. A method for constructing Saint-Venant's solutions for a cylinder whose cross section may be composed of any number of perfectly bonded anisotropic materials was presented by Dong, Kosmatka, and Lin [1], hereinafter called DKL. Their method was based on semi-analytical finite elements with the displacement field expressed in terms of nodal distributions and their corresponding amplitudes. These amplitudes are related to the force and moment resultants in the form of *cross-sectional stiffness relations*. In this paper, attention is devoted to defining section properties, such as centroids, principal bending axes, line of extension-bending centers, center of twist, and shear center. The Saint-Venant problem in itself does not address these topics, but the Saint-Venant solutions provide means for defining these cross-sectional properties.

A body of literature on finite element analyses of Saint-Venant's problem exists for homogeneous isotropic and anisotropic cylinders. Herrmann [2] and Mason and Herrmann [3] analyzed the torsion and flexure of homogeneous, isotropic beams of arbitrary cross-sectional shapes; Tolf [4] explored the bending of a homogeneous, orthotropic beam; Wörndle [5] considered a restricted class of inhomogeneous, monotropic beams; and Kosmatka and Dong [6] considered the extension-bending-torsion and flexure of a homogeneous, anisotropic beam. Many analytical issues on the Saint-Venant problem for anisotropic cylinders were discussed by Lekhnitskii [7], but no solutions were given for a completely anisotropic cylinder. Ieşan [8–10] also studied this problem and set forth a general procedure for generating the proper displacement fields for the extension-bending-torsion and flexure problems. His methodology was applied in DKL [1].

To define the various cross-sectional properties, two general coordinate transformations are needed, a translation and a rotation about an axis normal to the cross-sectional plane. These transformations are used for locating the centroid and principal bending axes. It will be seen that two distinct sets of these properties may be possible that are based either on the cross-sectional stiffness or flexibility relation. From the flexibility relation, a line of extension-bending centers can be defined. The location of the shear center is considered. The traditional definition by Griffith-Taylor [11] of this property does not take the warpage of the cross section into account. Because the Saint-Venant solutions provide warpages, other definitions of the shear center are possible. Lastly, the center of twist is discussed. Examples are given to illustrate these cross-sectional properties for two inhomogeneous, anisotropic cross sections.

Recapitulation of the Saint-Venant Solutions

For extension-bending-torsion of a uniform cantilevered cylinder of length L with an arbitrary cross section of perfectly bonded linearly elastic materials as shown in Fig. 1 of DKL [1], the finite element nodal displacement field has the form

$$\mathbf{U}(z)=[\mathbf{\Phi}_I(z)+\mathbf{\Psi}_I]\mathbf{a}_I+\mathbf{\Phi}_{RB}(z)\mathbf{a}_{RB} \tag{1}$$

where $\mathbf{U}(z)$ is an ordered $3m$ array of the nodal displacements of the finite element model with N number of nodes; $\mathbf{\Phi}_I$, $\mathbf{\Psi}_I$ and $\mathbf{\Phi}_{RB}$ are $3N\times 6$ matrices representing the primal field, cross-sectional warpages and rigid-body displacements, respectively; and $\mathbf{a}_I$ and $\mathbf{a}_{RB}$ are 6×1 arrays of amplitude coefficients. The

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 7, 1999; final revision, July 21, 2000. Associate Editor: J. W. Ju. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

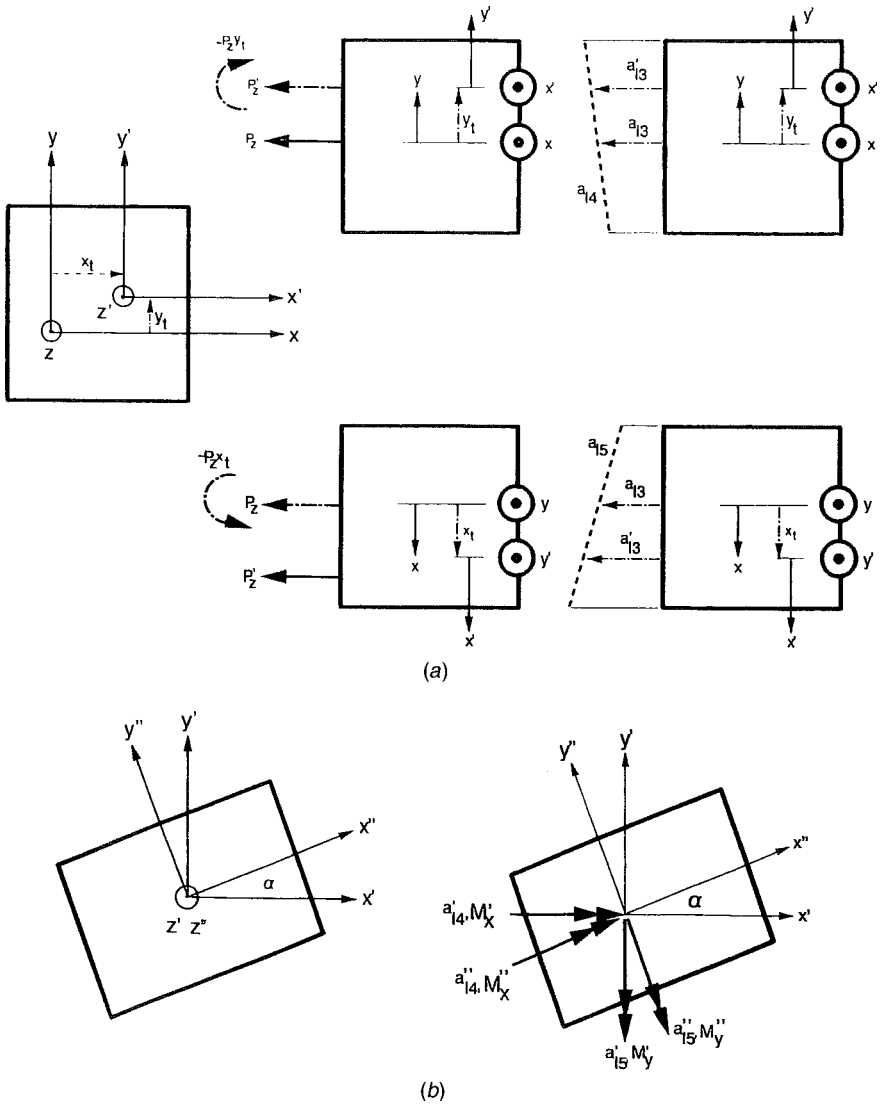


Fig. 1 (*a*) Coordinate transformation—translation; (*b*) coordinate transformation—rotation

stresses of this displacement field enables the construction of the *cross-sectional stiffness relation* $\boldsymbol{\kappa}_{Ibb}\mathbf{a}_{Ib}=\mathbf{F}_{Ib}$ relating the end force and moments to their corresponding deformational coefficients, that in expanded form is

$$\begin{bmatrix} \kappa_{I33} & \kappa_{I34} & \kappa_{I35} & \kappa_{I36} \\ \kappa_{I34} & \kappa_{I44} & \kappa_{I45} & \kappa_{I46} \\ \kappa_{I35} & \kappa_{I45} & \kappa_{I55} & \kappa_{I56} \\ \kappa_{I36} & \kappa_{I46} & \kappa_{I56} & \kappa_{I66} \end{bmatrix} \begin{Bmatrix} a_{I3} \\ a_{I4} \\ a_{I5} \\ a_{I6} \end{Bmatrix} = \begin{Bmatrix} P_z \\ M_x \\ -M_y \\ M_z \end{Bmatrix}. \tag{2}$$

The diagonal coefficients κ_{I33}, κ_{I44}, etc., of $\boldsymbol{\kappa}_{Ibb}$ are the extensional, bending, and torsional rigidities; $\kappa_{I34}, \kappa_{I35}$ are extension-bending coupling terms and κ_{I45} express mutual bending coupling about the (x,y)-axes; κ_{I36} is the extension-torsional coupling; and $(\kappa_{I46}, \kappa_{I56})$ are bending-torsional coupling terms. The inverse of Eq. (2) gives the cross-sectional *compliance* or *flexibility relation* $\mathbf{s}_{Ibb}\mathbf{F}_{Ib}=\mathbf{a}_{Ib}$ or in expanded form as

$$\begin{bmatrix} s_{I33} & s_{I34} & s_{I35} & s_{I36} \\ s_{I34} & s_{I44} & s_{I45} & s_{I46} \\ s_{I35} & s_{I45} & s_{I55} & s_{I56} \\ s_{I36} & s_{I46} & s_{I56} & s_{I66} \end{bmatrix} \begin{Bmatrix} P_z \\ M_x \\ -M_y \\ M_z \end{Bmatrix} = \begin{Bmatrix} a_{I3} \\ a_{I4} \\ a_{I5} \\ a_{I6} \end{Bmatrix}. \tag{3}$$

For flexure, two transverse forces, P_x and P_y, act on the tip end cross section but not necessarily through the shear center. The flexural displacement field has the form

$$\mathbf{U}(x,y,z)=[\boldsymbol{\Phi}_{II}(z)+z\boldsymbol{\Psi}_I+\boldsymbol{\Psi}_{II}]\mathbf{a}_{II}+[\boldsymbol{\Phi}_I(z)+\boldsymbol{\Psi}_I]\mathbf{b}_{II} +\boldsymbol{\Phi}_{RB}(z)\mathbf{a}_{RB} \tag{4}$$

where $\boldsymbol{\Phi}_I(z)$ and $\boldsymbol{\Psi}_I$ are the primal field and cross-sectional warpages of extension-bending-torsion, $\boldsymbol{\Phi}_{II}(z)$ and $\boldsymbol{\Psi}_{II}$ with $\mathbf{a}_{II}$ and $\mathbf{b}_{II}$ contain the additional sets of the primal field and cross-sectional warpages for flexure and their associated displacement amplitudes. Global equilibrium and end conditions give two relations for relating the end forces and moments to the displacement amplitudes, i.e., $\boldsymbol{\kappa}_{Ibb}\mathbf{a}_{IIb}=\mathbf{F}'_{IIb}$ and $\boldsymbol{\kappa}_{Ibb}\mathbf{b}_{IIb}=\mathbf{F}_{IIb}-\boldsymbol{\kappa}_{IIbb}\mathbf{a}_{IIb}$ which, in expanded form, are

$$\begin{bmatrix} \kappa_{I33} & \kappa_{I34} & \kappa_{I35} & \kappa_{I36} \\ \kappa_{I34} & \kappa_{I44} & \kappa_{I45} & \kappa_{I46} \\ \kappa_{I35} & \kappa_{I45} & \kappa_{I55} & \kappa_{I56} \\ \kappa_{I36} & \kappa_{I46} & \kappa_{I56} & \kappa_{I66} \end{bmatrix} \begin{Bmatrix} a_{II3} \\ a_{II4} \\ a_{II5} \\ a_{II6} \end{Bmatrix} = \begin{Bmatrix} 0 \\ P_y \\ P_x \\ 0 \end{Bmatrix} \tag{5}$$

$$\begin{bmatrix} \kappa_{I33} & \kappa_{I34} & \kappa_{I35} & \kappa_{I36} \\ \kappa_{I34} & \kappa_{I44} & \kappa_{I45} & \kappa_{I46} \\ \kappa_{I35} & \kappa_{I45} & \kappa_{I55} & \kappa_{I56} \\ \kappa_{I36} & \kappa_{I46} & \kappa_{I56} & \kappa_{I66} \end{bmatrix} \begin{Bmatrix} b_{II3} \\ b_{II4} \\ b_{II5} \\ b_{II6} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ P_y e_x + P_x e_y \end{Bmatrix}$$

$$- \begin{bmatrix} 0 & \kappa_{II34} & \kappa_{II35} & \kappa_{II36} \\ -\kappa_{II34} & 0 & \kappa_{II45} & \kappa_{II46} \\ -\kappa_{II35} & -\kappa_{II45} & 0 & \kappa_{II56} \\ -\kappa_{II36} & -\kappa_{II46} & -\kappa_{II56} & 0 \end{bmatrix} \begin{Bmatrix} a_{II3} \\ a_{II4} \\ a_{II5} \\ a_{II6} \end{Bmatrix} \quad (6)$$

where e_x and e_y are the x and y distances to the shear center. Coefficients (a_{II1}, a_{II2}) and (b_{II1}, b_{II2}) in $\mathbf{a}_{II}$ and $\mathbf{b}_{II}$ have no role in flexure and may be ignored.

Coordinate Transformations

To determine the various cross-sectional properties, two coordinate transformations are needed, viz., a translation (x_t, y_t) within the x-y plane and a counterclockwise rotation α about an axis normal to it as shown in Figs. 1(a) and 1(b). These two transformations will be taken in sequence, translation first and then rotation. The unprimed quantities are referenced to the original coordinate system, those with one prime are for the translated axes, and those with the double primes are for the rotated axes. For translation, the relation between (x,y) and (x',y') of a generic point in the old and new systems is given by the equations

$$x = x' + x_t; \quad y = y' + y_t. \quad (7)$$

For a counterclockwise rotation of angle α, the transformation has the form

$$x'' = x' \cos\alpha + y' \sin\alpha; \quad y'' = -x' \sin\alpha + y' \cos\alpha. \quad (8)$$

Referring to Fig. 1(a), the formulas for transformations of displacements and their corresponding force and moment resultants by translation between the two systems are

$$\mathbf{a}'_{Ib} = \mathbf{T}_1 \mathbf{a}_{Ib} \leftrightarrow \mathbf{a}_{Ib} = \mathbf{T}_2 \mathbf{a}'_{Ib} \quad (9)$$

$$\mathbf{F}'_{Ib} = \mathbf{T}_2^T \mathbf{F}_{Ib} \leftrightarrow \mathbf{F}_{Ib} = \mathbf{T}_1^T \mathbf{F}'_{Ib}, \quad (10)$$

where matrices $\mathbf{T}_1$ and $\mathbf{T}_2$ are given in the Appendix. Applying translation to Eqs. (2) and (3) gives the new cross-sectional stiffness and compliance matrices in terms of those of the old system as

$$\boldsymbol{\kappa}'_{Ibb} = \mathbf{T}_2^T \boldsymbol{\kappa}_{Ibb} \mathbf{T}_2 \quad (11)$$

$$\mathbf{s}'_{Ibb} = \mathbf{T}_1 \mathbf{s}_{Ibb} \mathbf{T}_1^T \quad (12)$$

where the components of $\boldsymbol{\kappa}'_{Ibb}$ and $\mathbf{s}'_{Ibb}$ are given in Table 1.

Referring to Fig. 1(b), rotation of displacements and their corresponding forces and moments by a counterclockwise angle α about the z-axis takes the forms

$$\mathbf{a}''_{Ib} = \mathbf{T}_3 \mathbf{a}'_{Ib} \leftrightarrow \mathbf{a}'_{Ib} = \mathbf{T}_4 \mathbf{a}''_{Ib} \quad (13)$$

$$\mathbf{F}''_{Ib} = \mathbf{T}_4^T \mathbf{F}'_{Ib} \leftrightarrow \mathbf{F}'_{Ib} = \mathbf{T}_3^T \mathbf{F}''_{Ib} \quad (14)$$

where $\mathbf{T}_3$ and $\mathbf{T}_4$ are given in the Appendix. Applying rotation to Eqs. (11) and (12) gives

$$\boldsymbol{\kappa}''_{Ibb} = \mathbf{T}_4^T \boldsymbol{\kappa}'_{Ibb} \mathbf{T}_4 \quad (15)$$

$$\mathbf{s}''_{Ibb} = \mathbf{T}_3 \mathbf{s}'_{Ibb} \mathbf{T}_3^T \quad (16)$$

where the components of $\boldsymbol{\kappa}''_{Ibb}$ and $\mathbf{s}''_{Ibb}$ are given in Table 2.

Transformations (9) (10) and (13) (14) form a *contragradient pair*. Tables 1 and 2 show that $\kappa_{I33}, \kappa_{I36}, \kappa_{I66}$ remain invariant in both coordinate transformations, but only s_{I66} is invariant in both transformation.

Table 1 Relations of κ'_{ij} and S'_{ij} in terms of κ_{ij} and S_{ij} for translation

$\kappa'_{I33} = \kappa_{I33}$	$s'_{I33} = s_{I33} + s_{I44} y_t^2 + s_{I55} x_t^2 + 2(s_{I34} y_t + s_{I35} x_t + s_{I45} x_t y_t)$
$\kappa'_{I44} = \kappa_{I44} + \kappa_{I33} y_t^2 - 2\kappa_{I34} y_t$	$s'_{I44} = s_{I44}$
$\kappa'_{I55} = \kappa_{I55} + \kappa_{I33} x_t^2 - 2\kappa_{I35} x_t$	$s'_{I55} = s_{I55}$
$\kappa'_{I66} = \kappa_{I66}$	$s'_{I66} = s_{I66}$
$\kappa'_{I34} = \kappa_{I34} - \kappa_{I33} y_t$	$s'_{I34} = s_{I34} + s_{I44} y_t + s_{I45} x_t$
$\kappa'_{I35} = \kappa_{I35} - \kappa_{I33} x_t$	$s'_{I35} = s_{I35} + s_{I45} y_t + s_{I55} x_t$
$\kappa'_{I36} = \kappa_{I36}$	$s'_{I36} = s_{I36} + s_{I46} y_t + s_{I56} x_t$
$\kappa'_{I45} = \kappa_{I45} + \kappa_{I33} x_t y_t - \kappa_{I34} x_t - \kappa_{I35} y_t$	$s'_{I45} = s_{I45}$
$\kappa'_{I46} = \kappa_{I46} - \kappa_{I36} y_t$	$s'_{I46} = s_{I46}$
$\kappa'_{I56} = \kappa_{I56} - \kappa_{I36} x_t$	$s'_{I56} = s_{I56}$

where x_t and y_t are the translation from the old to new axes.

Table 2 Relations of κ''_{ij} and S''_{ij} in terms of κ'_{ij} and S'_{ij} for rotation

$\kappa''_{I33} = \kappa_{I33}$	$s''_{I33} = s'_{I33}$
$\kappa''_{I44} = \kappa'_{I44} c^2 + \kappa'_{I55} s^2 + 2\kappa'_{I45} sc$	$s''_{I44} = s'_{I44} c^2 + s'_{I55} s^2 - 2s'_{I45} sc$
$\kappa''_{I55} = \kappa'_{I55} c^2 + \kappa'_{I44} s^2 + 2\kappa'_{I45} sc$	$s''_{I55} = s'_{I55} c^2 + s'_{I44} s^2 + 2s'_{I45} sc$
$\kappa''_{I66} = \kappa_{I66}$	$s''_{I66} = s_{I66}$
$\kappa''_{I34} = \kappa'_{I34} c - \kappa'_{I35} s$	$s''_{I34} = s'_{I34} c - s'_{I35} s$
$\kappa''_{I35} = \kappa'_{I35} c + \kappa'_{I34} s$	$s''_{I35} = s'_{I35} c + s'_{I34} s$
$\kappa''_{I36} = \kappa_{I36}$	$s''_{I36} = s'_{I36}$
$\kappa''_{I45} = \kappa'_{I45}(c^2 - s^2) + (\kappa'_{I44} - \kappa'_{I55}) sc$	$s''_{I45} = s'_{I45}(c^2 - s^2) + (s'_{I44} - s'_{I55}) sc$
$\kappa''_{I46} = \kappa'_{I46} c - \kappa'_{I56} s$	$s''_{I46} = s'_{I46} c - s'_{I56} s$
$\kappa''_{I56} = \kappa'_{I56} c + \kappa'_{I46} s$	$s''_{I56} = s'_{I56} c + s'_{I46} s$

where $s = \sin\alpha$ and $c = \cos\alpha$; α = angle measured counterclockwise from old to new axes. Unprimed quantities indicate invariance to both translation and rotation.

Modulus and Compliance-Weighted Centroids and Principal Bending Axes

The stiffness and compliance relations (2) and (3) depend on a coordinate system with origin located at some point on the tip end cross-sectional plane. There are no *a priori* provisions in the Saint-Venant problem for choosing an origin. Therefore, the stiffness and compliance matrices in general will be densely populated evincing full coupling of extensional, bending, and torsional behaviors. Certain terms, nevertheless, can be annihilated by suitable choices of the origin and orientation of the coordinate axes. Herein, transformations are presented to rid the stiffness and compliance relations of coupling between extension and bending and between the bending effects about the two orthogonal axes. This location and orientation express the *centroid* and the *principal bending* axes of the cross section. For a general anisotropic inhomogeneous cylinder, two distinct centroids, (1) a *modulus-weighted centroid* and (2) a *compliance-weighted centroid*, as well as two sets of principal bending axes are possible.

Translation of $\boldsymbol{\kappa}_{Ibb}$ to the *modulus-weighted centroid* requires that $\kappa'_{I34} = \kappa'_{I35} = 0$. Setting these terms equal to zero from their expressions in Table 1 gives

$$x_t = \frac{\kappa_{I35}}{\kappa_{I33}} \quad ; \quad y_t = \frac{\kappa_{I34}}{\kappa_{I33}} \quad (17)$$

so that

$$\begin{bmatrix} \kappa_{I33} & \kappa_{I34} & \kappa_{I35} & \kappa_{I36} \\ & \kappa_{I44} & \kappa_{I45} & \kappa_{I46} \\ & & \kappa_{I55} & \kappa_{I56} \\ sym. & & & \kappa_{I66} \end{bmatrix} \rightarrow \begin{bmatrix} \kappa_{I33} & 0 & 0 & \kappa_{I36} \\ & \kappa'_{I44} & \kappa'_{I45} & \kappa'_{I46} \\ & & \kappa'_{I55} & \kappa'_{I56} \\ sym. & & & \kappa_{I66} \end{bmatrix} \quad (18)$$

where the other κ'_{Iij} coefficients are given in Table 1. The unprimed quantities on the right-hand side of Eq. (18) denote an invariance with respect to translation, i.e., the terms κ_{I33}, κ_{I36}, and κ_{I66}.

Rotation of the coordinate axes to the principal bending axes eliminates mutual coupling of bending effects about both orthogonal axes, so that κ''_{I45} must vanish. Setting $\kappa''_{I45}=0$ from Table 2 gives

$$\tan 2\alpha = \frac{2\kappa'_{I45}}{\kappa'_{I55}-\kappa'_{I44}} \tag{19}$$

so that

$$\begin{bmatrix} \kappa_{I33} & 0 & 0 & \kappa_{I36} \\ & \kappa'_{I44} & \kappa'_{I45} & \kappa'_{I46} \\ & & \kappa'_{I55} & \kappa'_{I56} \\ sym. & & & \kappa_{I66} \end{bmatrix} \rightarrow \begin{bmatrix} \kappa_{I33} & 0 & 0 & \kappa_{I36} \\ & \kappa''_{I44} & 0 & \kappa''_{I46} \\ & & \kappa''_{I55} & \kappa''_{I56} \\ sym. & & & \kappa_{I66} \end{bmatrix} \tag{20}$$

where the κ''_{Iij} coefficients are given in Table 2. The unprimed quantities on the right-hand side of Eq. (20) indicates an invariance with respect to both translation and rotation. Consider the form of the stiffness matrix given by the right-handed side of Eq. (20) as the *canonical form*, where all extensional and flexural coupling are absent.

Translation of $\mathbf{s}_{Ibb}$ to the *compliance-weighted centroid* yields $s'_{I34}=s'_{I35}=0$, which from Table 1 gives the following equations and solution:

$$\begin{bmatrix} s_{I44} & s_{I45} \\ s_{I45} & s_{I55} \end{bmatrix} \begin{Bmatrix} y_t \\ x_t \end{Bmatrix} = \begin{Bmatrix} s_{I34} \\ s_{I35} \end{Bmatrix}$$

$$\begin{Bmatrix} x_t \\ y_t \end{Bmatrix} = \frac{1}{s_{I44}s_{I55}-s_{I45}^2} \begin{Bmatrix} s_{I34}s_{I45}-s_{I44}s_{I35} \\ s_{I35}s_{I45}-s_{I55}s_{I34} \end{Bmatrix} \tag{21}$$

so that

$$\begin{bmatrix} s_{I33} & s_{I34} & s_{I35} & s_{I36} \\ & s_{I44} & s_{I45} & s_{I46} \\ & & s_{I55} & s_{I56} \\ sym. & & & s_{I66} \end{bmatrix} \rightarrow \begin{bmatrix} s'_{I33} & 0 & 0 & s'_{I36} \\ & s_{I44} & s_{I45} & s_{I46} \\ & & s_{I55} & s_{I56} \\ sym. & & & s_{I66} \end{bmatrix} \tag{22}$$

where the formulas for s'_{I33} and s'_{I36} are given in Table 1.

A counterclockwise α-rotation of the z'-axis through compliance-weighted centroid to principal bending axes yields $s''_{I45}=0$. By the formula for s''_{I45} from Table 2, α is given by

$$\tan 2\alpha = \frac{2s_{I45}}{s_{I55}-s_{I44}} \tag{23}$$

so that

$$\begin{bmatrix} s'_{I33} & 0 & 0 & s'_{I36} \\ & s_{I44} & s_{I45} & s_{I46} \\ & & s_{I55} & s_{I56} \\ sym. & & & s_{I66} \end{bmatrix} \rightarrow \begin{bmatrix} s'_{I33} & 0 & 0 & s'_{I36} \\ & s''_{I44} & 0 & s''_{I46} \\ & & s''_{I55} & s''_{I56} \\ sym. & & & s_{I66} \end{bmatrix} \tag{24}$$

where s''_{Iij}'s are given in Table 2. Analogous to the canonical form of the stiffness matrix, the right-hand side of Eq. (24) is the canonical form of the compliance matrix.

The coefficient enabling two distinct centroids is the term, κ_{I36} or s_{I36}. In canonical forms, the presence of either of these coefficients is dependent upon the following characteristics of the cross section. For all homogeneous cross sections, regardless of the extent of the anisotropy in the material properties, κ_{I36} or s_{I36} will not occur, and there is no distinction between the two centroids. The extension-torsion coupling coefficient is present only for an *inhomogeneous cross section* that also lacks *material symmetry* with respect to a generic cross-sectional plane. For the most general form of anisotropy possessing symmetry about a cross section plane, the elastic moduli **C** matrix contains the following terms:

$$\mathbf{C}=\begin{bmatrix} C_{11} & C_{12} & C_{13} & \cdot & \cdot & C_{16} \\ \cdot & C_{22} & C_{23} & \cdot & \cdot & C_{26} \\ \cdot & \cdot & C_{33} & \cdot & \cdot & C_{36} \\ \cdot & \cdot & \cdot & C_{44} & C_{45} & \cdot \\ \cdot & \cdot & \cdot & \cdot & C_{55} & \cdot \\ & sym. & & & & C_{66} \end{bmatrix} \tag{25}$$

which is that for a monotropic material. Thus, the appearance of any of the coefficients beyond those shown in Eq. (25) in one of the materials in an inhomogeneous cross section precipitates κ_{I36}, i.e., the presence of any one of the coefficients $C_{14}, C_{15}, C_{24}, C_{25}, C_{34}, C_{35}, C_{46}, C_{56}$. Note that for an inhomogeneous, anisotropic cross section with κ_{I36} present, the inverse of the canonical form of the stiffness matrix will result, in general, in a densely populated compliance matrix, i.e.,

$$\begin{bmatrix} \kappa_{I33} & 0 & 0 & \kappa_{I36} \\ & \kappa_{I44} & 0 & \kappa_{I46} \\ & & \kappa_{I55} & \kappa_{I56} \\ sym. & & & \kappa_{I66} \end{bmatrix}^{-1} = \begin{bmatrix} s_{I33} & s_{I34} & s_{I35} & s_{I36} \\ & s_{I44} & s_{I45} & s_{I46} \\ & & s_{I55} & s_{I56} \\ & sym. & & s_{I66} \end{bmatrix} \tag{26}$$

so that the presence of s_{I34} and s_{I35} enables the possibility of two centroids.

The bending-torsion coupling coefficients, κ_{I46} and κ_{I56}, can occur in both homogeneous and inhomogeneous cross-sections lacking *material symmetry* with any cross-sectional plane, i.e., with the presence of any coefficient beyond those shown in Eq. (25).

Line of Extension-Bending Centers

Flexibility relation (3) at the compliance-weighted centroid has the form

$$\begin{bmatrix} s'_{I33} & 0 & 0 & s'_{I36} \\ & s_{I44} & s_{I45} & s_{I46} \\ & & s_{I55} & s_{I56} \\ sym. & & & s_{I66} \end{bmatrix} \begin{Bmatrix} P_z \\ M_x \\ -M_y \\ M_z \end{Bmatrix} = \begin{Bmatrix} a_{I3} \\ a_{I4} \\ a_{I5} \\ a_{I6} \end{Bmatrix}. \tag{27}$$

This relation shows that an axial force P_z by itself with resultant passing through the compliance-weighted centroid causes extensional, bending, and twisting deformations. By shifting P_z to a point with coordinates (x_{eb}, y_{eb}), twisting can be suppressed leaving it to produce extensional and bending deformations only. There exist an infinity of such points, however, lying on a straight line that can be called the *line of extension-bending centers*. This line is defined by setting $a_{I6}=0$ in the fourth line of Eq. (27) and letting $P_z=1$, so that $M_x=-y_{eb}$ and $M_y=-x_{eb}$, to give

$$y'_{eb}=\frac{s_{I56}}{s_{I46}}x'_{eb}+\frac{s'_{I36}}{s_{I46}}. \tag{28}$$

For a homogeneous cross section, the line of extension-bending centers passes through the centroid as $s_{I36}=0$, see Kosmatka and Dong [6]. Otherwise, this line is offset from the compliance-weighted centroid when s_{I36} occurs.

Shear Center

The shear center is that point in the cross-sectional plane through which the application of a transverse force causes no twist. It is generally regarded as a property of the cross section and is independent of the fixity conditions at the end of the cylinder; see Timoshenko and Goodier [12] for a discussion of this point.

For shear center calculations, use the modulus-weighted centroid in the tip cross-sectional plane as the origin. The shear center location requires definition of the *rotation rate* or the *unit angle of twist* θ as

$$\theta(x,y) \overset{\text{def}}{\equiv} \frac{1}{2}\frac{\partial}{\partial z}\left[\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right]. \tag{29}$$

Partial differentiations, $\partial v/\partial x$ and $\partial u/\partial y$ of Eq. (18*a*) of DKL [1] according to definition (29), give the rotation rate θ for torque M_z as

$$\theta(x,y) = a_{I6} \quad \text{a constant.} \tag{30}$$

Under torque M_z by itself, the rotation rate θ is constant, and a_{I3}, a_{I4}, and a_{I5} can be expressed directly in terms of a_{I6} from stiffness matrix (20).

$$a_{I3} = -\frac{\kappa_{36}}{\kappa_{33}}a_{I6};\quad a_{I4} = -\frac{\kappa_{46}}{\kappa_{44}}a_{I6};\quad a_{I5} = -\frac{\kappa_{56}}{\kappa_{55}}a_{I6} \tag{31}$$

Then, by the fourth equation of stiffness relation (2), the following torque-twist relation is obtained:

$$M_z = \overline{GJ}\,\theta \tag{32}$$

where $\overline{GJ}$ is the *composite torsional rigidity*[1] of the cross section given by

$$\overline{GJ} = \kappa_{66} - \frac{\kappa_{36}^2}{\kappa_{33}} - \frac{\kappa_{46}^2}{\kappa_{44}} - \frac{\kappa_{56}^2}{\kappa_{55}}. \tag{33}$$

The rotation rate θ for flexure by applying Eq. (29) to Eq. (29) of DKL [1] has the form

$$\theta(x,y) = a_{II6}z + b_{II6} + \frac{1}{2}\sum_{i=3}^{6} a_{IIi}(\psi_{Iiv,x}(x,y) - \psi_{Iiu,y}(x,y)). \tag{34}$$

Here, $\theta = \theta(x,y)$ due to cross-sectional warpage, in stark contrast to that of M_z. Note that even though a transverse shear force is involved in the shear center, only the warpages of the extension-bending-torsion problem participate in Eq. (34) as none of the warpages ψ_{IIi}'s from the flexure problem enter into the rotation rate.

After these explanatory comments, a general definition of the shear center can now be stated. It is that point in any cross-sectional plane through which the application of a transverse force produces no θ. Consider a unit force $P_x = 1$ passing through the shear center that is at a distance e_y from the centroid. A transverse force P_y at a distance e_x from the centroid can be similarly treated in a separate calculation. Translation of this unit force from the shear center to the centroid causes an accompanying torsional moment $M_z = e_y$. Superposition of rotation rates from pure torsion with $M_z = e_y$ and flexure with $P_x = 1$ passing through the modulus-weighted centroid yields the following equation:

$$\theta|_{\text{torsion}} + \theta|_{\text{flexure}} = 0. \tag{35}$$

Inserting Eqs. (32) and Eq. (34) into Eq. (35) and solving for e_y give

$$e_y(x,y,z) = -\overline{GJ}\left[a_{II6}z + b_{II6} + \frac{1}{2}\sum_{i=3}^{6} a_{IIi}(\psi_{Iiv,x}(x,y) - \psi_{Iiu,y}(x,y))\right]. \tag{36}$$

This equation shows e_y to the shear center as a *pointwise* function of all three coordinates (x,y,z). One possible way to proceed is to define a *zero-mean rotation shear center* e_{y_M}, obtained by averaging e_y over the cross-section.[2]

$$e_{y_M} \equiv \frac{1}{A}\int\int_A e_y(x,y,z)\,dx\,dy = -\overline{GJ}\left[a_{II6}z + b_{II6} + \sum_{i=3}^{6} a_{IIi}\Omega_i\right] \tag{37}$$

where A is the cross-sectional area and Ω_i is given by

$$\Omega_i = \frac{1}{2A}\int\int_A (\psi_{Iiv,x}(x,y) - \psi_{Iiu,y}(x,y))\,dx\,dy \quad (i = 3 \text{ to } 6). \tag{38}$$

The line of shear centers is parallel to the z-axis if a_{II6} is absent from Eq. (37), i.e., for a cross-sectional plane also as a material symmetry plane. For cylinders without such material symmetry, the line of shear centers varies linearly along the z-axis.[3]

The classical definition of the shear center due to Griffith-Taylor [11] does not consider warpage. Disregarding warpage in Eq. (36) gives this shear center as

$$e_{y_{GT}} = -\overline{GJ}[a_{II6}z + b_{II6}]. \tag{39}$$

Reissner's shear center definition [14,15] rests on the condition that it coincides with the center of twist at the root end. But, the center of twist (discussed in the next section) depends upon boundary conditions. Hence, Reissner's definition would also depend on boundary conditions, so that it will not be a property solely of the cross section. But the boundary conditions at the root end involves a particular value of the rigid-body displacement only. Since any two Saint-Venant displacement fields can only differ by a rigid-body displacement, Reissner's line of shear centers will be shifted but remain parallel to that of Eq. (37), a result shown by Kosmatka [16,17].

Since coefficients b_{IIi}'s in Eq. (6) depend on the choice of the origin, a consistency check of $\mathbf{F}_{II}$ in Eq. (39) of DKL [1] may be carried out by taking

$$\mathbf{F}_{II} = [P_1 = 1,\ P_2 = 0,\ 0,\ 0,\ 0,\ e_{y_i}] \tag{40}$$

where e_{y_i} is the y-distance to the shear center by either one of the two definitions, $e_{y_i} = e_{y_M}$, or $e_{y_{GT}}$, or e_{y_R}. This calculation will yield different values for coefficients a_{II3} to a_{II6} and b_{II6}, but it should give the same physical location of the shear center.

Center of Twist

The center of twist is that point at rest in every cross section of the cylinder subjected to a torque at the tip end and restrained from translation and rotation at the root end. It is dependent upon the root end boundary conditions; see Timoshenko and Goodier [12] for a discussion of this point. Therefore, it is necessary to discuss these conditions first. For a cantilevered cylinder of length L under a unit torque, the rotation rate θ is constant over its entire

[1] Lekhnitskii [7], (p. 180) calls this quantity the *generalized torsional rigidity*. In his case, the generalized torsional rigidity accounts for the torsion/bending coupling but not the torsion/axial force coupling inasmuch as he only considered a homogeneous anisotropic cylinder where $\kappa_{36} = 0$.

[2] Novozhilov [13] employed the same averaging procedure in his discussion of the shear center of a homogeneous, isotropic cylinder. He further showed that the location of the shear center depended upon the torsion problem only, and that the flexure problem need not be involved.

[3] Libove [18], in his investigation of twist rates in thin-walled members, offered a definition of a mean value shear center over the length of the beam by integrating an equation similar to Eq. (37) over the length, which essentially gives the shear center location at midlength, $z = L/2$. In his discussion of this point, he immediately questions the utility of such a definition.

length and equal to $1/\overline{GJ}$. Displacement field (18a) of DKL [1] with a_{Ii}'s given by Eqs. (30) and (31) takes the form

$$u(x,y,z)=\frac{1}{\overline{GJ}}\left[-yz+\frac{\kappa_{I56}}{\kappa_{I55}}\frac{z^2}{2}-\frac{\kappa_{I36}}{\kappa_{I33}}\psi_{I3u}(x,y)-\frac{\kappa_{I46}}{\kappa_{I44}}\psi_{I4u}(x,y)-\frac{\kappa_{I56}}{\kappa_{I55}}\psi_{I5u}(x,y)+\psi_{I6u}(x,y)\right]-\omega_3 y-\omega_2 z+u_o$$

$$v(x,y,z)=\frac{1}{\overline{GJ}}\left[xz+\frac{\kappa_{I46}}{\kappa_{I44}}\frac{z^2}{2}-\frac{\kappa_{I36}}{\kappa_{I33}}\psi_{I3v}(x,y)-\frac{\kappa_{I46}}{\kappa_{I44}}\psi_{I4v}(x,y)-\frac{\kappa_{I56}}{\kappa_{I55}}\psi_{I5v}(x,y)+\psi_{I6v}(x,y)\right]+\omega_3 x-\omega_1 z+v_o \tag{41}$$

$$w(x,y,z)=\frac{1}{\overline{GJ}}\left[-\frac{\kappa_{I36}}{\kappa_{I33}}z-\frac{\kappa_{I56}}{\kappa_{I55}}xz-\frac{\kappa_{I46}}{\kappa_{I44}}yz-\frac{\kappa_{I36}}{\kappa_{I33}}\psi_{I3w}(x,y)-\frac{\kappa_{I46}}{\kappa_{I44}}\psi_{I4w}(x,y)-\frac{\kappa_{I56}}{\kappa_{I55}}\psi_{I5w}(x,y)+\psi_{I6w}(x,y)\right]+\omega_1 y+\omega_2 x+w_o$$

where constants u_o, v_o, w_o, ω_1, ω_2, and ω_3 are rigid-body translations and rotations.

In principle, the six rigid-body displacement components allow sufficient mathematical freedom to satisfy any arbitrary condition but at one point only for each of the translations and rotations. Lekhnitskii [7] used the following conditions in his pure bending example, which at (x_o,y_o,L) of the root end, are

$$u(x_o,y_o,L)=v(x_o,y_o,L)=w(x_o,y_o,L)=0$$

$$\frac{\partial u(x_o,y_o,L)}{\partial z}=\frac{\partial v(x_o,y_o,L)}{\partial z}=\frac{\partial v(x_o,y_o,L)}{\partial x}-\frac{\partial u(x_o,y_o,L)}{\partial y}=0. \tag{42}$$

He has also used the conditions $\omega_1=0$ and $\omega_2=0$ for other examples elsewhere in his text. Restraint conditions (42) will not be used since (1) fixing a point from both translation and rotation is not physically realizable, and (2) conditions $\partial u(x_o,y_o,L)/\partial z=0$ and $\partial v(x_o,y_o,L)/\partial z=0$ do not strictly express full rotational restraint. Note also that applying Eq. (42) to another point only change the rigid-body components in Eq. (41), with the net effect being merely another rigid-body displacement. In a prototype physical case, many points may be partially and/or completely restrained. One way to rationally represent this condition is to insist on the vanishing of the mean cross-sectional translations, $\overline{u}$ and $\overline{v}$, and rotation about the cylinder's generator, $(\overline{v}_{,x}-\overline{u}_{,y})/2$ at the root end $z=L$, where the mean values are defined as their cross-sectional averages.

$$\overline{u}(z)\equiv\frac{1}{A}\int\int u(x,y,z)dxdy=0;$$

$$\overline{v}(z)\equiv\frac{1}{A}\int\int v(x,y,z)dxdy=0$$

$$\frac{1}{2}(\overline{v}_{,x}(z)-\overline{u}_{,y}(z))\equiv\frac{1}{2A}\int\int\{v_{,x}(x,y,z)-u_{,y}(x,y,z)\}dxdy=0 \tag{43}$$

Substituting displacement field (41) to Eq. (43) gives the following equations in terms of components $(u_o,v_o,\omega_1,\omega_2,\omega_3,)$

$$u_o-\omega_2 L-\omega_3\overline{y}=-\frac{1}{\overline{GJ}}\left[-\overline{y}L+\frac{\kappa_{I56}}{\kappa_{I55}}\frac{L^2}{2}-\frac{\kappa_{I36}}{\kappa_{I33}}U_3-\frac{\kappa_{I46}}{\kappa_{I44}}U_4-\frac{\kappa_{I56}}{\kappa_{I55}}U_5+U_6\right] \tag{44}$$

$$v_o-\omega_1 L+\omega_3\overline{x}=-\frac{1}{\overline{GJ}}\left[\overline{x}L+\frac{\kappa_{I46}}{\kappa_{I55}}\frac{L^2}{2}-\frac{\kappa_{I36}}{\kappa_{I33}}V_3-\frac{\kappa_{I46}}{\kappa_{I44}}V_4-\frac{\kappa_{I56}}{\kappa_{I55}}V_5+V_6\right] \tag{45}$$

$$\omega_3=-\frac{L}{\overline{GJ}}\left[-\frac{\kappa_{I36}}{\kappa_{I33}}\Omega_3-\frac{\kappa_{I46}}{\kappa_{I44}}\Omega_4-\frac{\kappa_{I56}}{\kappa_{I55}}\Omega_5+\Omega_6\right] \tag{46}$$

where $\overline{x}$ and $\overline{y}$ are the coordinates of the geometric centroid and

$$\left.\begin{aligned}U_i&=\frac{1}{A}\int\int\psi_{Iiu}(x,y)dxdy\\V_i&=\frac{1}{A}\int\int\psi_{Iiv}(x,y)dxdy\end{aligned}\right\}\ (i=3\ \text{to}\ 6) \tag{47}$$

and Ω_i's was given previously by Eq. (38). Additional equations come from minimizing the integral J of square of the displacement $w(x,y,L)$ at the root end where

$$J\equiv\int\int w^2(x,y,L)dxdy \tag{48}$$

with respect to ω_1, ω_2, and w_o, i.e., $\partial J/\partial w_o=0$, $\partial J/\partial\omega_2=0$, and $\partial J/\partial\omega_1=0$,

$$\int\int w(x,y,L)dxdy=0;$$

$$\int\int xw(x,y,L)dxdy=0;\quad\int\int yw(x,y,L)dxdy=0. \tag{49}$$

These minimizations supplant the conditions of the vanishing of the axial displacement w and slopes, $\partial w/\partial x$ and $\partial w/\partial y$, over the root end cross section. Weinstein [19] used this procedure, and he based it on an idea of Cicala [20] and Trefftz [21]. Substituting the third equation of (41) into Eq. (49) gives the following three equations, which together with Eqs. (44)–(46) enable the solution of the rigid-body displacement

$$w_o+\overline{y}\omega_1+\overline{x}\omega_2=-\frac{1}{\overline{GJ}}\left[-\frac{\kappa_{I36}}{\kappa_{I33}}(L+W_3)-\frac{\kappa_{I46}}{\kappa_{I44}}L(\overline{y}L+W_4)-\frac{\kappa_{I56}}{\kappa_{I55}}L(\overline{x}L+W_5)+W_6\right] \tag{50}$$

$$\overline{y}w_o+\frac{I_{xx}}{A}\omega_1+\frac{I_{xy}}{A}\omega_2=-\frac{1}{\overline{GJ}}\left[-\frac{\kappa_{I36}}{\kappa_{I33}}(\overline{x}L+W_{3\overline{x}})-\frac{\kappa_{I56}}{\kappa_{I55}}\left(\frac{I_{yy}}{A}L+W_{5\overline{x}}\right)-\frac{\kappa_{I46}}{\kappa_{I44}}\left(\frac{I_{xy}}{A}L+W_{4\overline{x}}\right)+W_{6\overline{x}}\right] \quad (51)$$

$$\overline{x}w_o+\frac{I_{xy}}{A}\omega_1+\frac{I_{yy}}{A}\omega_2=-\frac{1}{\overline{GJ}}\left[-\frac{\kappa_{I36}}{\kappa_{I33}}(\overline{y}L+W_{3\overline{y}})-\frac{\kappa_{I56}}{\kappa_{I55}}\left(\frac{I_{xy}}{A}L+W_{5\overline{y}}\right)-\frac{\kappa_{I46}}{\kappa_{I44}}\left(\frac{I_{yy}}{A}L+W_{4\overline{y}}\right)+W_{6\overline{y}}\right] \quad (52)$$

where I_{xx}, I_{yy}, and I_{xy} are planar moments and product of inertia with respect to the modulus-weighted centroidal axes and

$$(W_i, W_{i\overline{x}}, W_{i\overline{y}}) = \frac{1}{A}\int\int (\psi_{Iiw}(x,y), x\psi_{Iiw}(x,y), y\psi_{Iiw}(x,y))dxdy \quad (i=3 \text{ to } 6). \quad (53)$$

Returning to the center of twist location, its definition should be based on torsional deformations only. Hence, the extensional and bending terms in Eq. (41) should be discarded.

$$u(x,y,z)=\frac{1}{\overline{GJ}}[-yz+\psi_{I6u}(x,y)]-\omega_3 y-\omega_2 z+u_o$$

$$v(x,y,z)=\frac{1}{\overline{GJ}}[xz+\psi_{I6v}(x,y)]+\omega_3 x-\omega_1 z+v_o \quad (54)$$

$$w(x,y,z)=\frac{1}{\overline{GJ}}\psi_{I6w}(x,y)\Big]+\omega_1 y+\omega_2 x+w_o$$

or in terms of the nodal displacements of the finite element model as

$$\mathbf{U}(z)=\frac{1}{\overline{GJ}}[z\boldsymbol{\phi}_6+\boldsymbol{\psi}_6]+\boldsymbol{\Phi}_{RB}(z)\mathbf{a}_{RB} \quad (55)$$

where $\boldsymbol{\phi}_6$, $\boldsymbol{\psi}_6$ and $\boldsymbol{\Phi}_{RB}(z)$ are given by Eq. (15) of DKL [1].

Referring to Fig. 2 and according to Eq. (55), the total displacement Δ_i of ith node in the finite element model and its orientation angle η_i with respect to the x-axis arc

$$\Delta_i(z)=\sqrt{u_i^2(z)+v_i^2(z)};\quad \eta_i(z)=\tan^{-1}[u_i(z)/v_i(z)]. \quad (56)$$

Since θ is the unit angle of twist, the total twist at station z is $\theta(L-z)$, so that the ith node rotated about the center of twist by the amount

$$R_i\theta(L-z)=\Delta_i \quad (57)$$

where R_i is the distance between point (x_i, y_i) and the center of twist of coordinates (x_{ct}, y_{ct}).

$$R_i^2=(x_i-x_{ct})^2+(y_i-y_{ct})^2 \quad (58)$$

Different nodes yield different values of x_{ct} and y_{ct}. A least-squares solution for x_{ct} and y_{ct} is possible. Define the error ϵ_i of the ith node as $R_i\theta(L-z)-\Delta_i$, whose sum of its squares ϵ^2 over the N nodes of the finite element is

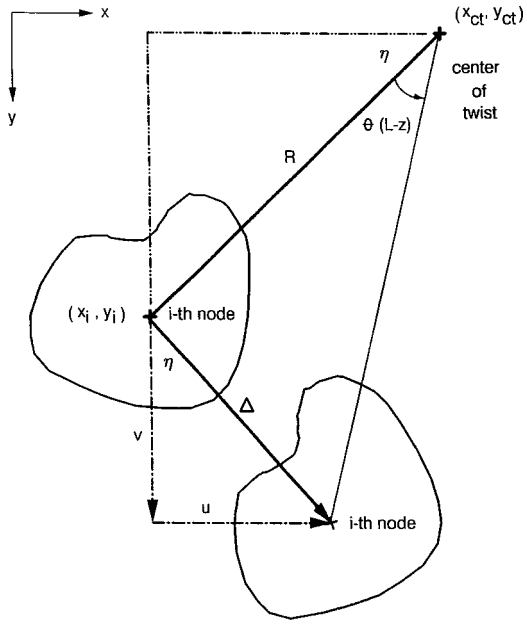


Fig. 2 Center of twist location

$$\epsilon^2=\sum_{i=1}^{N}[R_i\theta(L-z)-\Delta_i]^2. \quad (59)$$

Minimizing ϵ^2 with respect to x_{ct} and y_{ct} gives the following two nonlinear algebraic equations

$$x_{ct}=\frac{\overline{GJ}}{N(L-z)}\sum_{i=1}^{N}\frac{\Delta_i x_i}{R_i};\quad y_{ct}=\frac{\overline{GJ}}{N(L-z)}\sum_{i=1}^{N}\frac{\Delta_i y_i}{R_i} \quad (60)$$

where the nonlinearity emanates from x_{ct} and y_{ct} in R_i. An iterative solution can be obtained by *regula falsi*, i.e., by assuming x_{ct} and y_{ct} on the right-hand sides of Eq. (60) and checking their values by performing the indicated summations.

The warpage is the source of difficulty in defining the center of twist in an inhomogeneous, anisotropic cross section. When this point is not within the cross section, one cannot set u and v equal to zero and solve for x_{ct} and y_{ct} inasmuch as $\psi_{I6u}(x_{ct}, y_{ct})$ and $\psi_{I6v}(x_{ct}, y_{ct})$ are not defined. One way to remove this (x,y) dependency is to adopt some average values, for example, use U_6 and V_6 of Eq. (47).

$$u(x,y,z)=\frac{1}{\overline{GJ}}[-yz+U_6]-\omega_3 y-\omega_2 z+u_o$$

$$v(x,y,z)=\frac{1}{\overline{GJ}}[xz+V_6]+\omega_3 x-\omega_1 z+v_o \quad (61)$$

Setting $u=0$ and $v=0$ and solving for x and y give the following center of twist coordinates which may also act as initial estimates in the iterative solution of Eq. (59).

$$x_{ct}=-\frac{\overline{GJ}(v_o-\omega_1 z)+V_6}{z+\overline{GJ}\omega_3};\quad y_{ct}=\frac{\overline{GJ}(u_o-\omega_1 z)+U_6}{z+\overline{GJ}\omega_3} \quad (62)$$

Examples

Two laminated anisotropic cross sections are considered, viz., (1) an angle with legs of 1.0 and 1.5 units and (2) an uneven channel with sides of 1.0, 1.5 and 0.75 units as shown in Figs. 3 and 4, respectively. Both are fabricated from a two-ply sheet of

± 30 deg angle-ply with 0.1 unit total thickness and 0.05-unit-ply sheet thickness. The material properties are typical of a graphite-epoxy composite, i.e.,

$$\frac{E_L}{E_T}=10;\quad \frac{G_{LT}}{E_T}=0.4;\quad \frac{G_{TT}}{E_T}=0.3;\quad \nu_{LT}=0.3;\quad \nu_{TT}=0.25 \tag{63}$$

and the C_{ij} properties, in the ± 30 deg orientations with respect to the coordinate axes, have the values

$$\mathbf{C}=E_T\begin{bmatrix} 1.7033 & 0.3148 & 2.0804 & \cdot & \pm 1.0194 & \cdot \\ 0.3148 & 1.0831 & 0.3781 & \cdot & \pm 0.0549 & \cdot \\ 2.0804 & 0.3781 & 6.2848 & \cdot & \pm 2.9483 & \cdot \\ \cdot & \cdot & \cdot & 0.3750 & \cdot & \mp 0.0433 \\ \pm 1.0194 & \pm 0.0549 & \pm 2.9483 & \cdot & 2.0705 & \cdot \\ \cdot & \cdot & \cdot & \mp 0.0433 & \cdot & 0.3250 \end{bmatrix}. \tag{64}$$

These two cross sections are formed by folding two-layer ± 30 deg composite sheets into angle and channel shapes as shown in Figs. 3 and 4. Therefore, the above properties must be transformed accordingly to accommodate the fixed Cartesian coordinate system.

With the initial origin of the coordinate systems taken at the lower left corner of the cross section in both cases, their cross-sectional stiffness matrices $\boldsymbol{\kappa}_{Ibb}$

$$\boldsymbol{\kappa}_{Ibb}=\frac{E_T}{10^3}\begin{bmatrix} 871.14 & 421.98 & 202.58 & -19.17 \\ 421.98 & 399.34 & 29.07 & -9.12 \\ 202.58 & 29.07 & 116.37 & -4.30 \\ -19.17 & -9.12 & -4.30 & 1.20 \end{bmatrix}_{\text{angle}} \tag{65}$$

$$\boldsymbol{\kappa}_{Ibb}=\frac{E_T}{10^3}\begin{bmatrix} 1014.44 & 624.75 & 253.92 & -28.25 \\ 624.75 & 686.17 & 101.33 & -22.41 \\ 253.92 & 103.33 & 137.55 & -7.75 \\ -28.25 & -22.41 & -7.75 & 1.52 \end{bmatrix}_{\text{channel}}. \tag{66}$$

Translation to the centroids and rotations to principal bending axes lead to the following stiffness and flexibility matrices:

$$\boldsymbol{\kappa}''_{Ibb}=\frac{E_T}{10^3}\begin{bmatrix} 871.14 & \cdot & \cdot & -19.17 \\ \cdot & 225.47 & \cdot & 0.08 \\ \cdot & \cdot & 38.73 & 0.21 \\ -19.17 & 0.08 & 0.21 & 1.20 \end{bmatrix}_{\text{angle}} \tag{67}$$

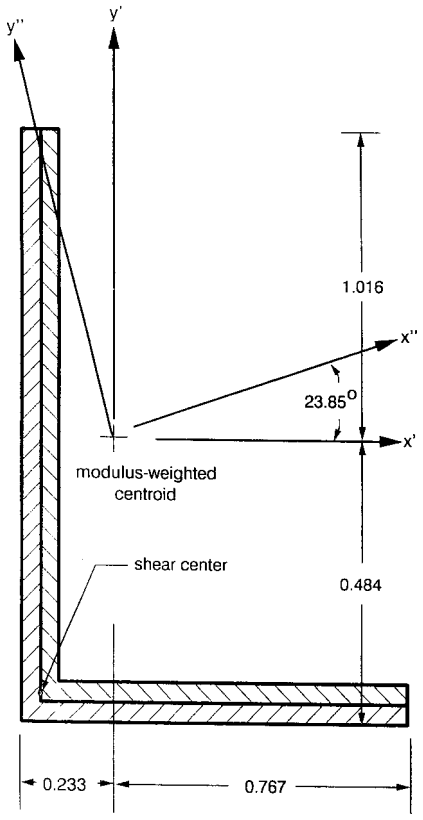


Fig. 3 Two-layer ±30 deg angle-ply angle cross section

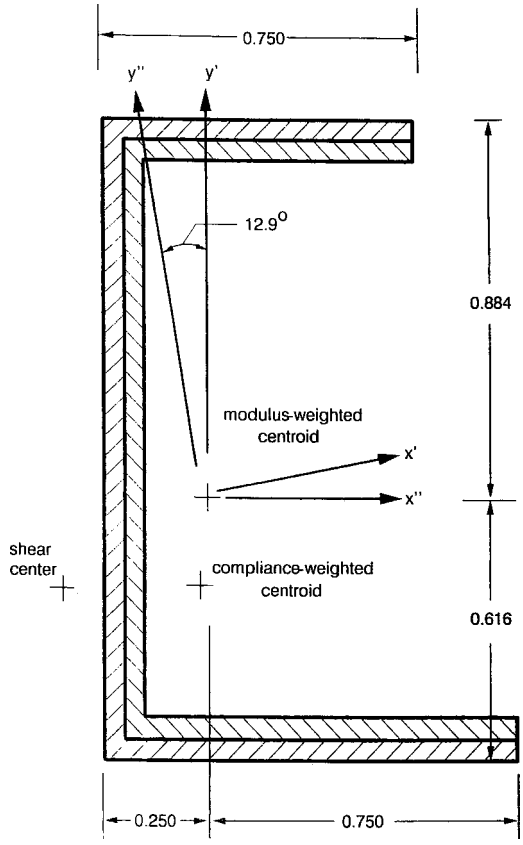


Fig. 4 Two-layer ±30 deg angle-ply channel cross section

Table 3 Summary of section properties

Cross Section →	Angle			Channel		
Centroids	$\bar{x}_t$	$\bar{y}_t$	θ_t	$\bar{x}_t$	$\bar{y}_t$	θ_t
Modulus-weighted	0.233	0.484	23.85 deg	0.250	0.616	12.92 deg
Compliance-weighted	0.237	0.489	23.86 deg	0.224	0.425	15.83 deg
Composite Torsional Rigidity $\overline{GJ}/E_T$	0.781×10^{-3}			0.607×10^{-3}		
Line of Extension-Bending Centers	$y'=1.906x'-8.657$			$y'=1.179x'+1.094$		
Shear Centers						
Griffith-Taylor	$e_x=-0.1891+0.0127z$ $e_y=-0.4069-0.01059z$			$e_x=-0.4169-0.00242z$ $e_y=-0.2740+0.01248z$		
Mean-valued	$e_x=-0.1855+0.00127z$ $e_y=-0.4064-0.01059z$			$e_x=-0.4232-0.00242z$ $e_y=-0.2776+0.01248z$		

$$\mathbf{s}''_{Ibb}=\frac{1}{E_T}\begin{bmatrix}1.77 & \cdot & \cdot & 28.13\\ \cdot & 4.44 & \cdot & -0.47\\ \cdot & \cdot & 25.86 & -6.98\\ 28.13 & -0.47 & -6.98 & 1280.75\end{bmatrix}_{\text{angle}} \tag{68}$$

$$\boldsymbol{\kappa}''_{Ibb}=\frac{E_T}{10^3}\begin{bmatrix}1014.44 & \cdot & \cdot & -28.25\\ \cdot & 314.03 & \cdot & -4.73\\ \cdot & \cdot & 61.37 & -1.78\\ -28.25 & -4.73 & -1.78 & 1.52\end{bmatrix}_{\text{channel}} \tag{69}$$

$$\mathbf{s}''_{Ibb}=\frac{1}{E_T}\begin{bmatrix}2.05 & \cdot & \cdot & 38.14\\ \cdot & 3.52 & \cdot & 22.34\\ \cdot & \cdot & 17.72 & 49.07\\ 38.14 & 22.34 & 49.07 & 1647.23\end{bmatrix}_{\text{channel}}. \tag{70}$$

The modulus and compliance-weighted centroids, composite torsional rigidity, lines of extension-bending centers, and shear center locations according to Griffith-Taylor and mean-valued definitions for both cross sections are tabulated in Table 3 and a number of these properties are shown in Figs. 3 and 4. The two centroidal values and principal bending axes for the angle are in close proximity and orientation with each other, but there are significant differences in these quantitities for the channel as seen in Fig. 4. The composite torsional rigidity $\overline{GJ}$ is based on Eq. (33), and that for the angle is approximately 65 percent of its κ_{I66} value. However, for the channel, there is an even greater reduction to 40 percent of its κ_{I66} value. For both of these cross sections, there is only a small difference in the shear center locations between the Griffith-Taylor and zero mean rotation values.

Concluding Remarks

Methods for determining cross-sectional properties for an inhomogeneous, anisotropic cross section based on Saint-Venant's solutions were presented. It was shown that for an inhomogeneous, anisotropic cylinder lacking a cross-sectional material symmetry plane, there are distinctions between the modulus-weighted and compliance-weighted centroids and their principal bending axes. The line of extension-bending centers was defined, which is based on the flexibility relation. Both Griffith-Taylor and zero-mean rotation definitions of the shear center were given, with the Griffith-Taylor definition ignoring warpages, while the zero-mean rotation shear center relied on integrals of the warpages over the cross section. Likewise, a center of twist was set forth, a cross-sectional property that is dependent upon the boundary conditions on the root end. These conditions were enforced on the basis of a number of mean-valued displacements found by integration of their pointwise values over the cross section, as seen in Eqs. (37), (43), and (49), and in the formula for J given by Eq. (48). While the definitions herein for both of these properties were rationally justified, other averaging procedures are possible so that issues regarding their *best definitions* remain open.

Appendix

The transformation matrices $\mathbf{T}_1$, $\mathbf{T}_2$, $\mathbf{T}_3$, and $\mathbf{T}_4$ for translation and rotation of coordinate axes are listed here.

$$\mathbf{T}_1=\begin{bmatrix}1 & y_t & x_t & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix};\quad \mathbf{T}_2=\begin{bmatrix}1 & -y_t & -x_t & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{A1}$$

$$\mathbf{T}_3=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & \cos\alpha & -\sin\alpha & 0\\ 0 & \sin\alpha & \cos\alpha & 0\\ 0 & 0 & 0 & 1\end{bmatrix};$$

$$\mathbf{T}_4=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & \cos\alpha & \sin\alpha & 0\\ 0 & -\sin\alpha & \cos\alpha & 0\\ 0 & 0 & 0 & 1\end{bmatrix}. \tag{A2}$$

These transformation matrices satisfy the equations

$$\mathbf{T}_1\mathbf{T}_2=\mathbf{T}_2\mathbf{T}_1=\mathbf{T}_2^T\mathbf{T}_1^T=\mathbf{T}_1^T\mathbf{T}_2^T=1 \tag{A3}$$

$$\mathbf{T}_3\mathbf{T}_4=\mathbf{T}_4\mathbf{T}_3=\mathbf{T}_3^T\mathbf{T}_4^T=\mathbf{T}_4^T\mathbf{T}_3^T=1. \tag{A4}$$

References

[1] Dong, S. B., Kosmatka, J. B., and Lin, H. C., 2001, "On Saint-Venant's Problem for an Inhomogeneous, Anisotropic Cylinder—Part I: Methodology for Saint-Venant Solutions," ASME J. Appl. Mech., **67**, pp. 376–381.

[2] Herrmann, L. R., 1965, "Elastic Torsional Analysis of Irregular Shapes," J. Eng. Mech., **91**, No. EM6, pp. 11–19.

[3] Mason, W. E., and Herrmann, L. R., 1968, ''Elastic Shear Analysis of General Prismatic Shaped Beams,'' J. Eng. Mech., **94**, No. EM4, pp. 965–983.
[4] Tolf, G., 1985, ''Saint-Venant Bending of an Orthotropic Beam,'' Comput. Struct., **4**, pp. 1–14.
[5] Wörndle, R., 1982, ''Calculation of the Cross-Section Properties and the Shear Stresses of Composite Rotor Blades,'' Vertica, **6**, pp. 111–129.
[6] Kosmatka, J. B., and Dong, S. B., 1991, ''Saint-Venant Solutions for Prismatic Anisotropic Beams,'' Int. J. Solids Struct., **28**, No. 7, pp. 917–938.
[7] Lekhnitskii, S. G., 1963, *Theory of Elasticity of an Anisotropic Body*, 1st Ed., Holden-Day, San Francisco; also 2nd Ed. (1977), Mir Publishers, Moscow.
[8] Ieşan, D., 1976, ''Saint-Venant's Problem for Inhomogeneous and Anisotropic Elastic Elastic Bodies,'' J. Elast., **24**, pp. 277–294.
[9] Ieşan, D., 1986, ''On Saint-Venant's Problem,'' Arch. Ration. Mech. Anal., **91**, pp. 363–373.
[10] Ieşan, D., 1987, *Saint-Venant's Problem* (Lectures Notes in Mathematics), Springer-Verlag, Heidelberg.
[11] Griffith, A. A., and Taylor, G. I., 1917, ''The Problem of Flexure and Its Solution by the Soap Film Method,'' Technical Report, National Advisory Committee on Aeronautics, Vol. 3, p. 950.
[12] Timosheko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd Ed., McGraw-Hill, New York.
[13] Novozhilov, V. V., 1961, *Theory of Elasticity*, Israel Program for Scientific Translation, Jerusalem, pp. 248–253 (translated from the Russian edition *Teoriya Uprugosti* (1958), by J. J. Shorr-Kon and edited by L. Meroz).
[14] Reissner, E., 1989, ''The Center of Shear as a Problem of the Theory of Plates of Variable Thickness,'' Ing. Archiv., **59**, pp. 324–332.
[15] Reissner, E., 1991, ''Approximate Determination of the Center of Shear by Use of the Saint-Venant Solution for the Flexure Problem of Plates of Variable Thickness,'' Ing. Archiv., **61**, pp. 555–566.
[16] Kosmatka, J. B., 1993, ''Flexure-Torsion Behavior of Prismatic Beams, Part I: Section Properties via Power Series,'' AIAA J., **31**, No. 1, pp. 170–179.
[17] Kosmatka, J. B., 1994, ''General Behavior and Shear Center Location of Prismatic Anisotropic Beams via Power Series,'' Int. J. Solids Struct., **31**, No. 3, pp. 417–439.
[18] Libove, C., 1988, ''Stresses and Rate of Twist in Single-Cell Thin-Walled Beams With Anisotropic Walls,'' AIAA J., **26**, No. 9, pp. 1107–1118.
[19] Weinstein, A., 1947, ''The Center of Shear and the Center of Twist,'' Q. Appl. Math., **5**, pp. 97–99.
[20] Cicala, P., 1935, ''II Centro di Taglio nei Solidi Cilindrici,'' Atti della Acad. Sci. Torino, **70**, pp. 356–371.
[21] Trefftz, E., 1935, ''Über den Schubmittelpunkt in einem durch eine Einzellast gebogenen Balken,'' Z. Angew. Math. Mech., **15**, pp. 220–225.

On Saint-Venant's Problem for an Inhomogeneous, Anisotropic Cylinder—Part III: End Effects

H. C. Lin

S. B. Dong
Mem. ASME

Civil and Environmental Engineering Department, University of California, Los Angeles, CA 90095-1593

J. B. Kosmatka
Department of Applied Mechanics and Engineering Science, University of California, San Diego, CA 92093-0085
Mem. ASME

End effects or displacements and stresses of a self-equilibrated state in an inhomogeneous, anisotropic cylinder are represented by eigendata extracted from an algebraic eigensystem. Such states are typical of traction and/or displacement boundary conditions that do not abide by the distributions according to Saint-Venant's solutions, whose construction were discussed in the first paper of this series of three. This type of analysis of end effects quantitifies Saint-Venant's principle, and the algebraic eigensystem providing the eigendata is based on homogeneous displacement equations of equilibrium with an exponential decaying displacement form. The real parts of the eigenvalues convey information on the inverse decay lengths and their corresponding eigenvectors are displacement distributions of self-equilibrated states. Stress eigenvetors can be formed by appropriate differentiation of the displacement eigenvectors. The eigensystem and its adjoint system provide complete sets of right and left-handed eigenvectors that are interrelated by two bi-orthogonality relations. Displacement and stress end effects can be represented by means of an expansion theorem based on these bi-orthogonality relations or by a least-squares solution. Two examples, a beam with a homogeneous, isotropic cross section and the other of a two layer beam with a ±30 deg angle-ply composite cross section, are given to illustrate the representation of various end effects. [DOI: 10.1115/1.1363597]

Introduction

In Dong, Kosmatka, and Lin [1], the first paper in this series of three and hereinafter called DKL [1], Saint-Venant solutions were constructed for a prismatic cylinder with an inhomogeneous, anisotropic cross section. This construction was made on the basis of the relaxed formulation of Saint-Venant's problem where the end conditions are prescribed in terms of integral resultants of axial force, bending moments, torque, and flexure force.

Saint-Venant solutions are rigorously valid everywhere except in the neighborhoods of the two ends of a cylinder. Saint-Venant [2,3] conjectured that his solutions differed with all other solutions possessing equipollent forces and moments only in these neighborhoods, i.e., in what has since been called Saint-Venant's Principle. For any solution with an end force or moment equal to that of a Saint-Venant solution, the difference between them is a self-equilibrated state whose effects are confined to near the ends of the cylinder, i.e., end effects. Our understanding of Saint-Venant's celebrated principle has been predominantly on a qualitative rather than a quantitative basis.

The first quantitative analyses of end effects were given by Johnson and Little [4] and Little and Childs [5] on homogeneous, isotropic plane-strain strips and circular cylinders, respectively. From their boundary value problems, characteristic equations were formed whose eigenvalues and eigenfunctions represented inverse decay lengths and associated distributions of self-equilibrated states. Moreover, these eigenfunctions are complete and may be used in an eigenfunction expansion for representing any self-equilibrated state. It is of interest to note that Synge [6] gave the essence of the complete analysis by posing the eigenproblem and indicating the exponential form of the decay. But, he gave no solutions. The quantification of Saint-Venant's principle was put on firm footing by Toupin [7] and Knowles [8] with their strain energy inequality theorems. These theorems bounded the attenuation rates to no greater than exponential decay. For an exceptionally comprehensive summary of the developments in Saint-Venant's principle, the reader is referred to Horgan and Knowles [9] and Horgan [10].

A considerable body of literature on finite element calculations of decay rates and their corresponding eigenvectors is available, viz., Rao and Valsarajan [11] on sandwich strip, Dong and Goetschel [12] on laminated, anisotropic strip, Giavotto et al. [13] on three-dimensional anisotropic cylinder, Huang and Dong [14] on laminated, anisotropic circular cylinder, Okumura et al. [15] on laminated strip, Goetschel and Hu [16] on general cross-section, and Kazic and Dong [17] on restrained torsion. Of these, only Kazic and Dong [17] considered the solution a three-dimensional solution of a particular end restraint condition by means of an eigenvector expansion of the end effects with the established eigenvectors. Besides these numerical studies, mention should be made of analytical representations of end effects based on analytical solutions of boundary value problems; see, for example, Gregory and Gladwell [18] on the clamped isotropic strip, Horgan [19] on composites, and Savoia and Tulluni [20] and Vel and Batra [21] on orthotropic strips. More information on such previous work can be found in Horgan's review state-of-the-art ([10]). Herein, Kazic and Dong's approach [17] is followed for appending end solutions to Saint-Venant solutions in order to satisfy any prescribed pointwise tractions at the tip end and restraint conditions at the root end.

In the next section, the exponential solution form is substituted into the governing equation of the previous paper to give an algebraic eigenproblem. The solutions to this equation and its adjoint in terms of their right and left-handed eigenvectors are stated. Bi-orthogonality properties on these two systems of eigenvectors are discussed. An expansion theorem that employs these bi-orthogonality conditions is given. Then, the analyses of end effects at both the tip and root ends are addressed. Examples are offered to illustrate the solution method.

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. 7, 1999; final revision, July 21, 2000. Associate Editor: J. W. Ju. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

The Algebraic Eigenproblem

For determining eigendata to represent end effects, the following discrete system of second-order ordinary differential equations in terms of nodal displacement $\mathbf{U}(\mathbf{z})$ is used.

$$\mathbf{K}_1\mathbf{U}_{,zz}+\mathbf{K}_2\mathbf{U}_{,z}-\mathbf{K}_3\mathbf{U}=\mathbf{0} \tag{1}$$

This equation was given in DKL [1] as Eq. (7). Recall that $\mathbf{K}_1$ and $\mathbf{K}_3$ are symmetric and $\mathbf{K}_2$ is antisymmetric. The solution to Eq. (1) for a self-equilibrated state has the form

$$\mathbf{U}(z)=\mathbf{U}_0e^{-\gamma z} \tag{2}$$

where $\mathbf{U}_0$ is a vector describing the nodal displacement distribution and γ is the inverse decay length. Substituting Eq. (2) into Eq. (1) and suppressing the multiplicative factor $e^{-\gamma z}$ lead to a second-order algebraic eigenvalue problem.

$$(\gamma^2\mathbf{K}_1-\gamma\mathbf{K}_2-\mathbf{K}_3)\mathbf{U}_0=0 \tag{3}$$

This equation may be reduced to first-order form by introducing $\mathbf{U}_1$ as

$$\mathbf{K}_3\mathbf{U}_1=\gamma\mathbf{K}_3\mathbf{U}_0\,. \tag{4}$$

Inserting Eq. (4) into Eq. (3) and combining the results yield

$$\begin{bmatrix}0 & \mathbf{K}_3\\ \mathbf{K}_3 & \mathbf{K}_2\end{bmatrix}\begin{bmatrix}\mathbf{U}_0\\ \mathbf{U}_1\end{bmatrix}=\gamma\begin{bmatrix}\mathbf{K}_3 & 0\\ 0 & \mathbf{K}_1\end{bmatrix}\begin{bmatrix}\mathbf{U}_0\\ \mathbf{U}_1\end{bmatrix}\quad\text{or}\quad\mathbf{AQ}_r=\gamma\mathbf{BQ}_r \tag{5}$$

where $\mathbf{Q}_r$ is the right-handed generalized coordinate state vector.

If M denotes the dimension of $\mathbf{U}_0$, then $2M$ eigenvalues γ's are contained in algebraic eigensystem (5). They can be real, complex conjugate pairs as well as have zero value. Nonzero roots represent attenuation rates of self-equilibrated effects into the interior of a cylinder with real γ's expressing monotonic decay and complex γ's sinusoidal decay. The real roots occur in positive and negative pairs, i.e., $\pm\gamma_j$, and the complex conjugate roots are foursomes with positive and negative real parts, i.e., $\pm\gamma_{jR}$ $\pm i\gamma_{jI}$. Positive real γ_j's and complex roots with positive real parts γ_{jR}'s represent decay into the region $z\geqslant 0$, which in our coordinate system (see [1], Fig. 1) applies to tip end conditions. The other subset of eigendata with negative real $-\gamma_j$'s and negative real parts $-\gamma_{jR}$'s of complex γ_j's are for root end conditions. The eigenvalue with the smallest magnitude real part is important, as it defines the inverse decay length with the furthest penetration into the interior. Zero roots represent rigid-body translations and rotations and the extension, bending, and torsion fields of Saint-Venant's solutions. Certain computer algorithms require $\mathbf{B}$ of algebraic eigensystem (5) to be nonsingular, so that the rigid-body modes in $\mathbf{K}_3$ must be suppressed prior to the eigensolution process.

The solution to Eq. (5) may be stated as a transformation to a right-handed system of normal coordinates $\mathbf{X}$, i.e.,

$$\mathbf{Q}_r=\mathbf{\Phi X} \tag{6}$$

where $\mathbf{\Phi}$ is the right modal matrix. In view of Eq. (4), $\mathbf{\Phi}$ may be partitioned into upper and lower parts.

$$\mathbf{Q}_r=\begin{bmatrix}\mathbf{U}_0\\ \mathbf{U}_1\end{bmatrix}=\begin{bmatrix}\mathbf{\Phi}_u\\ \mathbf{\Phi}_l\end{bmatrix}\mathbf{X}=\begin{bmatrix}\mathbf{\Phi}_u\\ \mathbf{\Phi}_u\mathbf{\Gamma}\end{bmatrix}\mathbf{X} \tag{7}$$

where $\mathbf{\Gamma}$ is a diagonal matrix of the eigenvalues.

Adjoint Eigenproblem, Bi-Orthogonality, and Eigenvector Expansion

The adjoint problem to Eq. (5) in terms of left-handed generalized coordinates $\mathbf{V}_0$ and $\mathbf{V}_1$ is

$$\begin{bmatrix}\mathbf{0} & \mathbf{K}_3\\ \mathbf{K}_3 & -\mathbf{K}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_0\\ \mathbf{V}_1\end{bmatrix}=\gamma\begin{bmatrix}\mathbf{K}_3 & 0\\ 0 & \mathbf{K}_1\end{bmatrix}\begin{bmatrix}\mathbf{V}_0\\ \mathbf{V}_1\end{bmatrix}\quad\text{or}\quad\mathbf{A}^T\mathbf{Q}_l=\gamma\mathbf{BQ}_l\,. \tag{8}$$

Solution of Eq. (8) consists of a transformation to left-handed normal coordinates $\mathbf{Y}$

$$\mathbf{Q}_l=\mathbf{\Psi Y}. \tag{9}$$

The eigenvalues of Eq. (8) are the same as those of Eq. (5) with $\mathbf{\Psi}$ as the left modal matrix. The left modal matrix $\mathbf{\Psi}$ may be partitioned into upper and lower parts.

$$\mathbf{Q}_l=\begin{bmatrix}\mathbf{V}_0\\ \mathbf{V}_1\end{bmatrix}=\begin{bmatrix}\mathbf{\Psi}_u\\ \mathbf{\Psi}_l\end{bmatrix}\mathbf{Y}=\begin{bmatrix}\mathbf{\Psi}_u\\ \mathbf{\Psi}_u\mathbf{\Gamma}\end{bmatrix}\mathbf{Y} \tag{10}$$

The right and left-handed eigenvectors satisfy the bi-orthogonality relations

$$\mathbf{\Psi}^T\mathbf{B\Phi}=\mathrm{diag}(B_n);\quad\mathbf{\Psi}^T\mathbf{A\Phi}=\mathrm{diag}(\gamma_nB_n). \tag{11}$$

In view of the upper and lower half forms given by Eqs. (7) and (10), bi-orthogonality relations (11) expressed entirely in terms of their upper half forms appear as

$$\begin{aligned}&\mathbf{\Psi}_u^T\mathbf{K}_3\mathbf{\Phi}_u+\gamma_m\gamma_n\mathbf{\Psi}_u^T\mathbf{K}_1\mathbf{\Phi}_u=\delta_{mn}B_n\\ &(\gamma_m+\gamma_n)\mathbf{\Psi}_u^T\mathbf{K}_3\mathbf{\Phi}_u+\gamma_m\gamma_n\mathbf{\Psi}_u^T\mathbf{K}_2\mathbf{\Phi}_u=\delta_{mn}\gamma_nB_n\,.\end{aligned} \tag{12}$$

With these bi-orthogonality relations, a dual expansion theorem can be formulated. Let $\mathbf{F}$ be an arbitrary displacement vector, which expressed in terms of the right-handed eigenvectors, takes the form

$$\mathbf{F}=\mathbf{\Phi}_u\boldsymbol{\alpha} \tag{13}$$

with $\boldsymbol{\alpha}$ representing the amplitudes. These amplitudes are then given with the aid of Eq. (12) as

$$\boldsymbol{\alpha}=\mathrm{diag}(1/B_n)(\mathbf{\Psi}_u^T\mathbf{K}_3+\gamma_n^2\mathbf{\Psi}_u^T\mathbf{K}_1)\mathbf{F}. \tag{14}$$

Analysis of End Effects

Let $\mathbf{U}(z)$ and $\boldsymbol{\sigma}(z)$ denote the total displacement and stress fields of a given problem that satisfy the pointwise prescribed tractions and restraints at the two ends. In our discussion, it should be understood that $\boldsymbol{\sigma}(z)$ and any other stress vector to be introduced denote an array by concatenation of the six components at each gaussian integration point taken over all such points of an element and over all elements comprising the discretized model. The total displacement and stress fields, $\mathbf{U}(z)$ and $\boldsymbol{\sigma}(z)$, can be written as

$$\mathbf{U}(z)=\mathbf{U}_{SV}(z)+\mathbf{U}_{\mathrm{end}}(z);\quad\boldsymbol{\sigma}(z)=\boldsymbol{\sigma}_{SV}(z)+\boldsymbol{\sigma}_{\mathrm{end}}(z) \tag{15}$$

where $\mathbf{U}_{SV}(z)$ and $\boldsymbol{\sigma}_{SV}(z)$ denote the displacement and stress fields of the Saint-Venant solution having equipollent resultant forces and moments as those of $\mathbf{U}(z)$ and $\boldsymbol{\sigma}(z)$ and $\mathbf{U}_{\mathrm{end}}(z)$ and $\boldsymbol{\sigma}_{\mathrm{end}}(z)$ are self-equilibrated displacement and stress fields of the end effects. The Saint-Venant stress field is given by

$$\boldsymbol{\sigma}_{SV}=\mathbf{C}(\mathbf{b}_1\mathbf{u}_{SV}+\mathbf{b}_2\mathbf{u}_{SV,z}) \tag{16}$$

where $\mathbf{u}_{SV}$ is the element level form of the displacement vector and $\mathbf{b}_1$ and $\mathbf{b}_2$ are strain-transformation matrices given by Eq. (5) of DKL [1]. From Eq. (15), the end effects are seen as the difference between the Saint-Venant solution and that for the pointwise prescription of the end conditions. Moreover, these end effects may be separated into two parts, one for decay from the tip end and the other for decay from the root end. Expressing these end effects in terms of a modal expansion of their eigenvectors enables Eq. (15) to be recast as

$$\mathbf{U}(z)=\mathbf{U}_{SV}(z)+[\mathbf{\Phi}_1\mathbf{E}_1[\gamma_jz]\quad\mathbf{\Phi}_2\mathbf{E}_2[\gamma_j(L-z)]]\begin{Bmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\end{Bmatrix} \tag{17}$$

where $\mathbf{\Phi}_1$, $\mathbf{\Phi}_2$ and $\boldsymbol{\sigma}_1$, $\boldsymbol{\sigma}_2$ contain the modal displacement and stress columns of the self-equilibrated states, $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ are unknown amplitudes, and $\mathbf{E}_1[\gamma_jz]$ and $\mathbf{E}_2[\gamma_j(L-z)]$ are diagonal matrices containing the inverse decay length data, i.e.,

$$\mathbf{E}_1[\gamma_jz]=\mathrm{diag}(e^{-\gamma_jz});\quad\mathbf{E}_2[\gamma_jz]=\mathrm{diag}(e^{-\gamma_j(L-z)}). \tag{19}$$

According to Saint-Venant's principle, the end effects are reasonably confined to the neighborhood at each end. First, consider the case of a sufficiently long cylinder where it is possible to consider the effects on one end completely uncoupled from the other end. Then, the two ends are treated as follows.

Displacement Boundary Conditions at Root End $z=L$. Representations (17) and (18) appropriate for the root end have the forms

$$\mathbf{U}(z)=\mathbf{U}_{SV}(z)+\mathbf{\Phi}_2\mathbf{E}_2[\gamma_j(L-z)]\boldsymbol{\alpha}_2;$$
$$\boldsymbol{\sigma}(z)=\boldsymbol{\sigma}_{SV}(z)+\boldsymbol{\sigma}_2\mathbf{E}_2[\gamma_j(L-z)]\boldsymbol{\alpha}_2. \tag{20}$$

Let $\mathbf{U}_L$ be the nodal values of the prescribed nodal restraints at the root end, i.e., $\mathbf{U}_L=U(L)$. Setting $z=L$ in Eq. (20) and noting that $\mathbf{E}_2=\mathbf{I}$ for $\mathbf{z}=L$ give

$$\mathbf{U}_L=\mathbf{U}_{SV}(L)+\mathbf{\Phi}_2\boldsymbol{\alpha}_2 \quad \text{or} \quad \mathbf{\Phi}_2\boldsymbol{\alpha}_2=\mathbf{U}_L-\mathbf{U}_{SV}(L). \tag{21}$$

The solution for the amplitudes $\boldsymbol{\alpha}_2$ can be obtained by means of Eq. (14).

$$\boldsymbol{\alpha}_2=\text{diag}(1/B_n)(\mathbf{\Psi}_2^T\mathbf{K}_3+\gamma_n^2\mathbf{\Psi}_2^T\mathbf{K}_1)[\mathbf{U}_L-\mathbf{U}_{SV}(L)] \tag{22}$$

It may also be obtained by a least-squares solution by premultiplying both sides of Eq. (21) by $\mathbf{\Phi}_2^H$, where superscript H denotes conjugate transpose, to square up the coefficient matrix of $\boldsymbol{\alpha}_2$ and then inverting.

$$\boldsymbol{\alpha}_2=[\mathbf{\Phi}_2^H\mathbf{\Phi}_2]^{-1}\mathbf{\Phi}_2^H[\mathbf{U}_L-\mathbf{U}_{SV}(L)] \tag{23}$$

A least-squares solution was used by Kazic and Dong [17].

Traction Conditions at Tip End $z=0$. The displacement and stress representations for the tip end have the form

$$\mathbf{U}(z)=\mathbf{U}_{SV}(z)+\mathbf{\Phi}_1\mathbf{E}_1[\gamma_j z]\boldsymbol{\alpha}_1; \quad \boldsymbol{\sigma}(z)=\boldsymbol{\sigma}_{SV}(z)+\boldsymbol{\sigma}_1\mathbf{E}_1[\gamma_j z]\boldsymbol{\alpha}_1. \tag{24}$$

At the tip end, only stress components $(\sigma_{xz},\sigma_{yz},\sigma_{zz})$ are involved in the prescribed traction. Let $\mathbf{S}$ denote these prescribed stress components in $\boldsymbol{\sigma}$ of solution vector (24) evaluated at $z=0$ for the N Gaussian points of the total discretized model of the cross section.

$$\mathbf{S}^T=[\sigma_{xz1},\sigma_{yz1},\sigma_{zz1},\sigma_{xz2},\sigma_{yz2},\sigma_{zz2},\cdots,\sigma_{xzN},\sigma_{yzN},\sigma_{zzN}]|_{z=0} \tag{25}$$

Extracting the same components from $\boldsymbol{\sigma}_{SV}$ and $\boldsymbol{\sigma}_1$, denoting them as $\mathbf{S}_{SV}$ and $\mathbf{S}_1$, and invoking $\mathbf{E}_1(0)=\mathbf{I}$ allow the traction boundary conditions at $z=0$ to be written as

$$\mathbf{S}=\mathbf{S}_{SV}(0)-\mathbf{S}_1\boldsymbol{\alpha}_1 \quad \text{or} \quad \mathbf{S}_1\boldsymbol{\alpha}_1=\mathbf{S}_{SV}(0)-\mathbf{S}. \tag{26}$$

Note that $\mathbf{S}_1$ is rectangular whose row and column dimensions are N and the number of eigenvectors from $\mathbf{\Phi}_1$ used in the representation. The accuracy of this representation is dependent on the number of modes used. It is usually possible to truncate a number of the higher modes from $\mathbf{\Phi}_1$ without affecting the solution accuracy of the dominant coefficients in $\boldsymbol{\alpha}_1$. However, the expression of prescribed pointwise tractions at the Gaussian integration points demands a significantly greater number of modes than that to merely define the coefficients $\boldsymbol{\alpha}_1$. The least squares solution of Eq. (26) has the form

$$\boldsymbol{\alpha}_1=[\mathbf{S}_1^H\mathbf{S}_1]^{-1}\mathbf{S}_1^H[S_{SV}(0)-\mathbf{S}]. \tag{27}$$

For mixed boundary conditions, it is possible to devise a system of equations, where displacement conditions are enforced at the nodal points and the stresses at the Gaussian integration points within the elements. Therefore, any combination of end displacement and traction state is amenable for representation.

In the case of a short cylinder and/or a cylinder in which the fundamental inverse decay length is such that it is not possible to uncouple the effects on the two ends, then the coefficients $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ must be solved simultaneously using Eqs. (17) and (19). Such cases occur in highly anisotropic materials (see Choi and Horgan [22,23], for example) and thin-walled members. Enforcing appropriate conditions on both ends provides the equations for solution of the $\boldsymbol{\alpha}_j$. For a short cylinder, it should be recognized that a three-dimensional finite element analysis of the cylinder may be numerically more expedient. The only drawback to a three-dimensional analysis is that it does not yield any information on the inverse decay lengths.

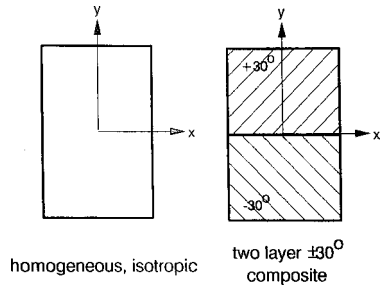


Fig. 1 Cross sections of two beams

Examples

Consider the two rectangular cross sections as shown in Fig. 1: (1) a homogeneous, isotropic beam and (2) a two-layer ± 30 deg angle-ply composite beam, both of unit height H and the same height-to-width ratio H/W of 1.5. Let E and $\nu=0.25$ denote Young's modulus and Poisson's ratio of the isotropic material. The material properties of the composite material are

$$\frac{E_L}{E_T}=10; \quad \frac{G_{LT}}{E_T}=0.4; \quad \frac{G_{TT}}{E_T}=0.3; \quad \nu_{LT}=0.3; \quad \nu_{TT}=0.25 \tag{28}$$

and the C_{ij} properties, transformed to the ± 30 deg orientations with respect to the coordinate axes, have the values

$$\mathbf{C}=E_T\begin{bmatrix} 1.7033 & 0.3148 & 2.0804 & \cdot & \pm 1.0194 & \cdot \\ 0.3148 & 1.0831 & 0.3781 & \cdot & \pm 0.0549 & \cdot \\ 2.0804 & 0.3781 & 6.2848 & \cdot & \pm 2.9483 & \cdot \\ \cdot & \cdot & \cdot & 0.3750 & \cdot & \mp 0.0433 \\ \pm 1.0194 & \pm 0.0549 & \pm 2.9483 & \cdot & 2.0705 & \cdot \\ \cdot & \cdot & \cdot & \mp 0.0433 & \cdot & 0.3250 \end{bmatrix}. \tag{29}$$

Table 1 Subset of eigenvalues for isotropic beam for decay from tip end

Extension		Bending/y-axis	
Mode	γ	Mode	γ
1,2	$4.2090 \pm 2.2752i$	1	4.0816
3,4	$6.3453 \pm 3.3558i$	2,3	$6.2120 \pm 1.9922i$
5	7.7293	4,5	$11.2386 \pm 4.1136i$
6,7	$8.8186 \pm 2.6749i$	6	11.2815
8,9	$10.7673 \pm 3.0861i$	7,8	$11.5921 \pm 3.0396i$
10	12.5522	9,10	$12.6693 \pm 3.5974i$
Bending/x-axis		**Torsion**	
Mode	γ	Mode	γ
1,2	$6.7078 \pm 3.0853i$	1,2	$4.1194 \pm 1.1796i$
3,4	$7.6709 \pm 3.3558i$	3	8.2404
5	8.0149	4,5	$8.6649 \pm 2.5180i$
6	10.3409	6,7	$11.5973 \pm 3.9789i$
7,8	$11.1915 \pm 2.9693i$	8	12.7278
9,10	$13.9402 \pm 3.3478i$	9,10	$14.2279 \pm 3.6504i$

Table 2 Subset of eigenvalues for two-layer beam for decay from tip end

Mode	γ	Mode	γ
1	1.8492	11,12	$2.9624 \pm 1.1922i$
2,3	$2.1092 \pm 0.7653i$	13,14	$3.1151 \pm 2.8740i$
4	2.2599	15	3.2565
5,6	$2.5288 \pm 0.9451i$	16,17	$3.7464 \pm 4.1427i$
7,8	$2.6339 \pm 2.7230i$	18,19	$3.8590 \pm 1.9040i$
9,10	$2.9581 \pm 3.0187i$	20,21	$4.0367 \pm 1.4298i$

Completely fixed conditions at the root end were assumed for both cases, i.e., all displacement components are zero. Loading at the tip end consisted of extension, bending, and torsion, and their assumed distributions are shown in Fig. 5.

For the isotropic beam, one-quarter of the cross section was modeled (64 elements/225 nodes/675 degrees-of-freedom) together with combinations of symmetry/antisymmetry interface conditions about the two planes of structural symmetry. In the two-layer cross section, the entire cross section was modeled (128 elements/433 nodes/1299 degrees-of-freedom).

The subsets of the lowest eigenvalues for these two beams are given in Tables 1 and 2. These eigenvalues define the inverse decay lengths from the tip end. There are an equal number of eigenvalues of the same magnitude but with negative real parts applicable for decay from the root end. Since the dimension H for both cross sections was unity, the eigenvalues are normalized with respect to the depth, i.e., $\gamma = \overline{\gamma}/H$ where $\overline{\gamma}$ is that for a cross section with a specific height and a H/W ratio of 1.5. Comparing the real parts of the eigenvalues between Tables 1 and 2 shows that those for the two-layer angle-ply composite cross section are smaller, indicating longer inverse decay lengths or greater distances needed to attenuate end effects, an observation of the effect of anisotropy that has been reported by other researchers.

For a fully restraint condition at the root end, all displacement components must vanish. In Fig. 2 is shown a measure of the deviation from zero of this condition as a function of the number of modes used in the modal representation. The cumulative *Error* is defined as the sum over all nodes of the finite element model of the normalized displacement, i.e.,

$$\text{Error} = \sum^{\text{nodes}} \sqrt{\frac{(u_{\text{end}}+u_{SV})^2+(v_{\text{end}}+v_{SV})^2+(w_{\text{end}}+w_{SV})^2}{U_{\max}^2}} \tag{32}$$

where $(u_{\text{end}}, v_{\text{end}}, w_{\text{end}})$ are components of the end solution, and $U_{\max}$ is the largest total displacement of the Saint-Venant solution for a particular load condition. In the case of the two-layer beam, the designations of extension, bending, and torsion modes are meant as the predominant behavior inasmuch as there is coupling of all of these behaviors in this cross-section. It is seen that with an increasing number of modes, the cumulative error is reduced accordingly and their magnitudes are reasonably small.

Bar graphs of dimensionless amplitudes of $\boldsymbol{\alpha}_1$, or amplitude ratios $\alpha_i/\alpha_{\max}$ where $\alpha_{\max}$ is the largest amplitude, are shown in Figs. 3 and 4 to show the participation of each mode in the modal summation. The occurrence of a pair of identical bars indicates complex conjugate roots, and a single bar is for a real eigenvalue. For the isotropic beam in Fig. 3, recall that one-fourth of the cross section was modeled with symmetry/antisymmetry interface conditions along the two planes of structural symmetry so that the amplitudes for each case are associated with its own set of eigenvectors. The amplitudes in Fig. 4 for the two-layer beam for all three loading cases are associated with the same set of eigendata.

For loading conditions at the tip end, an axial force P_z, a bending moment M_x about the x-axis and a torque M_z were considered. Unit values of these loads were applied in the form of uni-

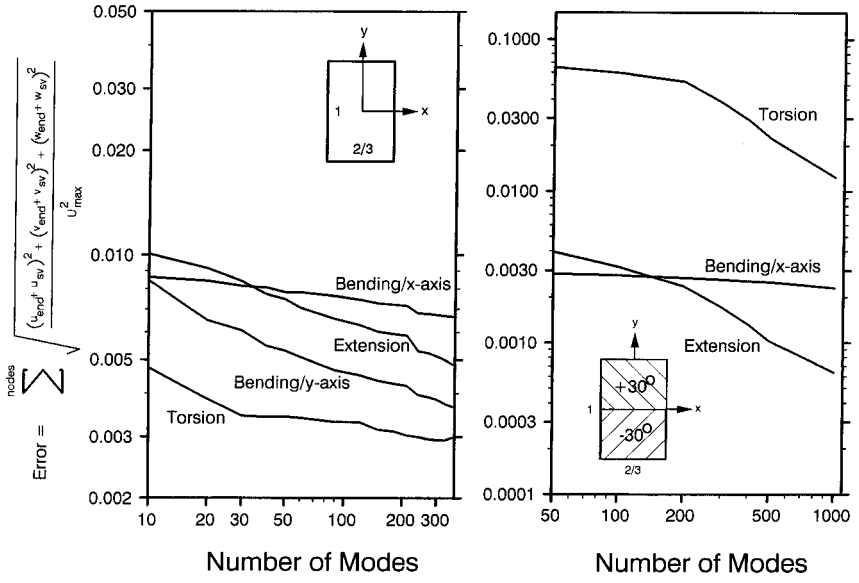


Fig. 2 Errors in representation of fully restraint displacement conditions

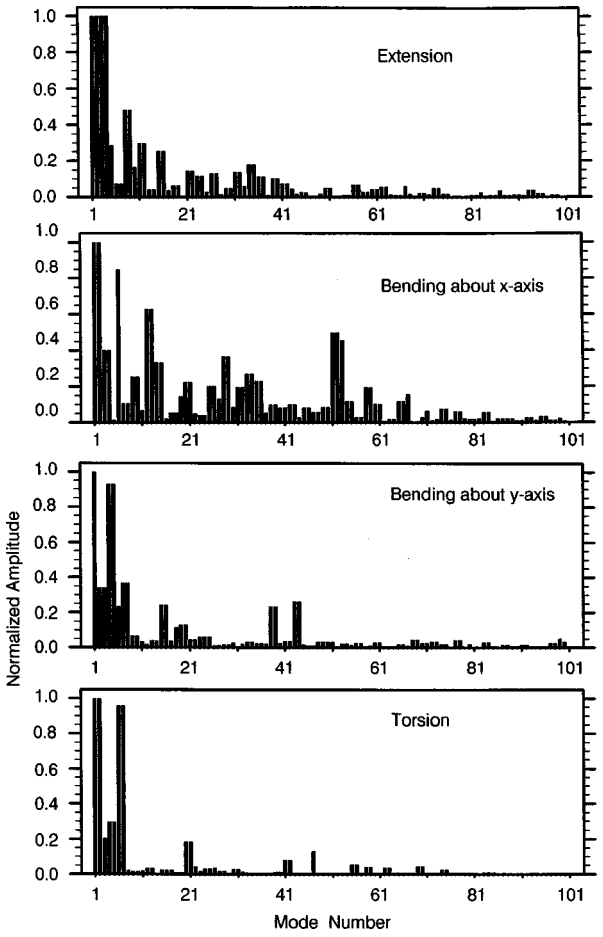


Fig. 3 Normalized amplitudes for homogeneous, isotropic beam

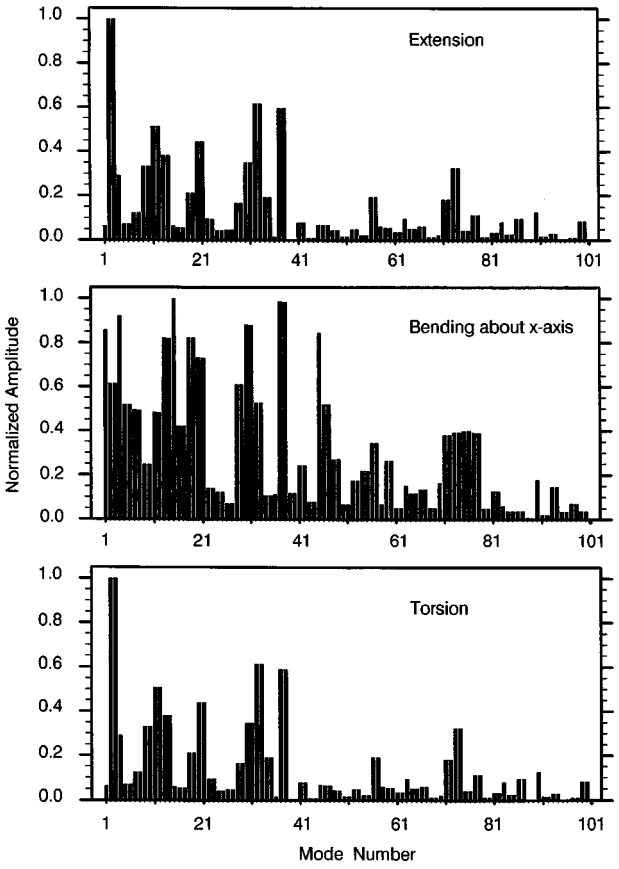


Fig. 4 Normalized amplitudes for two-layer ±30 deg composite beam

form stress blocks over various portions of the cross section as shown in Fig. 5. Also, shown on this figure are the self-equilibrated stress stated for extension and bending. That for torsion is not given because of the difficulty in presenting the Saint-Venant torsional stress distribution clearly, however, it was nevertheless a self-equilibrated stress state. While both cross sections were analyzed, only results for the two-layer composite cross section are shown herein. Results for the isotropic cross section describe the same sort of behavior as the two-layer beam and do not enhance the discussion, so they are not included here.

A plot of σ_{zz} for extension along the y-axis at $x=0$ is shown in Fig. 6, and that for bending in Fig. 7. Two sets of shear stresses, σ_{yz} and σ_{xz}, on either sides of the lines $y=0$ and $x=0$ along the x and y-axes are shown in Figs. 8 and 9. With an increasing the number of modes, the representations, converge onto their respective prescribed loading conditions. It is remarked that since stresses are obtained from differentiation of the displacement, the stress eigenvectors are inherently less accurate than the displacement eigenvectors. Therefore, more modes are needed for a comparable accuracy in the representations of stresses than of the displacements. No plots of the amplitudes of the various modes participating in the representations are given, but they would appear very much like those in Fig. 4 for representation of the displacement end conditions where the fundamental modes show a greater presence in comparison with the higher modes.

Concluding Remarks

This paper was devoted to the representation of end effects or the displacements and stresses of any arbitrary self-equilibrated state. Such states are typical at the tip and root ends, when the boundary conditions on the applied tractions and/or prescribed displacements do not follow the distributions of the Saint-Venant solutions. As both fields are associated with the same resultant forces and/or moments, the difference between them must be a self-equilibrated state.

The end effects are represented by eigendata extracted from an algebraic eigensystem. This eigenproblem is formed by using an exponential decaying displacement form in the homogeneous dis-

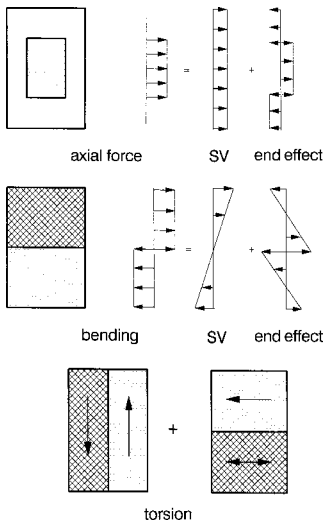


Fig. 5 Prescribed traction conditions at tip end

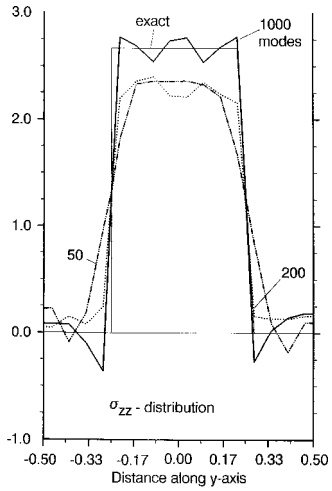


Fig. 6 Stress σ_{zz} for $P_z=1$ in two-layer ±30 deg composite beam

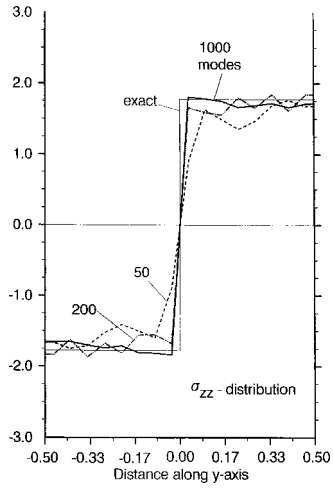


Fig. 7 Stress σ_{zz} for $M_x=1$ in two-layer ±30 deg composite beam

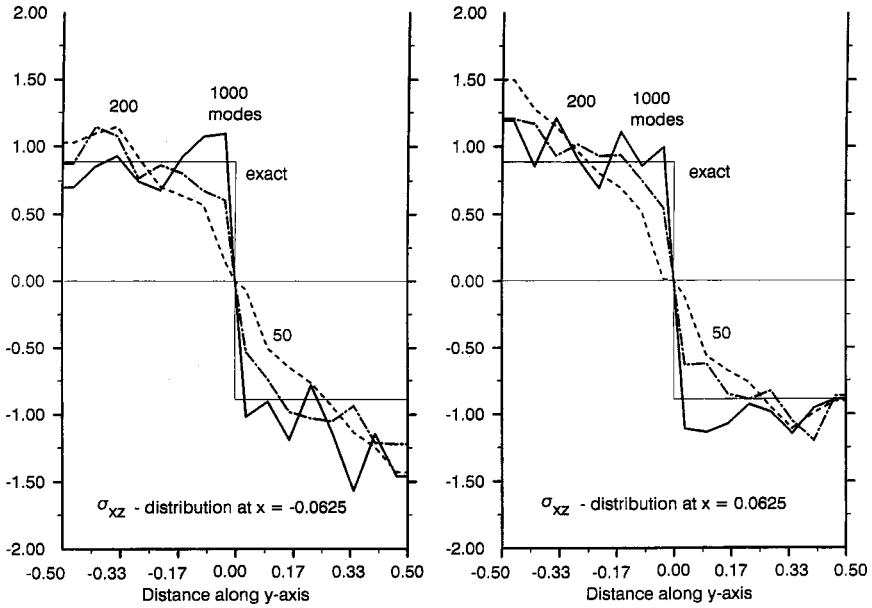


Fig. 8 Stress σ_{xz} for $M_z=1$ in two-layer ±30 deg composite beam

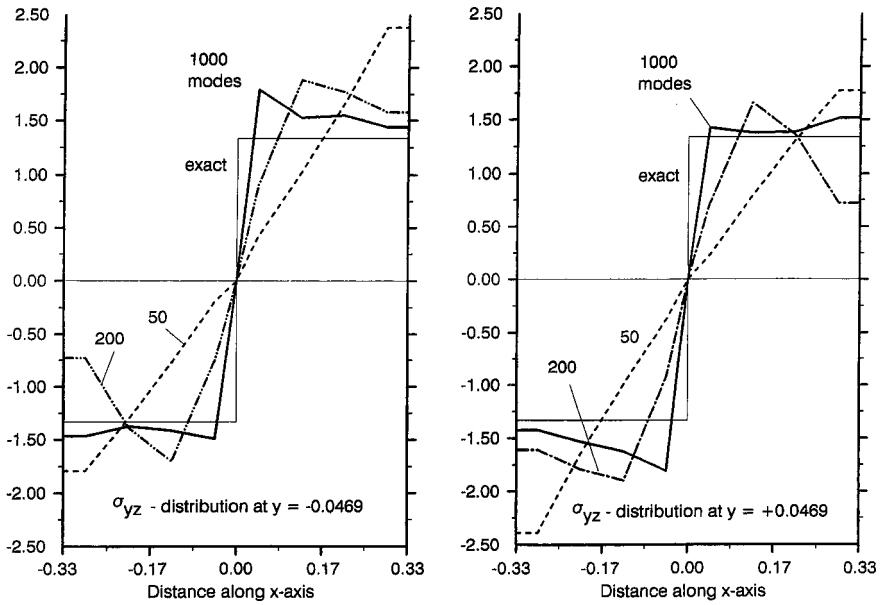


Fig. 9 Stress σ_{yz} for $M_z=1$ in two-layer ±30 deg composite beam

placement equations of equilibrium. The real parts of the eigenvalues convey information on the inverse decay lengths and their corresponding eigenvectors are displacement distributions of the self-equilibrated states. Stress eigenvectors can be formed by differentiation of the displacement eigenvectors according to the strain-displacement equations. The algebraic eigensystem and its adjoint system provide complete sets of right and left-handed eigenvectors which are interconnected by bi-orthogonality relations. An dual expansion theorem can be stated with the bi-orthogonality relations. The representation of displacement end effects can be made by means of this expansion theorem, or in the case of end tractions by means of a least-squares solution.

Two example beams, one with a homogeneous, isotropic cross section and the other with a two-layer ± 30 deg angle-ply composite cross section, were given to illustrate the representation of various end effects.

References

[1] Dong, S. B., Kosmatka, J. B., and Lin, H. C., 2001, ''On Saint-Venant's Problem for an Inhomogeneous, Anisotropic Cylinder, Part I: Methodology for Saint-Venant Solutions,'' ASME J. Appl. Mech., **68**, 376—381.

[2] de Saint-Venant, A. J. C. B., 1856, ''Memoire sur la Torsion des Prismes,'' Mem. Savants Etrangers, **14**, pp. 233–560.

[3] de Saint-Venant, A. J. C. B., 1856, ''Memoire sur la Flexion des Prismes,'' J. Math. Liouville, Ser. II, **1**, pp. 89–189.

[4] Johnson, M. W., and Little, R. W., 1965, ''The Semi-Infinite Strip,'' Quarterly Appl. Math., **22**, pp. 335–344.

[5] Little, R. W., and Childs, S. B., 1967, ''Elastostatic Boundary Region Problem in Solid Cylinders,'' Q. Appl. Math., **25**, pp. 261–274.

[6] Synge, J. L., 1945, ''The Problem of Saint Venant for a Cylinder With Free Sides,'' Q. Appl. Math., **2**, pp. 307–317.

[7] Toupin, R. A., 1965, ''Saint-Venant's Principle,'' Arch. Ration. Mech. Anal., **18**, pp. 83–96.

[8] Knowles, J. K., 1966, ''On Saint-Venant's Principle in the Two-Dimensional Linear Theory of Elasticity,'' Arch. Ration. Mech. Anal., **21**, pp. 1–22.

[9] Horgan, C. O., and Knowles, J. K., 1983, ''Recent Developments Concerning Saint-Venant Principle,'' Adv. Appl. Mech., **23**, pp. 179–269.

[10] Horgan, C. O., 1989, ''Recent Developments Concerning Saint-Venant Principle: An Update,'' Appl. Mech. Rev., **42**, No. 11, pp. 295–303.

[11] Rao, N. R., and Valsarajan, K. V., 1980, ''Saint-Venant's Principle in Sandwich Strip,'' Comput. Struct., **12**, pp. 185–188.

[12] Dong, S. B., and Goetschel, D. B., 1982, ''Finite Element Analysis of Edge Effects in Laminated Composite Plates,'' ASME J. Appl. Mech., **49**, pp. 129–135.

[13] Glavotto, V., Borri, M., Mantegazza, P., Ghiringhelli, G., Carmaschi, V., Maffioli, G. C., and Mussi, F., 1983, ''Anisotropic Beam Theory and Applications,'' Comput. Struct., **16**, pp. 403–413.

[14] Huang, K. H., and Dong, S. B., 1984, ''Propagating Waves and Standing Vibrations in a Composite Cylinder,'' J. Sound Vib., **96**, No. 3, pp. 363–379.

[15] Okumura, H., Yokouchi, Y., Watanabe, K., and Yamada, Y., 1985, ''Local Stress Analysis of Composite Materials Using Finite Element Methods: 1st Report, Saint-Venant End Effects in Laminate Media,'' Trans. Jpn. Soc. Mech. Eng., **51**, pp. 563–570.

[16] Goetschel, D. B., and Hu, T. H., 1985, ''Quantification of Saint-Venant's Principle for a General Prismatic Member,'' Comput. Struct., **21**, pp. 869–874.

[17] Kazic, M., and Dong, S. B., 1990, ''Analysis of Restrained Torsion,'' J. Eng. Mechanics Division **116**, No. 4, pp. 870–891.

[18] Gregory, R. D., and Gladwell, I., 1982, ''The Cantilever Beam Under Tension, Bending or Flexure at Infinity,'' J. Elast., **4**, pp. 317–343.

[19] Horgan, C. O., 1982, ''Saint-Venant End Effects in Composites,'' J. Compos. Mater., **16**, pp. 411–422.

[20] Savoia, M., and Tullini, N., 1996, ''Beam Theory for Strongly Orthotropic Materials,'' Int. J. Solids Struct., **33**, No. 17, pp. 2459–2484.

[21] Vel, S. S., and Batra, R. C., 2000, ''The Generalized Plane Strain Deformation of Thick Anisotropic Composite Plates,'' Int. J. Solids Struct., **37**, No. 5, pp. 715–733.

[22] Choi, I., and Horgan, C. O., 1977, ''Saint-Venant's Principle and End Effects in Anisotropic Elasticity,'' ASME J. Appl. Mech., **44**, pp. 424–430.

[23] Choi, I., and Horgan, C. O., 1978, ''Saint-Venant End Effects for Plane Deformation Sandwich Strips,'' Int. J. Solids Struct., **14**, pp. 187–195.

F. G. Rammerstorfer
Professor and Head,
Institute of Lightweight Structures
and Aerospace Engineering,
Vienna University of Technology,
Gusshausstr. 27-29,
A-1040 Vienna, Austria
e-mail: ra@ilfb.tuwien.ac.at

F. D. Fischer
Professor and Head,
Institute of Mechanics,
Montanuniversitat Leoben,
Franz-Josef-Strasse 18,
A-8700 Leoben, Austria
e-mail: mechanik@unileoben.ac.at
Mem. ASME

N. Friedl
University Assistant,
Institute of Lightweight Structures
and Aerospace Engineering,
Vienna University of Technology,
A-1040 Vienna, Austria

Buckling of Free Infinite Strips Under Residual Stresses and Global Tension

Long free strips of small thickness frequently show a wavy surface being the consequence of buckling due to residual stresses. The paper deals with the derivation of the corresponding relations under conditions which are frequently met in the case of edge wave buckling of rolled strips or in deployable structures. It is shown that an increasing global tension force does not only lead to increased critical residual stress intensity but also produces shorter buckling waves concentrated towards the edges of the strip. By introducing dimensionless quantities, diagrams and formulas are provided which allow the determination of critical loading combinations. Asymptotic considerations are also presented. [DOI: 10.1115/1.1357519]

1 Introduction

In a recent paper [1] the authors have presented analytical and numerical considerations of buckling phenomena in thin plates or strips under in-plane loads which typically appear during rolling and leveling of sheet metal. Buckling due to self-equilibrating residual stresses, caused by the rolling process, possibly in conjunction with global tensile stresses (denoted there as ''rolling buckling'') as well as buckling during the levelling process (denoted there as ''stretching buckling'' or ''towel buckling'') were considered. Analytical estimates were derived and compared against results of numerical simulations and field observations. Different buckling phenomena appearing during the rolling process have been investigated in several other papers, as, for instance, in the classical one by U. Fischer [2], or more recently by Y. Tomita and H. Shao [3] or K. Komori [4].

In [1] jumps of buckling modes due to variation of the global strip tension were found on the basis of the derived analytical solutions. A more detailed analysis, which is presented now, confirms the estimations derived in [1] and focuses on the variation of the buckling pattern with increased global strip tension. It is shown that, in addition to mode jumps, the buckling pattern changes continuously over a wide range of the tension force, which was not found in [1]. This continuous mode change is mathematically described by the conditions for envelopes of functions with varying parameters. Further extensions in comparison to [1] are that a wide variety of different characters of residual stress distributions are considered and analytical closed-form solutions are provided for describing the buckling modes.

In [5] strip buckling under homogeneous longitudinal tensile stesses is considered under specific boundary conditions. Here the boundary conditions are rather trivial (infinitely long strip) and the residual membrane stress distributions are assumed to be self-equilibrated. Despite this, buckling can happen under external tension.

Although the problem treated here is of general interest in plate buckling analysis the results are of practical relevance, too; for instance, with respect to the control of the rolling process in strip rolling mills or the reliability of the deployment of deployable structures containing strip shaped members. Regarding flatness in cold rolling of strip metal an overview on the current literature can be found in [6].

2 Formulation of the Mathematical Model

A strip of infinite length (representing, for example, a sheet metal in the rolling path) with the width B and a plate bending stiffness $K=Et^3/(12(1-\nu^2))$ (with E being the Young's modulus, ν the Poisson's ratio, and t the thickness), is loaded by a self-equilibrating residual membrane force distribution ${}^{R}n_{xx}(y)=N\hat{g}(y)$ and a constant global tensile membrane force N_0. Thus, the membrane force distribution is (independent of the longitudinal coordinate x) given by

$$n_{xx}(y)=N\hat{g}(y)+N_0, \tag{1}$$

(see Fig. 1).

The following dimensionless quantities are introduced:

$$\eta=y/B, \text{ with } -B/2\leqslant y\leqslant B/2, -\tfrac{1}{2}\leqslant\eta\leqslant\tfrac{1}{2},$$
$$\hat{g}(y)\rightarrow g(\eta), n_{xx}\rightarrow n_{xx}(\eta)=Ng(\eta)+N_0, \tag{2}$$
$$\tilde{N}=\frac{NB^2}{K\pi^2}, \quad \tilde{N}_0=\frac{N_0B^2}{K\pi^2}.$$

Because ${}^{R}n_{xx}$ must be self-equilibrated the following condition must hold:

$$\int_{-1/2}^{+1/2} g(\eta)\,d\eta=0, \text{ i.e., for symmetrical } g(\eta):$$
$$\int_{0}^{1/2} g(\eta)\,d\eta=0. \tag{3}$$

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, June 14, 2000; final revision, September 15, 2000. Associate Editor: S. Kyriakides. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

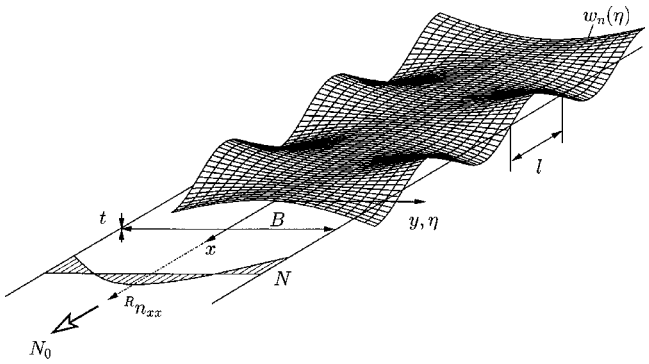


Fig. 1 The strip buckled under residual membrane forces and global tension

In order to capture a wide range of symmetrical residual membrane force distributions which typically lead to ''edge wave buckling,'' i.e., a pattern showing waves with maximum wave heights at the edges of the strip, the following distributions are assumed:

(a) Cosine distribution:

$$g^c(\eta)=C_m\cos^m\pi\eta\ -1 \quad \text{with} \quad m=1,2,\ldots \quad \text{and} \quad -\tfrac{1}{2}\leqslant\eta\leqslant\tfrac{1}{2}. \tag{4}$$

The equilibrium condition for $^R n_{xx}(\eta)$ leads to

$$C_m=\frac{1}{2}\left[\int_0^{1/2}\cos^m\pi\eta\ d\eta\right]^{-1}. \tag{5}$$

(b) Polynomial distribution:

$$g^p(\eta)=\frac{1}{m}[1-(m+1)(2|\eta|)^m] \quad \text{with} \quad m=1,2,\ldots. \tag{6}$$

which also complies with the equilibrium condition (3).

The superscripts ''c'' and ''p'' refer to the cosine and the parabolic residual membrane force distributions, respectively.

Due to the variable exponent m in $g^c(\eta)$ and $g^p(\eta)$ the required large variability of residual stress distributions can be provided (see Fig. 2). It should be noted that for the cosine distribution increasing exponents m lead to wider regions of compression

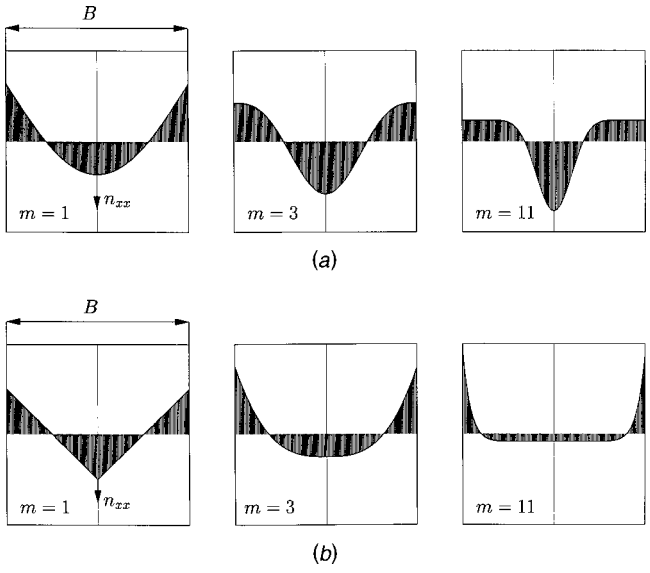


Fig. 2 Typical residual membrane force distributions; (*a*) according to (4), (*b*) according to (6).

membrane forces, while in the case of polynomial distributions higher m-values are typical for residual compressive forces concentrated in regions close to the edges of the strip. For $m=2$ both types of force distributions look very similar. As a consequence of this fact, for $m=2$ the results for cosine and polynomial distributions are almost identical.

Both distributions, i.e., the cosine and the polynomial one, lead to

$$g(\eta=\pm\tfrac{1}{2})=-1. \tag{7}$$

This means that a positive value of N in Eq. (1) corresponds to compressive residual membrane forces in the edge regions of the strip.

In order to apply the Ritz approach as described in [1] we introduce a Ritz-ansatz, i.e., a trial function, for the buckling pattern

$$w(x,\eta)=qw_n(x,\eta)=q\cos\frac{\pi x}{l}(2|\eta|)^n(\operatorname{sign}\eta)^k, \tag{8}$$

with $k=1\ldots.$ antisymmetrical mode; $k=2\ldots$ symmetrical mode.

As shown in [1] corresponding symmetrical and antimetrical modes lead to the same critical residual stress intensities and, therefore, $(\operatorname{sign}\eta)^k$ in Eq. (8) can be omitted.

Both the half-wave length l and the exponent n, which allow a wide variety of the buckling pattern, have to be determined appropriately in order to estimate the relevant, i.e., the minimum, buckling force amplitude as a function of the global tension force: $N^c(N_0)$ or $N^p(N_0)$ for a given exponent m in Eq. (4) or (6), respectively.

3 Analysis

We follow the ''classical'' concept of minimization of the total potential energy ϕ of a thin plate deformed also in the direction normal to the middle surface and subjected to a membrane force state $^R n_{xx}+N_0$. The expression for the total potential energy can be taken from the literature; see, e.g., [7], and specialized to this problem reads as follows:

$$\phi_B=\frac{K}{2}\int_\Omega\left\{\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}\right)^2-2(1-\nu)\times\left[\frac{\partial^2 w}{\partial x^2}\frac{\partial^2 w}{\partial y^2}-\left(\frac{\partial^2 w}{\partial x\partial y}\right)^2\right]\right\}d\Omega, \tag{9}$$

$$\phi_M=\frac{1}{2}\int_\Omega(^R n_{xx}+N_o)\left(\frac{\partial w}{\partial x}\right)^2 d\Omega, \tag{10}$$

$$\phi_{N_0}=-\frac{N_0^2 LB}{2Et}, \quad \phi=\phi_B+\phi_M+\phi_{N_0}. \tag{11}$$

The application of Gauss's theorem leads to the following modified formulation of ϕ_B

$$\phi_B=\frac{K}{2}\int_\Omega\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}\right)^2 d\Omega+K(1-\nu)\int_{\partial\Omega}\frac{\partial w}{\partial x}\frac{\partial}{\partial y}\left(\frac{\partial w}{\partial s}\right)ds. \tag{12}$$

ϕ_B is the contribution to ϕ due to bending, ϕ_M that due to the membrane forces, and ϕ_{N_0} is the potential energy of the boundary force N_0 at $x=0$ and $x=L\gg B$, which, however, does not play any further role in the linearized theory.

According to the procedure with a single-term shape function we have just one degree-of-freedom, which is related to the amplitude q of the buckling mode (see Eq. (8)). Stationarity of $\phi(qw_n)$ requires $\partial\phi/\partial q=0$, which immediately leads to the equation

$$\phi_B(w_n)+\phi_M(w_n)=0. \tag{13}$$

If we now insert the membrane state according to Eq. (4) or (6) which depends on a positive constant N (see Eq. (1)), the Eq. (13) can be rewritten with the dimensionless variables defined by Eq. (2):

$$\tilde{N}\phi'_{M,N}(w_n)+\tilde{N}_0\Phi'_{M,N_0}(w_n)+\phi_B(w_n)=0, \tag{14}$$

where $\phi'_{M,N}$ and ϕ'_{M,N_0} are the contributions to ϕ_M due to $\tilde{N}=1$ and $\tilde{N}_0=1$, respectively. This leads to

$$\tilde{N}_n=-\frac{\phi_B(w_n)+\tilde{N}_0\phi'_{M,N_0}(w_n)}{\phi'_{MN}(w_n)} \tag{15}$$

as critical residual membrane force intensity corresponding to the Ritz-Ansatz with exponent n.

With Eq. (8) after some algebraic manipulation the following mathematical structure of Eq. (15) is obtained:

$$\tilde{N}_n=-\left[\tilde{N}_0\frac{1}{2n+1}+\frac{(B/l)^2}{2n+1}-\frac{8n(n-1)}{\pi^2(2n-1)}+\frac{16(l/B)^2n^2(n-1)^2}{\pi^4(2n-3)}+\frac{8(1-\nu)n}{\pi^2}\right]\Bigg/\left[2\int_0^{1/2}g(\eta)(2\eta)^{2n}d\eta\right]. \tag{16}$$

The right-hand side of Eq. (16) contains a further unknown entity, namely the half-wavelength l in the x-direction—expressed in the dimensionless quantity l/B. Since we look for the minimum possible value of the buckling parameter $\tilde{N}_n$, we find l by minimizing $\tilde{N}_n$ with respect to this quantity:

$$\frac{d\tilde{N}_n}{d(l/B)}=0. \tag{17}$$

This additional relation allows us to find $l_{\min}$ and, as a consequence, $\tilde{N}_{n^{\min}}(\tilde{N}_0)$. In the following text these quantities will be used without the subscript ''min.'' Equation (17) yields, after simple calculation,

$$l/B=\sqrt{\kappa} \tag{18}$$

with

$$\kappa=\frac{\pi^2}{4}\left(\frac{2n-3}{n^2(n-1)^2(2n+1)}\right)^{1/2}. \tag{19}$$

The relevant exponent n will be determined below.

It should be mentioned that l/B does not depend on the assumed residual membrane force distribution. This is because $g(\eta)$, which determines the membrane force distribution, appears only in the denominator of the r.h.s. of Eq. (16) which does not contain l/B. Hence it disappears when Eq. (17) is evaluated.

It can be shown that the contribution due to $\partial^2 w/\partial y^2$ in the bending related portion of the potential ϕ (see Eq. (9)) vanishes for $n=0$ and for $n=1$. For other values of the exponent $n\in\Re_+$ in Eq. (8) the condition $n>\frac{3}{2}$ must be met in order to achieve physically meaningful results (i.e., no singularities in the corresponding strain energy density).

Figure 3 shows this result in graphical form.

For $n>1.73$ increasing n leads to decreasing l. The interval [1.5, 1.73] appears to be aphysical; however, it is of no relevance because, as will be shown below, the solutions either with $n=1.0$ or $n>1.73$ lead to smaller critical residual membrane force intensities than those with $n\in[1.5,1.73]$.

For the cosine distribution the solution of the Ritz approach described above leads to

$$\tilde{N}_n^c(\tilde{N}_0)=f_1^c(n)+f_2^c(n)\tilde{N}_0 \tag{20}$$

with

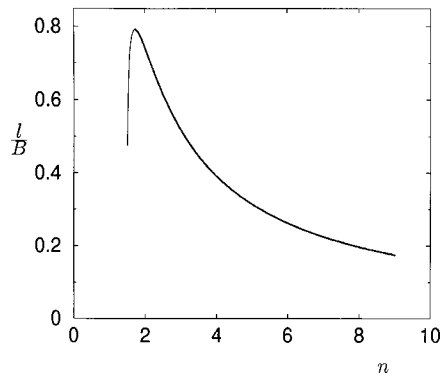


Fig. 3 Half-wavelength of the relevant buckling mode as a function of the exponent n in Eq. (8)

$$f_1^c(n,m)=-\left[\frac{1}{2n+1}\frac{1}{\kappa}-\frac{8}{\pi^2}\frac{n(n-1)}{2n-1}+\frac{16}{\pi^4}\frac{n^2(n-1)^2}{2n-3}\kappa+\frac{8(1-\nu)n}{\pi^2}\right]\Bigg/\left[2\int_0^{1/2}[C_m(\cos\pi\eta)^m-1]\times(2\eta)^{2n}d\eta\right] \tag{21}$$

and

$$f_2^c(n,m)=-1\Bigg/\left[2(2n+1)\int_0^{1/2}[C_m(\cos\pi n)^m-1]\times(2\eta)^{2n}d\eta\right], \tag{22}$$

with κ given in Eqs. (18) and (19).

For the parabolic distribution of the residual membrane force we obtain

$$\tilde{N}_n^p(\tilde{N}_0)=f_1^p(n)+f_2^p(n)\tilde{N}_0, \tag{23}$$

with

$$f_1^p(n,m)=\frac{2}{\pi^2}(2n+m+1)\left\{2(n-1)\left[\left(\frac{2n+1}{2n-3}\right)^{1/2}-\frac{2n+1}{2n-1}\right]+2(1+\nu)(2n+1)\right\} \tag{24}$$

and

$$f_2^p(n,m)=\frac{(2n+m+1)}{2n}. \tag{25}$$

These functions are shown in graphical form in Fig. 4.

For a given m, representing the character of the residual membrane force distribution, the critical intensity $\tilde{N}_n(\tilde{N}_0)$ represents for every exponent n a straight line in the $\tilde{N}-\tilde{N}_0$-diagram (Eq. (23) is a linear relation). Figure 5 shows this situation for $m=1$ in the case of a cosine distribution of the residual membrane force. This example is comparable with one of those considered in [1].

This figure makes clear how the proper exponent n has to be found as a function of the global tension $\tilde{N}_0$ and the given exponent m: Finally, the critical residual membrane force intensity must be

$$\tilde{N}(\tilde{N}_0)=\min_n\tilde{N}_n(\tilde{N}_0). \tag{26}$$

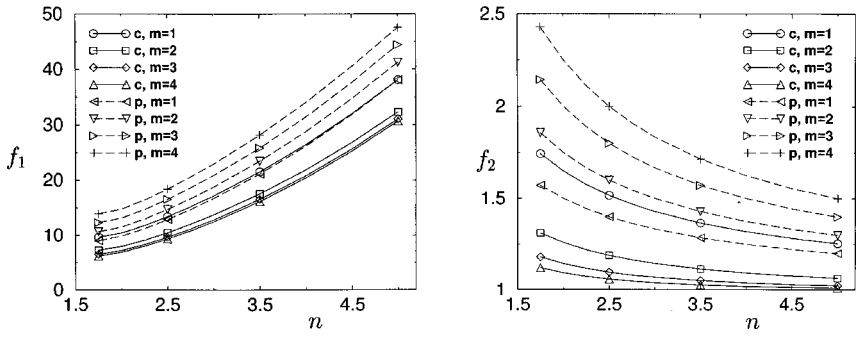


Fig. 4 Functions $f_{1,2}^{c,p}(n,m)$ for different exponents m; full lines refer to cosine distributions, broken lines to parabolic distribution of the residual membrane force

This means $\tilde{N}(\tilde{N}_0)$ is represented by the inner envelope as formed by the manifold of straight lines $\tilde{N}(\tilde{N}_0)$ with a continuously varying exponent n.

As long as the line $\tilde{N}_{n=1.0}(\tilde{N}_0)$ is below any other $\tilde{N}_n(\tilde{N}_0)$ line with $n>1.73$, i.e., for $0\leqslant\tilde{N}_0\leqslant\tilde{N}_0^*$, this inner envelope is given by this $\tilde{N}_{n=1.0}(\tilde{N}_0)$ line. The value $\tilde{N}_0=\tilde{N}_0^*$, corresponds with the situation where the buckling mode jumps from a pure twisting to a wavy pattern; see point A in the detail in Fig. 5. For $\tilde{N}_0>\tilde{N}_0^*$, i.e., for the region in which $\tilde{N}_n(\tilde{N}_0)$ with $n>1.73$ is smaller than $\tilde{N}_{n=1.0}(\tilde{N}_0)$, this envelope is, for given m, implicitly described by

$$F(\tilde{N},\tilde{N}_0,n)=\tilde{N}-f_1(n)-f_2(n)\tilde{N}_0=0. \tag{27}$$

Therefore, the condition for the envelope

$$\frac{\partial F}{\partial n}=-\frac{df_1}{dn}-\frac{df_2}{dn}\tilde{N}_0=0, \tag{28}$$

yields

$$\tilde{N}_0=-\frac{df_1}{dn}\Big/\frac{df_2}{dn}=\tilde{N}_0(n). \tag{29}$$

This way the critical intensity can be expressed as a function of n only instead of $\tilde{N}(\tilde{N}_0)$:

$$\tilde{N}(n)=f_1(n)-f_2(n)\left(\frac{df_1}{dn}\Big/\frac{df_2}{dn}\right) \tag{30}$$

Finally, $\tilde{N}(\tilde{N}_0)$ can be determined by evaluating Eqs. (29) and (30) for continuously varying n.

Equations (29) and (30) require the first derivatives of $f_1(n)$ and $f_2(n)$ which are calculated numerically for cosine distributions of the residual membrane force and analytically for parabolic distributions. In the latter case these expressions read:

$$\frac{df_1^p}{dn}=\frac{4}{\pi^2}n\left\{\left(4+\frac{m-1}{n}\right)\left[\left(1+\frac{4}{2n-3}\right)^{1/2}-\frac{2}{2n-1}+\frac{3(1-\nu)}{n-1}+(1-2\nu)\right]+\frac{1}{n}\left(2+\frac{m-1}{n}-\frac{m+1}{n^2}\right)\times\left[-4\left(1+\frac{4}{2n-3}\right)^{-1/2}\frac{1}{(2-3/n)^2}+\frac{4}{(2-1/n)^2}-\frac{3(1-\nu)}{(1-1/n)^2}\right]\right\}$$

$$\frac{df_2^p}{dn}=-\frac{m+1}{2n^2}. \tag{31}$$

Figure 6 shows the residual membrane force intensities $\tilde{N}(\tilde{N}_0)$ for different exponents m, and Fig. 7 shows how the half-wave length l/B decreases with increasing $\tilde{N}_0$ for different exponents m.

Figure 5 shows that a sudden mode jump from the pure twisting mode, i.e., $n=1$, to a mode with $n>1.73$ appears at a certain value of $\tilde{N}_0$, denoted as $\tilde{N}_0^*$. In the case of a cosine-distribution of the residual membrane force with $m=1$ this jump appears at $\tilde{N}_0^*\approx 10$. According to Eq. (19) the half-wavelength assigned to the pure twisting mode ($n=1$) goes to infinity; that means that a very long strip is just tilted over the whole length, compare [1].

This is the reason why the lines in Fig. 7, showing the half-wavelength l/B in dependence of the global strip tension $\tilde{N}_0$ for different exponents m, must not be extrapolated for small values

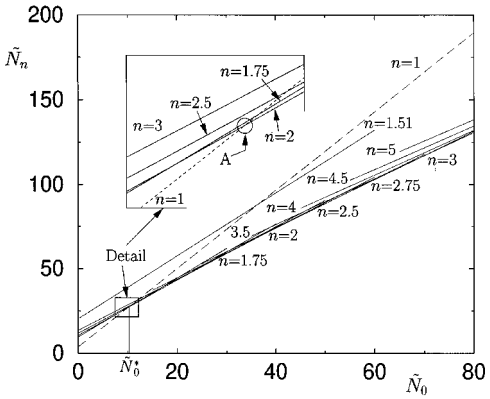


Fig. 5 Dependence of $\tilde{N}_n$ on $\tilde{N}_0$ for a cosine distribution of the residual membrane force with $m=1$. The detail shows that $n\in[1.5,1.73]$ is not relevant.

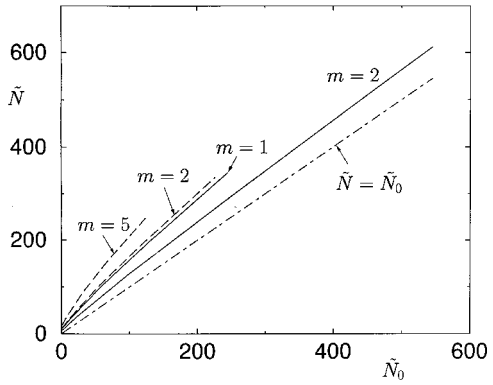


Fig. 6 Dependence of the critical residual membrane force intensity $\tilde{N}$ on the global strip tension $\tilde{N}_0$ for different exponents m; full lines refer to cosine distributions, broken lines to parabolic distributions of the residual membrane force

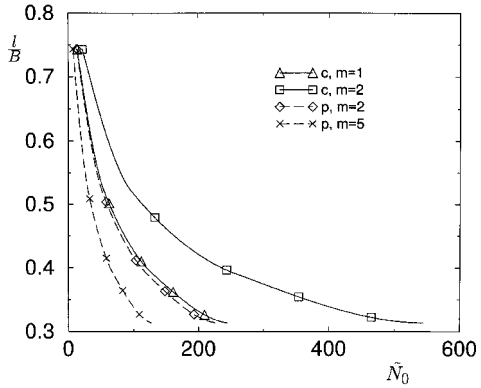


Fig. 7 Dependence of the half-wavelength l/B on the global strip tension $\tilde{N}_0$ for different exponents m; full lines refer to cosine distributions, broken lines to parabolic distributions of the residual membrane force

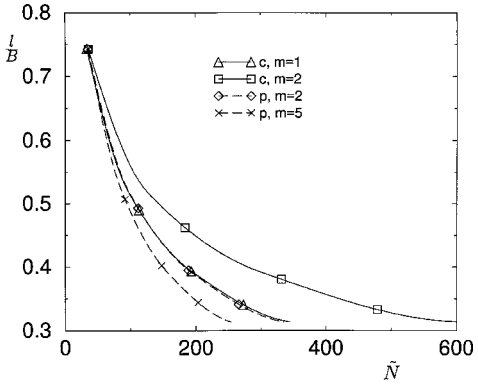


Fig. 8 Dependence of the half-wavelength l/B on the intensity of the residual membrane force distribution $\tilde{N}$ for different exponents m; full lines refer to cosine distributions, broken lines to parabolic distributions of the residual membrane force

of $\tilde{N}_0$; they end at $\tilde{N}_0=\tilde{N}_0^*$, where a jump of the half-wavelength l/B from a finite value to infinity (for an infinitely long strip) happens.

For $\tilde{N}_0>\tilde{N}_0^*$ the buckling mode changes continuously, i.e., n increases monotonically and smoothly. Since for the parabolic distribution the relations $\tilde{N}_{n=1.0}(\tilde{N}_0)$ and $\tilde{N}(\tilde{N}_0)$, i.e., the envelope, can be expressed in analytical form, $\tilde{N}_0^*$ can be determined also in analytical form. This is, however, not so easy for the cosine distribution, for which closed-form expressions are not available.

4 Asymptotic Considerations

The shape of the envelope might lead to the expectation of an asymptotic approach of the critical states to the boundary lines $\tilde{N}\to\tilde{N}_0$ for very large $\tilde{N}_0$; see [1].

For parabolic distributions of the residual stresses asymptotic considerations can be performed analytically. As mentioned above, increasing $\tilde{N}_0$ leads to increasing n and, due to Eq. (19), to decreasing l/B.

Due to this fact considerations of large $\tilde{N}_0$ can be performed on the basis of expressions for large n. For very large n, $n\gg 1$, we come up with the following simplified formulas:

$$f_1\to\frac{16}{\pi^2}(1-\nu)n^2, \tag{32}$$

$$f_2\to 1, \tag{33}$$

and

$$\tilde{N}_0(n)\to\frac{64}{\pi^2}\frac{1-\nu}{m+1}n^3, \tag{34}$$

and finally

$$\tilde{N}(\tilde{N}_0)\to\frac{16(1-\nu)}{\pi^2}\left(\frac{\pi^2(m+1)}{64(1-\nu)}\right)^{2/3}\tilde{N}_0^{2/3}+\tilde{N}_0. \tag{35}$$

Hence, it can be seen that the difference $\tilde{N}-\tilde{N}_0$ grows with increasing global strip tension $\tilde{N}_0$ in the order of $\tilde{N}_0^{2/3}$. Therefore, the above-mentioned expectation $\tilde{N}\to\tilde{N}_0$ for very large values of $\tilde{N}_0$, as argued in [1], could not be confirmed.

5 Practical Observations

During strip metal rolling processes the global tension force usually is large enough to prevent the strip from buckling under residual stresses. However, if this tension force is reduced a critical situation can be reached in which suddenly a wavy deformation pattern can be observed. This pattern represents the post-buckling deformations initiated by bifurcation at $\tilde{N}_0(\tilde{N})$.

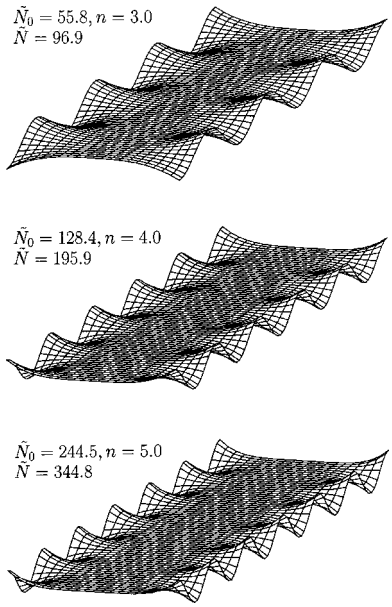


Fig. 9 Buckling patterns for different global tension forces $\tilde{N}_0$

The corresponding buckling mode shows different half-wavelengths, depending on the distribution and on the intensity of the residual membrane force distribution, see Fig. 8.

Figure 9 shows how the buckling pattern changes with increasing global tension force $\tilde{N}_0$. Increasing $\tilde{N}_0$ leads to a concentration of the buckles towards the edges of the strip and to a decrease of the buckling wavelength.

6 Conclusions

The buckling and post-buckling patterns of thin strips under the action of residual membrane forces and global tension show interesting features: With rising global strip tension a single-mode jump can be observed from the pure twisting mode to a wavy mode; then, as the global tension continues to increase, the buckling mode remains wavy with continuously decreasing wavelengths and waves concentrating more and more towards the edges of the strip.

It has been shown that the amplitudes, i.e., the intensities, of different residual membrane force distributions which are critical

with respect to buckling of a strip can be calculated analytically as functions of the global tension force. A wide range of characteristic residual membrane force distributions has been considered.

Formulas and diagrams are provided which, in combination with the dimensionless formulation presented in the paper, ensure quick and simple estimation of the critical residual membrane force intensities. This is of practical importance, for instance, in the case of buckling during the rolling process of thin strip metal.

References

[1] Fischer, F. D., Rammerstorfer, F. G., Friedl, N., and Wieser, W., 2000, ''Buckling Phenomena Related to Rolling and Levelling of Sheet Metal,'' Int. J. Mech. Sci., **42**, pp. 1887–1910.

[2] Fischer, U., 1976, ''Stabilität des freien Plattenstreifens mit eine Gleichgewichtsgruppe bildenden Längseigenspannungen,'' Z. Angew. Math. Mech., **56**, p. 331.

[3] Tomita, Y., and Shao, H., 1993, ''Buckling Behavior in Thin Sheet Metal Subjected to Nonuniform Membrane-Type Deformation,'' *Advances in Engineering Plasticity and its Applications* W. B. Lee, ed., Elsevier, Amsterdam, pp. 923–930.

[4] Komori, K., 1998, ''Analysis of Cross and Vertical Buckling in Sheet Metal Rolling,'' Int. J. Mech. Sci., **40**, pp. 1235–1246.

[5] Friedl, N., Rammerstorfer, F. G., and Fischer, F. D., 1999, ''Zum Beulen von Platten unter globalem Zug,'' Z. Angew. Math. Mech., **79**, Supplement 2, pp. 545–546.

[6] Tarnopolskaya, T., and de Hoog, F. R., 1998, ''An Efficient Method for Strip Flatness Analysis in Cold Rolling,'' Math. Engng. Ind., **7**, pp. 71–95.

[7] Yuan, S., and Jin, Y., 1998, ''Computation of Elastic Buckling Loads of Rectangular Thin Plates Using the Extended Kantorovich Method,'' Comput. Struct., **66**, pp. 861–867.

Stress Concentration Reduction at a Reinforced Hole Loaded by a Bonded Circular Inclusion

K. T. Chau
Mem. ASME

X. X. Wei

Department of Civil and Structural Engineering,
The Hong Kong Polytechnic University,
Kowloon, Hong Kong, China
e-mail: cektchau@polyu.edu.hk

This paper considers analytically the stress concentration in an infinite plane loaded by a circular inclusion, which is bonded to a reinforced hole in the plane. The pulling force of the inclusion is modeled by distributed body force. The infinite plane, the reinforced ring, and the circular inclusion can be of different elastic properties. Airy stress function with body force potential was used to solve the problem analytically. Numerical results show that the maximum tensile hoop stress at the hole boundary in the plane can be reduced to becoming negligible if an optimum stiffness ratio between the plane and the rivet is chosen (normally a harder material for the reinforced ring comparing to the plane is needed). An optimum thickness of the reinforced ring can also be determined to further reduce the hoop stress concentration. Therefore, the results of the present study provide a new theoretical basis for designing a reinforced rivet hole. [DOI: 10.1115/1.1357869]

1 Introduction

Stress concentration at a hole loaded by an elastic inclusion is one of the classical problems in linear elasticity ([1]). One main application of the inclusion problems is in the design of rivet-connected structures and components (e.g., [2,3]). Since the early 20th century, engineers, mathematicians, and experimentalists have worked on the inclusion problems in various branches of engineering and mechanics.

One of the first experimental studies on the stress concentration due to rivet load is the photoelastic experiment done by Coker [4] while the first theoretical treatment is apparently done by Brickley [5]. Among the circular inclusion problems, an elastic plane loaded by a pinned-rivet has received the most attention (e.g., [6–15], and others). One simplified approach is to assume the circular inclusion or the rivet as rigid ([9,16–18]). When the elastic circular inclusion is bonded to the elastic plane, exact solution can normally be obtained (e.g., [19–23]). However, when separation is allowed to develop between the plane and the circular inclusion under external loads, numerical approach is normally required to solve the resulting equations ([18,24–31]). Because the maximum tensile stress concentration is located at the hole boundary, the idea of reinforcing the hole boundary has been proposed ([3,32–38]) and has long been used in practice. The reinforcement is normally made of a thicker rim of the same material around the rivet hole ([3,39,40]).

In this paper, we propose a new approach that a different material for the reinforced ring is used, instead of using a ''compact'' rim approach ([3]). Another new feature of the present analysis is the use of body force in modeling the rivet-load, instead of assuming a concentrated load located at the center of the circular inclusion (e.g., [6,18]). The idea of modeling rivet force as uniform distribution body force is apparently from Hyer and Klang [14], and this approach has further been extended to case of non-uniform distributed load by Ho and Chau [23]. The same approach has also been used to consider the stress in finite strip loaded by a bonded-rivet made of a different material ([41]). In fact, except for the recent studies by Ho and Chau [23,41], none of the previous studies allow the rivet and the elastic plane to have different elastic properties. The main motivation of assuming a different material for the rivet steps from the application of the elastic solution to steel rivet-rock panel systems in cladding wall design ([23]).

To simplify our analysis and to model the epoxy sealed connections in rock panel systems, only the case of bonded inclusion was considered. And our main focus will be on the tensile hoop stress at the reinforced hole, as the hoop stress has long been recognized as the most important stress around the hole ([42]).

2 Mathematical Formulation

For two-dimensional stress analysis in cylindrical coordinates (r,θ), we adopt the Airy stress function φ with nonzero conservative body force ([22]):

$$\sigma_{rr}=\frac{1}{r}\frac{\partial\varphi}{\partial r}+\frac{1}{r^2}\frac{\partial^2\varphi}{\partial r^2}+V,$$

$$\sigma_{r\theta}=-\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\varphi}{\partial\theta}\right),\quad \sigma_{\theta\theta}=\frac{\partial^2\varphi}{\partial r^2}+V \tag{1}$$

where V is the conservative body force potential. For isotropic elastic materials, substitution of these stress components into the compatibility equation leads to the following governing equation for φ:

$$\nabla^4\varphi=\nabla^2\nabla^2\varphi=-2\left(\frac{\kappa-1}{\kappa+1}\right)\nabla^2V \tag{2}$$

where

$$\nabla^2=\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2} \tag{3}$$

is the plane Laplacian operator in polar form. The two-dimensional elastic constant κ equals to $3-4\nu$ for plane strain and $(3-\nu)/(1+\nu)$ for generalized plane stress, where ν is the Poisson's ratio of the material.

For the reinforced rivet problem shown in Fig. 1, the stress must be bounded in Domain II when r approaches zero, while the far-field stress must decay to zero as r approaches infinity. On the interface between Domains I and III (i.e., the infinite plane and the reinforced ring, respectively), the continuities between tractions and displacements must be enforced. The rivet load is transmitted from the rivet (Domain II) to the infinite plane (Domain I) through the contact with the reinforced ring (Domain III). In particular, the following conditions are prescribed on $r=R$:

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 5, 2000; final revision, Oct. 19, 2000. Associate Editor: J. R. Barber. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

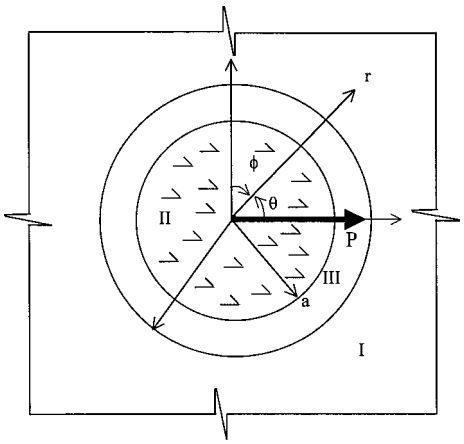


Fig. 1 A sketch for a circular elastic inclusion embedded into a reinforced hole in an elastic plane. The rivet load is modeled by uniformly distributed body force. Domains I, II, and III are the elastic plane, the circular inclusion, and the reinforced ring.

$$\sigma_{rr}^{(3)}(R,\theta)=\sigma_{rr}^{(1)}(R,\theta), \tag{4}$$

$$\sigma_{r\theta}^{(3)}(R,\theta)=\sigma_{r\theta}^{(1)}(R,\theta), \tag{5}$$

$$\boldsymbol{u}_{r}^{(3)}(R,\theta)=\boldsymbol{u}_{r}^{(1)}(R,\theta), \tag{6}$$

$$\boldsymbol{u}_{\theta}^{(3)}(R,\theta)=\boldsymbol{u}_{\theta}^{(1)}(R,\theta), \tag{7}$$

where **u** is the displacement and the bracketed superscripts denote the Domain number. The first two of these equations represent the continuities in normal and shear tractions, whilst the latter two represent continuities in radial and tangential displacements. Similar continuity conditions between Domains II and III on $r=a$ can be obtained by replacing the superscript (1) by ''(2)'', and R by ''a'' in (4)–(7).

3 Determination of Stresses

As shown in Fig. 1, one new feature of the present analysis is that the rivet load is modeled by uniform distributed body force on the rivet instead of considering a concentrated force at the center of the rivet (e.g., [6]). The body force potential for this uniform distributed body force was given by Karasudhi [22]:

$$V=-f_0 r\cos\theta \tag{8}$$

where f_0 is the magnitude of the uniform distributed load (in force per unit volume). The total rivet force per unit thickness P (shown in Fig. 1) can be determined by $f_0\pi a^2$. The normal stresses induced by this body force potential V is proportional to $\cos\theta$. In view of this, the following forms of Airy stress functions for Domains I–III are proposed ([23]):

$$\varphi^{(1)}=f_0R^2\left[Ar\theta\sin\theta+Br\ln r\cos\theta+\frac{CR^2\cos\theta}{r}\right], \tag{9}$$

$$\varphi^{(2)}=f_0Dr^3\cos\theta, \tag{10}$$

$$\varphi^{(3)}=f_0\left[Er^3\cos\theta+FR^2r\theta\sin\theta+GR^2r\ln r\cos\theta+\frac{HR^4\cos\theta}{r}\right], \tag{11}$$

where A, B, C, D, E, F, G, and H are unknown constants to be determined by the continuity conditions. Note that both the boundedness condition at the center and the decay condition at infinity are identically satisfied by these choices.

Substitution of (8)–(11) into (1) leads to the stresses in terms of the unknown constants. Integration of these resulting stresses yields the expressions of displacements in terms of the unknown constants. More specifically, it is straightforward to show the validity of the following expressions:

Domain I:

$$\sigma_{rr}^{(1)}=f_0\left[2AR^2\frac{\cos\theta}{r}+BR^2\frac{\cos\theta}{r}-2CR^4\frac{\cos\theta}{r^3}\right], \tag{12}$$

$$\sigma_{r\theta}^{(1)}=f_0\left[BR^2\frac{\sin\theta}{r}-2CR^4\frac{\sin\theta}{r^3}\right], \tag{13}$$

$$\sigma_{\theta\theta}^{(1)}=f_0\left[6BR^2\frac{\cos\theta}{r}+2CR^4\frac{\cos\theta}{r^3}\right], \tag{14}$$

$$2\mu_1u_r^{(1)}=\frac{f_0R^2}{2}\{[A(\kappa_1+1)+B(\kappa_1-1)]\ln r\cos\theta-(A+B)\cos\theta+[A(\kappa_1-1)+B(\kappa_1+1)]\theta\sin\theta\}+f_0CR^4\frac{\cos\theta}{r^2} \tag{15}$$

$$2\mu_1u_\theta^{(1)}=\frac{f_0R^2}{2}\{-[A(\kappa_1+1)+B(\kappa_1-1)]\ln r\sin\theta-(A+B)\sin\theta+[A(\kappa_1-1)+B(\kappa_1+1)]\theta\cos\theta\}+f_0CR^4\frac{\sin\theta}{r^2} \tag{16}$$

Domain II:

$$\sigma_{rr}^{(2)}=f_0(2D-1)r\cos\theta, \tag{17}$$

$$\sigma_{\theta\theta}^{(2)}=f_0(6D-1)r\cos\theta, \tag{18}$$

$$\sigma_{r\theta}^{(2)}=2f_0Dr\sin\theta, \tag{19}$$

$$2\mu_2u_r^{(2)}=f_0\left[D(\kappa_2-2)-\frac{\kappa_2-1}{4}\right]r^2\cos\theta+f_0a^2C_1\sin\theta+f_0a^2C_2\cos\theta, \tag{20}$$

$$2\mu_2u_\theta^{(2)}=f_0\left[D(\kappa_2+2)-\frac{\kappa_2-1}{4}\right]r^2\sin\theta+f_0a^2C_1\cos\theta-f_0a^2C_2\sin\theta+C_3r, \tag{21}$$

Domain III:

$$\sigma_{rr}^{(3)}=f_0\left[2Er\cos\theta+2FR^2\frac{\cos\theta}{r}+GR^2\frac{\cos\theta}{r}-2HR^4\frac{\cos\theta}{r^3}\right], \tag{22}$$

$$\sigma_{r\theta}^{(3)}=f_0\left[2Er\sin\theta+GR^2\frac{\sin\theta}{r}-2HR^4\frac{\sin\theta}{r^3}\right], \tag{23}$$

$$\sigma_{\theta\theta}^{(3)}=f_0\left[6Er\cos\theta+GR^2\frac{\cos\theta}{r}+2HR^4\frac{\cos\theta}{r^3}\right], \tag{24}$$

$$2\mu_3u_r^{(3)}=\frac{f_0R^2}{2}\{[F(\kappa_3+1)+G(\kappa_3-1)]\ln r\cos\theta-(F+G)\cos\theta+[F(\kappa_3-1)+G(\kappa_3+1)]\theta\sin\theta\}+f_0HR^4\frac{\cos\theta}{r^2}+f_0E(\kappa_3-2)r^2\cos\theta+C_4R^2\sin\theta+C_5R^2\cos\theta \tag{25}$$

$$2\mu_3 u_\theta^{(3)}=\frac{f_0R^2}{2}\{-[F(\kappa_3+1)+G(\kappa_3-1)]\ln r\sin\theta$$
$$-(F+G)\sin\theta+[F(\kappa_3-1)+G(\kappa_3+1)]\theta\cos\theta\}$$
$$+f_0HR^4\frac{\sin\theta}{r^2}+f_0E(\kappa_3+2)r^2\sin\theta+C_4R^2\cos\theta$$
$$-C_5R^2\sin\theta+C_6Rr \tag{26}$$

where μ_i and κ_i $(i=1,2,3)$ are the shear modulus and plane elastic constant for Domain i. The additional constants C_i $(i=1,\ldots,6)$ are resulted from the integration process and correspond to the magnitudes of rigid-body displacements.

4 Determination of Unknown Coefficients

4.1 Single-valued Condition. By considering the uniqueness of the physical stress and displacement components, the multivalued terms involving $\theta\cos\theta$ and $\theta\sin\theta$ in (15)–(16) and (25)–(26) must be identically zero. This leads to the following conditions:

$$[A(\kappa_1-1)+B(\kappa_1+1)]=0 \tag{27}$$
$$F(\kappa_3-1)+G(\kappa_3+1)=0. \tag{28}$$

4.2 Continuity Conditions. Considering the continuity of normal stress on $r=R$ between Domains I and III, we substitute (12) and (22) into (4). This leads to

$$2A+B-2C-2E-2F-G+2H=0. \tag{29}$$

For the continuity of shear stress on $r=R$, we substitute (13) and (23) into (5) and obtain

$$B-2C-2E-G+2H=0. \tag{30}$$

Subtracting (30) from (29) yields

$$A=F. \tag{31}$$

Similar consideration for the normal and shear tractions on the interface between Domains II and III (i.e., on $r=a$) leads to the following equations:

$$2D-2E-\frac{2F}{\rho}-\frac{G}{\rho}+\frac{2H}{\rho^2}-1=0 \tag{32}$$
$$2D-2E-\frac{G}{\rho}+\frac{2H}{\rho^2}=0 \tag{33}$$

where $\rho=(a/R)^2$. Subtracting (33) from (32) immediately yields

$$F=-\frac{\rho}{2}. \tag{34}$$

The continuity of radial displacement between Domains I and III leads to

$$C_4=0, \tag{35}$$
$$\zeta_1A[(\kappa_1+1)\ln R-1]+\zeta_1B[(\kappa_1-1)\ln R-1]$$
$$+2\zeta_1C-2E(\kappa_3-2)-F[(\kappa_3+1)\ln R-1]$$
$$-G[(\kappa_3-1)\ln R-1]-2H-2C_5=0 \tag{36}$$

where $\zeta_1=\mu_3/\mu_1$. The continuity of tangential displacement between Domains I and III leads to the following equations:

$$C_4=0,\quad C_6=0, \tag{37}$$
$$\zeta_1A[-(\kappa_1+1)\ln R-1]-\zeta_1B[(\kappa_1-1)\ln R+1]$$
$$+2\zeta_1C-2E(\kappa_3+2)+F[(\kappa_3+1)\ln R+1]$$
$$+G[(\kappa_3-1)\ln R+1]-2H+2C_5=0. \tag{38}$$

Similarly, the continuity of radial displacement between Domains II and III gives

$$C_1=0, \tag{39}$$
$$\zeta_2D(\kappa_2-2)-\frac{1}{4}\zeta_2(\kappa_2-1)+\zeta_2C_2-E(\kappa_3-2)$$
$$-\frac{F}{2\rho}[(\kappa_3+1)\ln a-1]-\frac{G}{2\rho}[(\kappa_3-1)\ln a-1]$$
$$-\frac{C_5}{\rho}-\frac{H}{\rho^2}=0 \tag{40}$$

where $\zeta_2=\mu_3/\mu_2$. By considering the continuity of tangential displacement between Domains II and III yields the following equations:

$$C_1=C_3=0, \tag{41}$$
$$\zeta_2D(\kappa_2+2)-\frac{1}{4}\zeta_2(\kappa_2-1)-\zeta_2C_2-E(\kappa_3+2)$$
$$+\frac{F}{2\rho}[(\kappa_3+1)\ln a+1]+\frac{G}{2\rho}[(\kappa_3-1)\ln a+1]$$
$$+\frac{C_5}{\rho}-\frac{H}{\rho^2}=0. \tag{42}$$

Substitution of (34) into (28) and (31), then the result into (27) gives

$$B=\frac{\rho(\kappa_1-1)}{2(\kappa_1+1)}. \tag{43}$$

To eliminate C_5, we add (36) and (38) to obtain

$$\zeta_1A+\zeta_1B-2\zeta_1C+2E\kappa_3-F-G+2H=0. \tag{44}$$

Eliminating C_2 and C_5 from (40) and (42), we have

$$2\zeta_2D\kappa_2-\frac{1}{2}\zeta_2(\kappa_2-1)-2E\kappa_3+\frac{F}{\rho}+\frac{G}{\rho}-\frac{2H}{\rho^2}=0. \tag{45}$$

So far, we have already determined A, B, F, and G [see (28), (31), (34), and (43)]. Then, C, D, E, and H can first be solved by using (30), (33), (44), and (45). Then, they can be substituted into the following equations, which are from subtraction of (36) from (38) and (40) from (42),

$$\ln R\{-\zeta_1[A(\kappa_1+1)+B(\kappa_1-1)]+F(\kappa_3+1)+G(\kappa_3-1)\}$$
$$-4E+2C_5=0, \tag{46}$$
$$2\rho\zeta_2(2D-C_2)-4E\rho+\ln a[F(\kappa_3+1)+G(\kappa_3-1)]+2C_5=0, \tag{47}$$

to yield C_2 and C_5. In summary, the following solutions for the unknown constants are obtained:

$$A=F=-\frac{\rho}{2},\quad B=\frac{\rho(\kappa_1-1)}{2(\kappa_1+1)},\quad G=\frac{\rho(\kappa_3-1)}{2(\kappa_3+1)}, \tag{48}$$

$$C=\frac{\rho(1-\zeta_1-\kappa_3+\kappa_1)}{2(\kappa_1+1)(\zeta_1+\kappa_3)}-\frac{\rho^2}{4(\zeta_1+\kappa_3)}\left[\frac{2\rho(\zeta_1-1)(\zeta_2\kappa_2-\kappa_3)+(\zeta_1+\kappa_3)(\zeta_2\kappa_3-2\zeta_2\kappa_2+\zeta_2-2)}{\rho^2(\zeta_1-1)(\zeta_2\kappa_2-\kappa_3)-(\zeta_2\kappa_2+1)(\zeta_1+\kappa_3)}\right] \tag{49}$$

$$D=\frac{\rho(\zeta_1-1)}{2(\kappa_3+1)(\zeta_1+\kappa_3)}+\frac{(\kappa_3-1)}{4(\kappa_3+1)}-\frac{\rho^2}{4(1+\kappa_3)}\left(\frac{\zeta_1-1}{\zeta_1+\kappa_3}-\frac{1}{\rho^2}\right)\left[\frac{2\rho(\zeta_1-1)(\zeta_2\kappa_2-\kappa_3)+(\zeta_1+\kappa_3)(\zeta_2\kappa_3-2\zeta_2\kappa_2+\zeta_2-2)}{\rho^2(\zeta_1-1)(\zeta_2\kappa_2-\kappa_3)-(\zeta_2\kappa_2+1)(\zeta_1+\kappa_3)}\right] \tag{50}$$

$$E=\frac{\rho^2(\zeta_1-1)}{4(\kappa_3+1)(\zeta_1+\kappa_3)}\left[\frac{2}{\rho}-\frac{2\rho(\zeta_1-1)(\zeta_2\kappa_2-\kappa_3)+(\zeta_1+\kappa_3)(\zeta_2\kappa_3-2\zeta_2\kappa_2+\zeta_2-2)}{\rho^2(\zeta_1-1)(\zeta_2\kappa_2-\kappa_3)-(\zeta_2\kappa_2+1)(\zeta_1+\kappa_3)}\right] \tag{51}$$

$$H=-\frac{\rho^2}{4(1+\kappa_3)}\left[\frac{2\rho(\zeta_1-1)(\zeta_2\kappa_2-\kappa_3)+(\zeta_1+\kappa_3)(\zeta_2\kappa_3-2\zeta_2\kappa_2+\zeta_2-2)}{\rho^2(\zeta_1-1)(\zeta_2\kappa_2-\kappa_3)-(\zeta_2\kappa_2+1)(\zeta_1+\kappa_3)}\right] \tag{52}$$

$$C_2=\ln\left(\frac{R}{a}\right)\frac{\kappa_3}{\zeta_2(\kappa_3+1)}-\ln R\,\frac{\kappa_1\zeta_1}{\zeta_2(1+\kappa_1)}+\frac{2(\zeta_1-1)[1+\rho(\zeta_2-1)]+\zeta_2(\kappa_3-1)(\kappa_3+\zeta_1)}{2\zeta_2(\kappa_3+1)(\kappa_3+\zeta_1)}+2\left\{\frac{(\zeta_1-1)[1+\rho(\zeta_2-1)]}{\rho\zeta_2(\kappa_3+\zeta_1)}-\frac{1}{\rho^2}\right\}H \tag{53}$$

$$C_5=\rho\ln R\left(\frac{\kappa_3}{\kappa_3+1}-\frac{\kappa_1\zeta_1}{\kappa_1+1}\right)+\frac{\rho(\zeta_1-1)}{(\kappa_3+1)(\kappa_3+\zeta_1)}+2\left(\frac{\zeta_1-1}{\kappa_3+\zeta_1}\right)H. \tag{54}$$

The final stress and displacement expressions can be obtained by back substituting (48)–(54) into (12)–(26). Of particular interest are the stresses in the elastic plane, which are given by

$$\sigma_{\theta\theta}^{(1)}=f_0R\left\{L\frac{\rho}{\xi}+K\frac{\rho}{\xi^3}\right\}\cos\theta, \tag{55}$$

$$\sigma_{r\theta}^{(1)}=f_0R\left\{L\frac{\rho}{\xi}-K\frac{\rho}{\xi^3}\right\}\sin\theta, \tag{56}$$

$$\sigma_{rr}^{(1)}=-f_0R\left\{I\frac{\rho}{\xi}+K\frac{\rho}{\xi^3}\right\}\cos\theta, \tag{57}$$

Table 1 The identifications made and the value of constant *K* in (55)–(57) for the solutions of problems given in Fig. 2

Case	Identifications	K
I	$\zeta_2=0,\ \kappa_2=\overline{\kappa}$	$\frac{1-\zeta_1-\kappa_3+\kappa_1}{(\kappa_1+1)(\zeta_1+\kappa_3)}-\frac{\rho}{(\zeta_1+\kappa_3)}\times\left\{\frac{\rho(\zeta_1-1)+\zeta_1+\kappa_3}{\rho^2(\zeta_1-1)+\zeta_1+\kappa_3}\right\}$
II	$\rho=1,\ \zeta_2=1,\ \zeta_1=\zeta,\ \kappa_2=\kappa_3$	$\frac{1+\kappa_1-2(\zeta+\kappa_2)}{2(\kappa_1+1)(\zeta+\kappa_2)}$
III	$\rho=0,\ \zeta_1=\zeta,\ \zeta_2=1,\ \kappa_2=\kappa_3$	$\frac{1+\kappa_1-(\zeta+\kappa_2)}{(\kappa_1+1)(\zeta+\kappa_2)}$
IV	$\zeta=1$ in Case III, $\kappa_2=\kappa_1$	0
V	$\zeta_1=\zeta_2=0,\ \kappa_1=\kappa_2=\kappa$	$\frac{1+\overline{\kappa}-\kappa_3(1+\rho)}{\kappa_3(\overline{\kappa}+1)(1+\rho)}$
VI	$\zeta\to\infty$ in Case II or III	$-\frac{1}{(\kappa_1+1)}$

Note: $\overline{\kappa}$ equals 5/3 for plane stress and 1 for plane strain

where $\xi=r/R$ is the normalized r-coordinate and

$$L=\frac{(\kappa_1-1)}{2(\kappa_1+1)},\quad I=\frac{(\kappa_1+3)}{2(\kappa_1+1)}, \tag{58}$$

$$K=\frac{(1-\zeta_1-\kappa_3+\kappa_1)}{(\kappa_1+1)(\zeta_1+\kappa_3)}-\frac{\rho}{2(\zeta_1+\kappa_3)}\times\left[\frac{2\rho(\zeta_1-1)(\zeta_2\kappa_2-\kappa_3)+(\zeta_1+\kappa_3)(\zeta_2\kappa_3-2\zeta_2\kappa_2+\zeta_2-2)}{\rho^2(\zeta_1-1)(\zeta_2\kappa_2-\kappa_3)-(\zeta_2\kappa_2+1)(\zeta_1+\kappa_3)}\right]. \tag{59}$$

Before considering the stresses at the hole boundary, it is instructive to consider various special cases of the present solution. Table 1 tabulates the solutions for various problems shown in Fig. 2 which are special cases of the solution for the problem shown in Fig. 1.

Of more practical importance is the tensile stress concentration at the rivet hole in the infinite plane. By choosing appropriate stiffness and thickness of the reinforced material (i.e., Domain III in Fig. 1), it is hoped that the concentration of the tensile hoop stress can be alleviated (as we will show later that this is indeed the case). Fatigue crack initiations can therefore be avoided or, at least, deferred. Therefore, the stresses in Domain I on $r=R$ are evaluated next.

5 Numerical Results and Discussion

As remarked earlier, our main objective is to investigate what kind of reinforced material can be used to reduce the tensile hoop stress concentration at the rivet hole boundary on $r=R$. For illustrative purposes, only the plane-stress condition is considered in our numerical calculations given in Figs. 3–8. In addition, since the range of Poisson's ratio is relatively small comparing to the range of stiffness, we have assumed $\nu_1=\nu_2=\nu_3=0.25$ in all our calculations.

For steel structures, both the connecting plates and the rivets are made of steel. Therefore, we have set $\zeta_1=\zeta_2$ in the numerical calculations given in Figs. 3–5. Figure 3 plots the normalized hoop stress $\sigma_{\theta\theta}(\pi R/P)$ against the angular coordinate ϕ defined in Fig. 1 for various values of $\zeta(=\zeta_1=\zeta_2)$ with $\rho=0.25$. For $\zeta=5.5$ (i.e., the reinforced material is about five times stiffer than the plane and rivet), the tensile stress drops to becoming negligible for all values of ϕ between $-\pi/2$ and $\pi/2$. A more refined calculation yields an optimum value of 5.35 should be used. It means that we can choose a reinforced material such that the hoop stress can be significantly reduced. However, when ζ further increases, tensile hoop stress develops again but at the opposite side of the hole (i.e., at $\phi=-\pi/2$). Thus, an optimum value of stiffness ratio must be selected. This result provides a new theoretical basis for designing an efficient reinforced hole. Figures 4 and 5 show that the normalized radial stress $\sigma_{rr}(\pi R/P)$ and shear stress $\sigma_{r\theta}(\pi R/P)$ decreases slightly and increases with the increase in ζ, respectively. But, the increases in the magnitude of the normal-

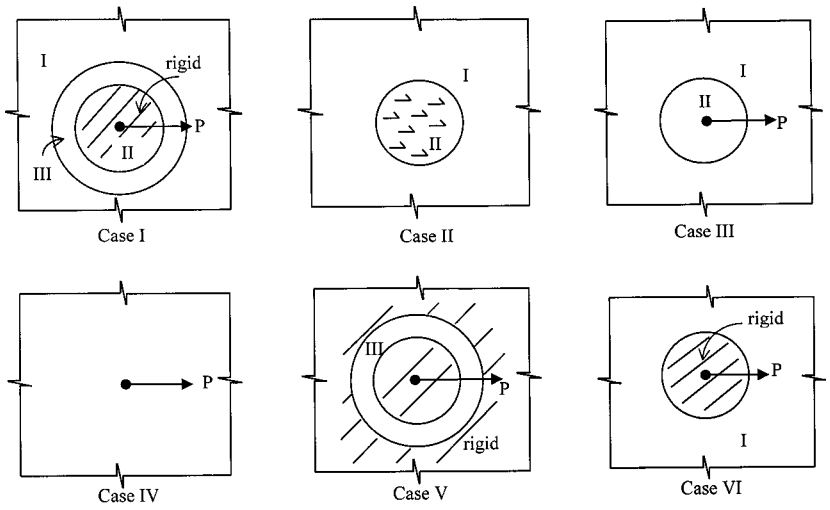


Fig. 2 Six special cases of the problem showed in Fig. 1

ized shear stress is about 23 percent for ζ increasing from 2.5 to 5.5 which is considered insignificant comparing to the drastic drop in the hoop stress shown in Fig. 3.

As considered by Ho and Chau ([23,41]), different materials may also be used for the circular inclusion and the plane. Thus, ζ_1/ζ_2 is allowed to vary in Fig. 6 for $\zeta_1=2.5$ and $\rho=0.25$. In particular, Fig. 6 plots the normalized hoop stress $\sigma_{\theta\theta}(\pi R/P)$ against ϕ for various values of ζ_1/ζ_2. The hoop stress in general drops with the increase of ζ_1/ζ_2, but the drop is relatively small comparing to the effect of reinforced material shown in Fig. 3. In

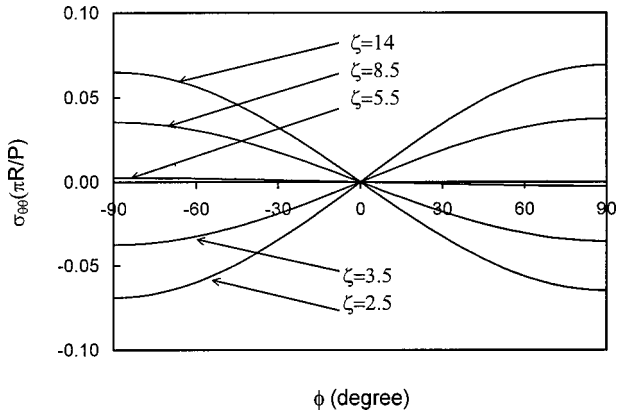


Fig. 3 The normalized hoop stress $\sigma_{\theta\theta}(\pi R/P)$ against ϕ on $r=R$ for various stiffness ratio $\zeta_1=\zeta_2=\zeta$ under the plane stress condition with $\nu_1=\nu_2=\nu_3=0.25$, and $\rho=0.25$

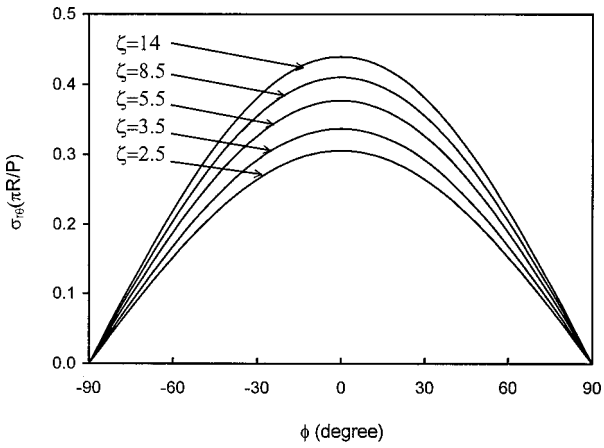


Fig. 5 The normalized shear stress $\sigma_{r\theta}(\pi R/P)$ against ϕ on $r=R$ for various stiffness ratio $\zeta_1=\zeta_2=\zeta$ under the plane stress condition with $\nu_1=\nu_2=\nu_3=0.25$, and $\rho=0.25$

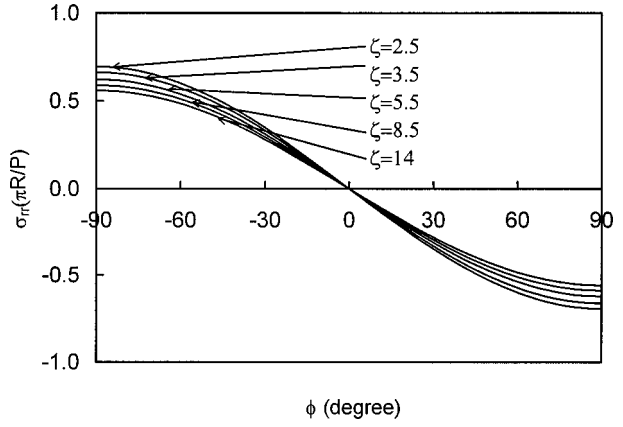


Fig. 4 The normalized normal stress $\sigma_{rr}(\pi R/P)$ against ϕ on $r=R$ for various stiffness ratio $\zeta_1=\zeta_2=\zeta$ under the plane stress condition with $\nu_1=\nu_2=\nu_3=0.25$, and $\rho=0.25$

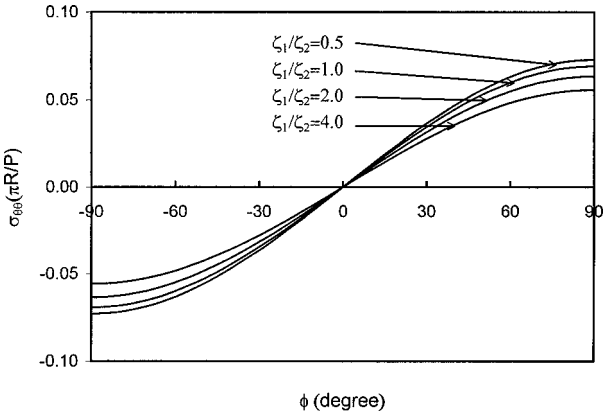


Fig. 6 The normalized hoop stress $\sigma_{\theta\theta}(\pi R/P)$ against ϕ on $r=R$ for various ζ_1/ζ_2 with $\nu_1=\nu_2=\nu_3=0.25$, $\rho=0.25$ and $\zeta_1=2.5$

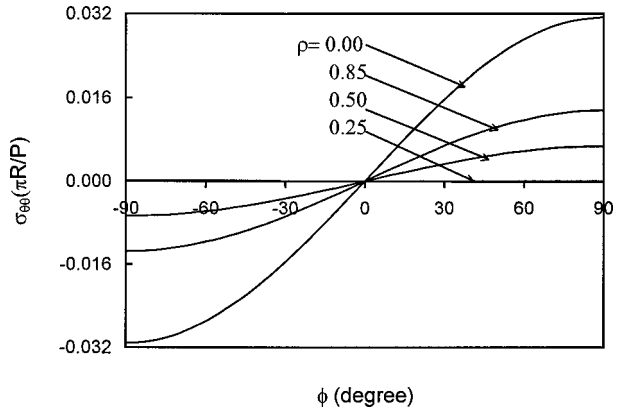


Fig. 7 The normalized hoop stress $\sigma_{\theta\theta}(\pi R/P)$ against ϕ on $r=R$ for various ρ with $\nu_1=\nu_2=\nu_3=0.25$, and $\zeta_1=\zeta_2=5.35$

addition, we have also plotted the radial and shear stresses versus ζ_1/ζ_2. The shapes of these curves are similar to those given in Figs. 4–5, and thus they will not be given here. We found that the radial stress is basically independent of the value of ζ_1/ζ_2 while the normalized shear stress increases only slightly with ζ_1/ζ_2.

In addition to the choice of material for the ring, we can also select the thickness of the ring (Domain III in Fig. 1). Figure 7 plots the normalized hoop stress against ϕ for various values of $\rho=(a/R)^2$ on $r=R$ for $\zeta_1=\zeta_2=\zeta=5.35$ (the optimum stiffness ratio leading to the minimum hoop stress). It appears that the tensile hoop stress does not decrease monotonically with ρ; thus, it is possible to find an optimum thickness of the ring to reduce the hoop stress. Again, although the details are not given here, we found that on the hole boundary both the radial and shear stresses are relatively insensitive to the value of ρ.

To further investigate the optimum thickness, Fig. 8 plots the maximum normalized hoop stress $\sigma_{\theta\theta}(\pi R/P)$ against $\sqrt{\rho}$ for various values of ζ. For small ζ (say $\zeta=2.5$, 3.5), a thinner reinforced ring always leads to a smaller hoop stress concentration. However, for larger ζ (say $\zeta=5.35$, 8.4), local minimums of the hoop stress can be found for some optimum values of $\sqrt{\rho}$ (or the optimum thickness). Therefore, this suggests another opportunity to further reduce the tensile hoop stress at the hole boundary by choosing an optimum thickness of the reinforced ring for a specific ring material (i.e., for a specific value of ζ).

6 Conclusion

We have derived analytically the stresses induced in an infinite elastic plane containing circular inclusion (or rivet) bonded perfectly to a reinforced hole in the plane. The inclusion is assumed being bonded perfectly to the elastic plane through a reinforced ring. The circular inclusion, the reinforced ring and the plane can be made of different materials. In contrast to most of the previous studies, the rivet force is applied through a uniformly distributed body force on the inclusion instead of assuming a point force at the center of the inclusion. For the stiffness ratio between the reinforced material and the plane equal to or larger than one ($\zeta \geqslant 1$), the tensile hoop stress at the hole boundary increases initially with ζ, achieves a minimum value (close to zero) at an optimum value of ζ, and increases again with further increase in ζ. Whereas, both the shear and radial stresses are relatively insensitive to the changing values of ζ. Thus, the tensile hoop stress can be reduced to becoming negligible (very close to zero) if an appropriate stiffness is assigned to the reinforced ring, which is stiffer than the plane. We also found that the tensile hoop stress in general decreases with the increase in stiffness of the inclusion. In addition, for a fixed stiffness ratio ζ an optimum thickness of the ring can be determined such that the magnitude of the tensile hoop stress is further reduced. In short, we have illustrated that a new opportunity exists to reduce tensile hoop stress at the hole boundary of the elastic plane by choosing the appropriate material for the reinforcement and the appropriate thickness of the reinforced ring.

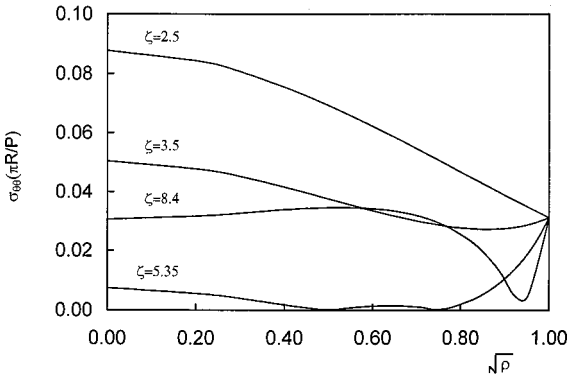


Fig. 8 The maximum normalized hoop stress $\sigma_{\theta\theta}(\pi R/P)$ against $\sqrt{\rho}$ on $r=R$ for various values of ζ with $\nu_1=\nu_2=\nu_3=0.25$

Acknowledgment

This research was supported by funding from the Research Grants Council (RGC) and the Hong Kong Polytechnic University (HKPolyU). Some conceptual changes to an earlier draft of the paper were proposed by Professor J. R. Barber and an anonymous reviewer.

References

[1] Mura, T., 1987, *Micromechanics of Defects in Solids*, 2nd and rev. ed., Martinus Nijhoff Dordrecht.
[2] Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd ed., McGraw-Hill, New York.
[3] Peterson, R. E., 1974, *Stress Concentration Factors*, John Wiley and Sons, New York.
[4] Coker, E. G., 1925, ''Photo-elasticity,'' J. Franklin Inst., **199**, pp. 289–331.
[5] Brickley, W. G., 1928, ''The Distribution of Stress Round a Circular Hole in a Plate,'' Philos. Trans. R. Soc. London, Ser. A, **A227**, pp. 383–415.
[6] Knight, R. C., 1935, ''The Action of a Rivet in a Plate of Finite Breadth,'' Philos. Mag. Ser. 7, **19**, pp. 517–540.
[7] Atsumi, A., 1952, ''On the Stress Distributions in a Strip With a Circular Hole Fitted by a Rivet,'' Trans. Jpn. Soc. Mech. Eng., **18**, No. 68, pp. 28–40 (in Japanese with English abstract).
[8] Lambert, T. H., and Brailey, R. J., 1962, ''The Influence of the Coefficient of Friction on the Elastic Stress Concentration Factor for a Pin-Jointed Connection,'' Aeronaut. Q., **13**, No. 2, pp. 17–29.
[9] Sharfuddin, S. M., 1967, ''Loaded Loose-Fitted Rough Circular Rigid Pin in a Circular Hole,'' Acta Mech., **3**, pp. 376–384.
[10] de Jong, Th., 1977, ''Stress Around Pin-Loaded Holes in Elastically Orthotropic or Isotropic Plates,'' J. Compos. Mater., **11**, pp. 313–331.
[11] de Jong, Th., 1982, ''The Influence of Friction on the Theoretical Strength of Pin-Loaded Holes Orthotropic Plates,'' *Report LR-350, Department of Aerospace Engineering, Delft University of Technology*, pp. 1–36.
[12] Eshwar, V. A., 1978, ''Analysis of Clearance Fit Pin Joints,'' Int. J. Mech. Sci., **20**, pp. 477–484.
[13] Zhang, K. D., and Ueng, C. E. S., 1984, ''Stresses Around a Pin-Loaded Hole in Orthotropic Plates,'' J. Compos. Mater., **18**, pp. 432–446.
[14] Hyer, M. W., and Klang, E. C., 1985, ''Contact Stresses in Pin-Loaded Orthotropic Plates,'' Int. J. Solids Struct., **21**, pp. 957–975.
[15] Madenci, E., Ileri, L., and Starnes, Jr., J. H., 1995, ''Analysis of Pin-Loaded Holes in Composite Laminates Under Combined Bearing-Bypass and Shear Loading,'' Int. J. Solids Struct., **32**, pp. 2053–2062.
[16] Greenwood, J. A., 1989, ''Exact Formulas for Stresses Around Circular Holes and Inclusions,'' Int. J. Mech. Sci., **31**, pp. 219–227.
[17] Tiffen, R., and Sharfuddin, S. M., 1964, ''A Mixed Boundary Value Problem of Two-Dimensional Elasticity Theory,'' Mathematika, **11**, pp. 151–154.
[18] Mizushima, I., and Hamada, M., 1983, ''Stress Analysis Around Circular Hole in Infinite Plate With Rigid Disk (Case of Load Applied to Disk),'' Bull. JSME, **26**, pp. 1296–1301.
[19] Hetenyi, M., and Dundurs, J., 1962, ''The Elastic Plane With a Circular Insert, Loaded by a Tangentially Directed Force,'' ASME J. Appl. Mech., **29**, pp. 362–368.

[20] Dundurs, J., 1963, ''Concentrated Force in an Elastically Embedded Disk,'' ASME J. Appl. Mech., **30**, pp. 568–570.
[21] Fukui, T., 1990, ''The Stresses of an Infinite Plane With an Elastic Circular Insert Loaded by a Concentrated Force (The Case Having a Crevice)'', Trans. Jpn. Soc. Mech. Eng., **56**, No. 527, 1009–7, pp. 133–136.
[22] Karasudhi, P., 1991, *Foundations of Solid Mechanics*, Kluwer Academic Publishers, Dordrecht.
[23] Ho, K. C., and Chau, K. T., 1997, ''An Infinite Plane Loaded by a Rivet of a Different Material,'' Int. J. Solids Struct., **34**, pp. 2477–2496.
[24] Stippes, M., Wilson, H. B., Jr., and Krull, F. N., 1962, ''A Contact Stress Problem for a Smooth Disk in an Infinite Plate,'' *Proceedings of the 4th U.S. National Congress of Applied Mechanics*, pp. 799–807.
[25] Wilson, H. B. Jr., 1964, ''Approximation Determination of Contact Stresses in an Infinite Plate With a Smooth Circular Insert,'' *Proceedings of the 2nd Southeastern Conference on Theoretical and Applied Mechanics*, Oxford, Pergamon, Oxford, pp. 147–159.
[26] Wilson, Jr., H. B., 1964b, ''Contact Stresses in an Infinite Plate Containing a Smooth Rigid Elliptic Insert,'' ASME J. Appl. Mech., **31**, pp. 703–704.
[27] Keer, L. M., Dundurs, J., and Kiattikomol, K., 1973, ''Separation of a Smooth Circular Inclusion From a Matrix,'' Int. J. Eng. Sci., **11**, pp. 1221–1233.
[28] Noble, B., and Hussain, M. A., 1967, ''Angle of Contact for Smooth Elastic Inclusions,'' *Proceedings of the 10th Midwestern Mechanics Conference*, pp. 459–476.
[29] Noble, B., and Hussain, M. A., 1969, ''Exact Solution of Certain Dual Series for Indentation and Inclusion Problems,'' Int. J. Eng. Sci., **7**, pp. 1149–1161.
[30] Hussain, M. A., and Pu, S. L., 1971, ''Slip Phenomenon for a Circular Inclusion,'' ASME J. Appl. Mech., **38**, pp. 627–633.
[31] Sheremetev, M. P., 1961, ''The Solution of the Equations of Certain Contact Problems of the Theory of Elasticity (Equations of the Prandtl type),'' *Problems of Continuum Mechanics* J. R. M. Rodak ed., SIAM, Philadelphia, PA, pp. 419–439.
[32] Beskin, L., 1944, ''Strengthening of Circular Holes in Plates Under Edge Loads,'' Transactions of ASME Applied Mechanics Section, **66**, pp. A140–A148.
[33] Levy, S., and Smith, F. C., 1949, ''Stress Distribution Near Reinforced Circular Hole Loaded by Pin,'' J. Res. Natl. Bur. Stand., **42**, pp. 397–404.
[34] Wells, A. A., 1950, ''On the Plane Stress-Distribution in an Infinite Plane With a Rim-Stiffened Elliptical Opening,'' Q. J. Mech. Appl. Math., **3**, pp. 23–31.
[35] Mansfield, E. H., 1953, ''Neutral Holes in Plane Sheet-Reinforced Holes Which are Elastically Equivalent to the Uncut Sheet,'' Q. J. Mech. Appl. Math., **6**, pp. 370–378.
[36] Hicks, R., 1957, ''Reinforced Elliptical Holes in Stressed Plates,'' Journal of the Royal Aeronautical Society, **61**, pp. 688–693.
[37] Savin, G. N., 1961, *Stress Concentration Around Holes*, (translated to English by E. Gros), Pergamon, Oxford, UK.
[38] Davies, G. A. O., 1967, ''Plate-Reinforced Holes,'' Aeronaut. Q., **18**, pp. 43–54.
[39] Timoshenko, S. P., 1924, ''On Stresses in a Plate With a Circular Hole,'' J. Franklin Inst., **197**, pp. 505–516.
[40] Houghton, D. S., and Rothwell, A., 1961, ''The Analysis of Reinforced Circular and Elliptical Cutouts Under Various Loading Conditions,'' *Report No. 151 The College of Aeronautics Cranfield*. pp. 1–10.
[41] Ho, K. C., and Chau, K. T., 1999, ''A Finite Strip Loaded by a Bonded-Rivet of a Different Material,'' Comput. Struct., **70**, No. 2, pp. 203–218.
[42] Kienzler, R., and Duan, Z., 1987, ''On the Distribution of Hoop Stresses Around Circular Holes in Elastic Sheets,'' ASME J. Appl. Mech., **54**, pp. 110–114.

On Bridgman's Stress Solution for a Tensile Neck Applied to Axisymmetrical Blunt Notched Tension Bars

A. Valiente
Departamento de Ciencia de Materiales,
Universidad Politécnica de Madrid,
E.T.S. Ingenieros de Caminos,
Ciudad Universitaria s/n,
28040 Madrid, Spain

The displacement field at the minimum cross section of an axisymmetrical notched tensile bar is analytically related to the notch root radius for large geometry changes of the notch profile. This relationship is used to complete Bridgman's formula for the tensile load of a tensile necking in order to predict the entire load-minimum diameter curve of an axisymmetrical blunt notched bar under tension. A particular case is solved by a finite element modeling to check the theoretical results derived at different stages of the analysis. [DOI: 10.1115/1.1360689]

1 Introduction

The stress and strain distributions derived by Bridgman [1,2] for the neck of a tensile round bar are found quite often in books and treatises on plasticity. This is largely due to the interest of the subject, but in part it is also due to the intrinsic value of the analysis since few problems of plasticity can offer a solution as elegant as Bridgman's. It should be noted that in spite of the fact that the neck development is a nonlinear problem owing to the plastic behavior of the material and the large geometry changes, Bridgman arrived at simple formulas, analytically derived, for the stress distribution across the minimum cross section. This distribution is given as a function of the tensile load acting on the bar, the neck diameter, and the radius of curvature of the neck profile at its root.

Bridgman's work on tensile necking has given rise to several contributions on the same subject. Soon after Bridgman, Davidenkov and Spiridonova [3] proposed a similar analysis with quantitative differences in some of the assumptions. Subsequently, Kaplan [4] extended Bridgman's analysis beyond the minimum cross section and was able to predict the form of the neck profile from the same parameters as those of Bridgman. Jones, Gillis, and Shalaby [5] used the method of Kaplan to obtain a more complete solution, and Eisenberg and Yen [6] generalized Bridgman's analysis for orthotropic bars with the longitudinal axis parallel to two planes of orthotropy.

Several interesting consequences can be drawn from Bridgman's results, those referring to the triaxiality of the stress state at the neck being worthy of mention: His solution ascertains that the stress state is no longer uniaxial when the neck begins to form; it shows that the triaxiality is maximum at the axis of the bar; and it predicts the corresponding value as a function of the main variables (the tensile load, the neck diameter, and the radius of curvature of the neck). Perhaps the idea of producing controlled triaxial stress states by means of artificial necks was suggested by these conclusions, and hence that of assimilating a circumferential notch machined on a cylindrical tension specimen to the tensile necking of a round bar, once general yielding has taken place in the notched specimen.

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. 15, 1998; final revision, Nov. 29, 2000. Associate Editor: K. T. Ramesh. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

As a consequence, a considerable amount of investigation on ductile steels has been conducted using circumferentially notched cylindrical tension specimens in conjunction with Bridgman's formulas, and research works have been reported relating the stress state to the fracture mechanisms ([7]), determining experimentally the influence of the stress triaxiality on the fracture strain ([8]), and predicting the upper shelf static fracture toughness ([9]). In recent years, testing with circumferentially notched cylindrical tension specimens has become the usual method of characterizing the local approach theories. Often Bridgman's solution is used as a reference in these tests ([10]), although sometimes it is complemented with finite element calculations ([11,12]).

Bridgman's solution was derived for a necking with a given profile, so it cannot be applied to notched specimens that have experienced large geometry changes unless the shape of the deformed notch be known. This paper shows that the notch geometry changes suitable for the application of Bridgman's formulas can be predicted from the first hypothesis assumed by Bridgman. According to this hypothesis, which is based on experimental evidence ingeniously obtained, the axial strain rate is uniform at the minimum cross section of the necking. The prediction of the geometry changes requires an equation previously derived in the paper which determines the notch root radius (the curvature radius of the notch profile at the notch root) from the displacement field of the minimum cross section. When this equation is particularized for the displacement field corresponding to a uniform axial strain rate, a biunivocal relation results between the notch root radius and the radius of the minimum cross section (minimum radius).

Once the notch root radius is given as a function of the minimum radius, Bridgman's formula for the load that produces a given necking can be applied to a notched specimen so that the load sustained by the specimen becomes dependent on only one kinematical variable, the minimum radius. A second hypothesis assumed by Bridgman is involved in his formula and concerns the isostatic lines in the axial plane of the necking, which were approximated by circumferences in the neighborhood of the minimum cross section. When Bridgman's formula is employed in that way, it transforms into a theoretical prediction of the load-minimum radius curve of a notched specimen, whose maximum, the plastic instability load, becomes also a theoretical prediction.

All these theoretical predictions are developed in the paper and are used for the assessment of Bridgman's solution as an approximation to describe the plastic behavior of tensile round specimens with circumferential blunt notches. The assessment is based on the comparison of the theoretical predictions with the numerical re-

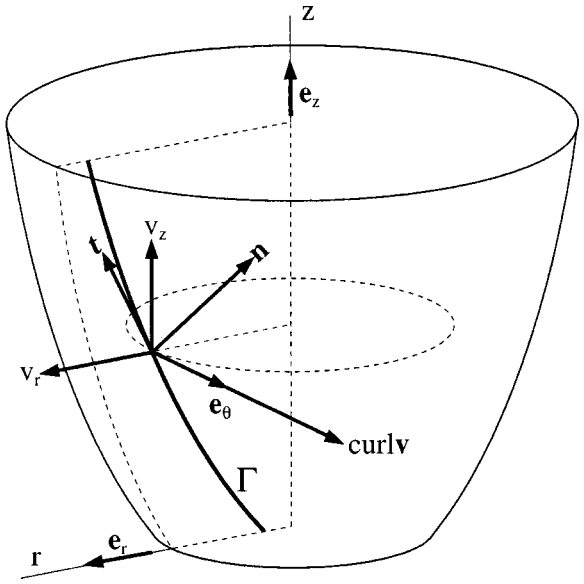


Fig. 1 Axisymmetrical motion of a continuous medium: material line coplanar with the axis

sults obtained in a finite element modeling of the tensile loading process of a circumferentially notched specimen.

2 Evolution of the Notch Root Radius

In a tensile round bar of isotropic and homogeneous material, the necking development, and hence the deformation of a circumferential notch, consists of geometry changes maintaining the axial symmetry about the longitudinal axis of the bar, as well as the mirror symmetry to the plane of the minimum cross section area. No other condition is considered in the analysis, but the classical theory of plasticity (the Von Mises yield criterion with isotropic hardening and the Prandtl-Reuss equations) is applied. The theoretical basis of an equation for the evolution of the notch profile radius is the condition of the notch profile of being simultaneously a traction-free contour and a material line. The latter means that it is formed by the same material points at all time, which is a consequence of being the intersection of a material surface (the bar surface) with a plane (an axial plane) that the material points can not leave, due to the axisymmetry.

The starting point is the kinematical equation derived in the Appendix for the time derivative of the curvature of the material lines of a continuous medium in motion. In this equation, the time rate $\dot{\kappa}$ of the curvature at a given point of a material line is expressed as a function of the strain rate tensor field, the velocity field, and the curvature κ and the torsion τ of the material line at the given point

$$\dot{\kappa}=\frac{1}{2}\frac{\partial\dot{\gamma}_{tn}}{\partial s}-\frac{1}{2}\tau\dot{\gamma}_{tb}+\frac{1}{2}\frac{\partial(\mathrm{curl}\,\mathbf{v})}{\partial s}\cdot\mathbf{b}-\dot{\varepsilon}_t\kappa. \tag{1}$$

The amounts involved in Eq. (1) are the components $\dot{\varepsilon}_t$, $\dot{\gamma}_{tn}$, and $\dot{\gamma}_{tb}$ of the strain rate tensor (normal and shear strain rates) in the tangent (**t**), principal normal (**n**) and binormal (**b**) directions of the line, the directional derivatives of these components in the tangent direction of the line ($\partial/\partial s$), and the curl of the velocity field **v**. For a material line Γ coplanar with the axis of revolution in an axisymmetrical motion of the continuous medium (Fig. 1), Eq. (1) can be simplified. First, since the material line remains plane, its torsion τ vanishes. Further, if the axis of revolution is chosen as the *z*-axis of a cylindrical coordinate system r,θ,z and the principal normal is the tangent rotated 90 degrees counterclockwise (Fig. 1), the binormal **b** coincides with the negative of the unit base vector $\mathbf{e}_\theta$. Finally, the axisymmetry of the velocity field **v** implies only radial and axial components independent of θ, v_r, and v_z, which leads to the following expression of curl **v**:

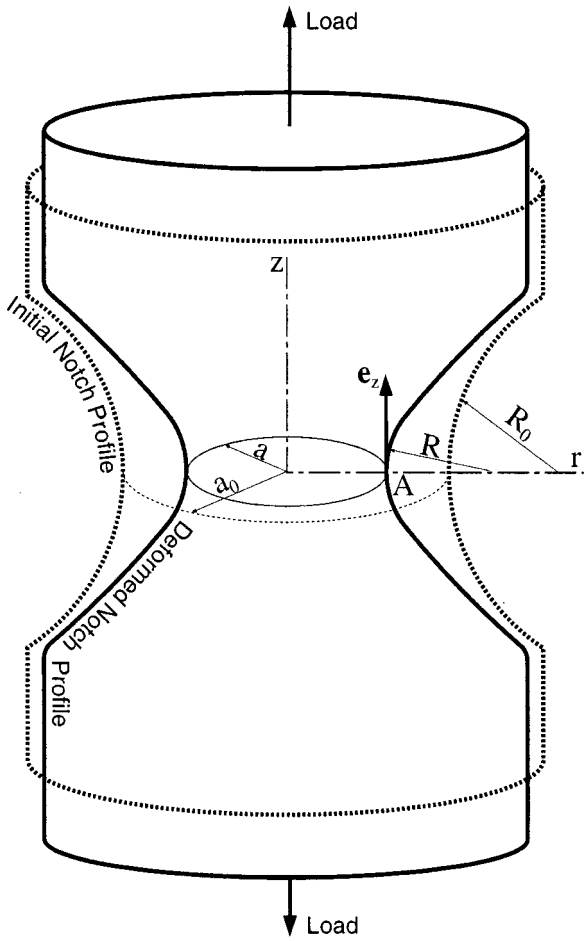


Fig. 2 Deformed and undeformed notch profiles

$$\mathrm{curl}\mathbf{v}=2\omega\mathbf{e}_\theta \quad \text{with} \quad \omega=\frac{1}{2}\left(\frac{\partial v_r}{\partial z}-\frac{\partial v_z}{\partial r}\right). \tag{2}$$

By virtue of these particular conditions, the directional derivative of curl **v** along Γ may be put in the form

$$\frac{1}{2}\frac{\partial(\mathrm{curl}\,\mathbf{v})}{\partial s}\cdot\mathbf{b}=\frac{\partial(\omega\mathbf{e}_\theta)}{\partial s}\cdot(-\mathbf{e}_\theta)=-\frac{\partial\omega}{\partial s}=-\mathrm{grad}\,\omega\cdot\mathbf{t} \tag{3}$$

and Eq. (1) becomes

$$\dot{\kappa}=\frac{1}{2}\frac{\partial\dot{\gamma}_{tn}}{\partial s}-\mathrm{grad}\,\omega\cdot\mathbf{t}-\dot{\varepsilon}_t\kappa. \tag{4}$$

Furthermore, the gradient of ω may be expressed as a function of the strain rate field by applying the formulas for this operator in cylindrical coordinates and then eliminating the derivatives of the velocity components through the expressions which relate the strain rates to the velocity field, also in cylindrical coordinates. Denoting the radial and axial unit vectors of the cylindrical coordinates by $\mathbf{e}_r$ and $\mathbf{e}_z$, the result is

$$\mathrm{grad}\,\omega=\left(\frac{\partial\dot{\varepsilon}_r}{\partial z}-\frac{1}{2}\frac{\partial\dot{\gamma}_{rz}}{\partial r}\right)\mathbf{e}_r+\left(\frac{1}{2}\frac{\partial\dot{\gamma}_{rz}}{\partial z}-\frac{\partial\dot{\varepsilon}_z}{\partial r}\right)\mathbf{e}_z. \tag{5}$$

On applying Eq. (4) to the deformed notch profile Γ (Fig. 2), the shear strain rate $\dot{\gamma}_{tn}$ remains null along Γ, and as a consequence its derivative in the tangent direction vanishes. This is due to the absence of shear stresses acting on the notch profile, which im-

plies the absence of the corresponding shear strain rates according to the Prandtl-Reuss equations. Then for the notch profile Eq. (4) becomes

$$\dot{\kappa} = -\operatorname{grad}\omega\cdot\mathbf{t} - \dot{\varepsilon}_t\kappa. \tag{6}$$

At the root of the notch (point A of Fig. 2), where the value of the curvature κ is denoted by $c(c=-1/R)$ and its initial value by $c_0(c_0=-1/R_0)$, the direction of the tangent to Γ coincides with that of the axis of axisymmetry, so the vector $\mathbf{t}$ and the strain rate $\dot{\varepsilon}_t$ coincide with $\mathbf{e}_z$ and $\dot{\varepsilon}_z$, respectively. Then, substitution of these particular conditions besides Eq. (5) into Eq. (6) yields

$$\dot{c} = \frac{1}{2}\frac{\partial\dot{\gamma}_{rz}}{\partial z} + \frac{\partial\dot{\varepsilon}_z}{\partial r} - \dot{\varepsilon}_z c. \tag{7}$$

A form of this equation involving still fewer variables can be obtained. According to the axisymmetry, a principal direction of the strain rate tensor is $\mathbf{e}_\theta$, and the other two, $\mathbf{e}_{\mathrm{I}}$ and $\mathbf{e}_{\mathrm{II}}$, are contained in the axial plane. By applying the Mohr's circle procedure, the strain rate $\dot{\gamma}_{rz}$ can be expressed as a function of the strain rates $\dot{\varepsilon}_{\mathrm{I}}$, $\dot{\varepsilon}_{\mathrm{II}}$ in the directions $\mathbf{e}_{\mathrm{I}}$ and $\mathbf{e}_{\mathrm{II}}$ and of the angle φ between the radial direction and $\mathbf{e}_{\mathrm{I}}$, so that

$$\dot{\gamma}_{rz} = (\dot{\varepsilon}_{\mathrm{I}} - \dot{\varepsilon}_{\mathrm{II}})\operatorname{sin} 2\varphi \tag{8}$$

$$\frac{\partial\dot{\gamma}_{rz}}{\partial z} = \frac{\partial(\dot{\varepsilon}_{\mathrm{I}} - \dot{\varepsilon}_{\mathrm{II}})}{\partial z}\operatorname{sen} 2\varphi + 2(\dot{\varepsilon}_{\mathrm{I}} - \dot{\varepsilon}_{\mathrm{II}})\frac{\partial\varphi}{\partial z}\cos 2\varphi. \tag{9}$$

Since at the root of the notch the axial and radial directions are principal directions of stress and strain rate, the conditions to be particularized in Eq. (4) for point A are

$$\varphi = \frac{\pi}{2}\quad \frac{\partial\varphi}{\partial z} = c\quad \dot{\varepsilon}_{\mathrm{I}} = \dot{\varepsilon}_z\quad \dot{\varepsilon}_{\mathrm{II}} = \dot{\varepsilon}_r.$$

Thus, for the root of the notch, Eq. (9) and subsequently Eq. (7) become

$$\frac{\partial\dot{\gamma}_{rz}}{\partial z} = -2(\dot{\varepsilon}_z - \dot{\varepsilon}_r)c \tag{10}$$

$$\dot{c} = -\dot{\varepsilon}_r c + \frac{\partial\dot{\varepsilon}_z}{\partial r}. \tag{11}$$

An undeformed elemental length dr_0 lying on the r-axis transforms into dr after deformation and remains on the r-axis since this axis moves along itself due to the axial and mirror symmetry. Let λ be the value of the ratio dr/dr_0 at point A. Hence, at this point

$$\dot{\varepsilon}_r = \frac{\dot{\lambda}}{\lambda}\quad \frac{\partial\dot{\varepsilon}_z}{\partial r} = \frac{1}{\lambda}\frac{\partial\dot{\varepsilon}_z}{\partial r_0} = \frac{1}{\lambda}\left(\frac{\partial\varepsilon_z}{\partial r_0}\right)^{\cdot} \tag{12}$$

and Eq. (11) transforms into

$$\dot{c} = -\frac{\dot{\lambda}}{\lambda}c + \frac{1}{\lambda}\left(\frac{\partial\varepsilon_z}{\partial r_0}\right)^{\cdot} \tag{13}$$

which, after some rearrangement, can be integrated to yield

$$\lambda c - \frac{\partial\varepsilon_z}{\partial r_0} = cte = c_0 \tag{14}$$

where the constant has been determined from the undeformed specimen, for which the curvature of the notch is c_0, λ is equal to unity and ε_z vanishes. In terms of the notch root radius R and its initial value R_0 ($c=-1/R$, $c_0=-1/R_0$, Fig. 2):

$$\frac{R_0}{R} = \frac{1}{\lambda} - \frac{R_0}{\lambda}\frac{\partial\varepsilon_z}{\partial r_0} \tag{15}$$

If the elastic compressibility is neglected, incompressibility follows, and for the points of the r-axis it takes the form

$$1 = e^{\varepsilon_z}\frac{r}{r_0}\frac{dr}{dr_0} = e^{\varepsilon_z}\frac{r}{r_0}\Lambda \Rightarrow \varepsilon_z = -\ln\left(\Lambda\frac{r}{r_0}\right) \tag{16}$$

where $\Lambda=\Lambda(r_0)$ is the radial stretch ratio dr/dr_0 along this axis. Then

$$\frac{\partial\varepsilon_z}{\partial r_0} = -\frac{1}{\Lambda}\frac{d\Lambda}{dr_0} - \frac{1}{r}\frac{dr}{dr_0} + \frac{1}{r_0} = -\frac{1}{\Lambda}\frac{d\Lambda}{dr_0} - \frac{\Lambda}{r} + \frac{1}{r_0} \tag{17}$$

and at point A

$$\frac{\partial\varepsilon_z}{\partial r_0} = -\frac{\lambda'}{\lambda} - \frac{\lambda}{a} + \frac{1}{a_0}, \tag{18}$$

λ' being the value of the derivative $d\Lambda/dr_0$ at this point, namely, at $r_0=a_0$. Thus, Eq. (15) becomes

$$\frac{R_0}{R} = \frac{1}{\lambda} + \frac{R_0\lambda'}{\lambda^2} + \frac{R_0}{a_0}\left(\frac{a_0}{a} - \frac{1}{\lambda}\right). \tag{19}$$

Therefore, the notch root radius R is determined by three values dependent on the displacement field of the minimum cross section of the notch: the radius a of the section and the values λ and λ' taken at the notch root by the radial stretch ratio and its derivative along the r-axis. Equation (19) is a kinematical relationship that allows the notch root radius to be found from the single function of the distance to the center of the cross section, to which that displacement field reduces.

3 Application to Bridgman's Solution

The first specific hypothesis of Bridgman's analysis determines the displacement field at the minimum cross section of the tensile neck. Therefore, when Bridgman's solution is applied to a tensile notched specimen, the result of particularizing Eq. (19) for this displacement field is being assumed as valid for the specimen and it provides a means to check such application of Bridgman's solution.

According to experimental evidence, Bridgman assumed the axial strain rate $\dot{\varepsilon}_z$ to be uniform across the minimum cross section, which leads to the uniformity of the overall strain rate and strain states, including the standard equivalent plastic strain $\overline{\varepsilon_p}$. Indeed, if $\dot{\varepsilon}_z$ is not dependent on r, from the conditions of axisymmetry and incompressibility it follows that

$$\dot{\varepsilon}_r = \dot{\varepsilon}_\theta = -\frac{1}{2}\dot{\varepsilon}_z = -\frac{1}{2}\dot{\overline{\varepsilon}}_p = \frac{\dot{a}}{a} \tag{20}$$

$$\varepsilon_r = \varepsilon_\theta = -\frac{1}{2}\varepsilon_z = -\frac{1}{2}\overline{\varepsilon_p} = \ln\frac{a}{a_0} \tag{21}$$

$$\frac{r}{r_0} = e^{\varepsilon_\theta} = \frac{a}{a_0}\quad \Lambda(r_0) = \frac{dr}{dr_0} = \frac{a}{a_0} \tag{22}$$

$$\lambda = \Lambda(a_0) = \frac{a}{a_0}\quad \lambda' = \left.\frac{d\Lambda}{dr_0}\right|_{a_0} = 0. \tag{23}$$

Particularizing these values of λ and λ' in Eq. (19), it simplifies to

$$\frac{R}{R_0} = \frac{a}{a_0}, \tag{24}$$

i.e., the notch root radius and the minimum specimen radius vary proportionally during the deformation process. This must be satisfied by the tensile notched specimens for which Bridgman's solution is applicable.

The second specific hypothesis of Bridgman's analysis was assumed in the development of the formula that gives the tensile load as a function of the neck geometry and the stress-strain curve of the material. The development of the formula is next summarized as an introduction to the hypothesis.

The stresses σ_r, σ_θ, and σ_z are the principal ones at the minimum cross section due to the axial and mirror symmetry, and accordingly the r-axis is an isostatic line of the plane rz. These conditions give rise to a particular equilibrium equation along the r-axis that can be obtained from the general equilibrium equation along the isostatic lines of the axial planes for axisymmetry. Let σ_{I} and σ_{II} be the principal stresses contained in the axial plane, s_{I} the distance along the isostatic line tangent to σ_{I}, R_{I} the curvature radius along the isostatic line tangent to σ_{II}, and φ the angle between the direction of σ_{I} and the r-axis. The general equation is

$$\frac{\partial \sigma_{\mathrm{I}}}{\partial s_{\mathrm{I}}} - \frac{\sigma_{\mathrm{I}} - \sigma_{\mathrm{II}}}{R_{\mathrm{II}}} + \frac{\sigma_{\mathrm{I}} - \theta_\theta}{r}\cos\varphi = 0, \tag{25}$$

and for the r-axis

$$\frac{d\sigma_r}{dr} - \frac{\sigma_r - \sigma_z}{\rho} + \frac{\sigma_r - \sigma_\theta}{r} = 0, \tag{26}$$

ρ being the curvature radius of the isostatic lines normal to the r-axis at the points of intersection with it. Furthermore, the equality of the stresses σ_r and σ_θ arises from that of the corresponding strain rates, and as a consequence the difference $\sigma_z - \sigma_r$ is the equivalent stress $\overline{\sigma}$, which remains uniform in all the section due to isotropic hardening (such a property implies a biunivocal relation $\overline{\sigma} = \mathrm{H}(\overline{\varepsilon_p})$ between this stress and the equivalent plastic strain $\overline{\varepsilon_p}$, so that if one of them is uniform so is the other). Thus Eq. (26) becomes

$$\frac{d\sigma_z}{dr} + \frac{\overline{\sigma}}{\rho} = 0 \quad \text{with} \quad \overline{\sigma} = H\left(2\ln\frac{a_0}{a}\right) \tag{27}$$

and can be integrated with the boundary condition $\sigma_z = \overline{\sigma}$ at $r = a$ (since $\sigma_r = 0$ at this point) to give

$$\sigma_z = \overline{\sigma}\left(1 + \int_r^a \frac{dr}{\rho}\right). \tag{28}$$

The tensile load P sustained by the specimen is the resultant force of the axial stresses σ_z acting on any cross section. Then, P can be expressed as a function of the variables introduced up to now by integrating the value given by Eq. (28) over the minimum cross section

$$\begin{aligned} P &= \int_0^a \sigma_z 2\pi x dx = \int_0^a \overline{\sigma}\left(1 + \int^a \frac{dr}{\rho}\right) 2\pi x dx \\ &= \pi a^2 \overline{\sigma}\left(1 + \frac{2}{a^2}\int_0^a \left[\frac{1}{\rho}\int_0^r x dx\right] dr\right) \\ &= \pi a^2 \overline{\sigma}\left(1 + \frac{1}{a^2}\int_0^a \frac{r^2}{\rho} dr\right) \\ &= \pi a^2 \overline{\sigma}\left(1 + \int_0^1 \frac{a}{\rho}\left(\frac{r}{a}\right)^2 d\left(\frac{r}{a}\right)\right). \end{aligned} \tag{29}$$

The second hypothesis formulated by Bridgman states the variation of the curvature radius ρ along the r-axis. Apart from Bridgman, other researchers ([3,6]) have proposed or suggested alternative hypotheses, but they all fall under the general formulation

$$\rho = a f\left(\frac{r}{a}, \frac{R}{a}\right). \tag{30}$$

As a consequence, the integral of the right-hand side of Eq. (29) depends only on the ratio a/R, which according to relation Eq. (24) holds constant during the deformation; so that the expression of the tensile load transforms into

$$P = \pi a^2 \overline{\sigma} F\left(\frac{a_0}{R_0}\right) = \pi a^2 H\left(2\ln\frac{a_0}{a}\right) F\left(\frac{a_0}{R_0}\right) \tag{31}$$

Table 1 Curvature radii and notch factors

	Radius ρ	Notch Factor $F(a_0/R_0)$
Bridgman	$\frac{aR}{r} + \frac{a^2}{2r} - \frac{r}{2}$	$\left(1 + \frac{2R_0}{a_0}\right)\ln\left(1 + \frac{a_0}{2R_0}\right)$
Davidenkov and Spiridonova	$\frac{r}{R}$	$1 + \frac{a_0}{4R_0}$
Eisenberg and Yen	$\frac{aR}{r} + \frac{a^2}{r} - r$	$\frac{1}{2} + \left(1 + \frac{R_0}{2a_0}\right)\ln\left(1 + \frac{a_0}{R_0}\right)$

where

$$F\left(\frac{a_0}{R_0}\right) \equiv 1 + \int_0^1 \frac{x^2 dx}{f\left(x, \frac{a_0}{R_0}\right)}. \tag{32}$$

With the exception of the factor $F(a_0/R_0)$ (henceforth, notch factor) this relation between the tensile load P and the radius a is that of an unnotched round bar of initial radius a_0 made from the same material as the circumferentially notched round specimen. Therefore, these two relations can be predicted one from another by dividing or multiplying the tensile load by the notch factor and considering the radius a as that of an unnotched bar or as that of the minimum cross section of a notched one. In particular, the maximum loads of these two tensile bars are the same except for this factor, whereupon the load of plastic collapse P_m of a circumferentially notched round specimen should be given by

$$P_m = \pi a_0^2 R_m F\left(\frac{a_0}{R_0}\right), \tag{33}$$

R_m being the tensile strength of the material. Another important consequence of Eq. (31) is the possibility of determining the stress-strain curve $\overline{\sigma} = H(\overline{\varepsilon_p})$ by testing notched round specimens and recording the tensile load as a function $P = P(a)$ of the radius of the minimum cross section. The function $H(\cdot)$ would be

$$H(x) = \frac{P(a_0 e^{-x/2})}{\pi a_0^2 e^{-x} F\left(\frac{a_0}{R_0}\right)} \tag{34}$$

and would be valid for a larger range of strain than that provided by a standard tensile test, since the plastic instability of a notched specimen would not mean the end of the test for the purpose of determining the stress-strain curve because it does not produce a geometrical configuration change as radical as the necking of a smooth specimen.

Table 1 shows different functions giving the curvature radius of the isostatic lines along the r axis according to Eq 30, as well as the corresponding notch factors as calculated from Eq. 32.

4 Finite Element Solution

As previously stated, Eqs. (24), (31), and (33) provide support to assess the validity of Bridgman's solution for tensile blunt notched specimens by checking them with numerical or experimental results. The numerical results were preferred to the experimental ones due to the difficulties of performing the measurements involved in Eqs. (24) and (31), since there are no simple methods to measure radii of curvature in a contour or tensile stress-strain curves beyond the tensile strength.

A large number of finite element solutions have been computed for elastic-plastic circumferentially notched tensile round bars ([13,14]). In general the results approach Bridgman's solution as the sharpness of the notch profile decreases, but the reported details do not allow Eq. (24) to be checked, so a finite element solution was specifically obtained for this purpose. The length of the tensile bar numerically modeled, the radius of the gross

cross section and the radius of the notch profile were 9.82, 1.23, and 0.90 times the radius a_0 of the minimum cross section. An elastoplastic strain-hardening material obeying the isotropic Von Mises yield criterion and the Prandlt-Reuss constitutive equations was assumed. The mechanical properties of this material, taken from an A533 vessel steel, are 200 GPa (Young's modulus), 0.3 (Poisson' ratio), 470 MPa (0.2 percent yield stress), and $\overline{\varepsilon_p}=(\overline{\sigma}/\sigma_0)^{8.5}$ with $\sigma_0=900$ MPa (Ramberg-Osgood's strain hardening curve). An axisymmetrical mesh of 431 nodes and 393 eight-node and six-node isoparametric elements, moderately refined at the crack profile and at the *r*-axis (with nine nodes on the former and 14 on the latter), was used for modeling a quarter of the bar, as shown in Fig. 3. According to symmetry, the nodes on the *r* and *z*-axis were constrained to move along their respective axis. Loading was simulated by imposing a single value of the displacement parallel to the *z*-axis on all the nodes on the base of the bar, this displacement being gradually increased until the maximum equivalent plastic strain in the bar roughly reached a value of 1, which occurred well beyond maximum load. The numerical computation was made with a commercial finite element code allowing large geometry changes. The Ramberg-Osgood's curve was piecewise implemented in the code for a plastic strain range from 0 to 1.5.

Special attention was paid to the numerical results concerned with Eqs. (22) and (24) in order to assess the accuracy of Bridgman's solution complemented with Eq. (19) for axisymmetrical notched bars. Figures 4, 5, and 6 are intended for this purpose. In Fig. 4 the displacements of the nodes on the *r*-axis are plotted in nondimensional and absolute values as a function of their initial position for a number of load levels covering elastic regime, contained yielding, general yielding, extended yielding, plastic instability (maximum load), and plastic unloading after plastic instability.

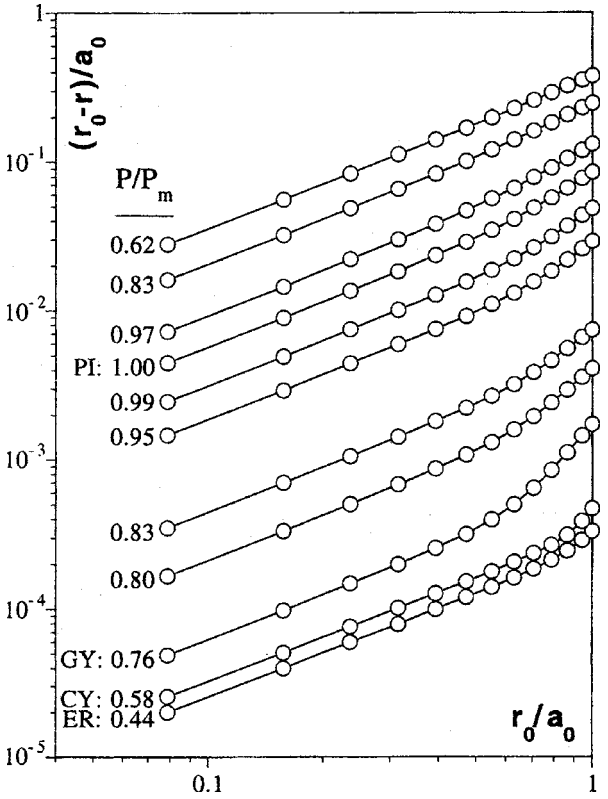


Fig. 4 Displacements of the *r*-axis nodes at different levels of the maximum load P_m. (ER: elastic regime; CY: contained yielding; GY: general yielding; PI: plastic instability)

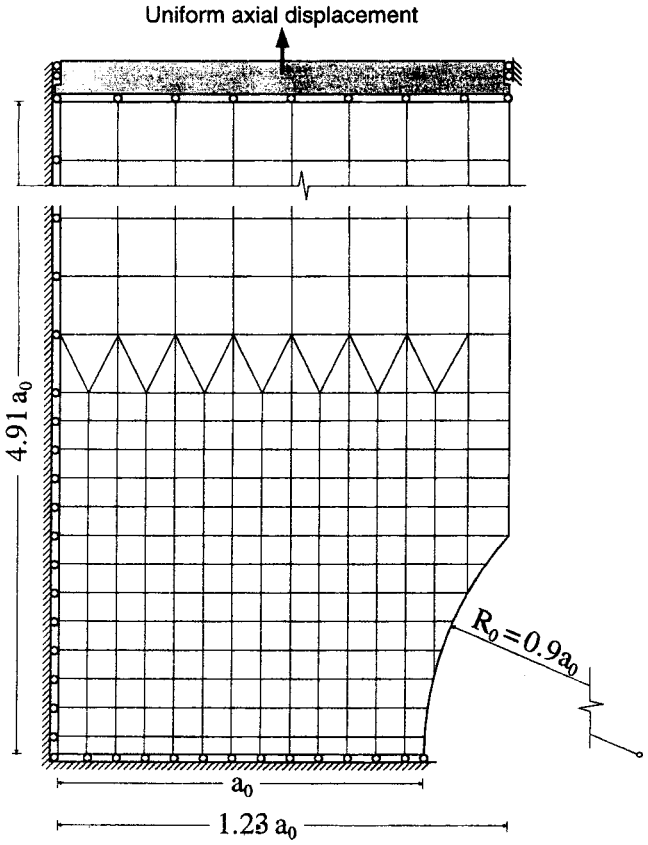


Fig. 3 Notch profile and finite element mesh used for the numerical solution

According to Eq. (22), with the logarithmic scales used in Fig. 4, the points from a same load level should lie on parallel straight lines of slope unity. As expected, in the elastic regimen the numerical displacements near the notch profile fail to follow this trend. The deviation increases as plasticity extends over the notch sections and is maximum at general yielding, but it decreases as plastic instability approaches and continues decreasing

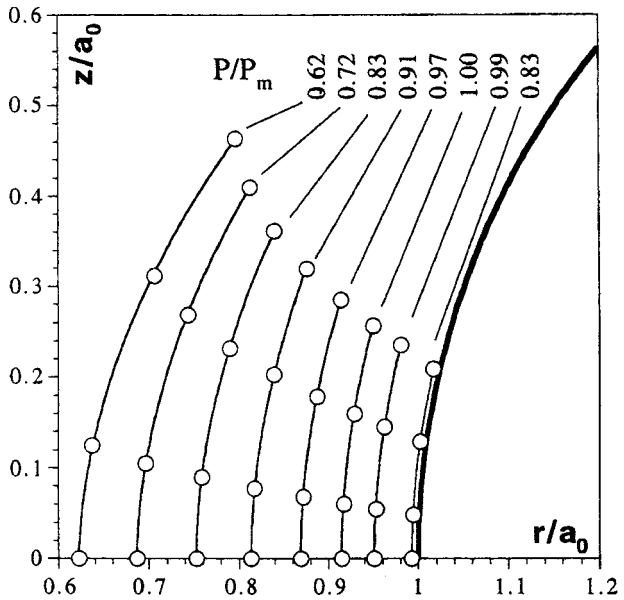


Fig. 5 Deformed notch profiles at different load levels (P_m is the maximum load)

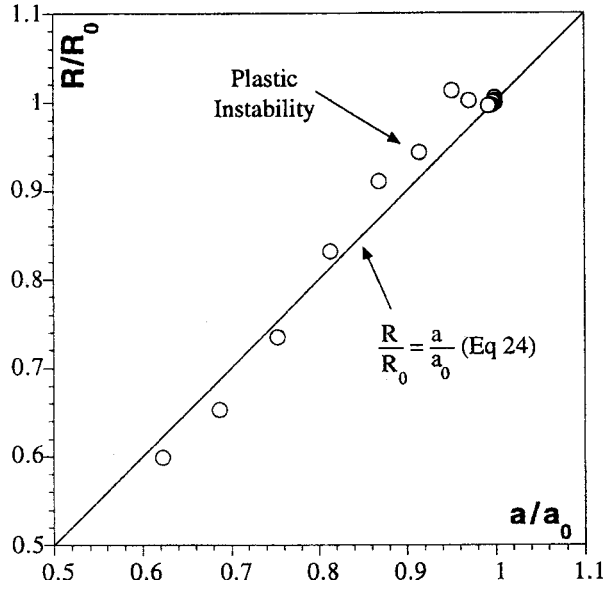


Fig. 6 Notch root radius as a function of deformation

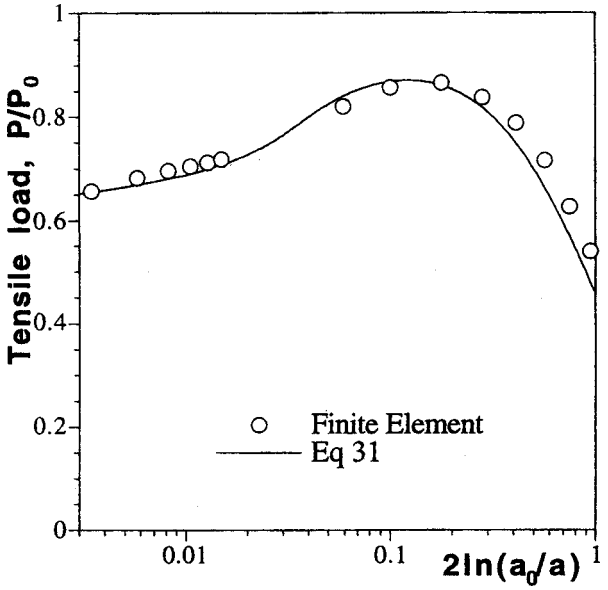


Fig. 7 Tensile load as a function of deformation

with plastic unloading until practically vanishing. With regard to Eq. (24), the notch profile in the neighborhood of the notch root has been plotted at different load levels in Fig. 5 from the displacements of four nodes lying on it. The corresponding values of the notch root radius R were determined as that of the circumference symmetric to the r-axis which passes through the node on the notch root and the nearest on the notch profile.

The resulting R/R_0 values are plotted in Fig. 6 as a function of the minimum area expressed by the ratio a/a_0. The exact solution for incompressible material provided by Eq. (19) allows the deviation from Eq. (24) to be explained. For small strain, the hoop and radial stretches a/a_0 and λ hardly differ from unity, so Eq. (19) predicts no significant difference between the radii R and R_0 unless a strain gradient giving rise to a value of λ' not much below $1/R_0$ be produced at the notch root. As the displacements plotted in Fig. 4 indicate, this requires a notch profile very much sharper than the finite element modeled one, so that the change of curvature cannot precede that of the overall dimensions. Since large deformations initiate at the early stages of plastic instability, Eq. (24) is verified until then simply because there is no geometry change, even though the radial displacements do not satisfy entirely Eq. (22) (Fig. 4). When the large geometry changes do occur, the notch profile is deformed according to Eq. (22) as shown in Fig. 6 by the points clustered around the straight line representing this equation. This is in agreement with the trend of the radial displacements to fulfill Eq. (22) as plastic unloading increases (Fig. 4). The maximum deviation from Eq. (24) takes place just before plastic instability, when strains are already large, but the radial displacements still differ from Eq. (22) (Fig. 4).

Finally, the tensile load obtained by the finite element modeling is plotted in Fig. 7 against the logarithmic strain $2\ln(a_0/a)$. Equation (31) is also plotted in Fig. 7 with Bridgman's notch factor of Table 1. The unit load used in the plots is $P_0=\pi a_0^2\sigma_0$, namely, the area of the minimum cross section multiplied by the stress constant σ_0 of the strain hardening curve.

5 Discussion

The adequacy of Bridgman's solution for axisymmetrical notched tensile bars will be discussed by comparing the finite element solution of Section 4 with the main theoretical predictions derived in Section 3, which give the notch root radius (Eq. (24)) and the tensile load (Eq. (31)) as a function of the specimen minimum diameter. The two results are a consequence of combining Eq. (19) with Bridgman's solution, but whereas the first is based only on the linear displacement distribution assumed at the minimum cross section of the bar, the second comes also from the curvatures assumed for the isostatic lines along the r-axis. So the errors that the deviations from these two assumptions produce on the notch root radius (a local effect) and on the tensile load (a global effect) will be assessed.

For large plastic deformation, the linear displacement distribution is a good approximation, as shown by the finite element calculated displacement field (Fig. 4). According to the absence of deviations, in this regime no disagreement should be expected between the theoretical prediction of the notch root radius given by Eq. (24) and the finite element results; indeed no significant disparity is detected, since the slight scatter of Fig. 6 is attributable to the numerical errors arising from the finite discretization of the bar. On the contrary, in the previous elastic and elastic-plastic regimes (roughly up to plastic instability) the displacements deviate from the linear distribution near the notch root (Fig. 4). This has no effect on the notch root radius for small strains, as already stated, and only at the transition to large deformation (the beginning of plastic instability), do the data in Fig. 6 indicate some disagreement between the theoretical results and the numerical ones.

As far as the tensile load is concerned, Fig. 7 shows good agreement between the values given by the finite element modeling and Eq. (31) in the range covering general and extended yielding up to plastic instability. The two results differ from maximum load, i.e., when the notch root radius largely changes. The tensile load being affected, the discrepancy is a global effect and cannot be due the displacement field at the minimum cross section, since in this range these displacements deviate so little from the linear distribution as to produce no local effect (Eq. (24) is satisfied). So, it can only be attributed to the inaccuracy of the curvatures adopted for the isostatic lines. The alternatives to Bridgman's curvatures given in Table 1 would not improve the agreement, because they would only modify the notch factor and would produce a displacement of the curve in Fig. 7 parallel to the ordinate axis. In fact, the curvatures derived by Bridgman seem to be an excellent solution of the type defined by Eq. (29), since this can be derived not only from the circumferences assumed by Bridgman to be the family of isostatic lines to which the axis r belongs, but also from other families of curves as ellipses or parabolas. A plausible explanation of the inaccuracy arises when Eq. (24) is substituted into Eq. (29) and it becomes apparent that the assumed

curvature distribution is the same through the deformation process but a scale factor. Therefore, as a consequence of Eq. (24), the root notch radius R cannot account for the influence of the ratio a/a_0 on the curvature distribution, so this dependence must be explicitly incorporated into Eq. (30).

The failure of Eq. (31) to predict the tensile load beyond plastic instability does not allow an unbounded range of the stress-strain curve to be measured by applying Eq. (34) to tests carried out with axisymmetrical blunt notched specimens. Nevertheless, Fig. 7 indicates that the tensile strength and even the yield stress might be determined by this procedure. Indeed, reported tensile tests covering a wide range of these properties and performed with two types of specimens provide experimental support for this possibility ([15]).

6 Concluding Remarks

A theoretical development was carried out aimed at deriving a relationship between the notch root radius of an axisymmetrical notched tensile bar and the displacement field at the minimum cross section for large geometry changes of the notch profile and for elastic-plastic incompressible material. Bridgman's displacement and stress solution for a tensile neck was examined in combination with that relationship with regard to their application to axisymmetrical blunt notched bars under tension, and this allowed the entire load-minimum diameter curve of the bar to be predicted. A finite element solution of a particular case was computed for comparison, and different steps of the analytical development were checked, the following conclusions being drawn as to the effects produced by the inaccuracy of Bridgman's solution for the notched bar on the predicted notch root radius and tensile load: no significant error of the former was observed over the entire range of plastic deformation (small and large), but significant differences were found in the latter after plastic instability.

Acknowledgments

The financial support of the Spanish Dirección General de Enseñanza Superior (Grant PB 95-0238) is gratefully acknowledged. Thanks are also due to Prof. C. Navarro for his valuable assistance in the finite element modeling.

Appendix

The aim of this section is to derive a kinematical equation relating the curvature of a deformed material line to the velocity field in a body which is being subjected to deformation. The equation is obtained for the general case and particularized for that of axisymmetry in Section 2 of the main text.

Let Γ be the curve which represents the deformed material line at present time. The position vector of a material point is $\mathbf{r}$, $\mathbf{v}$ is the velocity field, s is the distance along Γ, κ and τ are the curvature and the torsion of Γ, and $\mathbf{t}$, $\mathbf{n}$, and $\mathbf{b}$ are the vectors of an orthonormal basis in the tangent, principal normal, and binormal directions of Γ (Fig. 8). All these quantities and their directional derivatives along Γ are related by the Frenet equations:

$$\frac{\partial \mathbf{r}}{\partial s} = \mathbf{t} \tag{A1}$$

$$\frac{\partial \mathbf{t}}{\partial s} = \kappa \mathbf{n} \tag{A2}$$

$$\frac{\partial \mathbf{n}}{\partial s} = -\kappa \mathbf{t} + \tau \mathbf{b} \tag{A3}$$

$$\frac{\partial \mathbf{b}}{\partial s} = -\tau \mathbf{n}. \tag{A4}$$

In order to introduce the velocity field, a rule is required by which the order of differentiation between the directional derivative and the material time one (denoted by a superimposed dot) can be exchanged. Let $\mathbf{q}$ be the vector amount to be differentiated and $\dot{\varepsilon}_t$ the normal strain rate in the direction tangent to Γ; given that e^{ε_t} is the ratio ds/ds_0 of the deformed and undeformed length of an elemental arc of the material line, such a rule would be

$$\left(\frac{\partial \mathbf{q}}{\partial s}\right)^{\cdot} = \left(\frac{1}{e^{\varepsilon_t}}\frac{\partial \mathbf{q}}{\partial s_0}\right)^{\cdot} = \frac{1}{e^{\varepsilon_t}}\left(\frac{\partial \mathbf{q}}{\partial s_0}\right)^{\cdot} - \frac{\dot{\varepsilon}_t}{e^{\varepsilon_t}}\frac{\partial \mathbf{q}}{\partial s_0}$$

$$= \frac{1}{e^{\varepsilon_t}}\frac{\partial \dot{\mathbf{q}}}{\partial s_0} - \dot{\varepsilon}_t \frac{\partial \mathbf{q}}{\partial s} = \frac{\partial \dot{\mathbf{q}}}{\partial s} - \dot{\varepsilon}_t \frac{\partial \mathbf{q}}{\partial s}. \tag{A5}$$

Now, taking into account that any unit vector is normal to its derivatives, using the Frenet equations and applying Eq. (A5) to the tangent vector $\mathbf{t}$, the following expression for the time rate $\dot{\kappa}$ of the curvature is obtained:

$$\dot{\kappa} = \mathbf{n}\cdot(\kappa \mathbf{n})^{\cdot} = \mathbf{n}\cdot\left(\frac{\partial \mathbf{t}}{\partial s}\right)^{\cdot} = \mathbf{n}\cdot\left(\frac{\partial \dot{\mathbf{t}}}{\partial s} - \dot{\varepsilon}_t \frac{\partial \dot{\mathbf{t}}}{\partial s}\right) = \mathbf{n}\cdot\frac{\partial \dot{\mathbf{t}}}{\partial s} - \dot{\varepsilon}_t \kappa \tag{A6}$$

A further development of this expression requires an explicit form to be found of the time derivative of the tangent vector. This can be done by applying Eq. (A5) to the position vector $\mathbf{r}$, decomposing subsequently the velocity gradient into the sum of its symmetric (the strain rate tensor $\dot{\boldsymbol{\varepsilon}}$) and skew part (which is equivalent to the vector product with $1/2\,\mathrm{curl}\,\mathbf{v}$ as first factor), and finally introducing the shear strain rates $\dot{\gamma}_{tn}$ and $\dot{\gamma}_{tb}$ for the tangent, principal normal, and binormal directions. All this yields

$$\dot{\mathbf{t}} = \left(\frac{\partial \mathbf{r}}{\partial s}\right)^{\cdot} = \frac{\partial \dot{\mathbf{r}}}{\partial s} - \dot{\varepsilon}_t \frac{\partial \mathbf{r}}{\partial s} = \frac{\partial \mathbf{v}}{\partial s} - \dot{\varepsilon}_t \mathbf{t} = (\mathrm{grad}\,\mathbf{v})\mathbf{t} - \dot{\varepsilon}_t \mathbf{t}$$

$$= \dot{\boldsymbol{\varepsilon}}\mathbf{t} + \frac{1}{2}(\mathrm{curl}\,\mathbf{v})\times\mathbf{t} - \dot{\varepsilon}_t \mathbf{t} = \frac{1}{2}\dot{\gamma}_{tn}\mathbf{n} + \frac{1}{2}\dot{\gamma}_{tb}\mathbf{b} + \frac{1}{2}(\mathrm{curl}\,\mathbf{v})\times\mathbf{t}. \tag{A7}$$

The differentiation of this expression and the subsequent use of the Frenet equations provide the directional derivative of the time derivative of the tangent vector. Then, its projection onto the normal vector $\mathbf{n}$ can be found and substituted into Eq A6, which finally gives:

$$\dot{\kappa} = \frac{1}{2}\frac{\partial \dot{\gamma}_{tn}}{\partial s} - \frac{1}{2}\tau\dot{\gamma}_{tb} + \frac{1}{2}\frac{\partial(\mathrm{curl}\,\mathbf{v})}{\partial s}\cdot\mathbf{b} - \dot{\varepsilon}_t \kappa. \tag{A8}$$

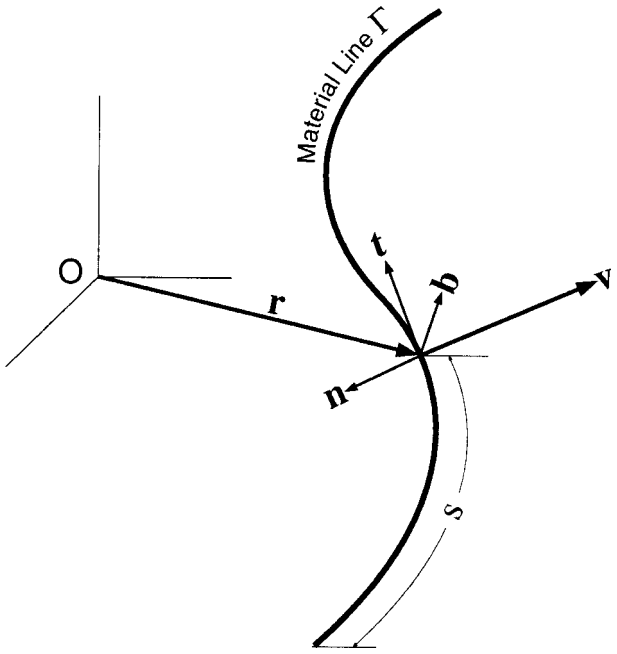


Fig. 8 Vectors concerning the motion of a material line

References

[1] Bridgman, P. W., 1944, ''The Stress Distribution at the Neck of a Tension Specimen,'' Trans. ASME, **32**, pp. 553–574.

[2] Bridgman, P. W., 1964, *Studies in Large Plastic Flow and Fracture*, Harvard University Press, Cambridge, MA.

[3] Davidenkov, N. N., and Spiridonova, N. I., 1946, ''Analysis of Tensile Stress in the Neck of an Elongated Test Specimen,'' Proc. A. S. T. M. **46**, pp. 1147–1158.

[4] Kaplan, A. M., 1973, ''The Stress and Deformation in Mild Steel During Axisymmetric Necking,'' ASME J. Appl. Mech., **40**, pp. 271–276.

[5] Jones, S. E., Gillis, P. P., and Shalaby, A. H., 1979, ''Stress Distributions in the Vicinity of a Neck,'' J. Appl. Phys., **50**, No. 5, pp. 3168–3173.

[6] Eisenberg, M. A., and Yen, C. F., 1983, ''An Anisotropic Generalization of the Bridgman Analysis of Tensile Necking,'' J. Eng. Mater. Technol., **105**, pp. 264–267.

[7] Hancock, J. W., and Mackenzie, A. C., 1976, ''On the Mechanisms of Ductile Fracture in High Strength Steels Subjected to Multiaxial Stress States,'' J. Mech. Phys. Solids, **24**, pp. 147–169.

[8] Mackenzie, A. C., Hancock, J. W., and Brown, D. K., 1977, ''On the Influence of State of Stress on Ductile Failure Initiation in High Strength Steels,'' Eng. Fract. Mech., **9**, pp. 167–188.

[9] Ritchie, R. O., Server, W. L., and Wullaert, R. A., 1979, ''Critical Fracture Stress and Fracture Strain Models for the Prediction of Lower and Upper Shelf Toughness in Nuclear Pressure Vessels Steels,'' Met. Trans., **10A**, pp. 1557–1570.

[10] Chaouadi, R., De Meester, P., and Vandermeulen, W., 1994, ''Damage Work as Ductile Fracture Criterion,'' Int. J. Fract., **66**, pp. 155–164.

[11] Amar, E., and Pineau, A., 1985, ''An Interpretation of Ductile Fracture Toughness Temperature Dependence of a Low Strength Steel in Terms of a Local Approach,'' Eng. Fract. Mech., **22**, pp. 1061–1071.

[12] Fontaine, A., and Jeunehomme, S., 1989, ''Numerical Study and Local Approach of Fracture in a Perlitic Steel,'' *Advances in Fracture Research (ICF7)*, Pergamon Press, Oxford, UK, pp. 3865–3871.

[13] Moran, B., and Norris, M. D., 1978, ''Computer Simulation of the Round Notched Tension Test of Cox and Low,'' Met. Trans., **9A**, pp. 1685–1687.

[14] Hancock, J. W., and Brown, D. K., 1983, ''On the Role of Strain and Stress State in Ductile Failure,'' J. Mech. Phys. Solids, **31**, pp. 1–24.

[15] Valiente, A., and Lapeña, J., 1996, ''Measurement of the Yield and Tensile Strength of Neutron Irradiated and Post-Irradiation Recovered Vessel Steels With Notched Specimens,'' Nucl. Eng. Des., **167**, pp. 11–22.

Cavitation and Mushrooming in Attack of Thick Targets by Deforming Rods

M. Lee
School of Mechanical and
Aerospace Engineering,
Sejong University,
98 Kwangjin-Gu,
Kunja-Dong, Seoul 143747, Korea
e-mail: mlee@sejong.ac.kr

This paper analyzes the cavitation and mushrooming phenomena occurring in a metal target during the high-speed penetration by a deforming rod. It is motivated by the fact that two mechanisms are involved in the formation of cavity by a deforming rod. First, the flow of the deformed rod products exerting radial stress on the target opens a cavity (mushrooming). Second, the radial inertia of the target as it flows around the head plays another part in the formation of the cavity (cavitation). By examining the dynamics of the flow of deformed rod products, the proposed model can estimate the extent of cavity growth due to the mushrooming effect. Predicted results for the final cavity growth are found to match well with the corresponding test values. [DOI: 10.1115/1.1360690]

Introduction

Penetration mechanics involves the behavior of materials under extreme loading conditions, such as plasticity, fracture, and hydrodynamics. Among a variety of interests, there have been several previous attempts to explain the formation of cavity in a steel target that is penetrated by high-speed projectiles ([1,2]). As shown in Fig. 1, Hill [3] investigated the cavitation phenomena produced by a *nondeforming* projectile traveling through an infinite medium (any ductile metal) at constant velocity. If the cavity radius is λa, he provides the following equation:

$$\lambda^2 = \frac{P}{p} = 1 + \frac{C_d \rho U^2}{2p} \tag{1}$$

where a is the radius of a rigid projectile, P is the resistance pressure, p can be interpreted as the theoretical work per unit volume to make cylindrical or spherical holes ([4]), ρ is the target density, C_d is the drag coefficient depending on the headshape, and U is the penetration velocity. For a rigid body penetration, the impact velocity V is equal to the penetration velocity. According to Hill [3], the dependence of p on headshape is almost negligible and it is evidenced by quasi-static experiments with long punches. For the static test where there is no cavitation, the resistance load is pA, where A is the projected area of the projectile.

More pertinent to the present investigation is the analysis of the hypervelocity penetration, in which the projectile itself deforms. This is usually applicable to military terminal ballistics. Walters and Scott [5] formulated a theory of the cavity growth rate under ballistic impact conditions. However, this model is applicable to thin plates and surface penetration/perforation where a plane stress condition is expected to exist. For a deforming rod in attack of thick targets at high speed, Miller [6] presented an engineering description of the deforming rod front and kinematics. Without the presence of material strengths, however, the model is not able to predict the final cavity growth.

Although one way to analyze a detailed penetration problem is to use computer predictions, this approach is computationally expensive. By using an analytical model, significant parameters and how they vary over a wide velocity range can be obtained. Hence, the purpose of this paper is to analyze the cavitation process occurring in a metal target during penetration by a deforming rod. The analysis includes the centrifugal force exerted by the flow of deforming rod material, radial inertia of the target, and the strength of the target. An axisymmetric coordinate system moving at the constant penetration velocity U is used as reference, in which the bottom of the contact surface is not moving. Some assumptions are made, many of which are motivated by the results of experiments and numerical simulations.

Observations and Assumptions

It is worth noting that for hypervelocity penetration the final cavity diameter can be two to three times larger than the projectile diameter. This is due to the fact that two distinct mechanisms are involved in the radial motion of target material resulting from hypervelocity penetration. First, the deformed rod elements, accompanied by a very high stress level in the target material, play a dominant role in opening the cavity (mushrooming). This is only observed during hypervelocity penetration ([7]). Second, the target inertia is responsible for further cavity formation until the target strength forces it to come to rest (cavitation). The term ''cavitation'' is borrowed from Hill's work ([3]) which deals with the attack of thick targets by nondeforming projectiles, in which only the cavitation effect is investigated.

The theory must account for the previous two mechanisms. Indeed, since the coupling contribution of the two effects, if any, could not be discerned from the available data, it is not clear how to model the coupling effects of the two mechanisms. This difficulty can be eliminated by assuming that the total cavity growth is due to the sum of the cavity growth produced by each effect acting independently, regardless of the order of application of loads. In other words, the principle of superposition is used for the total cavity growth. It is now necessary to evaluate the two effects individually.

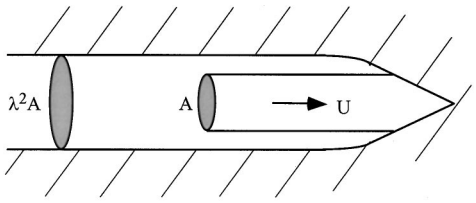


Fig. 1 Cavitation in a thick target by a nondeforming projectile

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Feb. 16, 1999; final revision, Nov. 27, 2000. Associate Editor: K. T. Ramesh. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Mushrooming

The formation of cavity solely due to mushrooming can be derived from Miller's analysis ([6]) with the presence of target strength. This is due to the fact that the target strength will eventually halt the cavity growth. The geometry of the mushrooming process is shown in Fig. 2. The target is assumed to be semi-infinite. Note that only half of the rod is displayed in the figure. The length is scaled by the radius of rod, such that the scaled radius of the rod is one. Now we focus on the dynamics of the flow of deforming rod elements during penetration. The flow is considered as incompressible, steady, and inviscid relative to the stagnation region (head). The trace of the centerline of the deforming rod elements allows us to determine the cavity growth. Since an axisymmetric coordinate system moving at the constant penetration velocity (U) is used as reference, the speed of the flowing material is always $(V-U)$ regardless of θ and r. Hence, this model is not valid for the entrance and end phases and is valid for the steady-state phase. We can also obtain a time-dependent model using a variable penetration velocity, if necessary. As shown in Fig. 2, we consider two contact regions between the head and target: a finite stagnation region ($r<r_1$, $\pi/2\leqslant\theta\leqslant\pi$), over which the resistant pressure equals the stagnation pressure $1/2\rho_t U^2$, and the domain outside of the stagnation region ($0\leqslant\theta\leqslant\pi/2$).

First, consider the case when the flowing rod element remains outside of the stagnation region. From the rod mass conservation in a coordinate system located at the head, the nondimensional thickness of a flowing rod element is given by

$$h(\theta)=\frac{1}{2r(\theta)}. \tag{2}$$

Note that the radius is nondimensional. The resistant pressure profile exerted by the target due to inertia is assumed to be $1/2\rho_t U^2\sin^3\theta$ as provided by Miller [6], where θ is the angle between the tangent direction of the centerline of the rod flow and the axis of symmetry. Although the Newtonian pressure $1/2\rho_t U^2\sin^2\theta$ allows a simple closed-form solution it overpredicts the pressure profile. The target resistance that is equal to pressure force plus target strength must balance the centrifugal force generated by the flow of a deforming element,

$$R_t+\frac{1}{2}\rho_t U^2\sin^3\theta=h(\theta)\rho_p\frac{(V-U)^2}{R(\theta)}, \tag{3}$$

where R_t is the work expended per unit volume to open a spherical or cylindrical cavity, which can be interpreted as p in Eq. (1), and $R(\theta)$ is the scaled local radius of curvature of the centerline curve. There are no axial stress gradients along the centerline since the speed of the flowing material is always $(V-U)$ regardless of θ and r. From geometric relations,

$$R(\theta)=-\frac{1}{\sin\theta}\frac{dr}{d\theta}. \tag{4}$$

Combining Eqs. (2), (3), and (4) gives the differential equation describing the trajectory of a flowing element,

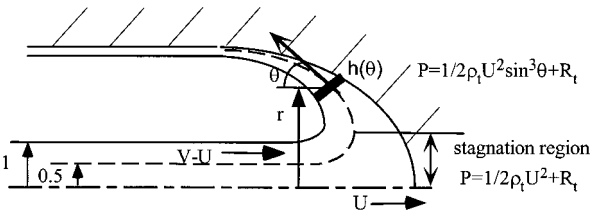


Fig. 2 Mushrooming and region of the target resistant pressure

$$R_t+\frac{1}{2}\rho_t U^2\sin^3\theta=-\frac{1}{2r}\rho_p(V-U)^2\sin\theta\left(\frac{dr}{d\theta}\right)^{-1}. \tag{5}$$

The pressure on the left-hand side in Eq. (5) is assumed to act normal to the flow at radius r, where r is the radius of the centerline of a flowing rod element, although the resistance pressure actually acts normal to the interface between the rod and target. By introducing two constants,

$$K_1=\frac{1/2\rho_p(V-U)^2}{R_t},\quad K_2=\frac{1/2\rho_t U^2}{R_t}, \tag{6}$$

Eq. (5) becomes

$$2rdr=-\frac{2K_1\sin\theta}{1+K_2\sin^3\theta}d\theta. \tag{7}$$

To determine the trajectory of the flowing element as a function of θ, this equation is to be integrated with an initial condition $r=r_1$ at $\theta=\pi/2$, which yields

$$r(\theta)=\sqrt{r_1^2-\int_{\pi/2}^{\theta}\frac{2K_1\sin\theta}{1+K_2\sin^3\theta}d\theta}. \tag{8}$$

Since the deformed rod element starts to flow from $\theta=\pi/2$ to 0, the numerical integration of Eq. (8) for the domain of $0\leqslant\theta\leqslant\pi/2$ allows us to determine the cavity growth (a_{c1}) which is solely due to the mushrooming. Note that if the inertia pressure had been approximated to be $1/2\rho_t U^2\sin^2\theta$, there would be a simple closed-form solution given by

$$a_{c1}=\sqrt{r_1^2+\frac{K_1}{\sqrt{K_2+K_2^2}}\ln[1+2K_2+2\sqrt{K_2+K_2^2}]}. \tag{9}$$

Second, consider the case when the flowing rod element remains inside of the stagnation region. For this region, the location of the center of deformed rod material is expected to vary from half of the radius ($\theta=\pi$), where the rod material is not initially deformed, to $r_1(\theta=\pi/2)$. In order to find the initial condition r_1, by the same reasoning with above, the pressure balance equation (Eq. (3)) can be written for inside of the stagnation region,

$$R_t+\frac{1}{2}\rho_t U^2=h(\theta)\rho_p\frac{(V-U)^2}{R(\theta)}. \tag{10}$$

Equation (10) implies that the actual resistance pressure, which does not have angle variations, is acting on the centerline of the rod elements. So, the stagnation pressure is always acting normal to the penetration direction. The existence of stagnation region will compensate the assumption that the resistant pressure is assumed to act normal to the centerline. By using the radius of curvature and the thickness of the flowing element, Eq. (10) is to be integrated from $r=0.5$ at $\theta=\pi$ to r_1 at $\theta=\pi/2$ resulting in

$$r_1=\sqrt{\frac{1}{4}+\frac{2K_1}{1+K_2}}. \tag{11}$$

At this point it is possible to calculate the cavity growth via Eq. (9). However, the calculation requires the values of R_t and U. For steady-state penetration of the stagnation head, the modified bernoulli equation applies ([8–10])

$$\frac{1}{2}\rho_p(V-U)^2+Y_p=\frac{1}{2}\rho_t U^2+R_t. \tag{12}$$

Y_p is the pressure at which the rod begins to flow hydrodynamically. It is assumed that the projectile is rigid except for an infinitesimally thin region near the target-projectile interface where erosion is occurring. Then Tate [9] gives U,

$$U=\frac{V-\alpha\sqrt{V^2+2(1-\alpha^2)(R_t-Y_p)/\rho_t}}{(1-\alpha^2)}\quad\text{for }\alpha\neq1 \tag{13}$$

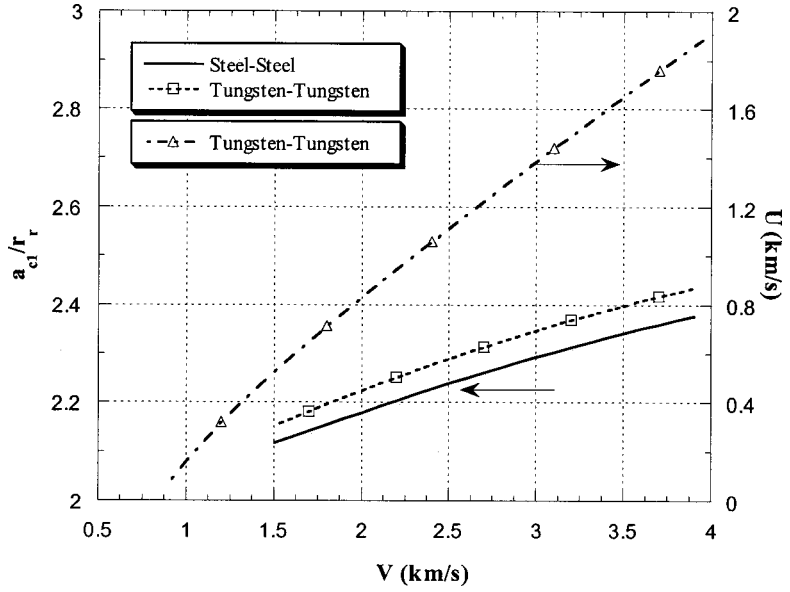


Fig. 3 The ratio of cavity growth due to mushrooming to the rod radius with impact velocity, and the corresponding penetration velocity

$$U=\frac{V}{2}-\frac{(R_t-Y_p)}{\rho V} \quad \text{for } \alpha=1 \tag{14}$$

where $\alpha=\sqrt{\rho_t/\rho_p}$. For $Y_p<R_t$, there is a critical velocity given by $[2(R_t-Y_p)/\rho_p]^{1/2}$ below which penetration no longer occurs. The projectile will however continue to erode until it comes to rest. For $Y_p>R_t$, there is a critical velocity given by $[2(Y_p-R_t)/\rho_t]^{1/2}$ below which the projectile behaves as a rigid body ([11]). The R_t value determined by the classical spherical cavity expansion model as presented by Satapathy and Bless [12] is used in the present work.

In order to gain an insight into the influence of mushrooming on formation of the cavity, the calculated values of the cavity size for the penetration of a steel rod into a steel target and a tungsten rod into a tungsten target are shown in Fig. 3. The remaining undetermined values, (R_t-Y_p), used in the calculations are 2.65 GPa for the first case and 4.17 GPa for the second case ([13]). The corresponding penetration velocity for the impact of tungsten rod into tungsten target is also shown in the figure. The critical impact velocity below which the projectile can not penetrate is about 700 m/s. The penetration velocity then increases with increasing impact velocity. Generally, it can be shown that the ratio of cavity radius to rod radius increases with impact velocity, however, the ratio is less than 2.5 even up to 4 km/s. As expected, it can be concluded that another mechanism should be taken into account to explain the actual cavity measured in the experiments. This will be described in the next section.

Cavitation

We now turn to the cavitation stage in which the inertia imparted to the target plays a role for further cavity expansion. This inertia is created by flow in the target around a rod head moving at constant velocity U. To estimate the additional cavity formation due to the target inertia, we consider the target response to an ''equivalent rod''—a rigid body of the same shape as the deforming rod head and moving at a constant penetration velocity ([3]). In other words, the eroding rod resembles a spherical nose rigid projectile. Actually, due to deceleration of a rigid body there is no steady state unless the body is self-propelled. However, steady state, except during the initial and end phases of the penetration, can be achieved for deforming projectiles. When a rigid projectile is moving through the ductile target, the force on projectile (work per unit length) is put in the Poncelet form,

$$\frac{F}{(\pi a_{c1}{}^2)}=R_t+\frac{1}{2}C_d\rho_t U^2. \tag{15}$$

It is explicit in the right-hand side of Eq. (15) that the first term is static given by cavity expansion, and the second term is the drag force that comes about due to the convective inertia effects ([14]). Since the cavity grows from a_{c1} to a final size a_c due to the inertia effect, the energy (per unit length) balance equation becomes

$$\frac{1}{2}(\pi a_{c1}{}^2)C_d\rho_t U^2=\pi(a_c{}^2-a_{c1}{}^2)R_t. \tag{16}$$

The simple energy balance equation implies that the inertia is utilized to open a cavity from a_{c1} to a_c against the resistant pressure R_t. The solution to the final cavity size is then given by

$$\lambda=\frac{a_c}{a_{c1}}=\sqrt{1+\frac{1/2C_d\rho_t U^2}{R_t}}. \tag{17}$$

When the drag coefficient is assumed to be 0.5 (for a hemispherical nose), this equation coincides with Eq. (1) but with a_{c1} replacing projectile radius. This is due to the fact that there is no mushrooming for a rigid-body penetration. Again the current study is mainly focused on deforming projectiles, in which the modified Bernoulli equation holds ([8–10]).

As an illustration, Fig. 4 shows the variation of $(a_{c1}/a_c)^2$ with impact velocity. For example, (a_{c1}/a_c) at an impact velocity of 2.7 km/s is about 0.7. This means that 70 percent of the cavity radius growth is due to mushrooming. This ratio at an impact velocity of 1.5 km/s is about 0.9 (tungsten rod into tungsten target case). In this case, 90 percent of the cavity radius growth is due to mushrooming. Hence, the extent of the cavity ''overshoot'' due to target inertia increases with impact velocity since the U^2 term in Eq. (17) becomes more important in the high impact velocity limit.

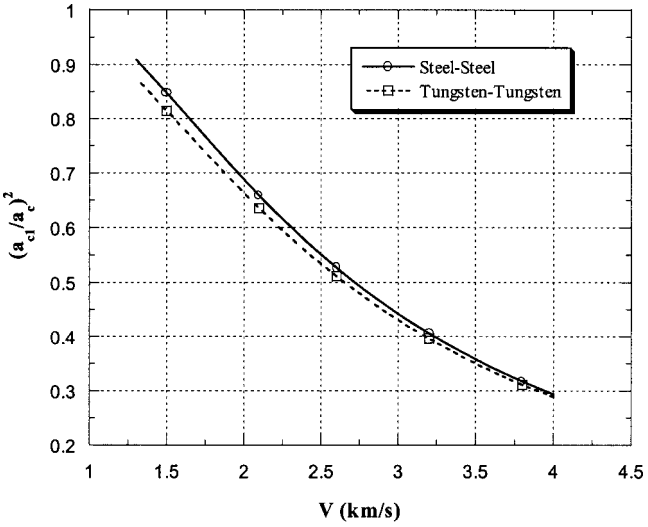


Fig. 4 Square of the ratio of cavity growth due to mushrooming to final growth with impact velocity

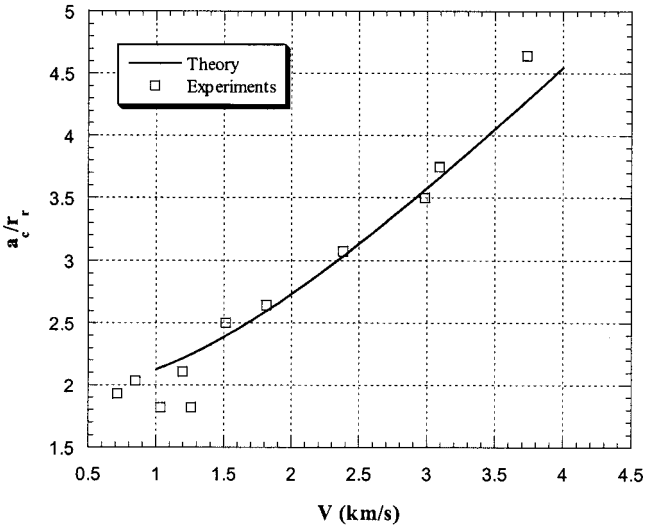


Fig. 6 Ratio of cavity radius to rod radius with impact velocity, experimental data for tungsten rod striking into tungsten target provided by Hohler and Stilp [15]

Comparison of Theory and Experiments

A cavity growth model has been derived considering both the mushrooming and cavitation effects. Since the accuracy of the current model relies on the validity of the assumptions as well as the accuracy of the selected values of the parameters, it is tested with the experimental data. Comparison between the predictions and the corresponding experimental data are made in Figs. 5–7. Figure 5 compares Hohler and Stilp's data ([15]) to the current model for the impact of steel rods into steel targets. The cavity diameter was measured at the mid depth of the cavity. It can be concluded that by including the cavitation effect, the predictions compare well to the experimental data at high velocities as well.

Even for the case of penetration of tungsten rods ($\rho_p = 17400\ \mathrm{kg/m^3}$) into tungsten targets ([15]), as shown in Fig. 6, the agreement is quite good. A further comparison for the case of aluminum alloy projectiles ($\rho_p = 2800\ \mathrm{kg/m^3}$) into lead targets ([13]), which produces quite a large crater size, is also shown in Fig. 7. As shown in the figure, most of the cavity production is due to the cavitation effect in this case. The analytical predictions compare well with the experimental data. Table 1 summarizes the material data used in the analysis. The material data for the aluminum alloy and lead are obtained from Refs. [8,16].

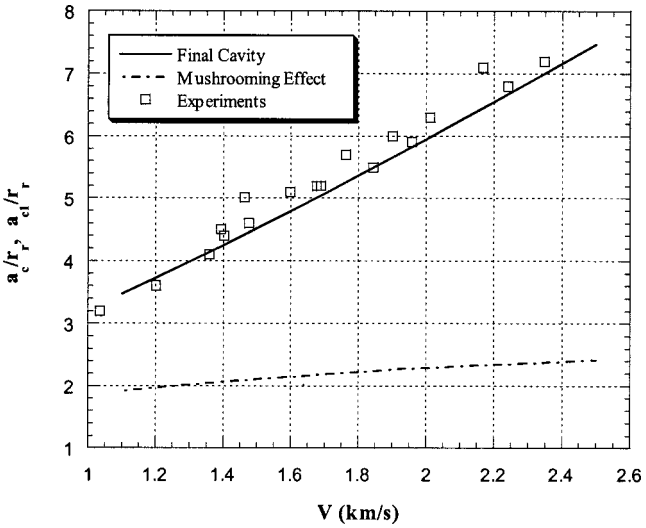


Fig. 7 Ratio of cavity radius to rod radius with impact velocity, experimental data for aluminum alloy rod striking into lead target provided by Tate [9]

Table 1 Material parameters used in the analysis

Projectile→Target	Y_p (GPa)	$(R_t - Y_p)$ (GPa)
Steel→Steel	1.6	2.65
Tungsten→Tungsten	2	4.17
Aluminum Alloy→Lead	0.8	−0.55

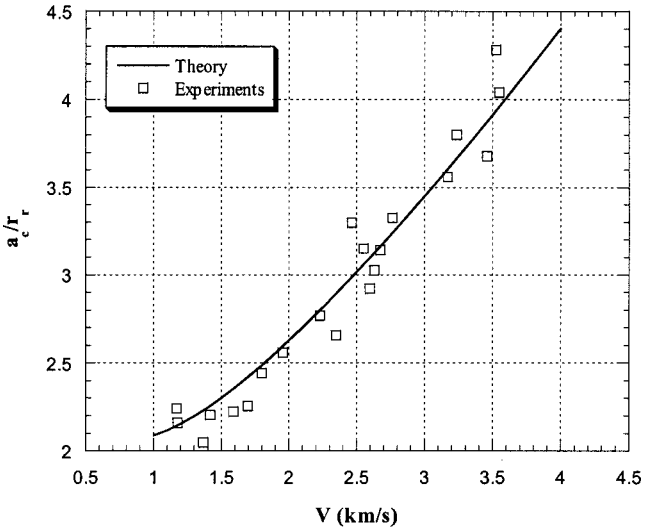


Fig. 5 Ratio of cavity radius to rod radius with impact velocity, experimental data for steel rod striking into steel target provided by Hohler and Stilp [15]

Conclusions

An analytical model for cavity formation in a ductile target by hypervelocity penetration of a deforming long rod is presented and the predictions are compared with the experimental data. It has been demonstrated for deforming projectiles that by considering both the mushrooming and cavitation effects the predictions compare well with the experimental data. An estimate of the amount of mushrooming with respect to the impact velocity is a

valuable contribution of the current analysis. The model is also found to be accurate for rod projectiles for many different material systems. For a deforming projectile at high impact velocity, some of the kinetic energy is absorbed by the enhanced expansion of a hole otherwise it goes towards penetration. The application of the current cavity formation model to a nonhomogeneous rod, such as jacketed rods, has been investigated ([17]).

Acknowledgment

The reported work was initiated when the author was with the Institute for Advanced Technology USA, and supported by Korea Research Foundation Grant (KRF-99-041-E00015).

References

[1] Bethe, H. A., 1941, ''Attempt of a Theory of Armor Penetration,'' *Frankford Arsenal Report*, Philadelphia, PA.
[2] Taylor, G. I., 1941, ''The Formation and Enlargement of a Circular Hole in a Thin Plastic Sheet,'' Q. J. Mech. Appl. Math., **1**.
[3] Hill, R., 1980, ''Cavitation and the Influence of Headshape in Attack of Thick Targets by Non-Deforming Projectiles,'' J. Mech. Phys. Solids, **28**, pp. 249–263.
[4] Bishop, R. F., Hill, R., and Mott, N. F., 1945, ''The Theory of Indentation and Hardness Tests,'' Proc. Phys. Soc., **57**, p. 148.
[5] Walters, W. P., and Scott, B. R., 1985, ''The Crater Radial Growth Rate Under Ballistic Impact Conditions,'' Comput. Struct., **20**, No. 1–3, pp. 641–648.
[6] Miller, C. W., 1995, ''Two-Dimensional Engineering Model of Jet Penetration,'' *Proceedings of the 15th International Symposium on Ballistics*, American Defense Preparedness Association, pp. 8–21.
[7] Shinar, G. I., Barnea, N., Ravid, M., and Hirsch, E., 1995, ''An Analytical Model for the Cratering of Metallic Targets by Hypervelocity Long Rods,'' *Proceedings of the 15th International Symposium on Ballistics*, Israel, American Defense Preparedness Association, pp. 59–66.
[8] Tate, A., 1967, ''A Theory for the Deceleration of Long Rods after Impact,'' J. Mech. Phys. Solids, **15**, pp. 387–399.
[9] Tate, A., 1967, ''Further Results in the Theory of Long Rod Penetration,'' J. Mech. Phys. Solids, **17**, pp. 141–150.
[10] Alekseevskii, V. P., ''Penetration of a Rod into a Target At High Velocity,'' *Combustion, Explosion and Shock Waves*, Vol. 2, Faraday Press, New York, pp. 63–66 (translation from the Russian).
[11] Walker, J. D., and Anderson, C. E., 1995, ''A Time-Dependent Model for Long-Rod Penetration,'' Int. J. Impact Eng., **16**, No. 1, pp. 19–48.
[12] Satapathy, S., and Bless, J. S., 1996, ''Quasi-Static Penetration Tests of PMMA: Analysis of Strength and Crack Morphology,'' Mech. Mater., **23**, No. 4, p. 323.
[13] Tate, A., 1986, ''Long Rod Penetration Models—Part II. Extensions to the Hydrodynamic Theory of Penetration,'' Int. J. Mech. Sci., **28**, No. 9, pp. 599–612.
[14] Bless, S., 1995, ''Impact and Penetration Mechanics,'' short course notes.
[15] Hohler, A., and Stilp, A., 1987, ''Hypervelocity Impact of Rod Projectiles With L/D From 1 to 32,'' Int. J. Impact Eng., **5**, pp. 323–331.
[16] Szendrei, T., 1995, ''Analytical Model for High-Velocity Impact Cratering With Material Strengths: Extensions and Validation,'' *15th International Symposium on Ballistics*, Israel, American Defense Preparedness Association, pp. 123–131.
[17] Lee, M., 2000, ''Analysis of Jacketed-Rod Penetration,'' Int. J. Impact Eng., **24**, No. 9, pp. 891–905.

N. Yoshida
General Manager,
Engineering Research Institute,
Sato Kogyo Co., Ltd.
Nihonbashi Honcho 4-12-20,
Chuo-ku, Tokyo 103-8639, Japan

T. Nonaka[1]
Professor,
Chubu University,
1200 Matsumoto-cho,
Kasugai, Aichi Prefecture 487-8501, Japan

Hysteretic Behavior of a Bar Under Repeated Axial Loading: An Extended History

Analytical study is made of an elastic-perfectly plastic bar under repeated axial loading. A previous formulation on a pin-ended bar is extended here to include the effects of load eccentricity and rotational constraint at the bar ends. Basic equations are derived, based on the assumptions of planar and small deflection, and of symmetry with respect to the bar center. The end spring is allowed to yield. Numerical examples are presented to demonstrate the application of the basic equations, and adequacy is shown for any specified history of axial displacement. Diagrammatical representation of state variation provides a better understanding of the hysteretic behavior as well as the applicability of the basic equations. [DOI: 10.1115/1.1360691]

Introduction

Axially loaded members play an important role in such structures as trusses or braced frames. Clarification of the performance of these structures requires knowledge of the load-deformation characteristics of the members. Plastic action in axially loaded members ordinarily takes precedence over flexurally loaded members because of the predominance in stiffness of the former over the latter. Structural members are often subjected to variable repeated loading; load due to winds, earthquakes, cranes, transportation vehicles, and some machine parts are applied repeatedly in nature, and they may act in different or opposite directions. An axially resistant member in such a condition may buckle under compression, deform plastically, but may partially recover in a subsequent tension. It may undergo plastic elongation and as a result become loosened, reducing the overall stiffness of the structure.

Because of the lack of knowledge on the hysteretic behavior of axially loaded structural members, a series of investigations was begun a few decades ago to determine the load-deformation relationship of a steel bar. Aside from computational or experimental studies, previous analyses have provided the derivation of basic equations in a closed form ([1–7]). On the basis of a one-dimensional idealization and of elastic-perfectly plastic behavior of its element, the basic equations are adequate to formulate the load-deformation characteristics of a simply supported prismatic bar for any specified history of axial loading. It was assumed that an effective length could be found such that the bar had inflections or was free to rotate at both ends of the length under the action of the centroidal axial forces.

While the results are of great value when these conditions are met with sufficient accuracy, there are cases where great eccentricity is brought about by the connection and cases where the effective length cannot be approximated as constant or is difficult to estimate. An example is a braced frame in which braces may have eccentricity, or bracing and neighboring members are of comparable flexibility. End supporting gusset plates deform elastically or may yield due to the rotation of the ends of the brace, or braces may have initial deflections. Although a study of post-buckling behavior of steel braces in frames shows that an elastically restrained brace can be treated as a pin-ended strut of half length ([8]), this is uncertain for repeated loading ([9]). Thus, an attempt has been made in the present paper to develop an extended theory that allows for externally applied end moments as well as load eccentricity on a bar.

Two papers have come to the authors' attention, dealing with the effects of end restraints on the hysteretic behavior of steel braces ([8,10]). Prathuangsit et al. have assumed elastic-perfectly plastic bending, and made use of the notion of a yield hinge ([8]). They take into account that axial force reduces the fully plastic moment but not that plastic axial deformation takes place at a yield hinge in addition to bending or rotation. While the latter fact is well understood from the plastic flow or normality rule under combined action of bending moments and axial forces, this is not recognized either in earlier papers of their group ([11,12]). It may be worthwhile to point out that an account of the plastic axial deformation at yield hinges in a kinematical approach leads to a reduction in the plastic collapse load of a frame ([13]); this is the same as the effect caused by the reduction of fully plastic moments due to the existence of axial force in a statical approach. In the papers cited above ([1–7]), the plastic axial deformation has been shown to play an important role in the hysteretic behavior of an axially loaded bar of moderate length. Mitani has modified Nonaka's formulation to analyze eccentrically loaded bars with elastic restraints against rotation at the bar ends ([10]). With recourse to step-by-step integration, he has considered some numerical examples of steel bars of rectangular and wide-flange cross sections.

The present paper is concerned with a more general formulation to allow for nonlinear rotational restraint as well as load eccentricity by extending the previously established basic equations. Emphasis is placed on closed-form formulation, even with the drawback of drastic assumptions.

Assumptions

Supposing an initially straight bar of length L in equilibrium with a slowly varying eccentric axial load, analysis is made on the basis of simplifying assumptions, in the same line as before except for supporting conditions. They are summarized by the following items:

1 A uniform cross section is assumed, with dimensions sufficiently small in comparison with L that the bar may well be

[1]Formerly at Kyoto University.

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 24, 1999; final revision, Nov. 28, 2000. Associate Editor: R. C. Benson. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

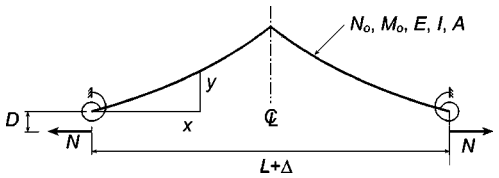


Fig. 1 Problem under consideration

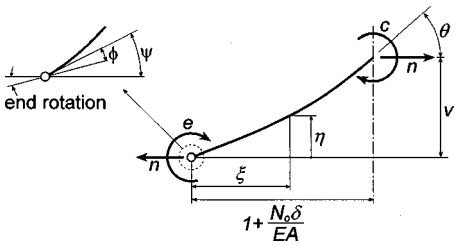


Fig. 2 Dimensionless positive quantities

treated as a one-dimensional continuum, but the slenderness is moderate so that the strength in compression is of the same order of magnitude as that in tension.

2 The cross section is assumed to be of such symmetry that the bar deforms only in a definite plane without torsion.

3 Plane cross sections are assumed to remain plane without shear deformation. Elemental behavior is assumed to be elastic-perfectly plastic under combined action of axial force and bending moment, and the yield condition is such that the bending moment is a function of the axial force.

4 The load is composed of a pair of equal and opposite forces with intensity N applied at the ends in the direction parallel to the original axis of the bar with eccentricity D in the plane of deformation. The load N is taken positive when tensile and negative when compressive. An assumption is made of symmetry with respect to the center of the bar so that the two halves behave in an identical manner.

5 Constraint from neighboring members at the connection is represented by spiral springs as shown in Fig. 1. (The deflection curve in this figure represents a generic state with previous hinging action.) It is assumed that the resisting moment is a piecewise linear or polylinear function of the angle of end rotation; this includes the elastic-perfectly plastic spring as a special case.

6 Two assumptions of small deformation are made: Change in length is negligibly small as compared with the original length of the bar; deflection is small enough so that the square of the slope in the deflection curve is negligible in comparison with unity. (See the Appendix for discussion on this point.)

7 Material ductility and absence of local instability are assumed.

Assumption 1 excludes extremely stubby and slender bars. If cross-sectional dimensions are of the same order of magnitude as the bar length as in the former case, shear effect may become important; on the other hand, if the bar is so slender that its compressive strength is negligibly small as compared with the tensile, it may well be treated as a tension-carrying member like a string or cable. The Assumptions of 2 and 7 are rather stringent. Thin-walled cross sections tend to be distorted and/or lose local stability, when subjected to significant plastic straining during repeated loading ([14,15]). Nevertheless, past experimental studies have shown a reasonable agreement with the basic formulation of the present theory under limited cycles of loading for negligible eccentricity and end restraint ([16]).

The displacement Δ of an end takes place relative to the other end along the original bar axis; Δ is taken positive for end separation so that the distance between the ends amounts to $L+\Delta$. The x-y coordinates are taken to indicate the deflection curve $y(x)$ with the origin at an end. Positive direction of y is taken opposite to the load eccentricity D (Fig. 1). Young's modulus is denoted by E, cross-sectional area by A, moment of inertia by I, limit moment in pure bending by M_o, and limit load in pure tension by N_o.

Derivation of Basic Equations

It is convenient to treat the load N as the independent variable and to determine the deformations as its functions with due account of the nonuniqueness and history-dependence of their relations. The relative displacement Δ is caused by change in the length of the bar due to elastic and plastic deformations, and by change in geometry accompanied by lateral deflection (See the Appendix for the separability of components.) The relative displacement is nondimensionalized by the elastic limit displacement. It follows that

$$\delta=\frac{EA\Delta}{N_oL}=\delta^e+\delta^p+\delta^t+\delta^g. \tag{1}$$

On the basis of Assumption 6, the axial force is approximated as constant along the bar axis, and equal to N. The component δ^e due to elastic deformation equals the dimensionless load $n\equiv N/N_o$. Plastic pure elongation induces the component δ^t and plastic deformation at yield hinges induces the component δ^p. These plastic components as well as the component δ^g due to geometry change depend on the deflected configuration. It turns out through examples of analytical behavior that in an initially straight bar yield hinges appear at the center or the ends or both, so that half of the deflection curve is determined from the elastic equilibrium equation which, on the basis of Assumption 6, reads

$$\frac{d^2\eta}{d\xi^2}-\nu^2\eta=e \quad \text{for } 0<\xi<1 \tag{2}$$

where

$$\xi=\frac{2x}{L}, \quad \eta=\frac{4EIy}{M_oL^2}, \quad \nu=\sqrt{\frac{NL^2}{4EI}} \tag{3}$$

and e is the ratio of the end moment to M_o. The combined action of the load N with eccentricity D and the resistant moment M_or at the end is replaced, based on Assumption 3, by a statically equivalent system of the centroidal axial load N and end moment $M_oe=M_or+ND$. Positive directions of the dimensionless quantities are indicated in Fig. 2. Because of symmetry, only the left half of the dimensionless deflection curve is shown. Let ϕ denote the angle of rotation at the end yield-hinge, and let ψ denote the slope angle at the point just to the right of the end hinge. Assumption 5 permits the dimensionless resistant moment r to be linearly related to the angle of end rotation $\psi-\phi$, so that

$$r=k_e(\psi-\phi)+k_p \tag{4}$$

where k_e and k_p are piecewise constant factors and are to be evaluated in each range of linearity. If the spiral spring is plastic, these factors depend on the history of loading in the context of yielding in plastic theory, taking distinct values for loading and unloading processes. For perfectly plastic action with limit moment M_p, for example, $k_e=0$ and $k_p=\pm M_p/M_o$.

Equation (2) is integrated with end conditions $\eta=0$ at $\xi=0$ and $d\eta/d\xi\to\psi$ as $\xi\to0$, to give

$$\eta=\psi\frac{\sinh(\nu\xi)}{\nu}+e\frac{\cosh(\nu\xi)-1}{\nu^2} \tag{5}$$

Note that ν is treated as known and, ψ and e are determined below, by noting that bending moments are known at acting yield hinges and that the corresponding rotations are predetermined in elastic processes. When the bending moment M_oc at the center

reaches a value such that together with the axial force, it satisfies the yield condition, plastic action takes place with a yield hinge at the center (see Fig. 2 for positive direction of c). In this plastic process c is expressed as a function of n on the basis of Assumption 3, and the boundary condition at the center $d^2\eta/d\psi^2 \to -c$ as $\xi \to 1$ reads

$$\psi \nu \sinh \nu + e \cosh \nu = -c. \tag{6}$$

Let 2θ be the angle of hinge rotation and θ is the slope angle at the point just to the left of the center in the dimensionless deflection curve. The hinge rotation during this process is determined from the condition that $d\eta/d\phi \to \theta$ as $\xi \to 1$, which becomes

$$\psi \cosh \nu + e \frac{\sinh \nu}{\nu} = \theta. \tag{7}$$

Equation (7) also serves as the boundary condition at the center for elastic action, in which θ is known from the preceding plastic straining and remains constant under varying load.

Similarly, when plastic action takes place at the end, e is given from the yield condition, and ψ can be found directly either from Eq. (6) or from Eq. (7), depending upon central plastic or elastic action, respectively, in terms of known quantities. During this process the angle of rotation ϕ at the end hinge is determined from the end condition

$$e = k_e(\psi - \phi) + k_p + nd \tag{8}$$

where $d \equiv N_o D/M_o$. For elastic action at the end, i.e., when the end hinge is inactive, the boundary condition is given by Eq. (8), in which ϕ is known from the preceding plastic straining; Eq. (8) is combined with Eq. (6) for plastic action at the center to provide

$$\psi = \frac{[k_e\phi - (k_p + nd)]\cosh \nu - c}{\nu \sinh \nu + k_e \cosh \nu} \tag{9}$$

$$e = \frac{(k_e + nd)\nu \sinh \nu - k_e(c + \phi \nu \sinh \nu)}{\nu \sinh \nu + k_e \cosh \nu} \tag{10}$$

and combined with Eq. (7) for elastic action at the center to provide

$$\psi = \frac{\theta + k_e\phi \dfrac{\sinh \nu}{\nu} - (k_p + nd)\dfrac{\sinh \nu}{\nu}}{\cosh \nu + k_e \dfrac{\sinh \nu}{\nu}} \tag{11}$$

$$e = \frac{k_e\theta + (k_p + nd - k_e\phi)\cosh \nu}{\cosh \nu + k_e \dfrac{\sinh \nu}{\nu}}. \tag{12}$$

With the deflection curve of Eq. (5) thus determined, δ^g is evaluated by noting that the total difference in length between the arc and chord of the entire deflection curve is twice the integration of $\sqrt{(dx)^2 + (dy)^2} - dx \approx (dy/dx)^2 dx/2$ from $x = 0$ to $x = L/2$ on the basis of Assumption 6. Thus,

$$\begin{aligned} \delta^g &= -\frac{A}{8EN_o}\left(\frac{M_oL}{I}\right)^2 \int_0^1 \left(\frac{d\eta}{d\xi}\right)^2 d\xi \\ &= -\frac{\alpha}{4n}\left[(\nu^2\psi^2 - e^2) + (\nu^2\psi^2 + e^2)\frac{\sinh(2\nu)}{2\nu} \right. \\ &\quad \left. + \psi e\{\cosh(2\nu) - 1\}\right] \end{aligned} \tag{13}$$

where $\alpha \equiv (A/I)(M_o/N_o)^2$ is a dimensionless parameter depending only on the cross section shape.

In the presence of axial force, plastic action at yield hinges causes not only hinge rotation or relative rotation across a hinge but also plastic axial deformation or relative displacement, which is responsible for the component δ^p. Suppose that the end yield-hinge is activated with the coordinates (e, n) of the dimensionless stress resultant changing to $(e + de, n + dn)$ by satisfying the yield condition. The value of δ^p changes accordingly by an amount proportional to the increment $d\phi$ in hinge rotation occurring during this process. The flow or normality rule of the theory of plasticity stipulates that the factor of proportionality depends on the stress state and equals $-(\alpha/2)(de/dn)$, which is a function of n according to Assumption 3. Similarly, the increment $d\theta$ of the central yield hinge induces an increment in δ^p during the change from (c, n) to $(c + dc, n + dn)$ at yield. By noting the same contribution from the other half of the bar, it is seen that

$$d\delta^p = -\alpha[e'(n)d\phi + c'(n)d\theta] \tag{14}$$

where the primes indicate differentiation with respect to the argument n. Thus δ^p is given by integrating Eq. (14) for the ranges where ϕ, and/or θ varies, i.e., for each plastic process. If, for example, plastic action takes place only at the bar center with θ varying from θ_c to θ for n varying from n_c to n, integration by parts leads to

$$\begin{aligned} \delta^p &= -\alpha \int_{\theta_c}^{\theta} c'(n)d\theta + \delta_c^p \\ &= \alpha\left[\theta_c c'(n_c) - \theta c'(n) + \int_{n_c}^{n} \theta c''(n)dn\right] + \delta_c^p \end{aligned} \tag{15}$$

where δ_c^p is an integral constant during this process and is determined from previous plastic processes starting from zero at the virgin state. In the same manner,

$$\begin{aligned} \delta^p &= -\alpha \int_{\phi_e}^{\phi} e'(n)d\phi + \delta_e^p \\ &= \alpha\left[\phi_e e'(n_e) - \phi e'(n) + \int_{n_e}^{n} \phi e''(n)dn\right] + \delta_e^p \end{aligned} \tag{16}$$

if plastic action takes place only at the bar ends with ϕ varying from ϕ_e to ϕ for n varying from n_e to n, where δ_e^p is an integral constant. Plastic action can take place both at the center and ends simultaneously. Both the components then add up to produce δ^p with a properly predetermined integral constant.

It is to be noted that the present formulation allows for some variables to take both signs. When N is negative, ν becomes imaginary and certain equations involve complex functions. Because the variables introduced in this paper are of physical meaning for negative load, these equations can be expressed in terms of real numbers and functions. In fact, this is accomplished by replacing $\cosh\nu$ by $\cos|\nu|$ and $\sinh\nu/\nu$ by $\sin|\nu|/|\nu|$; $n = 0$ is an isolated removable singularity, and hence continuity is assured by supplementing the values of the variables in the limit $n \to 0$ as the values for $n = 0$. This was discussed and confirmed in an earlier paper ([4]).

Plastic elongation causes a change $d\delta^t \geqslant 0$ in $\delta^t \geqslant 0$, without interaction with bending in a perfectly straight configuration. This is possible in the case of $d = 0$, with $n = 1$. This can take place also for $d \neq 0$ when $e = 0$ and $\psi = 0$, in which case the bar becomes straight with end kink $\phi \neq 0$, after undergoing plastic distortion. It may also be possible for the bar to remain straight under the yield load in pure compression. However, the bar will most likely buckle before undergoing appreciable plastic contraction except for extremely short or stubby bars, which are not considered in this paper. The buckling or bifurcation load is found by regarding centroidal compression as eccentric compression in the limiting case of small eccentricity, to be the smaller of the elastic buckling load and the yield load in pure compression. The elastic bucking load is determined from the condition for a nonzero deflection to occur with $d = \theta = \phi = k_p = 0$, to be $N_e \equiv 4EI\nu_e^2/L^2$, where ν_e is the solution of

$$\nu_e + k_e \tanh \nu_e = 0 \tag{17}$$

dependent upon the spring constant k_e.

Another deformation parameter of importance is the maximum deflection. The dimensionless central deflection v is found from rotational equilibrium of the half bar in Fig. 2, to be

$$v = -\frac{c+e}{n} \tag{18}$$

The aforederived basic equations are adequate to determine the hysteretic behavior of a restrained bar under any specified history of axial loading with fixed end eccentricity. It is to be noted, however, that load specification does not guarantee a unique determination of the corresponding deformation. Under existence of plasticity, loading and unloading processes bring about entirely different subsequent behavior; it is necessary to identify which path to follow at a bifurcation or fork. In contrast, specifying history of axial displacement determines the history of the load and hysteretic behavior.

Examples

Two examples are presented of the application of the basic equations for a specified history of axial displacement to show characteristic features of hysteretic behavior. The spiral spring is assumed to have an elastic-perfectly plastic moment-rotation relationship with elastic spring constant $2EIk/L$ and yield moment $m_p M_o$. Among components of δ in Eq. (1), δ^e equals the independent variable n, δ^g is given from Eq. (13) and δ^p from Eq. (15), or (16), or from both Eqs. (15) and (16). Therefore, unless δ increases under constant load N_o, δ is given from variables θ, ϕ, ψ, c and e, which are to be determined below.

Example 1. Consider repetitions with a fixed dimensionless displacement amplitude $\delta_A = 5$ in both sides. Let $k = 0.45$, $m_p = 0.5$, $N_e/N_o = 0.8$. Let eccentricity be absent, and $d = 0$ so that $r = e$. With weak-axis bending of a wide-flange cross section in mind, a piecewise-linear yield condition is assumed such that

$$|c| = 1 \quad (|n| \leqslant 0.5)$$
$$\frac{1}{2}|c| + |n| = 1 \quad (|n| \geqslant 0.5) \tag{19}$$

corresponding to a hexagonal yield curve, when expressed in terms of the central moment (see Fig. 3(*b*)). This approximation simplifies integration as in Eqs. (15) and (16), since the second derivative of moments with respect to the axial force is zero in plastic processes.

Numerical results are illustrated in Fig. 3 with $\alpha = 0.6$. For the abscissas of (*a*) to (*g*), quantities δ, c, e, θ, ϕ, ψ, and v are taken, respectively, as against the common ordinate of the dimensionless load n. Figure 3(*h*) shows the moment-rotation relationship for the end spring. Figure 4 is a sketch of deflection curves, in which a solid circle indicates a yield hinge, and a cross indicates yielding of the end spring. The encircled numerals, which are consistent through Figs. 3 and 4, indicate the hysteretic series in order. The same numerals are used in the subscripts of variables for corresponding states. The dotted lines in Figs. 3(*b*) and (*c*) indicate the yield curve. The dotted-and-dashed lines are drawn in Fig. 3(*c*) for yielding of end spring, i.e.,

$$e = \pm m_p. \tag{20}$$

Loading is initiated with tension. Linearly elastic response up to ① is described by the relation $\delta = \delta^e = n$. The process ①→② is plastic elongation with increasing δ^t. Displacement specification gives

$$\delta^t_2 - \delta^t_1 = \delta^t_2 = \delta_2 - \delta_1 = \delta_A - 1. \tag{21}$$

Unloading from ② is associated with elastic contraction, until the bar buckles at $n = -N_e/N_o$ at ③. Lateral deflection increases elastically as seen in Fig. 3(*g*), until plastic action takes place at ④, when the state of stress at the bar center reaches the yield limit as in Fig. 3(*b*). The load variation in ③→④ is small for bars of moderate slenderness ([4]), and neglected in the present formulation. Deflection is expressed in terms of end slope ψ, which is determined later. With $\theta = 0$, Eqs. (7) and (17) are combined to give

$$e = k_e \psi. \tag{22}$$

Equation (6) gives

$$c = -\frac{k_e \psi}{\cos \nu_e} \tag{23}$$

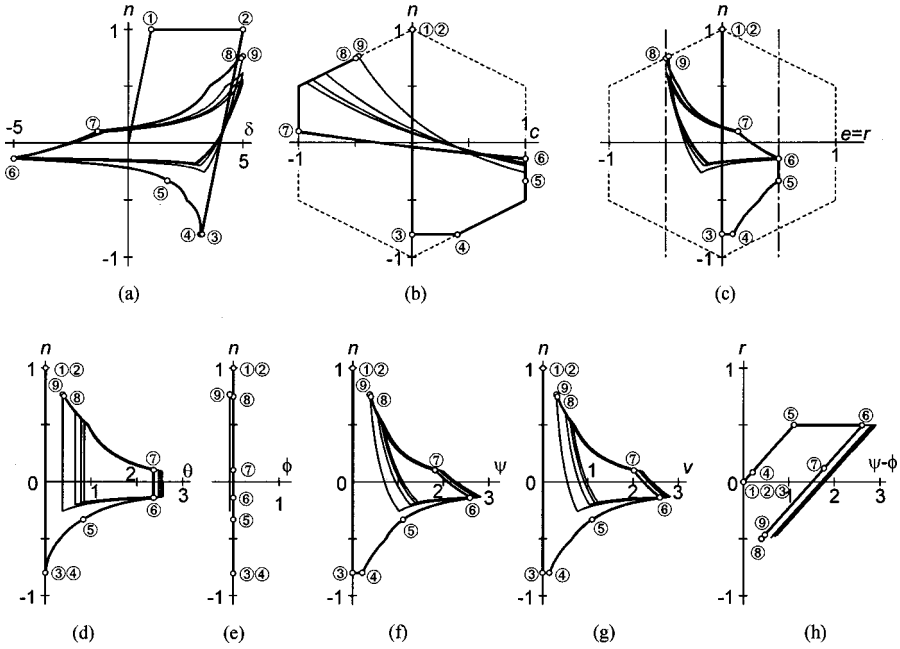


Fig. 3 Behavioral diagram of Example 1

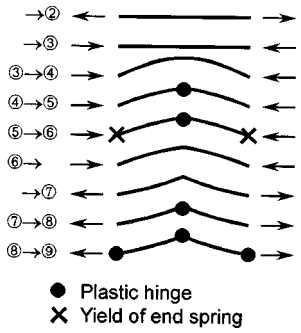


Fig. 4 Deflection curves for Example 1

and Eq. (13) gives

$$\delta^g = -\frac{\alpha}{4n}\psi^2 \nu_e \frac{\cos(2\nu_e)(\cos\nu_e \sin\nu_e - \nu_e)}{\sin\nu_e}. \tag{24}$$

It is only δ^g among the components of δ that varies in the process ③→④. Specified values of δ, therefore, determine ψ with recourse to Eq. (24), and hence determines e and c on the basis of Eqs. (22) and (23), respectively. The buckling load depends on the stiffness of the spring. The particular case of a pinned end is expressed by $k_e=0$, and leads to $e=0$, $c=\pi\psi/2$. End fixty $k_e=\infty$ leads to $e=c$ with $\cos\nu_e=-1$. For intermediate stiffness, $-1<\cos\nu_e<0$, and $c>e>0$. The process between ② and ④ is reversible.

Plastic action takes place at the bar center in the process ④→⑤. The yield condition determines c in terms of n. The stress trajectory follows the yield curve as shown in Fig. 3(*b*). With $\phi=0$, $k_p=0$ and $k_e=k$, Eq. (9) gives ψ, Eq. (10) gives e, and Eq. (7) gives θ in succession. With $\theta_4=0$, Eq. (15) gives the component

$$\delta^p = \alpha\left[\int_{n_4}^{n}\theta c''(n)dn - \theta c'(n)\right] \tag{25}$$

in which the integrand vanishes. Plastic action occurs not only at the center but also at the end spring from ⑤. The process ⑤→⑥ is followed similarly with relations $k_e=0$ and $k_p=m_p$.

The reversed displacement for the process ⑥→⑦ is realized by elastic recovery with $\theta=\theta_6$, $\phi=\phi_5=0$ and $\kappa_e=k$. The end spring undergoes residual rotation. Equation (8) is solved for k_p by noting that plastic rotation terminates at ⑥ to give

$$k_p = m_p - k(\psi_6 - \phi_5). \tag{26}$$

Equation (11) determines ψ, and Eq. (12) e. Equation (6) then determines c.

The central yield hinge is again activated at ⑦. The process ⑦→⑧ is analyzed in a manner similar to the process ④→⑤ in distinction with a tension as against a compression in the latter process. Additional plastic action starts at the bar end at ⑧ as well as at the central yield hinge. The yield condition gives c and e, Eq. (6) ψ, Eq. (7) θ, and Eq. (8) ϕ. Since θ and δ^p have remained constant during ⑥→⑦, Eq. (15) gives

$$\delta^p = \alpha\left[\theta_6 c'(n_7) - \theta c'(n) + \int_{n_7}^{n}\theta c''(n)dn\right] + \delta_6^p. \tag{27}$$

State ⑨ indicates another reversal in displacement loading. A cycle in displacement variation in ②→⑨ results in positive δ^p and hence in elongation of the bar ([5–7]). This effect gives rise to a reduction in load, as seen in the difference in the ordinates of ② and ⑨ in Fig. 3(*a*). This feature plays an important role in interpreting experimentally observed hysteretic behavior of steel members under axially repeated loading ([5–7,17]). Further repetitions of displacement cycling induce further reduction in the peak strength. Thin lines are drawn in Fig. 3 for five consecutive cycles after ⑨.

Example 2. A stubby bar of the case $N_e/N_o=4$ is considered. It is taken that $k=1.7$, $m_p=0.7$, and $d=1$. The yield condition is taken to express full plasticity for a rectangular cross section. It reads, e.g., for central moment

$$|c| + n^2 = 1. \tag{28}$$

It is taken that $\alpha=0.75$ corresponding to the rectangular cross section. Variation of δ is specified as $0\rightarrow2\rightarrow0\rightarrow3$. Numerical results are shown in Fig. 5. Variables r and e are distinguished due to eccentricity, which leads to slanting lines in Fig. 5(*c*) for yielding of the end spring, as described by

$$e = nd \pm m_p. \tag{29}$$

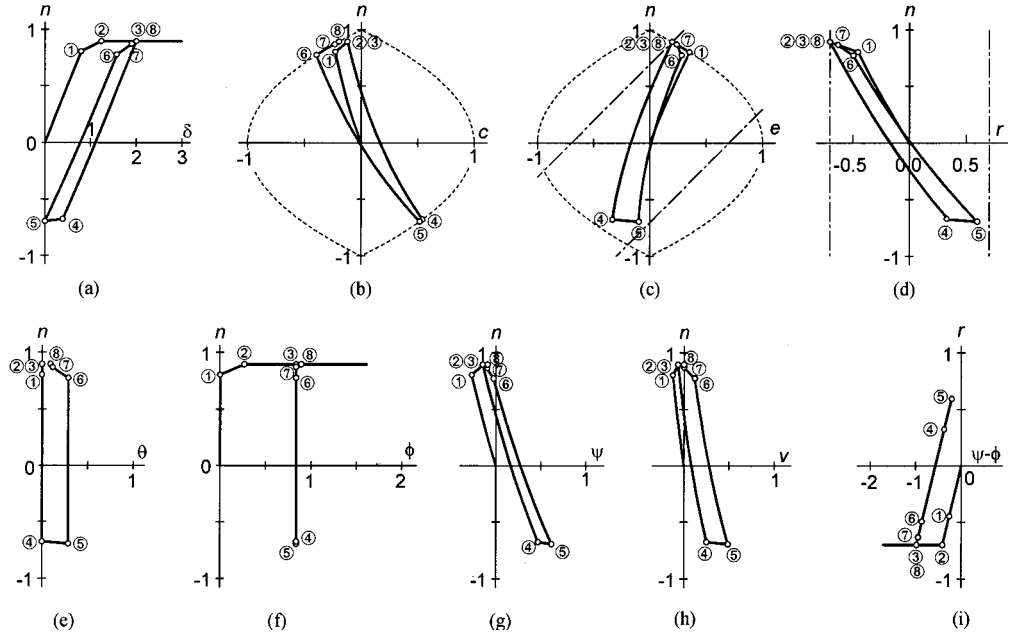


Fig. 5 Behavioral diagram of Example 2

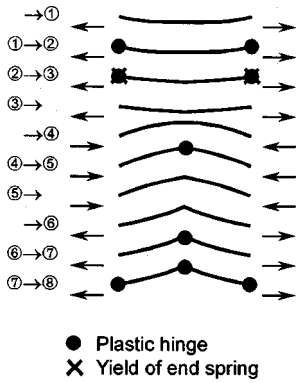


Fig. 6 Deflection curves for Example 2

This relation is derived by setting $k_e=0$ and $k_p=\pm m_p$ in Eq. (8). The state of stress of the bar end is represented by a point between this pair of parallel lines and within the yield curve of dotted lines. The fact that the allowable domain excludes the stress point (0, 1), gives $\delta^t=0$. Deflection curves are sketched in Fig. 6.

Eccentric tension induces bending so that the bar starts to elongate with a deflected configuration. The elastic behavior up to ① is determined simply with $k_e=k$, $k_p=0$, $\theta=0$, and $\phi=0$; ψ is given by Eq. (11), e by Eq. (12), and then c by Eq. (6); δ^e and δ^g are the only nonzero components of δ. Yielding takes place at the bar end after ①. The yield condition gives e in terms of n. Substitution of $\theta=0$ in Eq. (7) gives ψ; Eq. (8) gives ϕ and Eq. (6) c. Equation (16) provides

$$\delta^p=\alpha\left[\int_{n_1}^{n}\phi e''(n)dn-\phi e'(n)\right] \tag{30}$$

as the remaining nonvanishing component of δ. The end spring starts yielding at ② as well as at the bar end, having e of Eq. (29). The load cannot increase beyond the ordinate of ②. The end spring and bar end undergo plastic rotation in opposite directions. Since δ^e and δ^g remain constant, Eq. (14) gives plastic deformation at the unloading point ③ as

$$\delta_3^p-\delta_2^p=\delta_3-\delta_2=-\alpha e'(n_2)(\phi_3-\phi_2). \tag{31}$$

Specified δ therefore determines ϕ as well as δ^p. Deflection does not vary with constant ψ; $\psi_3=\psi_2$.

Elastic recovery from ③ is described by $\theta=\theta_3$ and $\phi=\phi_3$. The end spring returns elastic also with $k_e=k$. Spring factor k_p is found with recourse to Eq. (8), and hence from

$$e_3=-m_p+n_3d=k(\psi_3-\phi_3)+k_p+n_3d. \tag{32}$$

Equations (6), (11), and (12) give c, ψ and e, respectively. Deflection is so small during the elastic processes that the corresponding portions of the n-δ curve are almost straight.

The bar center yields at ④, with c determined from the yield condition (28). Since ϕ remains constant at ϕ_3, Eqs. (7), (9), and (10) give θ, ψ, and e, respectively. Change in the direction of varying δ at ⑤ gives rise to elastic behavior. Process ⑤→⑥ is analyzed with $\theta=\theta_5$ in a manner similar to ③→④.

After undergoing central hinging, both the center and end start yielding at ⑦. The yield condition gives c and e. Equation (6) provides ψ, Eq. (7) θ, and Eq. (8) ϕ. With $\theta_6=\theta_5$, and $\phi_7=\phi_3$, Eqs. (15) and (16) are combined to provide

$$\delta^p=\alpha\left[\theta_5c'(n_6)-\theta c'(n)+\int_{n_6}^{n}\theta c''(n)dn+\phi_3e'(n_7)-\phi e'(n)+\int_{n_7}^{n}\phi e''(n)dn\right]+\delta_6^p. \tag{33}$$

The process subsequent to ⑧ is close to that in ②→③, the difference being that the bar center yields with vanishing central deflection $v=0$, as shown in Fig. 5(*h*). Deflection remains constant, and so is the central hinge rotation. The end hinge is activated in addition to the end spring, and contributes to changes in δ^p and ϕ.

Concluding Remarks

Basic equations are first derived and it has been shown through examples that they are adequate to determine the hysteretic behavior of an elastic-perfectly plastic bar under a repeated eccentric axial load. Both ends are supported by piecewise-linear springs. Symmetry is assumed with respect to the bar center, with deflection in a definite plane due to restrictions set forth earlier. The primary disadvantage in this closed-form formulation is that the independent variable is not the displacement but the load. There is a need to judge which path to follow, whenever bifurcation takes place, such as at reversals in displacement direction, or at the initiation and termination of yielding. Restriction on the magnitude of deflection is not stringent in engineering application. Small deflection analysis as in this formulation retains sufficient accuracy, unless lateral deflection reaches a magnitude of the original length ([4]).

Appendix

Separability of Displacement Components and Effects of Large Deformation. It has been taken for granted in the text that the relative displacement is divided into four components. The validity of this treatment was confirmed previously ([3]), and is repeated here briefly with due regard to the effect of large deformation.

When the length changes from L_o to L, it follows from the definition

$$L_o+\Delta=\int_0^{L_o+\Delta}dx$$

that

$$\Delta=\int_0^L\cos\Phi ds-L_o=\left(\int_0^L\cos\Phi ds-L\right)+(L-L_o)$$

where s is the arc length measured along the deflection curve from the origin of the Cartesian coordinates (x, y) and Φ is the slope angle. The quantity in the first pair of parentheses is caused by a change in geometry, and is the displacement which would take place if a straight bar of length L were deflected laterally; the quantity in the second pair of parentheses is the change in length. The separability of the changes in the geometry and in the length is thus established without any restriction on the magnitude of deformation.

When deflection is so large that the square of the slope in the deflection curve attains a magnitude of the order of unity, the curvature cannot be approximated by the second derivative of y with respect to x, and pertinent distinction has to be made between the x and s-coordinates. Equations (2) and (13) are no longer valid in theory. Exact analysis of large deformation has shown that this condition of large deflection is realized when the decrement in the distance between the bar ends attains a magnitude of the order of the bar length, and, through numerical examples of moderate slenderness, the effect of large deflection has been found not to be significant even under such a condition. Exceptional cases are for extremely slender bars or bars deflected into loops ([3,5–7]).

Separability of the change in the length into elastic and plastic components is a basic premise in the theory of plasticity. Further separation of the plastic deformation into two components is made, because a change in the component δ^p is accompanied by changes in load and other variables, as against the component δ^t. The latter is associated with a straight configuration; unlimited

plastic elongation can take place under $N=N_o$, being controlled only by displacement constraint at the ends of the bar. When elongation attains a magnitude of the order of L_o before violating the condition of Assumption 7, then the length L has to replace L_o in the consideration of geometry of the deformed bar. A simple account of this effect as well as the consequent changes in cross-sectional dimensions was made previously by resorting to the assumption of the similar figure of cross section together with plastic incompressibility ([3]). As a result it has been found that by reducing the buckling load in the proportion of $(L_o/L)^4$, the weakening effect of great elongation is somewhat significant compared with the effect of large deflection, under alternate repeated loading with the same extreme in the tension and compression sides ([5–7]).

References

[1] Nonaka, T., 1973, ''An Elastic-Plastic Analysis of a Bar Under Repeated Axial Loading,'' Int. J. Solids Struct., **9**, pp. 569–580, (with Erratum in No. 10).

[2] Nonaka, T., 1977, ''Approximation of Yield Condition for the Hysteretic Behavior of a Bar Under Repeated Axial Loading,'' Int. J. Solids Struct., **13**, pp. 637–643.

[3] Nonaka, T., 1977, ''An Analysis for Large Deformation of an Elastic-Plastic Bar Under Repeated Axial Loading, Part—I Derivation of Basic Equations,'' Int. J. Mech. Sci., **19**, pp. 619–627.

[4] Nonaka, T., 1977, ''An Analysis for Large Deformation of an Elastic-Plastic Bar Under Repeated Axial Loading, Part—II Correlation with Small Deformation Theory,'' Int. J. Mech. Sci., **19**, pp. 631–638.

[5] Nonaka, T., 1983, ''Closed Form Formulation for the Hysteretic Behavior of a Bar Under Repeated Axial Loading, Part 1—Derivation of Basic Equations,'' Trans. Arch. Inst. Jpn., No. **334**, pp. 1–8 (in Japanese).

[6] Nonaka, T., 1984, ''Closed Form Formulation for the Hysteretic Behavior of a Bar Under Repeated Axial Loading, Part 2—Analytic Solution,'' Trans. Arch. Inst. Jpn., No. **338**, pp. 29–35 (in Japanese).

[7] Nonaka, T., 1984, ''Closed Form Formulation for the Hysteretic Behavior of a Bar Under Repeated Axial Loading, Part 3—Plastic Straining and Characteristic Features,'' Trans. Arch. Inst. Jpn., No. **343**, pp. 42–50 (in Japanese).

[8] Prathuangsit, P., Goel, S. C., and Hanson, R. D., 1978, ''Axial Hysteresis Behavior With End Restraints,'' J. Struct. Div. ASCE, **104**, No. ST6, pp. 2261–2277.

[9] Fukuta, T., and Yamanouchi, H., 1986, ''Post-Buckling Behavior of Steel Braces With Elastically Restrained Ends,'' J. Struct. Construct. Eng., Trans. Arch. Inst. Jpn., No. **364**, pp. 10–22 (in Japanese).

[10] Mitani, I., 1978, ''An Elastic-Plastic Analysis of a Restrained Steel Bar Under Repeated Eccentrical Axial Loading,'' Trans. Arch. Inst. Jpn., No. **274**, pp. 65–73 (in Japanese).

[11] Higginbotham, A. B., and Hanson, R. D., 1976, ''Axial Hysteresis Behavior of Steel Members,'' J. Struct. Div. ASCE, **102**, No. ST7, Proc. Paper 12245, pp. 1365–1381.

[12] Kahn, L. F., and Hanson, R. D., 1976, ''Inelastic Cycles of Axially Loaded Steel Members,'' J. Struct. Div. ASCE, **102**, No. ST5, Proc. Paper 12111, pp. 947–959.

[13] Nonaka, T., 1972, ''A Kinematic Approach to the Plastic Collapse Load of a Rectangular Frame With Consideration of Axial Forces,'' *Proc., Symposium on Plastic Analysis of Structures*, Iassy, Rumania, Sept., pp. 192–220; *Theoretical and Applied Mechanics*, Vol. 23, University of Tokyo Press, Tokyo, pp. 301–315.

[14] Nonaka, T., and Iwai, S., 1989, ''Failure of Bar Structures Under Repeated Loading,'' *Structural Failure*, T. Wierzbicki and N. Jones, ed., John Wiley and Sons, New York, pp. 389–433.

[15] Iwai, S., Nonaka, T., Park, Y.-S., and Kameda, H., 1993, ''Experimental Observation on Fatigue of Steel Members and Elements Under Very Low Load Cycles,'' *Structural Dynamics-Eurodyn'93* (Proc. 2nd Euro. Conf. on Structural Dynamics, Tronheim), Vol. 2, A. A. Balkema, Rotterdam, pp. 771–778.

[16] Nonaka, T., 1989, ''Elastic-Plastic Bar Under Changes in Temperature and Axial Load,'' Jour. of Structural Engineering, ASCE, **115**, No. 12, pp. 3059–3075.

[17] Wakabayashi, M., Nonaka, T., Nakamura, T., Morino, S., and Yoshida, N., 1973, ''Experimental Studies on the Behavior of Steel Bars Under Repeated Axial Loading—Part 1, Rectangular Cross-Section,'' Bulletin, Disaster Prevention Research Institute, Kyoto University, No. 16B, pp. 113–125.

K. Kumar
Professor,
Department of Aerospace Engineering,
Indian Institute of Technology,
Kanpur, India

J. Botsis
Mem. ASME,
Professor and Director,
Laboratory of Applied Mechanics
And Reliability Analysis,
Swiss Federal Institute of Technology,
Lausanne, Switzerland

Contact Stresses in Multilayered Strands Under Tension and Torsion[1]

First, an attempt is made to experimentally test the validity of the linear deformation derivative results earlier developed for the multilayered wire-rope strands under tension and torsion. The theoretical results are next utilized to obtain analytical expressions for the maximum contact stresses induced in the multilayered strands with metallic wire core. These closed-form solutions provide some useful design insights into the influence of several important cable parameters and material properties on the resulting contact stresses. The strong influence of the material modulus of elasticity on the critical stresses is highlighted. Significantly, the analysis brings out how the contact stresses can rise to an order of magnitude higher levels than that of the nominal stresses.
[DOI: 10.1115/1.1355777]

Introduction

The problem of stresses and strains in conventional cables under static loads has been analyzed by several authors. The earlier theoretical investigations were extensively reviewed by Costello [1] and later by Kumar and Cochran [2]. These analyses are characterized by several simplifying assumptions of questionable validity. Costello and Phillips [3–5] gave a new direction to these studies by adopting a more comprehensive and fundamental approach to analyze the static behavior of the cables. They treated the cables as a group of separate curved rods in the form of helices. Their ''linearized'' deformation relations are applicable to the the wire-ropes of arbitrary cross sections [6,7]. These approximate relations lead to considerable economy of computational time in the study of the cable characteristics although the analysis remains essentially numerical in character. Jiang, et al. [8] recently showed the validity of these relations through their finite element analysis. They also found excellent agreement between the deformation results of the wire-rope strands under axial tensile loads based on Costello's analysis and experimental results of Utting and Jones [9,10].

Kumar and Cochran [2,11] introduced some additional approximations based on order-of-magnitude considerations while applying Costello's ''linear theory'' to the wire-rope analysis. This approach was instrumental in achieving considerable simplifications and enabled an analytical prediction of the deformation characteristics of single strand cables with metallic as well as fibrous cores. Later, Kumar, et al. [12] also developed analytical expressions for critical contact stresses induced by tension and torsion in single strand cables with fibrous cores. It may be pointed out, however, that so far no attempt appears to have been made for experimental verification of these theoretical results.

In this paper, first, an attempt is made to experimentally test the validity of the deformation derivative results earlier obtained for multilayered wire-rope strands with metallic core. Subsequently, these earlier deformation relations are utilized for closed-form prediction of the resulting maximum contact stresses in these cables under tension and torsion.

[1]The work for this paper was carried out at the Laboratory of Applied Mechanics and Reliability Analysis of the Swiss Federal Institute of Technology, Lausanne.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 12, 1999; final revision, Feb. 24, 2000. Associate Editor: J. W. Ju. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Summary of Earlier Deformation Derivative Results

Let us consider a single strand cable made up of n successive layers of helical wires. Let the layer i have m_i wires, each of radius R_i, helix radius r_i, and helix angle α_i. Since the metallic core is treated as the first layer in the analysis, we have

$$m_1 = 1 \quad \text{and} \quad \alpha_1 = \pi/2.$$

As discussed in the earlier paper [2], the diameter of the core is assumed to be large enough to prevent the helical wires in the same layer from touching each other. Instead, the wires in a layer remain in contact with those in the adjacent layers only.

The deformation relations earlier obtained by Kumar and Cochran [2] for the single strand cable under consideration may be summarized as follows:

$$\hat{F} = F/(A_n E) = F_\epsilon \epsilon + F_\beta \beta$$

$$\hat{M} = M/(ER^3) = M_\epsilon \epsilon + M_\beta \beta \tag{1}$$

where

$$F_\epsilon = \pi \Sigma_{i=1}^{n} [m_i R_i^2 \sin\alpha_i (\sin^2\alpha_i - \nu \cos^2\alpha_i)]/A_n$$

$$F_\beta = \pi \Sigma_{i=1}^{n} [m_i R_i^2 (r_i/R) \sin^2\alpha_i \cos\alpha_i]/A_n$$

$$M_\epsilon = \pi \Sigma_{i=1}^{n} [m_i \hat{R}_i^3 \cos\alpha_i \{(r_i/R_i)(\sin^2\alpha_i - \nu\cos^2\alpha_i) - (1/4) R_i/r_i\}]$$

$$M_\beta = \pi \Sigma_{i=1}^{n} [m_i \hat{R}_i^4 \sin\alpha_i \{(r_i/R_i)^2 \cos^2\alpha_i + (1/4)(1+\nu_f \sin^4\alpha_i \cos 2\alpha_i)\}] \tag{2}$$

and

A_n = metallic area of cross-section; $\Sigma_{i=1}^{n}(m_i \pi R_i^2)$
E = Young's modulus of elasticity of the cable material
F = tensile Force, cable is subjected to
L = cable length for the test specimen
M = torsional moment, cable is subjected to
R = cable radius; $(R_1 + 2\Sigma_{j=2}^{n-1} R_j + R_n)$
R_i^* = $(R_i + r_i)$; $i = 2,3,..n$
$\hat{R}_i$ = R_i/R; $i = 1,2,..n$
n = number of layers of wires in the cable including the core
r_i = helix radius for wires in the ith layer; $(R_1 + 2\Sigma_{j=2}^{i-1} R_j + R_i)$; $i = 2,3,..n$
ϵ = tensile strain for the cable; $\Delta L/L$
β = torsional strain for the cable; $R\partial\phi/\partial L$

ΔL = elongation
ϕ = angular twist over the cable length
ν = Poisson's ratio of the cable material
$\nu_f = \nu/(1+\nu)$.

It may be pointed out that the above analytical relations for the load deformation derivatives are based on several simplifying approximations. Of these, the expressions for the cross-derivatives F_β and M_ϵ do not exactly satisfy the equations relating them as earlier developed by Costello [13] from consideration of ''independence of path'' for work done during loading. This was not unexpected, however.

Experimental Validation of Deformation Relations

Materials. The single-strand wire-rope experimentally investigated was obtained from Brugg Cables Acier SA (CH-1023, Crissier, Switzerland). It had a metallic core and one surrounding layer of six wires with a common helix angle of 80.4 deg and an overall diameter of 9 mm (Fig. 1). The core and the helical wires had a diameter of 3 mm each. The Young's modulus of the material was determined by testing the core of the rope and was found to be 157 GPa.

Experimental Methods. Specimens with length of 500 mm were cut from the as-received rope. To grip the specimens, special hollow steel cylinders, 50 mm in length, were machined with an internal diameter of about 9 mm, equal to the effective diameter of the rope, and an outside diameter of 15 mm to fit the hydraulic grips of the MTS machine. To assure firm mounting of the specimens and avoid any slippage of the rope during testing, the ends of the rope were brasured at the outside ends of the steel cylinders. Figure 1 shows the schematic diagram of a typical specimen along with the cable cross section.

The experiments were performed on an MTS 809 Servohydraulic axial-torsional machine with a TestWare-SX software to record the data. Three types of experiments were carried out: a simple tension, a simple torsion, and the combined tension-torsion. All experiments were deformation controlled with rates of 1 mm/min in tension and 20 deg/min in torsion. As explained at the end of

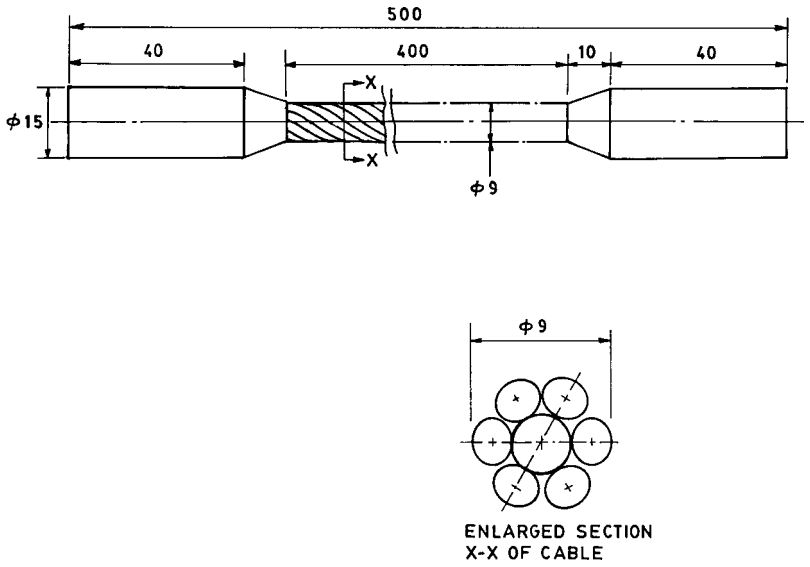


Fig. 1 Schematic diagram showing test specimen along with the cable cross section

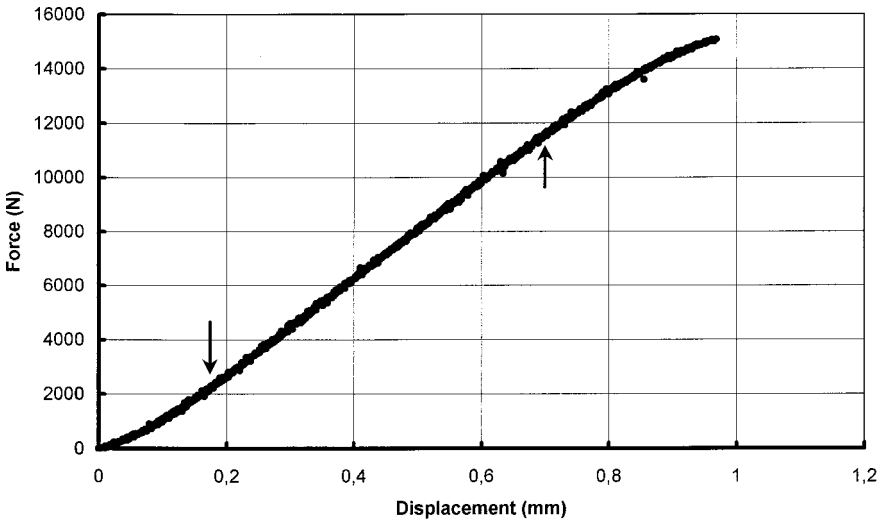


Fig. 2 Load-deformation curve obtained through simple tension test

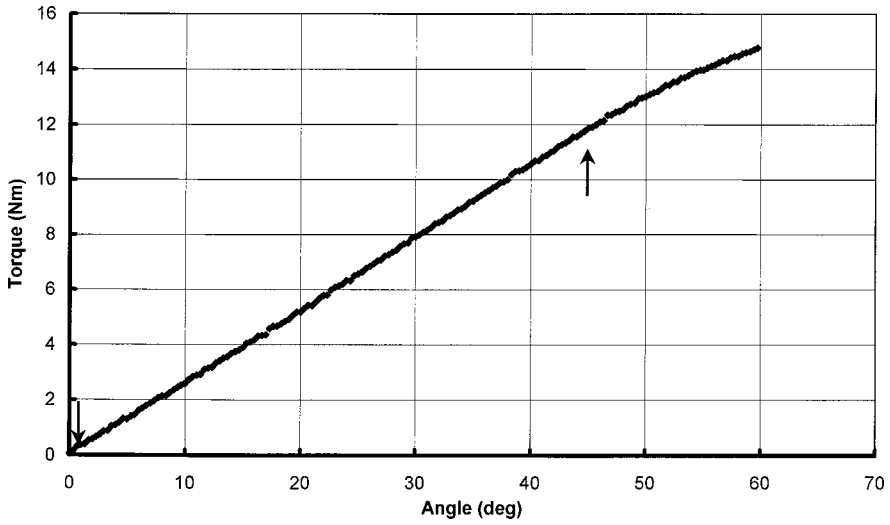


Fig. 3 Torque-deformation curve obtained through simple torsion test

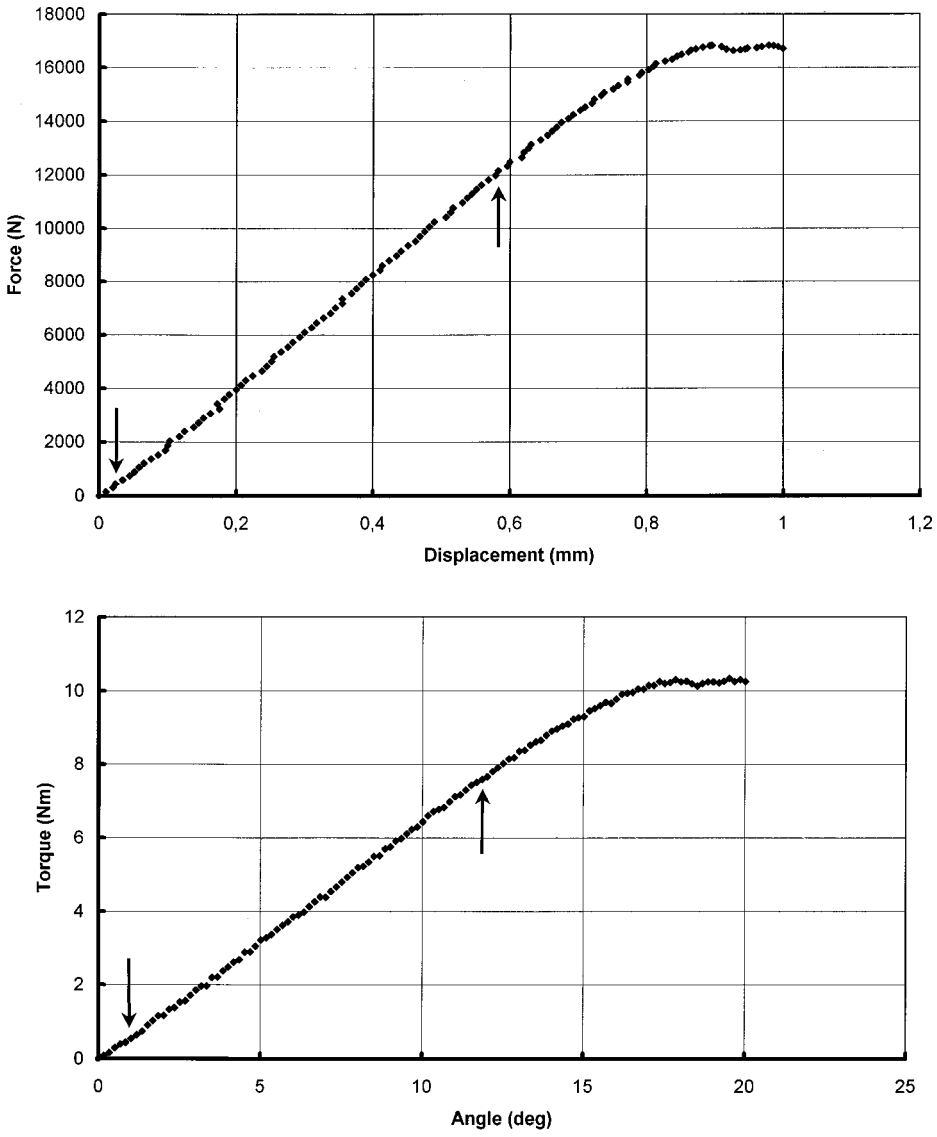


Fig. 4 (*a*) Force-displacement plot obtained through combined tension and torsion test, (*b*) torque-angle plot obtained through combined tension and torsion test

Table 1 Comparison between analytical and experimental values for the deformation derivatives of a simple 7×1 cable

Deformation Derivatives	F_ϵ	F_β	M_ϵ	M_β
Theoretical Values	0.96	0.093	0.210	0.079
Experimental Values	0.92	0.107	0.215	0.095

this section, a judicious choice of these rates is rather important for achieving higher accuracies in the calculations of the parameters. The tension and torsional loads were recorded as a function of displacement and the angular twist. The tests were repeated for various traction and/or torsion conditions. The excellent reproducibility observed established the reliability of the experimental data. The influence of loading rates on the deformation characteristics, however, was not considered in the present investigation.

Figures 2–4 show the experimentally obtained load-deformation curves. The first curve presented in Fig. 2 shows the variation of tensile force with increasing displacement for the experiment wherein no torsional twist was permitted. The slope of the ''linear'' segment of this curve—indicated by the two arrows—was obtained using least square fit. The slope of this straight line is required for computation of the deformation derivative F_ϵ using the relation

$$F_\epsilon=[L/(AE)].(\partial F/\partial L).$$

Figure 3 represents the variation of the torsional moment with angular twist for the next experiment with fixed specimen length. The slope of the ''linear'' part of the curve shown between the arrows is utilized to compute the derivative M_β as follows:

$$M_\beta=[L/(ER^4)].(\partial M/\partial \phi).$$

The tensile and torsional loads are now applied to a third specimen when the displacement as well as angular twist were controlled to increase simultaneously at suitably selected constant rates, as pointed out earlier. The data thus gathered are presented as force versus displacement (Fig. 4(*a*)) and the torsional moment versus angular twist (Fig. 4(*b*)). The slopes of the straight line fits for the ''linear'' parts lying between the two arrows are needed to obtain the remaining two deformation derivatives F_β and M_ϵ. These ''cross derivatives'' are evaluated using the following relations deduced from Eqs. (1):

$$F_\beta=[\{L/(AE)\}.(\partial F/\partial L)-F_\epsilon].(1/R).(\dot{L}/\dot{\phi})$$

$$M_\epsilon=[\{L/(ER^4)\}.(\partial M/\partial \phi)-M_\beta].R.(\dot{\phi}/\dot{L}).$$

It may be pointed out that the expressions within the square brackets in the numerators in these two relations for F_β and M_ϵ appear as a difference between two terms which are in general of the same order of magnitude. Furthermore, in each case, the tensile and torsional loading rates directly influence the two magnitudes and their relative closeness and hence are likely to have significant influence on sensitivity of the experimental parameter values thus computed. It is therefore imperative to choose the tensile and torsional rates judiciously so that the two terms are ''wide apart'' so as to ensure higher computational accuracies.

Comparison of Experimental and Analytical Results. The theoretical values of the force and moment deformation derivatives are computed for the 7×1 single strand cable used for the investigation. The values of these derivatives are also obtained from the estimated slopes of the ''linear segments'' of the set of the load-deformation curves generated experimentally as described earlier. A comparison of the theoretical and corresponding experimental results is presented in Table 1. Evidently, the theoretical results based on the Costello's ''linear theory'' and the additional simplifying approximations of Kumar and Cochran [2] compare quite favorably with the corresponding experimental values. The discrepancy between theoretical and experimental results is found to be maximum for the parameter M_β. Even here, the error seems to be well within the normally acceptable limits. We therefore conclude that the approximate analytical procedure earlier followed is reasonably accurate for design and other practical applications. Next, an attempt is made to utilize the closed-form deformation results for predicting the maximum normal contact stresses in the multilayered strands with metallic core.

Analysis for Contact Forces and Stresses

Contact Forces. On consideration of equilibrium of helical wires, the normal force per unit length on the wires (Q_i) as caused by the normal and binormal force components in the ith layer through contact with those in the adjacent inner layer, can be written as ([2])

$$Q_i=T_{ia}(\cos^2\alpha_i'/r_i')-T_{ib}(\sin\alpha_i'\cos\alpha_i'/r_i') \tag{3}$$

where

$$\begin{aligned}
T_{ia} &= [\pi R_i^2 E(\epsilon-\Delta\alpha_i\cot\alpha_i)]\\
T_{ib} &= [\pi R_i^4 E/(4r_i^2)\cos^2\alpha_i][(\Delta\alpha_i(1-\nu_f\cos 2\alpha_i)\\
&\quad +\nu_f\delta_i\sin\alpha_i\cos\alpha_i]\\
\Delta\alpha_i &= \text{increment in the helix angle of wires in the } i\text{th layer under loads}\\
&= \epsilon(1+\nu)(1-\nu_i)\sin\alpha_i\cos\alpha_i-\beta(r_i/R)(1-\nu_i)\sin^2\alpha_i\\
\delta_i &= -\nu\epsilon+(\nu/r_i)[\Sigma_{j=2}^{i-1}(2\Delta\alpha_j R_j\cot\alpha_j)+\Delta\alpha_i R_i\cot\alpha_i]\\
\nu_j &= [\nu(R_j/r_j)\cos^2\alpha_j]\\
(.)' &= (.)\text{ after deformation under applied loads.}
\end{aligned}$$

On substitution of $\Delta\alpha_i$ from above and making several simplifying approximations based on order of magnitude of the various terms, the earlier expression for the contact force per unit length can be written in the form

$$Q_i=[\pi(R_i^2/r_i)]E\cos^2\alpha_i[F_\epsilon M_\beta-M_\epsilon F_\beta]^-[A_{Fi}\hat{F}+A_{Mi}\hat{M}];$$
$$i=2,3,..,n \tag{4}$$

where

$$\begin{aligned}
A_{Fi} &= [M_\beta(\sin^2\alpha_i-\nu\cos^2\alpha_i)-M_\epsilon(r_i/R)\sin\alpha_i\cos\alpha_i.\\
&\quad \{1+0.25(R_i/r_i)^2\sin^2\alpha_i\}].\\
A_{Mi} &= [-F_\beta(\sin^2\alpha_i-\nu\cos^2\alpha_i)+F_\epsilon(r_i/R)\sin\alpha_i\cos\alpha_i.\\
&\quad \{1+0.25(R_i/r_i)^2\sin^2\alpha_i\}].
\end{aligned}$$

It may be pointed out that the interaction between wires of the adjacent layers is characterized by two distinct types of contacts. The wires in the second layer touch the metallic core continuously along a helical line. In contrast, the wires in the third and higher layers cross their neighbors in the adjacent layers—above and below—at an angle and thus make only ''point contacts.'' The two situations, being characteristically different, are treated separately.

Case 1 Contact Forces Between Core and the Adjacent Layer, i.e., for i=2. Here, the wires in the second layer make a ''line contact'' with the metallic core. In order to use the Hertz theory for determination of the contact stresses, the expression for the corresponding effective normal contact force per unit length on wires in the second layer due to axial and binormal force components, Q_{2-n}, is estimated as

$$Q_{2-n}=\Sigma_{j=2}^{n}[m_j Q_j/\sin\alpha_j][\sin\alpha_2/m_2]. \tag{5}$$

Case 2 Contact Forces in the Outer Layers, i.e., for i=3..,n In contrast, these layers are characterized by ''point contacts'' and hence the normal compressive forces induced at the points of contact are determined first. For the outermost layer of wires, i.e., $i=n$, the effective normal force at the contact points is estimated by multiplying Q_i by the length of the wire segment in the ith layer between its two successive points of contact with the wires in the adjacent inner layer ([14]). Hence,

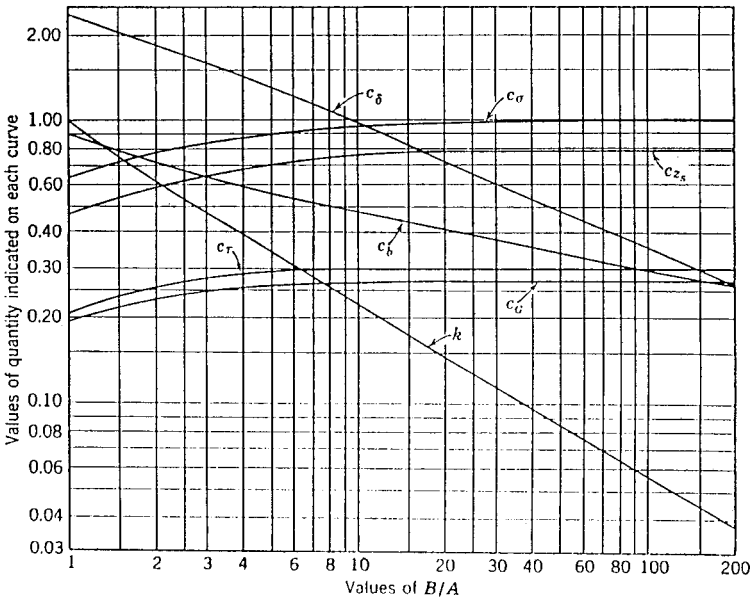


Fig. 5 Stresses and deflections between two bodies in contact at a point ([15])

$$P_n = Q_n[(2\pi R^*_{n-1}/m_{n-1})\sin\alpha_{n-1}/|\sin(\alpha_n-\alpha_{n-1})|]. \quad (6)$$

It may be emphasized that for the general ith layer lying inside the outermost one, the normal contact forces develop due to contributions through axial and binormal force components induced in wires in the layers i, $(i+1)$, $(i+2)$.. and n. The resulting normal force, P_i, generated at the ''contact points'' between the wires in this layer and the adjacent inner layer is obtained as ([14])

$$P_i = \Sigma^n_{j=i}[m_j Q_j/\sin\alpha_j].[(2\pi R^*_{i-1}\sin\alpha_i \sin\alpha_{i-1}]/ [m_i m_{i-1}\sin|\alpha_i-\alpha_{i\text{-}1}|] \quad i=3,.,n. \quad (7)$$

Contact Stresses

Case 1 Contact Stresses Between Core and the Adjacent Layer, i.e., for i=2. A direct application of the Hertz theory of contact stresses leads to the following relation for maximum compressive stresses in the region of contact between the wires in the second layer and the metallic core wire

$$\sigma_2 = \sqrt{E}\sqrt{Q_{2-n}[(\sin^2\alpha_2/R_1)+(1/R_2)]}/\sqrt{[2\pi(1-\nu^2)]} \quad (8)$$

In dimensionless form, the above result simplifies to:

$$\hat{\sigma}_2 = \sigma_2/\sqrt{E\sigma_{nom}} = \sqrt{Q_{2-n}[(\sin^2\alpha_2/R_1)+(1/R_2)]/\sigma_{nom}}/\sqrt{[2\pi(1-\nu^2)]} \quad (9)$$

where

$$\sigma_{nom} = (F/A_n).$$

Case 2 Contact Stresses in the Outer Layers, i.e., for i=3..,n As was pointed out earlier, the helical wires in the third and other outer layers make ''point contacts'' with those in their inner adjacent layers. The corresponding expressions for the maximum compressive stresses at these ''points,'' also given by Hertz theory, however, are relatively more involved and are given by ([15])

$$\sigma_i = c_\sigma(b/\Delta) \quad (10)$$

where

a = semi-major axis of the contact ellipse

b = $[1.5kE(k')P_i\Delta/\pi]^{1/3}$; semi-minor axis of the contact ellipse

c_σ = a variable fraction dependent on d, Fig. 5 ([15])

d = ratio (B/A) also equal to $[(1/k^2)E(k')-K(k')/[K(k')-E(k')]$

k = (b/a)

k' = $\sqrt{(1-k^2)}$

A = $(1/4)[(1/R_i)-(\cos^2\alpha_i/R^*_{i-1})+(1/R_{i-1})+(\cos^2\alpha_{i-1}/R^*_{i-1})]-(1/4)[(1/R_i+\cos^2\alpha_i/R^*_{i-1}+1/R_{i-1}-\cos^2\alpha_{i-1}/R^*_{i-1})^2-4(1/R_i-\cos^2\alpha_i/R^*_{i-1})\times(1/R_{i-1}+\cos^2\alpha_{i-1}/R^*_{i-1})\sin^2\psi]^{1/2}$

B = $(1/4)[(1/R_i)-(\cos^2\alpha_i/R^*_{i-1})+(1/R_{i-1})+(\cos^2\alpha_{i-1}/R^*_{i-1})]+(1/4)[(1/R_i+\cos^2\alpha_i/R^*_{i-1}+1/R_{i-1}-\cos^2\alpha_{i-1}/R^*_{i-1})^2+4(1/R_i+\cos^2\alpha_i/R^*_{i-1})\times(1/R_{i-1}-\cos^2\alpha_{i-1}/R^*_{i-1})\sin^2\psi]^{1/2}$

$E(k')$ = complete elliptic integral of second kind

$K(k')$ = complete elliptic integral of first kind

ψ = $(\alpha_i-\alpha_{i-1})$

Δ = $[2(1-\nu^2)/E]/(A+B)$.

For applications, the value of d first determined from the wire-rope data is made use of to read the corresponding values of k, k', and c_σ from the plots in Fig. 5 ([15]). The need for such graphical estimation of c_σ appears to pose a major challenge, especially in view of the intended analytical representation of the contact stresses. To overcome this difficulty, an attempt is made to first obtain an explicit approximate expression for d by solving the earlier equation relating it with k, k', and the two elliptic integrals of k'. Undertaking the series expansions of the elliptic integrals ([16]), the expression obtained for this variable can be written as

$$d = (c_d/k)^2 \quad (11)$$

where

$$c_d = [1+(1/2)^2(1/2)k'^2+(1/2.3/4)^2(1/3)k'^4+.]/[1+(1/2)^2(3/2)k'^2+(1/2.3/4)^2(5/3)k'^4+].$$

Although the fraction c_d is itself a function of k' and hence k, it may be observed that its values would always lie within a narrow band. Therefore, it is possible to take

$$k=(c_d/\sqrt{d}) \tag{12}$$

while treating the parameter c_d as a constant, it being only marginally sensitive to the values of d. On substitution of this value of k, the expression for the maximum contact stresses takes the form

$$\sigma_i=c_o P_i^{1/3} d^{-1/6}\Delta^{-2/3} \tag{13}$$

where

σ_i = maximum contact stresses on contacting wires in ith and $(i-1)$th layers

$c_o = c_\sigma[0.75c_d(1-k'^2/4-3k'^4/64-\ldots.)]^{1/3}$.

It so turns out that for the entire range of d-values corresponding to the practical values of ψ lying in the interval $(0,\pi/3)$, the coefficient c_o hovers around the value (2/3) within less than five percent. Accepting this mean value of c_o, relatively much simpler expression is obtained for the contact stresses as indicated below:

$$\sigma_i=(2/3)P_i^{1/3}d^{-1/6}[2(1-\nu^2)/\{(A+B)E\}]^{-2/3} \tag{14}$$

where

$$P_i=E\pi(F_\epsilon M_\beta-M_\epsilon F_\beta)^{-1}[\hat{F}\Sigma_{j=i}^{n}\{A_{F_j}(m_jR_j^2/r_j)\times(\cos^2\alpha_j/\sin\alpha_j)\}+\hat{M}\Sigma_{j=i}^{n}\{A_{M_j}m_j(R_j^2/r_j)\times(\cos^2\alpha_j/\sin\alpha_j)\}].$$

Example

For a better appreciation of the influence of various parameters on the contact stresses, it is now proposed to take up a particular example of a three-layered cable with the geometry and the corresponding deformation derivative data given as follows ([2]):

$$R_1=R_2=R_3=a;$$

$$m_1=1,\quad m_2=6,\quad m_3=12;$$

$$\alpha_1=\pi/2,\quad \alpha_2=\alpha(\text{right lay}),\quad \alpha_3=(\pi-\alpha)\ (\text{lang lay});$$

$$F_\epsilon=(1/19)[1+18\sin\alpha(\sin^2\alpha-\nu\cos^2\alpha)];$$

$$F_\beta=-(36/95)\sin^2\alpha\cos\alpha;$$

$$M_\epsilon=-(36\pi/125)(\sin^2\alpha-\nu\cos^2\alpha)\cos\alpha;$$

$$M_\beta=(216\pi/625)[\sin\alpha\cos^2\alpha+(1/48)\sin\alpha(1+\nu_f\sin^4\alpha\cos 2\alpha)+(1/864)/(1+\nu)].$$

It may be pointed out that wires within the second layer are just in contact with their adjacent neighbors in the same layer, however, these contacts are ignored for application of the results of the analysis. In other words, the treatment followed here assumes the outside wires to have ''slightly'' lower diameter than that of the core.

On substitution of the above data, the following expressions for the normal force per unit length of contact between the core and the wires in the second layer as well as that for the contact force at the points of contact between the wires of the second and third layer are obtained:

$$Q_{2-3}=(\pi a\sigma_{nom}\cos^2\alpha\sin\alpha/2)[1+\sin\alpha\{1+18\sin^3\alpha+48\cos^2\alpha(1+12\sin^3\alpha)\}]^{-1}[19(\sin^2\alpha-\nu\cos^2\alpha)\times(25/12+81\cos^2\alpha+\nu_f\sin^4\alpha\cos 2\alpha)+(125/36)\times(\tau_{nom}/\sigma_{nom})\cos\alpha\{306\sin\alpha(\sin^2\alpha-\nu\cos^2\alpha)-15\}] \tag{15}$$

$$P_3=(\pi a^2\sigma_{nom}\sin\alpha\cos\alpha)[1+\sin\alpha\{1+18\sin^3\alpha+48\cos^2\alpha\times(1+12\sin^3\alpha)\}]^{-1}[19(\sin^2\alpha-\nu\cos^2\alpha)\{\cos^2\alpha+(1/16)\times(1+\nu_f\sin^4\alpha\cos 2\alpha)+1/288\}-(125/18)\times(\tau_{nom}/\sigma_{nom})\cos\alpha\{1+9\sin\alpha(\sin^2\alpha-\nu\cos^2\alpha)\}] \tag{16}$$

where

$$\tau_{nom}=M/(\pi R^3).$$

The subsequent application of Hertz theory of contact stresses leads to their critical values in the cables as indicated below.

Case 1 Contact Stresses Between Core and the Second Layer. Here, the maximum contact stresses obtained can be written as

$$\sigma_2=(1/2)\sqrt{E\sigma_{nom}}\cos\alpha\sqrt{\sin\alpha(1+\sin^2\alpha)/(1-\nu^2)}[1+\sin\alpha.\{1+18\sin^3\alpha+48\cos^2\alpha(1+12\sin^3\alpha)\}]^{-1/2}[19(\sin^2\alpha-\cos^2\alpha)(25/12+81\cos^2\alpha+\nu_f\sin^4\alpha\cos 2\alpha)+(125/36)\times(\tau_{nom}/\sigma_{nom})\cos\alpha\{306\sin\alpha(\sin^2\alpha-\cos^2\alpha)-15\}]^{1/2} \tag{17}$$

In dimensionless form, the corresponding expressions take the form

$$\hat{\sigma}_2=(1/2)\sqrt{E\sigma_{nom}}\cos\alpha\sqrt{\sin\alpha(1+\sin^2\alpha)}[(1-\nu^2)]^{-1/2}\times[1+\sin\alpha\{1+18\sin^3\alpha+48\cos^2\alpha(1+12\sin^3\alpha)\}]^{-1/2}[19(\sin^2\alpha-\cos^2\alpha)(25/12+81\cos^2\alpha+\nu_f\sin^4\alpha\cos 2\alpha)+(125/36)(\tau_{nom}/\sigma_{nom})\times\cos\alpha\{306\sin\alpha(\sin^2\alpha-\cos^2\alpha)-15\}]^{1/2}. \tag{18}$$

Case 2 Contact Stresses Between Wires in the Second and Third Layer. In this case, the maximum contact stresses are given by

$$\hat{\sigma}_3=(1/3)[E^2\sigma_{nom}]^{1/3}[2\pi\sin\alpha\cos\alpha/(1-\nu^2)^2]^{1/3}d^{-1/6}\times[1+\sin\alpha\{1+18\sin^3\alpha+48\cos^2\alpha(1+12\sin^3\alpha)\}]^{-1/3}\times[19(\sin^2\alpha-\cos^2\alpha)\{\cos^2\alpha+(1/16)(1+\nu_f\sin^4\alpha\cos 2\alpha+1/288\}-(125/18)(\tau_{nom}/\sigma_{nom})\cos\alpha\{1+9\sin\alpha\times(\sin^2\alpha-\cos^2\alpha)\}]^{1/3}. \tag{19}$$

In dimensionless form, the corresponding expressions take the form

$$\hat{\sigma}_3=(1/3)[2\pi\sin\alpha\cos\alpha/(1-\nu^2)^2]^{1/3}d^{-1/6}[1+\sin\alpha\{1+18\sin^3\alpha+48\cos^2\alpha(1+12\sin^3\alpha)\}]^{-1/3}[19(\sin^2\alpha-\cos^2\alpha)\{\cos^2\alpha+(1/16)(1+\nu_f\sin^4\alpha\cos 2\alpha+1/288\}-(125/18)(\tau_{nom}/\sigma_{nom})\cos\alpha\{1+9\sin\alpha(\sin^2\alpha-\cos^2\alpha)\}]^{1/3} \tag{20}$$

where

$$\hat{\sigma}_3=\sigma_3/[E^2\sigma_{nom}]^{1/3}$$

$$d=[1+\sqrt{1-(1-\cos^4\alpha/9)\sin^2 2\alpha}]^2/[(1-\cos^4\alpha/9)\sin^2 2\alpha].$$

Results and Discussion

First, an attempt is made to experimentally test the validity of the linear deformation derivative results earlier developed for the multilayered wire-rope strands under tension and torsion. For this, we take up a particular example of a simple strand with metallic core and an adjacent layer of six helical wires as described in the earlier section on experimental validation of deformation relations. Its deformation derivative values based on the theoretical analysis are found to be in reasonable agreement with the corresponding experimentally obtained results. The experimental tests thus support the validity of the earlier analytical approach proposing the use of suitable approximations based on order-of-

magnitude considerations while applying Costello's ''linear theory'' for prediction of deformations in the strands under tension and torsion.

The theoretical analysis subsequently undertaken utilizes these closed-form deformation derivative results together with Hertz theory of contact stresses in order to develop analytical expressions for the maximum contact stresses in the cables. The results are rather general and apply to any arbitrary multilayered strand with metallic core. The importance of such explicit analytical predictions cannot be overemphasized. The results show that the Poisson's ratio of the cable material does not have much effect on the critical stresses. In contrast, the material modulus of elasticity is an important parameter with considerable influence on the contact stresses. For the cables made of material such as steel, under ''modest'' loading, the hike in the stress levels in the contact regions is likely to be rather large.

It may be noted that the closer the layer of wires lies to the core, the higher would be the contact forces induced due to tension and torsion. Thus, the forces would be highest between the core and its adjacent layer and significantly lower on the outermost layer. Yet, the maximum contact stresses on wires in the core and the adjacent layer are unlikely to be at the highest level. This happens by virtue of the ''line contact'' interaction between these innermost mating layers unlike all other outer layers characterized by ''point contacts.'' The contact stresses between the core and the adjacent layer of helical wires are found to be proportional to the square-root of the rigidity modulus while those among wires of the adjacent outer layers are proportional to $[E]^{2/3}$. As a consequence, in spite of the lower contact forces involved, the outer layers, especially the third and its inner contacting layer may suffer higher contact stresses. Interestingly, it also follows that under ''modest'' loading of the cables made of materials like steel which have rather ''large'' values of the rigidity modulus, these critical stresses are likely to be the highest at the points of contact between the second and third layers.

For a better assessment of the effect of the various important parameters on rope design, the critical stresses are computed for the typical three-layered strand with metallic core considered earlier. Here, the wires of the second and third layers have same helix angles but opposite lays. The plots showing the influence of some important design parameters on maximum dimensionless contact stresses in the cables are obtained (Figs. 6–8). Figure 6 shows the effect of the helix angle parameter α on the nondimensionalized critical stress values when the cables are subjected to pure tensile loads as well as the case while under combined tension and torsion. Of the two stresses, the values of $\hat{\sigma}_3$ remain significantly higher as expected. Their values, however, continually decrease as the helix angle increases from 60 to 90 deg for ropes with regular lay in the second and lang lay in the third layer. Almost similar behavior is observed as the helix angle α decreases from 120 to 90 deg.

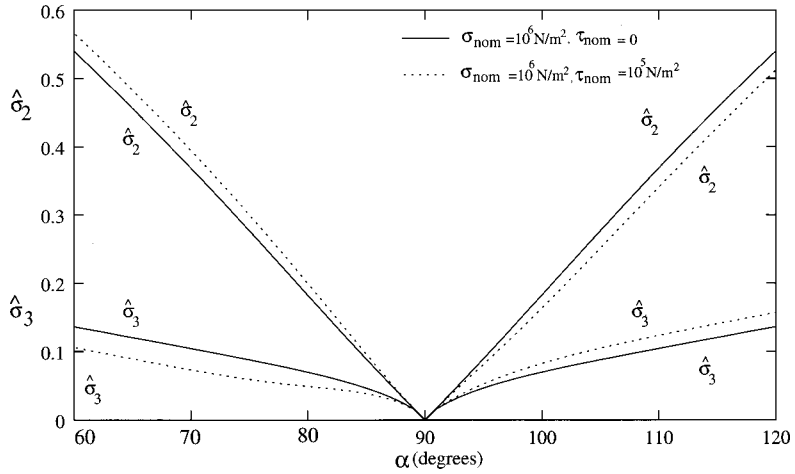


Fig. 6 Effect of α on dimensionless contact stresses

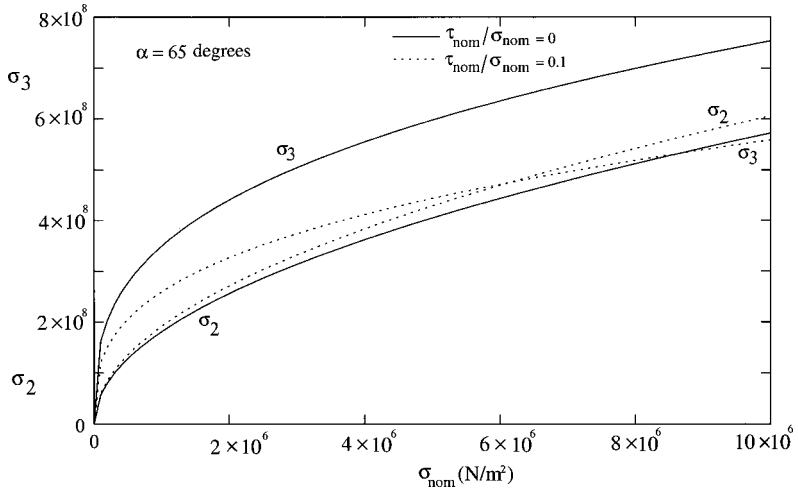


Fig. 7 Variation of critical stresses as influenced by nominal tensile stress

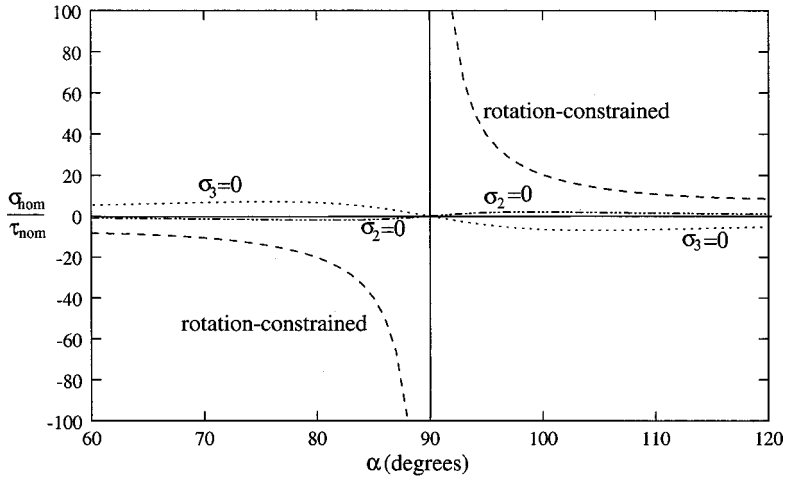


Fig. 8 Prediction of the ratio (σ_{nom}/τ_{nom}) for zero critical stresses and for rotation-constrained case

For a better appreciation on the order of contact stresses induced in the cables, their values are now obtained in dimensional form (Fig. 7). As expected, for relatively modest loads, the maximum stress σ_3 induced in the contacting region between the third and the second layer is higher than σ_2, i.e., the stress between the core and its adjacent layer. However, for relatively larger tensile loads and/or in the presence of significant torsional loads, this trend gets reversed. Interestingly, both the maximum contact stresses turn out to be virtually an order of magnitude higher than the nominal stresses applied. Consequently, even for the modest nominal cable stresses, the level of compressive stresses in the interior can reach that of the ultimate tensile strength of the cable material. That perhaps provides an explanation for the broken inner wires observed in the cables after long use even under modest loading.

Interestingly, through a judicious choice of the nominal tensile stress to torsional stress ratio based on the helix angle α, it seems theoretically possible to achieve a zero contact stress level for either of the two cases considered—not both. Figure 8 brings out how the helix angle α influences this choice for which the generally undesirable contact stresses $\hat{\sigma}_2$ or $\hat{\sigma}_3$ become zero. Also presented here is the plot of $(\sigma_{nom}/\tau_{nom})$ versus α for the rotation-constrained case. It may be noted that for the case $\tau_{nom}=0$, both the contact stresses would be nonzero and significantly large. Similarly, for the rotation-constrained case, none of the two contact stresses would vanish regardless of the choice of α. However, the situation may undergo qualitative change in cases of strands with more than two outer layers.

Concluding Remarks

First, the experimental test results support the general validity of the analytical expressions for the deformation derivatives earlier developed for single strand cable with multiple layers of helical wires around the metallic core. The theoretical analysis subsequently undertaken utilizes the Hertz theory of contact stresses together with these analytical expressions in order to obtain the closed-form expressions for the maximum contact stresses in the cable. The critical compressive stresses are found to be strongly dependent upon the material modulus of rigidity and the helix angles of wires in the various layers. In view of the point contacts between the wires in the second and other outer layers, the critical stresses in the strands with metallic core and surrounded by two or more layers are likely to be significantly higher than the corresponding values in those with fibrous cores.

Through application of the general results to a three-layered single strand cable with metallic core, it is shown that even though the contact forces on the core and its adjacent layer would be highest, in general, the critical stresses are likely to attain the highest levels in the contacting region on wires between the second and third layers. It is interesting to note that even for the modest loading of cables made of materials like steel with rather high values of E, the compressive stresses in the region of contact between the wires can attain the ultimate tensile stress levels.

For cable designs with regular and lang lay in the alternate layers, the critical stresses are relatively smaller when the helix angles are close to $\pi/2$. Hence such ''large'' helix angles are likely to promote longer life-span. In contrast, in the wire rope applications for dissipating energy, the smaller helix angles in the vicinity of $+\pi/3$ or $-\pi/3$ that result in larger contact stresses may be recommended.

Acknowledgment

The first author thanks the Department of Mechanical Engineering at the Swiss Federal Institute of Technology, Lausanne, for the financial support during his sabbatical stay. Thanks are also due to Mr. M. Studer, a doctoral candidate in the Laboratory of Applied Mechanics and Reliability Analysis of the department for his help in performing the experiments.

References

[1] Costello, G. A., 1978, ''Analytical Investigation of Wire Rope,'' ASME Appl. Mech. Rev., **31**, pp. 897–900.
[2] Kumar, K., and Cochran, Jr., J. E., 1987, ''Closed-Form Analysis for Elastic Deformations of Multilayered Strands,'' ASME J. Appl. Mech., **54**, pp. 898–903.
[3] Costello, G. A., and Phillips, J. W., 1973, ''Contact Stresses in Thin Twisted Rods,'' ASME J. Appl. Mech., **40**, pp. 629–630.
[4] Costello, G. A., and Phillips, J. W., 1974, ''A More Exact Theory for Twisted Wire Cables,'' ASCE J. Eng. Mech., **100**, pp. 1096–1099.
[5] Costello, G. A., and Phillips, J. W., 1976, ''Effective Modulus of Twisted Wire Cables,'' ASCE J. Eng. Mech., **102**, pp. 171–181.
[6] Velinsky, S. A., Anderson, G. L., and Costello, G. A., 1984, ''Wire Rope With Complex Cross Sections,'' ASCE J. Eng. Mech., **110**, pp. 380–391.
[7] Phillips, J. W., and Costello, G. A., 1985, ''Analysis of Wire Ropes With Internal Wire-Rope Cores,'' ASME J. Appl. Mech., **52**, pp. 510–516.
[8] Jiang, W. G., Yao, M. S., and Walton, J. M., 1999, ''A Concise Finite Element Model for Simple Straight Wire Rope Strand,'' Int. J. Mech. Sci., **41**, pp. 143–161.
[9] Utting, W. S., and Jones, N., 1987, ''The Response of Wire Rope Strands to Axial Tensile Loads—Part I: Experimental Results and Theoretical Predictions,'' Int. J. Mech. Sci., **29**, pp. 605–619.
[10] Utting, W. S., and Jones, N., 1987, ''The Response of Wire Rope Strands to Axial Tensile Loads—Part II: Comparison of Experimental Results and Theoretical Predictions,'' Int. J. Mech. Sci., **29**, pp. 621–636.
[11] Kumar, K., and Cochran, Jr., J. E., 1990, ''Analytical Estimation for Static

Elastic Deformations of Wire Ropes with Fibrous Core,'' ASME J. Appl. Mech., **57**, pp. 1000–1003.
[12] Kumar, K., Cochran, Jr., J. E., and Cutchins, M. A., 1997, ''Contact Stresses in Cables due to Tension and Torsion,'' ASME J. Appl. Mech., **64**, pp. 935–939.
[13] Costello, G. A., 1999, *Theory of Wire Rope*, Second ed., Springer-Verlag, New York, pp. 110–111.
[14] Vygodsky, M., 1975, *Mathematical Handbook*, pp. 514–534, (translated from Russian by G. Yankovsky).
[15] Seely, S. B., and Smith, J. O., 1963, *Advanced Mechanics of Materials*, John Wiley and Sons, New York, pp. 342–378.
[16] Gradshteyn, I. S., and Ryzhik, I. M., 1994, *Tables of Integrals, Series and Products*, 5th ed., Academic Press, San Diego, CA, pp. 907–910, (translated from Russian by Scripta Technical Inc.).

Elastic Fields in a Polyhedral Inclusion With Uniform Eigenstrains and Related Problems

H. Nozaki
Faculty of Education,
Ibaraki University,
2-1-1 Bunkyo
Mito, Ibaraki 310-8512, Japan

M. Taya
Department of Mechanical Engineering,
University of Washington,
Box 352600,
Seattle, WA 98195-2600
Fellow ASME

In this paper, the elastic field in an infinite elastic body containing a polyhedral inclusion with uniform eigenstrains is investigated. Exact solutions are obtained for the stress field in and around a fully general polyhedron, i.e., an arbitrary bounded region of three-dimensional space with a piecewise planner boundary. Numerical results are presented for the stress field and the strain energy for several major polyhedra and the effective stiffness of a composite with regular polyhedral inhomogeneities. It is found that the stresses at the center of a polyhedral inclusion with uniaxial eigenstrain do not coincide with those for a spherical inclusion (Eshelby's solution) except for dodecahedron and icosahedron which belong to icosidodeca family, i.e., highly symmetrical structure.
[DOI: 10.1115/1.1362670]

Introduction

The elastic field due to inclusions in an infinitely extended elastic media has been extensively investigated following Eshelby's pioneering work ([1]), and its summary is given in a book by Mura [2]. However, there have been only a limited number of analytic solutions for problems of nonellipsoidal inclusions, cases of cuboidal ([3–8]), rectangular ([9,10]), and cylindrical ([11–14]) shapes.

Rodin [15] considered a problem of Eshelby's tensors for polygonal and polyhedral inclusions where the Eshelby's tensors were given as a double or triple integral, which was evaluated by subdividing the inclusions into two or three-dimensional rectangular simplexes. He showed an approach to calculate the Eshelby's tensor for a polyhedral inclusion without explicit solutions. Rodin [15] and Markenscoff [16] proved that polyhedral inclusions with constant Eshelby's tensor do not exist. Lubarda and Markenscoff [17] showed that the Eshelby property (constancy of the stress for uniform eigenstrain) does not hold for any inclusion bounded by a polynomial surface of higher than the second degree, or any inclusion bounded by a nonconvex surface. They also showed that inclusions bounded by segments of two or more different surfaces are also precluded. Recently, Nozaki and Taya [18] analyzed the elastic field in a polygonal inclusion in an infinite body and proposed a method to estimate the effective stiffness of a composite with nonellipsoidal inhomogeneities. They carried out an area integral to obtain the displacement field ([2]) by subdividing a polygonal region into triangles and differentiated the displacement field to obtain the strain and the stress fields. They noted the special characteristic with a regular polygonal inclusion. More recently, Ru [19] obtained the analytic solution of elastic fields in a plane or half-plane containing an inclusion of arbitrary shape using the techniques of analytical continuation and conformal mapping.

The present paper is concerned with the solutions for the elastic field arising from an arbitrary polyhedral inclusion with uniformly distributed eigenstrains in an infinite elastic body. In this paper, the displacement is first given as a volume integral over the inclusion ([2]), which is then transformed to a surface integral over the surface of the inclusion by integration by parts. The resulting surface integral is evaluated by subdividing the surface into rectangular triangles ([15,20]). The solutions for strain, stress, and Eshelby's tensors will be obtained by analytical differentiation of the displacement field. Subsequently, we will calculate the strain energy of the body by using a numerical integration formula for three-dimensional simplexes ([21]). Numerical results will be presented for the stress field and the strain energy in a body containing a polyhedral inclusion. The results of the effective stiffness of a composite with polyhedral inhomogeneities will also be presented.

After the publication of our previous paper ([18]), Rodin [22] claimed that the method to estimate the effective stiffness of a composite in the paper contains a serious flaw and hence the results concerning composite stiffness are incorrect. In this paper we will show that our method provides still a good approximation to the stiffness of a composite with polygonal inhomogeneities. The degree of accuracy of the present method will be shown to increase as the polygonal inhomogeneity approaches to a circle and also, as the stiffness difference between the inhomogeneity and the matrix becomes small. This will be proved numerically by comparison between the solutions based on the present model and by the boundary element method (BEM).

2 Formulation

2.1 Statement of the Problem. Consider an infinite, elastic, homogeneous and isotropic domain D having arbitrary shaped polyhedral inclusion Ω with uniform eigenstrain ε_{ij}^* in a Cartesian coordinate system (x_1,x_2,x_3). The matrix domain is denoted as $D-\Omega$. The boundary of Ω ($|\Omega|$) is assumed to be formed by N faces $|\Omega|_I$ ($I=1-N$), $|\Omega|_I$ being a polygon with $M_I \geqslant 3$ vertices. For each face the outer unit normal $\mathbf{n}^I$, $|\mathbf{n}^I|=1$ is introduced.

The eigenstrain ε_{ij}^* is given some constant value in Ω but vanishes in $D-\Omega$. For prescribed eigenstrain ε_{ij}^* in Ω, the resulting stress σ_{ij} is given by

$$\sigma_{ij}=C_{ijkl}(\varepsilon_{kl}-\varepsilon_{kl}^*) \quad \text{in} \quad D, \tag{1}$$

where C_{ijkl} is the elastic stiffness tensor of the body and ε_{kl} is the strain induced by the inclusion. Using Green's function $G_{ij}(\mathbf{x}-\mathbf{x}')$, the displacement field inside and outside of Ω can be written as ([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, May 13, 1999; final revision, April 14, 2000. Editor: L. T. Wheeler. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

$$u_i(\mathbf{x}) = -C_{jkmn}\varepsilon^*_{mn}\int_\Omega G_{ij,k}(\mathbf{x}-\mathbf{x}')d\mathbf{x}', \tag{2}$$

where $\mathbf{x}$ is an arbitrary point in D and $\mathbf{x}'$ is a point inside Ω. To evaluate the integral, it is convenient to rewrite Eq. (2) by integration by parts. The equation is transformed to

$$u_i(\mathbf{x}) = C_{jkmn}\varepsilon^*_{mn}\int_{|\Omega|} G_{ij}(\mathbf{x}-\mathbf{x}')n_k dS(\mathbf{x}')$$

$$= C_{jkmn}\varepsilon^*_{mn}\sum_{I=1}^{N} n_k^I \int_{|\Omega|_I} G_{ij}(\mathbf{x}-\mathbf{x}')dS(\mathbf{x}'), \tag{3}$$

where dS is the surface element. For isotropic media, the Green's function is given as ([2])

$$G_{ij}(\mathbf{x}-\mathbf{x}') = \frac{1}{16\pi\mu(1-\nu)|\mathbf{x}-\mathbf{x}'|} \times\left[(3-4\nu)\delta_{ij}+\frac{(x_i-x_i')(x_j-x_j')}{|\mathbf{x}-\mathbf{x}'|^2}\right], \tag{4}$$

where μ is the shear modulus, ν is the Poisson's ratio, and δ_{ij} is the Kronecker delta. Substituting Eq. (4) into Eq. (3), we obtain

$$u_i(\mathbf{x}) = D_{imn}\varepsilon^*_{mn} \tag{5}$$

where

$$D_{imn} = \frac{C_{ikmn}}{16\pi\mu(1-\nu)}\sum_{I=1}^{N} n_k^I[(3-4\nu)\delta_{ij}I_1^I + I_2^{I(ij)}] \tag{6}$$

$$I_1^I = \int_{|\Omega|_I}\frac{1}{|\mathbf{x}-\mathbf{x}'|}dS(\mathbf{x}') \tag{7}$$

$$I_2^{I(ij)} = \int_{|\Omega|_I}\frac{(x_i-x_i')(x_j-x_j')}{|\mathbf{x}-\mathbf{x}'|^3}dS(\mathbf{x}'). \tag{8}$$

2.2 Subdivision of the Surface of the Polyhedron. To perform integrals I_1^I and $I_2^{I(ij)}$, we subdivide the surfaces of a polyhedron into rectangular triangles. Let now P_I be the orthogonal projection of $\mathbf{x}$ onto the face $|\Omega|_I$, and let C_J^I ($J=1,2,\ldots,M_I$: $C_{M_I+1}^I = C_1^I$) be a vertex of $|\Omega|_I$ (see Fig. 1). The orthogonal projection of P_I onto the edge $C_J^I C_{J+1}^I$ will be denoted by P_J^I. Using P_I as a pivot, $|\Omega|_I$ is now subdivided into pairs of rectangular triangles $(T_{I1},T_{I1}'),(T_{I2},T_{I2}'),\ldots,(T_{IM_1},T_{IM_1}')$. For example, the one associated with the edge $C_J^I C_{J+1}^I$ is $(T_{IJ}:P_IP_J^IC_J^I, T_{IJ}':P_IP_J^IC_{J+1}^I)$. Also, we introduce vectors shown in Fig. 1 as

$\boldsymbol{\rho}^I$: initial point $\mathbf{x}$, terminal point P_I,

$\boldsymbol{\rho}_J^I$: initial point $\mathbf{x}$, terminal point P_J^I,

$\boldsymbol{\gamma}_J^I$: initial point $\mathbf{x}$, terminal point C_J^I,

$\mathbf{a}_J^I$: initial point P_I, terminal point P_J^I,

$\mathbf{b}_J^I$: initial point P_J^I, terminal point C_J^I,

$\mathbf{c}_J^I$: initial point P_I, terminal point C_J^I,

$\boldsymbol{\gamma}_{J+1}^I$: initial point $\mathbf{x}$, terminal point C_{J+1}^I,

$\mathbf{b}_{J+1}^I$: initial point P_I^J, terminal point C_{J+1}^I,

$\mathbf{c}_{J+1}^I$: initial point P_I, terminal point C_{J+1}^I.

The integral I_1^I, $I_2^{I(ij)}$ can be expressed as

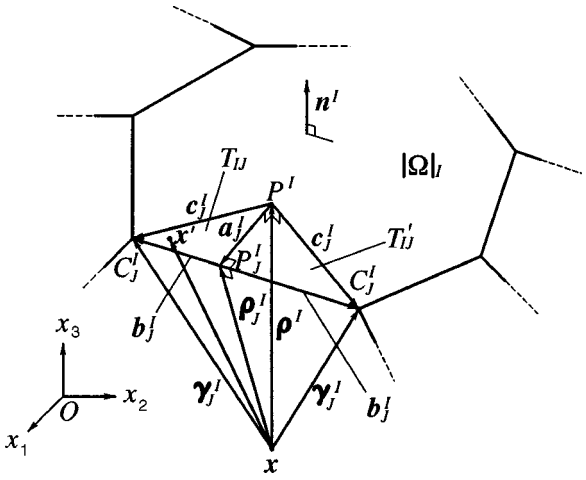


Fig. 1 Subdivision of a side of the polyhedron

$$I_1^I = \sum_{J=1}^{M_I}[s_1 I_1^{IJ} + s_1'(I_1^{IJ})'], \tag{9}$$

$$I_2^{I(ij)} = \sum_{J=1}^{M_I}[s_2 I_2^{IJ(ij)} + s_2'(I_2^{IJ(ij)})'], \tag{10}$$

where

$$I_1^{IJ} = \int_{T_{IJ}}\frac{1}{|\mathbf{x}-\mathbf{x}'|}dS(\mathbf{x}'), \tag{11}$$

$$(I_1^{IJ})' = \int_{T_{IJ}'}\frac{1}{|\mathbf{x}-\mathbf{x}'|}dS(\mathbf{x}'), \tag{12}$$

$$I_2^{IJ(ij)} = \int_{T_{IJ}}\frac{(x_i-x_i')(x_j-x_j')}{|\mathbf{x}-\mathbf{x}'|^3}dS(\mathbf{x}'), \tag{13}$$

$$(I_2^{IJ(ij)})' = \int_{T_{IJ}'}\frac{(x_i-x_i')(x_j-x_j')}{|\mathbf{x}-\mathbf{x}'|^3}dS(\mathbf{x}'), \tag{14}$$

In Eqs. (9) and (10), s_1, s_1' and s_2, s_2' are the signs of I_1^{IJ}, $(I_1^{IJ})'$ and $I_2^{IJ(ij)}$, $(I_2^{IJ(ij)})'$, respectively, which are determined by the locations of the point $\mathbf{x}$ and the rectangular triangles T_{IJ}, T_{IJ}' on which the integrals are defined ([15,20]).

At first we perform the integrals in Eqs. (11) and (13) over T_{IJ}. We introduce new parameters (ξ,η,ζ) defined by

$$\mathbf{x}'-\mathbf{x} = \xi\boldsymbol{\rho}_J^I + \eta\boldsymbol{\gamma}_J^I + \zeta\boldsymbol{\rho}^I \quad (0\leqslant\xi,\eta,\zeta\leqslant 1). \tag{15}$$

In Eqs. (11)–(14), $\mathbf{x}'$ is the point on surface of the polyhedron so ξ, η, and ζ have the following relationship:

$$\xi+\eta+\zeta = 1. \tag{16}$$

Using Eq. (16), we can modify Eq. (15) as

$$\mathbf{x}'-\mathbf{x} = \xi\mathbf{a}_J^I + \eta\mathbf{c}_J^I + \boldsymbol{\rho}^I \quad (0\leqslant\xi,\eta\leqslant 1). \tag{17}$$

The surface element $dS(\mathbf{x}')$ is transformed to

$$dS(\mathbf{x}') = |\mathbf{a}_J^I||\mathbf{b}_J^I|d\xi d\eta. \tag{18}$$

The integrals I_1^{IJ} and $I_2^{IJ(ij)}$ are reduced to

$$I_1^{IJ} = \int_{T_{IJ}}\frac{1}{|\mathbf{x}-\mathbf{x}'|}dS(\mathbf{x}') = \int_0^1\int_0^{1-\eta}\frac{1}{|\xi\mathbf{a}_J^I+\eta\mathbf{c}_J^I+\boldsymbol{\rho}^I|}|\mathbf{a}_J^I||\mathbf{b}_J^I|d\xi d\eta, \tag{19}$$

$$I_2^{IJ(ij)}=\int_{T_{IJ}}\frac{(x_i-x_i')(x_j-x_j')}{|\mathbf{x}-\mathbf{x}'|^3}dS(\mathbf{x}')$$

$$=\int_0^1\int_0^{1-\eta}\frac{\{\xi(a_J^I)_i+\eta(c_J^I)_i+(\rho^I)_i\}\{\xi(a_J^I)_j+\eta(c_J^I)_j+(\rho^I)_j\}}{|\xi\mathbf{a}_J^I+\eta\mathbf{c}_J^I+\boldsymbol{\rho}^I|^3}$$

$$\times|\mathbf{a}_J^I||\mathbf{b}_J^I|d\xi d\eta. \tag{20}$$

Using the integral formulas given in the Appendix and the orthogonality of the vectors, we can perform the above integrals to arrive at

$$I_1^{IJ}=|\mathbf{a}_J^I|\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}+|\boldsymbol{\rho}^I|F_{IJ}, \tag{21}$$

$$I_2^{IJ(ij)}=(I_2^{IJ(ij)})_1+(I_2^{IJ(ij)})_2+(I_2^{IJ(ij)})_3, \tag{22}$$

where

$$(I_2^{IJ(ij)})_1=(a_J^I)_i(a_J^I)_j\left[\frac{2(|\boldsymbol{\gamma}_J^I|-|\boldsymbol{\rho}_J^I|)}{|\mathbf{a}_J^I||\mathbf{b}_J^I|}-\frac{|\boldsymbol{\gamma}_J^I|-|\boldsymbol{\rho}^I|}{|\mathbf{a}_J^I||\mathbf{b}_J^I|}\right.$$
$$\left.+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\mathbf{c}_J^I||\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|^2|\mathbf{a}_J^I|^2}F_{IJ}\right], \tag{23}$$

$$(I_2^{IJ(ij)})_2=-\{(a_J^I)_i(c_J^I)_j+(a_J^I)_j(c_J^I)_i\}\left[\frac{|\boldsymbol{\rho}^I|-|\boldsymbol{\rho}_J^I|}{|\mathbf{a}_J^I||\mathbf{b}_J^I|}\right.$$
$$+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|^2}F_{IJ}\Bigg]$$
$$-\{(a_J^I)_i(\rho^I)_j+(a_J^I)_j(\rho^I)_i\}\left[\frac{1}{|\mathbf{a}_J^I|}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}\right.$$
$$\left.-\frac{|\mathbf{c}_J^I|}{|\mathbf{a}_J^I||\mathbf{b}_J^I|}\log\left\{\frac{|\mathbf{c}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}^I|}\right\}-\frac{1}{|\mathbf{b}_J^I|}\log\left\{\frac{|\boldsymbol{\rho}^I|}{|\boldsymbol{\rho}_J^I|+|\mathbf{a}_J^I|}\right\}\right], \tag{24}$$

$$(I_2^{Ij(ij)})_3=(c_J^I)_i(c_J^I)_j\left[-\frac{|\mathbf{a}_J^I|(|\boldsymbol{\gamma}_J^I|-|\boldsymbol{\rho}^I|)}{|\mathbf{b}_J^I||\mathbf{c}_J^I|^2}+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\right.$$
$$\times\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|^2}F_{IJ}\Bigg]-\{(c_J^I)_i(\rho^I)_j$$
$$+(c_J^I)_j(\rho^I)_i\}\left[\frac{|\mathbf{a}_J^I|}{|\mathbf{c}_J^I||\mathbf{b}_J^I|}\log\left\{\frac{|\mathbf{c}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}^I|}\right\}\right.$$
$$\left.+\frac{1}{|\mathbf{b}_J^I|}\log\left\{\frac{|\boldsymbol{\rho}^I|}{|\boldsymbol{\rho}_J^I|+|\mathbf{a}_J^I|}\right\}\right]-\frac{(\rho^I)_i(\rho^I)_j}{|\boldsymbol{\rho}^I|}F_{IJ}. \tag{25}$$

In the above equations, F_{IJ} is

$$F_{IJ}=\tan^{-1}\frac{|\mathbf{b}_J^I||\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}^I||\mathbf{a}_J^I|}-\tan^{-1}\frac{|\mathbf{b}_J^I||\mathbf{a}_J^I|}{|\boldsymbol{\rho}^I||\boldsymbol{\gamma}_J^I|}-\tan^{-1}\frac{|\mathbf{b}_J^I|}{|\mathbf{a}_J^I|}. \tag{26}$$

Integrals $(I_1^{IJ})'$ and $(I_2^{IJ(ij)})'$ in Eqs. (9) and (10) are obtained by replacing the vectors $\mathbf{b}_J^I$, $\mathbf{c}_J^I$, $\boldsymbol{\gamma}_J^I$ by $\mathbf{b}_{J+1}^I$, $\mathbf{c}_{J+1}^I$, $\boldsymbol{\gamma}_{J+1}^I$ in Eq. (21) and (23)–(26) (see Fig. 1).

Strain fields can be obtained by differentiating Eq. (5) as

$$\varepsilon_{ij}(\mathbf{x})=\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)=\frac{1}{2}(D_{imn,j}+D_{jmn,i})\varepsilon_{mn}^*=S_{ijmn}\varepsilon_{mn}^*, \tag{27}$$

where S_{ijmn} is the Eshelby's tensor for a polyhedral inclusion. The explicit expression of S_{ijmn} is given in the Appendix. The stress field in D is obtained by substituting Eq. (27) into Eq. (1).

To obtain the solutions for the elastic field on the boundary of the inclusion, we calculate the limiting solutions shown in the Appendix, which may also be required to calculate the elastic field inside or outside the inclusion when the absolute values of the vectors are equal to zero.

The elastic strain energy in the infinite domain D is obtained as ([2])

$$W^*=\frac{1}{2}\int_D\sigma_{ij}(\mathbf{x})\{\varepsilon_{ij}(\mathbf{x})-\varepsilon_{ij}^*\}d\mathbf{x}=-\frac{1}{2}\int_\Omega\sigma_{ij}(\mathbf{x})\varepsilon_{ij}^*d\mathbf{x}$$

$$=-\frac{1}{2}\{C_{ijkl}(\overline{S}_{klmn}-I_{klmn})\varepsilon_{mn}^*\}\varepsilon_{ij}^*V_\Omega, \tag{28}$$

where $\overline{S}_{klmn}$ is the averaged Eshelby's tensor ([18]) defined by

$$\overline{S}_{klmn}=\frac{1}{V_\Omega}\int_\Omega S_{klmn}(\mathbf{x})d\mathbf{x}, \tag{29}$$

where the bar over the quantity means the average over the volume of the inclusion (V_Ω).

3 Response to Rodin's Question

In our previous paper ([18]), we analyzed the elastic field in a polygonal inclusion and proposed a method to calculate the effective stiffness of a composite reinforced by nonellipsoidal inhomogeneities as an extension of Eshelby's equivalent inclusion method. However, Rodin [22] claimed that the following relationship in our paper does not hold in general except for the case of ellipsoidal inclusions and hence our result for composite stiffness is incorrect.

$$\overline{S_{ijkl}(\mathbf{x})\varepsilon_{kl}^*(\mathbf{x})}=\overline{S_{ijkl}(\mathbf{x})}\ \overline{\varepsilon_{kl}^*(\mathbf{x})}. \tag{30}$$

In Eq. (30), $\mathbf{x}$ is a point inside the inclusion, $\varepsilon_{ij}^*(\mathbf{x})$ is the equivalent eigenstrain, and $S_{ijkl}(\mathbf{x})$ is the Eshelby's tensor for a polygonal inclusion. The bar over the quantities means the average over the volume of inclusion. In this section, we calculate the average stress

$$\overline{\sigma}_{ij}=\frac{1}{V_{\Omega'}}\int_{\Omega'}\sigma_{ij}(\mathbf{x})d\mathbf{x} \tag{31}$$

inside several polygonal inhomogeneities Ω' (volume $V_{\Omega'}$) by the method proposed by the authors and the boundary element method (BEM). Even though Eq. (30) does not hold for nonellipsoidal inclusion in an exact sense, by comparing the two solutions we will show that Eq. (30) for polygonal inclusions will still provide a good estimate for the composite stiffness, and becoming a better approximation as the number of polygon's side increases and the difference between the stiffness of the inhomogeneity and the matrix becomes small.

We consider an infinite body D with a regular polygonal inhomogeneity Ω' as shown in Fig. 2. In this section, we compute the average stress in triangular, square, pentagonal, and hexagonal inhomogeneities inscribed to a circle of radius a. All regular polygons are centered at the origin O of the Cartesian coordinate system (x_1, x_2, x_3) and one of the vertices is located on the x_1-axis. Plane strain in x_1–x_2 plane is assumed. Using the method proposed by the authors ([18]), the average stress inside the inhomogeneity is obtained as

$$\overline{\sigma}_{ij}=\{C_{ijkl}^m+C_{ijpq}^m(\overline{S}_{pqmn}-I_{pqmn})B_{mnkl}\}\varepsilon_{kl}^\infty, \tag{32}$$

where C_{ijkl}^m is the elastic stiffness tensor of the matrix, I_{ijkl} is the identity tensor, ε_{ij}^∞ is an uniform farfield strain, and B_{ijkl} is a tensor defined by

$$B_{ijkl}=\{(C_{ijpq}^m-C_{ijpq}^f)\overline{S}_{pqmn}-C_{ijmn}^m\}^{-1}(C_{mnkl}^f-C_{mnkl}^m). \tag{33}$$

In Eq. (33), C_{ijkl}^f is the stiffness tensor of the inhomogeneity.

To solve the above two-dimensional problem by BEM, a finite size body is treated instead, but the width and height of the body

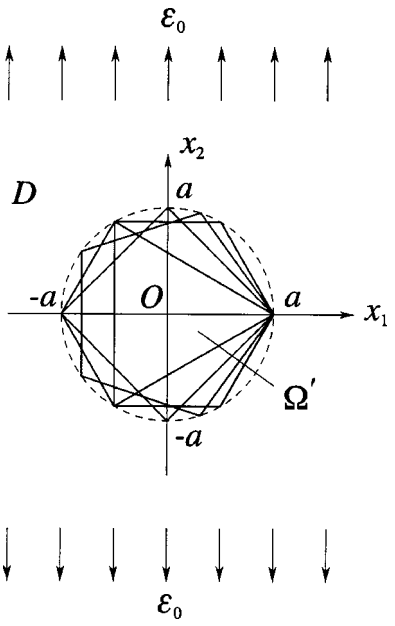


Fig. 2 A polygonal inhomogeneity in an infinite body subjected to far-field strain $(\varepsilon_{11}^{\infty}, \varepsilon_{12}^{\infty}, \varepsilon_{22}^{\infty}) = (0, 0, \varepsilon_0)$

Table 1 Comparison of averaged stress $\bar{\sigma}_{22}/\sigma_0$ inside Ω' calculated by the present method and BEM with $E^f/E^m = 0.001-1000$ and $\nu^f = \nu^m = 0.3$. The result for square is averaged over the rotation angle around the origin.

E^f/E^m	Method	Triangle	Square (averaged)	Pentagon	Hexagon
0.001	Present	3.292×10^{-3}	3.303×10^{-3}	3.292×10^{-3}	3.292×10^{-3}
	BEM	4.991×10^{-3}	3.893×10^{-3}	3.582×10^{-3}	3.455×10^{-3}
0.01	Present	3.225×10^{-2}	3.235×10^{-2}	3.225×10^{-2}	3.225×10^{-2}
	BEM	4.748×10^{-2}	3.774×10^{-2}	3.491×10^{-2}	3.375×10^{-2}
0.1	Present	2.678×10^{-1}	2.683×10^{-1}	2.678×10^{-1}	2.678×10^{-1}
	BEM	3.291×10^{-1}	2.923×10^{-1}	2.802×10^{-1}	2.749×10^{-1}
0.5	Present	7.661×10^{-1}	7.663×10^{-1}	7.661×10^{-1}	7.661×10^{-1}
	BEM	7.826×10^{-1}	7.731×10^{-1}	7.698×10^{-1}	7.683×10^{-1}
2	Present	1.181	1.181	1.181	1.181
	BEM	1.200	1.189	1.185	1.183
10	Present	1.383	1.384	1.383	1.383
	BEM	1.515	1.433	1.408	1.335
100	Present	1.438	1.440	1.438	1.438
	BEM	1.664	1.518	1.477	1.460
1000	Present	1.444	1.446	1.444	1.444
	BEM	1.688	1.528	1.484	1.466

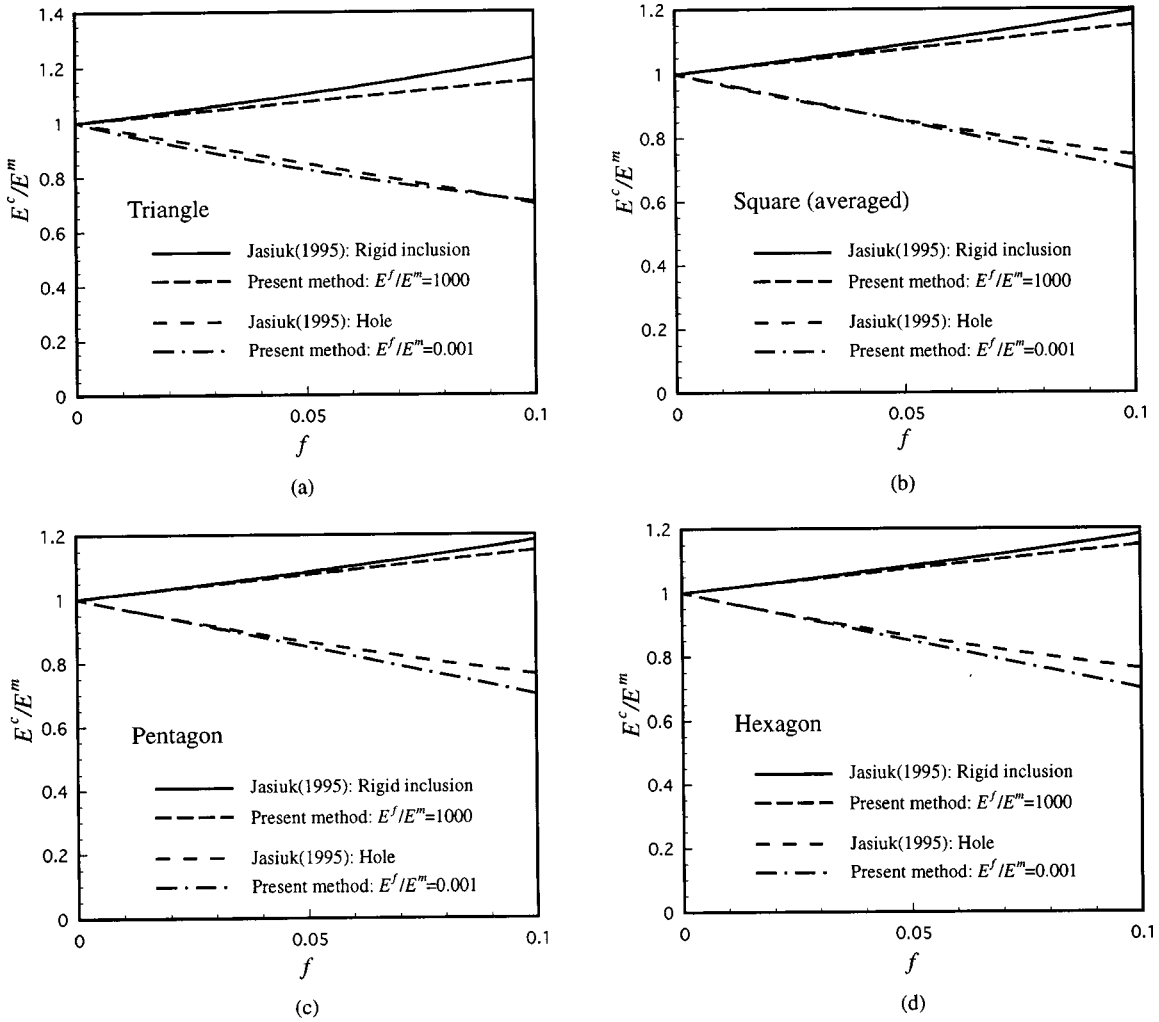


Fig. 3 Comparison of effective two-dimensional Young's moduli calculated by present method and Jasiuk's method ([25]): (*a*) triangle, (*b*) square, (*c*) pentagon, (*d*) hexagon. $\nu^f = \nu^m = 0.3$ for present method and $\nu^m = 0.3$ for Jasiuk's method.

are taken to be 100 times the radius of the circumcircle of the regular polygon where sharp corners are replaced by sufficiently small arcs. The radius of the arc is taken to be 0.01 times the radius of circumcircle of the regular polygons. This approximation gives sufficient accuracy. The boundary element models have totals of 234 (triangle model), 290 (square model), 378 (pentagon model), 402 (hexagon model) nodes for the two subregions which represent the inhomogeneity and surrounding infinite body. For all models, quadratic elements are used. For the evaluation of the average stress, we divide the polygon into triangles whose vertices are the center and two adjacent vertices. Then an integration formula for two-dimensional simplexes by Hammer et al. [21] is used. In each triangle, the number of integration points is 12.

Hereafter, we assume that the infinite body is subjected to a farfield strain $(\varepsilon_{11}^{\infty}, \varepsilon_{12}^{\infty}, \varepsilon_{22}^{\infty})=(0, 0, \varepsilon_0)$. Table 1 presents the values of the average stress $\overline{\sigma_{22}}$ obtained by the present method and BEM for the different E^f/E^m ratio, where E^f, E^m are the Young's modulus of the inhomogeneity and the matrix. The Poisson's ratios of the inhomogeneity and the matrix are set to $\nu^f=\nu^m=0.3$. The average stress is normalized by $\sigma_0=\varepsilon_0 E^m(1-\nu^m)/(1-\nu^m-2(\nu^m)^2)$, where σ_0 is the corresponding stress component along the x_2-direction for the given far-field strain $(\varepsilon_{11}^{\infty}, \varepsilon_{12}^{\infty}, \varepsilon_{22}^{\infty})=(0, 0, \varepsilon_0)$. Our solutions and those obtained by BEM disagree. However, it can be seen that the difference is becoming smaller as the number of polygon's side increases. The difference is also becoming smaller as the difference between E^f and E^m becomes small. Thus we can conclude that our method will give a practically reasonable result for the effective stiffness of a composite reinforced by polygonal inhomogeneities, especially for a composite having stiff matrix such as a metal matrix composite or a ceramic matrix composite. Here, we should comment that a square is not an isotropic shape ([18,23–25]). To obtain the average values over rotation angle from 0 deg, to 90 deg, we calculated the averages over 2 (0 deg, 45 deg), 4 (0 deg, 22.5 deg, 45 deg, 67.5 deg), 8 (0 deg, 11.25 deg,$\cdots$, 78.75 deg) and 16 (0 deg, 5.625 deg,$\cdots$, 84.375 deg) positions. It was confirmed that all results are the same to 14 figures for the present method and 6 figures for BEM. Hence we concluded that we can adopt the average over two positions 0 deg and 45 deg as the average over rotation angle from 0 deg to 90 deg. In Table 1, the result for the square is obtained by averaging $\overline{\sigma_{22}}$ over two positions of rotation angle.

Next, we compare the effective two-dimensional Young's modulus (E^c) of a body with dilute distribution of polygonal holes or rigid inclusions calculated by Jasiuk's method ([24,25]) and that of a body with dilute distribution of very compliant $(E^f/E^m=0.001)$ or very stiff $(E^f/E^m=1000)$ inhomogeneities calculated by the present method ([18]):

$$C_{ijkl}^c=C_{ijkl}^m+f(C_{ijmn}^f-C_{ijmn}^m)A_{mnkl}^{dil},$$

$$A_{ijkl}^{dil}=\{I_{ijkl}+\overline{S}_{ijmn}(C_{mnpq}^m)^{-1}(C_{pqkl}^f-C_{pqkl}^m)\}^{-1}. \quad (34)$$

The Poisson's ratios are again set to $\nu^f=\nu^m=0.3$. Figures 3(*a*)–(*d*) show the results for low volume fraction ($f\leqslant 0.1$) of holes, rigid inclusions, and inhomogeneities. The difference between the two methods is small. In the calculation for square, we used average $\overline{S}_{ijkl}$ over two positions of rotation angle 0 deg and 45 deg.

4 Numerical Results and Discussion for Polyhedral Inclusions

Numerical results of the stress distribution, the elastic strain energy for a body containing a polyhedral inclusion and the effective stiffness of a composite with polyhedral inhomogeneities are shown in this section. The Poisson's ratio of the body ν is assumed to be 0.3 for the stress and the strain energy calculations and an SiC particle-reinforced Al matrix composite is chosen for the effective stiffness calculation. Material properties of SiC and Al are shown in Table 2. We assume a polyhedral inclusion centered at the origin *O* of the Cartesian coordinate system. The five regular polyhedra shown in Fig. 4 are considered for the stress and the effective stiffness calculations. Furthermore, three inclusions of icosidodeca family shown in Fig. 5 are additionally considered

Table 2 Material properties of SiC particle and Al matrix

Material	$C_{1111}(\lambda+2\mu)$	$C_{1212}(\mu)$
Al	110.5 (GPa)	26.5 (GPa)
SiC	474.2 (GPa)	188.1 (GPa)

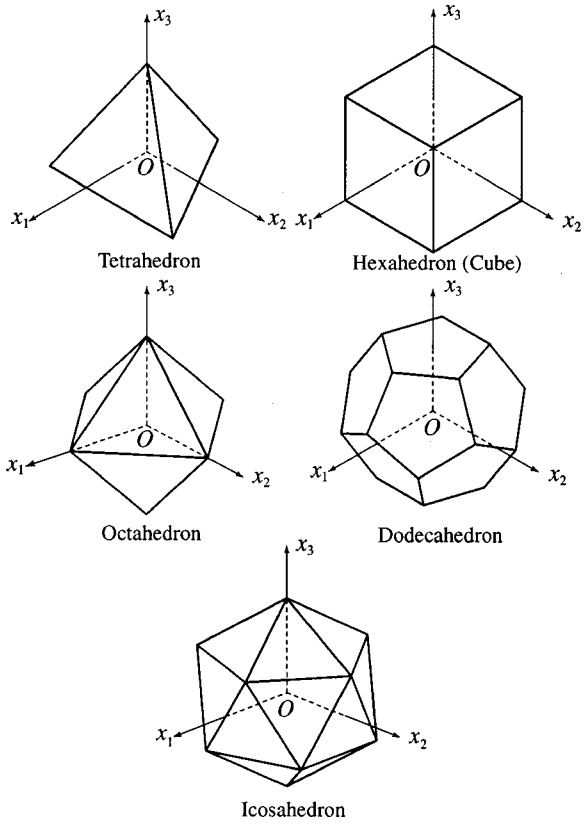


Fig. 4 Regular polyhedral inclusions

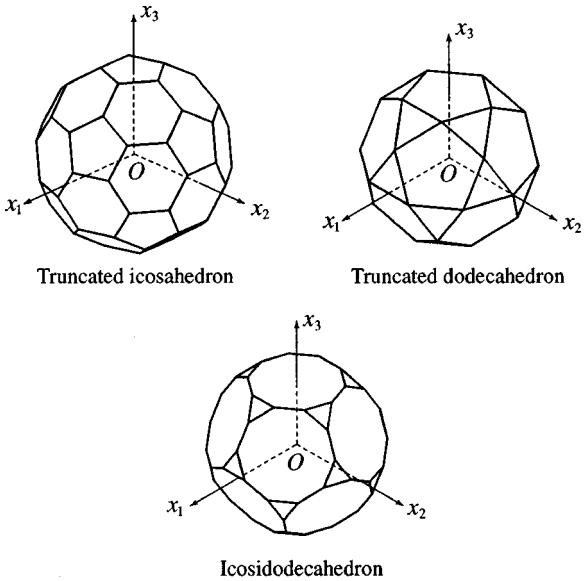


Fig. 5 Polyhedral inclusions belonging to icosidodeca family

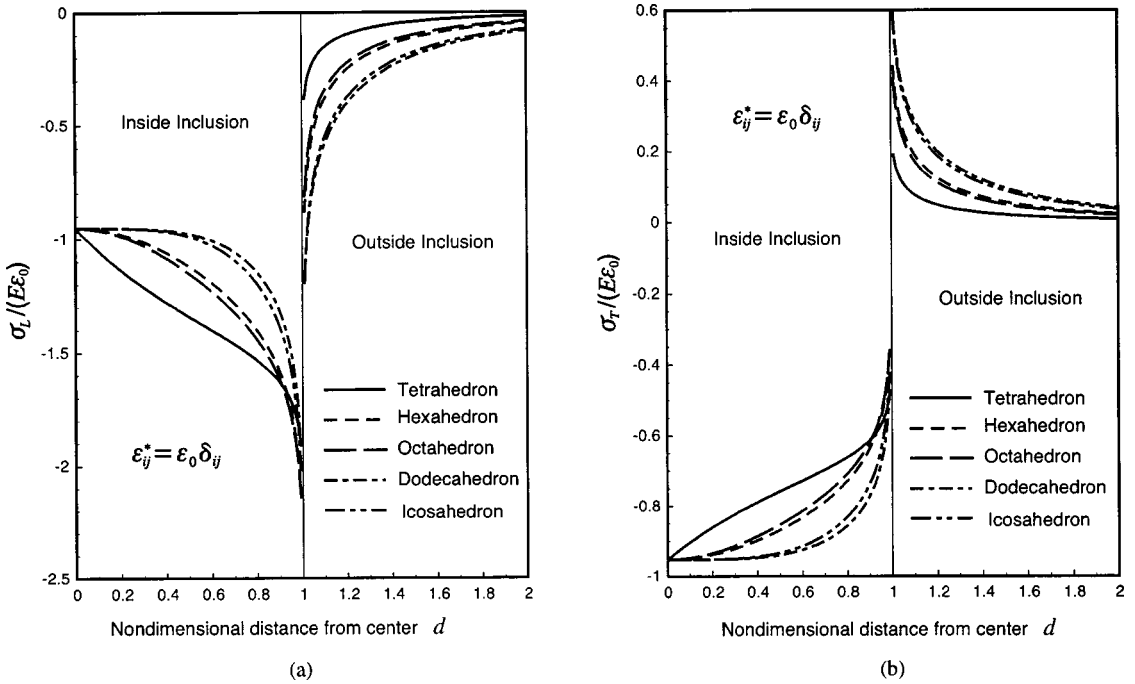


Fig. 6 Variation of normalized stress field in a polyhedral inclusion with a dilatational eigenstrain ($\varepsilon_{ij}^* = \varepsilon_0 \delta_{ij}$) along a line from the center through a vertex: (*a*) σ_L : normal stress on a plane normal to the line. (*b*) σ_T: normal stress on a plane parallel to the line. $d=1$ indicates the position of the vertex.

for the strain energy calculation. First, we examined the stress distribution for the cubic inclusion examined by Chiu [6] and confirmed that our results coincide with his.

Figures 6(a) and (b) show the distributions of the normalized stresses $\sigma_L/(E\varepsilon_0)$ and $\sigma_T/(E\varepsilon_0)$, respectively, inside and outside inclusion along a line from the center through a vertex for a dilatational eigenstrain ($\varepsilon_{ij}^* = \varepsilon_0 \delta_{ij}$). σ_L and σ_T are the normal stresses on a plane normal and parallel to the line, respectively, and E is the Young's modulus. Horizontal axis is normalized by the distance from the center to a vertex, hence $d=1$ indicates the position of the vertex. These figures show that the stresses have the logarithmic singularity at vertices of the polyhedra. Figures 7(a) and (b) show the distributions of the stresses along x_1-axis for a uniaxial eigenstrain ($\varepsilon_{11}^* = \varepsilon_0$ and other components are

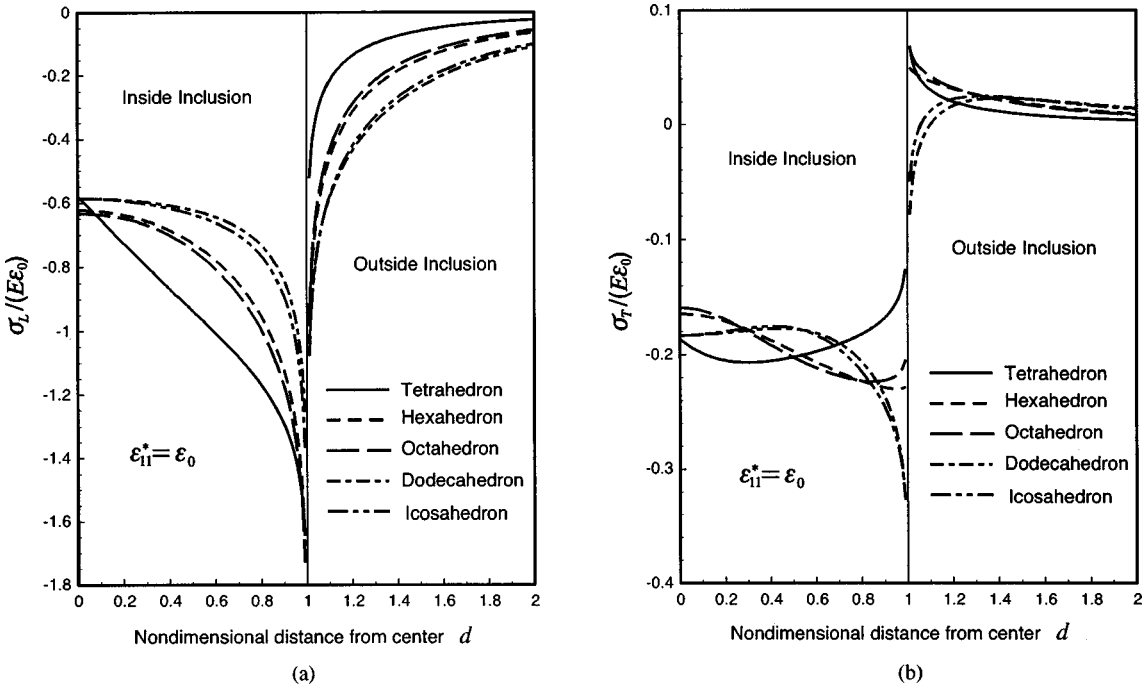


Fig. 7 Variation of normalized stress field in polyhedral inclusions with a uniaxial eigenstrain along x_1-axis, $\varepsilon_{11}^* = \varepsilon_0$: (*a*) $\sigma_L = \sigma_{11}$. (*b*) σ_T: normal stress on a plane parallel to the x_1-axis. $d=1$ indicates the position of the vertex.

Table 3 Normalized strain energy $\bar{W}^* = W^*/(\varepsilon_0^2 E V_\Omega)$ in an infinite body with a polyhedral inclusion for a dilatational eigenstrain ($\varepsilon_{ij}^* = \varepsilon_0 \delta_{ij}$) and a uniaxial eigenstrain along x_1-axis ($\varepsilon_{11}^* = \varepsilon_0$)

Inclusion shape	$\bar{W}^*$	
	Dilatational	Uniaxial
Sphere	1.42857	0.293040
Tetrahedron	1.42857	0.302871
Hexahedron (cube)	1.42857	0.285081
Octahedron	1.42857	0.287812
Dodecahedron	1.42857	0.293040
Icosahedron	1.42857	0.293040
Truncated icosahedron	1.42857	0.293040
Truncated dodecahedron	1.42857	0.293040
Icosidodecahedron	1.42857	0.293040

zero). In this case, we take one vertex of the regular polyhedra on the x_1-axis, so σ_L and σ_T are invariant for the rotation around the x_1-axis. Similar to the case of dilatational eigenstrain, it is noted here that the stresses have the logarithmic singularity at the vertex. However, the values of stresses at the center do not coincide with the sphere except for the dodecahedron and the icosahedron. It can be seen that the stress distributions inside the inclusion tend to be flat (or constant) as the shape of the inclusion approaches to the sphere through Figs. 6 and 7.

To obtain the strain energy, we have to evaluate the averaged Eshelby's tensor, Eq. (29) numerically. For this purpose, we divide the polyhedron into three-dimensional simplexes whose vertices are denoted by a vertex, a midedge, a midface, and the center *O*. Then a numerical integration formula by Hammer et al. [21] are used. The number of the integration points inside a simplex is 15. Table 3 shows normalized strain energy $\bar{W}^* = W^*/(\varepsilon_0^2 E V_\Omega)$ for an infinite body with a polyhedral inclusion with a dilatational eigenstrain ($\varepsilon_{ij}^* = \varepsilon_0 \delta_{ij}$) and a uniaxial eigenstrain ($\varepsilon_{11}^* = \varepsilon_0$ and other components are zero). The strain energies for all the polyhedron with a dilatational eigenstrain are the same as the sphere. On the other hand, the strain energies for the uniaxial eigenstrain are different but those for the dodecahedron, icosahedron, truncated icosahedron, truncated dodecahedron, and icosidodecahedron are the same as the sphere's. We note that they belong to the icosidodeca family which has the highly symmetrical structure. This finding is seems to be relevant to the fact that the values of stress at the center of these polyhedra are the same as those of the sphere (see Figs. 7(*a*) and (*b*) for the dodecahedron and icosahedron). This is interesting because the famous C_{60}: Buckminsterfullerene has the truncated icosahedral structure ([26]) and most quasi-crystals have icosahedral phase ([27]).

Finally, we calculate stiffness of an SiC particle-reinforced Al composite by the Mori-Tanaka method ([18,28]):

$$C_{ijkl}^c = C_{ijkl}^m + f(C_{ijmn}^f - C_{ijmn}^m) A_{mnkl}^{MT},$$

$$A_{ijkl}^{MT} = A_{ijmn}^{dil}\{(1-f) I_{mnkl} + f A_{mnkl}^{dil}\}^{-1}. \tag{35}$$

As seen in Table 3, strain energies of the polyhedra belonging to the icosidodeca family are the same as the sphere's. On the other hand, strain energies of tetrahedron, hexahedron (cube), and octahedron vary with their orientation (the numerical results are omitted). This means that polyhedra belonging to the icosidodeca family are the isotropic shapes and tetrahedron, hexahedron (cube), and octahedron are the anisotropic shapes. Therefore, for these three shapes, averaging over the orientation is needed to calculate the isotropic stiffness. However, in the case of polyhedra, it is difficult to calculate the average values. Thus, in this example, we assume that all inhomogeneities' orientations are the same: we take one vertex of the regular polyhedra on the x_1-axis. So resulting stiffness tensors have weak anisotropy in the case of tetrahedral, hexahedral (cubic), and octahedral inhomogeneities. We conjecture that if we can average $\bar{S}_{ijkl}$ over orientation of the above three polyhedra, the orientation-dependent part of $\bar{S}_{ijkl}$ will be canceled like the two-dimensional polygon case ([29]) and we will obtain the same $\bar{S}_{ijkl}$ as a sphere. Figures 8(*a*) and (*b*) show the variation of the stiffness C_{1111}^c and C_{1212}^c versus volume fraction of the particle *f*. It can be seen that the effect of particle shape is comparatively small, thus practically, use of a sphere for pre-

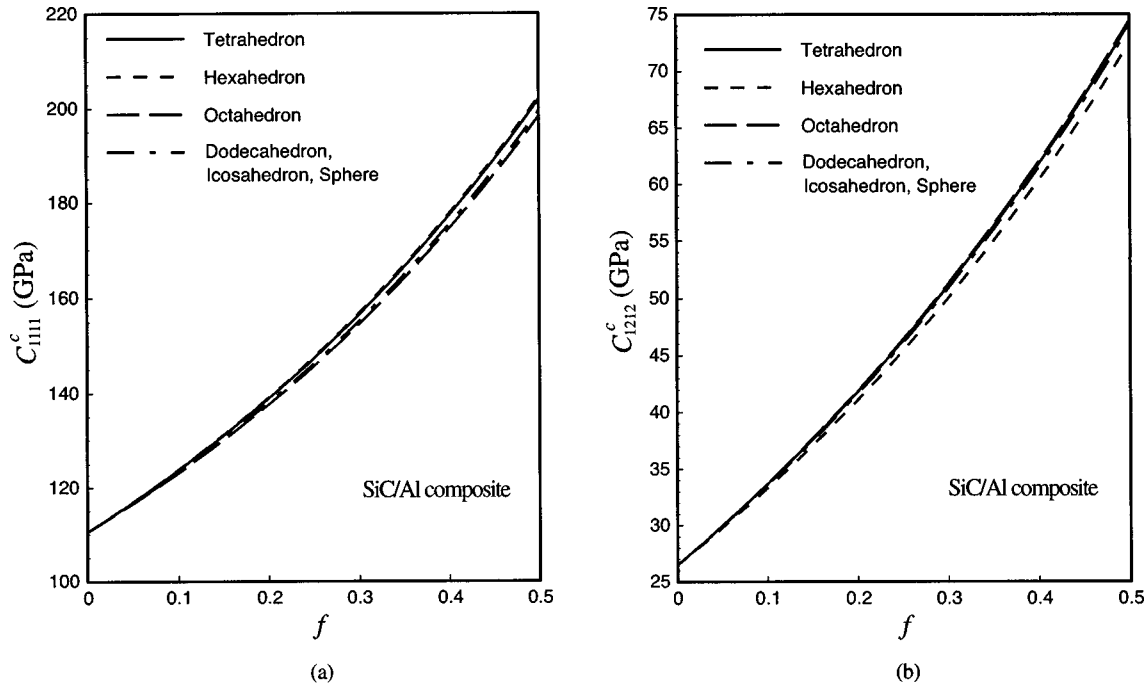


Fig. 8 Effective stiffness of SiC-Al composite versus volume fraction of inhomogeneities *f*: (*a*) C_{1111}^c, (*b*) C_{1212}^c

diction of the stiffness of a composite with regular polyhedral inhomogeneities may be justified as a good approximation.

Appendix

Integral Formulas.

$$\int \frac{1}{\sqrt{ax^2+bx+c}}dx=\frac{1}{\sqrt{a}}\log\left|2ax+b+2\sqrt{a(ax^2+bx+c)}\right| \quad (a>0), \quad (A1)$$

$$\int \frac{1}{(ax^2+bx+c)^{3/2}}dx=\frac{2(2ax+b)}{(4ac-b^2)\sqrt{ax^2+bx+c}}, \quad (A2)$$

$$\int \frac{x}{(ax^2+bx+c)^{3/2}}dx=\frac{2(bx+2c)}{(b^2-4ac)\sqrt{ax^2+bx+c}}, \quad (A3)$$

$$\int \frac{x^2}{(ax^2+bx+c)^{3/2}}dx=\frac{(2b^2-4ac)x+2bc}{a(4ac-b^2)\sqrt{ax^2+bx+c}}+\frac{1}{a\sqrt{a}}\log\left|2ax+b+2\sqrt{a(ax^2+bx+c)}\right| \quad (a>0), \quad (A4)$$

$$\int \frac{x}{(x^2+p^2)\sqrt{ax^2+c}}dx=\begin{cases}\dfrac{1}{p\sqrt{c-ap^2}}\tan^{-1}\left(\dfrac{x}{p}\sqrt{\dfrac{c-ap^2}{ax^2+c}}\right) & (c>ap^2),\\ \dfrac{1}{2p\sqrt{ap^2-c}}\log\left|\dfrac{x\sqrt{ap^2-c}+p\sqrt{ax^2+c}}{x\sqrt{ap^2-c}-p\sqrt{ax^2+c}}\right| & (c<ap^2)\end{cases}, \quad (A5)$$

$$\int \frac{x}{(x^2+p^2)\sqrt{ax^2+c}}dx=\begin{cases}\dfrac{1}{\sqrt{ap^2-c}}\tan^{-1}\sqrt{\dfrac{ax^2+c}{ap^2-c}} & (c<ap^2),\\ \dfrac{1}{2\sqrt{c-ap^2}}\log\left|\dfrac{\sqrt{ax^2+c}+\sqrt{c-ap^2}}{\sqrt{ax^2+c}-\sqrt{c-ap^2}}\right| & (c>ap^2)\end{cases}, \quad (A6)$$

Explicit Expression of the Eshelby's Tensor.

$$S_{ijmn}=\frac{1}{2}(D_{imn,j}+D_{jmn,i}) \quad (A7)$$

In Eq. (A7),

$$D_{imn,j}=\frac{C_{klmn}}{16\pi\mu(1-\nu)}\sum_{I=1}^{N} n_l^I\left[(3-4\nu)\delta_{ik}\frac{\partial I_1^I}{\partial x_j}+\frac{\partial I_2^{I(ik)}}{\partial x_j}\right]$$

$$=\frac{C_{klmn}}{16\pi\mu(1-\nu)}\sum_{I=1}^{N}\sum_{J=1}^{M_I} n_l^I\left[(3-4\nu)\delta_{ik}\left\{s_1\frac{\partial I_1^{IJ}}{\partial x_j}+s_1'\frac{\partial (I_1^{IJ})'}{\partial x_j}\right\}+\left\{s_2\frac{\partial I_2^{IJ(ik)}}{\partial x_j}+s_2'\frac{\partial (I_2^{IJ(ik)})'}{\partial x_j}\right\}\right], \quad (A8)$$

where $\partial I_1^{IJ}/\partial x_j$ is

$$\frac{\partial I_1^{IJ}}{\partial x_j}=|\mathbf{a}_J^I|_{,j}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}+|\mathbf{a}_J^I|\left\{\frac{|\mathbf{b}_J^I|_{,j}+|\boldsymbol{\gamma}_J^I|_{,j}}{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}-\frac{|\boldsymbol{\rho}_J^I|_{,j}}{|\boldsymbol{\rho}_J^I|}\right\}+|\boldsymbol{\rho}^I|_{,j}F_{IJ}+|\boldsymbol{\rho}^I|F_{IJ,j}, \quad (A9)$$

$$F_{IJ,j}=\frac{\partial F_{IJ}}{\partial x_j}=f_1^{IJ(j)}+f_2^{IJ(j)}+f_3^{IJ(j)}, \quad (A10)$$

$$f_1^{IJ(j)}=\frac{(|\mathbf{b}_J^I|_{,j}|\boldsymbol{\gamma}_J^I|+|\mathbf{b}_J^I||\boldsymbol{\gamma}_J^I|_{,j})|\boldsymbol{\rho}^I||\mathbf{a}_J^I|-|\mathbf{b}_J^I||\boldsymbol{\gamma}_J^I|(|\boldsymbol{\rho}^I|_{,j}|\mathbf{a}_J^I|+|\boldsymbol{\rho}^I||\mathbf{a}_J^I|_{,j})}{|\boldsymbol{\rho}^I|^2|\mathbf{a}_J^I|^2+|\mathbf{b}_J^I|^2|\boldsymbol{\gamma}_J^I|^2},$$

$$f_2^{IJ(j)}=-\frac{(|\mathbf{b}_J^I|_{,j}|\mathbf{a}_J^I|+|\mathbf{b}_J^I||\mathbf{a}_J^I|_{,j})|\boldsymbol{\rho}^I||\boldsymbol{\gamma}_J^I|+|\mathbf{b}_J^I||\mathbf{a}_J^I|(|\boldsymbol{\rho}^I|_{,j}|\boldsymbol{\gamma}_J^I|+|\boldsymbol{\rho}^I||\boldsymbol{\gamma}_J^I|_{,j})}{|\boldsymbol{\rho}^I|^2|\boldsymbol{\gamma}_J^I|^2+|\mathbf{b}_J^I|^2|\mathbf{a}_J^I|^2}, \quad (A11)$$

$$f_3^{IJ(j)}=-\frac{|\mathbf{b}_J^I|_{,j}|\mathbf{a}_J^I|-|\mathbf{b}_J^I||\mathbf{a}_J^I|_{,j}}{|\mathbf{c}_J^I|^2},$$

and $\partial I_2^{IJ(ik)}/\partial x_j$ is given by

$$\frac{\partial I_2^{IJ(ik)}}{\partial x_j}=\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_1+\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_2+\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_3, \quad (A12)$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_1=2|\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,j}|\mathbf{a}_J^I|_{,k}\left[2\frac{|\boldsymbol{\gamma}_J^I|-|\boldsymbol{\rho}_J^I|}{|\mathbf{b}_J^I|}-\frac{|\boldsymbol{\gamma}_J^I|-|\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|}+\frac{|\mathbf{a}_J^I|^2}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\mathbf{c}_J^I|^2|\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|^2|\mathbf{a}_J^I|}F_{IJ}\right]$$
$$+|\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,k}\left[2\frac{(|\boldsymbol{\gamma}_J^I|_{,j}-|\boldsymbol{\rho}_J^I|_{,j})|\mathbf{a}_J^I||\mathbf{b}_J^I|-(|\boldsymbol{\gamma}_J^I|-|\boldsymbol{\rho}_J^I|)(|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|+|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j})}{|\mathbf{b}_J^I|^2}\right.$$
$$\left.-\frac{(|\boldsymbol{\gamma}_J^I|_{,j}-|\boldsymbol{\rho}^I|_{,j})|\mathbf{a}_J^I||\mathbf{b}_J^I|-(|\boldsymbol{\gamma}_J^I|-|\boldsymbol{\rho}^I|)(|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|+|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j})}{|\mathbf{b}_J^I|^2}+|\mathbf{a}_J^I|^2\frac{|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|-2|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j}}{|\mathbf{b}_J^I|^3}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}\right.$$
$$\left.+\frac{|\mathbf{a}_J^I|^3}{|\mathbf{b}_J^I|^2}\left\{\frac{|\mathbf{b}_{J,j}^I|+|\boldsymbol{\gamma}_{J,j}^I|}{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}-\frac{|\boldsymbol{\rho}_J^I|_{,j}}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\mathbf{c}_J^I|^2|\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|^2|}F_{IJ,j}\right.$$
$$\left.+\frac{(2|\mathbf{c}_J^I||\mathbf{c}_{J,j}^I|\boldsymbol{\rho}^I|+|\mathbf{c}_J^I|^2|\boldsymbol{\rho}^I|_{,j})|\mathbf{b}_J^I||\mathbf{a}_J^I|-2c_J^I|^2|\boldsymbol{\rho}^I|(|\mathbf{b}_{J,j}^I|\mathbf{a}_J^I|+|\mathbf{b}_J^I||\mathbf{a}_{J,j}^I|)}{|\mathbf{b}_J^I|^3|\mathbf{a}_J^I|}F_{IJ}\right], \tag{A13}$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_2=-\{|\mathbf{a}_J^I|_{,i}|\mathbf{c}_J^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\mathbf{c}_J^I|_{,i}\}|\mathbf{a}_J^I|_{,j}|\mathbf{c}_J^I|\left[\frac{|\boldsymbol{\rho}^I|-|\boldsymbol{\rho}_J^I|}{|\mathbf{a}_J^I||\mathbf{b}_J^I|}+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|^2}F_{IJ}\right]-\{|\mathbf{a}_J^I|_{,k}(\delta_{ij}-|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,j})$$
$$-|\mathbf{a}_J^I|_{,i}(\delta_{kj}-|\boldsymbol{\rho}^I|_{,k}|\boldsymbol{\rho}^I|_{,j})\}|\mathbf{a}_J^I|\left[\frac{|\boldsymbol{\rho}^I|-|\boldsymbol{\rho}_J^I|}{|\mathbf{a}_J^I||\mathbf{b}_J^I|}+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|^2}F_{IJ}\right]-\{|\mathbf{a}_J^I|_{,i}|\mathbf{c}_J^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\mathbf{c}_J^I|_{,i}\}|\mathbf{a}_J^I||\mathbf{c}_J^I|$$
$$\times\left[\frac{(|\boldsymbol{\rho}^I|_{,j}-|\boldsymbol{\rho}_J^I|_{,j})|\mathbf{a}_J^I||\mathbf{b}_J^I|-(|\boldsymbol{\rho}^I|-|\boldsymbol{\rho}_J^I|)(|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|+|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j})}{|\mathbf{a}_J^I|^2|\mathbf{b}_J^I|^2}+\frac{|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|-2|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j}}{|\mathbf{b}_J^I|^3}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}\right.$$
$$\left.+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\left\{\frac{|\mathbf{b}_{J,j}^I|+|\boldsymbol{\gamma}_{J,j}^I|}{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}-\frac{|\boldsymbol{\rho}_J^I|_{,j}}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\boldsymbol{\rho}^I|_{,j}|\mathbf{b}_J^I|-2|\boldsymbol{\rho}^I||\mathbf{b}_J^I|_{,j}}{|\mathbf{b}_J^I|^3}F_{IJ}+\frac{|\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|^2}F_{IJ,j}\right]-\{|\mathbf{a}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}|\}\{|\boldsymbol{\rho}^I||\mathbf{a}_{J,j}^I|+|\boldsymbol{\rho}^I|_{,j}|\mathbf{a}_J^I|\}$$
$$\times\left[\frac{1}{|\mathbf{a}_J^I|}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}-\frac{|\mathbf{c}_J^I|}{|\mathbf{a}_J^I||\mathbf{b}_J^I|}\log\left\{\frac{|\mathbf{c}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}^I|}\right\}-\frac{1}{|\mathbf{b}_J^I|}\log\left\{\frac{|\boldsymbol{\rho}^I|}{|\boldsymbol{\rho}_J^I|+|\mathbf{a}_J^I|}\right\}\right]-\{|\mathbf{a}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}|\}|\boldsymbol{\rho}^I||\mathbf{a}_J^I|$$
$$\left[-\frac{|\mathbf{a}_J^I|}{|\mathbf{a}_J^I|^2}\log\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}+\frac{1}{|\mathbf{a}_J^I|}\left\{\frac{|\mathbf{b}_{J,j}^I|+|\boldsymbol{\gamma}_{J,j}^I|}{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}-\frac{|\boldsymbol{\rho}_J^I|_{,j}}{|\boldsymbol{\rho}_J^I|}\right\}-\frac{|\mathbf{c}_J^I|_{,j}|\mathbf{a}_J^I||\mathbf{b}_J^I|-|\mathbf{c}_J^I|(|\mathbf{a}_{J,j}^I|\mathbf{b}_J^I|+|\mathbf{a}_J^I||\mathbf{b}_{J,j}^I|)}{|\mathbf{a}_J^I|^2|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{c}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}^I|}\right\}\right.$$
$$\left.-\frac{|\mathbf{c}_J^I|}{|\mathbf{a}_J^I||\mathbf{b}_J^I|}\left\{\frac{|\mathbf{c}_{J,j}^I|+|\boldsymbol{\gamma}_{J,j}^I|}{|\mathbf{c}_J^I|+|\boldsymbol{\gamma}_J^I|}-\frac{|\boldsymbol{\rho}^I|_{,j}}{|\boldsymbol{\rho}^I|}\right\}+\frac{|\mathbf{b}_J^I|_{,j}}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\boldsymbol{\rho}^I|}{|\boldsymbol{\rho}_J^I|+|\mathbf{a}_J^I|}\right\}-\frac{1}{|\mathbf{b}_J^I|}\left\{\frac{|\boldsymbol{\rho}^I|_{,j}}{|\boldsymbol{\rho}^I|}-\frac{|\boldsymbol{\rho}_J^I|_{,j}+|\mathbf{a}_J^I|_{,j}}{|\boldsymbol{\rho}_J^I|+|\mathbf{a}_J^I|}\right\}\right], \tag{A14}$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_3=\{(\delta_{ij}-|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,j})|\mathbf{c}_J^I|_{,k}+|\mathbf{c}_J^I|_{,i}(\delta_{kj}-|\boldsymbol{\rho}^I|_{,k}|\boldsymbol{\rho}^I|_{,j})\}|\mathbf{c}_J^I|\left[-\frac{|\mathbf{a}_J^I|(|\boldsymbol{\gamma}_J^I||\boldsymbol{\rho}^I|)}{|\mathbf{b}_J^I||\mathbf{c}_J^I|^2}+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|^2}F_{IJ}\right]$$
$$+|\mathbf{c}_J^I|_{,i}|\mathbf{c}_J^I|_{,k}|\mathbf{c}_J^I|^2\left[-\frac{\{|\mathbf{a}_{J,j}^I(|\boldsymbol{\gamma}_J^I|-|\boldsymbol{\rho}^I|)+|\mathbf{a}_J^I|(|\boldsymbol{\gamma}_{J,j}^I|-|\boldsymbol{\rho}^I|_{,j})\}|\mathbf{b}_J^I||\mathbf{c}_J^I|-|\mathbf{a}_J^I|(|\boldsymbol{\gamma}_J^I|-|\boldsymbol{\rho}^I|)\{|\mathbf{b}_{J,j}^I|\mathbf{c}_J^I|-2|\mathbf{b}_J^I||\mathbf{c}_{J,j}^I|\}}{|\mathbf{b}_J^I|^2|\mathbf{c}_J^I|^3}\right.$$
$$\left.+\frac{|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|-2|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j}}{|\mathbf{b}_J^I|^3}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\left\{\frac{|\mathbf{b}_{J,j}^I|+|\boldsymbol{\gamma}_{J,j}^I|}{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}-\frac{|\boldsymbol{\rho}_J^I|_{,j}}{|\boldsymbol{\rho}_J^I|}\right\}+\frac{|\boldsymbol{\rho}^I|_{,j}|\mathbf{b}_J^I|-2|\boldsymbol{\rho}^I||\mathbf{b}_{J,j}^I|}{|\mathbf{b}_J^I|^3}F_{IJ}+\frac{|\boldsymbol{\rho}^I|}{|\mathbf{b}_J^I|^2}F_{IJ,j}\right]$$
$$-\left[\{(\delta_{ij}-|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,j})|\boldsymbol{\rho}^I|_{,k}+(\delta_{kj}-|\boldsymbol{\rho}^I|_{,k}|\boldsymbol{\rho}^I|_{,j})|\boldsymbol{\rho}^I|_{,i}\}|\boldsymbol{\rho}^I|+\{|\mathbf{c}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}+|\mathbf{c}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k}|\}|\boldsymbol{\rho}^I|_{,j}|\mathbf{c}_J^I\right|\left[\frac{|\mathbf{a}_J^I|}{|\mathbf{c}_J^I||\mathbf{b}_J^I|}\log\left\{\frac{|\mathbf{c}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}^I|}\right\}\right.$$
$$\left.+\frac{1}{|\mathbf{b}_J^I|}\log\left\{\frac{|\boldsymbol{\rho}^I|}{|\boldsymbol{\rho}_J^I|+|\mathbf{a}_J^I|}\right\}\right]-\{|\mathbf{c}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k}+|\mathbf{c}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}\}|\mathbf{c}_J^I||\boldsymbol{\rho}^I|\left[\frac{|\mathbf{a}_{J,j}^I|\mathbf{c}_J^I||\mathbf{b}_J^I|-|\mathbf{a}_J^I|\{|\mathbf{c}_{J,j}^I|\mathbf{b}_J^I|+|\mathbf{c}_J^I||\mathbf{b}_{J,j}^I|\}}{|\mathbf{c}_J^I|^2|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{c}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}^I|}\right\}\right.$$
$$\left.+\frac{|\mathbf{a}_J^I|}{|\mathbf{c}_J^I||\mathbf{b}_J^I|}\left\{\frac{|\mathbf{c}_{J,j}^I|+|\boldsymbol{\gamma}_{J,j}^I|}{|\mathbf{c}_J^I|+|\boldsymbol{\gamma}_J^I|}-\frac{|\boldsymbol{\rho}^I|_{,j}}{|\boldsymbol{\rho}^I|}\right\}-\frac{|\mathbf{b}_J^I|_{,j}}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\boldsymbol{\rho}^I|}{|\boldsymbol{\rho}^I|+|\mathbf{a}_J^I|}\right\}+\frac{1}{|\mathbf{b}_J^I|}\left\{\frac{|\boldsymbol{\rho}^I|_{,j}}{|\boldsymbol{\rho}^I|}-\frac{|\boldsymbol{\rho}_J^I|_{,j}+|\mathbf{a}_J^I|_{,j}}{|\boldsymbol{\rho}_J^I|+|\mathbf{a}_J^I|}\right\}\right]$$
$$-\{|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,k}+|\boldsymbol{\rho}^I|_{,k}|\boldsymbol{\rho}^I|_{,i}\}|\boldsymbol{\rho}^I|_{,j}F_{IJ}-|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,k}(F_{IJ,j}|\boldsymbol{\rho}^I|-F_{IJ}|\boldsymbol{\rho}^I|_{,j}). \tag{A15}$$

$\partial(I_1^{IJ})'/\partial x_j$ and $\partial(I_2^{IJ(ik)})'/\partial x_j$ are obtained from Eqs. (A9)–(A15) by replacing the vectors related with T_{IJ} by those related with T'_{IJ}.

Limiting Solutions. (i) When $|\mathbf{a}_J^I|\to 0$,

$$I_1^{IJ}\to 0,$$

$$I_2^{IJ(ij)}\to 0, \tag{A16}$$

$$\frac{\partial I_1^{IJ}}{\partial x_j}\to|\mathbf{a}_J^I|_{,j}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}-|\mathbf{a}_J^I|_{,j}\left\{\frac{|\boldsymbol{\rho}_J^I|^2}{|\mathbf{b}_J^I||\boldsymbol{\gamma}_J^I|}+\frac{|\mathbf{b}_J^I|}{|\boldsymbol{\gamma}_J^I|}-\frac{|\boldsymbol{\rho}_J^I|}{|\mathbf{b}_J^I|}\right\},$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_1\to-|\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,j}|\mathbf{a}_J^I|_{,k}\left[\frac{|\boldsymbol{\gamma}_J^I|}{|\mathbf{b}_J^I|}-\frac{|\boldsymbol{\rho}_J^I|^2}{|\mathbf{b}_J^I||\boldsymbol{\gamma}_J^I|}-\frac{|\mathbf{b}_J^I|}{|\boldsymbol{\gamma}_J^I|}\right],$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ij)})_2 \to 0,$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ij)})_3 \to |\mathbf{b}_J^I|_{,i}|\mathbf{b}_J^I|_{,k}|\mathbf{a}_J^I|_{,j}\left[-\frac{|\boldsymbol{\gamma}_J^I|-2|\boldsymbol{\rho}_J^I|}{|\mathbf{b}_J^I|}+\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}\right.$$
$$\left.-\frac{|\boldsymbol{\rho}_J^I|^2}{|\mathbf{b}_J^I||\boldsymbol{\gamma}_J^I|}-\frac{|\mathbf{b}_J^I|}{|\boldsymbol{\gamma}_J^I|}\right]-\{|\mathbf{b}_J^I|_{,i}|\boldsymbol{\rho}_J^I|_{,k}+|\mathbf{b}_J^I|_{,k}|\boldsymbol{\rho}_J^I|_{,i}\}|\mathbf{a}_J^I|_{,j}$$
$$\times\left[\frac{|\boldsymbol{\rho}_J^I|}{|\mathbf{b}_J^I|}\log\left\{\frac{|\mathbf{b}_J^I|+|\boldsymbol{\gamma}_J^I|}{|\boldsymbol{\rho}_J^I|}\right\}-1\right]$$
$$-|\boldsymbol{\rho}_J^I|_{,i}|\boldsymbol{\rho}_J^I|_{,k}|\boldsymbol{\rho}_J^I||\mathbf{a}_J^I|_{,j}\left[\frac{1}{|\mathbf{b}_J^I|}-\frac{|\boldsymbol{\rho}_J^I|}{|\mathbf{b}_J^I||\boldsymbol{\gamma}_J^I|}-\frac{|\mathbf{b}_J^I|}{|\boldsymbol{\rho}_J^I||\boldsymbol{\gamma}_J^I|}\right]. \tag{A17}$$

(ii) When $|\mathbf{b}_J^I|\to 0$,

$$I_1^{IJ}\to 0,$$
$$I_2^{IJ(ik)}\to 0, \tag{A18}$$
$$\frac{\partial I_1^{IJ}}{\partial x_j}\to\frac{(|\boldsymbol{\rho}_J^I|-|\boldsymbol{\rho}^I|)|\mathbf{b}_J^I|_{,j}}{|\mathbf{a}_J^I|},$$
$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_1 \to |\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,k}\left[\frac{|\boldsymbol{\rho}_J^I||\mathbf{b}_J^I|_{,j}}{|\mathbf{a}_J^I|}\right.$$
$$\left.+(|\mathbf{b}_J^I|_{,j}+|\boldsymbol{\rho}_J^I|_{,j})\left(\frac{1}{|\mathbf{b}_J^I|}-\frac{1}{|\boldsymbol{\rho}_J^I|}\right)\frac{|\mathbf{a}_J^I|^3}{|\boldsymbol{\rho}_J^I|^2}\right],$$
$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_2 \to -2|\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,k}|\mathbf{a}_J^I|^2\left[\left(\frac{1}{2|\mathbf{a}_J^I||\boldsymbol{\rho}_J^I|}+\frac{|\boldsymbol{\rho}^I|}{|\mathbf{a}_J^I|}\right)|\mathbf{b}_J^I|_{,j}\right.$$
$$\left.+(|\mathbf{b}_J^I|_{,j}+|\boldsymbol{\rho}_J^I|_{,j})\left(\frac{1}{|\mathbf{b}_J^I|}-\frac{1}{|\boldsymbol{\rho}_J^I|}\right)\frac{|\mathbf{a}_J^I|}{|\boldsymbol{\rho}_J^I|^2}\right]$$
$$-\{|\mathbf{a}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}\}\left[\frac{1}{|\boldsymbol{\rho}_J^I|}\right.$$
$$\left.-\frac{1}{2|\mathbf{a}_J^I|^2}\log\left\{\frac{|\mathbf{a}_J^I|+|\boldsymbol{\rho}_J^I|}{|\boldsymbol{\rho}^I|}\right\}\right]|\boldsymbol{\rho}^I||\mathbf{b}_J^I|_{,j},$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_3 \to |\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,k}\left[\frac{|\boldsymbol{\rho}_J^I||\mathbf{b}_J^I|_{,j}}{|\mathbf{a}_J^I|}\right.$$
$$\left.+(|\mathbf{b}_J^I|_{,j}+|\boldsymbol{\rho}_J^I|_{,j})\left(\frac{1}{|\mathbf{b}_J^I|}-\frac{1}{|\boldsymbol{\rho}_J^I|}\right)\frac{|\mathbf{a}_J^I|^3}{|\boldsymbol{\rho}_J^I|^2}\right]$$
$$+\{|\mathbf{a}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}\}$$
$$\times\frac{|\boldsymbol{\rho}^I||\mathbf{b}_J^I|_{,j}}{2|\mathbf{a}_J^I|^2}\log\left\{\frac{|\mathbf{a}_J^I|+|\boldsymbol{\rho}_J^I|}{|\boldsymbol{\rho}^I|}\right\}-|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,k}\left(\frac{|\boldsymbol{\rho}_J^I|}{|\mathbf{a}_J^I|}\right.$$
$$\left.-\frac{|\mathbf{a}_J^I|}{|\boldsymbol{\rho}_J^I|}-\frac{|\boldsymbol{\rho}^I|}{|\mathbf{a}_J^I|}\right). \tag{A19}$$

(iii) When $|\mathbf{c}_J^I|\to 0$,

$$I_1^{IJ}\to 0,$$
$$I_2^{IJ(ik)}\to 0, \tag{A20}$$
$$\frac{\partial I_1^{IJ}}{\partial x_j}\to 0,$$
$$\frac{\partial I_2^{IJ(ik)}}{\partial x_j}\to 0. \tag{A21}$$

(iv) When $|\boldsymbol{\rho}^I|\to 0$,

$$I_1^{IJ}\to|\mathbf{a}_J^I|\log\left\{\frac{|\mathbf{b}_J^I||\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\},$$

$$(I_2^{IJ(ik)})_1\to|\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,k}|\mathbf{a}_J^I|\left[\frac{|c_J^I|-2|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|}+\frac{|\mathbf{a}_J^I|^2}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}\right],$$

$$(I_2^{IJ(ik)})_2\to\{|\mathbf{a}_J^I|_{,i}|\mathbf{c}_J^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\mathbf{c}_J^I|_{,i}\}|\mathbf{c}_J^I|\left[\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|}-\frac{|\mathbf{a}_J^I|^2}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}\right],$$

$$(I_2^{IJ(ik)})_3\to|\mathbf{c}_J^I|_{,i}|\mathbf{c}_J^I|_{,k}|\mathbf{c}_J^I|\left[-\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|}+\frac{|\mathbf{a}_J^I||\mathbf{c}_J^I|}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}\right], \tag{A22}$$

$$\frac{\partial I_1^{IJ}}{\partial x_{,j}}\to|\mathbf{a}_J^I|_{,j}\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}+|\mathbf{a}_J^I|\left\{\frac{|\mathbf{b}_J^I|_{,j}+|\mathbf{c}_J^I|_{,j}}{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}-\frac{|\mathbf{a}_J^I|_{,j}}{|\mathbf{a}_J^I|}\right\}+|\boldsymbol{\rho}^I|_{,j}\tan^{-1}\frac{|\mathbf{b}_J^I|}{|\mathbf{a}_J^I|},$$

$$\frac{\partial}{\partial x_{,j}}(I_2^{IJ(ik)})_1 \to 2|\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,j}|\mathbf{a}_J^I|_{,k}\left[\frac{|\mathbf{c}_J^I|-2|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|}+\frac{|\mathbf{a}_J^I|^2}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}\right]$$
$$+|\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,k}|\mathbf{a}_J^I|^2\left[\frac{(|\mathbf{c}_J^I|_{,j}-2|\mathbf{a}_J^I|_{,j}+|\boldsymbol{\rho}^I|_{,j})|\mathbf{a}_J^I||\mathbf{b}_J^I|-(|\mathbf{c}_J^I|-2|\mathbf{a}_J^I|)(|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|+|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j})}{|\mathbf{b}_J^I|^2}\right.$$
$$\left.+\frac{|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|-2|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j}}{|\mathbf{b}_J^I|^3}|\mathbf{a}_J^I|^2\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}+\frac{|\mathbf{a}_J^I|^3}{|\mathbf{b}_J^I|^2}\left\{\frac{|\mathbf{b}_J^I|_{,j}+|\mathbf{c}_J^I|_{,j}}{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}-\frac{|\mathbf{a}_J^I|_{,j}}{|\mathbf{a}_J^I|}\right\}+\frac{|\mathbf{c}_J^I|^2|\boldsymbol{\rho}^I|_{,j}}{|\mathbf{b}_J^I|^2}\tan^{-1}\frac{|\mathbf{b}_J^I|}{|\mathbf{a}_J^I|}\right],$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_2 \to -\{|\mathbf{a}_J^I|_{,i}|\mathbf{c}_J^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\mathbf{c}_J^I|_{,j}\}|\mathbf{a}_J^I|_{,j}|c_J^I|\left[-\frac{1}{|\mathbf{b}_J^I|}+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}\right]-\{|\mathbf{a}_J^I|_{,k}(\delta_{ij}-|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,j})$$
$$+|\mathbf{a}_J^I|_{,i}(\delta_{kj}-|\boldsymbol{\rho}^I|_{,k}|\boldsymbol{\rho}^I|_{,j})\}|\mathbf{a}_J^I|\left[-\frac{1}{|\mathbf{b}_J^I|}+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}\right]-\{|\mathbf{a}_J^I|_{,i}|\mathbf{c}_J^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\mathbf{c}_J^I|_{,i}\}|\mathbf{a}_J^I||\mathbf{c}_J^I|$$
$$\times\left[\frac{(|\boldsymbol{\rho}^I|_{,j}-|\mathbf{a}_J^I|_{,j})|\mathbf{b}_J^I|+(|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|+|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j})}{|\mathbf{a}_J^I||\mathbf{b}_J^I|^2}+\frac{|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|-2|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j}}{|\mathbf{b}_J^I|^3}\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\left\{\frac{|\mathbf{b}_J^I|_{,j}+|\mathbf{c}_J^I|_{,j}}{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}-\left|\frac{\mathbf{a}_J^I|_{,j}}{|\mathbf{a}_J^I|}\right\}\right.$$
$$\left.+\frac{|\boldsymbol{\rho}^I|_{,j}}{|\mathbf{b}_J^I|^2}\tan^{-1}\frac{|\mathbf{b}_J^I|}{|\mathbf{a}_J^I|}\right]-\{|\mathbf{a}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}+|\mathbf{a}_J^I|_{,i}|\rho^I|_{,k}\}|\boldsymbol{\rho}^I|_{,j}\left[\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}-\frac{|\mathbf{c}_J^I|}{|\mathbf{b}_J^I|}\log\frac{2|\mathbf{c}_J^I|}{|\boldsymbol{\rho}^I|}-\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|}\log\frac{|\boldsymbol{\rho}^I|}{2|\mathbf{a}_J^I|}\right]$$
$$-\{|\mathbf{a}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}\}\left[\frac{|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}-1\right]\frac{|\mathbf{a}_J^I||\boldsymbol{\rho}^I|_{,j}}{|\mathbf{b}_J^I|},$$

$$\frac{\partial}{\partial x_j}(I_3^{IJ(ik)})_2 \to \{(\delta_{ij}-|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,j})|\mathbf{c}_J^I|_{,k}+|\mathbf{c}_J^I|_{,i}(\delta_{kj}-|\boldsymbol{\rho}^I|_{,k}|\boldsymbol{\rho}^I|_{,j})\}\left[-\frac{|\mathbf{a}_J^I|}{\mathbf{b}_J^I}+\frac{|\mathbf{a}_J^I||\mathbf{c}_J^I|}{|\mathbf{b}_J^I|^2}\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}\right]+|\mathbf{c}_J^I|_{,i}|\mathbf{c}_J^I|_{,k}|\mathbf{c}_J^I|^2$$
$$\left[-\frac{\{|\mathbf{a}_J^I|_{,j}|\mathbf{c}_J^I|+|\mathbf{a}_J^I|(|\mathbf{c}_J^I|_{,j}-|\boldsymbol{\rho}^I|_{,j})\}|\mathbf{b}_J^I|-|\mathbf{a}_J^I|\{|\mathbf{b}_J^I|_{,j}|\mathbf{c}_J^I|-2|\mathbf{b}_J^I||\mathbf{c}_J^I|_{,j}\}}{|\mathbf{b}_J^I|^2|\mathbf{c}_J^I|^2}+\frac{|\mathbf{a}_J^I|_{,j}|\mathbf{b}_J^I|-2|\mathbf{a}_J^I||\mathbf{b}_J^I|_{,j}}{|\mathbf{b}_J^I|^3}\log\left\{\frac{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}{|\mathbf{a}_J^I|}\right\}\right.$$
$$\left.+\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|^2}\left\{\frac{|\mathbf{b}_J^I|_{,j}+|\mathbf{c}_J^I|_{,j}}{|\mathbf{b}_J^I|+|\mathbf{c}_J^I|}-\frac{|\mathbf{a}_J^I|_{,j}}{|\mathbf{a}_J^I|}\right\}-\frac{|\boldsymbol{\rho}^I|_{,j}}{|\mathbf{b}_J^I|^2}\tan^{-1}\frac{|\mathbf{b}_J^I|}{|\mathbf{a}_J^I|}\right]-\{|\mathbf{c}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}+|\mathbf{c}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k}\}|\boldsymbol{\rho}^I|_{,j}\left[\frac{|\mathbf{a}_J^I|}{|\mathbf{b}_J^I|}\log\frac{2|\mathbf{c}_J^I|}{|\boldsymbol{\rho}^I|}+\frac{|\mathbf{c}_J^I|}{|\mathbf{b}_J^I|}\log\frac{|\boldsymbol{\rho}^I|}{2|\mathbf{a}_J^I|}\right]$$
$$+\{|\mathbf{c}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k}+|\mathbf{c}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}\}[|\mathbf{a}_J^I|-|\mathbf{c}_J^I|]\frac{|\boldsymbol{\rho}^I|_{,j}}{|\mathbf{b}_J^I|}+|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,j}|\boldsymbol{\rho}^I|_{,k}\tan^{-1}\frac{|\mathbf{b}_J^I|}{|\mathbf{a}_J^I|}. \quad (A23)$$

(v) When $|\boldsymbol{\rho}_J^I|\to 0$,

$$I_1^{IJ}\to 0,$$
$$I_2^{IJ(ik)}\to 0, \quad (A24)$$
$$\frac{\partial I_1^{IJ}}{\partial x_j}\to|\mathbf{a}_J^I|_{,j}\left[\log\frac{2|\mathbf{b}_J^I|}{|\boldsymbol{\rho}_J^I|}-1\right]-|\boldsymbol{\rho}^I|_{,j}\theta,$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_1 \to |\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,j}|\mathbf{a}_J^I|_{,k}(1-\cos^2\theta)-|\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,j}|\boldsymbol{\rho}^I|_{,k}$$
$$\times(\theta-\cos\theta\sin\theta),$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_2 \to -(|\mathbf{a}_J^I|_{,i}|\mathbf{b}_J^I|_{,k}+|\mathbf{a}_J^I|_{,k}|\mathbf{b}_J^I|_{,i})(|\boldsymbol{\rho}^I|_{,j}-|\mathbf{a}_J^I|_{,j}\sin\theta$$
$$-|\boldsymbol{\rho}^I|_{,j}\cos\theta)-(|\mathbf{a}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}+|\mathbf{a}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k})$$
$$\times[|\boldsymbol{\rho}^I|_{,j}\log(\cos\theta)-\sin\theta(|\boldsymbol{\rho}^I|_{,j}\sin\theta$$
$$-|\mathbf{a}_J^I|_{,j}\cos\theta)],$$

$$\frac{\partial}{\partial x_j}(I_2^{IJ(ik)})_3 \to |\mathbf{b}_J^I|_{,i}|\mathbf{b}_J^I|_{,k}\left[|\mathbf{a}_J^I|_{,j}\left(\log\frac{2|\mathbf{b}_J^I|}{|\boldsymbol{\rho}_J^I|}-2\right)\right.$$
$$\left.-|\boldsymbol{\rho}^I|_{,j}\theta\right]-(|\mathbf{b}_J^I|_{,k}|\boldsymbol{\rho}^I|_{,i}+|\mathbf{b}_J^I|_{,i}|\boldsymbol{\rho}^I|_{,k})$$
$$\times\left[|\boldsymbol{\rho}^I|_{,j}\left(\log\frac{\cos\theta}{1+\sin\theta}+1\right)-\{|\mathbf{a}_J^I|_{,j}(1+\sin\theta)\right.$$
$$\left.+|\boldsymbol{\rho}^I|_{,j}\cos\theta\}\frac{\cos\theta}{1+\sin\theta}\right]+|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,j}|\boldsymbol{\rho}^I|_{,k}\theta$$
$$+|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,k}\cos\theta(|\mathbf{a}_J^I|_{,j}\cos\theta-|\boldsymbol{\rho}^I|_{,j}\sin\theta), \quad (A25)$$

where $\theta=\tan^{-1}|\mathbf{a}_J^I|/|\boldsymbol{\rho}^I|$

(vi) When $|\boldsymbol{\gamma}_J^I|\to 0$,

$$I_1^{IJ}\to 0,$$
$$I_2^{IJ(ik)}\to 0. \quad (A26)$$

Limiting solutions for $\partial I_1^{IJ}/\partial x_j$ and $\partial I_2^{IJ(ik)}/\partial x_j$ are obtained by replacing $|\mathbf{a}_J^I|$, $|\mathbf{b}_J^I|$, $|\mathbf{c}_J^I|$, $|\boldsymbol{\rho}^I|$, $|\boldsymbol{\rho}_J^I|$, $|\boldsymbol{\gamma}_J^I|$, $|\boldsymbol{\rho}_J^I|_{,j}$, $|\mathbf{c}_J^I|_{,j}$, and $|\boldsymbol{\gamma}_J^I|_{,j}$ in Eq. (A13) to (A15) by $|\mathbf{a}_J^I|'$, $|\mathbf{b}_J^I|'$, $|\mathbf{c}_J^I|'$, $|\boldsymbol{\rho}^I|'$, $|\boldsymbol{\rho}_J^I|'$, $|\boldsymbol{\gamma}_J^I|'$, $|\boldsymbol{\rho}_J^I|'_{,}$, $|\mathbf{c}_J^I|'_{,j}$, and $|\boldsymbol{\gamma}_J^I|'_{,j}$ shown below.

$$|\mathbf{a}_J^I|'=\sin\phi\cos\varphi,$$

$$|\mathbf{b}_J^I|' = \sin\phi \sin\varphi,$$

$$|\mathbf{c}_J^I|' = \sin\phi,$$

$$|\boldsymbol{\rho}^I|' = \cos\phi,$$

$$|\boldsymbol{\rho}_J^I|' = \sqrt{\cos^2\phi + \sin^2\phi\cos^2\varphi},$$

$$|\boldsymbol{\gamma}_J^I|' = 1,$$

$$|\boldsymbol{\rho}_J^I|'_{,j} = \frac{(|\boldsymbol{\rho}^I|_{,j}\cos\phi + |\mathbf{a}_J^I|_{,j}\sin\phi\cos\varphi)}{\sqrt{\cos^2\phi + \sin^2\phi\cos^2\varphi}},$$

$$|\mathbf{c}_J^I|'_{,j} = |\mathbf{a}_J^I|_{,j}\cos\varphi + |\mathbf{b}_J^I|_{,j}\sin\varphi,$$

$$|\boldsymbol{\gamma}_J^I|'_{,j} = |\boldsymbol{\rho}^I|_{,j}\cos\phi + |\mathbf{a}_J^I|_{,j}\sin\phi\cos\varphi + |\mathbf{b}_J^I|_{,j}\sin\phi\sin\varphi. \tag{A27}$$

In equations (*A*25), $\phi = \tan^{-1}|\mathbf{c}_J^I|/|\boldsymbol{\rho}^I|$, $\varphi = \tan^{-1}|\mathbf{b}_J^I|/|\mathbf{a}_J^I|$.

In Eqs. (*A*9), (*A*13), (*A*14), (*A*15), (*A*17), (*A*19), (*A*22), (*A*23), (*A*25), and (*A*27), derivatives are given by

$$|\mathbf{a}_J^I|_{,i} = -(a_J^I)_i/|\mathbf{a}_J^I|,$$

$$|\mathbf{B}_J^I|_{,i} = -(b_J^I)_i/|\mathbf{b}_J^I|,$$

$$|\mathbf{c}_J^I|_{,i} = -(c_J^I)_i/|\mathbf{c}_J^I|,$$

$$|\boldsymbol{\rho}^I|_{,i} = -(\rho^I)_i/|\boldsymbol{\rho}^I|,$$

$$|\boldsymbol{\rho}_J^I|_{,i} = -(\rho_J^I)_i/|\boldsymbol{\rho}_J^I|,$$

$$|\boldsymbol{\gamma}_J^I|_{,i} = -(\gamma_J^I)_i/|\boldsymbol{\gamma}_J^I|, \tag{A28}$$

$$(a_J^I)_{i,j} = -|\mathbf{a}_J^I|_{,i}|\mathbf{a}_J^I|_{,j},$$

$$(c_J^I)_{i,j} = -(\delta_{ij} - |\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,j}),$$

$$(\rho^I)_{i,j} = -|\boldsymbol{\rho}^I|_{,i}|\boldsymbol{\rho}^I|_{,j}. \tag{A29}$$

References

[1] Eshelby, J. D., 1957, ''The Determination of the Elastic Field of an Ellipsoidal Inclusion and Related Problems,'' Proc. R. Soc. London, Ser. A, **A241**, pp. 376–396.

[2] Mura, T., 1987, *Micromechanics of Defects in Solids*, 2nd Ed., Martinus Nijhoff, Dordrecht.

[3] Faivre, G., 1964, ''Déformations de cohérence d'un précipité quadratique,'' Phys. Status Solidi, **35**, pp. 249–259.

[4] Sankaran, R., and Laird, C., 1976, ''Deformation Field of a Misfitting Inclusion,'' J. Mech. Phys. Solids, **24**, pp. 251–262.

[5] Lee, J. K., and Johnson, W. C., 1977, ''Elastic Strain Energy and Interactions of Thin Square Plates Which Have Undergone a Simple Shear,'' Scr. Metall., **11**, pp. 477–484.

[6] Chiu, Y. P., 1977, ''On the Stress Field Due to Initial Strains in a Cuboid Surrounded by an Infinite Elastic Space,'' ASME J. Appl. Mech., **44**, pp. 587–590.

[7] Lee, J. K., and Johnson, W. C., 1978, ''Calculation of the Elastic Strain Field of a Cuboidal Precipitate in an Anisotropic Matrix,'' Phys. Status Solidi, **46**, pp. 267–272.

[8] Chiu, Y. P., 1978, ''On the Stress Field and Surface Deformation in a Half Space with Cuboidal Zone in Which Initial Strains Are Uniform,'' ASME J. Appl. Mech., **45**, pp. 302–306.

[9] Owen, D. R. J., 1972, ''Analysis of Fibre-Reinforced Materials by an Initial Strain Method,'' Fibre Sci. Technol., **5**, pp. 37–59.

[10] Chiu, Y. P., 1980, ''On the Internal Stresses in a Half Plane and a Layer Containing Localized Inelastic Strains or Inclusions,'' ASME J. Appl. Mech., **47**, pp. 313–318.

[11] Takao, Y., Taya, M., and Chou, T. W., 1981, ''Stress Field Due to a Cylindrical Inclusion With Constant Axial Eigenstrain in an Infinite Elastic Body,'' ASME J. Appl. Mech., **48**, pp. 853–858.

[12] Hasegawa, H., Lee, V.-G., and Mura, T., 1992, ''The Stress Fields Caused by a Circular Cylindrical Inclusion,'' ASME J. Appl. Mech., **59**, pp. S107–S114.

[13] Wu, L., and Du, S. Y., 1995, ''The Elastic Field Caused by a Circular Cylindrical Inclusion—Part I: Inside the region $x_1^2+x_2^2<a^2$, $-\infty<x_3<\infty$ Where the circular Cylindrical Inclusion is Expressed by $x_1^2+x_2^2\leqslant a^2$, $-h\leqslant x_3\leqslant h$,'' ASME J. Appl. Mech., **62**, pp. 579–584.

[14] Wu, L., and Du, S. Y., 1995, ''The Elastic Field Caused by a Circular Cylindrical Inclusion—Part II: Inside the region $x_1^2+x_2^2>a^2$, $-\infty<x_3<\infty$ Where the Circular Cylindrical Inclusion is Expressed by $x_1^2+x_2^2\leqslant a^2$, $-h\leqslant x_3\leqslant h$,'' ASME J. Appl. Mech., **62**, pp. 585–589.

[15] Rodin, G. J., 1996, ''Eshelby's Inclusion Problem for Polygons and Polyhedra,'' J. Mech. Phys. Solids, **44**, pp. 1977–1995.

[16] Markenscoff, X., 1997, ''On the Shape of the Eshelby Inclusions,'' J. Elast., **49**, pp. 163–166.

[17] Lubarda, V. A., and Markenscoff, X., 1998, ''On the Absence of Eshelby Property for Non-Ellipsoidal Inclusions,'' Int. J. Solids Struct., **35**, pp. 3405–3411.

[18] Nozaki, H., and Taya, M., 1997, ''Elastic Fields in a Polygon-Shaped Inclusion With Uniform Eigenstrains,'' ASME J. Appl. Mech., **64**, pp. 495–502.

[19] Ru, C. Q., 1999, ''Analytic Solution for Eshelby's Problem of an Inclusion of Arbitrary Shape in a Plane or Half-Plane,'' ASME J. Appl. Mech., **66**, pp. 315–322.

[20] Waldvogel, J., 1979, ''The Newtonian Potential of Homogeneous Polyhedra,'' Jrnl. of Applied Math and Phys. (ZAMP), **30**, pp. 388–398.

[21] Hammer, P. C., Marlowe, O. J., and Stroud, A. H., 1956, ''Numerical Integration over Simplexes and Cones,'' Math. Tables Aids Comput., **10**, pp. 130–137.

[22] Rodin, G. J., 1998, discussion of ''Elastic Fields in a Polygon-Shaped Inclusion With Uniform Eigenstrains,'' by N. Nozaki and M. Taya, ASME J. Appl. Mech., **65**, p. 278.

[23] Kachanov, M., Tsukrov, I., and Shafiro, B., 1994, ''Effective Moduli of Solids With Cavities of Various Shapes,'' Appl. Mech. Rev., **47**, pp. S151–S174.

[24] Jasiuk, I., Chen, J., and Thorpe, M. F., 1994, ''Elastic Moduli of Two Dimensional Materials with Polygonal and Elliptical Holes,'' Appl. Mech. Rev., **47**, pp. S18–S28.

[25] Jasiuk, I., 1995, ''Cavities VIS-A-VIS Rigid Inclusions: Elastic Moduli of Materials With Polygonal Inclusions,'' Int. J. Solids Struct., **32**, pp. 407–422.

[26] Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F., and Smalley, R. E., 1985, ''C60: Buckminsterfullerene,'' Nature (London), **318**, pp. 162–163.

[27] Shechtman, D., Blech, I., Gratias, D., and Cahn, J. W., 1984, ''Metallic Phase With Long-Range Orientational Order and No Translational Symmetry,'' Phys. Rev. Lett., **53**, pp. 1951–1953.

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

[29] Kawashita, M., and Nozaki, H., 2001, ''Eshelby Tensor of a Polygonal Inclusion and Its Special Properties,'' J. Elast., submitted for publication.

Lyapunov Exponents and Moment Lyapunov Exponents of a Two-Dimensional Near-Nilpotent System

W.-C. Xie
Associate Professor,
Solid Mechanics Division,
Faculty of Engineering,
University of Waterloo,
Waterloo, ON N2L 3G1,
Canada

The Lyapunov exponents and moment Lyapunov exponents of a near-nilpotent system under stochastic parametric excitation are studied. The system considered is the linearized system of a two-dimensional nonlinear system exhibiting a pitchfork bifurcation. The effect of stochastic perturbation in the vicinity of static pitchfork bifurcation is investigated. Approximate analytical results of Lyapunov exponent are obtained. The eigenvalue problem for the moment Lyapunov exponent is converted to a two-point boundary value problem, which is solved numerically by the method of relaxation.
[DOI: 10.1115/1.1364491]

1 Introduction

Investigations of the dynamic stability of elastic systems, such as the transverse vibration of columns and flat plates under randomly fluctuating axial loading or end displacement, frequently lead to the study of the bifurcation behavior of the solution of a nondimensional second-order differential equation of the form

$$\ddot{q}+2\beta\dot{q}-[\gamma_0+\sigma\xi(t)]q+\delta q^3=0, \tag{1}$$

where q is the generalized coordinate, β the damping constant, γ_0 and σ loading and fluctuation parameters, δ a constant depending on the geometry of the system, and $\xi(t)$ a unit Gaussian white noise process with zero mean. In the absence of stochastic fluctuation, i.e., when $\sigma=0$, as the loading parameter γ_0 is increased from negative to positive values, the system undergoes a pitchfork bifurcation from the trivial equilibrium configuration into one of the two symmetric nontrivial equilibrium configurations.

It is of practical interest to study the shift in the point of bifurcation as a result of the small stochastic perturbation to the applied load or end displacement. This can be done by examining the stability of the trivial solution of the linearized equation

$$\ddot{q}+2\beta\dot{q}-[\gamma_0+\sigma\xi(t)]q=0. \tag{2}$$

The Lyapunov exponents, which characterize the average exponential rate of growth of the solutions of system (2) for t large, are defined as

$$\lambda_q=\lim_{t\to\infty}\frac{1}{t}\log\|\mathbf{q}(t)\|, \tag{3}$$

where $\mathbf{q}(t)=\{q(t),\dot{q}(t)\}^T$ and $\|\cdot\|$ denotes a suitable norm. The trivial solution is stable with probability one if the largest Lyapunov exponent is negative, whereas it is unstable with probability one if the largest Lyapunov exponent is positive. A dynamical or *D*-bifurcation from the trivial solution occurs when the largest Lyapunov exponent vanishes. The corresponding value of γ_0 gives the point of *D*-bifurcation. This concept of stochastic bifurcation is in accord with the concept of bifurcation of the corresponding deterministic system when the stochastic perturbation is set to zero.

A different criterion has been adopted by many physicists ([1]) based on the change in the form of the stationary probability density function of the system response. A stochastic phenomenological or *P*-bifurcation is said to occur when this probability density function undergoes a qualitative change from a unimodal to a bimodal or multimodal density function. The value of the bifurcation parameter at which such a transition in the probability density function occurs has been shown ([2]) to be related to the nontrivial zero of the pth moment Lyapunov exponent of the solution of system (2) defined as ([3])

$$\Lambda_q(p)=\lim_{t\to\infty}\frac{1}{t}\log E[\|\mathbf{q}(t)\|^p], \tag{4}$$

where $E[\cdot]$ denotes expected value. The value of $\Lambda_q(p)$ characterizes the pth moment stability of system (2). The pth moment Lyapunov exponent $\Lambda_q(p)$ is a convex analytic function in p with $\Lambda_q(0)=0$, $\Lambda_q'(0)=$the largest Lyapunov exponent λ_q. If $p=\delta(\gamma_0)\neq 0$ is the nontrivial zero of $\Lambda_q(p)$, i.e., $\Lambda_q(\delta)=0$, then the value γ_0^p at which *P*-bifurcation occurs satisfies the condition $\delta(\gamma_0^p)=-d$, where $d=2$ is the dimension of system (2). The value δ is called the stability index.

The concepts of using the Lyapunov exponents and moment Lyapunov exponents in the study of bifurcations in stochastically perturbed dynamical systems were presented in ([4]). To study the dynamical effect of stochastic perturbation in the bifurcation of system (1), it is important to determine the Lyapunov exponents and moment Lyapunov exponents of the linear system (2).

When $|\gamma_0|$ is finite and $\gamma_0<0$, by applying the time scaling $\tau=\sqrt{-\gamma_0}t$, Eq. (2) can be simplified as

$$q''+2\bar{\beta}q'+[1+\bar{\sigma}\eta(\tau)]q=0 \tag{5}$$

where $\bar{\beta}=\beta(-\gamma_0)^{-1/2}$, $\bar{\sigma}=\sigma(-\gamma_0)^{-3/4}$, $\eta(\tau)$ is a unit Gaussian white noise process, and a prime denotes differentiation with respect to τ. The Lyapunov exponent $\bar{\lambda}_q$ and the moment Lyapunov exponent $\bar{\Lambda}_q(p)$ of system (5) are related to those of (2), λ_q and $\Lambda_q(p)$, by

$$\lambda_q=\sqrt{-\gamma_0}\ \bar{\lambda}_q, \quad \Lambda_q(p)=\sqrt{-\gamma_0}\ \bar{\Lambda}_q(p).$$

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. 9, 1999; final revision, Nov. 8, 2000. Associate Editor: A. A. Ferri. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

The Lyapunov exponents and moment Lyapunov exponents of system (5) have been studied by many researchers, see e.g., ([5,6]).

When $|\gamma_0|$ is small, the above time scaling cannot be applied and the Lyapunov exponents and moment Lyapunov exponents obtained for system (5) are not applicable to system (2). However, this case is of particular importance in the studies of the effect of small stochastic perturbation in the dynamical bifurcations, which is expected to occur in the vicinity of the static pitchfork bifurcation, i.e., $\gamma_0=0$.

In this paper, the Lyapunov exponents and moment Lyapunov exponents of system (2) are studied for the case that $|\gamma_0|$ is small. Note that the nonlinear term in Eq. (1) is not used in the remaining of this paper; it only serves as a motivational example. The Lyapunov exponents and moment Lyapunov exponents obtained are for the linear system (2).

2 Formulation

For the linear system (2) under stochastic perturbation, the damping term can be removed by applying the transformation $q=xe^{-\beta t}$, which results in

$$\ddot{x}+[\gamma+\sigma\xi(t)]x=0 \tag{6}$$

where $\gamma=-\gamma_0-\beta^2$. Letting $x_1=x$, $x_2=\dot{x}$, Eq. (6) may be written in the form of state equations

$$\begin{aligned}\dot{x}_1&=x_2,\\ \dot{x}_2&=-\gamma x_1-\sigma x_1\xi(t).\end{aligned} \tag{7}$$

Following Ariaratnam and Xie [7] and introducing the scaling $x_1=y_1$, $x_2=\sigma^{\alpha_1}y_2$, $\gamma=\sigma^{\alpha_2}\hat{\gamma}$, Eqs. (7) become

$$\begin{aligned}\dot{y}_1&=\sigma^{\alpha_1}y_2,\\ \dot{y}_2&=-\sigma^{\alpha_2-\alpha_1}\hat{\gamma}y_1-\sigma^{1-\alpha_1}y_1\xi(t).\end{aligned}$$

For the right-hand sides of both equations to have comparable influence, it is required that $\sigma^{\alpha_1}=\sigma^{\alpha_2-\alpha_1}=\varepsilon$, $\sigma^{1-\alpha_1}=\varepsilon^{1/2}$, where ε is a small quantity, which leads to $\alpha_1=2/3$, $\alpha_2=4/3$, and $\varepsilon=\sigma^{2/3}$. Hence

$$\begin{aligned}dy_1&=\varepsilon y_2\,dt,\\ dy_2&=-\varepsilon\hat{\gamma}y_1\,dt-\varepsilon^{1/2}y_1\,dW,\end{aligned} \tag{8}$$

where $W(t)$ is the unit Wiener process.

It is easy to show that the Lyapunov exponents and the moment Lyapunov exponents of systems (2), (6), and (8) are related as

$$\begin{aligned}\lambda_q&=-\beta+\lambda_x, \quad \Lambda_q(p)=-p\beta+\Lambda_x(p),\\ \lambda_x&=\lambda_y, \quad \Lambda_x(p)=\Lambda_y(p),\end{aligned}$$

hence

$$\lambda_q=-\beta+\lambda_y, \quad \Lambda_q(p)=-p\beta+\Lambda_y(p). \tag{9}$$

It is well known that, to unfold a nilpotent singularity, two unfolding parameters are needed. In this problem, the damping constant β and the loading parameter γ_0 are the unfolding parameters. From the scaling, it is seen that γ is of the order of $\sigma^{4/3}$ or ε^2. From Eqs. (9) and (23), β is of the order of $\sigma^{2/3}$ or ε. Since $\gamma=-\gamma_0-\beta^2$, γ_0 is of the order of $\sigma^{4/3}$ or ε^2.

In general, consider the following two-dimensional Itô's linear stochastic system

$$d\mathbf{X}=\mathbf{A}_0\mathbf{X}\,dt+\sum_{i=1}^{m}\mathbf{A}_i\mathbf{X}\,dW_i, \tag{10}$$

where $\mathbf{X}=\{X_1,X_2\}^T$, $\mathbf{A}_i$, $i=0,1,\ldots,m$, are 2×2 matrices, and W_i, $i=1,2,\ldots,m$, are independent unit Wiener processes. Apply the Khasminskii transformation ([8])

$$\frac{X_1}{\|\mathbf{X}\|}=\cos\varphi, \quad \frac{X_2}{\|\mathbf{X}\|}=\sin\varphi,$$

where $\|\mathbf{X}\|$ is the Euclidean norm of $\mathbf{X}$, and denote

$$\mathbf{s}(\varphi)=\begin{Bmatrix}\cos\varphi\\ \sin\varphi\end{Bmatrix}, \quad \hat{\mathbf{s}}(\varphi)=\begin{Bmatrix}\sin\varphi\\ -\cos\varphi\end{Bmatrix}.$$

It is well known ([3,9]) that the moment Lyapunov exponent $\Lambda_X(p)$ of system (10) is the principal simple eigenvalue of the infinitesimal operator $\mathcal{L}_p$

$$\mathcal{L}_p f(\varphi,p)=\Lambda_X(p)f(\varphi,p),$$

where $f(\varphi,p)$ is a π-periodic function and

$$\mathbf{A}(\varphi)=\sum_{i=1}^{m}\mathbf{A}_i\mathbf{s}(\varphi)\mathbf{s}^T(\varphi)\mathbf{A}_i^T,$$

$$\alpha_i(\varphi)=\mathbf{s}^T(\varphi)\mathbf{A}_i\mathbf{s}(\varphi), \quad \beta_i(\varphi)=\hat{\mathbf{s}}^T(\varphi)\mathbf{A}_i\mathbf{s}(\varphi), \quad i=0,1,\ldots,m,$$

$$Q(\varphi)=\alpha_0(\varphi)+\frac{1}{2}\mathrm{tr}[\mathbf{A}(\varphi)]-\sum_{i=1}^{m}\alpha_i^2(\varphi),$$

$$k^2(\varphi)=\sum_{i=1}^{m}\beta_i^2(\varphi),$$

$$b(\varphi,p)=-\beta_0(\varphi)+(1-p)\sum_{i=1}^{m}\alpha_i(\varphi)\beta_i(\varphi),$$

$$c(\varphi,p)=pQ(\varphi)+\frac{1}{2}p^2\sum_{i=1}^{m}\alpha_i^2(\varphi),$$

$$\mathcal{L}_p f(\varphi,p)=\frac{1}{2}k^2(\varphi)f''(\varphi,p)+b(\varphi,p)f'(\varphi,p)+c(\varphi,p)f(\varphi,p),$$

in which the prime stands for differentiation with respect to φ. For system (8),

$$\mathbf{X}=\begin{Bmatrix}y_1\\ y_2\end{Bmatrix}, \quad \mathbf{A}_0=\begin{bmatrix}0 & \varepsilon\\ -\varepsilon\hat{\gamma} & 0\end{bmatrix}, \quad \mathbf{A}_1=\begin{bmatrix}0 & 0\\ -\varepsilon^{1/2} & 0\end{bmatrix}, \tag{11}$$

and it is easy to show that the moment Lyapunov exponent $\Lambda_y(p)$ is the principal simple eigenvalue of the following system:

$$\begin{aligned}&\frac{1}{2}\cos^4\varphi f''(\varphi,p)+[p\sin\varphi\cos^3\varphi-F_2(\varphi)]f'(\varphi,p)\\ &\quad+pF_1(\varphi)f(\varphi,p)=\hat{\Lambda}_y(p)f(\varphi,p),\end{aligned} \tag{12}$$

where

$$F_1(\varphi)=(1-\hat{\gamma})\sin\varphi\cos\varphi+\frac{1}{2}\cos^2\varphi[\cos^2\varphi+(p-1)\sin^2\varphi],$$

$$F_2(\varphi)=\sin^2\varphi+\hat{\gamma}\cos^2\varphi+\sin\varphi\cos^3\varphi,$$

and

$$\Lambda_y(p)=\varepsilon\hat{\Lambda}_y(p)=\sigma^{2/3}\hat{\Lambda}_y(p). \tag{13}$$

The Lyapunov exponent is then given by

$$\lambda_y(p)=\sigma^{2/3}\lim_{p\to0}\frac{\hat{\Lambda}_y(p)}{p}=\sigma^{2/3}\left.\frac{d\hat{\Lambda}_y(p)}{dp}\right|_{p=0}. \tag{14}$$

It is important to note that, although the small term ε appears in system (8), it does not appear in the eigenvalue problem (12). Hence, the method of perturbation cannot be applied and the difficulty in solving Eq. (12) analytically is increased significantly.

Furthermore, it has been shown ([10]) that the exact value of system (6) when $\beta=0$ and $\gamma_0=0$, is $\lambda_x=0.28931\sigma^{2/3}$. It is therefore expected that, in the vicinity of $\gamma=0$, λ_x varies as $\sigma^{2/3}$. Equation (14) certainly conforms with this expectation.

The eigenvalue problem (12) for the pth moment Lyapunov exponent can also be derived using Wedig's approach ([11]). Introduce the polar coordinates (a,φ) as $y_1=a\cos\varphi$, $y_2=a\sin\varphi$, and define a pth norm $P=a^p$. The Itô equations for P and φ can be found by Itô's lemma

$$\begin{aligned} dP&=\varepsilon pPF_1(\varphi)dt-\varepsilon^{1/2}pP\sin\varphi\cos\varphi dW,\\ d\varphi&=-\varepsilon F_2(\varphi)dt-\varepsilon^{1/2}\cos^2\varphi dW. \end{aligned} \tag{15}$$

Apply a linear stochastic transformation

$$S=T(\varphi)P,\quad P=T^{-1}(\varphi)S,\quad -\frac{1}{2}\pi\leqslant\varphi\leqslant\frac{1}{2}\pi.$$

The Itô equation for the new pth norm process S is given by, from Itô's lemma,

$$\begin{aligned} dS=\varepsilon\Big\{&\frac{1}{2}\cos^4\varphi T''(\varphi)+[p\sin\varphi\cos^3\varphi-F_2(\varphi)]T'(\varphi)\\ &+pF_1(\varphi)T(\varphi)\Big\}Pdt-\varepsilon^{1/2}[\cos^2\varphi T'(\varphi)\\ &+\sin\varphi\cos\varphi T(\varphi)]PdW. \end{aligned} \tag{16}$$

For bounded and nonsingular transformation $T(\varphi)$, both processes P and S are expected to have the same stability behavior. Therefore, $T(\varphi)$ is chosen so that the drift term of the Itô differential Eq. (16) is independent of the phase process φ, so that

$$dS=\Lambda Sdt+\varepsilon^{1/2}Sg(\varphi)dW. \tag{17}$$

Comparing Eqs. (16) and (17), it is seen that such a transformation $T(\varphi)$ is given by the following equation:

$$\begin{aligned} \varepsilon\Big\{&\frac{1}{2}\cos^4\varphi T''(\varphi)+[p\sin\varphi\cos^3\varphi-F_2(\varphi)]T'(\varphi)\\ &+pF_1(\varphi)T(\varphi)\Big\}=\Lambda T(\varphi),\quad -\frac{1}{2}\pi\leqslant\varphi\leqslant\frac{1}{2}\pi, \end{aligned} \tag{18}$$

which defines an eigenvalue problem for a second-order differential operator with $T(\varphi)$ as the unknown eigenfunction and Λ the associated eigenvalue. From Eq. (17), the eigenvalue Λ is seen to be the Lyapunov exponent of the pth moment, i.e., $\Lambda=\Lambda_y(p)$. Employing Eq. (13), Eq. (18) becomes Eq. (12).

To solve Eq. (12), consider a Fourier series expansion of the eigenfunction $f(\varphi,p)$ in the form

$$f(\varphi,p)=u_0+\sum_{k=1}^{\infty}(u_k\cos 2k\varphi+v_k\sin 2k\varphi), \tag{19}$$

since the coefficients in Eq. (12) are periodic with period π. Substituting Eq. (19) into (12), equating the coefficients of like trigonometric terms $\sin 2k\varphi$ and $\cos 2k\varphi$, $k=0,1,\ldots,$ results in a system of infinitely many homogeneous linear equations for the unknown coefficients u_0, u_k, v_k, $k=1,2,\ldots,$

$$\begin{bmatrix} a_{00}^u-\hat\Lambda_y & a_{01}^u & a_{01}^v & a_{02}^u & a_{02}^v & \cdots\\ a_{10}^u & a_{11}^u-\hat\Lambda_y & a_{11}^v & a_{12}^u & a_{12}^v & \cdots\\ b_{10}^u & b_{11}^u & b_{11}^v-\hat\Lambda_y & b_{12}^u & b_{12}^v & \cdots\\ a_{20}^u & a_{21}^u & a_{21}^v & a_{22}^u-\hat\Lambda_y & a_{22}^v & \cdots\\ b_{20}^u & b_{21}^u & b_{21}^v & b_{22}^u & b_{22}^v-\hat\Lambda_y & \cdots\\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}\begin{Bmatrix} u_0\\ u_1\\ v_1\\ u_2\\ v_2\\ \vdots \end{Bmatrix}=\mathbf{0}. \tag{20}$$

The existence of nontrivial solution requires that the determinant of the coefficient matrix be equal to zero, from which the eigenvalue $\hat\Lambda_y(p)$ can be obtained in principle.

In practice, only a finite number of terms is considered to obtain an approximate value for the eigenvalue $\hat\Lambda_y(p)$. If, instead of Eq. (19), $f(\varphi,p)$ is taken as follows:

$$f(\varphi,p)=u_0+\sum_{k=1}^{K}(u_k\cos 2k\varphi+v_k\sin 2k\varphi), \tag{21}$$

system (20) is truncated to a set of $2K+1$ homogeneous linear equations for u_0, u_1, v_1, $\ldots$, u_K, v_K. To obtain nontrivial solutions, the determinant is set to zero to yield an algebraic equation of degree $2K+1$ for $\hat\Lambda_y^{(K)}(p)$

$$a_{2K+1}^{(K)}[\hat\Lambda_y^{(K)}]^{2K+1}+a_{2K}^{(K)}[\hat\Lambda_y^{(K)}]^{2K}+\cdots+a_1^{(K)}\hat\Lambda_y^{(K)}+a_0^{(K)}=0. \tag{22}$$

3 Lyapunov Exponents

Equation (22) can be employed to determine the Kth-order approximation $\hat\lambda_y^{(K)}$ of the Lyapunov exponent $\hat\lambda_y$ easily. Since $\hat\Lambda_y(p)=O(p)$ as $p\to 0$, $[\hat\Lambda_y(p)]^k=o(p)$ for $k>1$. Dividing equation (22) by p, taking the limit $p\to 0$, and using Eq. (14) results in

$$\begin{aligned} \lambda_y^{(K)}=\sigma^{2/3}\hat\lambda_y^{(K)}&=\sigma^{2/3}\lim_{p\to 0}\frac{\hat\Lambda_y^{(K)}(p)}{p}\\ &=-\sigma^{2/3}\lim_{p\to 0}\frac{a_0^{(K)}}{pa_1^{(K)}}=\frac{N^{(K)}}{D^{(K)}}\frac{\sigma^{2/3}}{16}. \end{aligned} \tag{23}$$

The above calculations can be easily manipulated by a symbolic computation software such as *Maple*. Because of the limitation of the space, $\lambda_y^{(2)}$ and $\lambda_y^{(3)}$ are given in the following equation and $N^{(10)}$ and $D^{(10)}$ are presented in Appendix A,

$$\begin{aligned} \lambda_y^{(2)}&=\frac{\sigma^{2/3}}{16}\frac{9807+1016\hat\gamma+3828\hat\gamma^2-384\hat\gamma^3+96\hat\gamma^4}{2313+2544\hat\gamma+2248\hat\gamma^2+960\hat\gamma^3+144\hat\gamma^4},\\ \lambda_y^{(3)}&=\frac{\sigma^{2/3}}{16}\frac{204505+95464\hat\gamma+134748\hat\gamma^2+17920\hat\gamma^3+12832\hat\gamma^4-768\hat\gamma^5+128\hat\gamma^6}{62329+94768\hat\gamma+90728\hat\gamma^2+45248\hat\gamma^3+17808\hat\gamma^4+3584\hat\gamma^5+256\hat\gamma^6}. \end{aligned} \tag{24}$$

To check the accuracy of the Lyapunov exponents $\lambda_y^{(K)}$ obtained, $\lambda_y^{(K)}=\hat\lambda_y^{(K)}\sigma^{2/3}$ for the nilpotent system, i.e., $\beta=0$ and $\gamma_0=0$ in Eq. (6), which are obtained by setting $\hat\gamma=0$ in Eq. (23), are compared with the exact value $\lambda_y=0.28931\sigma^{2/3}$ obtained in Ariaratnam and Xie [10]. The results are presented in Table 1.

From Table 1, it is noted that an increase in the order K by 1 does not necessary increase the accuracy of the approximation, such as the cases for $K=4$, 5, and 6. However, it is usually true that a better approximation can be obtained when more terms are retained in Eq. (21).

For a near-nilpotent system, i.e., β and γ_0 small in system (2), which is an important case in the study of the dynamic effect of

Table 1 Comparison of the Lyapunov exponents for the nilpotent system

K	1	2	3	4	5
$\hat{\lambda}_y^{(K)}$	0.375	0.26500	0.29531	0.28957	0.28803
% error	29.62	8.40	2.07	0.089	0.44
K	6	7	8	9	10
$\hat{\lambda}_y^{(K)}$	0.28999	0.28918	0.28923	0.28937	0.28930
% error	0.24	0.043	0.026	0.020	0.0036

stochastic noise in the vicinity of the static pitchfork bifurcation, Eq. (23) gives an approximate result of the Lyapunov exponent λ_y. The Lyapunov exponent λ_q of system (2) can also be easily determined using a numerical scheme proposed by Wedig [12]. For the purpose of comparison, the Lyapunov exponents from the analytical approximate results (23) with $K=10$ are shown in Figs. 1 along with those obtained from the numerical simulation. It can be seen that both results agree extremely well.

The point of dynamic bifurcation γ_0^D as a function of the intensity of noise perturbation σ can be easily obtained by solving the equation

$$\lambda_q^{(K)} = -\beta + \frac{\sigma^{2/3}}{16}\frac{N^{(K)}(\hat{\gamma})}{D^{(K)}(\hat{\gamma})} = 0, \quad \hat{\gamma} = -\frac{\gamma_0 + \beta^2}{\sigma^{4/3}}, \qquad (25)$$

for γ_0. The results are shown in Fig. 2 for different values of damping coefficient β and $K=10$.

Using the numerical scheme proposed by Wedig [12], for given values of β and σ, the dynamic bifurcation point γ_0^D satisfies the condition $\lambda_q = 0$. The numerical results are shown in Figs. 3 together with those obtained from Eq. (25) with $K=10$. It can be seen that both results agree quite well.

The numerical scheme proposed by Wedig [12] has difficulties in determining the Lyapunov exponents in the vicinity of $\gamma_0 = 0$ and $\sigma = 0$. However, this region is smaller than the invalid region of the analytical results. In system (8), the parameter $\hat{\gamma}$ should be a small quantity, i.e., $|\gamma_0 + \beta^2|\sigma^{-4/3} < 1$. Hence, for the analytical results to be valid, it is required that $\beta^2 - \sigma^{4/3} < \gamma_0 < \beta^2 + \sigma^{4/3}$ or $\sigma > |\gamma_0 + \beta^2|^{3/4}$. By comparing the analytical results with those of numerical simulation, it is observed that σ or ε does not have to be small. In fact, the larger the value of σ, the smaller the value of $\hat{\gamma}$, leading to a better convergence of the result (23).

In the static case, $\beta = 0$ and $\sigma = 0$, the system undergoes a pitchfork bifurcation when γ_0 is increased from negative to positive values, or static bifurcation occurs at $\gamma_0 = 0$. For dynamic bifurcation, it is observed from Figs. 2 that, in the presence of damping, the dynamic bifurcation is delayed under the perturbation of small intensity noise (γ_0^D is shifted to the right), whereas the dynamic bifurcation is advanced (γ_0^D is shifted to the left) when the intensity of the noise is large.

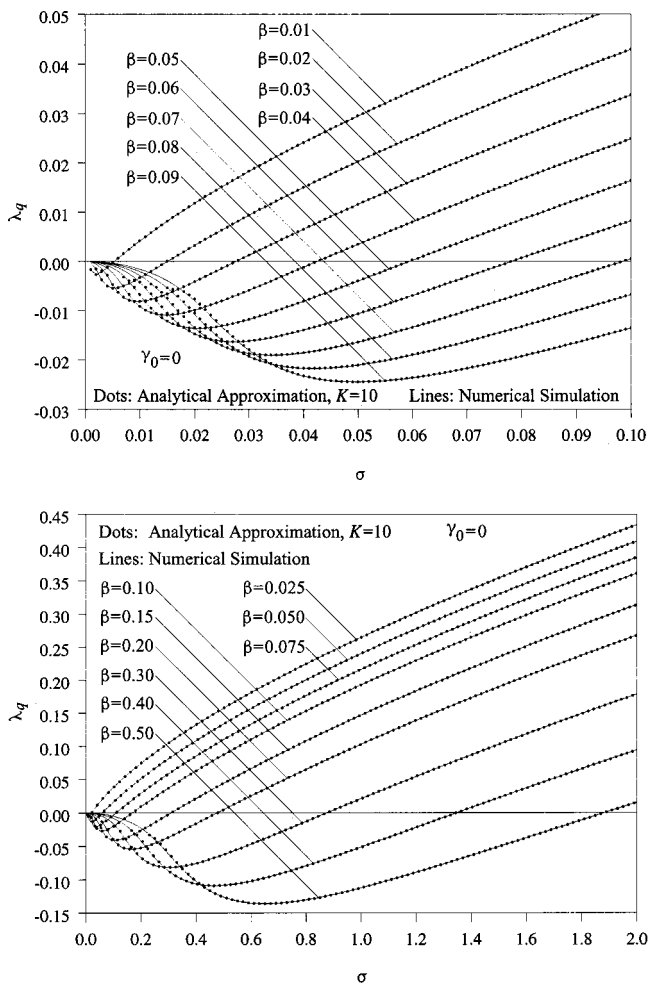


Fig. 1 Lyapunov exponents

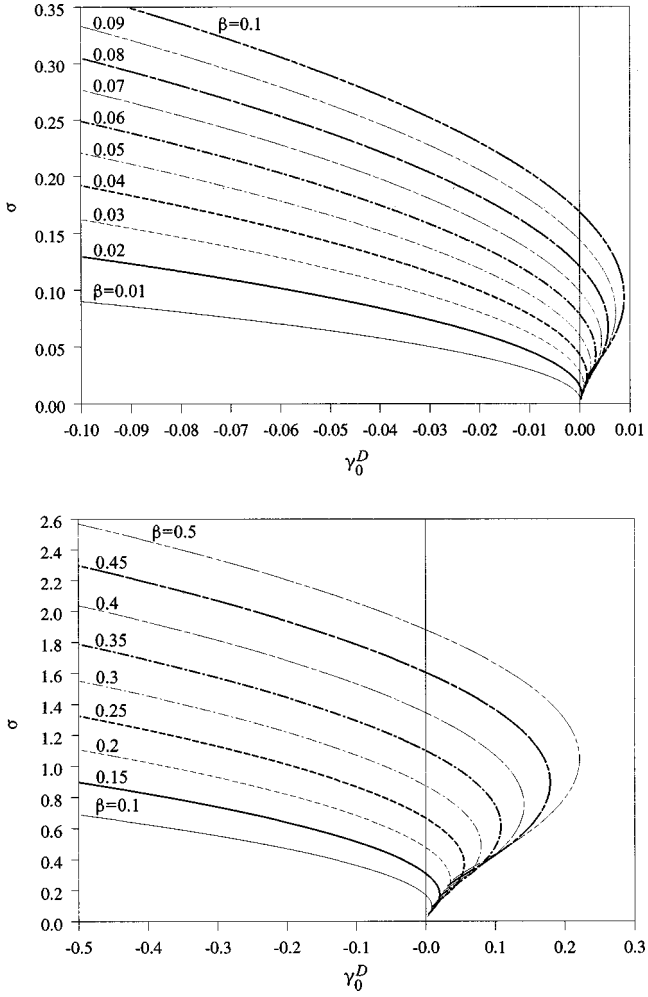


Fig. 2 Points of dynamic bifurcation

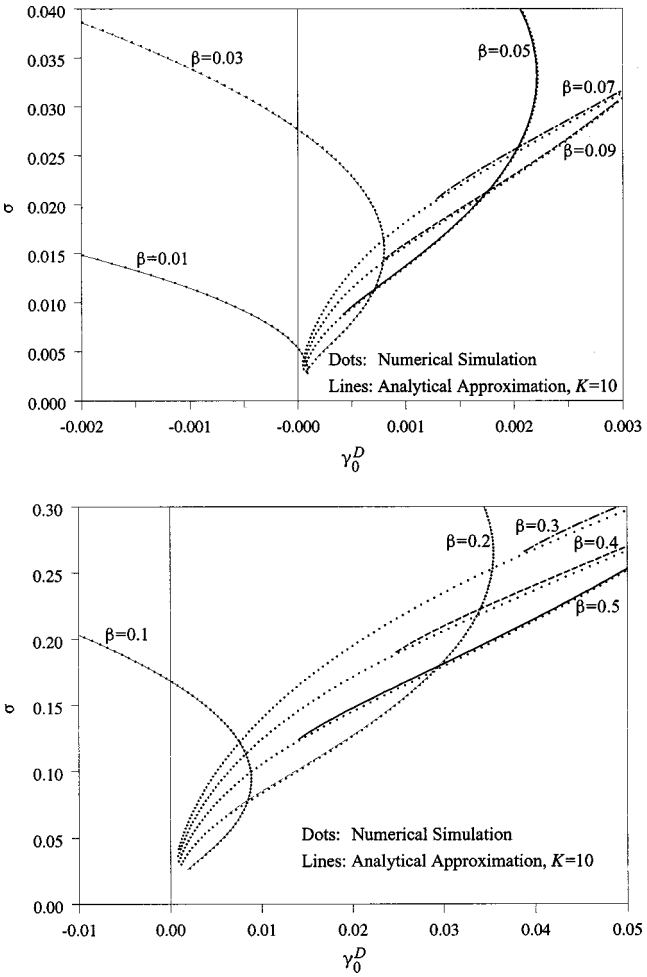


Fig. 3 Points of dynamic bifurcation

4 Moment Lyapunov Exponents

When $K=0$, i.e., when Eq. (19) is taken as $f(\varphi,p)=u_0$, Eq. (20) becomes

$$[a_{00}^u-\hat{\Lambda}_y(p)]=0,$$

where $a_{00}^u=p(p+2)/16$. Hence, the zeroth-order approximation of the moment Lyapunov exponent is

$$\hat{\Lambda}_y^{(0)}(p)=a_{00}^u=\frac{1}{16}p(p+2), \tag{26}$$

or

$$\Lambda_q^{(0)}(p)=-p\beta+\hat{\Lambda}_y^{(0)}(p)\sigma^{2/3}=-p\beta+\frac{1}{16}p(p+2)\sigma^{2/3}.$$

When $K=1$, the Eq. (22) for $\hat{\Lambda}_y^{(1)}(p)$ is a cubic equation, the real root of which can be easily obtained. However, because the expression is quite complicated and the accuracy of the result is not very high, it is not presented here.

In principle, the Kth-order approximation $\hat{\Lambda}_y^{(K)}(p)$ of the pth moment Lyapunov exponent $\hat{\Lambda}_y(p)$ can be obtained by solving Eq. (22). However, the analytical solution of Eq. (22) is not possible for $K>2$; numerical solution has to be resorted to.

The eigenvalue problem (12) can be converted to a two-point boundary value problem (41) of Appendix B. Letting $y_1=f(\varphi,p)$, $y_2=f'(\varphi,p)$, $y_3=\hat{\Lambda}_y$, Eq. (12) can be written in the following standard form:

$$y_1'=y_2,$$

$$y_2'=\frac{2}{\cos^4\varphi}\{[F_2(\varphi)-p\sin\varphi\cos^3\varphi]y_2-pF_1(\varphi)y_1+y_1y_3\}, \tag{27}$$

$$y_3'=0.$$

Since $f(\varphi,p)$ is a π-periodic function, the domain of φ can be chosen as $-1/2\pi\leqslant\varphi\leqslant1/2\pi$, as in Eq. (18). When $\varphi\to\pm1/2\pi$, $\cos\varphi\to0$ and $\sin\varphi\to\pm1$. Hence, at the boundary points $\varphi=\pm1/2\pi$, the eigenvalue problem (12) becomes

$$f'\left(\pm\frac{1}{2}\pi,p\right)=-\hat{\Lambda}_y(p)f\left(\pm\frac{1}{2}\pi,p\right).$$

Hence the boundary conditions are

$$\text{at } \varphi=-\frac{1}{2}\pi:\quad \begin{cases} y_2=-y_3y_1, \\ y_1=c, \end{cases} \tag{28}$$

$$\text{at } \varphi=\frac{1}{2}\pi:\quad \begin{cases} y_2=-y_3y_1, \\ y_1=c, \end{cases} \tag{29}$$

where c is a constant to be determined. Note that there are four boundary conditions in Eqs. (28) and (29), because the two conditions $\varphi(-1/2\pi)=c$ and $\varphi(1/2\pi)=c$ are equivalent to the condition of periodicity $\varphi(-1/2\pi)=\varphi(1/2\pi)$.

The method of relaxation can be applied to solve the two-point boundary value problem (27)–(29). The method of relaxation is advantageous when studying the variation of the moment Lyapunov exponent with the change of a parameter, since relaxation rewards a good initial guess with rapid convergence and the previous solution should be a good initial guess when the parameter is changed only slightly. Following the procedure as described in Appendix B, discretize the domain of φ into M grid points $\varphi_m=-1/2\pi+(m-1)h$, $m=1,2,\ldots,M$, where $h=\pi/(M-1)$. At the grid points φ_2, φ_3, $\ldots$, φ_M, the ordinary differential Eqs. (27) are replaced by the finite difference equations:

$$E_{1,m}\equiv y_{1,m}-y_{1,m-1}-h\bar{y}_{2,m}=0,$$

$$E_{2,m}\equiv y_{2,m}-y_{2,m-1}-\frac{2h}{\cos^4\bar{\varphi}_m}\{[F_2(\bar{\varphi}_m)-p\sin\bar{\varphi}_m\cos^3\bar{\varphi}_m]\bar{y}_{2,m}$$
$$-pF_1(\bar{\varphi}_m)\bar{y}_{1,m}+\bar{y}_{1,m}\bar{y}_{3,m}\}, \tag{30}$$

$$E_{3,m}\equiv y_{3,m}-y_{3,m-1},$$

where $\bar{\varphi}_m=1/2(\varphi_m+\varphi_{m-1})$, $\bar{y}_{i,m}=1/2(y_{i,m}+y_{i,m-1})$, $i=1,2,3$. At the first boundary point $\varphi=-1/2\pi$, there are two boundary conditions. From Eqs. (28),

$$E_{2,1}\equiv y_{2,1}+y_{1,1}y_{3,1}=0,$$
$$E_{3,1}\equiv y_{1,1}-c=0. \tag{31}$$

At the second boundary point $\varphi=1/2\pi$, there are also two boundary conditions. From Eqs. (29),

$$E_{1,M+1}\equiv y_{2,M}+y_{1,M}y_{3,M}=0,$$
$$E_{2,M+1}\equiv y_{1,M}-c=0. \tag{32}$$

The $3M$ corrections $\Delta y_{i,m}$, $i=1,2,3$, $m=1,2,\ldots,M$, and the undetermined constant c are given by the solution of the $3M+1$ linear algebraic Eqs. (46)–(48) with $n_1=n_2=2$. For $m=2,3,\ldots,M$, there are $3(M-1)$ equations in (46),

$$S_{1,1}^m=-1,\quad S_{1,2}^m=S_{1,5}^m=-\frac{1}{2}h,\quad S_{1,3}^m=S_{1,6}^m=0,\quad S_{1,4}^m=1,$$

$$S_{2,1}^m=-\frac{h}{\cos^4\bar{\varphi}_m}[-pF_1(\bar{\varphi}_m)+\bar{y}_{3,m}]=S_{2,4}^m,$$

$$S^m_{2,2}=-1-\frac{h}{\cos^4\overline{\varphi}_m}[F_2(\overline{\varphi}_m)-p\sin\overline{\varphi}_m\cos^3\overline{\varphi}_m]=S^m_{2,5}-2,$$

(33)

$$S^m_{2,3}=-\frac{h}{\cos^4\overline{\varphi}_m}\overline{y}_{1,m}=S^m_{2,6},$$

$$S^m_{3,1}=S^m_{3,2}=S^m_{3,4}=S^m_{3,5}=0,\quad S^m_{3,3}=-1,\quad S^m_{3,6}=1.$$

For the first boundary point, there are two equations in (47),

$$S^1_{2,1}=y_{3,1},\quad S^1_{2,2}=1,\quad S^1_{2,3}=y_{1,1},\tag{34}$$

$$S^1_{3,1}=1,\quad S^1_{3,2}=0,\quad S^1_{3,3}=0,\tag{35}$$

and for the second boundary point, there are two equations in (48),

$$S^{M+1}_{1,1}=y_{3,M},\quad S^{M+1}_{1,2}=1,\quad S^{M+1}_{1,3}=y_{1,M},\tag{36}$$

$$S^{M+1}_{2,1}=1,\quad S^{M+1}_{2,2}=0,\quad S^{M+1}_{2,3}=0.\tag{37}$$

By using the zeroth approximation $\hat{\Lambda}^{(0)}_y(p)=p(p+2)/16$ as an initial guess, solving the linear algebraic Eqs. (46)–(48) for $\Delta y_{i,m}$ and updating $y_{i,m}$ iteratively, an approximate value of $\hat{\Lambda}_y(p)=y_{3,1}$ is obtained. The moment Lyapunov exponent $\Lambda_q(p)$ is then determined as

$$\Lambda_q(p)=-p\beta+\hat{\Lambda}_y(p)\sigma^{2/3}.\tag{38}$$

Numerical results of the moment Lyapunov exponents $\Lambda_q(p)$ of the nilpotent system, i.e., system (2) with $\beta=0$ and $\gamma_0=0$, are plotted in Figs. 4 for various values of σ. As mentioned earlier, the slope of the moment Lyapunov exponent curve at $p=0$ is the Lyapunov exponent. Straight lines with slopes being the exact values of the Lyapunov exponent $0.28931\sigma^{2/3}$ and passing through the origin are plotted in Figs. 5 for the purpose of comparison. It is clearly seen that these straight lines are tangent to the moment Lyapunov exponent curves at the origin, which ascertains the correctness of the moment Lyapunov exponents obtained. Numerical results of the moment Lyapunov exponents $\Lambda_q(p)$ for $\beta=0.05$, $\gamma_0=0$ are plotted in Figs. 6.

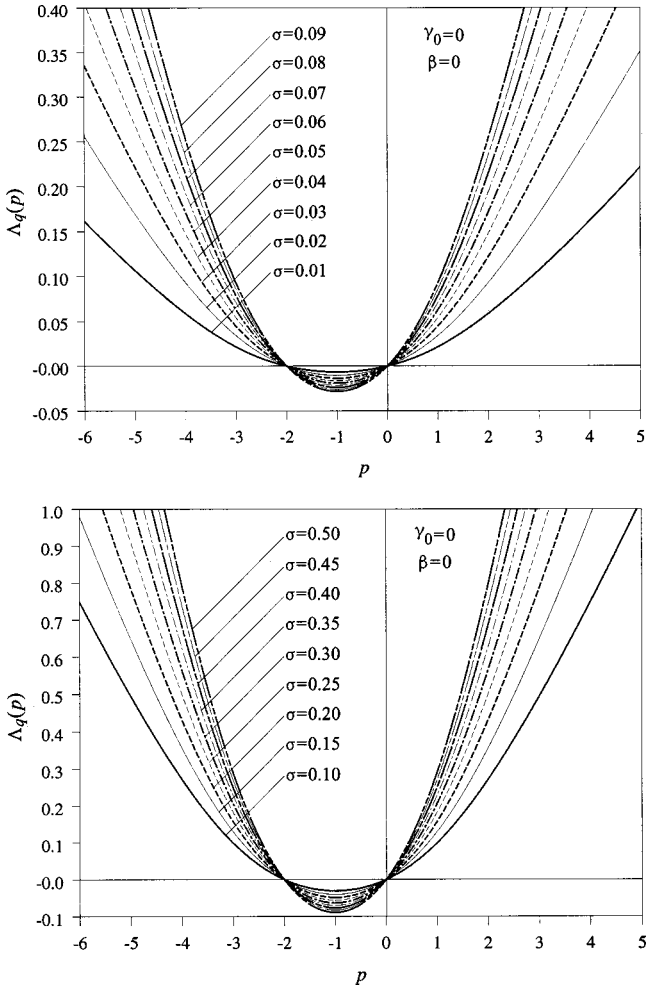


Fig. 4 Moment Lyapunov exponents of the nilpotent system

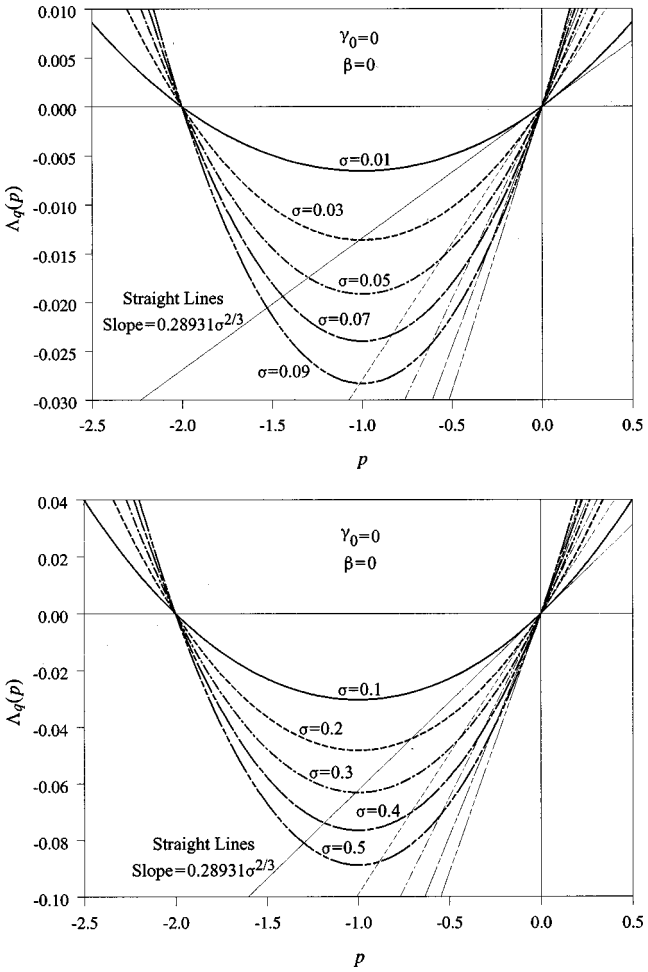


Fig. 5 Moment Lyapunov exponents of the nilpotent system

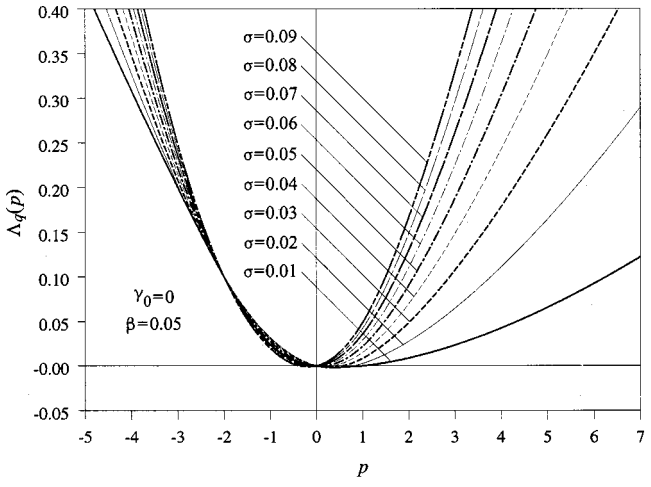
The stability index $p=\delta$ is the nontrivial zero of the moment Lyapunov exponent. For the given values of β, γ_0, and σ, the stability index δ can be determined as a root-finding problem such that $\Lambda_q(\delta)=0$. Numerical results of the stability index δ are shown in Figs. 7 for $\gamma_0=0$ and various values of β and σ.

For given values of β and σ, the point γ_0^P of P-bifurcation is given by the condition $\Lambda_q(\delta)=0$, $\delta(\gamma_0^P)=-d$, where $d=2$ is the dimension of system (2). Baxendale and Stroock [13] found a simple expression for the determination of $\Lambda(-d)$. Consider the Stratonovich stochastic differential equation

$$d\mathbf{X}=\mathbf{A}_0\mathbf{X}\,dt+\sum_{i=1}^{m}\mathbf{A}_i\mathbf{X}\circ dW_i,\tag{39}$$

in $\mathbb{R}^d$, which satisfies a Lie algebra condition as specified in Baxendale and Stroock [13]. Then

$$\Lambda_X(-d)=-\mathrm{tr}(\mathbf{A}_0)+\frac{1}{2}\sum_{i=1}^{m}[\mathrm{tr}(\mathbf{A}_i)]^2.\tag{40}$$



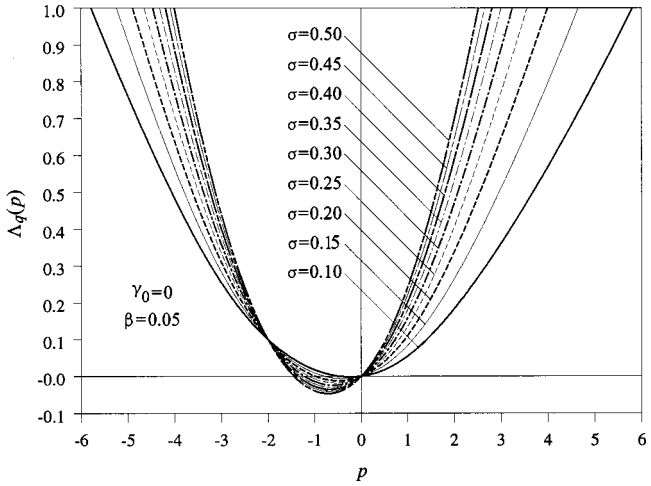
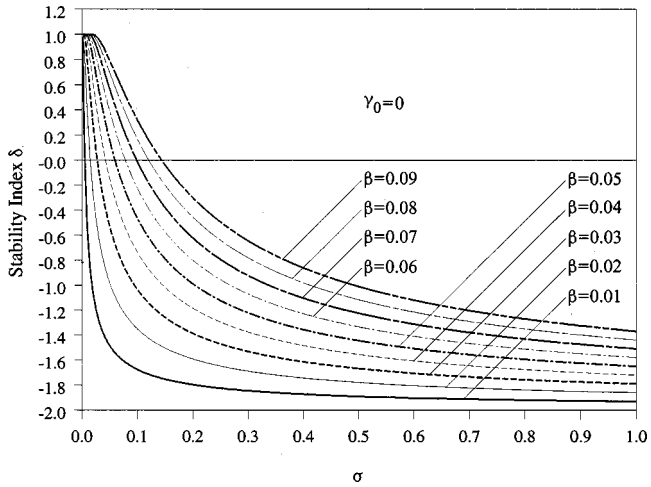


Fig. 6 Moment Lyapunov exponents



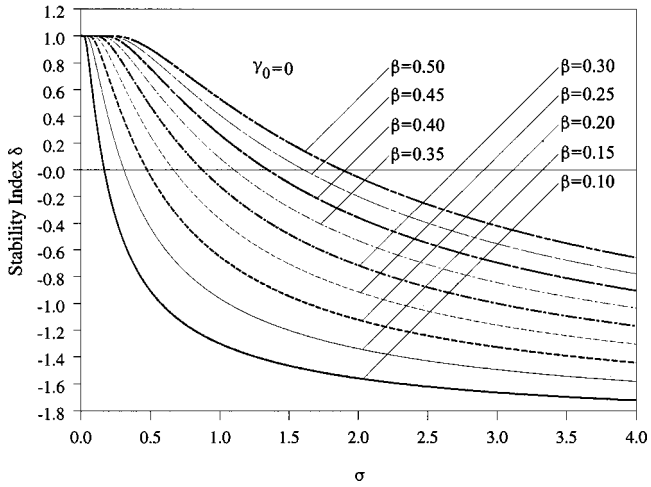


Fig. 7 Moment Lyapunov exponents

For system (8), the Stratonovich stochastic system and the associated Itô stochastic system have the same form with $\mathbf{A}_0$ and $\mathbf{A}_1$ given by Eq. (11). Using Eq. (40), $\Lambda_y(-2)=0$. Since $\Lambda_q(p)=-p\beta+\Lambda_y(p)$, hence $\Lambda_q(-2)=2\beta+\Lambda_y(-2)=2\beta$.

For a damped system, $\Lambda_q(-2)=2\beta\neq 0$, implying that the point of *P*-bifurcation does not exist or there is no *P*-bifurcation.

From Figs. 4 and 6, it is clearly seen that the numerical results of the moment Lyapunov exponent for system (2) show that $\Lambda_q(-2)=2\beta$.

5 Conclusions

In this paper, the Lyapunov exponents and moment Lyapunov exponents of a near-nilpotent system under stochastic parametric excitation is studied. The system considered is the linearized system of a two-dimensional nonlinear system exhibiting a pitchfork bifurcation. When the system is in the vicinity of the static bifurcation, the linearized system is near-nilpotent. The effect of stochastic perturbation in the vicinity of static pitchfork bifurcation is investigated. To obtain correct approximations of the Lyapunov exponent and moment Lyapunov exponents, a scaling is introduced. However, because of this scaling, the eigenvalue problem for the moment Lyapunov exponent does not contain small terms so that the method of perturbation is not applicable.

To determine the moment Lyapunov exponent, which is given by the eigenvalue of an eigenvalue problem, a Fourier series expansion of the eigenfunction is applied. Using a symbolic computation software, approximate analytical results of Lyapunov exponent can be easily obtained, which compare very well with the numerical results. For the nilpotent system, the approximate analytical Lyapunov exponents compare extremely well with the exact value of $0.28931\sigma^{2/3}$. For given values of β and σ, the point of dynamic bifurcation γ_0^D are easily determined.

On the other hand, the moment Lyapunov exponent can be determined by solving a polynomial algebraic equation. However, it is generally impossible to analytically solve an algebraic equation when the degree is higher than four. Therefore it is impossible to obtain an approximation of the moment Lyapunov exponent higher than order one. A numerical scheme for the determination of the moment Lyapunov exponents is proposed. The eigenvalue problem is converted to a two-point boundary value problem, which is solved by the method of relaxation. Numerical results of the moment Lyapunov exponent and stability index are obtained.

Acknowledgment

The research for this paper was supported, in part, by the Natural Sciences and Engineering Research Council of Canada through Grant No. OGPO131355. The author is grateful to the referees for the constructive comments which helped to improve the paper.

Appendix A

Tenth-Order Approximation of the Lyapunov Exponent.

$$N^{(10)}=8141637789159147344366 13+10506765138745300917186 64\hat{\gamma}+11281611723291441181749 88\hat{\gamma}^2$$
$$+7759432783461661497981 44\hat{\gamma}^3+4419682138565212210387 52\hat{\gamma}^4+1941952966447538155865 60\hat{\gamma}^5$$
$$+7223120244908297683596 8\hat{\gamma}^6+2195118290965633571225 6\hat{\gamma}^7+5732064023551329947648\hat{\gamma}^8$$
$$+1247296360504790646784\hat{\gamma}^9+2349681854787763077 12\hat{\gamma}^{10}+3683916017743108505 6\hat{\gamma}^{11}+4988807288488263680\hat{\gamma}^{12}$$
$$+550279102592057344\hat{\gamma}^{13}+51575234755100672\hat{\gamma}^{14}+3656803439083520\hat{\gamma}^{15}+216194152660992\hat{\gamma}^{16}$$
$$+6831118221312\hat{\gamma}^{17}+193257799680\hat{\gamma}^{18}-1845493760\hat{\gamma}^{19}+92274688\hat{\gamma}^{20}$$

$$D^{(10)}=1758912144932091905362 09+4305272579790855739167 20\hat{\gamma}+6050365730442241139675 92\hat{\gamma}^2$$
$$+5709298936214482533767 04\hat{\gamma}^3+4022525651312912764316 32\hat{\gamma}^4+2213635442961837156275 20\hat{\gamma}^5$$
$$+9844821261992515673881 6\hat{\gamma}^6+3606357775506843379302 4\hat{\gamma}^7+1105468555894003819724 8\hat{\gamma}^8$$
$$+2857986587819379523584\hat{\gamma}^9+6269373098317537935 36\hat{\gamma}^{10}+1170060259400883896 32\hat{\gamma}^{11}+1852501243764640972 8\hat{\gamma}^{12}$$
$$+2478260546110750720\hat{\gamma}^{13}+2776637563523563 52\hat{\gamma}^{14}+2545115484651520 0\hat{\gamma}^{15}+1868615977271296\hat{\gamma}^{16}$$
$$+103932625944576\hat{\gamma}^{17}+3781716606976\hat{\gamma}^{18}+71051509760\hat{\gamma}^{19}+507510784\hat{\gamma}^{20}$$

Appendix B

Two-Point Boundary Value Problems. Consider the following two-point boundary value problem of N dependent variables

$$\frac{dy_i}{dx}=g_i(x;y_1,y_2,\ldots,y_N),\quad i=1,2,\ldots,N, \tag{41}$$

with the n_1 boundary conditions at point $x=a$ and n_2 boundary conditions at point $x=b$:

$$\text{at } x=a:\quad B_j(x;y_1,y_2,\ldots,y_N)=0,\quad j=1,2,\ldots,n_1,$$

$$\text{at } x=b:\quad C_k(x;y_1,y_2,\ldots,y_N)=0,\quad k=1,2,\ldots,n_2,$$

where $n_1+n_2=N$, i.e., there are total N boundary conditions.

There are two well-documented methods for solving the two-point boundary values problem (41), i.e., the shooting method and the relaxation method ([14]). For the completeness of the presentation of this paper, the relaxation method is briefly reviewed.

Discretize the domain of solution $[a,b]$ into M grid points $x_1=a,x_2,\ldots,x_{M-1},x_M=b$. Using the backward difference

$$\left.\frac{dy_i}{dx}\right|_{x_m}\approx\frac{y_{i,m}-y_{i,m-1}}{x_m-x_{m-1}},$$

the ordinary differential Eqs. (41) are replaced by approximate finite difference equations on the $M-1$ grid points $x_2,x_3,\ldots,x_M$:

$$E_{i,m}\equiv y_{i,m}-y_{i,m-1}-(x_k-x_{k-1})g_i(\overline{x_m};\overline{y_{1,m}},\overline{y_{2,m}},\ldots,\overline{y_{N,m}})=0,$$
$$i=1,2,\ldots,N,\quad m=2,3,\ldots,M, \tag{42}$$

where $\overline{x_m}=1/2(x_m+x_{m-1})$, $\overline{y_{i,m}}=1/2(y_{i,m}+y_{i,m-1})$, and the boundary conditions are

$$E_{j,1}\equiv B_j(x_1;y_{1,1},y_{2,1},\ldots,y_{N,1})=0,$$
$$j=N-n_1+1,N-n_1+2,\ldots,N, \tag{43}$$

$$E_{j,M+1}\equiv C_k(x_M;y_{1,M},y_{2,M},\ldots,y_{N,M})=0,\quad k=1,2,\ldots,n_2, \tag{44}$$

in which, for efficient computation, for the first boundary point the last n_1 components are taken as the n_1 nonzero components, whereas for the second boundary point the first n_2 components are taken as the n_2 nonzero components.

The solution of the finite difference Eqs. (42)–(44) consists of a set of values $y_{i,m}$ of the N variables y_i, $i=1,2,\ldots,N$, at the M grid points x_m, $m=1,2,\ldots,M$. The numerical solution requires an initial guess for $y_{i,m}$. The corrections $\Delta y_{i,m}$ are determined such that $y_{i,m}+\Delta y_{i,m}$ is an improved approximation to the solution. The corrections $\Delta y_{i,m}$ can be obtained by expanding the finite difference Eqs. (42)–(44) in the first-order Taylor series with respect to the small changes $\Delta y_{i,m}$. At the grid points $x_2,x_3,\ldots,x_M$, from Eq. (42),

$$E_{i,m}(y_{1,m-1}+\Delta y_{1,m-1},\ldots,y_{N,m-1}+\Delta y_{N,m-1};y_{1,m}+\Delta y_{1,m},\ldots,y_{N,m}+\Delta y_{N,m})$$
$$\approx E_{i,m}(y_{1,m-1},\ldots,y_{N,m-1};y_{1,m},\ldots,y_{N,m})+\sum_{n=1}^{N}\frac{\partial E_{i,m}}{\partial y_{n,m-1}}\Delta y_{n,m-1}+\sum_{n=1}^{N}\frac{\partial E_{i,m}}{\partial y_{n,m}}\Delta y_{n,m}. \tag{45}$$

For the updated $y_{i,m}+\Delta y_{i,m}$ to be a solution, the updated value of $E_{i,m}$ must be zero, which leads to a set of $N(M-1)$ equations for $\Delta y_{i,m}$

$$\sum_{n=1}^{N}S_{i,n}^{m}\Delta y_{n,m-1}+\sum_{n=N+1}^{2N}S_{i,n}^{m}\Delta y_{n-N,m}=-E_{i,m}, \tag{46}$$

where

$$S_{i,n}^{m}=\frac{\partial E_{i,m}}{\partial y_{n,m-1}},\quad S_{i,n+N}^{m}=\frac{\partial E_{i,m}}{\partial y_{n,m}}.$$

Similarly, at the first boundary $x_1=a$, from Eq. (43), one has n_1 equations

$$\sum_{n=1}^{N}S_{j,n}^{1}\Delta y_{n,1}=-E_{j,1},\quad j=N-n_1+1,N-n_1+2,\ldots,N, \tag{47}$$

where

$$S_{j,n}^{1}=\frac{\partial E_{j,1}}{\partial y_{n,1}},$$

and at the second boundary $x_M=b$, from Eq. (44), one has n_2 equations

$$\sum_{n=1}^{N} S_{k,n}^{M+1}\Delta y_{n,M}=-E_{k,M+1}, \quad k=1,2,\ldots,n_2, \qquad (48)$$

where

$$S_{k,n}^{M+1}=\frac{\partial E_{j,M+1}}{\partial y_{n,M}}.$$

Equations (46)–(48) give total NM linear algebraic equations for the corrections $\Delta y_{i,m}$, $i=1,2,\ldots,N$, $m=1,2,\ldots,M$. By iteratively solving for $\Delta y_{i,m}$ and updating $y_{i,m}$, an approximate solution of the two-point boundary value problem (41) is obtained.

References

[1] Horsthemke, W., and Lefever, R., 1984, *Noise-Induced Transitions*, Springer-Verlag, Berlin.

[2] Baxendale, P. H., 1991, ''Invariant Measure for Nonlinear Stochastic Differential Equations,'' *Lyapunov Exponents* (Lecture Notes in Mathematics, 1486), L. Arnold, H. Crauel, and J.-P. Eckmann, eds., Springer-Verlag, Berlin, pp. 123–140.

[3] Arnold, L., 1984, ''A Formula Connecting Sample and Moment Stability of Linear Stochastic Systems,'' SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **44**, No. 4, pp. 793–801.

[4] Arnold, L., 1998, *Random Dynamical Systems*, Springer-Verlag, Berlin, Chapter 9.

[5] Khasminskii, R., and Moshchuk, N., 1998, ''Moment Lyapunov Exponent and Stability Index for Linear Conservative System With Small Random Perturbation,'' SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **58**, No. 1, pp. 245–256.

[6] Arnold, L., Doyle, M. M., and Sri Namachchivaya, N., 1997, ''Small Noise Expansion of Moment Lyapunov Exponents for Two-Dimensional Systems,'' Dyn. Stab. Syst., **12**, No. 3, pp. 187–211.

[7] Ariaratnam, S. T., and Xie, W.-C., 1993, ''Lyapunov Exponents and Stochastic Stability of Two-Dimensional Parametrically Excited Random Systems,'' Trans. ASME J. Appl. Mech., **60**, pp. 677–682.

[8] Khasminskii, R. Z., 1967, ''Necessary and Sufficient Conditions for the Asymptotic Stability of Linear Stochastic Systems,'' Theor. Probab. Appl., **12**, pp. 144–147 (English translation).

[9] Milstein, G. N., 1996, ''Evaluation of Moment Lyapunov Exponents for Second Order Stochastic Systems,'' Random Comput. Dyn., **4**, No. 4, pp. 301–315.

[10] Ariaratnam, S. T., and Xie, W.-C., 1990, ''Lyapunov Exponent and Rotation Number of a Two-Dimensional Nilpotent Stochastic System,'' Dyn. Stab. Syst., **5**, No. 1, pp. 1–9.

[11] Wedig, W., 1988, ''Lyapunov Exponent of Stochastic Systems and Related Bifurcation Problems,'' *Stochastic Structural Dynamics-Progress in Theory and Applications*, S. T. Ariaratnam, G. I. Schuëller, and I. Elishakoff, eds., Elsevier, New York, pp. 315–327.

[12] Wedig, W., 1995, ''Pitchfork and Hopf Bifurcations in Stochastic Systems—Effective Methods to Calculate Lyapunov Exponents,'' *Probabilistic Methods in Applied Physics*, P. Krée and W. Wedig, eds., Springer-Verlag, Berlin, pp. 120–148.

[13] Baxendale, P., and Stroock, D., 1988, ''Large Deviations and Stochastic Flows of Diffeomorphisms,'' Prob. Theory Related Fields, **80**, pp. 169–215.

[14] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, *Numerical Recipes in Fortran: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, Cambridge, UK.

Explicit Equations of Motion for Mechanical Systems With Nonideal Constraints

F. E. Udwadia
Professor of Civil Engineering, Aerospace and Mechanical Engineering, and Mathematics

R. E. Kalaba
Professor of Biomedical Engineering, Electrical Engineering, and Economics

University of Southern California, Los Angeles, CA 90089-1453

Since its inception about 200 years ago, Lagrangian mechanics has been based upon the Principle of D'Alembert. There are, however, many physical situations where this confining principle is not suitable, and the constraint forces do *work. To date, such situations are excluded from general Lagrangian formulations. This paper releases Lagrangian mechanics from this confinement, by generalizing D'Alembert's principle, and presents the explicit equations of motion for constrained mechanical systems in which the constraints are nonideal. These equations lead to a simple and new fundamental view of Lagrangian mechanics. They provide a geometrical understanding of constrained motion, and they highlight the simplicity with which Nature seems to operate.*
[DOI: 10.1115/1.1364492]

1 Introduction

One of the central problems in classical mechanics is the determination of the equations of motion for constrained systems. The importance of the problem stems from the fact that what makes a set of point masses and rigid bodies, a ''system,'' is the presence of constraints. When physical constraints are imposed on an unconstrained set of particles, forces of constraint are engendered which ensure the satisfaction of the constraints. The equations of motion developed to date for such constrained systems are based on a principle first enunciated by D'Alembert, and later elaborated by Lagrange [1] in his *Mechanique Analytique* which dates back to 1787. Today the principle is referred to as D'Alembert's principle, and it is the centerpiece of classical analytical dynamics. It states, simply, that the total work done by the forces of constraint under virtual displacements is always zero. Constraints for which D'Alembert's principle is applicable are referred to as *ideal* constraints.

Since its initial formulation by Lagrange more than 200 years ago, the problem of constrained motion has been vigorously and continuously worked on by numerous scientists including Volterra, Boltzmann, Hamel, Whittaker, and Synge, to name a few. In 1829, Gauss [2] provided a new general principle for the motion of constrained mechanical systems in what is today referred to as Gauss's Principle. About 100 years after Lagrange, Gibbs [3] and Appell [4] independently discovered what are known today as the Gibbs-Appell equations of motion ([3,4]). Pars ([5], p. 202) refers to the Gibbs-Appell equations as ([5]) '' . . . probably the most comprehensive equations of motion so far discovered.'' Dirac, because of his interest in constrained systems that arise in quantum mechanics, in a series of papers from 1951 to 1969 developed an approach for determining the Lagrange multipliers for constrained Hamiltonian systems ([6]). More recently, Udwadia and Kalaba [7] presented a simple, explicit, set of equations, applicable to general mechanical systems, with holonomic and nonholonomic constraints ([7,8]).

However, *all* these alternative descriptions of the motion of constrained systems discovered so far, as well as the numerous articles that have subsequently dealt with them, rely on D'Alembert's principle, and each of these mathematical formalisms is equivalent to the other. Despite the continuous and vigorous attention that this problem has received, the inclusion of situations where the physically generated forces of constraint in a mechanical system *do not* satisfy D'Alembert's principle has so far evaded Lagrangian dynamics. Yet, such forces of constraint are among those quite commonly found in nature. As stated by Goldstein ([9], p. 17), ''This [total work done by forces of constraint equal to zero] is no longer true if sliding friction is present, and we must exclude such systems from our [Lagrangian] formulation'' ([9]). And Pars in his treatise ([5]) on analytical dynamics (1979, p. 14) writes, ''There are in fact systems for which the principle enunciated [D'Alembert's principle] . . . does not hold. But such systems will not be considered in this book.''

In this paper we obtain the equations of motion for constrained systems where the forces of constraint indeed do not satisfy D'Alembert's principle, and the sum total of the work done by them under virtual displacements no longer need be zero.

The outline of the paper is as follows. In Section 2.1 we generalize D'Alembert's Principle to include constraint forces that *do* work. This leads us to a deeper understanding of the specification of constraints in mechanical systems. This we discuss in Section 2.2. Section 3 deals with the mathematical statement of the problem of constrained motion. Section 4 states and verifies the explicit equation of motion for constrained systems with nonideal equality constraints. This equation leads to a new and fundamental principle of Lagrangian mechanics. The proof we give here is simpler than the one given in ([10]), and it yields an important geometrical interpretation that we discuss later. Section 5 gives an example of a nonholonomically constrained system for which the constraints are nonideal. We show here the ease of applicability of the explicit equation of motion obtained in the previous section and point out the insights it provides into understanding constrained motion where the constraint forces do work. Lastly, Section 6 deals with the geometry of constrained motion and exhibits the simplicity and aesthetics with which Nature seems to operate.

2 Generalization of D'Alembert's Principle, Constraint Forces That *Do* Work, and Their Specification

2.1 Generalized D'Alembert's Principle. Consider an unconstrained system of particles, each particle having a constant mass. By ''unconstrained'' we mean that the number of generalized coordinates, n, used to describe the configuration of the sys-

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 2, 2000; final revision, Oct. 9, 2000. Associate Editor: N. C. Perkins. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

tem at any time, t, equals the number of degrees-of-freedom of the system. The Lagrangian equation of motion for such a system can be written in the form

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

where $q(t)$ is the n-vector (i.e., n by 1 vector) of generalized coordinates, M is an n by n symmetric, positive-definite matrix, Q is the ''known'' n-vector of impressed forces, and the dots refer to differentiation with respect to time. By ''known,'' we shall mean that Q is a known function of its arguments. The acceleration, a, of the unconstrained system at any time t is then given by the relation $a(q,\dot{q},t)=M^{-1}(q,t)Q(q,\dot{q},t)$.

We shall assume that this system is subjected to a set of $m=h+s$ consistent equality constraints of the form

$$\varphi(q,t)=0 \tag{2}$$

and

$$\psi(q,\dot{q},t)=0, \tag{3}$$

where φ is an h-vector and ψ an s-vector. Furthermore, we shall assume that the initial conditions q_0 and $\dot{q}_0$ satisfy these constraint equations at time $t=0$. Assuming that Eqs. (2) and (3) are sufficiently smooth, we differentiate Eq. (2) twice with respect to time, and Eq. (3) once with respect to time, to obtain the equation

$$A(q,\dot{q},t)\ddot{q}=b(q,\dot{q},t), \tag{4}$$

where the matrix A is m by n, and b is a suitably defined m-vector that results from carrying out the differentiations.

This set of constraint equations includes among others, the usual holonomic, nonholonomic, scleronomic, rheonomic, catastatic, and acatastatic varieties of constraints; combinations of such constraints may also be permitted in Eq. (4). It is important to note that Eq. (4), together with the initial conditions, is equivalent to Eqs. (2) and (3).

Consider now any instant of time t. When the equality constraints (Eqs. (2) and (3)) are imposed at that instant of time on the unconstrained system, the motion of the unconstrained system is, in general, altered from what it would have been (at that instant of time) in the absence of these constraints. We view this alteration in the motion of the unconstrained system as being caused by an additional set of forces, called the ''forces of constraint,'' acting on the system at that instant of time. The equation of motion of the constrained system can then be expressed as

$$M(q,t)\ddot{q}=Q(q,\dot{q},t)+Q^c(q,\dot{q},t), \quad q(0)=q_0,\dot{q}(0)=\dot{q}_0 \tag{5}$$

where the additional ''constraint force'' n-vector, $Q^c(q,\dot{q},t)$, arises by virtue of the constraints (2) and (3) imposed on the unconstrained system, which is described by Eq. (1). Our aim is to determine Q^c explicitly at time t in terms of the known quantities M, Q, A, b, ***and*** information about the nonideal nature of the constraint force, at time t. The latter comes from looking at the physics of the system.

A virtual displacement ([8]) at time t is any nonzero n-vector v such that $A(q,\dot{q},t)v=0$. When the constraint force n-vector does *no* work under virtual displacements v, we have $v^TQ^c=0$. This is also referred to as D'Alembert's principle, and it is the basis that underlies *all* the different formalisms ([1–8]) hereto developed of the equations of motion for mechanical systems subjected to the constraints described by Eqs. (2) and (3).

As demonstrated elsewhere ([7,8]), one formalism that yields an *explicit* equation describing the motion of such a constrained system that abides by D'Alembert's principle is given by $M\ddot{q}=Q+M^{1/2}B^{+}(b-Aa)=Q+Q^c$, where the m by n matrix $B=AM^{-1/2}$, and B^{+} stands for the Moore-Penrose generalized inverse ([11]) of the matrix B.

The central question that arises now is how to incorporate into the equation of motion, *constraints* that *do* do work under virtual displacements, thereby bringing such constraints within the Lagrangian framework. Such nonideal constraint forces (for example, sliding frictional forces) are in fact commonplace, and have to date defied ([5,9]) inclusion in a simple way within the general framework of analytical mechanics. The main reason for this difficulty is that three obstacles need to be simultaneously surmounted. Firstly, we require the specification of such constraint forces to be general enough so that they encompass problems of practical utility. Secondly, this specification must, in order to comply with physical observations, yield the accelerations of the constrained system *uniquely* when using the accepted mathware of analytical dynamics that has been developed over the last 250 years. And lastly, when the constraint forces do no work, we must obtain the usual formalisms/equations that have thus far been obtained (e.g., by Lagrange, Gibbs, Appell, and Gauss), and are known to be of practical value.

Clearly, the work done by such a constraint force under virtual displacements v at each instant of time needs to be known, and must therefore be specified using some known n-vector $C(q,\dot{q},t)$, as v^TC. Such an additional specification calls for a generalization of D'Alembert's principle. We make this generalization in the following manner:

For any virtual displacement v at time t, the constraint force n-vector Q^c at time t does a prescribed amount of work given by

$$v^TQ^c(t)=v^TC(q,\dot{q},t). \tag{6}$$

Here $C(q,\dot{q},t)$ is a known n-vector (i.e., a known function of q, $\dot{q}$, and t) that needs to be specified and depends on the physics of the situation, as discussed in the example below. The work done by the constraint force in a virtual displacement may thus be *positive, negative, or zero*.

Relation (6) constitutes a new principle. This principle requires a description of the nature of the nonideal constraint force at time t through a specification of the work it does during a virtual displacement at that time. It generalizes D'Alembert's principle, and when $C\equiv 0$, it reduces to it. In what follows we shall often refer to the constraint force n-vector, Q^c, as the constraint force.

2.2 Specification of Constraints. The equations of motion provide a mathematical model for describing the motion of any given physical mechanical system. The constraints specify the conditions that the generalized displacements and/or velocities must satisfy at each instant of time as the motion of the system ensues under the action of the impressed forces. However, the equations that state these conditions (Eqs. (2) and (3)) *do not completely specify the influence of these constraints on the motion of the mechanical system*. For short, we shall say that Eqs. (2) and (3) do not completely specify the constraints on the mechanical system. This is what the generalized D'Alembert's principle tells us.

There is a second part to the specification of the constraints, and this deals with the *nature of the forces that are created by virtue of the presence of the constraints*. For this, the mechanician who is modeling a specific mechanical system needs to study the system, possibly through experimentation, or otherwise. It is this information regarding the nonideal *nature* of the force of constraint that is encapsulated in the vector $C(q,\dot{q},t)$.

For example, consider a rigid block that is confined to move on a horizontal surface $z=0$. The specification of this relation (i.e., $z=0$) does *not* constitute a complete specification of the constraint. For, the presence of this constraint creates a constraint force, and this force influences the motion of the block. So to adequately model the motion of the block on the surface, one needs to prescribe the *nature* of this constraint force. Such a prescription is situation-specific and must be specified by the mechanician either by experimentation with the system, by observation, by analogy with other systems (s)he has experience with, or by some other means. For example, if the mechanician finds that the surfaces in contact are rough (s)he may want to perform some experiments to understand the nature of the forces created by the presence of this constraint. For a specific setup, (s)he may find

that the work done by the constraint force under virtual displacements is proportional to the speed of the block, or perhaps to the square of its speed. Thus, depending on the situation at hand, C would then be specified as $-a_0\left[\begin{smallmatrix}\dot{q}_x\\ \dot{q}_y\end{smallmatrix}\right]$ or $-a_0|\dot{q}|\left[\begin{smallmatrix}\dot{q}_x\\ \dot{q}_y\end{smallmatrix}\right]$ respectively, where a_0 may be a suitable constant whose value would also need to be prescribed (perhaps by performing more experiments). If, further, the roughness of the surface changes from location to location, additional experimentation may be warranted, and a further refinement may be required in specifying the vector C. Or, in some other situation, C may perhaps be modeled as $-a_0 Q_i^c(q,\dot{q},t)$ (see Eq. (11) below).

The invocation of D'Alembert's principle when modeling a mechanical system is then clear. D'Alembert's principle specifies the nature of the constraint forces by simply setting $C\equiv 0$. It points to the genius of Lagrange, for this specification accomplishes the following three things simultaneously.

1 It provides a condition that enables the accelerations of the constrained system to be *uniquely* determined, something desirable when dealing with mechanical systems.

2 It specifies the nature of the constraint force through the ad hoc specification of $C\equiv 0$. This allows the mechanician to model a given mechanical system *without* having to explicitly provide further information (beyond that contained in the constraint Eqs. (2) and/or (3)) on the *nature* of the constraint forces that are created by the presence of the constraints. Most importantly, it therefore obviates the need for situation-specific experimentation, observation, etc., that would have been otherwise necessary to specify C when modeling a specific mechanical system.

3 This specification of $C\equiv 0$ works well (or at least sufficiently well) in many practical situations. This is perhaps the most remarkable attribute of D'Alembert's principle, and it points to the genius of Lagrange.

All this becomes quite obvious, especially when modeling the problem of sliding friction where we immediately recognize that the equation that describes the motion of the block on a horizontal surface must depend not only of the constraint equation, $z=0$, but indeed also on the *nature* of the constraint force engendered by this constraint. And the latter depends on the physics of the specific situation—the materials in contact, the surface roughnesses, etc., and, of course, the intended use that the mechanician wants to put the model to.

But in analytical dynamics, we may have got so used to invoking D'Alembert's principle, which obviates the explicit need to specify the *nature* of the constraint force for any given mechanical system by implicitly taking $C\equiv 0$, that it is tempting to think that such a specification may be wholly unnecessary, even in general. One perhaps may then get the impression that the equations that specify the constraints (Eqs. (2) and/or (3)) are all that is necessary for properly posing the problem of constrained motion. *This indeed is not so*. Specification of the nature of the constraint forces is *always* necessary. The generalized D'Alembert's principle stated in Section 2.1 reminds us that, D'Alembert's principle provides, in fact, *one particular* specification for the nature of the constraint force. As in the case of sliding friction, C may not be zero, and its explicit specification is necessary, in general. Such a specification, as mentioned before, is situation-specific and relies on the discernment and discretion of the mechanician who is modeling the system.

Having explained what we mean by ''specification of constraints'' for a given, constrained mechanical system at hand, we now need to explicitly determine its equation of motion. We start by providing a statement of the problem of constrained motion.

3 General Statement of the Problem of Constrained Motion With Constraints That Do Work

In the notation that we have thus far developed, the problem of constrained motion can now be mathematically stated as follows.

We require to find the n-vector $Q^c(q,\dot{q},t)$ such that

1 $M(q,t)\ddot{q}=Q(q,\dot{q},t)+Q^c(q,\dot{q},t)$, with $q(0)=q_0$, $\dot{q}(0)=\dot{q}_0$, and Q a known function of q, $\dot{q}$, and t; (S1)

2 $\varphi(q,t)=0$, $\psi(q,\dot{q},t)=0$, with $\varphi(q_0,0)=0$; $\dot{\phi}(q_0,t)=0$, and $\psi(q_0,\dot{q}_0,0)=0$; and, (S2)

3 for all vectors v such that $A(q,\dot{q},t)v=0$, we require $v^TQ^c(t)=v^TC(q,\dot{q},t)$, where the n-vector $C(q,\dot{q},t)$ is a known function of its arguments. It specifies the *nature* of the constraint forces. (S3)

We remind the reader that item (S2) above is equivalent to Eq. (4), and item (S3) is our generalized D'Alembert's principle as stated in Section 2.

Next we shall provide the explicit equation of motion that emerges from the above mathematical statement, and furthermore show that the accelerations provided by it are unique. From here on, for clarity, we shall suppress the arguments of the various quantities.

4 Equation of Motion for Constrained System With Nonideal Constraints

Result 1. An equation of motion of the constrained mechanical system that satisfies conditions (S1)–(S3) given in the previous section is explicitly given by

$$M\ddot{q}=Q+Q^c=Q+M^{1/2}B^+(b-Aa)+M^{1/2}\{I-B^+B\}M^{-1/2}C. \tag{7}$$

Proof. We shall prove that the constraint force n-vector, Q^c, given by Eq. (7) satisfies (S1)–(S3).

(S1) The form of Eq. (7) shows that (S1) is satisfied.

(S2) Using $\ddot{q}$ from Eq. (7) in Eq. (4) gives

$$\begin{aligned} A\ddot{q}&=Aa+BB^+(b-Aa)+B(I-B^+B)M^{-1/2}C\\ &=Aa+BB^+b-BB^+BM^{1/2}a\\ &=Aa+BB^+b-BM^{1/2}a=BB^+b, \end{aligned} \tag{8}$$

where we have used the relations $a=M^{-1}Q$, $BB^+B=B$, and $B=AM^{-1/2}$. Equation (4) can be expressed as $B(M^{1/2}\ddot{q})=b$, and being consistent, implies ([8]) that $BB^+b=b$. Using this in the right-hand side in (11) proves that the acceleration $\ddot{q}$ satisfies Eq. (4). Hence (S2) is satisfied.

(S3) As seen from (7), the constraint force, Q^c, is given by

$$Q^c=Q_i^c+Q_{ni}^c=M^{1/2}B^+(b-Aa)+M^{1/2}\{I-B^+B\}M^{-1/2}C. \tag{9}$$

Since $B=AM^{-1/2}$, after setting $v=M^{-1/2}\mu$, (S3) is equivalent to proving that

$$\{\mu|B\mu=0,\mu\neq 0\}\}\Rightarrow \mu^TM^{-1/2}Q^c=\mu^TM^{-1/2}C. \tag{10}$$

But $B\mu=0$ implies $\mu^+B^+=0$, and this ([8]) implies $\mu^TB^+=0$. By Eq. (9) we then have $\mu^TM^{-1/2}Q^c=\mu^TB^+(b-Aa)+\mu^T\{I-B^+B\}M^{-1/2}C=\mu^TM^{-1/2}C$, which is the required result (S3).□

Result 2. The equation of motion for the constrained system given by (7) is unique.

Proof. Assume there exists another set of solution vectors $\ddot{q}+\ddot{e}$ and Q^c+R such that (S1)–(S3) are also satisfied. We must then have $M(\ddot{q}+\ddot{e})=Q+Q^c+R$, and by (5), $M\ddot{e}=R$. Similarly, $A(\ddot{q}+\ddot{e})=b$, and by Eq. (4), $A\ddot{e}=0$. So the n-vector $\ddot{e}$ qualifies as a virtual displacement. Also, for all virtual displacements v, we must have $v^T(Q^c+R)=v^TC$, so that $v^TR=0$. Thus $\ddot{e}^TR=\ddot{e}^TM\ddot{e}=0$, and hence $\ddot{e}=0$ because M is positive definite. Since $R=M\ddot{e}=0$, uniqueness follows. □

Thus Eq. (7) gives the *unique* equation of motion describing the acceleration of a constrained mechanical system where the constraints are nonideal and the constraint forces do an amount of work (under the virtual displacement, v) given by $v^TC(q,\dot{q},t)$,

with the *n*-vector C being known. We explain the salient features of Eqs. (7) and (9) in the following series of remarks.

Remark 1. The equation of motion, (7), for the constrained system does not contain any ''multipliers'' that need to be solved for, as found in Lagrange's equations that describe constrained motion with ideal constraints. □

Remark 2. No elimination of coordinates (or velocities) is done; therefore, no set of coordinates (or velocities) is singled out for special treatment, as in the Gibbs-Appell approach that is applicable for ideal constraints. The equation of motion is stated in the *same* coordinates as those describing the *unconstrained* system. This makes it simple to directly assess the influence that the presence of the constraints have on the accelerations of the unconstrained system. The next remarks deal with this. □

Remark 3. The total constraint force *n*-vector, Q^c, is given by $Q^c = Q_i^c + Q_{ni}^c$, and it is seen to be made up of *two additive* contributions. The first member on the right-hand side of Eq. (9) given by

$$Q_i^c = M^{1/2} B^+ (b - Aa) \tag{11}$$

is the constraint force *that would have been engendered were all the constraints ideal*, and $C \equiv 0$. This contribution is ever present, no matter whether the constraints are ideal or not.

The second member on the right-hand side of Eq. (9) given by

$$Q_{ni}^c = M^{1/2}\{I - B^+ B\} M^{-1/2} C \tag{12}$$

gives the *additional* contribution to the constraint force due to the presence of nonideal constraints where the constraint forces *do* work under virtual displacements. This breakdown of the total constraint force *n*-vector explicitly shows the way in which knowledge of the virtual work done by nonideal constraints enters the equation of motion of the constrained system. □

Remark 4. The contribution, Q_i^c, to total force of constraint, Q^c, does no work under virtual displacements. For, as in the proof of Result 1, $v^T Q_i^c = v^T M^{1/2} B^+ (b - Aa) = \mu^T B^+ (b - Aa) = 0$, for all μ such that $B\mu = 0$. Hence, at each instant of time $v^T Q^c = v^T Q_{ni}^c = v^T C$. □

Remark 5. The force $C(q, \dot{q}, t)$ provides a mathematical specification of the nonideal nature of the constraints by informing us of the work done by the constraint force *n*-vector, Q^c, under virtual displacements, v. Its specification depends on the physics of any given particular situation. It engenders a contribution, Q_{ni}^c, to the total constraint force, Q^c, but in general, this contribution is such that, $Q_{ni}^c \neq C$. As seen from Eq. (12), only at those instants of time when $M^{-1/2} C$ lies in the null space of the matrix B, does $Q_{ni}^c = C$.

Furthermore, at those instants of time when $M^{-1/2} C$ is such that it lies in the range space of B^T, then $Q_{ni}^c = 0$. For then, $M^{-1/2} C$ can be expressed as $B^T w$ for some suitable vector w, and by Eq. (12) we have, $M^{-1/2} Q_{ni}^c = (I - B^+ B) B^T w = [B^T - (B^+ B)^T B^T] w = [B^T - B^T (B^T)^+ B^T] w = 0$. Here, in the second and third equalities we use the properties of the Moore-Penrose inverse ([8]). □

Remark 6. When the constraints are ideal, $C \equiv 0$, and the equation of motion given by Eq. (7) reverts to one that is well known ([8]), and has been shown to be equivalent to the usual Lagrange equations with multipliers, and to the Gibbs-Appell equations, each of which is valid only for ideal constraints. □

5 Example

We illustrate the power of our result by considering a particle of unit mass moving in an inertial Cartesian frame subjected to a set of impressed forces $f_x(x,y,z,t)$, $f_y(x,y,z,t)$, $f_z(x,y,z,t)$ acting in the *x*, *y*, and *z*-directions, respectively. The particle is subjected to the nonholonomic, constraint $\dot{y} = z^2 \dot{x}$. *The presence of this nonideal constraint creates a force of constraint*. For the specific system at hand, we assume that this force of constraint does work under virtual displacements given by $v^T Q^c = -v^T (a_0 u^T u)(u/|u|)$, where u is the velocity of the particle and $|u| = +\sqrt{u^T u}$. Such a specification of the *nature* of the constraint force is left to the discretion of the mechanician who is modeling the system, and it would depend on the physics of any particular situation (see Section 2.2). What is the equation of motion of this nonholonomically constrained system in which the constraints create nonideal forces of constraint?

Using Eq. (7) we can write down an explicit equation for the motion of the particle as follows.

Differentiating the constraint equation $\dot{y} = z^2 \dot{x}$, we get

$$A = [-z^2 \;\; 1 \;\; 0], \tag{13}$$

with

$$b = 2\dot{x}\dot{z}z. \tag{14}$$

We note that it is the *existence* of the constraint $\dot{y} = z^2 \dot{x}$ that *creates* the force of constraint. This force of constraint is nonideal. It does work under virtual displacements; its magnitude is proportional to the square of the speed of the particle, and it opposes the particle's motion. It is *not* an ''impressed force'' on the particle. It would disappear in the absence of the constraint.

Since $M = I_3$, $B = A$. By Eq. (11) we then obtain

$$Q_i^c = [-z^2 \;\; 1 \;\; 0]^T \frac{(2z\dot{x}\dot{z} + z^2 f_x - f_y)}{(1 + z^4)}, \tag{15}$$

and, by Eq. (12),

$$Q_{ni}^c = -a_0 \begin{bmatrix} \dot{x} + z^2 \dot{y} \\ z^2 \dot{x} + z^4 \dot{y} \\ \dot{z}(1 + z^4) \end{bmatrix} \frac{(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)^{1/2}}{(1 + z^4)}. \tag{16}$$

The equation of motion of the nonholonomically constrained system with nonideal constraints then becomes

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = Q + Q_i^c + Q_{ni}^c = \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} + \frac{(2z\dot{x}\dot{z} + z^2 f_x - f_y)}{(1 + z^4)} \begin{bmatrix} -z^2 \\ 1 \\ 0 \end{bmatrix} - a_0 \begin{bmatrix} \dot{x} + z^2 \dot{y} \\ z^2 \dot{x} + z^4 \dot{y} \\ \dot{z}(1 + z^4) \end{bmatrix} \frac{(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)^{1/2}}{(1 + z^4)}. \tag{17}$$

The last member on the right-hand side of Eq. (17) exposes explicitly the contribution that the nonideal character of this nonholonomic constraint provides to the total constraint force, Q^c. The second member on the right informs us of the constraint force the particle *would* be subjected to, were the nonholonomic constraint $\dot{y} = z^2 \dot{x}$ ideal. As stated in Remark 5, in this example $Q_{ni}^c \neq C$.

Note that when $a_0 = 0$, the third member on the right of Eq. (17) disappears, and we get the correct equation of motion that is valid for ideal constraints. Then, our equation becomes equivalent to Lagrange's equation with multipliers and the Gibbs-Appell equation, both of which are valid only for ideal constraints.

In Ref. ([10]) we handle the sliding friction problem of a bead running down a wire. As expected, Eq. (7) indeed yields the proper equations of motion, which in this case are easy to verify using Newtonian mechanics.

Holonomically constrained systems where the constraint forces are nonideal, as in sliding friction, may at times be handled by the Newtonian approach. However, to the best of our knowledge there is no way to date to obtain the equations of motion for nonholonomically constrained systems where the constraint forces are nonideal. Thus, seemingly simple problems like the one considered in this section have so far been beyond the compass of the Lagrangian formulation (see Refs. [5] and [9] for a more extensive discussion).

6 The Geometry of Constrained Motion

The geometrical simplicity of the equation of motion (7) developed herein can perhaps be best captured by using the ''scaled'' accelerations $\ddot{q}_s = M^{1/2}\ddot{q}$, $a_s = M^{1/2}a = M^{-1/2}Q$, $\ddot{q}_s^c = M^{-1/2}Q^c$ and $c_s = M^{1/2}(M^{-1}C) = M^{-1/2}C$. The equation of motion (5) of the constrained system can then be written in terms of these scaled accelerations as

$$\ddot{q}_s(t) = a_s(t) + \ddot{q}_s^c(t), \tag{18}$$

and the problem of finding the equation of motion of the constrained system then reduces, as pointed out by Gauss [2], to finding the *deviation* $\Delta\ddot{q}_s \equiv \ddot{q}_s^c(t) = \ddot{q}_s(t) - a_s(t)$ of the scaled acceleration of the constrained system, $\ddot{q}_s(t)$, from its known, unconstrained, scaled acceleration, $a_s(t)$. Equation (7), then takes on the simple form

$$\ddot{q}_s = (I - B^+B)(a_s + c_s) + B^+b, \tag{19}$$

from which we can explicitly obtain the deviation, $\Delta\ddot{q}_s$, as

$$\Delta\ddot{q}_s = B^+(b - Ba_s) + (I - B^+B)c_s. \tag{20}$$

Let us denote $N = (I - B^+B)$, and $T = B^+B$. To understand the first member on the right-hand side of Eq. (20), we note that the extent to which the acceleration a of the unconstrained system does *not* satisfy the constraint Eq. (4) is given by

$$e = b - Aa = b - Ba_s. \tag{21}$$

Equations (19) and (20) can now be rewritten as

$$\ddot{q}_s = N(a_s + c_s) + B^+b \tag{22}$$

and

$$\Delta\ddot{q}_s = B^+b - Ta_s + Nc_s = B^+e + Nc_s. \tag{23}$$

Noting the definition of $\Delta\ddot{q}_s$, Eq. (23) can be expressed alternatively as

$$\ddot{q} - a = (M^{-1/2}B^+)e + (M^{-1/2}NM^{-1/2})C. \tag{24}$$

This form of our result leads to the following new fundamental principle of Lagrangian mechanics:

> *The motion of a discrete mechanical system subjected to constraints that are nonideal evolves, at each instant of time, in such a way that the deviation of its accelerations from those it would have at that instant if there were no constraints on it, is made up of two components. The first component is proportional to the extent to which the accelerations corresponding to the unconstrained motion, at that instant, do not satisfy the constraints; the matrix of proportionality is* $M^{-1/2}B^+$, *and the measure of the dissatisfaction of the constraints is provided by the vector e. The second component is proportional to the vector C that specifies the work done by the constraint forces under virtual displacements, at that instant, and the matrix of proportionality is* $(M^{-1/2}NM^{-1/2})$.

Now the operator N, being symmetric and idempotent, is an orthogonal projection operator on the null space of B, and the vector B^+b belongs to the range space of B^T. Furthermore, the two right-hand members of Eq. (22) constitute two n-vectors that are orthogonal to each other, because

$$N^TB^+ = (I - B^+B)^TB^+ = (I - B^+B)B^+ = B^+ - B^+BB^+ = 0, \tag{23}$$

since $B^+BB^+ = B^+$. Equation (22) thus informs us that the scaled acceleration of the constrained system is simply the sum of two *orthogonal* vectors, one belonging to the null space of B—denoted $\mathcal{N}(B)$, and the other belonging to the range space of B^T—denoted $\mathcal{R}(B^T)$. Figure 1 depicts relations (22) and (23) pictorially, and reveals the geometrical elegance with which Nature appears to operate. It generalizes the results obtained in Ref. ([12]) to include systems in which nonideal forces of constraint exist.

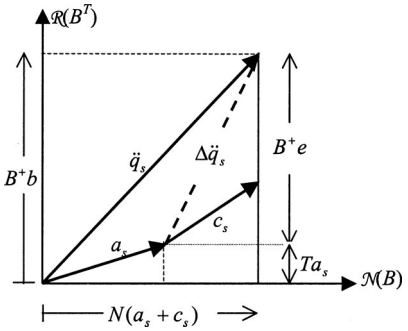


Fig. 1 The geometry of constrained motion is depicted using projections on $\mathcal{N}(B)$ and $\mathcal{R}(B^T)$. The projection of $\ddot{q}_s$ on $\mathcal{N}(B)$ is the same as that of $(a_s + c_s)$ because $N\ddot{q}_s = N(a_s + c_s)$. The vector B^+b is orthogonal to this projection.

It should come as no surprise that the vectors a_s and c_s enter Eq. (19) in the same way. Though their genesis is vastly different, they come, after all, from forces that act on the system. Notice, however, that the sum $(a_s + c_s)$ does not enter directly. The matrix $N = (I - B^+B)$ is a projection on the null space of B, and hence it is the sum's projection on this space that enters the equation of motion.

Conclusions

We summarize the contribution in this paper as follows.

1 To date, Largangian mechanics has been built upon the Principle of D'Alembert. This principle restricts Lagrangian mechanics to situations where the work done by the forces of constraint under virtual displacements is zero. In this paper we relax this restriction and thereby release Lagrangian mechanics from this confinement.

2 We have generalized D'Alembert's principle to include situations in which the constraints are not ideal, and the forces of constraint may do positive, negative, or zero work under virtual displacements. The generalized principle reduces to the usual D'Alembert's principle when the constraints are ideal.

3 The generalized D'Alembert's principle highlights the fact that the description of the motion of a constrained mechanical system requires more than just a statement of the equations of constraint, i.e., Eqs. (2) and/or (3). It *always* also requires a specification of the *nature* of the forces of constraint that the constraints engender. This is done in terms of the work done by the forces of constraint under virtual displacements, through a prescription of the n-vector $C(q, \dot{q}, t)$. D'Alembert's principle is thus seen as *one particular way* of specifying the nature of the forces of constraint, for it prescribes the vector $C(q, \dot{q}, t)$ to be identically zero. In general, one has to rely on the discretion of the mechanician to specify the vector $C(q, \dot{q}, t)$ upon examination of the specific system whose motion needs to be modeled. When D'Alembert's principle is invoked while dealing with a given constrained mechanical system—and this is most often the case in analytical dynamics, to date—the burden of this specification ''seems'' lifted from the shoulders of the mechanician, for the principle simply sets $C(q, \dot{q}, t)$ to the zero vector. However, the conscientious mechanism needs to examine if, and how well, the forces of constraint (in the given physical system being modeled) exhibit the behavior subsumed by this principle.

4 The framework of Lagragian mechanics is used to show that this generalized D'Alembert's principle provides just the right extent of information to yield the accelerations of the constrained system *uniquely*, as demanded by practical observation. In the situation that the constraints are ideal, these accelerations agree

with those determined using formalisms developed by Lagrange, Gibbs, and Appell, each of these being applicable only to the case of ideal constraints.

5 We have presented here the general, explicit, equations of motion for mechanical systems with nonideal, equality, constraints. They lead to a new and fundamental understanding of constrained motion. To the best of our knowledge, these equations are arguably the simplest and most comprehensive so far discovered. They will aid in understanding the dynamics of mechanical systems in various fields such as biomechanics, robotics, and multibody dynamics, where such nonideal constraints abound.

6 Our equations show that the constraint force n-vector is made up of two additive contributions: $Q^c = Q_i^c + Q_{ni}^c$. Explicit expressions for each of these contributions are given in this paper. The contribution Q_i^c *always* exists whether or not the constraints are ideal, and it is dictated by the kinematic nature of the constraints. The contribution Q_{ni}^c arises from a *specification* by the mechanician of the nonideal nature of the constraints that may be involved in any particular situation; it prevails when the constraint forces do work under virtual displacements.

7 We have provided an insight into the geometry of constrained motion revealing the simplicity and elegance with which Nature seems to operate.

References

[1] Lagrange, J. L., 1787, *Mecanique Analytique*, Mme Ve Courcier, Paris.
[2] Gauss, C. F., 1829, ''Uber Ein Neues Allgemeines Grundgesetz der Mechanik,'' J. Reine Ang. Math.,, **4**, pp. 232–235.
[3] Gibbs, J. W., 1879, ''On the Fundamental Formulae of Dynamics,'' Am. J. Math., **2**, pp. 49–64.
[4] Appell, P., 1899, ''Sur une Forme Generale des Equations de la Dynamique,'' C. R. Acad. Sci., Paris,, **129**, pp. 459–460.
[5] Pars, L. A., 1979, *A treatise on Analytical Dynamics*, Oxbow Press, Woodridge, CT, p. 202.
[6] Dirac, P. A. M., 1964, *Lectures in Quantum Mechanics*, Yeshiva University, New York.
[7] Udwadia, F. E., and Kalaba, R. E., 1992, ''A New Perspective on Constrained Motion,'' Proc. R. Soc. London, Ser. A, **439**, pp. 407–410.
[8] Udwadia, F. E., and Kalaba, R. E., 1996, *Analytical Dynamics: A New Approach*, Cambridge University Press, Cambridge, UK.
[9] Goldstein, H., 1981, *Classical Mechanics*, Addison-Wesley, Reading, MA.
[10] Udwadia, F. E., and Kalaba, R. E., 2000, ''Non-ideal Constraints and Lagrangian Dynamics,'' J. Aerosp. Eng.,, **13**, pp. 17–22.
[11] Penrose, R., 1955, ''A Generalized Inverse of Matrices,'' Proc. Cambridge Philos. Soc., **51**, pp. 406–413.
[12] Udwadia, F. E., and Kalaba, R. E., 1995, ''The Geometry of Constrained Motion,'' Z. Angew. Math. Mech., **75**, pp. 637–640.

Chyanbin Hwu

J. Y. Wu

Institute of Aeronautics and Astronautics,
National Cheng Kung University,
Tainan, Taiwan, R.O.C.

C. W. Fan

Center for Aviation and Space Technology,
Industrial Technology Research Institute,
Hsinchu, Taiwan, R.O.C.

M. C. Hsieh

Institute of Aeronautics and Astronautics,
National Cheng Kung University,
Tainan, Taiwan, R.O.C.

Stroh Finite Element for Two-Dimensional Linear Anisotropic Elastic Solids

A general solution satisfying the strain-displacement relation, the stress-strain laws and the equilibrium conditions has been obtained in Stroh formalism for the generalized two-dimensional anisotropic elasticity. The general solution contains three arbitrary complex functions which are the basis of the whole field stresses and deformations. By selecting these arbitrary functions to be linear or quadratic, and following the direct finite element formulation, a new finite element satisfying both the compatibility and equilibrium within each element is developed in this paper. A computer windows program is then coded by using the FORTRAN and Visual Basic languages. Two numerical examples are shown to illustrate the performance of this newly developed finite element. One is the uniform stress field problem, the other is the stress concentration problem.
[DOI: 10.1115/1.1364497]

1 Introduction

Due to the anisotropy nature of composite materials, the mechanical behavior of composite structures is usually studied by using anisotropic elasticity. There are two main approaches dealing with the two-dimensional linear anisotropic elasticity. One is Lekhnitskii formalism ([1]) which begins with the stresses, the other is Stroh formalism ([2,3]) which starts with the displacements. Both of these two formalisms are formulated by complex variable functions. Although they are well known in the mechanics community, they are not very popular in the engineering society. On the other hand, the finite element method is an important and popular tool for mechanical analyses. The numerous approximate procedures discussed in the literature generally fall into three categories: the direct method, the method of weighted residuals, and the variational method ([4]). Among them the displacement-based variational formulation is the most popular one. To the authors' knowledge, none of the results concluded by the complex variable formulation has ever been employed in the displacement-based finite element formulation. In this paper, we try to build a bridge connecting these two main formulations.

It is known that the equilibrium is usually not satisfied within elements and between elements for a displacement-based finite element. However, for a displacement-based complex variable formulation—Stroh formalism, a general solution satisfying the strain-displacement relation, the stress-strain laws and the equilibrium equations has been obtained explicitly ([3]). Hence, it is expected to get some merits by imbedding the general solution of Stroh formalism into the finite element formulation, which will be called ''Stroh finite element'' in this paper.

In Stroh formalism, the general solution contains three arbitrary complex functions which are the basis of the whole field stresses and deformations. Like the shape functions used in the finite element method, the arbitrary functions can be chosen to be polynomials. Although they are complex variable functions, through the use of identities developed in the literature, for example ([5]), the entire formulation can be transformed into a real form expression like the usual finite element. Based upon this concept, the explicit expressions of the stiffness matrices for two basic elements: linear and quadratic elements, are derived in this paper. A finite element computer windows program is then coded by using the FORTRAN and Visual Basic languages. Due to the use of the Stroh formalism, the compatibility and the equilibrium conditions within each element are all satisfied. The accuracy and versatility of the elements are then shown through several numerical examples.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 4, 2000; final revision, November 30, 2000. Associate Editor: M. Ortiz. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

2 Two-Dimensional Linear Anisotropic Elasticity

In a fixed rectangular coordinate system x_i, $i=1,2,3$, let u_i, σ_{ij}, ε_{ij} be, respectively, the displacement, stress, and strain. The strain-displacement equations, the stress-strain laws, and the equations of equilibrium are

$$\varepsilon_{ij}=\frac{1}{2}(u_{i,j}+u_{j,i}), \tag{2.1a}$$

$$\sigma_{ij}=C_{ijks}\varepsilon_{ks}, \tag{2.1b}$$

$$\sigma_{ij,j}=C_{ijks}u_{k,sj}=0, \tag{2.1c}$$

where repeated indices imply summation, a comma stands for differentiation and C_{ijks} are the elastic constants which are assumed to be fully symmetric and positive definite. Assuming that u_i, $i=1,2,3$, depend on x_1 and x_2 only, the general solution to (2.1) can be written in matrix notation as ([3])

$$\begin{aligned}\mathbf{u}&=\mathbf{A}\mathbf{f}(z)+\overline{\mathbf{A}\mathbf{f}(z)}=2\,\mathrm{Re}\{\mathbf{A}\mathbf{f}(z)\},\\ \boldsymbol{\phi}&=\mathbf{B}\mathbf{f}(z)+\overline{\mathbf{B}\mathbf{f}(z)}=2\,\mathrm{Re}\{\mathbf{B}\mathbf{f}(z)\},\end{aligned} \tag{2.2a}$$

where

$$\begin{aligned}&\mathbf{A}=[\mathbf{a}_1\ \ \mathbf{a}_2\ \ \mathbf{a}_3],\quad \mathbf{B}=[\mathbf{b}_1\ \ \mathbf{b}_2\ \ \mathbf{b}_3],\\ &\mathbf{f}(z)=[f_1(z_1),\ f_2(z_2),\ f_3(z_3)]^T,\\ &z_\alpha=x_1+p_\alpha x_2.\end{aligned} \tag{2.2b}$$

An overbar denotes the complex conjugate and Re stands for the real parts. $f_\alpha(z_\alpha)$, $\alpha=1,2,3$, are holomorphic complex functions which will be determined by the satisfaction of the boundary conditions. $\boldsymbol{\phi}$ is the stress function vector which is related to the stresses by

$$\sigma_{i1}=-\phi_{i,2},\quad \sigma_{i2}=\phi_{i,1}. \tag{2.3}$$

p_α, $\alpha=1,2,3$, are the material eigenvalues whose imaginary parts are positive. $(\mathbf{a}_\alpha, \mathbf{b}_\alpha)$ are their associated eigenvectors. These values can be determined by the following eigenrelation:

$$\mathbf{N}\boldsymbol{\xi}=p\boldsymbol{\xi}, \tag{2.4a}$$

where

$$\mathbf{N}=\begin{bmatrix}\mathbf{N}_1 & \mathbf{N}_2\\ \mathbf{N}_3 & \mathbf{N}_1^T\end{bmatrix}, \quad \boldsymbol{\xi}=\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix}, \tag{2.4b}$$

and

$$\mathbf{N}_1=-\mathbf{T}^{-1}\mathbf{R}^T,\ \ \mathbf{N}_2=\mathbf{T}^{-1}=\mathbf{N}_2^T,\ \ \mathbf{N}_3=\mathbf{R}\mathbf{T}^{-1}\mathbf{R}^T-\mathbf{Q}=\mathbf{N}_3^T, \tag{2.4c}$$

$$Q_{ik}=C_{i1k1},\quad R_{ik}=C_{i1k2},\quad T_{ik}=C_{i2k2}. \tag{2.4d}$$

In (2.4*c*) the superscript T stands for the transpose, and the superscript -1 means inverse.

It should be noted that the solutions given in (2.2) are derived under the assumption that p_α, $\alpha=1,2,3$, are distinct. When the material eigenvalues are repeated, it is possible that we cannot find three independent eigenvectors to construct the eigenvector matrices **A** and **B**. To overcome this degenerate case, a small perturbation of the material constants is usually introduced ([6]) in numerical calculation, or a generalized eigenvector is introduced to modify the formulation given in (2.2) ([7]).

If **t** is the surface traction at a point on a curve boundary, then

$$\mathbf{t}=\partial\boldsymbol{\phi}/\partial s, \tag{2.5}$$

where s is the arc length measured along the curved boundary in the direction such that, when one faces the direction of increasing s, the material is located on the right-hand side. We see that (2.3) are special cases of (2.5) when the boundary is a plane parallel to the x_2-axis or the x_1-axis.

One of the special features of the Stroh formalism is the identities which transform the complex functions into real form expressions. With these identities, the mathematical manipulation becomes easier and the real form solution becomes possible. An identity which plays an important role in our later formulation is now listed below ([8]).

$$\begin{bmatrix}\mathbf{A}\langle\langle z_\alpha^k\rangle\rangle\mathbf{B}^T & \mathbf{A}\langle\langle z_\alpha^k\rangle\rangle\mathbf{A}^T\\ \mathbf{B}\langle\langle z_\alpha^k\rangle\rangle\mathbf{B}^T & \mathbf{B}\langle\langle z_\alpha^k\rangle\rangle\mathbf{A}^T\end{bmatrix}=\frac{1}{2}\hat{\mathbf{N}}^k(\mathbf{I}-i\mathbf{N}_o), \tag{2.6a}$$

where

$$\mathbf{N}_o=\frac{1}{\pi}\int_0^\pi \mathbf{N}(\omega)d\omega=\begin{bmatrix}\mathbf{S} & \mathbf{H}\\ -\mathbf{L} & \mathbf{S}^T\end{bmatrix}, \tag{2.6b}$$

and

$$\hat{\mathbf{N}}=x_1\mathbf{I}+x_2\mathbf{N}. \tag{2.6c}$$

In the above, the angular bracket $\langle\!\langle f_\alpha\rangle\!\rangle$ denotes the diagonal matrix with the diagonal components varied according to the subscript α. The real matrix function $\mathbf{N}(\omega)$ is related to **N** by $\mathbf{N}(0)=\mathbf{N}$. The general definition of $\mathbf{N}(\omega)$ is the same as that shown in (2.4) except that all the submatrices $\mathbf{N}_i(\omega)$ is calculated based upon the real matrices $\mathbf{Q}(\omega)$, $\mathbf{R}(\omega)$ and $\mathbf{T}(\omega)$ which are defined as

$$Q_{ik}(\omega)=C_{ijks}n_jn_s,\ \ R_{ik}(\omega)=C_{ijks}n_jm_s,\ \ T_{ik}(\omega)=C_{ijks}m_jm_s, \tag{2.7a}$$

where

$$\mathbf{n}(\omega)=(\cos\omega,\sin\omega,0),\ \ \mathbf{m}(\omega)=(-\sin\omega,\cos\omega,0), \tag{2.7b}$$

and ω is an arbitrary real parameter. When $\omega=0$, one can show that (2.7) is equivalent to (2.4*d*). It has been proved ([9]) that the three real matrices **S**, **H**, and **L** given in (2.6*b*) are also related to the material eigenvector matrices **A** and **B** by

$$\mathbf{H}=2i\mathbf{A}\mathbf{A}^T,\ \ \mathbf{L}=-2i\mathbf{B}\mathbf{B}^T,\ \ \mathbf{S}=i(2\mathbf{A}\mathbf{B}^T-\mathbf{I}), \tag{2.8}$$

where **I** is the unit matrix. It can be shown that **H** and **L** are symmetric and positive definite ([5]).

With $\hat{\mathbf{N}}$ defined in (2.6*c*), $\hat{\mathbf{N}}^k$ can be written explicitly as

$$\hat{\mathbf{N}}^k=(x_1\mathbf{I}+x_2\mathbf{N})^k=x_1^k\mathbf{I}+kx_1^{k-1}x_2\mathbf{N}+\ \ldots\ +C_r^kx_1^{k-r}x_2^r\mathbf{N}^r+\ \ldots\ +x_2^k\mathbf{N}^k=\begin{bmatrix}\hat{\mathbf{N}}_1^{\langle k\rangle} & \hat{\mathbf{N}}_2^{\langle k\rangle}\\ \hat{\mathbf{N}}_3^{\langle k\rangle} & \hat{\mathbf{N}}_1^{\langle k\rangle^T}\end{bmatrix} \tag{2.9}$$

where the 3×3 matrices $\hat{\mathbf{N}}_i^{\langle k\rangle}$, $i=1,2,3$ are the submatrices of the 6×6 matrix $\hat{\mathbf{N}}^k$; C_r^k is the binomial coefficient.

Before going further, we like to describe some important features of the identity (2.6) which may be helpful for the understanding of the following derivation: (1) All the matrices consisted in the left-hand side of the identity (2.6) are complex, while those in the right-hand side are real. (2) For the degenerate materials whose eigenvalues p_α are repeated, the eigenvector matrices **A** and **B** may not exist and hence it is difficult to calculate the left-hand side of the identity (2.6). However, all the real matrices in the right-hand side of the identity (2.6) are related directly to the material constants and no eigenvalue problems need to be solved. Hence, even for the degenerate materials, they can still be calculated easily without further efforts. Therefore, even our basic formulation given in (2.2) are derived based upon the assumption that the material eigenvalues are distinct, if our final results are all expressed in terms of the real matrices such as $\hat{\mathbf{N}}$ defined in (2.6*c*), they can be applied to any kind of anisotropic materials including the degenerate materials such as the isotropic materials.

If the undetermined function $\mathbf{f}(z)$ is approximated by a polynomial such as

$$\mathbf{f}(z)=\mathbf{c}_o+\langle\langle z_\alpha\rangle\rangle\mathbf{c}_1+\ \ldots\ +\langle\langle z_\alpha^n\rangle\rangle\mathbf{c}_n, \tag{2.10}$$

Eq. (2.2*a*) will become

$$\mathbf{u}=2\sum_{k=0}^{n}\ \mathrm{Re}\{\mathbf{A}\langle\langle z_\alpha^k\rangle\rangle\mathbf{c}_k\},\qquad \boldsymbol{\phi}=2\sum_{k=0}^{n}\ \mathrm{Re}\{\mathbf{B}\langle\langle z_\alpha^k\rangle\rangle\mathbf{c}_k\}. \tag{2.11}$$

Note that the coefficients $\mathbf{c}_k, k=0,1,\ldots,n$ are complex numbers. Realizing that any complex number contains two real numbers combined together by the imaginary unit $i=\sqrt{-1}$, for the convenience of derivation we may replace $\mathbf{c}_k$ by two real column vectors $\mathbf{g}_k$ and $\mathbf{h}_k$ such that

$$\mathbf{c}_k=\mathbf{A}^T\mathbf{g}_k+\mathbf{B}^T\mathbf{h}_k,\quad \mathbf{g}_k,\mathbf{h}_k:\text{real}. \tag{2.12}$$

The arbitrariness of the complex constants $\mathbf{c}_k$ will not be restricted by the replacement of two real constants $\mathbf{g}_k$ and $\mathbf{h}_k$, because the eigenvectors contained in the eigenvector matrices **A** and **B** are independent each other if their associated eigenvalues p_α are distinct.

Substituting (2.12) into (2.11) and using the identities shown in (2.6*a*), a real form solution for the displacements **u** and the stress functions $\boldsymbol{\phi}$ can be obtained as

$$\begin{Bmatrix}\mathbf{u}\\ \boldsymbol{\phi}\end{Bmatrix}=\sum_{k=0}^{n}\hat{\mathbf{N}}^k\begin{Bmatrix}\mathbf{h}_k\\ \mathbf{g}_k\end{Bmatrix}=\sum_{k=0}^{n}\begin{bmatrix}\hat{\mathbf{N}}_1^{\langle k\rangle} & \hat{\mathbf{N}}_2^{\langle k\rangle}\\ \hat{\mathbf{N}}_3^{\langle k\rangle} & \hat{\mathbf{N}}_1^{\langle k\rangle^T}\end{bmatrix}\begin{Bmatrix}\mathbf{h}_k\\ \mathbf{g}_k\end{Bmatrix}, \tag{2.13a}$$

or

$$\begin{Bmatrix}\mathbf{u}\\ \boldsymbol{\phi}\end{Bmatrix}=\begin{Bmatrix}\mathbf{h}_o\\ \mathbf{g}_o\end{Bmatrix}+\hat{\mathbf{N}}\begin{Bmatrix}\mathbf{h}_1\\ \mathbf{g}_1\end{Bmatrix}+\hat{\mathbf{N}}^2\begin{Bmatrix}\mathbf{h}_2\\ \mathbf{g}_2\end{Bmatrix}+\ \ldots\ +\hat{\mathbf{N}}^n\begin{Bmatrix}\mathbf{h}_n\\ \mathbf{g}_n\end{Bmatrix}. \tag{2.13b}$$

Now, all we have to do is determining the unknown real coefficients $\mathbf{h}_k, \mathbf{g}_k, k=0,1,2,\ldots,n$ by satisfying the boundary conditions for the given problems. The problem is now solved approximately in principle. An alternative approach which follows the

spirit of the finite element method is dividing the entire body into several small subdivisions and assuming the arbitrary function $\mathbf{f}(z)$ to be a lower-order polynomial in each subregion. This is what we like to explore in this paper. Based upon this concept, we will show two simplest elements in the next two sections. One is approximated by linear function in each subregion, the other is approximated by quadratic function in each subregion. That is,

$$\mathbf{f}(z)=\begin{cases}\mathbf{c}_o^1+\langle\langle z_\alpha\rangle\rangle\mathbf{c}_1^1, & \text{region } 1\\ \mathbf{c}_o^2+\langle\langle z_\alpha\rangle\rangle\mathbf{c}_1^2, & \text{region } 2\\ \quad\cdot\\ \quad\cdot\\ \quad\cdot\\ \quad\cdot\\ \mathbf{c}_o^n+\langle\langle z_\alpha\rangle\rangle\mathbf{c}_1^n, & \text{region } n\end{cases} \tag{2.14}$$

in linear element.

$$\mathbf{f}(z)=\begin{cases}\mathbf{c}_o^1+\langle\langle z_\alpha\rangle\rangle\mathbf{c}_1^1+\langle\langle z_\alpha^2\rangle\rangle\mathbf{c}_2^1, & \text{region } 1\\ \mathbf{c}_o^2+\langle\langle z_\alpha\rangle\rangle\mathbf{c}_1^2+\langle\langle z_\alpha^2\rangle\rangle\mathbf{c}_2^2, & \text{region } 2\\ \quad\cdot\\ \quad\cdot\\ \quad\cdot\\ \quad\cdot\\ \mathbf{c}_o^n+\langle\langle z_\alpha\rangle\rangle\mathbf{c}_1^n+\langle\langle z_\alpha^2\rangle\rangle\mathbf{c}_2^n, & \text{region } n\end{cases} \tag{2.15}$$

in quadratic element.

3 Linear Elements

By choosing the unknown function $\mathbf{f}(z)$ in each element to be a linear function as that shown in (2.14), from (2.13*b*) the displacements and stress functions for each element may be expressed as

$$\begin{Bmatrix}\mathbf{u}\\ \boldsymbol{\phi}\end{Bmatrix}=\begin{Bmatrix}\mathbf{h}_0\\ \mathbf{g}_o\end{Bmatrix}+\hat{\mathbf{N}}\begin{Bmatrix}\mathbf{h}_1\\ \mathbf{g}_1\end{Bmatrix}, \tag{3.1}$$

or

$$\mathbf{u}=\mathbf{h}_o+\hat{\mathbf{N}}_1\mathbf{h}_1+\hat{\mathbf{N}}_2\mathbf{g}_1,$$
$$\boldsymbol{\phi}=\mathbf{g}_o+\hat{\mathbf{N}}_3\mathbf{h}_1+\hat{\mathbf{N}}_1^T\mathbf{g}_1. \tag{3.2}$$

Before converting the above expression into the finite element formulation, the physical meaning of the zero and first-order terms have been studied (see Appendix). The discussion shown in the Appendix reveals that $\mathbf{h}_o$ represents rigid-body translation, and the second component of $\mathbf{h}_1$ represents rigid-body rotation. All the other components of $\mathbf{h}_1$ and $\mathbf{g}_1$ represent constant strain.

In addition to the above discussion, the continuity among elements is an important problem for the finite element formulation. In order to satisfy the continuity conditions, we try to represent the displacement and stress function in terms of the nodal displacement. To this end, the transformation matrix between $(\mathbf{h}_o,\mathbf{h}_1,\mathbf{g}_1)$ and the nodal displacements should be found. By using the triangular element with nodes 1, 2, and 3 shown in Fig. 1(*a*), and substituting the positions of nodes 1, 2, and 3 into Eq. $(3.2)_1$, we have

$$\mathbf{u}_n=\begin{Bmatrix}\mathbf{u}^{(1)}\\ \mathbf{u}^{(2)}\\ \mathbf{u}^{(3)}\end{Bmatrix}=\begin{bmatrix}\mathbf{I} & \hat{\mathbf{N}}_1^{(1)} & \hat{\mathbf{N}}_2^{(1)}\\ \mathbf{I} & \hat{\mathbf{N}}_1^{(2)} & \hat{\mathbf{N}}_2^{(2)}\\ \mathbf{I} & \hat{\mathbf{N}}_1^{(3)} & \hat{\mathbf{N}}_2^{(3)}\end{bmatrix}\begin{Bmatrix}\mathbf{h}_o\\ \mathbf{h}_1\\ \mathbf{g}_1\end{Bmatrix}, \tag{3.3a}$$

where

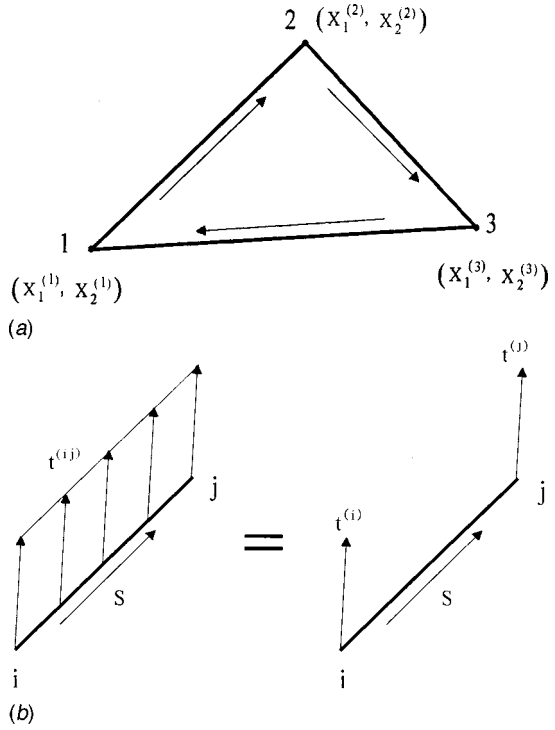


Fig. 1 (*a*) Linear triangular element, (*b*) traction distribution for the linear element

$$\hat{\mathbf{N}}_1^{(i)}=x_1^{(i)}\mathbf{I}+x_2^{(i)}\mathbf{N}_1,$$
$$\hat{\mathbf{N}}_2^{(i)}=x_2^{(i)}\mathbf{N}_2,\quad i=1,2,3. \tag{3.3b}$$

Throughout this paper, the symbols with the superscripts (i) denote the values calculated at node i. From ([10]), we know that the fundamental elasticity matrices $\mathbf{N}_1$, $\mathbf{N}_2$, and $\mathbf{N}_3$ have the units of $1,1/E$ and E, respectively. Here, E represents the Young's modulus. Therefore, the transformation matrix relating $\mathbf{u}_n$ and $(\mathbf{h}_o,\mathbf{h}_1,\mathbf{g}_1)$ is doomed to be numerically ill-conditioned. To avoid this situation, we nondimensionalize the matrix by the following way:

$$\mathbf{u}_n=\begin{Bmatrix}\mathbf{u}^{(1)}\\ \mathbf{u}^{(2)}\\ \mathbf{u}^{(3)}\end{Bmatrix}=\begin{bmatrix}\mathbf{I} & \hat{\mathbf{N}}_1^{(1)}/l_o & \hat{\mathbf{N}}_2^{(1)}E_o/l_o\\ \mathbf{I} & \hat{\mathbf{N}}_1^{(2)}/l_o & \hat{\mathbf{N}}_2^{(2)}E_o/l_o\\ \mathbf{I} & \hat{\mathbf{N}}_1^{(3)}/l_o & \hat{\mathbf{N}}_2^{(3)}E_o/l_o\end{bmatrix}\begin{Bmatrix}\mathbf{h}_o\\ l_o\mathbf{h}_1\\ \dfrac{l_o}{E_o}\mathbf{g}_1\end{Bmatrix}, \tag{3.4}$$

where l_o and E_o are, respectively, the reference length and Young's modulus. By inversion, we have

$$\begin{Bmatrix}\mathbf{h}_o\\ l_o\mathbf{h}_1\\ \dfrac{l_o}{E_o}\mathbf{g}_1\end{Bmatrix}=\tilde{\mathbf{N}}\mathbf{u}_n, \tag{3.5a}$$

where

$$\tilde{\mathbf{N}}=\begin{bmatrix}\tilde{\mathbf{N}}_{11} & \tilde{\mathbf{N}}_{12} & \tilde{\mathbf{N}}_{13}\\ \tilde{\mathbf{N}}_{21} & \tilde{\mathbf{N}}_{22} & \tilde{\mathbf{N}}_{23}\\ \tilde{\mathbf{N}}_{31} & \tilde{\mathbf{N}}_{32} & \tilde{\mathbf{N}}_{33}\end{bmatrix}=\begin{bmatrix}\mathbf{I} & \hat{\mathbf{N}}_1^{(1)}/l_o & \hat{\mathbf{N}}_2^{(1)}E_o/l_o\\ \mathbf{I} & \hat{\mathbf{N}}_1^{(2)}/l_o & \hat{\mathbf{N}}_2^{(2)}E_o/l_o\\ \mathbf{I} & \hat{\mathbf{N}}_1^{(3)}/l_o & \hat{\mathbf{N}}_2^{(3)}E_o/l_o\end{bmatrix}^{-1}. \tag{3.5b}$$

Back to (3.2), we now have

$$\mathbf{u}=[\mathbf{I}\ \ \hat{\mathbf{N}}_1/l_0\ \ \hat{\mathbf{N}}_2E_o/l_o]\tilde{\mathbf{N}}\mathbf{u}_n,$$

$$\boldsymbol{\phi}=\mathbf{g}_o+[\hat{\mathbf{N}}_3\ \ \hat{\mathbf{N}}_1^T]\begin{bmatrix}\tilde{\mathbf{N}}_{21}/l_o & \tilde{\mathbf{N}}_{22}/l_o & \tilde{\mathbf{N}}_{23}/l_o\\ \tilde{\mathbf{N}}_{31}E_o/l_o & \tilde{\mathbf{N}}_{32}E_o/l_o & \tilde{\mathbf{N}}_{33}E_o/l_o\end{bmatrix}\mathbf{u}_n. \tag{3.6}$$

With the stress function expressed by the nodal displacement, the traction along the line connecting nodes i and j (see Fig. 1(b)) can now be calculated by using (2.5) and (3.6). That is

$$\mathbf{t}^{(ji)}=\frac{d\boldsymbol{\phi}}{ds}=\begin{bmatrix}\frac{d}{ds}\tilde{\mathbf{N}}_3 & \frac{d}{ds}\tilde{\mathbf{N}}_1^T\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{N}}_{21}/l_o & \tilde{\mathbf{N}}_{22}/l_o & \tilde{\mathbf{N}}_{23}/l_o\\ \tilde{\mathbf{N}}_{31}E_o/l_o & \tilde{\mathbf{N}}_{32}E_o/l_o & \tilde{\mathbf{N}}_{33}E_o/l_o\end{bmatrix}\mathbf{u}_n, \tag{3.7a}$$

where

$$\frac{d}{ds}\hat{\mathbf{N}}_1^T=\frac{dx_1}{ds}\mathbf{I}+\frac{dx_2}{ds}\mathbf{N}_1^T,\qquad \frac{d}{ds}\hat{\mathbf{N}}_3=\frac{dx_2}{ds}\mathbf{N}_3, \tag{3.7b}$$

and

$$\frac{dx_1}{ds}=\cos\psi=\frac{x_1^{(j)}-x_1^{(i)}}{l_{ji}},\qquad \frac{dx_2}{ds}=\sin\psi=\frac{x_2^{(j)}-x_2^{(i)}}{l_{ji}}. \tag{3.7c}$$

l_{ji} is the length connecting nodes i and j. Since the traction is uniformly distributed along the straight boundary for the present linear element, the resultant forces along this line contributing to nodes i and j can be considered to be equal. That is,

$$\mathbf{t}^{(i)}=\mathbf{t}^{(j)}=\frac{l_{ji}}{2}\mathbf{t}^{(ji)}. \tag{3.8}$$

Therefore, the vector of the nodal forces $\mathbf{t}_n$ may be written as

$$\mathbf{t}_n=\begin{Bmatrix}\mathbf{t}^{(1)}\\ \mathbf{t}^{(2)}\\ \mathbf{t}^{(3)}\end{Bmatrix}=\begin{Bmatrix}\frac{1}{2}(l_{21}\mathbf{t}^{(21)}+l_{13}\mathbf{t}^{(13)})\\ \frac{1}{2}(l_{21}\mathbf{t}^{(21)}+l_{32}\mathbf{t}^{(32)})\\ \frac{1}{2}(l_{13}\mathbf{t}^{(13)}+l_{32}\mathbf{t}^{(32)})\end{Bmatrix}. \tag{3.9}$$

By employing (3.7) and (3.9), we now have

$$\mathbf{t}_n=\mathbf{K}\mathbf{u}_n, \tag{3.10a}$$

$$\mathbf{K}=\begin{bmatrix}\mathbf{K}_{11} & \mathbf{K}_{12} & \mathbf{K}_{13}\\ \mathbf{K}_{21} & \mathbf{K}_{22} & \mathbf{K}_{23}\\ \mathbf{K}_{31} & \mathbf{K}_{32} & \mathbf{K}_{33}\end{bmatrix}. \tag{3.10b}$$

The submatrices $\mathbf{K}_{ij}$ of the element stiffness $\mathbf{K}$ can be written explicitly as

$$\mathbf{K}_{ij}=\frac{1}{2l_o}(\Delta\hat{\mathbf{N}}_3^{(i)}\tilde{\mathbf{N}}_{2j}+E_o\Delta\hat{\mathbf{N}}_1^{(i)^T}\tilde{\mathbf{N}}_{3j}), \tag{3.10c}$$

where

$$\Delta\hat{\mathbf{N}}_1^{(i)^T}=(\Delta x_1)^{(i)}\mathbf{I}+(\Delta x_2)^{(i)}\mathbf{N}_1^T,\quad \Delta\hat{\mathbf{N}}_3^{(i)}=(\Delta x_2)^{(i)}\mathbf{N}_3, \tag{3.10d}$$

and

$$(\Delta x_i)^{(1)}=x_i^{(2)}-x_i^{(3)},\quad (\Delta x_i)^{(2)}=x_i^{(3)}-x_i^{(1)},\quad (\Delta x_i)^{(3)}=x_i^{(1)}-x_i^{(2)},\quad i=1,2. \tag{3.10e}$$

The internal stresses at any point within the element can now be calculated by using (2.3), (3.2), and (3.5a). The results are

$$\begin{Bmatrix}\boldsymbol{\sigma}_1\\ \boldsymbol{\sigma}_2\end{Bmatrix}=\begin{bmatrix}\frac{-1}{l_o}(\mathbf{N}_3\tilde{\mathbf{N}}_{21}+E_o\mathbf{N}_1^T\tilde{\mathbf{N}}_{31}) & \frac{-1}{l_o}(\mathbf{N}_3\tilde{\mathbf{N}}_{22}+E_o\mathbf{N}_1^T\tilde{\mathbf{N}}_{32}) & \frac{-1}{l_o}(\mathbf{N}_3\tilde{\mathbf{N}}_{23}+E_o\mathbf{N}_1^T\tilde{\mathbf{N}}_{33})\\ \frac{E_o}{l_o}\tilde{\mathbf{N}}_{31} & \frac{E_o}{l_o}\tilde{\mathbf{N}}_{32} & \frac{E_o}{l_o}\tilde{\mathbf{N}}_{33}\end{bmatrix}\mathbf{u}_n, \tag{3.11a}$$

where

$$\boldsymbol{\sigma}_1=\{\sigma_{11},\sigma_{12},\sigma_{13}\}^T,\quad \boldsymbol{\sigma}_2=\{\sigma_{21},\sigma_{22},\sigma_{23}\}^T. \tag{3.11b}$$

It should be noted that σ_{21} and σ_{12} obtained above look like to have different expressions. By knowing the structures of $\mathbf{N}_1$ and $\mathbf{N}_3$ shown in ([5]), it can easily be proved that they are really identical.

From the above derivation, we see that the displacements $\mathbf{u}$ within the element, the traction $\mathbf{t}$ along the boundary, and the internal stresses $(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2)$ can all be expressed in terms of the nodal displacement $\mathbf{u}_n$, as shown in $(3.6)_1$, (3.7), and (3.11). The element stiffness matrix $\mathbf{K}$ which is the core matrix in the finite element formulation is obtained in (3.10). It can also be seen that all these formulas are written explicitly and do not contain any integrals. This feature may save us a lot of computational effort. Moreover, the present element is valid for any kind of anisotropic materials and considers not only in-plane deformations and stresses $(u_1,u_2,\sigma_{11},\sigma_{12},\sigma_{22})$ but also the antiplane deformations and stresses $(u_3,\sigma_{13},\sigma_{23})$.

4 Quadratic Elements

We now consider the next higher order approximation, i.e., the second-order approximation. By (2.13b), we have

$$\begin{Bmatrix}\mathbf{u}\\ \boldsymbol{\phi}\end{Bmatrix}=\begin{Bmatrix}\mathbf{h}_o\\ \mathbf{g}_o\end{Bmatrix}+\hat{\mathbf{N}}\begin{Bmatrix}\mathbf{h}_1\\ \mathbf{g}_1\end{Bmatrix}+\hat{\mathbf{N}}^2\begin{Bmatrix}\mathbf{h}_2\\ \mathbf{g}_2\end{Bmatrix}, \tag{4.1a}$$

or

$$\mathbf{u}=\mathbf{h}_o+\hat{\mathbf{N}}_1\mathbf{h}_1+\hat{\mathbf{N}}_2\mathbf{g}_1+\hat{\mathbf{N}}_1^{\langle 2\rangle}\mathbf{h}_2+\hat{\mathbf{N}}_2^{\langle 2\rangle}\mathbf{g}_2,$$

$$\boldsymbol{\phi}=\mathbf{g}_o+\hat{\mathbf{N}}_3\mathbf{h}_1+\hat{\mathbf{N}}_1^T\mathbf{g}_1+\hat{\mathbf{N}}_3^{\langle 2\rangle}\mathbf{h}_2+\hat{\mathbf{N}}_1^{\langle 2\rangle^T}\mathbf{g}_2. \tag{4.1b}$$

Similar to the linear element, the next step is trying to transform the expression in terms of the nodal displacements. Since we have five sets of unknown coefficients $(\mathbf{h}_0,\mathbf{h}_1,\mathbf{g}_1,\mathbf{h}_2,\mathbf{g}_2)$, the most appropriate element should be four-noded quadrilateral element without internal node (see Fig. 2(a)). To be consistent with the

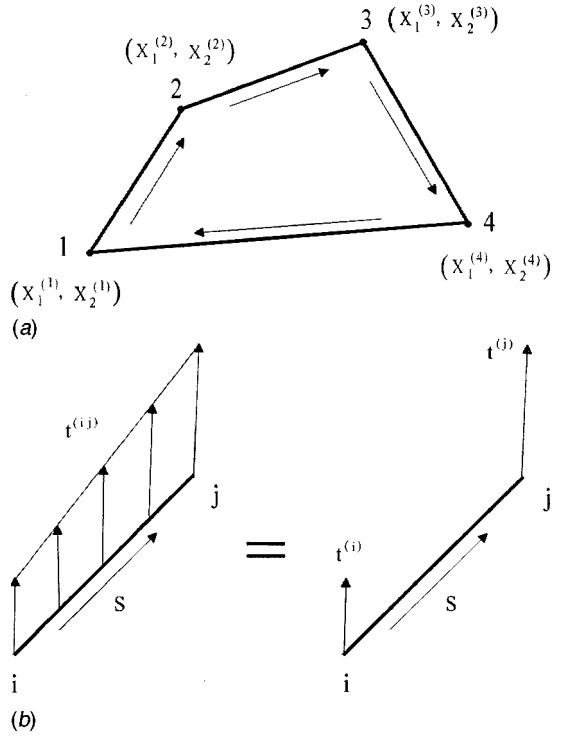


Fig. 2 (*a*) Quadratic quadrilateral element, (*b*) traction distribution for the quadratic element

node number, we need to reduce the number of unknown coefficients. The most common way in finite element formulation ([11]) is combining the last two terms by using one coefficient. Before combining the last two terms, one should note that $\hat{\mathbf{N}}_1^{\langle 2\rangle}$ and $\hat{\mathbf{N}}_2^{\langle 2\rangle}$ have different units. Knowing that ([10]) the fundamental matrices $\mathbf{N}_1$ and $\mathbf{N}_2$ have the units of 1 and $1/E$, respectively, the matrices $\hat{\mathbf{N}}_1^{\langle 2\rangle}$ and $\hat{\mathbf{N}}_2^{\langle 2\rangle}$ defined in (2.9) will have the units of l^2 and l^2/E. Therefore, to avoid adding the terms with different units, we let

$$l_o^2\mathbf{h}_2=\frac{l_o^2}{E_o}\mathbf{g}_2, \tag{4.2}$$

where l_o and E_o are, respectively, the reference length and Young's modulus. With this choice, we now have only four sets of unknown coefficients corresponding to four nodes, and the displacement field $\mathbf{u}$ can be expressed as

$$\mathbf{u}=[\mathbf{I}\ \ \hat{\mathbf{N}}_1/l_o\ \ \hat{\mathbf{N}}_2E_o/l_o\ \ (\hat{\mathbf{N}}_1^{\langle 2\rangle}+\hat{\mathbf{N}}_2^{\langle 2\rangle}E_o)/l_o^2]\begin{Bmatrix}\mathbf{h}_o\\ l_o\mathbf{h}_o\\ \dfrac{l_o}{E_o}\mathbf{g}_1\\ l_o^2\mathbf{h}_2\end{Bmatrix}. \tag{4.3}$$

In the following, all the derivations are similar to those described in Section 3. Hence, the detailed explanation will not be repeated here, only the final results are provided. Similar to (3.5b), the transformation matrix $\tilde{\mathbf{N}}$ for the quadratic element is now obtained as

$$\tilde{\mathbf{N}}=\begin{bmatrix}\tilde{\mathbf{N}}_{11}&\tilde{\mathbf{N}}_{12}&\tilde{\mathbf{N}}_{13}&\tilde{\mathbf{N}}_{14}\\ \tilde{\mathbf{N}}_{21}&\tilde{\mathbf{N}}_{22}&\tilde{\mathbf{N}}_{23}&\tilde{\mathbf{N}}_{24}\\ \tilde{\mathbf{N}}_{31}&\tilde{\mathbf{N}}_{32}&\tilde{\mathbf{N}}_{33}&\tilde{\mathbf{N}}_{34}\\ \tilde{\mathbf{N}}_{41}&\tilde{\mathbf{N}}_{42}&\tilde{\mathbf{N}}_{43}&\tilde{\mathbf{N}}_{44}\end{bmatrix}$$
$$=\begin{bmatrix}\mathbf{I}&\hat{\mathbf{N}}_1^{(1)}/l_o&\hat{\mathbf{N}}_2^{(1)}E_o/l_o&(\hat{\mathbf{N}}_1^{\langle 2\rangle^{(1)}}+\hat{\mathbf{N}}_2^{\langle 2\rangle^{(1)}}E_0)/l_0^2\\ \mathbf{I}&\hat{\mathbf{N}}_1^{(2)}/l_o&\hat{\mathbf{N}}_2^{(2)}E_o/l_o&(\hat{\mathbf{N}}_1^{\langle 2\rangle^{(2)}}+\hat{\mathbf{N}}_2^{\langle 2\rangle^{(2)}}E_0)/l_0^2\\ \mathbf{I}&\hat{\mathbf{N}}_1^{(3)}/l_o&\hat{\mathbf{N}}_2^{(3)}E_o/l_o&(\hat{\mathbf{N}}_1^{\langle 2\rangle^{(3)}}+\hat{\mathbf{N}}_2^{\langle 2\rangle^{(3)}}E_o)/l_o^2\\ \mathbf{I}&\hat{\mathbf{N}}_1^{(4)}/l_o&\hat{\mathbf{N}}_2^{(4)}E_o/l_o&(\hat{\mathbf{N}}_1^{\langle 2\rangle^{(4)}}+\hat{\mathbf{N}}_2^{\langle 2\rangle^{(4)}}E_o)/l_o^2\end{bmatrix}^{-1}. \tag{4.4}$$

The unknown coefficients $(\mathbf{h}_0,\mathbf{h}_1,\mathbf{g}_1,\mathbf{h}_2)$ may now be expressed as

$$\begin{Bmatrix}\mathbf{h}_o\\ l_o\mathbf{h}_1\\ \dfrac{l_o}{E_o}\mathbf{g}_1\\ l_o^2\mathbf{h}_2\end{Bmatrix}=\tilde{\mathbf{N}}\mathbf{u}_n. \tag{4.5}$$

Similar to (3.7), the traction $\mathbf{t}$ along the boundary connecting nodes i and j is (Fig. 2(*b*))

$$\mathbf{t}^{(ji)}=\frac{dx_2}{ds}\mathbf{N}_3\mathbf{h}_1+\left(\frac{dx_1}{ds}\mathbf{I}+\frac{dx_2}{ds}\mathbf{N}_1^T\right)\mathbf{g}_1+\left\{2\left[\left(x_1\frac{dx_2}{ds}+x_2\frac{dx_1}{ds}\right)\mathbf{N}_3+x_2\frac{dx_2}{ds}\hat{\mathbf{N}}_3^{\langle 2\rangle}\right]+\frac{2}{E_o}\left[x_1\frac{dx_1}{ds}\mathbf{I}+\left(x_1\frac{dx_2}{ds}+x_2\frac{dx_1}{ds}\right)\mathbf{N}_1^T+x_2\frac{dx_2}{ds}\hat{\mathbf{N}}_1^{\langle 2\rangle^T}\right]\right\}\mathbf{h}_2. \tag{4.6}$$

Similar to (3.9), the nodal force $\mathbf{t}_n$ may be written as

$$\mathbf{t}_n=\begin{Bmatrix}\mathbf{t}^{(1)}\\ \mathbf{t}^{(2)}\\ \mathbf{t}^{(3)}\\ \mathbf{t}^{(4)}\end{Bmatrix}=\begin{Bmatrix}\frac{1}{6}[l_{21}(2\mathbf{t}_1^{(21)}+\mathbf{t}_2^{(21)})+l_{14}(2\mathbf{t}_1^{(14)}+\mathbf{t}_4^{(14)})]\\ \frac{1}{6}[l_{32}(2\mathbf{t}_2^{(32)}+\mathbf{t}_3^{(32)})+l_{21}(2\mathbf{t}_2^{(21)}+\mathbf{t}_1^{(21)})]\\ \frac{1}{6}[l_{43}(2\mathbf{t}_3^{(43)}+\mathbf{t}_4^{(43)})+l_{32}(2\mathbf{t}_3^{(32)}+\mathbf{t}_2^{(32)})]\\ \frac{1}{6}[l_{14}(2\mathbf{t}_4^{(14)}+\mathbf{t}_1^{(14)})+l_{14}(2\mathbf{t}_4^{(43)}+\mathbf{t}_3^{(43)})]\end{Bmatrix} \tag{4.7}$$

where $\mathbf{t}_i^{(ji)}$ and $\mathbf{t}_j^{(ji)}$ denote, respectively, the values of the traction at nodes i and j along the boundary connecting nodes i and j.

Similar to (3.10), the element stiffness matrix $\mathbf{K}$ is obtained as

$$\mathbf{K}=\begin{bmatrix}\mathbf{K}_{11}&\mathbf{K}_{12}&\mathbf{K}_{13}&\mathbf{K}_{14}\\ \mathbf{K}_{21}&\mathbf{K}_{22}&\mathbf{K}_{23}&\mathbf{K}_{24}\\ \mathbf{K}_{31}&\mathbf{K}_{32}&\mathbf{K}_{33}&\mathbf{K}_{34}\\ \mathbf{K}_{41}&\mathbf{K}_{42}&\mathbf{K}_{43}&\mathbf{K}_{44}\end{bmatrix}, \tag{4.8a}$$

$$\mathbf{K}_{ij}=\Delta\hat{\mathbf{N}}_3^{(i)}\tilde{\mathbf{N}}_{2j}/l_o+\Delta\hat{\mathbf{N}}_1^{T^{(i)}}\tilde{\mathbf{N}}_{3j}E_o/l_o+(\Delta\hat{\mathbf{N}}_3^{\langle 2\rangle^{(i)}}+E_o\Delta\hat{\mathbf{N}}_1^{\langle 2\rangle^{T^{(i)}}})\tilde{\mathbf{N}}_{4j}/l_o^2 \tag{4.8b}$$

where

$$\Delta\hat{\mathbf{N}}_3^{(i)}=(\Delta x_2)^{(i)}\mathbf{N}_3,\quad \Delta\hat{\mathbf{N}}_1^{T^{(i)}}(\Delta x_1)^{(i)}\mathbf{I}+(\Delta x_2)^{(i)}\mathbf{N}_1^T,$$

$$\begin{aligned}\Delta\hat{\mathbf{N}}_3^{\langle 2\rangle^{(i)}}=&[\alpha_{i,i+1}(x_2^{(i+1)}-x_2^{(i)})+\beta_{i,i+1}(x_1^{(i+1)}-x_1^{(i)})\\&+\alpha_{i,i-1}(x_2^{(i)}-x_2^{(i-1)})+\beta_{i,i-1}(x_1^{(i)}-x_1^{(i-1)})]\mathbf{N}_3\\&+[\beta_{i,i+1}(x_2^{(i+1)}-x_2^{(i)})+\beta_{i,i-1}(x_2^{(i)}-x_2^{(i-1)})]\mathbf{N}_3^{\langle 2\rangle},\end{aligned}\tag{4.8c}$$

$$\begin{aligned}\Delta\hat{\mathbf{N}}_1^{\langle 2\rangle^{T^{(i)}}}=&[\alpha_{i,i+1}(x_1^{(i+1)}-x_1^{(i)})+\alpha_{i,i-1}(x_1^{(i)}-x_1^{(i-1)})]\mathbf{I}\\&+[\alpha_{i,i+1}(x_2^{(i+1)}-x_2^{(i)})+\beta_{i,i+1}(x_1^{(i+1)}-x_1^{(i)})\\&+\alpha_{i,i-1}(x_2^{(i)}-x_2^{(i-1)})+\beta_{i,i-1}(x_1^{(i)}-x_1^{(i-1)})]\mathbf{N}_1^T\\&+[\beta_{i,i+1}(x_2^{(i+1)}-x_2^{(i)})+\beta_{i,i-1}(x_2^{(i)}-x_2^{(i-1)})]\mathbf{N}_1^{\langle 2\rangle^T},\end{aligned}$$

and

$$(\Delta x_1)^{(i)}=\frac{1}{2}(x_1^{(i+1)}-x_1^{(i-1)}),\quad (\Delta x_2)^{(i)}=\frac{1}{2}(x_2^{(i+1)}-x_2^{(i-1)}),$$

$$\alpha_{i,j}=\frac{1}{3}(2x_1^{(i)}+x_1^{(j)}),\quad \beta_{i,j}=\frac{1}{3}(2x_2^{(i)}+x_2^{(j)}),\tag{4.8d}$$

$$\mathbf{N}_1^{\langle 2\rangle^T}=\mathbf{N}_3\mathbf{N}_2+\mathbf{N}_1^{T^2},\quad \mathbf{N}_3^{\langle 2\rangle}=\mathbf{N}_3\mathbf{N}_1+\mathbf{N}_1^T\mathbf{N}_3.$$

In the above, if the value of the superscript $(i-1)$ equals to zero, it should be replaced by 4 which is the last number of nodes of the present quadratic element.

Similar to (3.11*a*), the internal stresses $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ are

$$\begin{aligned}\boldsymbol{\sigma}_1=&-[\mathbf{N}_3\quad \mathbf{N}_1^T\quad 2[x_1\mathbf{N}_3+x_2\mathbf{N}_3^{\langle 2\rangle}]+2E_o[x_1\mathbf{N}_1^T+x_2\mathbf{N}_1^{\langle 2\rangle^T}]]\\&\times\begin{bmatrix}\tilde{\mathbf{N}}_{21}/l_o & \tilde{\mathbf{N}}_{22}/l_o & \tilde{\mathbf{N}}_{23}/l_o & \tilde{\mathbf{N}}_{24}/l_o\\ \tilde{\mathbf{N}}_{31}E_o/l_o & \tilde{\mathbf{N}}_{32}E_o/l_o & \tilde{\mathbf{N}}_{33}E_o/l_o & \tilde{\mathbf{N}}_{34}E_o/l_o\\ \tilde{\mathbf{N}}_{41}/l_o^2 & \tilde{\mathbf{N}}_{42}/l_o^2 & \tilde{\mathbf{N}}_{43}/l_o^2 & \tilde{\mathbf{N}}_{44}/l_o^2\end{bmatrix}\mathbf{u}_n,\\ \boldsymbol{\sigma}_2=&[\mathbf{0}\quad \mathbf{I}\quad 2\hat{\mathbf{N}}_3+2E_o\hat{\mathbf{N}}_1^T]\\&\times\begin{bmatrix}\tilde{\mathbf{N}}_{21}/l_o & \tilde{\mathbf{N}}_{22}/l_o & \tilde{\mathbf{N}}_{23}/l_o & \tilde{\mathbf{N}}_{24}/l_o\\ \tilde{\mathbf{N}}_{31}E_o/l_o & \tilde{\mathbf{N}}_{32}E_o/l_o & \tilde{\mathbf{N}}_{33}E_o/l_o & \tilde{\mathbf{N}}_{34}E_o/l_o\\ \tilde{\mathbf{N}}_{41}/l_o^2 & \tilde{\mathbf{N}}_{42}/l_o^2 & \tilde{\mathbf{N}}_{43}/l_o^2 & \tilde{\mathbf{N}}_{44}/l_o^2\end{bmatrix}\mathbf{u}_n.\end{aligned}\tag{4.9}$$

Although the formulas derived in the previous two sections look complicated, they are all explicit real closed-form expressions and do not contain any integrals. This really helps a lot in numerical programming shown next.

5 Numerical Examples

Based upon the usual finite element procedures and the formulations derived in Sections 3 and 4, a computer windows program was coded by using the FORTRAN and Visual Basic languages. To save the space of this paper, one may refer to ([12]) for the detail programming techniques and numerous examples. In this section, only two simple examples are shown to illustrate the performance of these newly developed elements. The first check for our written computer program is the problem with uniform stress field whose solution should be exact no matter what element subdivision is made. The second one is the stress concentration problem of which the fine meshes near holes are needed to get a convergent solution. To show the versatility, in these two problems we include the cases of in-plane and antiplane problems, and the anisotropic and isotropic materials. The main reason of including these special features is as follows: (1) the ''two-dimensional'' used in our paper title includes not only the in-plane but also the antiplane problems and the problems where in-plane and antiplane deformations couple each other; (2) the ''anisotrop-

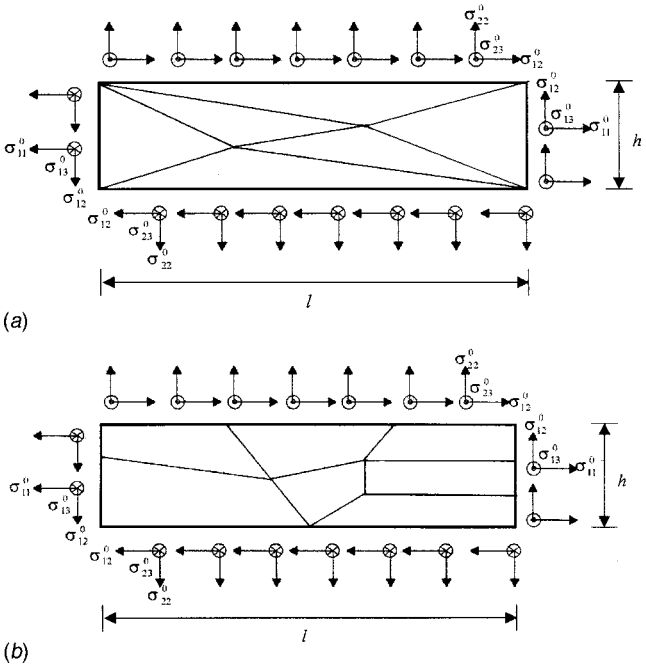


Fig. 3 (*a*) An anisotropic rectangular plate subjected to uniform tension, in-plane shear, and antiplane shear (patch 1); (*b*) an anisotropic rectangular plate subjected to uniform tension, in-plane shear, and antiplane shear (patch 2)

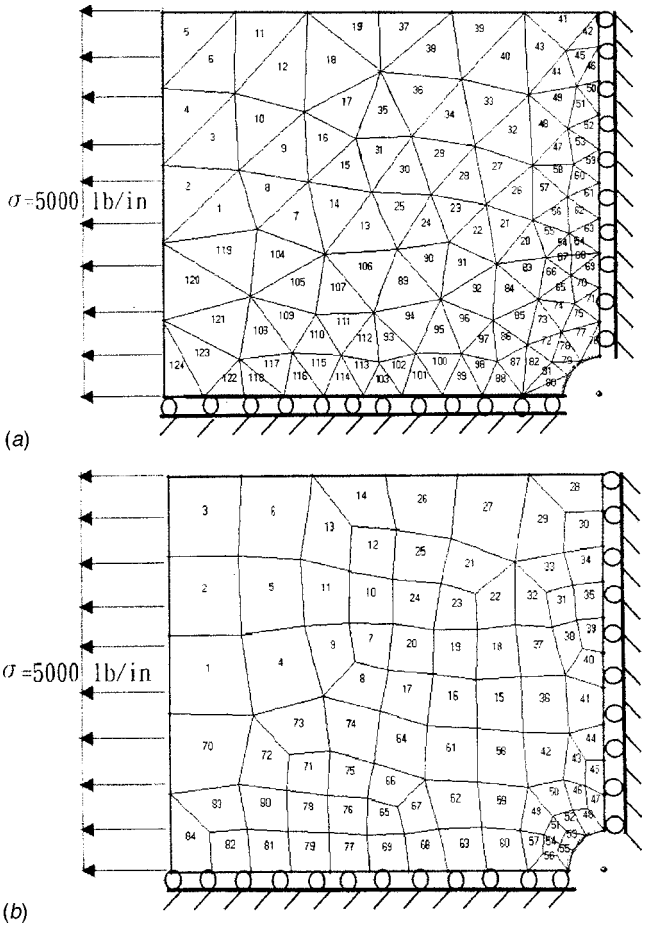


Fig. 4 (*a*) A linear finite element mesh for a quadrant of the plate with a circular hole under uniform tension, (*b*) a quadratic finite element mesh for a quadrant of the plate with a circular hole under uniform tension.

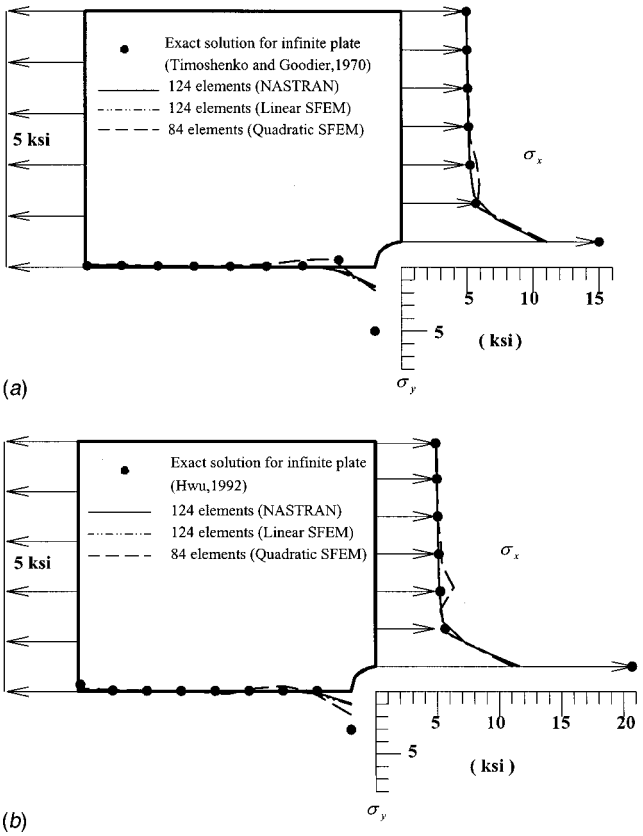


Fig. 5 (*a*) Comparison of the normal stresses on two edges for the plate made of isotropic materials, (*b*) comparison of the normal stresses on two edges for the plate made of orthotropic materials.

ic'' which need not have any material symmetry restrictions also includes the degenerate cases of which the material eigenvalues are repeated such as the isotropic materials.

Example 1: Uniform Stress Field. Consider a rectangular composite laminate subjected to a uniform load along the boundary edges (Fig. 3). The uniform load includes the in-plane tension $\sigma^0_{11}=1000$ lb./in., $\sigma^0_{22}=2000$ lb./in., in-plane shear $\sigma^0_{12}=3000$ lb./in. and antiplane shear $\sigma^0_{13}=4000$ lb./in., $\sigma^0_{23}=5000$ lb./in. The laminate has ply orientation $[0/30/-30]$ which will behave as an anisotropic material. Each lamina has a thickness of 0.04 in and is composed of graphite/epoxy whose material properties are: $E_1=26.25\times10^6$ psi, $E_2=1.49\times10^6$ psi, $G_{12}=1.04\times10^6$ psi, $\nu_{12}=0.28$. The length and width of the plate are $l=16$ in, $h=4$ in. To ensure the exactness of the solutions for the problems with uniform stress field, several different patches of elements (such as Fig. 3) have been tested and the Stroh finite element solutions (linear or quadratic elements discussed in this paper) really coincide exactly with the exact uniform stress solution ([12]).

Example 2: Stress Concentration. After the elements passed the simple test of uniform stress field, a more realistic test such as a circular hole in a rectangular plate subjected to uniform tension ($\sigma=5000$ lb./in.) is considered. Two different kinds of plate are tested. One is an isotropic plate made of aluminum whose properties are: $E=10^7$ psi, $G=3.85\times10^6$ psi, $\nu=0.3$. The other is a fiber-reinforced composite plate whose properties are: $E_1=7.8\times10^6$ psi, $E_2=2.6\times10^6$ psi, $G_{12}=1.25\times10^6$ psi, $\nu_{12}=0.25$. The length, width, and thickness of the plate are, respectively, $l=48$ in., $h=40$ in., $t=1$ in. Since the thickness of the plate is relatively thinner than the length and width, the plane stress condition will be assumed in our finite element modeling. Due to double symmetry, only a quadrant need be analyzed. A linear element mesh with 124 elements and 80 nodal points and a quadratic element mesh with 84 elements and 102 nodal points are shown in Figs. 4(*a*) and 4(*b*). The roller boundary conditions are also shown to simulate the symmetry condition and to restrain the rigid-body movement. The results for normal stresses on two edges are shown in Figs. 5(*a*) and 5(*b*) in which the exact solutions for infinite plates ([13,14]) and the numerical solutions from the commercial finite element code NASTRAN are available for

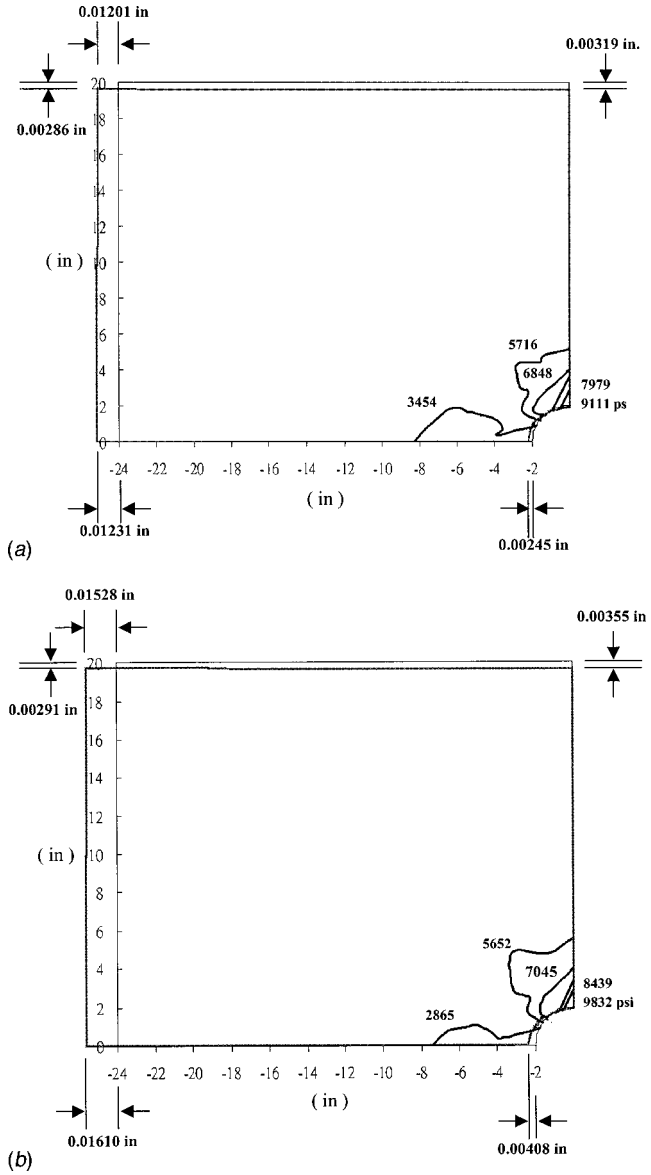


Fig. 6 (*a*) Deformed configuration and Von Mises stress contour for the plate containing a circular hole (isotropic case), (*b*) deformed configuration and Von Mises stress contour for the plate containing a circular hole (orthotropic case)

Table 1 Comparison of the computation time for the stress concentration problem (on K7-600 PC with 128MB RAM)

	Computation time	
	Isotropic plate	Orthotropic plate
124 elements (NASTRAN)	6 sec.	6 sec.
124 elements (Linear SFEM)	3 sec.	3 sec.
84 elements (Quadratic SFEM)	2 sec.	2 sec.

comparison. The agreement is satisfactory except for the points on the circular hole boundary, which can be improved by more fine meshes near holes. Figure 6 is a diagram showing the deformed configuration and the Von Mises stress contour, which is also agree with the plot shown by commercial finite element code NASTRAN. Since this simple example can also be solved very well by using the commercial finite element, to show the advantage of the present method their corresponding computation time is listed in Table 1. From this result, we see that our method really has the potential to improve the efficiency and accuracy.

6 Concluding Remarks

Through the numerical examples shown in the last section, we know that the Stroh finite element is accurate enough to perform the usual finite element task. Inheriting from the Stroh formalism, this newly developed finite element is valid for the generalized two-dimensional anisotropic elasticity problems which include in-plane and antiplane problems with the most general anisotropic (includes monoclinic, orthotropic, and isotropic, etc.) materials. Moreover, this element satisfies both of the compatibility and equilibrium conditions within each element. The derived results show that the stiffness matrix can be written explicitly and does not contain any integrals. Thus, the efficiency and accuracy compared with the conventional finite element is an interesting topic for the future study. Our preliminary numerical comparison for the problem of stress concentration supports our observation. It is also hoped that through this work similar studies combining the analytical work with the numerical techniques can be raised to improve the existing technology.

Acknowledgments

The authors would like to thank the support by National Science Council through grant NSC 82-0401-E006-088 and NSC 89-2212-E-006-022.

Appendix

From (2.3), we know that the stresses are the first derivative of the stress function. Thus, the stresses are identical to zero for zeroth-order approximation and the displacement is constant for the entire field. Therefore, the zeroth-order term represents the rigid-body translation.

We now consider the case of the uniform stress $\sigma_{11}=\sigma$ with all the other stress components $\sigma_{ij}=0$ and all the rigid-body translations are fixed ($\mathbf{h}_0=\mathbf{g}_0=\mathbf{0}$). By (3.2) we have

$$\boldsymbol{\phi}=\hat{\mathbf{N}}_3\mathbf{h}_1+\hat{\mathbf{N}}_1^T\mathbf{g}_1. \tag{A1}$$

Use of (2.3) may lead to

$$\begin{Bmatrix}\sigma_{11}\\ \sigma_{21}\\ \sigma_{31}\end{Bmatrix}=\begin{Bmatrix}\sigma\\ 0\\ 0\end{Bmatrix}=-\boldsymbol{\phi}_{,2}=-\mathbf{N}_3\mathbf{h}_1+\mathbf{N}_1^T\mathbf{g}_1,$$
$$\begin{Bmatrix}\sigma_{12}\\ \sigma_{22}\\ \sigma_{32}\end{Bmatrix}=\begin{Bmatrix}0\\ 0\\ 0\end{Bmatrix}=-\boldsymbol{\phi}_{,1}=\mathbf{g}_1. \tag{A2}$$

Hence,

$$\mathbf{g}_1=\mathbf{0},\quad \mathbf{N}_3\mathbf{h}_1=-\sigma\begin{Bmatrix}1\\ 0\\ 0\end{Bmatrix}. \tag{A3}$$

Note that $\mathbf{N}_3$ in (*A*3) is a sigular matrix. The existence of solutions to this kind of problems has been discussed in ([8]), and its solution is

$$\mathbf{h}_1=-\sigma\mathbf{N}_3^{\langle -1\rangle}\begin{Bmatrix}1\\ 0\\ 0\end{Bmatrix}+k\begin{Bmatrix}0\\ 1\\ 0\end{Bmatrix}, \tag{A4}$$

where k is an arbitrary constant and $\mathbf{N}_3^{\langle -1\rangle}$ is defined as

$$\mathbf{N}_3^{\langle -1\rangle}\mathbf{N}_3=\mathbf{N}_3\mathbf{N}_3^{\langle -1\rangle}=\mathbf{I}_2,\quad \mathbf{I}_2=\begin{bmatrix}1&0&0\\ 0&0&0\\ 0&0&1\end{bmatrix}. \tag{A5}$$

At first glance, it looks ridiculous to have a nonunique solution. By further studying the solution relating

$$\mathbf{h}_0=\mathbf{g}_0=\mathbf{g}_1=\mathbf{0},\quad \text{and}\quad \mathbf{h}_1=\begin{Bmatrix}0\\ k\\ 0\end{Bmatrix}, \tag{A6}$$

we see that

$$\mathbf{u}=x_1\mathbf{h}_1+x_2\mathbf{N}_1\mathbf{h}_1=\begin{Bmatrix}-kx_2\\ kx_1\\ 0\end{Bmatrix},$$
$$\boldsymbol{\phi}=x_2\mathbf{N}_3\mathbf{h}_1=\mathbf{0}, \tag{A7}$$

where the second equalities are obtained with the help of the structures of $\mathbf{N}_1$ and $\mathbf{N}_3$ ([5]). By the expression shown above, it can easily be understood why we have a nonunique solution because the rigid-body rotation k is not fixed for our present problem. With this presentation, it is noted that the rigid-body rotation is included in the first-order approximation. Therefore, when implementing the finite element procedure, not only the rigid-body translation $\mathbf{h}_0$ but also the rigid-body rotation $\mathbf{h}_1=(0,k,0)^T$ should be fixed in order to have a unique solution.

References

[1] Lekhnitskii, S. G., 1963, *Theory of Elasticity of an Anisotropic Body*, MIR, Moscow.
[2] Stroh, A. N., 1958, ''Dislocations and Cracks in Anisotropic Elasticity,'' Philos. Mag., **7**, pp. 625–646.
[3] Ting, T. C. T., 1996, *Anisotropic Elasticity—Theory and Applications*, Oxford Science Publications, New York.
[4] Zienkiewicz, O. C., and Taylor, R. L., 1989, *The Finite Element Method—Volume 1: Basic Formulation and Linear Problems*, 4rd ed., McGraw-Hill, New York.
[5] Ting, T. C. T., 1988, ''Some Identities and the Structures of N_i the Stroh Formalism of Anisotropic Elasticity,'' Q. Appl. Math., **46**, pp. 109–120.
[6] Hwu, C., and Yen, W. J., 1991, ''Green's Functions of Two-Dimensional Anisotropic Plates Containing an Elliptic Hole,'' Int. J. Solids Struct., **27**, No. 13, pp. 1705–1719.
[7] Ting, T. C. T., and Hwu, C., 1988, ''Sextic Formalism in Anisotropic Elasticity for Almost Non-semisimple Matrix N,'' Int. J. Solids Struct., **24**, No. 1, pp. 65–76.
[8] Hwu, C., and Ting, T. C. T., 1990, ''Solutions for the Anisotropic Elastic Wedge at Critical Wedge Angles,'' J. Elast., **24**, pp. 1–20.
[9] Barnett, D. M., and Lothe, J., 1973, ''Synthesis of the Sextic and the Integral Formalism for Dislocation, Green's Functions and Surface Waves in Anisotropic Elastic Solids,'' Phys. Norv., **7**, pp. 13–19.
[10] Hwu, C., 1993, ''Fracture Parameters for Orthotropic Bimaterial Interface Cracks,'' Eng. Fract. Mech, **45**, No. 1, pp. 89–97.
[11] Reddy, J. N., 1984, *An Introduction to the Finite Element Method*, McGraw-Hill, New York.
[12] Wu, J. Y., 1999, ''Stroh Finite Element Design for Two-Dimensional Anisotropic Elastic Bodies and Its Applications,'' *M.S. thesis*, Institute of Aeronautics and Astronautics, National Cheng Kung University, Taiwan, R.O.C.
[13] Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, McGraw-Hill, New York.
[14] Hwu, C., 1992, ''Polygonal Holes in Anisotropic Media,'' Int. J. Solids Struct., **29**, No. 19, pp. 2369–2384.

Scattering of a Rayleigh Wave by an Elastic Wedge Whose Angle is Greater Than 180 Degrees

A. K. Gautesen
Department of Mathematics
and Ames Laboratory,
Iowa State University,
136 Wilhelm Hall,
Ames, IA 50011

The steady-state problem of scattering of an incident Rayleigh wave by an elastic wedge whose angle is greater than 180 degrees is considered. The problem is reduced to the numerical solution of a pair of Fredholm integral equations of the second kind whose kernels consist of elementary functions. Numerical results are given for the amplitude and phase of the Rayleigh waves transmitted and reflected by the corner.
[DOI: 10.1115/1.1365156]

1 Introduction

Hitherto, there have been a number of investigations into the problem of scattering of Rayleigh waves by an elastic wedge. Generally these methods derive equations for the problem which are then solved either numerically or approximately. Knopoff [1] has reviewed many of the earlier efforts. Mal and Knopoff [2] give an approximate solution for wedge angles less that 180 degrees. Momoi [3] considered the 90-deg wedge. Gautesen [4,5] considered wedges whose angles are less than 180 deg using Fourier transform methods. Fujii [6] in an extension of his earlier work gives results for a wide range of wedge angles. Using the Sommerfeld-Malivzhinetz method, Budaev and Bogy [7,8] have derived exact equations for this problem. They give corrected numerical results in [9].

The method of solution used here is quite simple. First the problem is symmetrized. This reduces the number of the unknown displacements on the traction-free surfaces from four to two. Of course, the problem must then be solved twice—once for the symmetric case and once for the antisymmetric case. A free-space Green's function integral representation of the displacements is used. This representation is valid in the entire plane, and not just in the region occupied by the elastic material. From this integral representation, the tractions, dilatation, and rotation are computed on a line just below the x-axis ($x_2=0^-$)—see Fig. 1. From the integral representation of the displacements, it follows that these quantities vanish for $x_1>0$ and are unknown for $x_1<0$. Next the Fourier transform of these quantities is taken, followed by substitution of a suitable representation of the Fourier transform of the displacements $\overline{u_j}(\xi)$ on the traction-free surface $x_2=0$, $x_1>0$. The branch cut and Rayleigh poles of $\overline{u_j}(\xi)$ are shown in Fig. 2. A pair of Fredholm integral equations of the second kind is achieved by computing the jump across the branch cut of $\overline{u_j}(\xi)$. The kernels of these integral equations consist of elementary functions and are continuous except for an isolated square-root singularity. The Fourier transform of the unknown tractions, dilatation and rotation on the line $x_2=0$, $x_1<0$ are continuous across the branch cut of $\overline{u_j}(\xi)$ and therefore do not appear in the Fredholm integral equations.

The wedge angle is restricted to the range 189 deg to 327 deg. Beyond this range of wedge angles, numerical instabilities arose because a pole is approaching the contour of integration. This problem can be corrected by using a more sophisticated integration scheme to account for this phenomenon. These results appear to be approaching the known limiting values for a 180-deg wedge (an elastic half-space) and a 360-deg wedge (a semi-infinite crack).

In the next section the governing Fredholm integral equations are derived. In the last section the numerical results are presented and discussed.

2 Derivation of the Integral Equations

The diffraction of steady-state waves by an elastic wedge whose wedge angle α is greater than 180 degrees is considered. The location of the wedge; the coordinate system; and the incident, transmitted and reflected surface waves are shown in Fig. 1. By using the free-space Green's stress tensor, an integral representation of the total displacements $u_k(x_1,x_2)$ is given by

$$u_k(x_1,x_2)H(\mathbf{x}) = u_k^{in}(x_1,x_2)+\sum_{i=1}^{2}\int_0^{\infty}\left\{\tau_{i2;k}^{G}(x_1-y,x_2)u_i(y,0) + \sum_{j=1}^{2}\tau_{ij;k}^{G}(x_1-t_1y,x_2-t_2y)n_ju_i(t_1y,t_2y)\right\}dy, \quad k=1,2. \tag{2.1}$$

Here, $H(\mathbf{x})=1$ if (x_1,x_2) is a point in the elastic material and $H(\mathbf{x})=0$ otherwise. The incident field is denoted by $u_k^{in}(x_1,x_2)$ and since the incident field is a Rayleigh surface wave, $u_k^{in}(x_1,x_2)=0$. The free-space Green's stress tensor is denoted by $\tau_{ij;k}^{G}(x_1,x_2)$. The importance of (2.1) is that it holds for all points (x_1,x_2) and not just for points lying within the elastic material. The normal n_j and tangent t_j to the lower traction-free surface are given by

$$n_1=t_2=\sin\alpha, \quad n_2=-t_1=-\cos\alpha \tag{2.2}$$

where α is the wedge angle. To reduce the number of unknowns the problem is divided into an antisymmetric ($l=1$) and symmetric ($l=2$) problem. For the symmetric problem, the normal and tangential displacements on the traction-free surfaces are equal, and for the antisymmetric problem they have the opposite sign. Substitution of this result into (2.1) reduces the number of unknowns by one half, but then the problem must be solved twice (once for $l=1$ and once for $l=2$).

The rotation, dilatation, shear, and normal tractions are computed from (2.1) on the line $x_2=-\delta^2$, $-\infty<x_1<\infty$. These quantities vanish for $x_1>0$ and are unknown for $x_1<0$. Then the Fourier transform of these quantities is taken. The Fourier transform is defined by

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 7, 2000; final revision, Nov. 21, 2000. Associate Editor: A. K. Mal. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS.

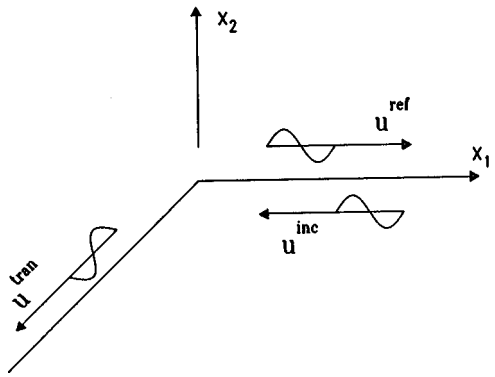


Fig. 1 Incident, transmitted, and reflected surface waves

$$\mathfrak{F}\{u(x_1))\}=k_L\int_0^\infty \exp[ik_L\xi x_1]u(x_1)dx_1 \tag{2.3}$$

where k_L denotes the wave number of longitudinal waves. After taking the limit as $\delta\to 0$, the result is

$$\sum_{j=1}^{2}(-1)^{i+j}B_{ij}(\xi,\gamma_1,\gamma_2)\overline{u_j}(\xi)+\gamma_i^{-1}U_i(\xi)=D_i^- \tag{2.4}$$

$$\gamma_i^{-1}R(\xi)\overline{u_i}(\xi)+\sum_{j=1}^{2}B_{ij}(\xi,\gamma_1,\gamma_2)U_j(\xi)=D_{i+2}^- \tag{2.5}$$

where

$$B_{ij}(\xi,z_1,z_2)=(-1)^i z_i^{-1}a(\xi)\delta_{ij}-2\xi(1-\delta_{ij}) \tag{2.6}$$

$$\gamma_i=(c_2/c_i+\xi)^{1/2}(c_2/c_i-\xi)^{1/2} \tag{2.7}$$

$$a(\xi)=\kappa^2-2\xi^2,\quad \kappa=c_2/c_1 \tag{2.8}$$

$$U_i(\xi)=(-1)^l\sum_{j=1}^{2}b_{ij}(\xi,z_1,z_2)\overline{u_j}(\xi_i),\quad z_j=\gamma_j \tag{2.9}$$

$$b_{ij}(\xi,z_1,z_2)=a(\xi_i)\delta_{ij}+2(-1)^i\xi_i(1-\delta_{ij})(\xi\sin\alpha+z_i\cos\alpha) \tag{2.10}$$

$$\xi_i=\xi\cos\alpha-z_i\sin\alpha \tag{2.11}$$

$$R(\xi)=a^2(\xi)+4\xi^2\gamma_1\gamma_2, \tag{2.12}$$

δ_{ij} denotes the Kronecker delta, and c_2 and c_1 denote the speed of longitudinal and transverse waves. The functions D_j^- represent the unknown Fourier transforms of the dilatation, rotation, and tractions for $x_1<0$. These functions are analytic in the lower half of the complex ξ-plane, while $\overline{u_j}(\xi)$, which denotes the Fourier transform of $u_j(x_1,0)$, $x_1>0$, is analytic in the upper half of the complex ξ-plane. The Rayleigh function $R(\xi)$ satisfies $R(\kappa_R)=0$ where $\kappa_R=c_2/c_R$ and c_R is the speed of Rayleigh waves. It is convenient to consider the following three linear combinations of (2.4)–(2.5):

$$\overline{u_i}(\xi)-\frac{\gamma_i}{R(\xi)}\sum_{j=1}^{2}B_{ij}(\xi,\gamma_1,\gamma_2)U_j(\xi)=\frac{1}{2\xi}\left(D_{3-i}^-+(-1)^i\frac{a(\xi)}{R(\xi)}D_{5-i}^-\right) \tag{2.13}$$

$$[a^4(\xi)-16\xi^4\gamma_1^2\gamma_2^2]\overline{u_1}(\xi)+4\xi^2\gamma_1^2\gamma_2 a(\xi)\times\sum_{j=1}^{2}(-1)^j B_{1j}(\xi,\gamma_1,\gamma_2)\overline{u_j}(\xi)-2\xi\gamma_1 R(\xi)U_2(\xi)=D \tag{2.14}$$

where

$$D=a^2(\xi)[\gamma_1 D_3^-+a(\xi)U_1(\xi)]-8\xi^3\gamma_1^2\gamma_2^2 D_2^-.$$

Figure 2 shows the poles and branch cut of $\overline{u_j}(\xi)$, the Fourier transform of the displacements on the upper traction-free surface. The following representation of $\overline{u_j}(\xi)$ is used:

$$\overline{u_j}(\xi)=\frac{u_j^i}{\xi-\kappa_R}+\frac{A_l u_j^r}{\xi+\kappa_R}-\frac{1}{2\pi i}\int_1^\infty\frac{w_j(z)}{z+\xi}dz+B_jF^+(\xi) \tag{2.15}$$

where

$$u_1^i=-u_1^r=2\kappa_R\gamma_1(\kappa_R)/\kappa^2,\quad u_2^i=u_2^r=a(\kappa_R)/\kappa^2. \tag{2.16}$$

The first term in (2.15) represents the incident surface wave. The next term represents the surface wave reflected and transmitted by the corner—recall that by symmetrizing the problem there is a surface wave incident on both of the traction-free surfaces. The constant A_l is then given by

$$A_l=R_c+(-1)^l T_c \tag{2.17}$$

where R_c and T_c denote the reflection and transmission coefficients for surface waves. The third term in (2.15) is analytic everywhere except along the branch cut shown on Fig. 2. The quantity $w_j(\xi)$ is to be determined. The constant B_j in the last term of (2.15) is to be chosen so that $w_j(\xi)$ is finite at $\xi=\kappa_R$. The function $F^+(\xi)$ is analytic in the upper half of the complex ξ-plane. It is easily computed in closed form by the Wiener-Hopf sum splitting:

$$F^+(\xi)+F^-(\xi)=\frac{\gamma_1(\xi)-\gamma_1(\kappa_R)}{\kappa_R^2-\xi^2}-\frac{\gamma_1(\xi)-\gamma_1(z_0)}{z_0^2-\xi^2} \tag{2.18}$$

where the choice $z_0=\kappa(1-i)^{1/2}$ has been made and $F^-(\xi)=F^+(-\xi)$.

Next (2.15) is substituted into (2.4), (2.5), (2.13), and (2.14). The result contains no poles in the lower half of the complex ξ-plane. Then the difference of (2.4) with $i=1$, (2.13) and (2.14) evaluated at $\xi=-\xi_0+i0$ and $\xi=-\xi_0-i0$ is taken. In (2.13) ξ_0 is restricted to the range $\xi_0>\kappa$, while in (2.4) and (2.14), $1<\xi_0<\kappa$. This yields the following Fredholm integral equations for $w_j(\xi)$:

$$w_j(\xi)+\sum_{k=1}^{2}\left\{-\frac{1}{2\pi i}\int_1^\infty W_{jk}(\xi,z)w_k(z)dz+W_{jk}(\xi,-\kappa_R)u_k^i+A_l W_{jk}(\xi,\kappa_R)u_k^r+B_k G_{jk}(\xi)\right\}=0,\quad c_j/c_1<\xi \tag{2.19}$$

$$a(\xi)w_1(\xi)+2\xi\gamma_1 w_2(\xi)=0,\quad 1<\xi<\kappa. \tag{2.20}$$

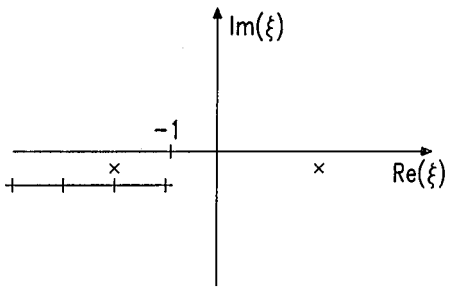


Fig. 2 Branch cut and location of Rayleigh poles of $\overline{u_j}(\xi)$

When $1<\xi<\kappa$, $w_2(\xi)$ is defined by (2.20). When the angle exterior to the wedge is smaller than the angle of head waves, some additional steps are required to arrive at (2.20). These are discussed at the end of this section. Here,

$$W_{jk}(\xi,z)=V_{jk}(-\xi,\gamma_1,\gamma_2,z)-V_{jk}(-\xi,s\gamma_1,-\gamma_2,z) \tag{2.21}$$

where $s=\mathrm{sgn}(\kappa-\xi)$ and

$$V_{jk}(\xi,z_1,z_2,z)=-\sum_{i=1}^{2}\frac{z_j B_{ji}(\xi,z_1,z_2)b_{ik}(\xi,z_1,z_2)}{(a^2(\xi)+4\xi^2 z_1 z_2))(z+\xi_i)},\quad \xi<-\kappa \tag{2.22}$$

$$V_{1k}(\xi,z_1,z_2,z)=-\frac{2\xi z_1}{a^2(\xi)-4\xi^2 z_1 z_2}\times\left(\frac{2\xi z_1 z_2 a(\xi)B_{1k}(\xi,z_1,z_2)}{(a^2(\xi)+4\xi^2 z_1 z_2)(z+\xi)}+\frac{b_{2k}(\xi,z_1,z_2)}{(z+\xi_2)}\right),\quad -\kappa<\xi<-1. \tag{2.23}$$

Also,

$$\hat{G}_{jk}(\xi)=G_{jk}(\xi)+\gamma_j\delta_{jk}(1+s)(z_0^2-\kappa_R^2)/[(\kappa_R^2-\xi^2)(z_0^2-\xi^2)] \tag{2.24}$$

where $G_{jk}(\xi)$ is the discontinuity across the branch cut of the function $F_{jk}(\xi,\gamma_1,\gamma_2)$ (see (2.21)). The function $F_{jk}(\xi,z_1,z_2)$ is defined by the right side of (2.22) and (2.23) with the quantities $(z+\xi_i)^{-1}$ and $(z+\xi)^{-1}$ replaced by $F^+(\xi_i)$ and $F^+(\xi)$, respectively.

For numerical stability of the solution to (2.19), $w_j(\kappa_R)$ is required to be bounded. Thus the constants B_j are determined by

$$\lim_{\xi\to\kappa_R}(\kappa_R^2-\xi^2)\sum_{k=1}^{2}\left\{-\frac{1}{2\pi i}\int_1^\infty W_{jk}(\xi,z)w_k(z)dz+W_{jk}(\xi,-\kappa_R)u_k^i+A_l W_{jk}(\xi,\kappa_R)u_k^r+B_k G_{jk}(\xi)\right\}=0,\quad j=1,2. \tag{2.25}$$

Two more useful relations follow from evaluating either (2.4) or (2.5) at $\xi=-c_2/c_i$. In order that the displacements on the traction-free surfaces have the right asymptotic behavior far from the corner (see Gautesen [10]), $\bar{u}_2(-1)$ and $\bar{u}_1(-\kappa)$ must be finite. The unknown quantities $D_j^-(\xi)$ are analytic near these points. Thus,

$$a(1)\bar{u}_2(-1)+(-1)^l[\sin 2\alpha\bar{u}_1(-\cos\alpha)+a(\cos\alpha)\bar{u}_2(-\cos\alpha)]=0 \tag{2.26}$$

$$\bar{u}_1(-\kappa)-(-1)^l[\cos 2\alpha\bar{u}_1(-\kappa\cos\alpha)+\sin 2\alpha\bar{u}_2(-\kappa\cos\alpha)]=0. \tag{2.27}$$

Next (2.15) is substituted into (2.26) and (2.27). None of the resulting integrals are Cauchy principal-value integrals.

To obtain the final integral equations the constants A_l, B_1, and B_2 are eliminated from (2.19) by using (2.25) and (2.27) when $j=1$ and by using (2.25) and (2.26) when $j=2$. From the integral equations it follows that $w_1(\xi)=O((c_2/c_j-\xi)^{1/2})$ near $\xi=c_2/c_j$, $j=1,2$ and that $w_2(\xi)=O((1-\xi)^{1/2})$ near $\xi=1$. It is convenient to introduce new unknown functions $\hat{w}_j(\xi)$ defined by

$$w_1(\xi)=\gamma_1\gamma_2\hat{w}_1(\xi)/\xi^2,\quad w_2(\xi)=\gamma_1\hat{w}_2(\xi)/\xi. \tag{2.28}$$

Then the kernels of resulting integral equations are continuous except for square-root singularities at $z=1,\kappa$.

Care must be exercised in devising a numerical integration scheme to solve the resulting Fredholm integral equations of the second kind. For example, the stresses are singular near the corner (see Karp and Karal [11]). This singularity is different for the symmetric ($l=2$) and antisymmetric ($l=1$) problems. This singularity reflects itself in the behavior of $\hat{w}_j(\xi)$ as $\xi\to\infty$ and has been accounted for.

When the wedge's exterior angle $\alpha_0=2\pi-\alpha$ is less than the head wave angle $\alpha_h=\arccos(\kappa^{-1})$, some extra steps are required to achieve (2.20). After substitution of (2.15) into (2.4) when $i=1$, taking the difference of the result evaluated at $\xi=-\xi_0+i0$ and $\xi=-\xi_0-i0$, $1<\xi_0<\kappa$ gives

$$w(\theta)-(-1)^l H(\theta-n\alpha_0)w(\theta-n\alpha_0)=0,\quad 0<\theta<\alpha_h \tag{2.29}$$

where $n=1$, $H(\theta)$ is the Heaviside function, and with $\xi=\kappa\cos\theta$, $w(\theta)$ is defined by

$$w(\theta)=a(\xi)w_1(\xi)+2\xi\gamma_1 w_2(\xi),\quad 1<\xi<\kappa. \tag{2.30}$$

Examination of (2.29) on the interval $\alpha_0<\theta<\min(2\alpha_0,\alpha_h)$ shows that $w(\theta)$ also vanishes on this interval. Thus either $w(\theta)=0$, $0<\theta<\alpha_h$, or $w(\theta)$ satisfies (2.29) with $n=2$. In the latter case this process is repeated until $n\alpha_0\geqslant\alpha_h$. Eventually this leads to the conclusion that $w(\theta)=0$, $0<\theta<\alpha_h$, which by (2.30) establishes (2.20).

3 Results and Discussion

The reflection coefficient R_c and transmission coefficient T_c are related to A_l by $A_l=R_c+(-1)^l T_c$, $l=1,2$. Once the solution $w_j(\xi)$ has been found, the constant A_l can be computed from either (2.26) or (2.27) after using (2.25) to eliminate B_1 and B_2. These values should of course be the same. Here the maximum value of the amplitude of their differences was found to be less that 0.0003 over the range of wedge angles considered. With Poisson's ratio equal to 0.25, the amplitude and phase of the reflection and transmission coefficients versus the wedge angle are plotted in Figs. 3 and 4, for 189 deg$<\alpha<$327 deg. Beyond these angles the numerical scheme became unstable. For wedge angles less than 189 deg a pole in the analytic continuation of the nonhomogeneous term approaches the contour of integration, while for wedge angles greater than 327 deg a pole in the analytic continuation of the kernel approaches the contour of integration. The numerical integration scheme that has been used did not account for this behavior. For a 180-deg wedge (an elastic half-space) $R_c=0$ and $T_c=1$, while for a 360-deg wedge (a semi-infinite

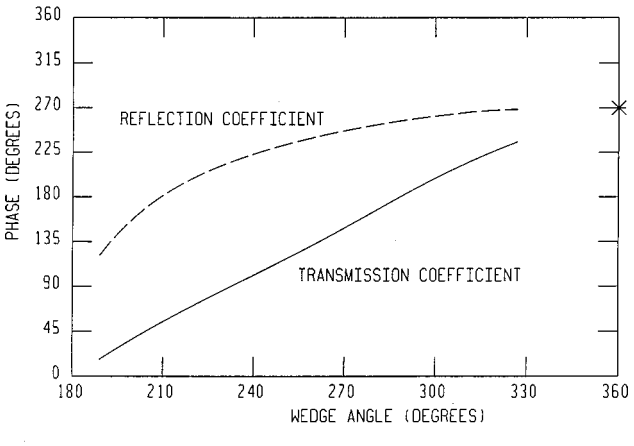


Fig. 3 Amplitude of the reflection and transmission coefficients versus wedge angle (in degrees) for Poisson's ratio =0.25. The exact values for a 360-deg wedge are indicated by an asterisk.

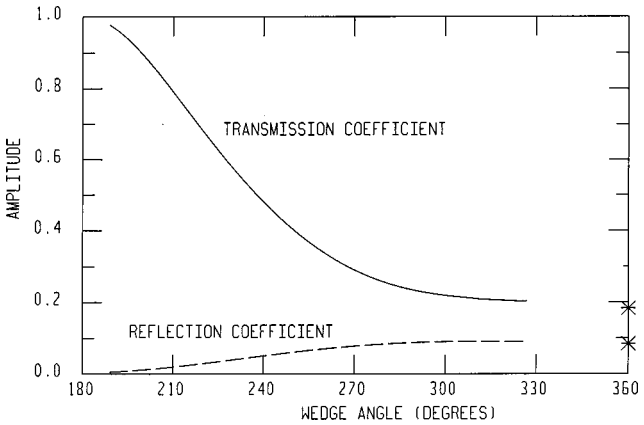


Fig. 4 Phase (in degrees) of the reflection and transmission coefficients versus wedge angle (in degrees) for Poisson's ratio=0.25. The exact values for a 360-deg wedge are indicated by an asterisk.

crack) $R_c = -0.0835\ldots i$ and $T_c = -0.1817\ldots i$ (see e.g., Achenbach, Gautesen, and McMaken [12]). The limiting values of the reflection and transmission coefficients for a 360-deg wedge are indicated by an asterisk in Figs. 3 and 4. In Fig. 3 it appears that as the wedge angle increases from 180 deg to 360 deg, the magnitude of the transmission coefficient decreases from 1 to $0.181\ldots$, while the magnitude of the reflection coefficient increases from 0 to $0.083\ldots$. In Fig. 4 it appears that as the wedge angle increases from 180 deg to 360 deg, the phase of the transmission coefficient increases from 0 deg to 270 deg, while the phase of the reflection coefficient increases from an apparent value of 90 deg to 270 deg.

The numerical results presented here are in good agreement with the corresponding results of Fujii [6], except for the phase of the reflection coefficient when the wedge angle is less than 220 deg. Over this range of wedge angles, the magnitude of the reflection coefficient is small. Also, when the method presented here is used to obtain numerical results for steel, good agreement is found with the corresponding results of Budaev and Bogg [9], except for the phase of the reflection coefficient. Note that in [6] and [7], the phase of the reflection and transmission coefficients are defined as the negative of that used herein.

Acknowledgment

Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under contract number W-7405-ENG-82. This research was supported in part by the Office of Basic Energy Science, Applied Mathematical Sciences Division.

References

[1] Knopoff, L., 1969, ''Elastic Wave Propagation in a Wedge,'' *Wave Propagation in Solids*, J. Miklowitz, ed., ASME, New York, pp. 3–42.
[2] Mal, A. K., and Knopoff, L., 1966, ''Transmission of Rayleigh Waves at a Corner,'' Bull. Seismol. Soc. Am., **56**, pp. 455–466.
[3] Momoi, H., 1980, ''Scattering of Rayleigh Waves in an Elastic Quarter Space,'' J. Phys. Earth, **28**, pp. 385–413.
[4] Gautesen, A. K., 1985, ''Scattering of a Rayleigh Wave by an Elastic Quarter Space,'' ASME J. Appl. Mech., **107**, pp. 664–668.
[5] Gautesen, A. K., 1987, ''Scattering of a Rayleigh Wave by an Elastic Wedge,'' Wave Motion, **9**, pp. 51–59.
[6] Fujii, K., 1994, ''Rayleigh-Wave Scattering at Various Wedge Corners: Investigation in the Wider Range of Wedge Angles,'' Bull. Seismol. Soc. Am., **84**, pp. 1916–1924.
[7] Budaev, B. V., and Bogy, D. B., 1995, ''Rayleigh Wave Scattering by a Wedge,'' Wave Motion, **22**, pp. 239–257.
[8] Budaev, B. V., and Bogy, D. B., 1996, ''Rayleigh Wave Scattering by a Wedge: II,'' Wave Motion, **24**, pp. 307–314.
[9] Budaev, B. V., and Bogy, D. B., 2001, ''Scattering of Rayleigh and Stoneley Waves by Two Adhering Elastic Wedges,'' Wave Motion, in press.
[10] Gautesen, A. K., 1979, ''On Matched Asymptotic Expansions for Two-Dimensional Elastodynamic Diffraction by Cracks,'' Wave Motion, **1**, pp. 17–140.
[11] Karp, S. N., and Karal, F. C., 1962, ''The Elastic-Field Behavior in the Neighborhood of a Crack of Arbitrary Angle,'' Commun. Pure Appl. Math., **15**, pp. 413–421.
[12] Achenbach, J. D., Gautesen, A. K., and McMaken, H., 1982, *Ray Methods for Waves in Elastic Solids with Application to Scattering by Cracks*, Pitman, Boston.

Boundary Element Formulation for Thermal Stresses During Pulsed Laser Heating

P. H. Tehrani
Mechanical Engineering Department,
Amirkabir University of Technology,
Tehran, Iran

L. G. Hector, Jr.[1]
Surface Science Division,
Alcoa Technical Center,
Alcoa Center, PA 15069

R. B. Hetnarski
Department of Mechanical Engineering,
Rochester Institute of Technology,
Rochester, NY
Fellow ASME

M. R. Eslami
Mechanical Engineering Department,
Amirkabir University of Technology,
Tehran, Iran
Mem. ASME

Pulsed lasers are used in a variety of materials processing applications that range from heating for metallurgical transformation to scribing vehicle identification numbers on anodized aluminum strips. These lasers are commonly configured to deliver a large quantity of heat energy in very short time intervals and over very small areas due to the manner in which radiant energy is stored within, and then released from, the laser resonator. At the present time, little is known about the effect of pulse duration on thermomechanical distortion during heating without phase change. To explore this issue, a boundary element method was developed to calculate temperature, displacement, and thermal stress fields in a layer that is rigidly bonded to an inert semi-space. The layer absorbs thermal energy from a repetitively pulsed laser in the plane of its free surface. The effects of two pulse durations, which differ by four-orders-of-magnitude, were examined in this work. The temporal profiles of ultrafast pulses of the order of ten picoseconds (such as those emitted by a mode-locked laser), and pulses of the order of tens-of-nanoseconds (such as those emitted by a Q-switched Nd:YAG laser) were mathematically modeled using a rectified sine function. The spatial profile of each pulse was shaped to approximate a Gaussian strip source. The equations of coupled thermoelasticity, wherein the speed of mechanical distortion due to material expansion during heat absorption is finite, but the speed of heat propagation within the layer is infinite, were solved for both pulse durations. The resulting temperature and stress fields were compared with those predicted in the limit of no thermomechanical coupling. [DOI: 10.1115/1.1365155]

1 Introduction

Theoretical and experimental investigations of laser heating of materials, and resulting thermal stresses, began to appear in the literature not long after the first working laser was invented in the early 1960s ([1]). Since that time, interest in the effects of both continuous wave and pulsed laser heating on the properties of metals, ceramics, and semiconductors has been maintained, in part, by the rapid growth of the microelectronics industry. An ever increasing demand for improved material properties at smaller and smaller length scales has also created an incentive for controlling thermal stress levels through control of critical laser processing parameters such as power density, pulse repetition rate, duty cycle, and depth of focus.

Absorption of radiant energy from a pulsed laser typically results in a very rapid rise in temperature in a material surface over a very short time interval. Thermal stresses result due to the rapid movement of heated material against the constraint of colder, albeit deformable surrounding material. Thermomechanical energy transport may occur through wave propagation ([2]): This depends upon the duration of each laser pulse and the efficiency with which electrons in the material transfer radiant energy to the surrounding lattice ([3]). If thermal stress levels are uncontrolled during laser processing, then a variety of material and process related problems, such as crack nucleation and propagation ([4–6]), and undesirable microstructural transformations ([7,8]), may result. It is therefore useful to predict the thermal stress levels that arise during a laser processing application to establish guidelines for appropriate beam settings and to ultimately achieve desired material properties.

Theoretical studies of thermal stresses in laser processing have been directed toward a wide variety of applications and materials. For example, Welsh et al. [9] calculated thermal stresses in the extreme cases of an elastic half-space and a thin film under steady-state heating with a Gaussian laser source. Uglov et al. [10] addressed the problem of predicting thermal stresses in a metallic cylinder due to pulsed laser heating. The thermoelastic response of a half-space subjected to volumetric heating by a concentrated heat flux for various values of absorption coefficient was derived by Germanovich et al. [11]. Suh and Burger [12] developed a model to predict thermal stresses in aluminum plates of various thicknesses due to a high power Nd:YAG laser pulse. Cohen et al. [13] calculated thermal stresses in thin films due to laser pulsing and developed a fatigue criterion based on the results of their model. Fesenko [14] developed a model to predict thermal stresses due to a periodic, volumetric heat source in a rectangular prism and applied his model to understand the thermomechanical response of a potassium dihydrophosphate monocrystal. The thermal stress field in an elastic half-space due to a single pulse from a laser was derived for the general case of a mixed-mode structure beam by Hector and Hetnarski [15]. Solutions for thermoelastic stress and displacement fields in thin films and thick layers that absorb thermal energy from a repetitively pulsed laser (with axisymmetric geometry) were reported by Kim et al. [16]. The coupled solution of a two-dimensional domain subjected to thermal shock from a laser source was studied by Tehrani and Eslami [17]. None of these papers considered the effects of different pulse durations (either with or without coupling of the thermal and mechanical fields) on predicted temperature and stress fields.

In the present paper, we calculate the thermal stresses in an elastic layer bonded to a thermomechanically inert semi-space using a boundary element formulation (see Kim et al. [16] for a discussion of the technological significance of this problem). The

[1]Currently at General Motors Research and Development Center, Mail Code 480-106-224, 30500 Mound Road, Warren, MI 48090-9055.

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 9, 2000; final revision, December 5, 2000. Associate Editor: A. K. Mal. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

free surface of the layer absorbs thermal energy from a repetitively pulsed laser in its surface plane. The energy equation, and equation of motion, are solved for the case of an ultrafast pulse train, wherein the pulse duration is of the order of ten picoseconds. These equations are also solved for a pulse duration which is of the order of tens-of-nanoseconds in order to compare differences in the material response due to the picosecond pulse duration. A boundary element method is applied to the solution of these equations for the temperature, displacement, and stress fields as functions of time and position assuming that the thermal and mechanical fields are coupled. Coupling of the thermal fields is due to finite transport of mechanical energy due to rapid absorption of thermal energy from the laser source followed by material expansion against colder surrounding material. Alternatively, heat transport is assumed to be infinite. Key differences in the layer displacements and stresses due to the picosecond and nanosecond pulse trains are carefully examined both with and without coupling.

2 Coupled Formulation

At $t>0$, the homogeneous, thermoelastic layer shown in Fig. 1 is exposed to a thermal load from a repetitively pulsed laser. The layer extends from $-\infty<y<\infty$, and has a width, l, along the positive x-direction. The layer is rigidly bonded to an inert semispace along $x=l$. Heat from a constant frequency laser pulse train is absorbed in the plane of the free surface and instantaneously diffuses throughout the material. The layer is assumed to be insulated at $x=l$.

In the absence of body forces and volumetric heat generation, the governing equations for dynamic, coupled thermoelasticity in the time domain may be written as

$$(\lambda+\mu)u_{j,ij}+\mu u_{i,jj}-\rho\ddot{u}_i-\gamma T_{,i}=0 \tag{1}$$

$$kT_{,jj}-\rho c_e\dot{T}-\gamma T_0\dot{u}_{j,j}=0 \tag{2}$$

where Eq. (1) is the equilibrium equation, Eq. (2) is the energy equation, and a dot denotes time differentiation ([18,19]). Note that λ, μ are the Lamé constants ([20]); u_i is the displacement vector which consists of two components, i.e., u, the displacement along the layer axis, x, and v, the displacement parallel to the layer surfaces along y; ρ, $T=T(x,y,t)$, T_0, k, c_e are the mass density, reference temperature, reference temperature, thermal conductivity, and specific heat at constant strain, respectively. Equations (1) and (2) are coupled through the stress-temperature modulus, γ, where

$$\gamma=\eta(3\lambda+2\mu) \tag{3}$$

and η is the coefficient of thermal expansion. The thermal boundary conditions are

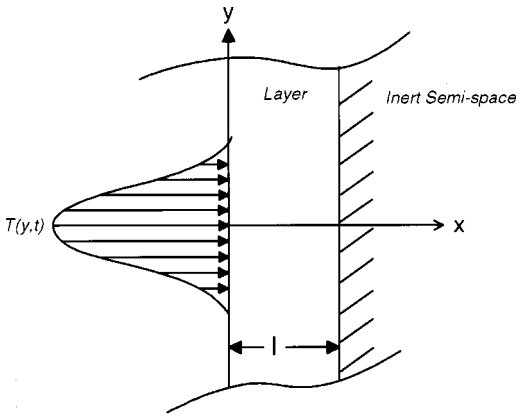


Fig. 1 Layer subjected to thermal load from pulsed laser

$$T=0 \quad \text{at} \quad t=0 \tag{4}$$

$$T=\left|\sin\left(\frac{2\pi t}{t_w}\right)\right|\exp\left\{-2.2\left(\frac{|y|}{l}\right)^3\right\} \quad \text{at } x=0 \tag{5}$$

$$\frac{\partial T}{\partial x}=0 \quad \text{at } x=l \tag{6}$$

$$T=0 \quad \text{at } y\to\pm\infty \tag{7}$$

$$\frac{\partial T}{\partial y}=0 \quad \text{at } y\to 0. \tag{8}$$

The coupling of Eqs. (1) and (2) implies that heat absorbed in the surface plane of the layer produces a longitudinal wave through dilatation: Material which lies ahead of the wave is in a state of compression, whereas material in the wake of the wave is in a state of tension. Generation of the longitudinal wave during pulsed laser heating is strongly dependent upon the pulse duration, t_w. Note that the energy equation does not include a thermal inertia term (i.e., $T_{,tt}$) since the speed of heat propagation in the layer is infinite. For pulse durations of the order of a few nanoseconds, the classical Fourier heat conduction model provides, in most cases, an adequate approximation of the thermal field ([3]). However, for a pulse durations of the order of a few picoseconds, it is likely that the thermal inertia term is important thereby requiring generalization of the equations of thermoelasticity ([21,22]). Nevertheless, we shall ignore the thermal inertia term for both pulse durations in order to reveal significant differences between the laser-irradiated material response as predicted by the present coupled and uncoupled formulations. Equation (5) represents the temperature profile on the free surface of the layer due to a repetitively pulsed laser: Thermal absorption into the surface is assumed to be perfect ([23]). It consists of an oscillatory temporal profile and a Gaussian spatial component. The temporal profile of each pulse is assumed to follow a rectified sine function: This profile approximates the nearly Gaussian temporal profile associated with many pulsed solid state lasers ([24,25]). Each pulse achieves peak power at $t=t_w/2$. Note that each pulse is immediately activated following deactivation of the preceeding pulse: The pulse duration and pulse train period are thus equivalent. The two-dimensional spatial profile of the beam is a Gaussian strip source since it consists of a Gaussian profile along the y-direction with no variation along z for any y. The surface temperature peaks along $y=0$. Clearly, this distribution differs from the more common cylindrically symmetric mode structure associated with many laser beams. It is, however, similar to the oblong beam profile emitted by repetitively pulsed excimer lasers ([26]). Alternatively, if a cylindrically symmetric laser beam is passed through a cylindrical lens, the output will resemble the spatial profile given by Eq. (5) ([27]). Beam spreading methodologies of this type are discussed in ([28]). Note that the 2.2 ''shape parameter'' in Eq. (5) guarantees that the spatial component of the temperature distribution approximates a Gaussian profile along the y-direction. Equation (6) states that the boundary along $x=l$ is insulated.

The mechanical boundary conditions are

$$u=v=0 \quad \text{at } x=l \tag{9}$$

$$u=v=0 \quad \text{at } y\to\pm\infty \tag{10}$$

where $u=u(x,y,t)$ and $v=v(x,y,t)$ are the displacement components along the positive x and y-directions, respectively, in Fig. 1.

The time for heat to diffuse across the layer thickness is given by l^2/α where α is thermal diffusivity. If

$$\frac{l^2}{\alpha t_w}\ll 1 \tag{11}$$

then heat conduction along the layer thickness is negligible and the layer can be treated as a thin film.

It is convenient for the numerical calculations to introduce the following dimensionless variables:

$$\hat{x}=\frac{x}{\beta};\quad \hat{y}=\frac{y}{\beta};\quad \hat{t}=\frac{tC_1}{\beta};\quad \hat{\sigma}_{ij}=\frac{\sigma_{ij}}{\gamma T_0}\quad \hat{u}_i=\frac{(\lambda+2\mu)u_i}{\beta\gamma T_0};$$

$$T=\frac{T-T_0}{T_0} \tag{12}$$

where

$$\beta=\frac{\alpha}{C_1} \tag{13}$$

is a unit length, and the longitudinal or thermoelastic shock wave speed is given by

$$C_1=\sqrt{\frac{(\lambda+2\mu)}{\rho}}. \tag{14}$$

A wavefront appears if

$$\frac{l}{C_1 t_w}>1. \tag{15}$$

Note that the layer thickness, l, can be written in terms of a dimensionless thickness, L, through

$$l=L\beta. \tag{16}$$

The dimensionless stress tensor is denoted by $\hat{\sigma}_{ij}=\hat{\sigma}_{ij}(x,y,t)$.

The dimensionless forms of Eqs. (1) and (2) are

$$\left(\frac{\mu}{\lambda+2\mu}\right)u_{i,jj}+\left(\frac{\lambda+\mu}{\lambda+2\mu}\right)u_{j,ij}-T_{,i}-\ddot{u}_i=0 \tag{17}$$

$$T_{,jj}-\dot{T}-C\dot{u}_{j,j}=0 \tag{18}$$

where for convenience we have dropped the hat notation in Eqs. (16) and (17). The dimensionless coupling coefficient is defined as

$$C=\frac{T_0\gamma^2}{\rho c_e(\lambda+2\mu)}. \tag{19}$$

Transferring Eqs. (17) and (18) to the Laplace transform domain yields

$$\left(\frac{\mu}{\lambda+2\mu}\right)u_{i,jj}+\left(\frac{\lambda+\mu}{\lambda+2\mu}\right)u_{j,ij}-T_{,i}-s^2u_i=0 \tag{20}$$

$$T_{,jj}-sT-Csu_{j,j}=0 \tag{21}$$

where s is the Laplace transform variable due to transformation with respect to the time variable. Note that the transformed functions could be written as u_i^* and T^* (for example) so as to distinguish them from their counterparts in the physical domain. We elected to retain the same notation for these functions as used for the physical domain since we will shortly bring them back to this domain through numerical inversion. Equations (20) and (21) can be written in matrix form as

$$\Theta_{ij}U_j=0 \tag{22}$$

where

$$U_i=[U\quad V\quad T]. \tag{23}$$

For a two-dimensional domain, the operator Θ_{ij} reduces to

$$\Theta_{ij}=\begin{bmatrix}\left(\frac{\mu}{\lambda+2\mu}\right)\Delta+\left(\frac{\lambda+\mu}{\lambda+2\mu}\right)D_1^2-s^2 & \left(\frac{\lambda+\mu}{\lambda+2\mu}\right)D_1D_2 & -D_1\\ \left(\frac{\lambda+\mu}{\lambda+2\mu}\right)D_1D_2 & \left(\frac{\mu}{\lambda+2\mu}\right)\Delta+\left(\frac{\lambda+\mu}{\lambda+2\mu}\right)D_2^2-s^2 & -D_2\\ -sCD_1 & -sCD_1 & \Delta-s\end{bmatrix} \tag{24}$$

where $D_i=\partial/\partial x_i$ for $(i=1,2)$ and Δ denotes the Laplacian.

In order to derive the boundary integral problem, we start with the following weak formulation of the differential equation set, Eqs. (22), for the fundamental solution tensor V_{ik}^*;

$$\int_\Omega(\Theta_{ij}U_j)V_{ik}^*d\Omega=0. \tag{25}$$

After integrating by parts over the domain and taking a limiting procedure with which an internal source approaches a boundary point, we obtain the following boundary integral equation:

$$C_{kj}U_k(y,s)=\int_\Gamma[\tau_\epsilon(x,s)V_{\epsilon j}^*(x,y,s)-U_\epsilon(x,s)\Sigma_{\epsilon j}^*(x,y,s)]d\Gamma(x)+\int_\Gamma[T_{,n}(x,s)V_{3j}^*(x,y,s)-T(x,s)V_{3j,n}^*(x,y,s)]d\Gamma(x) \tag{26}$$

where $U_\epsilon=u_\epsilon(\epsilon=1,2)$, $U_3=T$ and C_{kj} denotes the shape coefficient tensor. The kernel $\Sigma_{\epsilon j}^*$ in Eq. (26) is defined by

$$\Sigma_{\epsilon j}^*=\left[\left\{\left(\frac{\lambda}{\lambda+2\mu}\right)V_{kj,k}^*+sCV_{3j}^*\right\}\delta_{\epsilon\beta}+\frac{\mu}{\lambda+2\mu}(V_{\alpha j,\beta}^*+V_{\beta j,\epsilon}^*)\right]n_\beta \tag{27}$$

where n_β is a component of the unit outward normal vector, **n** to the boundary. The first through fourth terms in Eq. (26), represent, respectively, contributions from the product of the traction vector and the fundamental solution tensor, the product of the displacement vector and the derivative of the fundamental solution tensor, the product of the derivative of temperature and the fundamental solution tensor, and the product of the temperature and the derivative of the fundamental solution tensor. The fundamental solution tensor must satisfy the following differential equation:

$$\theta_{ij}V_{jk}^*=-\delta_{ik}\delta(x-y) \tag{28}$$

where $\delta(x-y)$ is the delta function and θ_{ij} is the adjoint operator of Θ_{ij} in Eq. (22). This is given by

$$\theta_{ij}=\begin{bmatrix} \left(\frac{\mu}{\lambda+2\mu}\right)\Delta+\left(\frac{\lambda+\mu}{\lambda+2\mu}\right)D_1^2-s^2 & \left(\frac{\lambda+\mu}{\lambda+2\mu}\right)D_1D_2 & -sCD_1 \\ \left(\frac{\lambda+\mu}{\lambda+2\mu}\right)D_1D_2 & \left(\frac{\mu}{\lambda+2\mu}\right)\Delta+\left(\frac{\lambda+\mu}{\lambda+2\mu}\right)D_2^2-s^2 & -sCD_2 \\ -D_1 & -D_2 & \Delta-s \end{bmatrix}. \tag{29}$$

3 Fundamental Solution

In order to construct the fundamental solution, we express the fundamental solution tensor, V_{ij}^*, from Eq. (28), in the following potential representation by using the transposed co-factor operator μ_{ij} of θ_{ij} and scalar function Φ^* ([29]):

$$V_{ij}^*(x,y,s)=\mu_{ij}\Phi^*(x,y,s). \tag{30}$$

After substitution of Eq. (30) into Eq. (28), we get the following differential equation:

$$\Lambda\Phi^*=-\delta(x-y) \tag{31}$$

where

$$\Lambda=\det(\theta_{ij})=\frac{\mu}{\lambda+2\mu}(\Delta-h_1^2)(\Delta-h_2^2)(\Delta-h_3^2) \tag{32}$$

and the h_i^2 are the solutions of

$$h_1^2=\left(\frac{\lambda+2\mu}{\mu}\right)s^2$$

$$h_2^2+h_3^2=s^2+s(1+C)$$

$$h_2^2h_3^2=s^3. \tag{33}$$

Note that h_1 is the longitudinal wave velocity, h_2 is the shear wave velocity, h_3 is the rotational wave velocity, and

$$\Phi^*=\left(\frac{\lambda+2\mu}{2\pi\mu}\right)\left[\frac{K_0(h_1r)}{(h_2^2-h_1^2)(h_3^2-h_1^2)}+\frac{K_0(h_2r)}{(h_3^2-h_2^2)(h_1^2-h_2^2)}+\frac{K_0(h_3r)}{(h_1^2-h_3^2)(h_2^2-h_3^2)}\right]. \tag{34}$$

Here K_0, K_1, and K_2 are modified Bessel functions of the second kind of zero, first, and second-order, respectively ([30]). The components of the fundamental solution tensor, V_{ij}^*, for the two-dimensional domain in Fig. 1, are as follows:

$$V_{\epsilon\phi}^*=\sum_{k=1}^{3}(\psi_k(r)\delta_{\epsilon\phi}-\kappa_k r_{,\epsilon}r_{,\phi}) \quad (\epsilon,\phi=1,2) \tag{35}$$

$$V_{3\epsilon}^*=\sum_{k=1}^{3}\xi_k(r)r_{,\epsilon} \tag{36}$$

$$V_{\epsilon 3}^*=\sum_{k=1}^{3}\xi_k(r)r_{,\epsilon} \tag{37}$$

$$V_{33}^*=\sum_{k=1}^{3}\zeta_k(r) \tag{38}$$

where

$$\psi_k(r)=\frac{W_k}{2\pi}\left[(h_k^2-m_2)(h_k^2-m_1)+\left(\frac{\lambda+\mu}{\mu}\right)\times\left\{h_k^2-m_1\left(1+\left(\frac{\lambda+2\mu}{\lambda+\mu}\right)C\right)\right\}h_k^2\right]K_0(h_kr)+\frac{W_k(\lambda+\mu)}{2\pi\mu}\left[h_k^2-m_1\left(1+\left(\frac{\lambda+2\mu}{\lambda+\mu}\right)C\right)\right]\frac{h_k}{r}K_1(h_kr) \tag{39}$$

$$\kappa_k(r)=\frac{W_k(\lambda+\mu)}{2\pi\mu}\left[h_k^2-m_1\left(1+\left(\frac{\lambda+2\mu}{\lambda+\mu}\right)C\right)\right]K_2(h_kr) \tag{40}$$

$$\xi_k(r)=\frac{W_k}{2\pi}\left[h_k^2-\left(\frac{\lambda+2\mu}{\mu}\right)s^2\right]h_kK_1(h_kr) \tag{41}$$

$$\xi_k(r)=\frac{W_k}{2\pi}\left[h_k^2-\left(\frac{\lambda+2\mu}{\mu}\right)s^2\right]h_ksCK_1(h_kr) \tag{42}$$

$$\zeta_k(r)=\frac{W_k}{2\pi}\left[h_k^2-\left(\frac{\lambda+2\mu}{\mu}\right)s^2\right](h_k^2-s^2)K_0(h_kr) \tag{43}$$

and

$$r=\|x-y\|; \quad m_1=s; \quad m_2=\left(\frac{\lambda+2\mu}{\mu}\right)s^2;$$

$$W_i=\frac{-1}{(h_i^2-h_j^2)(h_k^2-h_i^2)}$$

$$(i=1,2,3 \quad j=2,3,1 \quad k=3,2,1). \tag{44}$$

In order to solve Eq. (26), the standard boundary element procedure may be applied ([31]). When transformed numerical solutions are specified, transient solutions may be obtained using an appropriate numerical inversion technique. In this paper, the method presented by Durbin [32] is adopted for numerical inversion in the time domain.

4 Results and Discussion

We examine the influence of thermoelastic coupling due to the applied surface temperature in Eq. (5). We consider two values of pulse duration, t_w, for pulsed laser heating of a pure aluminum layer where $E=6\times10^{10}$ Pa, $\nu=0.33$, $\rho=2650$ kg/m^3, and $\alpha=9.4\times10^{-5}$ m^2/sec (extrapolated close to the melting temperature from [33]). Hence, the velocity of the longitudinal wave in aluminum is $C_1=5792$ m/sec. We must ensure that the layer criterion given by Eq. (11) is maintained. For the shorter of the two pulse durations, we choose $t_w=8.5\times10^{-12}$ sec or $\hat{t}_w=3$. We set $l=10\beta$: this gives $l=0.16\times10^{-6}$ m, $L=10$ and

$$\frac{l^2}{\alpha t_w}=33. \tag{45}$$

The time for the longitudinal wave to travel across the layer thickness is 27.6×10^{-12} sec. Hence, the pulse duration is less than three times the time required for the thermoelastic wave to propagate from the free surface to the bonded surface of the layer. Equation (15) implies that a longitudinal wave will be generated

within the layer. For the longer of the two pulse durations, we choose $t_w = 85\times10^{-9}$ sec (which is four-orders-of-magnitude longer than the shorter pulse duration) or $\hat{t}_w = 30{,}000$. We set $l = 200\beta$: this gives $l = 3.24\times10^{-6}$ m, $L = 200$ and

$$\frac{l^2}{\alpha t_w} = 1.34. \tag{46}$$

The time for the longitudinal wave to travel across the layer thickness in this case is 56×10^{-9} sec. Hence, the pulse duration is 150 times longer than the time it takes for the longitudinal wave to travel from the free surface to the bonded surface. Equation (15) implies that no longitudinal wave will be generated within the layer. The picosecond pulse duration is common to mode-locked Nd:YAG lasers ([24]), whereas the nanosecond duration is common to Q-switched Nd:YAG lasers ([25]). Note that both of these pulse durations are considerably shorter than that addressed by Kim et al. [16]. In that paper, the pulse duration was of the order of 10^{-4} sec. All quantities reported in the following figures are in terms of dimensionless variables defined by Eqs. (12): We have dropped the hat notation for convenience.

The temperature versus time profile due to the shorter of the two pulse trains is shown in Fig. 2. This consists of four pulses at $x = y = 0$. The dimensionless peak power per pulse delivered to the surface, P_p, has been arbitrarily set to one. Peak power in the pulse train is delivered to the surface of the layer at times t_p^N, where

$$t_p^N = \frac{3(2N-1)}{2} \tag{47}$$

and N is the pulse number (for example, $N = 3$ for the pulse that is activated at $t = 6$). The average power, P_a, is

$$P_a = \frac{P_p}{t_w}\int_0^{t_w}\left|\sin\left(\frac{2\pi t}{t_w}\right)\right| dt = \frac{2P_p}{\pi} = 0.64. \tag{48}$$

Note that we selected four pulses since this was deemed to be adequate for purposes of discerning differences between the stress and displacement fields due to the coupled and uncoupled solutions.

Figure 3 compares the temperature evolution, at specific depths within the layer, due to the uncoupled ($C = 0$) and coupled ($C = 0.3$) formulations for $0 \leq t \leq 12$. The uppermost curves (solid and dotted) show the temperature evolution at $x = 0.3\ L$, $y = 0.2\ L$. The lower two curves (dashed and dot-dashed) show the temperature evolution at $x = 0.6\ L$, $y = 0.2\ L$. The dimensionless quantities were selected in such a way that the thermoelastic shock front at $x = 0.3\ L$ occurs at time $t = 3$, and that at $x = 0.6\ L$ occurs at time

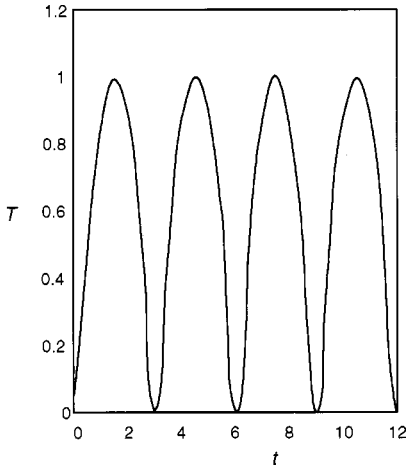


Fig. 2 Thermal load due to laser pulse train at $x = y = 0$

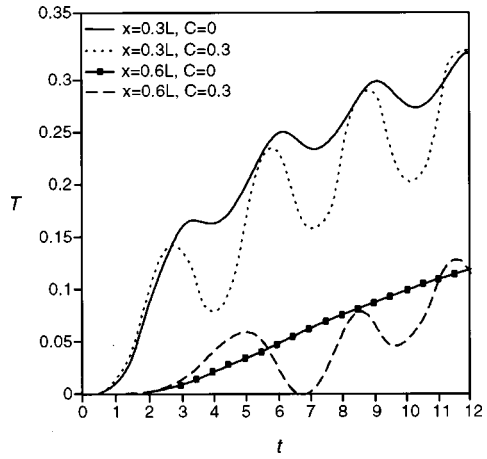


Fig. 3 T versus t, at $y = 0.2\ L$, for $t_w = 3$

$t = 6$. It is immediately apparent that the mean value of temperature at $x = 0.6\ L$ is much lower than that at $x = 0.3\ L$. The curves corresponding to $x = 0.3\ L$, $y = 0.2\ L$ show four crests in the temperature distribution. A comparison of Figs. 2 and 3 shows that each crest due to the uncoupled solution appears near to the time when the corresponding laser pulse that produced it is completely deactivated (i.e., $t = 3$, 6, 9, 12). However, each crest due to the coupled solution occurs during the time when the corresponding laser pulse that produced it is in the process of deactivation. The coupled solution predicts more extreme temperature fluctuations in the layer since the crest to trough deviations in the temperature profile between successive pulses exceed those predicted by the uncoupled formulation. The curve corresponding to the coupled solution at $x = 0.6\ L$, $y = 0.2\ L$ shows three temperature crests, whereas that corresponding to the uncoupled solution shows a continuous increase in temperature. It is interesting to note that for $C = 0.3$, the maximum temperature due to the third pulse occurs just before $t = 12$. As seen in Fig. 2, this corresponds to the time when the fourth pulse is deactivated. Clearly, the coupled solution is more sensitive to the temporal pulse profile than the uncoupled solution. Also, the point $x = 0.6\ L$, $y = 0.2\ L$ is deep enough within the layer so that the temperature has not significantly increased there by the time the fourth pulse is delivered to the surface. As additional pulses beyond the four shown in Fig. 2 are delivered to the layer surface, the insulated boundary will play a more important role in the temperature profile since the thermoelastic shock front will be reflected from this boundary. The temperatures will begin to increase, with the coupled solution continuing to show greater crest-to-trough deviations than those predicted by the uncoupled solution. Eventually, the surface will melt and/or ablate, resulting in different modes of thermal transport within the layer. At this point, however, the present model would have to be modified to account for the propagation of a phase change front. As a final note on Fig. 3, the shape of the curves for the coupled solution may in fact be subject to further scrutiny since thermal transport at such short pulse widths is a wave phenomena. Careful examination of this issue would require solution of the present problem with the generalized dynamical theory of thermoelasticity wherein both the thermal and mechanical disturbances are characterized by finite wave speeds ([22]).

Figures 4 and 5 show the evolution of the axial displacement, u, and the lateral displacement, v (i.e., parallel to the surface $x = 0$), respectively, for $0 \leq t \leq 12$. In Fig. 4, the curves are grouped according to the curves in Fig. 3 and compare differences between the coupled and uncoupled solutions. At $x = 0.3\ L$, $y = 0.2\ L$ (solid and dotted curves), the material response to the first pulse is an inward ''bulging'' (up to $t = 3.5$) since a positive displacement, u, follows the positive x-axis into the layer. This suggests that the axial stress is compressive. The displacement, u, changes sign at

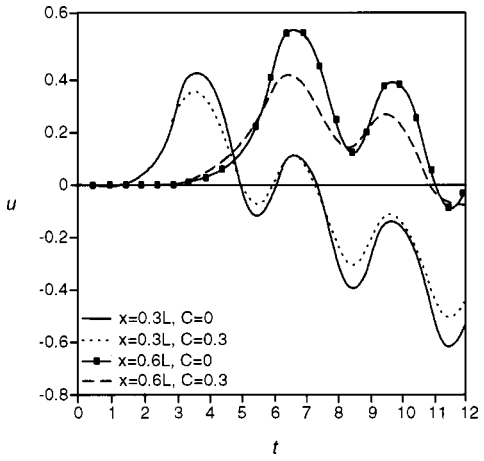


Fig. 4 ***u*** **versus** ***t*****, at** $y=0.2\,L$**, for** $t_w=3$

$t=4.6$, at which point the surface bulges outward. This suggests that the axial stress becomes tensile. As the material responds to the second pulse between $5.3\leqslant t\leqslant 8.3$, the surface again bulges inward (but only by a very small amount) and then continuously bulges outward beyond $t=7.2$. A similar behavior is noted for the two curves corresponding to $x=0.6\ L$, $y=0.2\ L$, except that the amplitude of the inward bulge at the first crest exceeds that at $x=0.3\ L$, $y=0.2\ L$. The coupled solution noticeably differs from the uncoupled solution at $x=0.6\ L$, $y=0.2\ L$ in that it tends to predict a less extreme variation in the axial displacement. Clearly, the coupled solution shows little difference between the uncoupled solution at $x=0.3\ L$, $y=0.2\ L$, and hence longitudinal wave propagation has little effect there. However, the same is not true for the deeper lying point at $x=0.6\ L$, $y=0.2\ L$ which is closer to the bonded interface. The reversal of sign in the curves of Fig. 4 results from layer deformation in response to the very short pulse duration. This behavior was not predicted Kim et al. [16] since the axial displacement was found to be compressive for all time. This prediction is consistent with experiments by Appollonov et al. [34], in which a continuous wave CO_2 laser beam was used to demonstrate outward bulging of a sapphire surface. Figure 5 shows that the lateral displacement is always outward from the point $x=y=0$. Note, however, that the displacement at $x=0.3\ L$, $y=0.2\ L$ is always greater than that at $x=0.6\ L$, $y=0.2\ L$: This is independent of the coupling constant, C. This reveals a behavior that differs from that shown in Fig. 4 due, for the most part, to the absence of a rigid boundary constraint along the y-direction. The curves corresponding to $x=0.3\,L, y=0.2\,L$ show only the tops of the crests that result from the pulses in Fig. 2. The temporal variation of the pulse train profile is thus smeared out (for the most part) in the lateral displacement. At $x=0.6\,L, y=0.2\,L$, however, the response to the first two pulses is slightly better developed for the coupled solution, although any hint of the pulse shape in the corresponding uncoupled solution is smeared out. At both points, the coupled solution predicts smaller displacements than those predicted by the uncoupled solution. This effect is most pronounced at longer times in the process.

Figures 6–8 show the evolution of the axial stress, σ_{xx}, lateral stress, σ_{yy}, and shear stress, σ_{xy}, respectively, over the time frame considered in Figs. 4–5. Each figure compares the coupled and uncoupled stresses at $x=0.3\,L$, $y=0.2\,L$ and $x=0.6\,L$, $y=0.2\,L$.

In Fig. 6, the minima in the two curves for which $C=0$ fall at the times when the longitudinal wave passes the points $x=0.3\,L$ and $0.6\,L$ (longitudinal wave propagation has no meaning for the uncoupled formulation). The same is not true for the two curves corresponding to $C=0.3$. The most extreme compressive stresses occur just before the longitudinal wave passes $x=0.3\,L$ and $0.6\,L$. At $x=0.3\,L, y=0.2\,L$, the material is in a state of tension for $3\leqslant t\leqslant 6$ since the wave at $t=3$ has passed. The tensile stress peaks at $t=4.8$, and then decreases as the second wave approaches. This behavior is consistent with the behavior of the axial displacement shown in Fig. 4. Similar observations can be made about the curves corresponding to $x=0.3\,L, y=0.2\,L$. At both points within the layer, the coupled solution predicts an axial stress variation

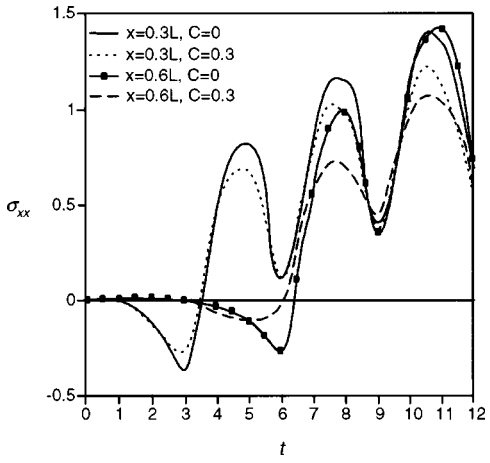


Fig. 6 σ_{xx} **versus** ***t*****, at** $y=0.2\,L$**, for** $t_w=3$

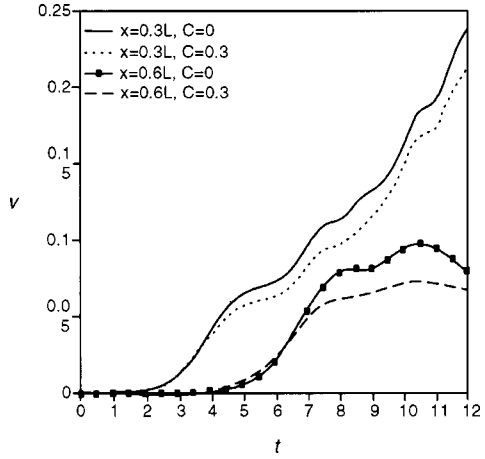


Fig. 5 ***v*** **versus** ***t*****, at** $y=0.2\,L$**, for** $t_w=3$

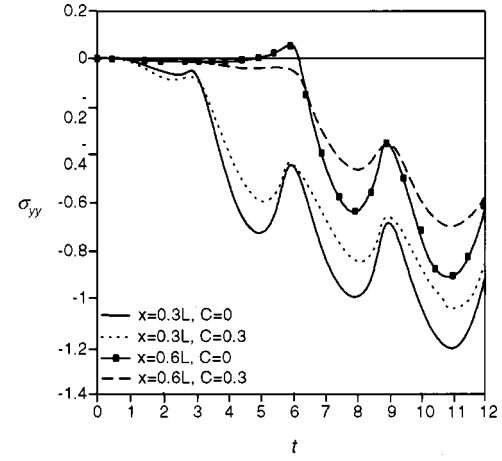


Fig. 7 σ_{yy} **versus** ***t*****, at** $y=0.2\,L$**, for** $t_w=3$

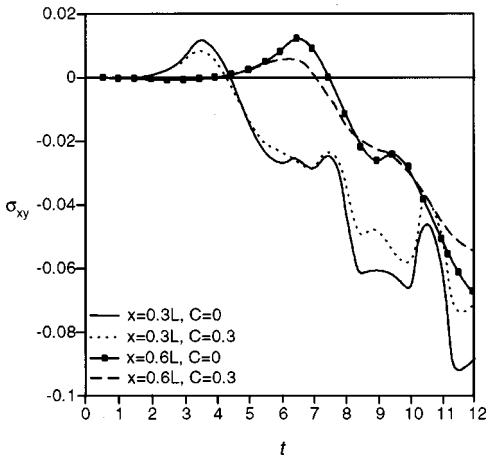


Fig. 8 σ_{xy} **versus** t, **at** $y=0.2\,L$, **for** $t_w=3$

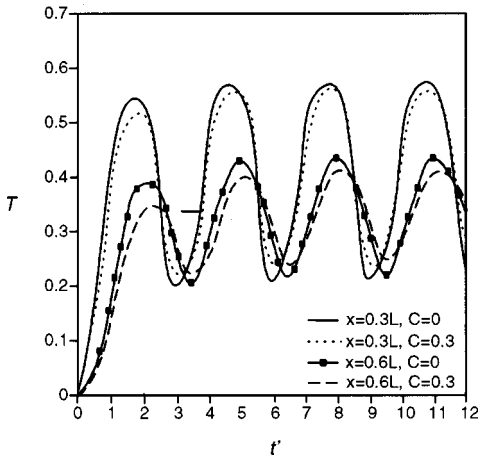


Fig. 9 T **versus** t', **at** $y=0.2$, **for** $t_w=30{,}000$

that is less extreme than that predicted by the uncoupled solution once the shock front has passed. Of the cases considered in Fig. 6, the most extreme tensile stress was found to occur at $x=0.6\,L$. Note that the wavefronts are not perpendicular to the time axis since heat is not instantaneously released into the material (this would require a delta function in the surface temperature boundary condition (Eq. (5)).

Figure 7 shows that the lateral stress, σ_{yy}, is compressive (other than a small excursion into the tensile region at $t=6$) from the onset of laser irradiation. The origin of the compressive lateral stress is the sudden expansion of heated material against cooler, albeit deformable substrate material that surrounds the heated region. Note that the coupled solution predicts that the lateral stress is less compressive than that predicted by the uncoupled solution. Of the cases considered in Fig. 7, the most extreme compressive stress was found to occur at $x=0.3\,L$. Note that the axial stress and displacement profiles (Figs. 4 and 6) tend to vary according to the pulse train profile. However, this is not the case with Figs. 5 and 7 for the lateral stress and displacement profiles: While the former denotes very little of the pulse train profile, the latter tends to display the characteristic rise and decay of the pulses delivered to the surface.

Figure 8 shows that the shear stress, σ_{xy}, crests at $x=0.3\,L$, $y=0.2\,L$, and $x=0.6\,L$, $y=0.2\,L$, respectively, a short time after the first and second pulses have been deactivated. The sign of the shear stress thus changes some time after the passage of the thermoelastic shock front. Note that the shear stress distribution is necessarily complicated since it is influenced by the variations of σ_{xx} and σ_{yy} as well as the bonded interface at $x=L$. For example, the shear stress at $x=0.6\,L$ is nearly constant between the crests due to the second and third pulses, i.e., $8\leqslant t\leqslant 10$. The coupled solution predicts smaller values of the shear stress than those predicted by the uncoupled solution. This difference becomes more apparent with increasing time.

Figures 9–13 explore the response of the layer material to the $t_w=30{,}000$ pulse duration at $x=0.3\,L$, $y=0.2\,L$, and $x=0.6\,L$, $y=0.2\,L$. For this purpose, we define a new dimensionless time variable, t', where

$$t'=10\,000t. \tag{49}$$

Each figure is considered over the $0\leqslant t'\leqslant 12$ time interval.

Figure 9 compares the temperature evolution, at depths $x=0.3\,L$ and $x=0.6\,L$, due to the uncoupled ($C=0$) and coupled ($C=0.3$) formulations. Note that the differences between the coupled and uncoupled solutions are substantially smaller compared with those due to the ultrafast pulse train displayed in Fig. 3. At both depths, the coupled solution predicts a slightly less extreme temperature profile. At $x=0.3\,L$, the temperature peaks close to the time when the corresponding pulses achieve peak power. At $x=0.6\,L$, the temperature achieves its extreme values midway between the time when the corresponding pulse achieves peak power and when it is deactivated. It is interesting to note that while the temperature maxima differ at the two depths, the temperature minima are roughly equivalent. This was clearly not the

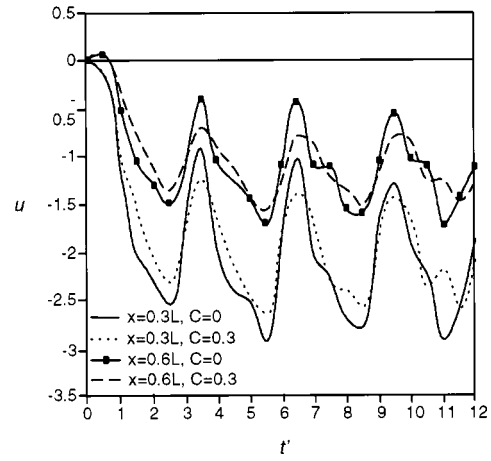


Fig. 10 u **versus** t', **at** $y=0.2$, **for** $t_w=30{,}000$

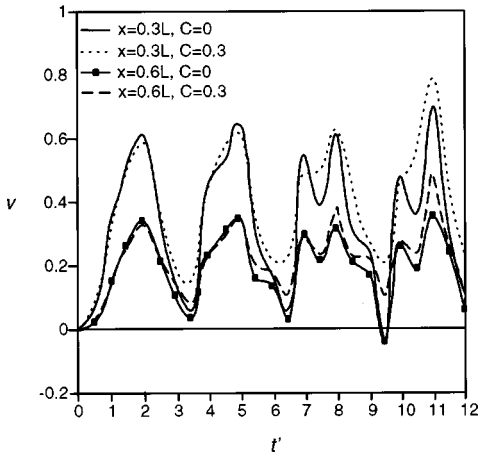


Fig. 11 v **versus** t', **at** $y=0.2\,L$, **for** $t_w=30{,}000$

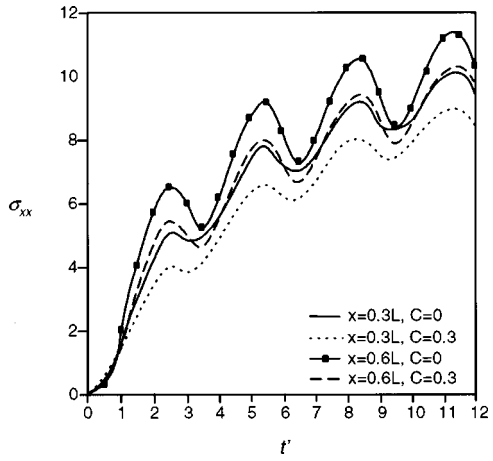


Fig. 12 σ_{xx} versus t', at $y=0.2\,L$, for $t_w=30{,}000$

case in Fig. 3. Based upon the results in Fig. 9, it is clear that thermoelastic coupling is much more important for the $t_w=3$ pulse duration.

Figures 10 and 11 show, respectively, the evolution of the axial and lateral displacement fields. In Fig. 10, the four trough regions occur in response to heat absorbed from the four pulses delivered to the free surface of the layer. Unlike the inward and outward bulging behavior due to the $t_w=3$ pulse duration (see Fig. 4), the layer surface only bulges outward in response to the $t_w=30{,}000$ pulse duration. The coupled solution predicts that the maximum outward displacement of the layer surface is less than that predicted by the uncoupled formulation: The extreme values occur at $t'=2.7$, 5.6, 8.5, and 11. However, as the displacement decreases (in response to deactivation of the corresponding pulse), thereby implying a reduction in the outward bulging, the coupled solution tends to cross-over the uncoupled solution. When the amplitude of the bulge drops to its smallest value (i.e., at the least negative values of displacement), the uncoupled solution predicts the smallest bulge amplitude. Hence, the extreme shapes of the bulged material are predicted to be smaller by the coupled formulation. Figure 11 shows that the coupled and uncoupled solutions for the lateral displacement differ insignificantly in response to the first two pulses. Extreme values during this time occur at $t'=2,5$. Like that predicted for the $t_w=3$ pulse duration in Fig. 5, the lateral displacement is outward from the center of the layer at $y=0$. As the material responds to the third and fourth pulses, the oscillatory behavior of the lateral displacement becomes more

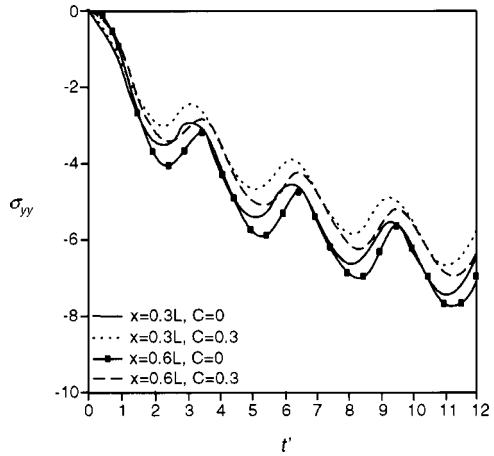


Fig. 13 σ_{yy} versus t', at $y=0.2\,L$, for $t_w=30{,}000$

complicated. The uncoupled solution predicts that the crests in the lateral displacement ''split,'' while this behavior tends to be smeared out by the coupled solution. The greatest differences between the two solutions occur at $t'=7.5,10.5$. The splitting of the third and fourth crests of the lateral displacement profile illustrates a fundamental difference between the material response to the nanosecond and picosecond pulse trains. For the nanosecond pulse train, the material tends to develop a vibration or elastic ''spring'' along the y-direction due to substantial expansion against the colder regions that surround the region exposed by the beam. Some indication of this is also seen in Fig. 10. For the picosecond pulse train, however, this behavior is not observed (see Fig. 5) since the effect of thermoelastic coupling is much more important and the average value of the displacement at both depths is considerably smaller than that for the longer pulse duration. The time scale for heat absorption and diffusion is considerably shorter for the picosecond pulse train and hence the material does not displace along the y-direction to the extremes that occur during exposure to the nanosecond pulse train.

Figures 12 and 13 show the evolution of the axial stress, σ_{xx}, and lateral stress, σ_{yy}, respectively, over the same time frame considered in Figs. 10 and 11. Each figure compares the coupled and uncoupled stresses at $x=0.3\,L$, $y=0.2\,L$ and $x=0.6\,L$, $y=0.2\,L$. Note that Figs. 10 and 12 show oscillatory behavior in the displacement and stress components along the layer axis. A similar observation can be made relative to Figs. 11 and 13. Maximum tensile stresses along the layer axis occur at $t'=2.3$, 5.3, 8.3, and 11.3 in Fig. 12. Maximum compressive stresses parallel to the free surface of the layer also occur at these times according to Fig. 13. The key difference between Fig. 6 and Fig. 11 is that the changeover in stress from compressive to tensile along the axial direction due to the $t_w=3$ pulse duration is not predicted for the $t_w=30{,}000$ pulse duration. The stress due to the $t_w=30{,}000$ pulse duration is always tensile along the layer axis. Note that the coupled solution predicts a smaller tensile stress than is predicted by the uncoupled solution in both Figs. 6 and 11. For the lateral stress, σ_{yy}, the uncoupled formulation shows more extreme compressive values than are predicted by the coupled formulation.

In the present model, we have assumed that the material does not melt or evaporate during the laser heating process. This of course has limited our discussion to heating without phase change and the thermal stress field that results in the material. Since pulsed lasers are commonly use to scribe and evaporatively cut materials in many industrial processes ([35,36]), some comment on the extent to which the temporal pulse widths considered in the present paper might result in ablation of a representative material are worthwhile. For this purpose, we choose to examine pure aluminum metal, and will therefore neglect the fact that all aluminum surfaces consist of an amorphous natural oxide layer which is 3–6 nm deep ([37]). An estimate of ablation depth, d, (which is useful for the purpose of preliminary process design) following delivery of a single pulse of temporal width, t_w, is given by the following lumped heat capacity expression ([28]):

$$d=\frac{t_w f}{\rho[c_p(T_v-T_i)+L_v]}. \tag{50}$$

We assume that all quantities in Eq. (50) are dimensional. Note that f is the absorbed power density, c_p is the heat capacity, T_v is the vaporization temperature, T_i is the initial temperature of the material, and L_v is its vaporization temperature. In writing Eq. (50), we have assumed that all material properties are temperature-independent, the laser pulses do not couple to ejected material or beam plasma (and hence nonlinear optical processes are neglected ([38])), and ablated material is removed instantaneously with no melting and subsequent fluid flow due to thermocapillary effects ([39]). The absorbed power density, f, is related to the peak power per pulse, P_p, via

$$f = \epsilon F_p = \frac{\epsilon P_p}{A} \tag{51}$$

wherein ϵ is the absorption coefficient (or fraction of the incident radiation that is absorbed by the surface), F_p is the peak power density per pulse, and A is the beam exposure area on the surface. Note that f contains information about the optical properties of the surface material as well as the surface microgeometry (which may adversely affect beam absorption due to scattering). Aluminum metal will, if polished to a mirror finish, reflect the majority of incident light energy at the near infrared wavelength of 1.0 μm emitted by pulsed Nd:YAG lasers. Under these conditions, we assume that only five percent of the incident light energy from a single pulse is absorbed in the surface plane of the metal. We further consider the case where each pulse carries a peak power of 20,000 W and irradiates a 0.5 mm^2 area. For aluminum initially at room temperature, we have $c_p = 900$ J/kg°C, $T_v = 2450$°C, $T_i = 25$°C, and $L_v = 9492$ J/kg ([28]). For $t_w = 85 \times 10^{-9}$ sec, Eq. (50) gives $d = 29$ nm, and hence after four pulses, the total ablation depth would be 117 nm. For the shorter pulse duration of 8.5×10^{-12} sec, Eq. (50) gives $d = 0.002$ nm. Clearly, ablation will be more of a concern with the longer of the two pulse trains considered in this paper (provided of course that power losses due to nonlinear optical phenomena, such as beam coupling to the plasma, and melting are ignored). For more detailed information on the relationship of pulse power density to phase change processes resulting from pulsed laser irradiation of metals, the interested reader is referred to Hector and Sheu [40].

5 Concluding Remarks

Temperature, displacement, and thermal stress fields due to heat absorption from pulses with temporal and spatial profiles that emulate pulse trains emitted by common materials processing lasers were calculated with a boundary element formulation. The temporal profile of the pulse train was modeled with a rectified sine wave, while the spatial profile was modeled as a Gaussian strip source. The thermal and mechanical fields predicted under the assumption of thermomechanical coupling were compared with those calculated assuming no coupling. Both nanosecond and picosecond pulse durations of 85×10^{-9} sec and 8.5×10^{-12} sec, respectively, were considered. In the case of the picosecond pulse duration, the layer thickness was adjusted so that the pulse duration was less than three times the time required for the thermoelastic wave to propagate from the free surface to the bonded surface of the layer. For the case where the thermal and mechanical fields are coupled, this led to the propagation of a longitudinal wavefront along the axis of the layer due to the dilatational effect of heat absorption from the laser pulse train. This front leaves a compressive stress field in its wake. On the other hand, the layer thickness associated with the 85×10^{-9} sec pulse duration was adjusted so that the pulse duration was 150 times longer than the time required for the longitudinal wave to travel from the free surface to the bonded surface. Examination of the stress and displacement fields demonstrated that differences between the coupled and uncoupled solutions were significant for the 8.5×10^{-12} sec. However, the uncoupled solution was sufficient for the 85×10^{-9} sec. An additional effect, which is best described as a thermoelastic springing of the material, was noted in the splitting of the crests associated with the lateral displacement field due to the 85×10^{-9} sec pulse duration. This was not observed in the analogous displacement profile due to the picosecond pulse train.

Current modeling efforts associated with thermomechanical effects in laser-irradiated materials are focused on examination of the conditions under which a thick layer will delaminate from a deformable substrate. This is an important design issue for many industrial applications in which a coated mechanical component is exposed to repetitive thermal loading from a pulsed laser. For this purpose, the present model is being reformulated to account for an elastic, conducting substrate. In addition, concepts from solid state physics are being incorporated into an extension of the present model which includes the thermal inertia term for ultrafast laser heating and related nonlinear optical effects. This will establish a link between electronic conduction due to heat absorption over very short time intervals and the resulting material deformation.

References

[1] Maiman, T. H., 1960, ''Stimulated Optical Radiation in Ruby,'' Nature (London), **187**, p. 493.
[2] Yamawaki, H., and Saito, T., 1992, ''Computer Simulation of Laser-Generated Elastic Waves in Solids,'' Nondestr. Test. Eval., **7**, pp. 165–177.
[3] Qui, T., Tien, C.-L., Shannon, M. A., and Russo, R. E., 1994, ''Thermal and Mechanical Responses of Gold Films During Nanosecond Laser-Pulse Heating,'' Exp. Heat Transfer, **7**, pp. 175–188.
[4] Zhou, Y. C., Duan, Z. P., and Yang, Q. B., 1998, ''Failure of SiC Particulate-Reinforced Metal Matrix Composites Induced by Laser Thermal Shock,'' Metall. Mater. Trans. A, **29A**, pp. 685–692.
[5] Mortia, N., Watanabe, T., and Yoshida, Y., 1991, ''Crack-Free Processing of Hot-Pressed Silicon Nitride Ceramics Using Pulsed YAG Laser,'' JSME Int. J., Ser. III, **34**, pp. 149–153.
[6] Kostrubiec, F., and Walczak, M., 1994, ''Thermal Stresses in the Tungsten and Molybdenum Surface Layer Following Laser Treatment,'' J. Mater. Sci. Lett., **13**, pp. 34–36.
[7] Masse, J.-E., and Barreau, G., 1993, ''Residual Stresses in Steel by Laser Shock Treatment,'' C. R. Acad. Sci., Ser. II: Mec., Phys., Chim., Sci. Terre Univers, **317**, pp. 1015–1018.
[8] Ashby, M. F., and Easterling, K. E., 1984, ''The Transformation Hardening of Steel Surfaces by Laser Beams—I. Hypo-eutectoid Steels,'' Acta Metall., **32**, pp. 1935–1948.
[9] Welsh, L. P., Tuchman, J. A., and Herman, J. P., 1988, ''The Importance of Thermal Stresses and Strains Induced in Laser Processing with Focused Gaussian Beams,'' J. Appl. Phys., **64**, pp. 6274–6286.
[10] Uglov, A. A., Kulik, A. N., Makorkin, I. N., and Senik, A. P., 1994, ''On the Calculation of Metallic Cylinder Thermal Stressed State Under Periodic Pulsed Laser Heating,'' Fiz. Khim. Obrab. Mater., **4**, pp. 12–18 (in Russian).
[11] Germanovich, L. N., Kill, I. D., and Tsodokova, N. S., 1988, ''Thermoelastic Stresses in a Half-Space Heated by Concentrated Energy Flux,'' Appl. Math. Mech., **52**, pp. 525–533 (in Russian).
[12] Suh, C. S., and Burger, C. P., 1998, ''Thermoelastic Modeling of Laser-Induced Stress Waves in Plates,'' J. Therm. Stresses, **21**, pp. 829–847.
[13] Cohen, S. S., Bernstein, J. B., and Wyatt, P. W., 1992, ''The Effect of Multiple Laser Pulses on Damage to Thin Metallic Films,'' J. Appl. Phys., **71**, pp. 630–637.
[14] Fesenko, V. M., 1995, ''Thermal-Stressed State of a Tetragonal Symmetry Crystal Under the Laser Radiation Effect,'' Inzh.-Fiz. Zh., **68**, pp. 160–161 (in Russian).
[15] Hector, Jr., L. G., and Hetnarski, R. B., 1996, ''Thermal Stresses Due to a Laser Pulse: Elastic Solution,'' ASME J. Appl. Mech., **63**, pp. 38–46.
[16] Kim, W. S., Hector, Jr., L. G., and Hetnarski, R. B., 1997, ''Thermoelastic Stresses in a Bonded Layer Due to Repetitively Pulsed Laser Radiation,'' Acta Mech., **125**, pp. 107–128.
[17] Tehrani, P. H., and Eslami, M. R., 1998, ''Two-Dimensional Time Harmonic Dynamic Coupled Thermoelasticity Analysis by BEM Formulation,'' Eng. Anal. Boundary Elem., **22**, pp. 245–250.
[18] Nowacki, W., 1986, *Thermoelasticity*, 2nd ed., Pergamon, New York.
[19] Nowacki, W., 1975, *Dynamics Problems of Thermoelasticity*, P. H. Francis and R. B. Hetnarski, eds., Noordhoff International Publishers, Leyden.
[20] Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd ed., McGraw-Hill, New York.
[21] Hector, Jr., L. G., Kim, W.-S., and Ozisik, M. N., 1992, ''Propagation and Reflection of Thermal Waves in a Finite Medium Due to Axisymmetric Surface Sources,'' Int. J. Heat Mass Transf., **35**, pp. 897–912.
[22] Chandrasekaraiah, D. S., 1989, ''Thermoelasticity With Second Sound: A Review,'' Appl. Mech. Rev., **39**, pp. 355–376.
[23] Sparks, M., 1976, ''Theory of Laser Heating of Solids,'' J. Appl. Phys., **47**, pp. 837–849.
[24] Hector, Jr., L. G., Kim, W.-S., and Osizik, M. N., 1992, ''Hyperbolic Heat Conduction due to a Mode Locked Laser Pulse Train,'' Int. J. Eng. Sci., **30**, pp. 1731–1744.
[25] Koechner, W., 1991, *Solid State Laser Engineering*, 2nd ed., Springer-Verlag, New York.
[26] Hecht, J., 1992, *The Laser Guidebook*, 2nd ed., McGraw-Hill, New York.
[27] Lim, G. C., 1982, ''Measurement of the Temporal and Spatial Power Distribution of a High-Power CO_2 Laser Beam,'' Opt. Laser Technol., **14**, pp. 149–153.
[28] Steen, W. M., 1991, *Laser Material Processing*, Springer-Verlag, New York.
[29] Dominguez, J., 1993, *Boundary Elements in Dynamics*, Computational Mech., Boston.
[30] Watson, G. N., 1966, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, New York.
[31] Brebbia, C. A., Telles, J. C. F., and Wrobel, L. C., 1984, *Boundary Element Techniques*, Springer-Verlag, New York.
[32] Durbin, F., 1974, ''Numerical Inversion of Laplace Transforms: An Efficient Improvement to Dubner and Abate's Method,'' Comput. J., **17**, pp. 371–376.

[33] Incropera, F. P., and DeWitt, D. P., 1981, *Fundamentals of Heat Transfer*, John Wiley and Sons, New York.
[34] Apollonov, V. V., Barchukhov, A. I., Prokhorov, A. M., 1972, ''Thermoelastic Deformations of a Solid Surface by a Laser Beam,'' JETP Lett., **15**, pp. 172–174.
[35] Nonhof, C. J., 1988, *Materials Processing With Nd with Nd-Lasers*, Electrochemical Publications Ltd., Ayr, Scotland.
[36] Powell, J., 1993, CO_2 *Laser Cutting*, Springer-Verlag, New York.
[37] Wefers, K., and Misra, C., 1987, ''Oxides and Hydroxides of Aluminum,'' Alcoa Technical Paper No. 19, Revised. Alcoa Technical Center, Alcoa Center, PA.
[38] Krokhin, O., 1972, ''Generation of High Temperature Vapors and Plasmas by Laser Radiation,'' *Laser Handbook*, F. T. Arrechi, and E. O. Schulz-Dubois, eds., North-Holland, Amsterdam, Chapter 2.
[39] Zehr, R. L., 1991, ''Thermocapillary Convection in Laser Melted Pools During Material Processing,'' Ph.D. thesis, Mechanical Engineering Department, University of Illinois at Urbana-Champaign, University of Microfilms International, Ann Arbor, MI.
[40] Hector, Jr., L. G., and Sheu, S., 1993, ''Focused Energy Beam Work Roll Surface Texturing Science and Technology,'' J. Mater. Process. Manuf. Sci., **2**, pp. 63–117.

L. M. Brock
Fellow ASME
Department of Mechanical Engineering,
University of Kentucky,
Lexington, KY 40506

H. G. Georgiadis
Mem. ASME
Mechanics Division,
National Technical University of Athens,
Zographou 15773, Greece

M. T. Hanson
Department of Mechanical Engineering,
University of Kentucky,
Lexington, KY 40506

Rapid Indentation of Transversely Isotropic or Orthotropic Half-Spaces

The canonical problems of rapid indentation by, respectively, a rigid smooth wedge and a rigid smooth cylinder, are examined for a transversely isotropic or orthotropic half-space in plane strain. An exact transient analysis based on integral transforms is carried out for the case of contact zone expansion at a constant subcritical rate. Certain functions in the transform space can be factored in such a manner that the resulting solutions, despite anisotropy, have rather simple forms. This factorization is also exploited to obtain a compact exact formula for the Rayleigh wave speed, which serves as the critical contact zone expansion rate. Aspects of contact zone behavior for the two problems are illustrated for five specific materials. [DOI: 10.1115/1.1365154]

Introduction

Indentation of elastic solids is a key problem in contact mechanics ([1]). For the static case, much is known about the formulation of relevant boundary value problems ([2]) and classes of exact solutions ([3,4]). An extensive literature for anisotropic solids also exists, e.g. ([5–7]). Results for transient dynamic indentation are not as common, and studies of isotropic solids, e.g. ([8–10]), predominate. A recent survey of anisotropic work ([11]) is available, although general formulas for the contact process are a primary focus.

In this light, the present article considers an exact transient analysis of indentation of a transversely isotropic or orthotropic linearly elastic half-space. The canonical situations of indentation by, respectively, a rigid smooth wedge and a rigid smooth cylinder are treated. The shapes imply deformations in plane strain, and, as a first step, the material symmetry axes are normal and tangential to the half-space surface and expansion of the contact zone is at a constant subcritical rate.

The analysis begins in the next section with the basic equations for the material, and identification of the related problem of arbitrary tractions applied to the half-space surface. The exact integral transform solution to the problem is obtained, and serves as the basis for construction of the indentation problem solutions.

Key steps in the analysis are factorizations of certain functions of the transform variable that give, despite the material anisotropy, rather simple expressions for the transform of the normal displacement of the half-space surface. Indeed, one of the factors is extracted from a function of the Rayleigh type, and allows a compact exact formula for the Rayleigh wave speed along the surface. To illustrate solution behavior, contact zone variation with indentation motion parameters is examined in the two cases for five specific materials ([12]).

Basic Problem

Consider a half-space, defined in terms of Cartesian coordinates (x,y,z) as the region $y>0$, initially at rest. The half-space material is of a class of linear homogeneous anisotropic solids whose nontrivial governing equations in plane strain in the absence of body forces have the form

$$c_{11}\frac{\partial^2 u_x}{\partial x^2}+c_{44}\frac{\partial^2 u_x}{\partial y^2}+(c_{13}+c_{44})\frac{\partial^2 u_y}{\partial x\,\partial y}=\rho\ddot{u}_x \tag{1a}$$

$$c_{44}\frac{\partial^2 u_x}{\partial x^2}+c_{33}\frac{\partial^2 u_y}{\partial y^2}+(c_{13}+c_{44})\frac{\partial^2 u_x}{\partial x\,\partial y}=\rho\ddot{u}_y \tag{1b}$$

with the associated stress-strain formulas

$$\rho_{xx}=c_{11}\frac{\partial u_x}{\partial x}+c_{13}\frac{\partial u_y}{\partial y} \tag{2a}$$

$$\sigma_{yy}=c_{13}\frac{\partial u_x}{\partial x}+c_{33}\frac{\partial u_y}{\partial y} \tag{2b}$$

$$\sigma_{xy}=\sigma_{yx}=c_{44}\left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right). \tag{2c}$$

These equations hold for both orthotropic and transversely isotropic materials, where the x and y-axes are axes of material symmetry. The (u_x,u_y) are the (x,y)-components of displacement, and are functions of (x,y) and time, where $(\cdot)$ denotes time differentiation. The constants $(c_{11},c_{13},c_{33},c_{44})$ are a subset of the elasticities $c_{ik}(i,k=1,2,\ldots,6)$ that appear in the generalized Hooke's law ([13]), and ρ is the mass density. Equations (1) are a special case of a more general form that involves four constants that can be linearly related to various subsets of c_{ik}. In addition to the anisotropies considered here, that form describes the plane-strain response of materials with various types of crystal structure ([14]). Overviews of the general relation between crystal structure and the c_{ik} can be found in ([15,16]).

The subset considered in (1) and (2) are constrained to guarantee either a positive-definite strain energy function or strong ellipticity. The former is generally more restrictive than the latter, which itself guarantees real nonzero wave speeds, and guarantees solution uniqueness. Its restrictions are

$$c_{11}>|c_{12}|,\quad (c_{11}+c_{12})c_{33}>2c_{13}^2,\quad c_{44}>0. \tag{3}$$

The isotropic limit case can be extracted by setting

$$c_{11}=c_{33}=\lambda+2\mu,\quad c_{12}=c_{13}=\lambda,\quad c_{44}=\mu \tag{4}$$

where (λ,μ) are the Lame' constants ([13]).

For convenience, the dimensionless quantities ([12])

$$\alpha=\frac{c_{33}}{c_{44}},\quad \beta=\frac{c_{11}}{c_{44}},\quad \gamma=1+\alpha\beta-m^2,\quad m=1+\frac{c_{13}}{c_{44}} \tag{5}$$

and the temporal variable $s=v_r\times(\text{time})$, where

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, October 2, 2000; final revision, December 16, 2000. Associate Editor: A. K. Mal. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

$$v_r = \sqrt{\frac{c_{44}}{\rho}} \tag{6}$$

are introduced, so that the independent variables (x,y,s) all have the dimensions of length, and (1) and (2) become

$$\beta \frac{\partial^2 u_x}{\partial x^2} + \frac{\partial^2 u_x}{\partial y^2} + m \frac{\partial^2 u_y}{\partial x \partial y} = \frac{\partial^2 u_x}{\partial s^2} \tag{7a}$$

$$\frac{\partial^2 u_y}{\partial x^2} + \alpha \frac{\partial^2 u_y}{\partial y^2} + m \frac{\partial^2 u_x}{\partial x \partial y} = \frac{\partial^2 u_y}{\partial s^2} \tag{7b}$$

$$\frac{1}{c_{44}} \sigma_{xx} = \beta \frac{\partial u_x}{\partial x} + (m-1) \frac{\partial u_y}{\partial y} \tag{7c}$$

$$\frac{1}{c_{44}} \sigma_{yy} = (m-1) \frac{\partial u_x}{\partial x} + \alpha \frac{\partial u_y}{\partial y} \tag{7d}$$

$$\frac{1}{c_{44}} \sigma_{xy} = \frac{1}{c_{44}} \sigma_{yx} = \frac{\partial u_y}{\partial x} + \frac{\partial u_y}{\partial x}. \tag{7e}$$

Equation (6) defines a speed that, in the isotropic limit, is the classical ([17]) rotational wave speed. For purposes of illustration, we consider in view of (3) the following constraints ([12]) on (α,β,γ):

$$2\sqrt{\alpha\beta} \leqslant \gamma \leqslant 1 + \alpha\beta \quad (1 < \beta < \alpha) \tag{8a}$$

$$\alpha + \beta \leqslant \gamma \leqslant 1 + \alpha\beta \quad (1 < \alpha < \beta) \tag{8b}$$

$$2\alpha \leqslant \gamma \leqslant 1 + \alpha^2 \quad (1 < \beta = \alpha). \tag{8c}$$

The class of anisotropic materials governed by (8) includes beryl, cobalt, ice, magnesium, and titanium, as well as the isotropic limit.

For $s>0$ the half-space surface is subjected to the traction boundary conditions

$$\sigma_{yy} = \sigma(x,s), \quad \sigma_{yx} = \tau(x,s) \tag{9}$$

for $y=0$. Here (σ,τ) are largely arbitrary, but must be continuous and bounded almost everywhere, i.e., integrably singular behavior is allowed at isolated points. The initial conditions for $y>0$ are

$$\left(u_x, u_y, \frac{\partial u_x}{\partial s}, \frac{\partial u_y}{\partial s} \right) \equiv 0 \quad (s \leqslant 0) \tag{10}$$

and it is expected that (u_x,u_y) are bounded and continuous for $y>0$ when $s>0$ is finite. The related Boussinesq problem is treated in ([12]). However, the role of the basic problem here in the dynamic indentation study, and manipulations that are convenient for that purpose, suggest that the solution process be briefly outlined.

Transform Solution

Consider the unilateral ([18]) and bilateral ([19]) Laplace transforms

$$\hat{F} = \int_0^\infty F(s) e^{-ps} ds, \tag{11a}$$

$$\tilde{F} = \int_{-\infty}^\infty \hat{F} e^{-pqx} dx \tag{11b}$$

and their corresponding inverse operations

$$F(s) = \frac{1}{2\pi i} \int \hat{F} e^{ps} dp, \tag{12a}$$

$$\hat{F} = \frac{p}{2\pi i} \int \tilde{F} e^{pqx} dq \tag{12b}$$

In (11*a*,*b*) p can be treated as real and positive, while q is imaginery. Integration in (12*a*,*b*) is over the Bromwich contours in the p and q-planes, respectively.

Application of (11*a*,*b*) to (7*a*,*b*) in light of (10) gives the general transform expressions

$$\tilde{u}_x = C_1 e^{-pay} + \frac{1}{mqb} (\alpha b^2 - B^2) C_2 e^{-pby} \tag{13a}$$

$$\tilde{u}_y = \frac{1}{mq\alpha a} (\alpha a^2 - A^2) C_1 e^{-pay} + C_2 e^{-pby} \tag{13b}$$

for $y>0$, where (C_1,C_2) are undetermined functions of (p,q) and

$$A = \sqrt{\alpha}\sqrt{1-\beta q^2}, \quad B = \sqrt{1-q^2} \tag{14a}$$

$$\sqrt{2\alpha}a = \sqrt{S - \sqrt{S^2 - 4A^2B^2}}, \quad \sqrt{2\alpha}b = \sqrt{S + \sqrt{S^2 - 4A^2B^2}} \tag{14b}$$

$$S = A^2 + B^2 + m^2q^2 = 1 + \alpha - \gamma q^2, \quad \alpha a b = AB. \tag{14c}$$

It can be shown ([12]) that (8) guarantees that the branch points $q = \pm(1, 1/\sqrt{\beta})$ of (B,A) lie on the Re(q)-axis, and also constitute branch points for (b,a). Boundedness of (13) for $y>0$ requires that Re$(A,B,a,b) \geqslant 0$ in the cut q-plane. Application of (11*a*,*b*) to (9) in view of (10) and (13) allows the coefficients to be determined as

$$\frac{c_{44} p R}{mq\alpha a} C_1 = -[(m-1)B^2 + \alpha b^2] q\tilde{\tau} - (B^2 - \alpha b^2 + mq^2) b\tilde{\sigma} \tag{15a}$$

$$\frac{c_{44} p R}{mqb} C_2 = -[(m-1)\alpha a^2 + A^2] q\tilde{\sigma} + (B^2 - \alpha b^2 + mq^2)\alpha a\tilde{\tau} \tag{15b}$$

$$R = [(m-1)\alpha a^2 + A^2][(m-1)B^2 + \alpha b^2] q^2 + (B^2 - \alpha b^2 + mq^2)^2 AB. \tag{15c}$$

In the sequel, the normal displacement along $y=0$ is required, and so (13) and (15) are combined to yield

$$\tilde{u}_x = \frac{\alpha a N_1}{R} \frac{\tilde{\tau}}{c_{44} p} - \frac{qM}{R} \frac{\tilde{\sigma}}{c_{44} p}, \tag{16a}$$

$$\tilde{u}_y = \frac{bN_2}{R} \frac{\tilde{\sigma}}{c_{44} p} + \frac{qM}{R} \frac{\tilde{\tau}}{c_{44} p} \tag{16b}$$

for $y=0$, where R is given by (15*c*) and

$$M = (A^2 - \alpha a^2)[(m-1)B^2 + \alpha b^2] + (B^2 - \alpha b^2 + mq^2) mAB \tag{17a}$$

$$N_1 = (B^2 - \alpha b^2)^2 + m^2 q^2 B^2, \tag{17b}$$

$$\alpha b^2 N_2 = -A^2 N_1 \tag{17c}$$

In the isotropic limit, $c_{44} = \mu$, $\alpha = \beta = 1+m$, $\gamma = 2(1+m)$ and (16*a*,*b*) reduce to forms

$$\tilde{u}_x = -\frac{b_o}{R_o} \frac{\tilde{\tau}}{\mu p} - \frac{N_o}{R_o} \frac{\tilde{\sigma}}{\mu p}, \quad \tilde{u}_y = -\frac{a_o}{R_o} \frac{\tilde{\sigma}}{\mu p} + \frac{N_o}{R_o} \frac{\tilde{\tau}}{\mu p} \tag{18}$$

that are consistent with those found in ([20]). In (18)

$$a_o = \sqrt{\frac{1}{1+m} - q^2}, \quad b_o = \sqrt{1-q^2} \tag{19a}$$

$$N_o = 2(q^2 + a_o b_o) - 1, \quad R_o = 4q^2 a_o b_o + (2q^2 - 1)^2 \tag{19b}$$

where R_o is a form ([17]) of the Rayleigh function.

To aid in the transform inversion process, some simplifications of (16) are made: First, (14*b*) gives the quantities

$$d_\pm = \sqrt{\alpha}(b \pm a) = \sqrt{S \pm 2AB} \tag{20}$$

and it follows that $a^2=b^2$ when $d_\pm=0$, i.e.,

$$q^2=\frac{\gamma(1+\alpha)-2\alpha(1+\beta)\pm i2m\sqrt{\alpha}\sqrt{\gamma-\alpha-\beta}}{\gamma^2-4\alpha\beta}. \qquad (21)$$

In view of (8) the denominator and, in the numerator, the real term and second radical in the imaginary term, are positive and vanish in the isotropic limit. However, the imaginary term itself vanishes as $O(\sqrt{\varepsilon})$, $\varepsilon\to 0$ while the others behave as $O(\varepsilon)$. Therefore, the apparent branch points defined in (21) would move to infinity in the q-plane. Additional cuts for these branch points must be introduced so that the restriction $\mathrm{Re}(a,b)\geqslant 0$ is still satisfied. To this end, we allow d_+, i.e., $b+a$, to be continuous across these cuts, although (a,b) themselves are multivalued. Thus, the additional cuts define branches of d_-, i.e., b-a. A similar situation arises for wave diffraction by a crack in a transversely isotropic media ([21]).

It can be shown that the branch points of d_- are also roots of the functions in (15c) and (17). In the Appendix, d_- is extracted as a factor from these functions so that it cancels out in the ratios displayed in (16). From (A4) and (A5) in the Appendix and use of (14c), the result

$$\tilde{u}_x=-\sqrt{\alpha}B\frac{d_+}{2D}\frac{\tilde{\tau}}{c_{44}p}-N\frac{q}{D}\frac{\tilde{\sigma}}{c_{44}p}, \qquad (22a)$$

$$\tilde{u}_y=-A\frac{d_+}{2\sqrt{\alpha}D}\frac{\tilde{\sigma}}{c_{44}p}+N\frac{q}{D}\frac{\tilde{\tau}}{c_{44}p} \qquad (22b)$$

for $y=0$ can be obtained, where

$$D=A+[A^2+(m-1)^2q^2]B,\quad N=A+(1-m)B. \qquad (23)$$

It is noted that, for the restrictions (8), the coefficients of $(\tilde{\sigma},\tilde{\tau})$ exhibit only the branch cuts $\mathrm{Im}(q)=0$, $|\mathrm{Re}(q)|>1$ and $\mathrm{Im}(q)=0$, $|\mathrm{Re}(q)|>1/\sqrt{\beta}$. In particular, D is analytic in the q-plane cut along $\mathrm{Im}(q)=0$, $|\mathrm{Re}(q)|>1/\sqrt{\beta}$ and exhibits the nonisolated real roots $q=\pm 1/c_R(0<c_R<1)$. Indeed, setting $D=0$ and rationalization gives a cubic equation in q^2 that is identical in form to that obtained in ([12]) as Eq. (4.3.22) for the roots of the transversely isotropic Rayleigh function. That is, c_Rv_r is the Rayleigh wave speed ([22]) parallel to the x-axis for the class of materials considered here, and D is itself the essential factor of the Rayleigh function. As an alternative to the aforementioned cubic equation, c_R can be, by following a general approach ([21,23]), obtained to within a simple quadrature as

$$c_R=\sqrt{\frac{\alpha\beta-(m-1)^2}{(1+\sqrt{\alpha})\sqrt{\alpha\beta}}}G_o \qquad (24a)$$

$$\ln G_o=\frac{1}{\pi}\int_{1/\sqrt{\beta}}^{1}\frac{dt}{t}\tan^{-1}\frac{\sqrt{1-t^2}}{\sqrt{\alpha}\sqrt{\beta t^2-1}}[\alpha+((m-1)^2-\alpha\beta)t^2] \qquad (24b)$$

where (8) guarantees that the coefficient of G_o is real-valued. With (22) available, the two dynamic indentation problems can be addressed.

Indentation by a Rigid Wedge

Consider the same half-space governed by (7), (8), and (10). In this case, however, a rigid wedge of half-angle $\frac{1}{2}\pi-\psi$ is pressed directly into the surface $y=0$ for $s>0$, as depicted schematically in Fig. 1. The indentation speed is constant, as is the rate at which the contact zone spreads symmetrically from the wedge apex. In view of (6), (C,c) are the two speeds, nondimensionalized with respect to v_r. The former dimensionless constant (indentation speed) is given, while the latter (contact zone expansion rate) is a priori unknown. A subcritical rate is assumed, however, so that the restriction

$$0<c<c_R \qquad (25)$$

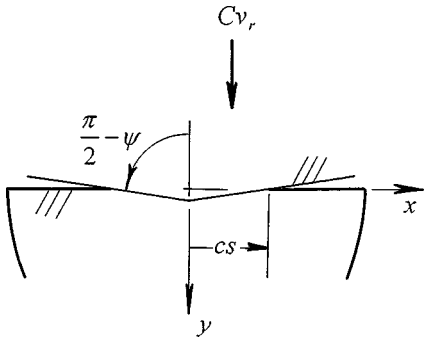


Fig. 1 Schematic of indentation by a rigid wedge

holds in view of (24). If contact is smooth, no shear stress exists along $y=0$. Therefore, in light of (9) and Fig. 1, the transform solution obtained above will also satisfy the indentation problem upon setting $\tau\equiv 0$ and finding σ such that the contact zone condition

$$u_y=Cs-\psi|\mathbf{x}| \quad (|x|<cs) \qquad (26)$$

is satisfied for $y=0$. Here the usual ([3]) assumption is made that the wedge angle is large enough to justify the approximation $\tan\psi\approx\psi$. In addition to (26), the choice of σ must satisfy the two Signorini conditions ([2]) of smooth contact: (I) the contact zone cannot be in tension anywhere, *i.e.*, $\sigma\leqslant 0$, $|x|<cs$, and (II) interpretation of the wedge and half-space surfaces should not occur outside ($|x|>cs$) the contact zone. Moreover, the solution should include determination of the zone expansion rate parameter c.

The indentation problem geometry has no characteristic length, and its only nonhomogeneous condition, (26), prescribes a displacement that is homogeneous of degree 1 in (x,s). The full-field displacement for $y>0$ should, therefore, be homogeneous of degree 1 in (x,y,s), and the corresponding analysis for the isotropic half-space ([20]) suggests the trial function

$$\sigma=\frac{\sigma_o}{\pi}\cosh^{-1}\frac{s}{c|x|} \quad (|x|<cs). \qquad (27)$$

Here σ_o is an unknown constant that can be shown to be, in fact, the average normal traction over the contact zone. That (27) vanishes continuously at the contact zone edges $|x|=cs$ essentially guarantees ([20]) that the second Signorini condition will be satisfied for (25).

Operating on (27) with (11) gives

$$\tilde{\sigma}=\frac{c\sigma_o}{p^2\sqrt{1-q^2c^2}} \qquad (28)$$

Substitution of (28) into (22b) and setting $\tilde{\tau}\equiv 0$ gives the transform of the normal displacement along $y=0$. The analyticity of the result indicates that the entire $\mathrm{Im}(q)$-axis can serve as the Bromwich contour in the inversion operation (12b). However, Cauchy theory can be used to switch the contour onto paths surrounding the branch cuts of (A,B,d_+) on the $\mathrm{Re}(q)$-axis. Thus, the exponential term can be made to be negative and real and, in a manner similar to the Cagniard-deHoop procedure ([24]), the inversion operation (12a) can be performed by inspection. Setting $\mathrm{Re}(q)=t$, $\mathrm{Im}(q)=0+$ then yields the formal integration

$$u_y=-\frac{c\sigma_o}{2\pi c_{44}}\mathrm{Im}\int_0^{s/|x|}\frac{s-t|x|}{\sqrt{1-t^2c^2}}\frac{Ad_+}{D}dt \qquad (29)$$

along the upper side of the $\mathrm{Re}(q)$-axis.

In view of (14a), (20), and (23), the integrand of (29) is purely real for $t>1/c$, so that when $|x|<cs$, $1/c$ becomes the effective upper limit of integration and (29) is, appropriately, linear in $(|x|,s)$. The factors of the s and $|x|$-terms in the integrand behave

as $O(t^{-2})$ and $O(t^{-1})$, respectively, as $|t|\to\infty$. Therefore, Cauchy theory can be used to transform the integration to the positive Im(q)-axis and a quarter-circle of infinite radius in the first quadrant of the q-plane. The result is that for $y=0$, $|x|<cs$,

$$u_y=\frac{c\sigma_o}{c_{44}}\left[-sI+\frac{|x|}{4c\chi\sqrt{\alpha}}\left(1+\sqrt{\frac{\gamma+2\sqrt{\alpha\beta}}{\gamma-2\sqrt{\alpha\beta}}}\right)\right] \tag{30}$$

where

$$\chi=\frac{\alpha\beta-(m-1)^2}{\sqrt{\gamma^2-4\alpha\beta}-\gamma+2\alpha\beta}\left(\frac{\sqrt{\alpha\beta}-1}{\sqrt{\gamma-2\sqrt{\alpha\beta}}}+\frac{\sqrt{\alpha\beta}+1}{\sqrt{\gamma+2\sqrt{\alpha\beta}}}\right) \tag{31a}$$

$$I=\frac{1}{\pi}\int_0^{\pi/2}\frac{\Omega\sqrt{T+2\Omega}}{\Omega(\Omega+\cos^2\phi)-(m-1)^2\sin^2\phi}\frac{d\phi}{\sqrt{\cos^2\phi+c^2\sin^2\phi}} \tag{31b}$$

$$T=(1+\alpha)\cos^2\phi+\gamma\sin^2\phi,\quad \Omega=\sqrt{\alpha}\sqrt{\cos^2\phi+\beta\sin^2\phi} \tag{31c}$$

and the integration variable change Im(q)=tan ϕ has been made for computational efficiency. Substitution of (30) into (26) gives the equations

$$\frac{\sigma_o}{c_{44}}=\frac{-4\chi\sqrt{\alpha}\psi}{1+\sqrt{\dfrac{\gamma+2\sqrt{\alpha\beta}}{\gamma-2\sqrt{\alpha\beta}}}} \tag{32a}$$

$$cI=\frac{C}{4\chi\sqrt{\alpha\psi}}\left(1+\sqrt{\frac{\gamma+2\sqrt{\alpha\beta}}{\gamma-2\sqrt{\alpha\beta}}}\right) \tag{32b}$$

for both the average normal traction σ_o needed to completely define the field (26) and the (dimensionless) contact zone expansion rate c.

For (8) the quantities on the right-hand side of (32a,b) are, respectively, negative and positive. This guarantees in light of (27) and Fig. 1 that the first Signorini condition is also satisfied. It is noted that the average contact zone stress is independent of both the indentation speed (C) and the rate of contact zone expansion (c). This feature arises from the homogeneous nature of the solution field, *i.e.*, the problem geometry has no characteristic length.

Indentation by a Rigid Cylinder

Consider the schematic in Fig. 2: A rigid smooth cylinder of radius r is pressed into the half-space at a constant acceleration a_o. The contact zone is again assumed to spread symmetrically from the initial contact point ($x=y=0$, $s=0$) with a constant subcritical speed, where c is again its value nondimensionalized with respect to v_r, and is restricted by (25). During the early part of the process, one can assume that $r\gg cs$, whereupon the contact zone condition is

$$u_y=\frac{a_o s^2}{2v_r^2}-\frac{x^2}{2r}\quad(|x|<cs). \tag{33}$$

As in the wedge study, the solution for this problem can be based on the transform solution (13) if a σ can be identified that allows the inverse of (22b) with $\tau\equiv 0$ to satisfy (33) and the two Signorini conditions, and that allows the expansion rate parameter c to be determined.

There is again no characteristic length in the problem geometry, and corresponding isotropic results ([20]) suggest the trial function

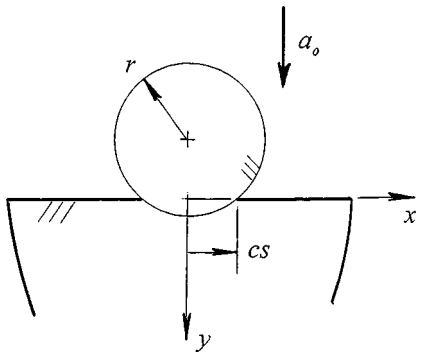


Fig. 2 Schematic of indentation by a rigid cylinder

$$\sigma=\frac{S_o}{\pi c}\sqrt{c^2s^2-x^2}\quad(|x|<cs). \tag{34}$$

Here S_o is an unknown constant and, as in the wedge case, the form of (34) essentially guarantees that the second Signorini condition, *i.e.*, interpenetration does not occur outside the contact zone for (25), is satisfied. Application of (11) to (34) gives

$$\tilde{\sigma}=\frac{c^2S_o}{p^3(1-q^2c^2)^{3/2}}. \tag{35}$$

Substitution of (35) into (22b) with $\tilde{\tau}\equiv 0$ gives a form that is apparently nonintegrable. However, the use of formal derivative operations and finite-part integration ([25]) shows that the inverse of (22b) exists and is indeed quadratic in (x,s) for $|x|<cs$. Substitution into (33) then yields the equations

$$\frac{rS_o}{c_{44}}=-\frac{1}{cI_c}, \tag{36a}$$

$$\frac{I_c}{I_s}=\frac{ra_o}{2\beta v_r^2} \tag{36b}$$

for both the constant S_o needed to completely define (34) and the (dimensionless) contact zone expansion rate c. In (36)

$$(I_c,I_s)=\frac{1}{\pi}\int_0^{\pi/2}\frac{\Omega\sqrt{T+2\Omega}}{\Omega(\Omega+\cos^2\phi)-(m-1)^2\sin^2\phi}\times\frac{(\cos^2\phi,\sin^2\phi)d\phi}{(\cos^2\phi+c^2\sin^2\phi)^{3/2}} \tag{37}$$

and, in light of (31b), it is seen that $I=I_c+c^2I_s$. It is also noted that $\sqrt{\beta}v_r$ is the dilatational wave speed parallel to the x-direction. The quantities $(I_c,I_s)>0$ so that, in view of (34) and (36a), the first Signorini condition is also satisfied. In contrast to the wedge case, (36a) shows that the defining constant S_o depends on the indentor shape (r) as well as the contact zone expansion rate (c).

Solution Aspects

In ([12]) properties for the materials mentioned earlier as being typical of the conditions (8)—beryl, cobalt, ice, magnesium, titanium—are listed. For insight into the response of such materials to indentation by the wedge, the data gathered in ([12]) is used in connection with (24) and (32a) to calculate the dimensionless Rayleigh speed c_R and the average normal contact zone stress σ_o. The results are presented in Table 1, where it is seen that the Rayleigh wave speed parallel to the x-direction varies little between materials as a fraction of the corresponding rotational wave speed. The average normal contact zone stress, however, varies widely. As noted earlier ψ (radians) should be small, but beryl,

Table 1 Dimensionless Rayleigh wave speed and average normal contact zone stress

	α	β	m	c_{44} (GPa)	c_R	σ_o (GPa)
Beryl	3.62	4.11	2.01	68.6	0.956	-67.6ψ
Cobalt	4.74	4.07	2.37	75.5	0.962	-85.9ψ
Ice	4.57	4.26	2.64	3.17	0.959	-3.47ψ
Magnesium	3.74	3.61	2.3	16.4	0.943	-19.5ψ
Titanium	3.88	3.47	2.48	46.7	0.936	-45.5ψ

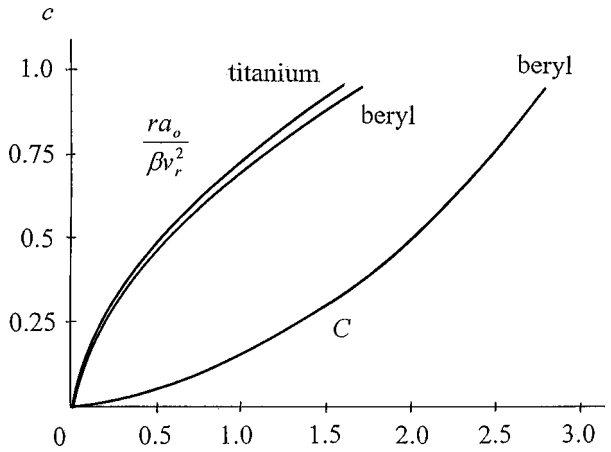


Fig. 3 Contact zone expansion rate variation with wedge and cylinder motion

cobalt and titanium—being stiffer in shear (c_{44})—must, nevertheless, withstand much larger average stress values for a given wedge profile.

Another feature of importance is the wedge speed/contact zone expansion rate relation (32*a*). Values of c that satisfy (25) are plotted versus C in Fig. 3 for beryl. Except for the small differences in the maximum allowable value of c, c_R, seen in Table 1, this curve would, on the scale chosen, essentially also serve for the other four materials. That is, although (31) and (32*b*) suggest a robust dependence on material constants (α,β,m,γ), the variations of expansion rate with wedge speed as fractions of the rotational wave speed are essentially the same. Figure 3 does indicate, nevertheless, that the rate of increase of zone expansion rate is higher for higher wedge speeds.

For insight into the short-time response of the five materials to indentation by the accelerating cylinder, (36*b*) is also plotted in Fig. 3. In this case, the two curves, for beryl and titanium, serve on the scale given, to illustrate the response for the materials. Figure 3 does show, nevertheless, that c is more sensitive to changes in the cylinder motion properties (r,a_o) than it is to the wedge motion properties (ψ,C).

Comments

This article considered the dynamic indentation of a class of transversely isotropic or orthotropic linearly elastic homogeneous half-spaces in plane strain. For illustration, indentation was either by a rigid smooth wedge, moving at a constant speed, or a rigid smooth cylinder in constant acceleration. The contact zone was assumed to expand symmetrically from the points of initial contact at constant subcritical rates.

A transient analysis provided exact transform solutions in terms of a contact zone normal stress distribution. The wedge solution showed for five specific materials in the class considered that the average normal contact zone stress is largely determined by the material shear modulus, i.e., materials stiffer in shear must withstand higher average stresses. The relation between wedge speed and contact zone expansion rate, as fractions of the rotational wave speed along the half-space surface, was approximately the same for all the materials considered. The average normal stress in the contact zone was, moreover, independent of the wedge speed and zone expansion rate. This lack of sensitivity arose, in part, because of the lack of a characteristic length in the problem geometry.

A study of the cylinder solution for the five materials showed more sensitivity: In particular, the contact zone expansion rate as a fraction of rotational wave speed depended on both motion parameters (acceleration, radius). Again, however, the lack of a characteristic length in problem geometry produced expansion rate/acceleration relations that were somewhat independent of the particular material.

The critical expansion rate was taken as the Rayleigh wave speed ([22]) associated with the half-space surface. A formula for this speed, analytical to within a simple quadrature, was obtained. A similar approach has been used in ([21]), but here a factorization of a function of the Rayleigh type yielded a more compact result. This factorization process also simplified the forms of the integral transforms of the solution. This was worthwhile because the anisotropic materials considered here give rise to additional branch points in the transform plane that do not occur in the isotropic limit ([12,21]). Such factorizations can also be performed in that limit (see the Appendix) but, because additional branch points do not arise, the advantage of the procedure is less relevant.

In summary, the present results are limited by the lack of a characteristic length. Nevertheless, they are exact, and in a somewhat simpler form than might be expected for an anisotropic material. They also do allow first-step insight into the response of transversely isotropic or orthotropic materials to rapid (dynamic) indentation.

Acknowledgment

The support of the National Technical University of Athens facilitated the visit of HGG with LMB at the University of Kentucky for preliminary discussions of this project. HGG also acknowledges the support of the Greek General Secretariat for Research and Technology under program PENED 99 DE 642.

Appendix

Consider N_2 defined in (17*c*). By using (17*b*) and (14*c*), it can be written as

$$N_2=(B^2-\alpha b^2)(A^2-\alpha a^2)-m^2q^2\alpha a^2. \qquad (A1)$$

Carrying out the multiplication and using (14*a*,*b*) and (20) gives

$$N_2=2A^2B^2-\frac{S^2}{2}+\frac{d_+^2d_-^2}{2}(m^2q^2+B^2-A^2). \qquad (A2)$$

But (14*a*,*b*) and (20) also show that

$$S^2-4A^2B^2=d_+^2d_-^2, m^2q^2+B^2=\frac{1}{2}(d_+^2+d_-^2)-A^2 \qquad (A3)$$

whereupon (*A*2) can be written as

$$N_2=\frac{d_+d_-}{4}[(d_+-d_-)^2-4A^2]=d_+d_-(\alpha a^2-A^2). \qquad (A4)$$

In a similar fashion, (17*b*) and (15*c*) can be written as

$$M=\frac{Bd_-}{\sqrt{\alpha}a}(A^2-\alpha a^2)N, \quad N=A+(1-m)B \qquad (A5a)$$

$$R=\frac{Bd_-}{\sqrt{\alpha}a}(A^2-\alpha a^2)D, \quad D=A+[A^2+(m-1)^2q^2]B. \quad (A5b)$$

These forms exhibit the common factor $d_-(A^2-\alpha a^2)$.

Similar factorizations of the terms (N_o,R_o) in (19) that arise in the isotropic limit can also be performed:

$$N_o=(b_o-a_o)\left(\frac{1+m}{m}a_o+\frac{1-m}{m}b_o\right) \quad (A6a)$$

$$R_o=\frac{2}{m}(b_o-a_o)[(1+m)a_o+(1+m-4mq^2)b_o]. \quad (A6b)$$

References

[1] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.

[2] Fichera, G., 1972, ''Boundary Value Problems of Elasticity With Unilateral Constraints,'' in *Handbuch der Physik*, VIa/2, Springer, Berlin, pp. 391–424.

[3] Muskhelishvili, N. I., 1975, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff, Leyden.

[4] Gladwell, G. M. L., 1980, *Contact Problems in the Classical Theory of Elasticity*, Sijthoff and Noordhoff, Alphen aan Rijn.

[5] Willis, J. R., 1966, ''Hertzian Contact of Anisotropic Bodies,'' J. Mech. Phys. Solids, **16**, pp. 163–176.

[6] Fan, H., and Keer, L. M., 1994, ''Two-Dimensional Contact on an Anisotropic Half-Space,'' ASME J. Appl. Mech., **61**, pp. 250–255.

[7] Hanson, M. T., 1992, ''The Elastic Field for Spherical Hertzian Contact Including Sliding Friction for Transverse Isotropy,'' ASME J. Tribol., **114**, pp. 606–611.

[8] Bedding, R. J., and Willis, J. R., 1973, ''Dynamic Indentation of an Elastic Half-Space,'' J. Elast., **3**, pp. 289–309.

[9] Georgiadis, H. G., and Barber, J. R., 1993, ''On the Super-Rayleigh/Subseismic Elastodynamic Indentation Problem,'' J. Elast., **31**, pp. 141–161.

[10] Brock, L. M., and Georgiadis, H. G., 1994, ''Dynamic Frictional Indentation of an Elastic Half-Plane by a Rigid Punch,'' J. Elast., **35**, pp. 223–249.

[11] Borodich, F. M., 2000, ''Some Contact Problems of Anisotropic Elastodynamics: Integral Characteristics and Exact Solutions,'' Int. J. Solids Struct., **37**, pp. 3345–3373.

[12] Payton, R. G., 1983, *Elastic Wave Propagation in Transversely Isotropic Media*, Martinus Nijhoff, The Hague.

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

[14] Scott, R. A., and Miklowitz, J., 1967, ''Transient Elastic Waves in Anisotropic Plates,'' ASME J. Appl. Mech., **34**, pp. 104–110.

[15] Nye, J. F., 1957, *Physical Properties of Crystals, Their Representation by Tensors and Matrices*, Clarendon Press, Oxford, UK.

[16] Theocaris, P. S., and Sokolis, D. P., 2000, ''Invariant Elastic Constants and Eigentensors of Orthorhombic, Tetragonal, Hexagonal and Cubic Crystalline Media,'' Acta Crystallogr., **A56**, pp. 319–331.

[17] Achenbach, J. D., 1973, *Wave Propagation in Elastic Solids*, North-Holland, Amsterdam.

[18] Sneddon, I. N., 1972, *The Use of Integral Transforms*, McGraw-Hill, New York.

[19] van der Pol, B., and Bremmer, H., 1950, *Operational Calculus Based on the Two-Sided Laplace Integral*, Cambridge University Press, Cambridge, UK.

[20] Brock, L. M., 1991, ''Exact Transient Results for Pure and Grazing Indentation With Friction,'' J. Elast., **33**, pp. 119–143.

[21] Norris, A. N., and Achenbach, J. D., 1984, ''Elastic Wave Diffraction by a Semi-Infinite Crack in a Transversely Isotropic Material,'' Q. J. Mech. Appl. Math., **37**, pp. 565–580.

[22] Buchwald, V. T., 1961, ''Rayleigh Waves in Transversely Isotropic Media,'' Q. J. Mech. Appl. Math., **14**, pp. 293–317.

[23] Brock, L. M., 1998, ''Analytical Results for Roots of Two Irrational Functions in Elastic Wave Propagation,'' J. Aust. Math. Soc. B, Appl. Math., **B40**, pp. 72–79.

[24] deHoop, A. T., 1960, ''A Modification of Cagniard's Method for Seismic Pulse Problems,'' Appl. Sci. Res., **B8**, pp. 349–356.

[25] Hadamard, J., 1908, ''Theories des Equations aux Derives Partielles Lineares Hyperboliques et du Problem de Cauchy,'' Acta Math., **31**, pp. 333–380.

Journal of Applied Mechanics

Brief Notes

A Brief Note is a short paper that presents a specific solution of technical interest in mechanics but which does not necessarily contain new general methods or results. A Brief Note should not exceed 1500 words *or equivalent* (a typical one-column figure or table is equivalent to 250 words; a one line equation to 30 words). Brief Notes will be subject to the usual review procedures prior to publication. After approval such Notes will be published as soon as possible. The Notes should be submitted to the Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME International, Three Park Avenue, New York, NY 10016-5990, or to the Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on Brief Notes appearing in this issue will be accepted until two months after publication. Readers who need more time to prepare a Discussion should request an extension of the deadline from the Editorial Department.

Elastic Multiscale Contact of Rough Surfaces: Archard's Model Revisited and Comparisons With Modern Fractal Models

M. Ciavarella
CNR-IRIS, str. Crocefisso 2/B, 70125 Bari, Italy
e-mail: M.Ciavarella@area.ba.cnr.it

G. Demelio
Dipartimento di Progettazione e Produzione Industriale, Politecnico di Bari, Viale Japigia 182, 70126 Bari, Italy
e-mail: Demelio@poliba.it

1 Introduction

Bowden and Tabor ([1] BT, in the following) state that friction is dictated by ''adhesion (cold weld)'' and ''ploughing (inelastic deformation term),'' between asperities. Amonton's law could easily be explained for the ploughing term, as the real area of contact would simply be $A=P/H$, where H is the hardness of the softer of the contacting bodies, and P is the applied load. However, for the elastic term, which in most cases would be the dominant one, Hertz' theory would not predict linearity with load. During the 1950s, several articles appeared in prestigious journals ([2–4]) where multiscale models were introduced to explain Amonton's and several connected well-known laws for friction, wear and electrical/thermal resistance, in terms of elastic deformations of multiscale, and rigorously ∞-scale model which we would now call a fractal ([4]), as depicted in Fig. 1—this is not the only possible choice, as the Archard model only takes into account of load redistribution and not of the actual geometry. These models found that the relation *real* contact area to *total* load for an ensemble of elastic asperities separated enough from each other to neglect interaction effects, is

$$A_n = K_n \left(\frac{W}{\tilde{E}} \right)^{\alpha_n} \tag{1}$$

where $\tilde{E}=E/(1-\nu^2)$, and K_n is a coefficient which depends on the number of scales introduced n, and was computed by Archard for the first few scales only of his model. Archard's main finding was that α_n tends rapidly to one as n is increased. No particular attention was, vice versa, paid by Archard to the coefficient K_n, which will be here recomputed in general and will be specialized for a fractal geometry.

Independently from these multiscale models, and actually extending these results, Greenwood and Williamson ([5], GW model, in the following) showed that statistical distribution of heights asperities leads (at least approximately) to linearity between τ and p *independently on the exact law relating locally* τ *with* p, i.e., including any arbitrary local elastoplastic constitutive and frictional laws. Only recent experiments at very small scales with just one asperity in contact under very carefully controlled conditions are having some success in explaining the intrinsic properties of friction (see [6]). It has been found in particular that the friction coefficient is a function of the size of the asperity, and varies from very high values close to the elastic moduli of the materials (around $G/30$, in particular, where G is the shear modulus of the material), for smallest size to Peierls stress values, comparable to yield limits at larger sizes. When this knowledge will be completed, the way towards quantitative predictions of the actual ''averaged'' friction coefficient will depend on the actual precise determination of the distribution of contact sizes. This in turn will need an accurate modelling of real surfaces. As recently proposed, fractal models seem to have a promising role in concisely describing the apparent self-affinity of roughness, i.e., with features repeating themselves at different scales ([7,8]) but early attempts to use measurements of real surfaces and modelling of contact ([9,10]) were somehow unsatisfactory because they only considered a geometrical method for computing the contact area from a ''bearing area'' assumption. Borri-Brunetto et al. [11], vice versa, created a finite numerical realization of a surface with appropriate fractal properties and then used a numerical method to solve the resulting elastic contact problem at various levels of spatial discretization, suggesting that in the fractal limit the contact may consist of an infinite number of infinitesimal contact areas of total area zero. In other words, the actual contact area appears to be a fractal with dimension below two. This originated a discussion between the authors during the process of writing a review paper ([12]), and then to develop a rigorous analysis, specialized to the plane contact for a Weierstrass profile ([13], CDBJ model in the following), demonstrating that extended regions of contact are not possible with this model.

After this effort, we moved back to the original Archard's work and recognized that, although its surface is not a fractal of well-known characteristics, and although the contact mechanics is not

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, May 22, 2000. Associate Editor: J. R. Barber.

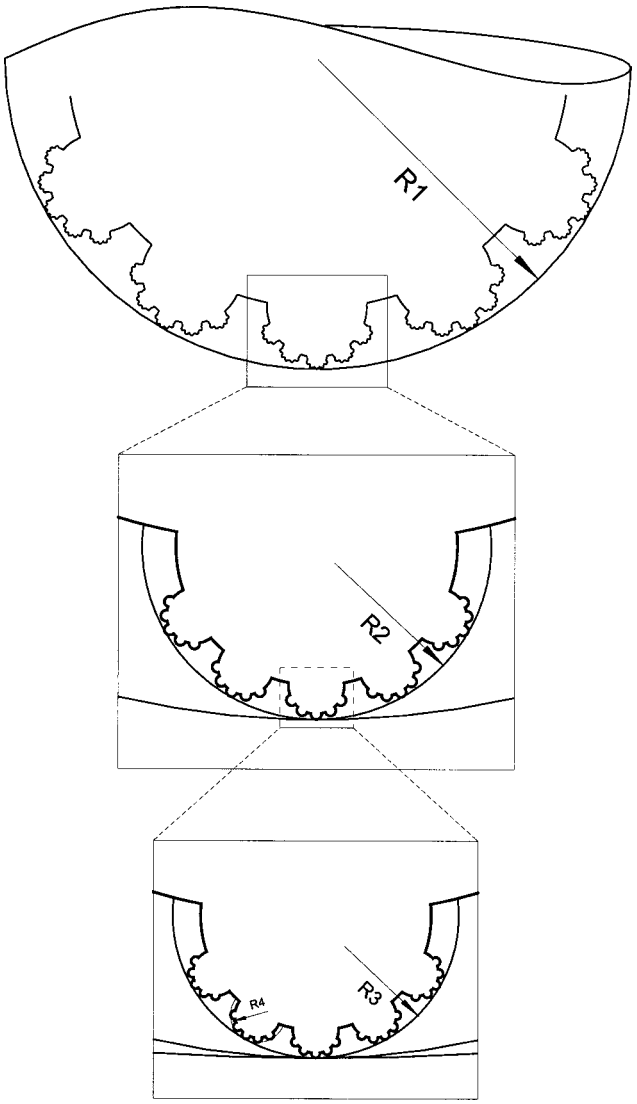


Fig. 1 Example of geometry for the Archard model

as rigorous as it can be with the Weierstrass profile, the result of fractal contact area can be reached very simply by extending the calculations of the original Archard paper ([4]) a little further, with the appropriate assumptions on the asperities geometry. The results, as shown in the present paper, are surprisingly similar to the Weierstrass profile CDBJ model, in many respects. A discussion of possible implications of these findings for friction theories follows.

2 The Revisited Archard's Model

Archard's model is based on the assumption that each asperity at scale n is replaced by many asperities at the higher scale. This permits an ''uncoupling'' of scales in the calculations of the redistribution of the pressure from one scale to the next, smaller one. Therefore, a very simple, analytical recursive argument can be developed. In order to consider ''interaction of scales'' one would need to consider the effect of smaller asperities in the calculation of the compliance of the larger ones, as Greenwood and Tripp [14] have done in the context of the statistical models. It is instructive to consider, however, that while the GT model seems to introduce some degree of multiscale features, the decision of having just two scales is somehow arbitrary, except for the case where the macroscopic roundness of the sphere is evident. Also, the GT model only obtains in most cases a minor modification of the pressure distribution at the largest scale, due to the small compliance effects, but loses the Archard feature of showing the pressure levels at smaller and smaller scales. It is clear that in principle a numerical model could take into account of all these effects simultaneously, but it remains a problem to build a multiscale model from a measured spectrum of the profile.

Turning back to the Archard model, at scale 0 Hertzian relations ([15], 4.2a) give the contact area as (1) with $K_0 = \pi(3R_0/4)^{2/3}$, $\alpha_0 = 2/3$, and R_0 being the radius of the sphere. The pressure distribution is $p(r) = 2\tilde{E}/\pi R\sqrt{r_0^2 - r^2}$. We now assume that in the area A_0 there is a uniform distribution of asperities having density m_1 (number of asperities per unit area). Extending Archard's procedure for a contact of always higher number of scales, we obtain the following general expressions of the contact area at the nth scale as a function of load:

$$A_n = \pi\left(\frac{3}{4}R_0\frac{W}{\tilde{E}}\right)^{(1-1/3^{n+1})} \times\prod_{k=1}^{n}\left[\frac{2}{3-3^{-k}}m_k^{1/3^{n+1-k}}\left(\frac{3}{2}\frac{R_k}{R_{k-1}}\right)^{1-1/3^{n+1-k}}\right]. \quad (2)$$

Notice that for $n->\infty$ the dependence on load becomes linear, independently on the assumptions on the geometrical quantities m_n, R_n. If, on the other hand, the ratios m_k/m_{k-1} and R_k/R_{k-1} are kept constant, we get

$$\frac{A_n}{A_{n-1}} = \frac{2}{3-3^{-n}}(A_0 m_0)^{3^{-n}}\left(\frac{3}{2}\frac{R_n}{R_{n-1}}\sqrt{\frac{m_n}{m_{n-1}}}\right)^{1-3^{-n}}. \quad (3)$$

If we now *assume* for the spacing λ_n between asperities and for the radii R_n a power-law function, i.e., $\lambda_{n-1}/\lambda_n = \gamma$ and $R_{n-1}/R_n = \gamma^D$, the density of asperities is obviously $m_n = 1/\lambda_n^2$, i.e., $m_{n-1}/m_n = (\lambda_{n+1}/\lambda_n)^2 = 1/\gamma^2$. Then, the ratio between contact area at subsequent scales is found from (3) to be

$$\frac{A_n}{A_{n-1}} = \frac{2}{3-3^{-n}}\left(\frac{A_0}{\lambda_0^2}\right)^{3^{-n}}\left(\frac{3}{2}\gamma^{1-D}\right)^{1-3^{-n}}. \quad (4)$$

The limit for $n\to\infty$ is

$$\frac{A_n}{A_{n-1}} = \left(\frac{1}{\gamma}\right)^{D-1}. \quad (5)$$

The tendency to power-law (5) indicates that the contact area tends to a fractal set, whereas the fractal dimension computed with the box-counting method (see CDBJ) is

$$d_A = -\frac{\ln\frac{N_{n+1}}{N_n}}{\ln\frac{\lambda_{n+1}}{\lambda_n}} = -\frac{\ln\left(\frac{A_n}{A_{n-1}}\frac{m_{n+1}}{m_n}\right)}{\ln\frac{1}{\gamma}} = -\frac{\ln\left[\left(\frac{1}{\gamma}\right)^{D-1}\left(\frac{1}{\gamma}\right)^{-2}\right]}{\ln\frac{1}{\gamma}} \quad (6)$$

indicating that the *limiting* fractal dimension is

$$d_A = 3 - D. \quad (7)$$

Considering that $D = 1-2$ by analogy to the Weierstrass case (see CDBJ), the contact area has dimension between 1 and 2, analogously to what found in CDBJ.

3 Discussion

The Archard model leads to asymptotic fractal behavior, under assumptions on radii of curvature for the asperities similar to the Weierstrass CDBJ model, and in contrast with Majumdar and Bhushan [9,10], and their ''bearing area'' geometrical assumption, with resulting finite area of contact. Obviously, we don't expect the contact area to be really zero, as real surfaces will have a truncation at one point, with asperities of given minimum size. Even if this was not the case, at a certain scale, deformations would be so intense that plastic deformation, nonlinearities, and other effects not included in our model would appear. It is clear that, in the spirit of Archard's model, if one asperity yields, this does not affect the load redistribution at the other asperities, as equilibrium is already accounted for. Therefore, an idea of the pressure level increase with scale can be found from the total contact area variation with n, as can be obtained from equations given above. The resulting trends are very similar to the ones obtained for the CDBJ model. In particular, the fractal dimension is a constant depending only on geometry and not on load level, but at the first few scale, both higher and lower apparent fractal dimension can occur. Also, as proved with the CDBJ model results, although both the contact areas and the distance between asperities become smaller, the ratio between the two decreases, so that interaction effects become increasingly smaller, and the Hertzian approximation becomes valid in any conditions, i.e., even in cases where high loads predict full contact at macroscopic scale (in this case, the present model is poor whereas the CDBJ model correctly considers the Westergaard solution for predicting contact area size).

4 Conclusions

The most striking conclusion of the calculations is that with multiscale models the contact area generally (if the radius of asperities decreases fast enough) tends to zero, i.e., is a fractal. The reason why such an implication had escaped the attention of researchers for more than 40 years, particularly as the model is quite well known. A possible explanation is that the main issue at that time was to find the linearity of relation (1), i.e., that $\alpha_n \rightarrow 1$, whereas the coefficients K_n were never computed for more than 2–3 scales. The results confirm the conclusion reached numerically by Borri-Brunetto et al. [11] that the contact area is defined by a fractal set—i.e., that contact is restricted to an infinite set of infinitesimal contact segments in the limit $n \rightarrow \infty$; there are no contact segments of finite dimension and the total contact area tends regularly to zero. In addition, the deviation from simple power-law fractal behavior at low wave numbers provides an explanation of their observation that the apparent fractal dimension is load-dependent. Even at large n, the splitting of segments of the contact area does not follow a ''simple'' rule for successive scales. Therefore, at successive scales, even if yielding is reached at one location, contact splitting will continue at other location, until yielding is reached even there. Therefore, it now becomes clear that the Archard model is in the limit compatible in a sense to the old Bowden-Tabor simple idea of contact area size given by $A = P/H$. Therefore, Greenwood-Williamson's model predicts Amonton's law from just the effect of randomness of the asperity height distribution, independently on the constitutive law at microscopic scale, Archard's model explains it as just an effect of load redistribution for a deterministic geometry, and leads in the limit to the other possible explanation (the Bowden-Tabor's fully plastic one), this goes some way in explaining why Amonton's law is so well hidden and intrinsic in the contact of any surface.

However, as Bowden-Tabor's theories and experiments show that full yield (ploughing term of friction) is negligible with respect to ''adhesive'' elastic term, particularly for hard materials and repeated sliding (shakedown), we can infer that the real case has a combination of features from all of the above models, and that the normal contact problem is largely unsolved. Future predictions of global friction coefficient depend crucially on better solutions and understanding of this problem.

Acknowledgments

The first author (M.C.) is pleased to acknowledge extensive discussions with Prof. J. R. Barber, Dr. J. A. Greenwood and Prof. K. L. Johnson, and support from CNR for a Short Term fellowship to travel to Oxford University in July 1999, where this work was completed, and for the network COMES (COmputational MEchanics of Solids) of CNR.

Appendix

Plane and Oversimplified Archard Models. In two dimension, with the same assumption for the ratios m_{k+1}/m_k and R_{k+1}/R_k, we have

$$\frac{A_n}{A_{n-1}} = \left(\frac{A_0}{\lambda_0}\right)^{1/2^n} \frac{\sqrt{\pi}}{2^{2-1/2^n}} \frac{\Gamma\left(\frac{3}{2} - \frac{1}{2^{n+1}}\right)}{\Gamma\left(2 - \frac{1}{2^{n+1}}\right)} \left(\frac{8}{\pi}\gamma^{1-D}\right)^{1-2^{-n}} \tag{8}$$

which, in the limit for $n \rightarrow \infty$ reduces to γ^{1-D}. Evaluating the fractal dimension as in the three-dimensional case the limiting fractal dimension is

$$d_A = 2 - D. \tag{9}$$

An oversimplified model of a surface could be imagined as having at scale n a number γ^{2n} of equal asperities of radius R_n not necessarily equally distributed leads to a fractal dimension

$$d_A = \frac{2}{3}(4 - D) \tag{10}$$

which ranges from 4/3 to 2. However, the relation real contact area versus load is still Hertzian at all scales, which is contradicting Amonton's law.

References

[1] Bowden, F. P., and Tabor, D., 1950, *Friction and Lubrication of Solids*, Clarendon Press, Oxford, UK.
[2] Lincoln, B., 1953, Nature (London), **172**, p. 169.
[3] Archard, J. F., 1953, ''Elastic Deformation and the Contact of Surfaces,'' Nature (London), **172**, No. 4385, pp. 918–919.
[4] Archard, J. F., 1957, ''Elastic Deformation and the Laws of Friction,'' Proc. R. Soc. London, Ser. A, **A243**, pp. 190–205.
[5] Greenwood, J. A., and Williamson, J. B. P., 1966, ''The Contact of Nominally Flat Surfaces,'' Proc. R. Soc. London, Ser. A, **A295**, pp. 300–319.
[6] Johnson, K. L., 2000, ''The contribution of Micro-Nano-Tribology to the Interpretation of Dry Friction,'' Proc. Inst. Mech. Eng., **214**, Part C, pp. 11–22.
[7] Mandelbrot, B. B., 1982, *The Fractal Geometry of Nature*, Freeman, San Francisco.
[8] Russ, J. C., 1994, *Fractal Surfaces*, Plenum Press, New York.
[9] Majumdar, A., and Bhushan, B., 1990, ''Role of Fractal Geometry in Roughness Characterization and Contact Mechanics of Surfaces,'' ASME J. Tribol., **112**, pp. 205–216.
[10] Majumdar, A., and Bhushan, B., 1995, Characterization and Modeling of Surface Roughness and Contact Mechanics. *Handbook of Micro/Nano Tribology*, CRC Press, Boca Raton, FL, pp. 109–165.
[11] Borri-Brunetto, M., Carpinteri, A., and Chiaia, B., 1998, ''Lacunarity of the Contact Domain Between Elastic Bodies With Rough Boundaries,'' *Probamat-21st Century: Probabilities and Materials*, G. Frantziskonis, ed., Kluwer, Dordrecht. pp. 45–64.
[12] Barber, J. R., and Ciavarella, M., 2000, ''Contact Mechanics,'' special issue ''Recent Trends in Solid Mechanics,'' Int. J. Solids Struct., **37**, pp. 29–43.
[13] Ciavarella, M., Demelio, G., Barber, J. R., and Jang, Y. H., 2000, ''Linear Elastic Contact of the Weierstrass Profile,'' Proc. R. Soc. London, Ser. A, **456**, No. 1994, pp. 387–405.
[14] Greenwood, J. A., and Tripp, J. H., 1967, ''The Elastic Contact of Rough Spheres,'' ASME J. Appl. Mech., **34**, p. 153.
[15] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.

Transfer Matrix Method of Wave Propagation in a Layered Medium With Multiple Interface Cracks: Antiplane Case

Y.-S. Wang
Institute of Engineering Mechanics, Northern Jiaotong University, Beijing 100044, P. R. China
e-mail: yswang@center.njtu.edu.cn

D. Gross
Institute of Mechanics, TU Darmstadt, Hochschulstr. 1, D-64289 Darmstadt, Germany

The paper develops a universal method for SH-wave propagation in a multilayered medium with an arbitrary number of interface cracks. The method makes use of the transfer matrix and Fourier integral transform techniques to cast the mixed boundary value problem to a set of Cauchy singular integral equations of the first type which can be solved numerically. The paper calculates the dynamic stress intensity factors for some simple but typical examples. [DOI: 10.1115/1.1360180]

1 Introduction

Wave propagation in a layered medium is of both theoretical and practical importance in such fields as composite materials, geophysics, etc. Since the 1970s, the problems of wave scattering from an interface crack between two bonded elastic solids have been widely investigated by many authors, for instance, Loeber and Sih [1,2], Takai, Shindo, and Atsumi [3] Srivastava, Palaiya, and Karaulia [4], and Boström [5] for a mode III Griffith or penny-shaped interface crack between two half-spaces; Srivastava, Gupta, and Palaiya [6,7], and Qu [8,9] for a Mode I or II interface crack; and Neerhaff [10] Kundu [11], Li and Tai [12], Yang and Bogy [13], and Gracewski and Bogy [14,15] for a Griffith interface crack of Mode I, II, and III in a layered plate or a layered half-space. However, one may note that only a few papers have considered multiple interface cracks. The published results are limited to some simple cases. Kundu [16] first discussed the interaction between two interface cracks in a layered half-space under antiplane transient loading, and then in a three-layered plate ([17]). Zhang [18,19] analyzed the SH-wave propagating through a periodic array of interface cracks between two bonded half-spaces. If the medium is composed of multiple layers and, furthermore, if cracks may occur in any interface with an arbitrary number, the associated wave propagation problems will become more difficult. Even by using numerical methods such as the finite element method and boundary element method, the problems cannot be solved easily. Here, in the present paper, we develop a universal method for wave propagation in a multilayered medium with multiple cracks distributed in different interfaces. The method makes use of the transfer matrix and singular integral equation techniques. As a preliminary analysis, we treat the SH-wave motion in this paper. But the method can be extended to the in-plane case in a straightforward manner.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 15, 1999; final revision, Aug. 10, 2000. Associate Editor: R. C. Benson.

2 Transfer Matrix and Dual Integral Equations

Consider the problem shown in Fig. 1. An elastic material is composed of $N+1$ layers bonded through N interfaces. Multiple Griffith cracks are distributed on $m(\leqslant N)$ interfaces. The interface between the rth and $(r+1)$th layer is denoted as the rth interface. We suppose that there are $n(p)$ collinear cracks on $r(p)$th interface. Take the x-axis along the 1st interface, and denote the x-coordinates of the tips of crack L_{pq} as a_{pq} and b_{pq} $(p=1\sim m,\ q=1\sim n(p))$ and the y-coordinate of the rth interface as $y=h_r$. This paper will consider the propagation of harmonic SH waves with frequency ω in such a layered medium. The harmonic term $e^{-i\omega t}$ will be omitted throughout paper.

We decompose the total displacement and stress wave fields $\{u, \sigma\}$ as the sum of the fields without cracks $\{u^{(0)},\sigma^{(0)}\}$ and those due to the scattering of the cracks $\{u^{(s)},\sigma^{(s)}\}$, i.e., $\{u,\sigma\}=\{u^{(0)},\sigma^{(0)}\}+\{u^{(s)},\sigma^{(s)}\}$, where $\{u^{(0)},\sigma^{(0)}\}$ can be obtained by the classical transfer matrix method or by other methods (cf. [20]). The following analysis will be focused on the solution of $\{u^{(s)},\sigma^{(s)}\}$. Without confusion, we omit the superscript (s).

The Helmholtz equation for SH-wave motion in the rth layer is

$$\nabla^2 w_r + K_{Tr}^2 w_r = 0, \quad r=1\sim N+1 \tag{1}$$

where w_r is the displacement component in z-direction; $K_{Tr}=\omega/C_{Tr}$ with $C_{Tr}=\sqrt{\mu_r/\rho_r}$ is the shear wave velocity; μ_r and ρ_r are, respectively, the shear modulus and mass density. We denote the displacement discontinuity on the pth interface as Δw_p which may be expressed as

$$\nabla w_p = \sum_{q=1}^{n(p)} \Delta w_{pq}[H(x-a_{pq})+H(x-b_{pq})],$$

with Δw_{pq} being the unknown tearing displacement of the crack L_{pq} $(q=1\sim n(p),\ p=1\sim m)$ and $H(\,)$ the Heaviside function. Then the boundary conditions may be written as

$$\tau_{yzr}=0, \quad y=h_0, h_{N+1}, \quad r=1,N+1 \tag{2}$$

$$\tau_{yzr}-\tau_{yzr+1}=0, \quad y=h_r, \quad r=1\sim N \tag{3}$$

$$w_r - w_{r+1} = \Delta w_p \delta_{rr(p)}, \quad y=h_r, \quad r=1\sim N \tag{4}$$

$$\tau_{yzr}=\tau_{yzr+1}=-\tau_{yzr}^{(0)}=-\tau_{yzr+1}^{(0)}, \quad x\in L_{pq}, \quad y=h_r, \quad r=r(p) \tag{5}$$

where $\delta_{rr(p)}$ is Kroneck symbol.

Applying Fourier integral transform to (1) with respect to x, we obtain its solution in the transformed space, which is written in the matrix form as

$$\{S_r\}=[T_r(y)]\{C_r\}, \quad r=1\sim N+1 \tag{6}$$

where

$$\{S_r\}=\{\overline{w}_r, \overline{\tau}_{yzr}\}, \quad \{C_r\}=\{C_{1r}, C_{2r}\}^T, \quad [T_r(y)]=[T_r^0][E_r(y)],$$

with

$$[T_r^0]=\begin{bmatrix} 1 & 1 \\ -\mu_r\beta_r & \mu_r\beta_r \end{bmatrix}, \quad [E_r(y)]=\begin{bmatrix} e^{-\beta_r y} & 0 \\ 0 & e^{\beta_r y} \end{bmatrix}$$

and $\beta_r=(s^2-K_{Tr}^2)^{1/2}$ of which the branch should be determined such that

$$\beta_r=(s^2-K_{Tr}^2)^{1/2}, \quad |s|\geqslant K_{Tr}; \beta_r=-i(K_{Tr}^2-s^2)^{1/2}, \quad |s|<K_{Tr}. \tag{7}$$

The bars appearing in above equations indicate the Fourier integral transforms, s is the parameter of the integral transforms, and C_{1r}, C_{2r} are undetermined functions of s.

The integral transforms of boundary conditions (2)–(4) can be written as

$$\{C_1\}=\{B\}C_{21}, \quad \{C_{N+1}\}=\{X\}C_{1,N+1}, \tag{8}$$

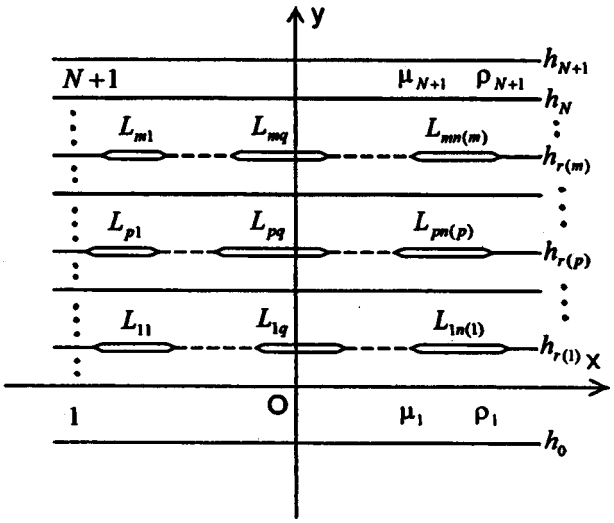


Fig. 1 A multilayered medium with multiple interface cracks

$$\{S_r\}-\{S_{r+1}\}=\{\Delta S_r\},\quad y=h_r,\quad r=1\sim N \tag{9}$$

where

$$\{B\}=\{e^{2\beta_1 h_0},1\}^T,\quad \{X\}=\{1,e^{-2\beta_{N+1}h_{N+1}}\}^T,$$
$$\{\Delta S_r\}=\{\Delta\overline{w}_p\delta_{rr(p)},0\}^T,\quad p=1\sim m.$$

The above equations involve the cases of $h_0\to-\infty$ and/or $h_{N+1}\to+\infty$. Equation (9) is a recurrence relation. Substituting (6) into this relation, we can express $\{C_r\}$ with $\{\Delta S_{r(p)}\}$ as

$$\{C_1\}=\sum_{p=1}^{m}[\overline{E}_{r(p)}]\{\Delta S_{r(p)}\}, \tag{10}$$

$$\{C_r\}=\sum_{p=1}^{m}([\overline{L}_{rr(p)}]+[\overline{K}_{rr(p)}]H(r-r(p)-1))\{\Delta S_{r(p)}\},$$
$$r=2\sim N+1 \tag{11}$$

where

$$[\overline{L}_{ri}]=[\overline{W}_r]^{-1}[\overline{E}_i];\ [\overline{K}_{ri}]=-[\overline{W}_r]^{-1}[L_i];$$
$$[\overline{E}_i]=\{B\}[1,0][\overline{W}]^{-1}[L_i];\ [L_i]=[\overline{W}_i][T_i(h_i)]^{-1};$$
$$[\overline{W}]=\{B\}[1,0]-[\overline{W}_{N+1}]\{X\}[0,1];$$
$$[\overline{W}_r]=[W_2]\cdots[W_r],\quad r>1;\ [\overline{W}_1]=[\mathrm{I}];$$
$$[W_{r+1}]=[T_r(h_r)]^{-1}[T_{r+1}(h_r)].$$

Substitution of (10) and (11) into (6) yields

$$\{S_r\}=\sum_{p=1}^{m}[M_{rr(p)}]\{\Delta S_{r(p)}\},\quad r=1\sim N+1 \tag{12}$$

where we have denoted

$$[M_{rr(p)}]=[T_1(y)][\overline{E}_{r(p)}],\quad r=1$$
$$[M_{rr(p)}]=[T_r(y)]([\overline{L}_{rr(p)}]+[\overline{K}_{rr(p)}]H(r-r(p)-1)),\quad r>1 \tag{13}$$

which is the transfer matrix of the multiple layered medium with interface cracks. Write the matrix $[M_{rr(p)}]$ as

$$[M_{rr(p)}]=\begin{bmatrix} * & * \\ m_{rr(p)} & * \end{bmatrix}, \tag{14}$$

where *s are the other elements of the matrix $[M_{rr(p)}]$ which are of no use in the following analysis. Then we have

$$\overline{\tau}_{yzr}=\sum_{p=1}^{m}m_{rr(p)}\Delta\overline{w}_p. \tag{15}$$

Inserting the inverse Fourier transform of (15) into the boundary condition (5), we obtain

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\sum_{p=1}^{m}\sum_{q=1}^{n(p)}\hat{m}_{kp}\Delta\overline{w}_{pq}e^{-isx}ds=-\tau_{yz}^{(0)}(x,h_{r(k)}), \tag{16}$$

where $\hat{m}_{kp}=m_{r(k)r(p)}|_{y=h_{r(k)}}$, $x\in L_{kj}$, $j=1\sim n(k)$, $k=1\sim m$. It is straightforward that the following relation holds:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\Delta\overline{w}_{kj}e^{-isx}ds=0,\quad x\notin L_{kj}. \tag{17}$$

Equations (16) and (17) are dual integral equations which will be transformed to a set of Cauchy singular integral equations in the next section.

3 Singular Integral Equations and Numerical Solution

Introduce the dislocation density function of the crack L_{pq},

$$\phi_{pq}(x)=\frac{\partial}{\partial x}(\Delta w_{pq}),\quad q=1\sim n(p),\quad p=1\sim m \tag{18}$$

By considering the differential properties of the Fourier transform, it is not difficult to obtain

$$\Delta\overline{w}_{pq}=is^{-1}\int_{L_{pq}}\phi_{pq}(u)e^{isu}du, \tag{19}$$

which, when substituted to (16) and (17), yields

$$\frac{i}{2\pi}\int_{-\infty}^{\infty}\sum_{p=1}^{m}\sum_{q=1}^{n(p)}s^{-1}\hat{m}_{kp}\int_{L_{pq}}\phi_{pq}(u)e^{is(u-x)}duds$$
$$=-\tau_{yz}^{(0)}(x,h_{r(k)}), \tag{20}$$

$$\int_{L_{kj}}\phi_{kj}(u)du=0 \tag{21}$$

where $x\in L_{kj}$, $j=1\sim n(k)$, $k=1\sim m$.

It is found that $s^{-1}\hat{m}_{kp}$ is an antisymmetric function of s and has the following asymptotic behavior as $s\to+\infty$,

$$s^{-1}\hat{m}_{kp}\to\begin{cases}-\dfrac{\mu_{r(k)}\mu_{r(k)+1}}{\mu_{r(k)}+\mu_{r(k)+1}}\triangleq\gamma_k, & k=p\\ 0, & k\neq p\end{cases}. \tag{22}$$

This result can be easily proved (e.g., by using Mathematica, Version 3.0) for fixed values of N, m, and p.

Due to (22), special care must be taken in interchanging the two integrations in (20). However, if we consider the following relation

$$\int_{-\infty}^{\infty}\mathrm{sgn}(s)e^{is(u-x)}ds=\frac{2i}{u-x}, \tag{23}$$

and denote

$$P_{kp}(u,x)=\frac{i}{2\pi}\int_{-\infty}^{\infty}[s^{-1}\hat{m}_{kp}-\gamma_k\delta_{kp}\,\mathrm{sgn}(s)]e^{is(u-x)}ds$$
$$=-\frac{1}{\pi}\int_0^{\infty}[s^{-1}\hat{m}_{kp}-\gamma_k\delta_{kp}]\sin[s(u-x)]ds, \tag{24}$$

we can transform (20) into Cauchy singular integral equations:

$$-\frac{\gamma_k}{\pi}\sum_{q=1}^{n(k)}\int_{L_{kq}}\frac{\phi_{kq}(u)}{u-x}du+\sum_{p=1}^{m}\sum_{q=1}^{n(p)}\int_{L_{pq}}\phi_{pq}(u)P_{kp}(u,x)du$$
$$=-\tau_{yz}^{(0)}(x,h_{r(k)}). \tag{25}$$

$P_p(u,x)$ is a Fredhelm kernel which has no singularity except when $\hat{M}_p(n)$ becomes infinite in some cases which we will discuss later.

By introducing the substitutions

$$\begin{cases} x=c_{kj}\xi+d_{kj}, \quad u=c_{kq}\eta+d_{kq} \\ \Phi_{pq}(\eta)=\phi_{pq}(c_{pq}\eta+d_{pq}) \\ Q_{pq}(\eta,\xi)=-\dfrac{1}{\gamma_k}P_{kp}(c_{pq}\eta+d_{pq},c_{kj}\xi+d_{kj}) \\ \qquad\qquad +\dfrac{1}{\pi}\dfrac{\delta_{kp}(1-\delta_{qj})}{c_{kq}\eta-c_{kj}\xi+d_{kq}-d_{kj}}, \end{cases} \tag{26}$$

with $c_{kj}=(b_{kj}-a_{kj})/2$ and $d_{kj}=(a_{kj}+b_{kj})/2$, Eq. (25) can be further converted to standard Cauchy singular integral equations

$$\frac{1}{\pi}\int_{-1}^{1}\frac{\Phi_{kj}(\eta)}{\eta-\xi}d\eta+\sum_{p=1}^{m}\sum_{q=1}^{n(p)}\int_{-1}^{1}c_{pq}\Phi_{pq}(\eta)Q_{pq}(\eta,\xi)d\eta$$
$$=\frac{1}{\gamma_k}\tau_{yz}^{(0)}(c_{kj}\xi+d_{kj},h_{r(k)}). \tag{27}$$

Meanwhile, (21) becomes

$$\int_{-1}^{1}\Phi_{kj}(\eta)d\eta=0. \tag{28}$$

The above equations can be solved numerically by the method developed by Erdogan and Gupta [21]. Set

$$\Phi_{kj}(\eta)=\frac{F_{kj}(\eta)}{\sqrt{1-\eta^2}}. \tag{29}$$

Then (27) and (28) reduce to

$$\frac{1}{M}\sum_{s=1}^{M}\left[\frac{F_{kj}(\eta_s)}{\eta_s-\eta_t}+\pi\sum_{p=1}^{m}\sum_{q=1}^{n(p)}c_{pq}Q_{pq}(\eta_s,\xi_t)F_{pq}(\eta_s)\right]$$
$$=\frac{1}{\gamma_k}\tau_{yz}^{(0)}(c_{kj}\xi_t+d_{kj},h_{r(k)}), \tag{30}$$

$$\frac{\pi}{M}\sum_{s=1}^{M}F_{kj}(\eta_s)=0, \tag{31}$$

where $\eta_s=\cos(\pi(2s-1)/2M)$, $\xi_t=\cos(\pi t/M)$, $t=1\sim M-1$; M is the number of the discrete points of $F_{kj}(\eta)$ in $(-1,1)$; and $j=1\sim n(k)$, $k=1\sim m$.

It is noted that difficulties may arise in evaluation of the semi-infinite integrals (24) because of the possible simple poles of the integrands along the integral path and located between $\min(K_{Tr})$ and $\max(K_{Tr})$. These poles correspond to the general Love-type surface waves. One should note that the path of the integration along the real s-axis is indeed the limit of the path as it approaches the real s-axis from below ([14,15]). Based on this fact, two different ways have been developed for dealing with these poles in the integrations. Kundu [22] developed a technique of removing the singularities from the integrands. In result, the original integrals are divided into two parts—the residues of the integrands at the poles and Cauchy principal integrals. The other technique deforms the contour of integration below the real axis so that no poles occur on the path of integration (cf. [13–15]). In this brief note we will employ the second method because it is easy for calculation.

4 Examples

The method developed above is quite general and can be applied to many complicated problems. In this brief note, we only present numerical results for some simple but typical examples. Our attention is focused on the dynamic stress intensity factors which are defined as

$$\begin{cases} K_{kj}^{+}=\lim\limits_{x\to b_{kj}^{+}}[\sqrt{2(x-b_{kj})}\tau_{yz}(x,h_{r(k)})] \\ K_{kj}^{-}=\lim\limits_{x\to a_{kj}^{-}}[\sqrt{2(a_{kj}-x)}\tau_{yz}(x,h_{r(k)})]. \end{cases} \tag{32}$$

The numerical results may be obtained by the following formula (cf. [23]):

$$K_{kj}^{\pm}=-\gamma_k\sqrt{c_{kj}}F_{kj}(\pm 1). \tag{33}$$

Example 1 Two Bonded Half-Spaces. In order to verify the validity and accuracy of our solution, we calculate the simplest case—an interface crack of length $2c$ lies between two dissimilar half-spaces and compare our results to those of Loeber and Sih [1]. The material constants are taken as $\mu_1/\mu_2=2$ and $\rho_1/\rho_2=1$. A harmonic SH-wave of the general form

$$w^{(i)}=A_0e^{iK_{T1}(x\sin\theta_0+y\cos\theta_0)-i\omega t}, \tag{34}$$

strikes the interface normally ($\theta_0=0$ deg) from medium 1 (see the sketch in Fig. 2). A_0 and θ_0 are the amplitude and incident angle, respectively. In computation we choose $M=30$ in Eq. (30). The variation of the dynamic stress intensity factor with normalized frequency $K_{T1}c$ is shown in Fig. 2, where the dynamic stress intensity factor is normalized by $\overline{\tau_0}\sqrt{c}$ with $\overline{\tau_0}$ being the shearing stress along the interface without the crack. The results of Loeber and Sih [1] are also plotted in Fig. 2. Good agreement between our results and theirs is observed.

Next we consider a more complex example—three cracks with the same length $2c$ and the same distance $2d$ between cracks lie on the interface (see the sketch in Fig. 3). The ratio of c and d is set to 1:1.25. The results are shown in Fig. 3, where the dynamic stress intensity factors are normalized by $\tau_0\sqrt{c}$ with $\tau_0=\mu_1A_0K_{T1}$. Resonance is observed at the lower frequencies, and the resonance peaks for the inner crack tips (tips 1 and 2) are more pronounced than that for the outer ones (tips 3). The later oscillates more pronouncedly at the higher frequencies. In this example we also take $M=30$. In order to check the convergence of the solution we calculate the dynamic stress intensity factor at $K_{T1}c=1$ for the crack tip 1 by choosing different values of M and list the results in Table 1 (see the first line for the present example). It is shown that $M=30$ can give good accuracy.

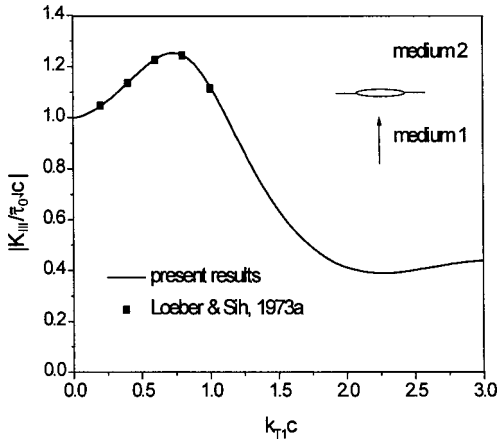


Fig. 2 Dynamic stress intensity factor for one interface crack in a system of two bonded half-spaces

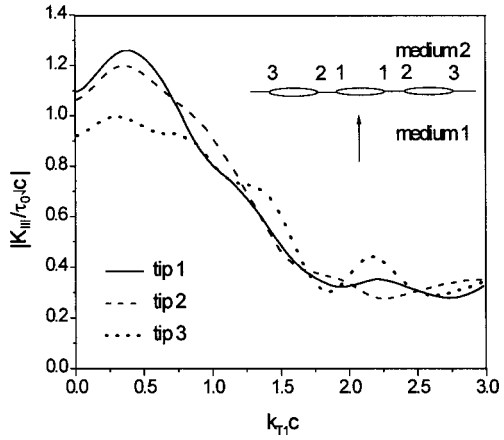


Fig. 3 Dynamic stress intensity factor for three interface cracks in a system of two bonded half-spaces

Table 1 Normalized dynamic stress intensity factors, $K_{III}/\tau_0\sqrt{c}$, at $K_{T1}c=1$ for crack tips 1 in the three examples by choosing different values of M

	$M=20$	$M=30$	$M=40$	$M=60$
Example 1	0.804341	0.804396	0.804422	0.804443
Example 2	0.684742	0.684793	0.684805	0.684822
Example 3	0.910096	0.910223	0.910245	0.910267

Example 2 A Layered Half-Space. Again consider the above example, but the upper half-space is of finite thickness h (see the sketch in Fig. 4). Figure 4 illustrates the normalized dynamic stress intensity factors $K_{III}/\tau_0\sqrt{c}$ versus the normalized frequency $K_{T1}c$ for $h/c=1$. The effects of the free surface can be observed in the figure. As the frequency increases from zero the dynamic stress intensity factors also increase from zero and reach resonant peak values at rather low frequencies. Contrary to Example I, the outer crack tips involve more pronounced peaks in the present case. It is worthy of note that a zero value of the dynamic stress intensity factors appears at a higher frequency ($k_{T1}c\approx 2.24$). All these features may be explained by the reflection of waves between the free surface and the interface.

Example 3 Two Half-Spaces Bonded Through a Layer Consider two half-spaces bonded through an interlayer of thickness h (see the sketch in Fig. 5). There are two cracks lying on each interface. The crack size and distribution are the same as those in Example 1. The material constants are taken as $\mu_1:\mu_2:\mu_3=2:1:2$ and $\rho_1=\rho_2=\rho_3$; and $h:c$ is set to 1 : 1. The incident SH-wave with the form (34) propagates normally ($\theta_0=0$ deg) to the interlayer in material 1. The normalized dynamic stress intensity factors $K_{III}/\tau_0\sqrt{c}$ for the four cracks are plotted versus $K_{T1}c$ in Fig. 5. As the frequency increases the dynamic stress intensity factors first decrease and then increase to peak values. Generally the cracks on the lower interface have higher dynamic stress intensity factors than those on the upper interfaces.

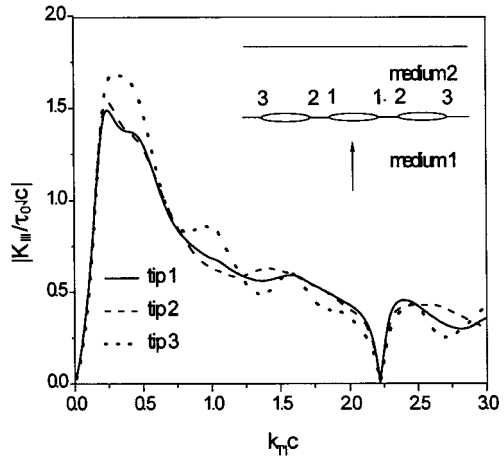


Fig. 4 Dynamic stress intensity factor for three interface cracks in a system of a layered half-space

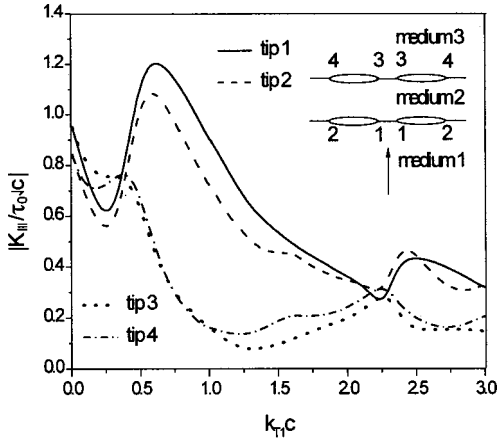


Fig. 5 Dynamic stress intensity factor for four interface cracks in a system of two half-spaces bonded through a layer

In the last two examples, we take $M=30$ as in the first example. The convergence is shown in Table 1.

Finally we mention that the method developed in this brief note is universal and can be used to solve many complex problems. However, we only give some simple examples. A lot of topics based on this piece of work are left for further investigation. The in-plane case that is more complicated will be explored in subsequent works.

Acknowledgment

The work was finished during the first author's stay in TU-Darmstadt. Support of the Alexander von Humboldt Foundation is gratefully acknowledged.

References

[1] Loeber, J. F., and Sih, G. C., 1973, ''Transmission of Anti-Plane Shear Waves past an Interface Crack in Dissimilar Media,'' Eng. Fract. Mech., **5**, pp. 699–725.
[2] Loeber, J. F., and Sih, G. C., 1973, ''Torsional Waves Scattering About a Penny-Shaped Crack Lying on a Bimaterial Interface,'' *Dynamic Crack Propagation*, G. C. Sih, ed., Noordhoff, Leydon, pp. 513–528.
[3] Takai, M., Shindo, Y., and Atsumi, A., 1982, ''Diffraction of Transient Horizontal Shear Waves by a Finite Crack at the Interface of Two Bonded Dissimilar Elastic Solids,'' Eng. Fract. Mech., **16**, pp. 799–807.
[4] Srivastava, K. N., Palaiya, R. M., and Karaulia, D. S., 1980, ''Interaction of Antiplane Shear Waves by a Griffith Crack at the Interface of Two Bonded Dissimilar Elastic Half-Spaces,'' Int. J. Fract., **16**, pp. 349–358.
[5] Boström, A., 1987, ''Elastic Wave Scattering From an Interface Crack: Anti-Plane Strain,'' ASME J. Appl. Mech., **54**, pp. 503–508.
[6] Srivastava, K. N., Gupta, O. P., and Palaiya, R. M., 1978, ''Interaction of Elastic Waves in Two Bonded Dissimilar Elastic Half-Spaces Having Griffith Crack at Interface—I,'' Int. J. Fract., **14**, pp. 145–154.
[7] Srivastava, K. N., Palaiya, R. M., and Gupta, O. P., 1979, ''Interaction of Longitudinal Wave With a Penny-Shaped Crack at the Interface of Two Bonded Dissimilar Elastic Solids—II,'' Int. J. Fract., **15**, pp. 591–599.
[8] Qu, J., 1994, ''Interface Crack Loaded by a Time-Harmonic Plane Wave,'' Int. J. Solids Struct., **31**, pp. 329–345.
[9] Qu, J., 1995, ''Scattering of Plane Waves from an Interface Crack,'' Int. J. Eng. Sci., **33**, pp. 179–194.
[10] Neerhaff, F. L., 1979, ''Diffraction of Love Waves by a Stress-Free Crack of Finite Width in the Plane Interface of a Layered Composite,'' Appl. Sci. Res., **35**, pp. 237–249.

[11] Kundu, T., 1986, ''Transient Response of an Interface Crack in a Layered Plate,'' ASME J. Appl. Mech., **53**, pp. 579–786.
[12] Li, D. H., and Tai, W. H., 1991, ''Elastodynamic Response of an Interface Crack in a Layered Composite Under Anti-Plane Shear Impact Load,'' Eng. Fract. Mech., **39**, pp. 687–693.
[13] Yang, H. J., and Bogy, D. B., 1985, ''Elastic Waves Scattering From an Interface Crack in a Layered Half-Space,'' ASME J. Appl. Mech., **52**, pp. 42–50.
[14] Gracewski, S. M., and Bogy, D. B., 1986, ''Elastic Wave Scattering From an Interface Crack in a Layered Half-Space Submerged in Water: Part I. Applied Tractions at the Liquid-Solid Interface,'' ASME J. Appl. Mech., **53**, pp. 326–332.
[15] Gracewski, S. M., and Bogy, D. B., 1986, ''Elastic Wave Scattering From an Interface Crack in a Layered Half-Space Submerged in Water: Part II. Incident Plane Waves and Bounded Beams,'' ASME J. Appl. Mech., **53**, pp. 333–338.
[16] Kundu, T., 1987, ''Transient Response Between Two Interface Cracks at the Interface of a Layered Half Space,'' Int. J. Eng. Sci., **25**, pp. 1427–1439.
[17] Kundu, T., 1988, ''Dynamic Interaction Between Two Interface Cracks in a Three-Layered Plate,'' Int. J. Solids Struct., **24**, pp. 27–39.
[18] Zhang, Ch., 1991, ''Dynamic Stress Intensity Factors for Periodically Spaced Collinear Antiplane Shear Cracks Between Dissimilar Media,'' Theor. Appl. Fract. Mech., **15**, pp. 219–277.
[19] Zhang, Ch., 1991, ''Reflection and Transmission of SH Wave by a Periodic Array of Interface Cracks,'' Int. J. Eng. Sci., **29**, pp. 481–491.
[20] Kennett, B. L. N., 1983, *Seismic Wave Propagation in Stratified Media*, Cambridge University Press, Cambridge, UK.
[21] Erdogan, F., and Gupta, G. D., 1972, ''On the Numerical Solution of Singular Integral Equations,'' Q. J. Appl. Math., **29**, pp. 525–539.
[22] Kundu, T., 1985, ''Elastic Waves in a Multilayered Solid due to a Dislocation Source,'' Wave Motion, **7**, pp. 459–471.
[23] Wang, Y. S., and Wang, D., 1996, ''Scattering of Elastic Waves by a Rigid Cylindrical Inclusion Partially Debonded From Its Surrounding Matrix—I. SH case,'' Int. J. Solids Struct., **33**, pp. 2789–2815.

Wave Propagation in Laminated Composite Plates Using Higher Order Theory

M. R. Chitnis
Graduate Student

Y. M. Desai
Assistant Professor

T. Kant
Professor

Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India

A higher order displacement based formulation has been developed to investigate wave propagation in fiber-reinforced polymer composite laminated (FRPCL) plates. The formulation has been applied, as an illustration, to a plate made up of transversely isotropic laminae with the axes of symmetry lying in the plane of the lamina. Results for the plane as well as the antiplane strain cases are shown to be in excellent agreement with the exact solutions for isotropic and transversely isotropic single layered plates. Also, numerical results have been obtained for crossply (0 deg/90 deg/0 deg/90 deg) laminated composite plates, which agree very well with the previously published numerical results. The formulation can be employed to expeditiously investigate the dispersion characteristics of waves propagating in a plate with an arbitrary number of anisotropic laminae. [DOI: 10.1115/1.1352062]

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 2, 1999; final revision, Oct. 6, 2000. Associate Editor: A. K. Mal.

Introduction

Wave propagation through laminated composite plates has been investigated extensively in the past by many researchers. A few exact solutions have been reported on wave propagation through isotropic plates (e.g., [1,2]). Datta et al. [3] investigated harmonic wave propagation through a laminated infinite medium in which each lamina was divided into sublayers and the displacement distribution through the thickness of each layer was approximated by cubic polynomial interpolation functions. The displacements and tractions at the interfaces of the adjoining sublayers were considered as generalized co-ordinates. Liu et al. [4] presented a numerical method for investigating wave propagation in anisotropic laminated plates. The displacement field in each sublayer of a lamina was approximated by a linear expansion in thickness direction and by a series in the width direction. Lih and Mal [5] developed a multiple transform technique coupled with a numerical evaluation scheme to analyze a composite laminate subjected to dynamic loads.

A simple approach with lesser complexities in formulation has been followed in the present work to solve the problem of wave propagation in laminated composite plates. A higher order cubic variation of displacement function is assumed through the thickness of each lamina in the laminate whereas displacement in longitudinal direction has been modeled exactly by an exponential function. Variational principle has been applied to each layer and the stiffness and mass matrices have been assembled by employing continuity of displacements and rotations at interfaces, thus making the formulation C^1 continuum. Although variation of displacement field has also been modeled using cubic polynomial in methods discussed above, the present method works with simple displacement based formulation with displacements and rotations as degrees-of-freedom instead of tractions and still produces equally accurate results. An isotropic plate has been analyzed and the results compared with the analytical results to prove the accuracy and validity of the numerical technique developed. Subsequently, a crossply laminated plate has been analyzed and the results discussed.

Formulation

A higher order displacement based formulation has been presented to analyze wave propagation through composite laminated plate. Displacement field is defined by

$$\begin{Bmatrix} u \\ w \end{Bmatrix} = [X]\{q\} \tag{1}$$

where

$$[X] = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & X_1 & X_2 & X_3 & X_4 \end{bmatrix} \tag{2}$$

and

$$\{q\}^t = [u_1 \;\; u_2 \;\; \theta_{x1} \;\; \theta_{x2} \;\; w_1 \;\; w_2 \;\; \theta_{z1} \;\; \theta_{z2}]. \tag{3}$$

Here, u_i and w_i, $i=1,2$ are the generalized displacements along the x and z-directions, and θxi, θzi, $i=1,2$, are rotation and transverse strain, respectively, at $z=(-1)^i h$. Moreover, X_i, $i=1,\ldots 4$ are expressed as cubic polynomials.

The dispersion equation for wave propagation in a lamina of unit width can be written as

$$[K]-\omega^2[M]=0 \tag{4}$$

where

$$[K]=[K_{11}]+i\lambda([K_{12}]-[K_{12}]^t)+\lambda^2[K_{22}]. \tag{5}$$

Here

$$[K_{ab}]=\int_{-h}^{h}([B_a]^t[C][B_b])dz, \quad a,b=1,2$$

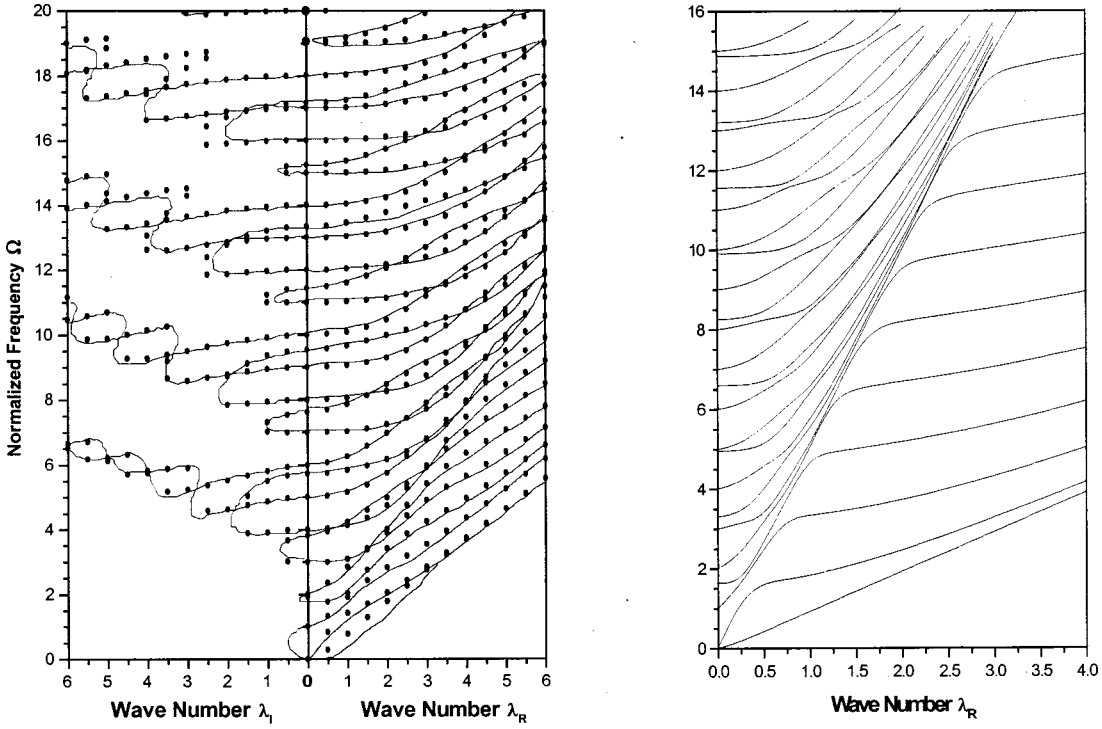


Fig. 1 (*a*) **Comparison of results obtained by the present method with the analytical results presented by Mindlin [1]** (*b*) **Results obtained by present method for crossply composite laminated plate of (0 deg/90 deg/0 deg/90 deg) lay-up**

and

$$[M]=\int_{-h}^{h}([X]^t\rho[X])dz.$$

Further, $[B_a]$, $a=1,2$ are strain displacement relationship matrices and $[C]$ is the constitutive matrix.

The general solution $\{q\}=\{q_0\}e^{i\lambda x}\sin(\omega t)$ has been employed to arrive at the dispersion Eq. (4). Here, $\{q_o\}$ is the amplitude vector, ω is circular frequency, and $\lambda=\zeta\pi/H$, where ζ is the wave number and H is the total thickness of the laminated plate.

Equation (4) can be solved as an eigenvalue problem and frequencies ω can be obtained for given value of λ.

Numerical Investigation

A general-purpose program has been developed on the basis of theoretical formulation discussed above, for determining natural frequencies of vibration of a composite laminate subjected to a time-harmonic wave. The program has been developed to deal with both real and imaginary wave numbers. The numerical method proposed has been validated by comparing results obtained with the analytical results presented by Mindlin [1] for an isotropic plate having properties as presented below:

thickness of plate: 25.4 mm, mass density: 2.7688×10^4 kg/m^3

$C_{11}=95.590$ GPa, $C_{13}=42.950$ GPa, $C_{33}=95.590$ GPa, $C_{55}=26.320$ GPa

The plate has been divided into 25 sublayers. The results have been presented graphically in Fig. 1(*a*) where the solid lines represent analytical results, whereas solid circles represent the results obtained by using the present formulation.

A crossply laminated plate of (0 deg/90 deg/0 deg/90 deg) lay-up and material properties as presented below has been solved wherein each lamina of the laminate was divided into 20 sublayers. The results obtained have been presented graphically in Fig. 1(*b*). Only frequencies for real wave numbers have been presented for brevity.

Thickness of each lamina: 6.35 mm, mass density: 1.0 kg/m^3

For 0 deg lamina, $C_{11}=146.83$ GPa, $C_{13}=4.08$ GPa, $C_{33}=15.99$ GPa, $C_{55}=5.86$ GPa

For 90 deg lamina, $C_{11}=15.99$ GPa, $C_{13}=4.08$ GPa, $C_{33}=146.83$ GPa, $C_{55}=5.86$ GPa

In Fig. 1(*a*,*b*) the natural frequencies of vibration and the wave number have been normalized to facilitate comparative study. Parameters Ω and ζ have been defined as

$$\Omega=\omega\frac{H}{\pi}\sqrt{\left(\frac{\rho}{C_{55}}\right)_{0^o}}\quad\text{and}\quad\zeta=\frac{\lambda H}{\pi}.$$

Here, ρ is the mass density, C_{55} is the in-plane shear modulus of elasticity of 0 deg lamina, and H is the total thickness of the plate. An excellent agreement of the results with the analytical results for isotropic plate shown in Fig. 1(*a*) demonstrates the accuracy of the proposed method. The dispersion curves shown in Fig. 1(*b*) for crossply plate differ considerably from those shown in Fig. 1(*a*) for isotropic plate. The cutoff frequencies are lowered in comparison with isotropic plate and in general the dispersion curves follow an asymptotic path. A thorough modal analysis is necessary for interpretation of these dispersion curves. Such analysis is not performed here for brevity, however.

Conclusions

A higher order displacement based formulation has been presented to analyze dispersion characteristics of guided waves in a laminated composite plate of infinite dimensions. The interlayer continuity of displacements and rotations has been imposed while assembling the stiffness and mass matrices of each layer, thus making the formulation C^1 continuous. The results for the isotropic plate have been shown to be in close agreement with the

analytical results. The proposed method has been demonstrated to be very simple and straightforward without any loss of accuracy in comparison with the available numerical methods. The results obtained for a crossply laminated composite plate have shown a considerable difference in comparison with the dispersion curves obtained for isotropic plate.

Acknowledgments

The work reported has been supported in part by a grant from the Ministry of Human Resource Development (Grant No. 98MH018).

References

[1] Mindlin, R. D., 1960, ''Waves and Vibrations in Isotropic Elastic Plates,'' *Proc. First Sympo. on Naval Struct. Mech.*, Pergamon, New York, pp. 199–232.

[2] Achenbach, J. D., 1976, *Wave Propagation in Elastic Solids*, North-Holland, New York.

[3] Datta, S. K., Shah, A. H., Bratton, R. L., and Chackraborty, T., 1988, ''Wave Propagation in Laminated Composite Plates,'' J. Acoust. Soc. Am., **83**, pp. 2020–2026.

[4] Liu, G. R., Tni, J., Watanabe, K., and Ohyoshi, T., 1990, ''Harmonic Wave-Propagation in Anisotropic Laminated Strips,'' J. Sound Vib., **139**, pp. 313–324.

[5] Lih, S., and Mal, A. K., 1995, On the Accuracy of Approximate Plane Theories for Wave-Field Calculations in Composite Laminates, Wave Motion, **21**, pp. 17–34.

Thickness Profiles for Rotating Circular Disks That Maximize Critical Speed

G. M. Warner
Mem. ASME

A. A. Renshaw
Mem. ASME

Department of Mechanical Engineering, Columbia University, M/C 4703, New York, NY 10027

[DOI: 10.1115/1.1360182]

1 Introduction

Radial thickness profiles of industrial circular saws often comprise a series of two or three uniform thickness annuli. Bird ([1]) claims that such thickness profiles increase the critical speed of a saw, which increases the operational speed range of the saw and helps maintain cutting straightness. The purpose of this brief note is to compute the optimal saw radial thickness profile for a fixed width cut and compare that profile to those currently in use.

Thickness variations of circular saws are bounded from above by the thickness of the carbide tips brazed to the saw teeth. Any thickness greater than the carbide tip thickness will rub against the wood during cutting, which will produce high cutting torques and burn and ruin the cut surface. Thickness variations are bounded from below by some minimum thickness capable of withstanding the in-plane stresses in the saw produced from cutting forces and rotation. Consequently, the numerical optimization routine needs only consider thickness variations between these two bounds. These bounds differentiate this optimization problem from those that have been previously studied. For example, Olhoff [2] and Thambiratnam and Thevendran [3] maximized the fundamental vibration frequency of a variable thickness, stationary, circular disk subject to a constraint on the disk volume. Optimal thickness profiles for minimum stress difference ([4]) and kinetic energy storage ([5]) have also been computed.

In this investigation, the optimal thickness profile for maximizing critical speed consists of two annuli, the inner one with thickness equal to the upper bound, the outer one with thickness equal to the stress-induced lower bound. The percentage increase in critical speed is a function of clamping ratio and the lower bound, and is greater than ten percent in some cases. However, while the radial profiles currently in use by industry are qualitatively similar to the optimal profile, the industrial profile examined here actually decreases the critical speed of the saw. This disparity suggests that critical speed may not be the critical design parameter for evaluating circular saw performance.

2 Modeling and Optimization Procedure

Modeling the saw as a Kirchhoff plate with in plane stresses, the dimensionless critical speed Ω_{cr} of an axisymmetric, centrally clamped, rotating circular saw is the lowest eigenvalue Ω of the eigenvalue problem derived by rendering the functional $R[w]$ stationary ([6])

$$R[w]=\frac{1}{2}\int \Omega^2 h(\sigma_r w_{,r}^2+\sigma_\theta w_{,\theta}^2/r^2-w_{,\theta}^2)+h^3\{(\nabla^2 w)^2 - 2(1-\nu)[w_{,rr}(w_{,r}/r+w_{,\theta\theta}/r^2)-((w_{,\theta}/r)_{,r})^2]\}dA \tag{1}$$

where $dA=rdrd\theta$ is the planar, unclamped area of the disk extending over $\kappa\leq r\leq 1$ and $0\leq\theta\leq 2\pi$, $w(r,\theta)$ is the transverse deflection of the disk, $h(r)$ is the disk thickness, σ_r and σ_θ are the radial and hoop stresses for unit rotation speed, ν is Poisson's ratio, ∇^2 is the Laplacian operator, and a comma indicates partial differentiation.

Table 1 Summary of optimal thickness profile designs and their properties

Clamping Ratio, κ	Minimum Thickness, h_{min}	Transition Radius, r_{opt}	Critical Speed, Ω_{cr}	Percent Increase
0.2	0.3	0.819	6.139	23.97%
0.2	0.5	0.811	5.514	11.35%
0.2	0.7	0.841	5.153	4.06%
0.2	1	--	4.952	--
0.3	0.3	0.834	6.502	13.12%
0.3	0.5	0.894	5.96	3.69%
0.3	0.7	0.947	5.784	0.63%
0.3	1	--	5.748	--
0.4	0.3	0.877	7.463	16.77%
0.4	0.5	0.871	6.821	6.73%
0.4	0.7	0.915	6.501	1.72%
0.4	1	--	6.391	--
0.5	0.3	0.878	8.607	13.47%
0.5	0.5	0.914	7.864	3.68%
0.5	0.7	0.959	7.624	0.51%
0.5	1	--	7.585	--
0.6	0.3	0.918	10.387	11.62%
0.6	0.5	0.932	9.605	3.21%
0.6	0.7	0.972	9.338	0.34%
0.6	1	--	9.306	--
0.7	0.3	0.938	13.442	10.04%
0.7	0.5	0.958	12.52	2.49%
0.7	0.7	0.984	12.287	0.58%
0.7	1	--	12.216	--

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. 20, 2000; final revision, Dec. 5, 2000. Associate Editor: R. C. Benson.

Table 2 Comparison of uniform thickness, optimal profile, and existing profile designs

Saw Profile Design	Clamping Ratio, κ	Minimum Thickness, h_{min}	Transition Radius, r_{opt}	Critical Speed, Ω_{cr}	Percent Increase
Uniform Thickness	0.625	1	--	9.854	--
Optimal Profile	0.625	0.702	0.96	9.926	0.73%
Existing Saw Profile	0.625	0.702	0.795	8.548	-13.25%

The axisymmetric stresses σ_r and σ_θ are determined prior to solving (1) by computing the in-plane radial displacement $u(r)$ that renders the functional $S[u]$ stationary ([7])

$$S[u]=\int \frac{1}{2}h[12(u_{,r}^2+u^2/r^2)+24\nu u_{,r}u/r]-rudA. \quad (2)$$

The stress are then

$$\sigma_r=12u_{,r}+12\nu u/r \quad \sigma_\theta=12\nu u_{,r}+12u/r. \quad (3)$$

While explicit equations and boundary conditions for u and w can be derived by setting the variation of S and R equal to zero, the fact that $h(r)$ is not uniform makes these calculations laborious. A more convenient solution approach is to render S and R stationary using the finite element method. With this approach, only the essential boundary conditions need be explicitly imposed:

$$u(\kappa)=w(\kappa)=w_{,r}(\kappa)=0. \quad (4)$$

The three natural boundary conditions—vanishing σ_r, moment, and shear at $r=1$—are satisfied as part of the finite element solution process. We use the three-node element described by Chen and Ren [8] with quadratic trial function for u and quintic trial functions for w and up to 21 elements between $\kappa \leqslant r \leqslant 1$. This number of elements gives results that are within 0.1 percent of those with twice as many elements.

In rendering (1) dimensionless, we taken $h(\kappa)=1$. We further bound the thickness variations by

$$h_{\min} \leqslant h \leqslant 1 \quad (5)$$

where $h_{\min}$ is a fixed constant. The optimization was performed using Powell's method which varied the nodal value of $h(r)$ in order to maximize Ω_{cr}.

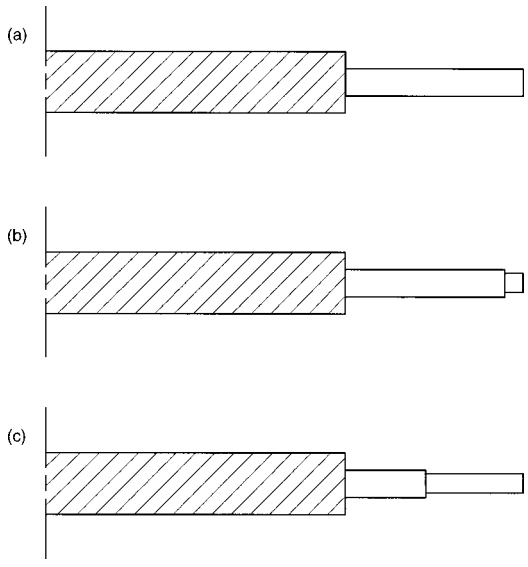


Fig. 1 Scale drawings of the radial profiles of the saw designs shown in Table 2. (*a*) Uniform thickness; (*b*) optimal profile; (*c*) existing profile. Cross-hatched area represents rigid clamping.

3 Results and Comparison to Existing Designs

Under all conditions tested, the optimum profile for maximizing critical speed consists of two, uniform thickness annuli:

$$h(r)=\begin{cases}1, & \kappa \leqslant r \leqslant r_{\text{opt}} \\ h_{\min}, & r_{\text{opt}} < r \leqslant 1\end{cases} \quad (6)$$

Table 1 gives a summary of specific r_{opt} and Ω_{cr} as functions of κ and $h_{\min}$. In general, for a given κ, as $h_{\min}$ increases, r_{opt} moves radially outward until the limiting condition $h_{\min}=r_{\text{opt}}=1$ of a uniform thickness disk is reached. Table 1 also reports the percentage increase in Ω_{cr} over the uniform thickness, $h=1$ case. These increases can be substantial for small values of $h_{\min}$. Further details concerning the optimal profile can be found in ([9]).

Table 2 compares the critical speed of three different saw designs with $\kappa=0.625$: a uniform thickness saw with $h=1$; the optimal design for $h_{\min}=0.702$; and an actual industrial design described by Bird [1] for which $h_{\min}=0.702$. Figure 1 gives scale drawings of these three radial profiles. While the optimal design predicted here raises Ω_{cr} slightly (0.73 percent), the industrial design profile actually decreases Ω_{cr} by over 13 percent. If Ω_{cr} were the principal design parameter for this particular industrial saw, the stepped saw profile would be significantly worse than the uniform thickness profile. Since this does not seem to be the case in practice, it suggests that Ω_{cr} is not the principal design parameter for this saw. Nevertheless, in the cases where Ω_{cr} is the principal design parameter, the optimal profiles described here can be used to improve rotating disk design.

4 Discussion

The optimal thickness profiles found here are in some respects counterintuitive. If one uses the analogy of an I-beam or a honeycomb panel, one might guess that removing mass from the midplane of the disk might improve its critical speed, which in fact, it does ([10]).

Accordingly, one might expect that removing mass from the exterior surfaces of the disk would lower critical speed. This reasoning, however, is incorrect, as the results reported here indicate. For rotating disks, removal of mass from either the midplane or the exterior surfaces can raise critical speed, provided, of course, that the removal is performed at the correct radial locations. Furthermore, the mass removal should not be gradual or even continuous, as is found in several related optimization problems ([11,12]).

From the perspective of the saw designer, the most important insight to be gained from this investigation is the fact that the uniform thickness profile is not optimal with regard to critical speed. In hindsight, this result is straightforward and, as one reviewer felt, perhaps even obvious. Historically, however, stepped circular saws did not become common in the industry until the 1980s, long after resources required to analyze and design such saws became readily available. If, in fact, this results were obvious, there would not have been such a delay between the analytic capabilities of the engineers and the industry practice.

Acknowledgment

The authors are grateful to Warren Bird for his many discussions and insights on circular saw design.

References

[1] Bird, W. M., 1990, ''Rotating Saw Blade Having Improved Critical Vibrational Speed,'' U.S. Patent Number 4,979,417.

[2] Olhoff, N., 1970, ''Optimal Design of Vibrating Circular Plates,'' Int. J. Solids Struct., **6**, pp. 139–156.

[3] Thambiratnam, D. P., and Thevendam, V., 1988, ''Optimum Vibrating Shapes of Beams and Circular Plates,'' J. Sound Vib., **121**, No. 1, pp. 13–23.

[4] Seireg, A., and Surana, K. S., 1970, ''Optimum Design of Rotating Disks,'' J. Eng. Ind., **92**, pp. 1–9.

[5] Berger, M., and Porat, I., 1988, ''Optimal Design of a Rotating Disk for Kinetic Energy Storage,'' ASME J. Appl. Mech., **55**, pp. 164–170.

[6] Renshaw, A. A., 1998, ''Critical Speed for Floppy Disks,'' ASME J. Appl. Mech., **65**, pp. 116–120.

[7] Sokolnikoff, I. S., 1983, *Mathematical Theory of Elasticity*, R. E. Krieger, Malabar, FL.

[8] Chen, D.-Y., and Ren, B.-S., 1998, ''Finite Element Analysis of the Lateral Vibration of Thin Annular and Circular Plates With Variable Thickness,'' ASME J. Vibr. Acoust., **120**, pp. 747–752.

[9] Warner, G. M., and Renshaw, A. A., 1999, ''Thickness Profiles for Rotating Circular Disks That Maximize Critical Speed,'' 1999 ASME Design Engineering Technical Conference, Las Vegas, NV.

[10] Renshaw, A. A., 1999, ''Increasing the Natural Frequencies of Circular Disks Using Internal Channels,'' J. Sound Vib., **229**, No. 2, pp. 355–375.

[11] Tadjbakhsh, I., and Keller, J. B., 1962, ''Strongest Columns and Isoperimetric Inequalities for Eigenvalues,'' ASME J. Appl. Mech., **29**, pp. 159–164.

[12] Keller, J. B., and Niordson, F. I., 1966, ''The Tallest Column,'' J. Math. Mech., **16**, pp. 433–446.