Journal of Applied Mechanics

TECHNICAL PAPERS

TECHNICAL BRIEFS

ERRATA

ANNOUNCEMENTS AND SPECIAL NOTES

Asymptotic Analysis of a Mode III Stationary Crack in a Ductile Functionally Graded Material

Dhirendra V. Kubair[1]

Philippe H. Geubelle

John Lambros[2]
e-mail: lambros@uiuc.edu

Aerospace Engineering,
University of Illinois at Urbana-Champaign,
306 Talbot Laboratory,
104 South Wright Street,
Urbana, IL 61801

The dominant and higher-order asymptotic stress and displacement fields surrounding a stationary crack embedded in a ductile functionally graded material subjected to anti-plane shear loading are derived. The plastic material gradient is assumed to be in the radial direction only and elastic effects are neglected. As in the elastic case, the leading (most singular) term in the asymptotic expansion is the same in the graded material as in the homogeneous one with the properties evaluated at the crack tip location. Assuming a power law for the plastic strains and another power law for the material spatial gradient, we derive the next term in the asymptotic expansion for the near-tip fields. The second term in the series may or may not differ from that of the homogeneous case depending on the particular material property variation. This result is a consequence of the interaction between the plasticity effects associated with a loading dependent length scale (the plastic zone size) and the inhomogeneity effects, which are also characterized by a separate length scale (the property gradient variation). [DOI: 10.1115/1.1876434]

1 Introduction

Numerous naturally occurring mechanical systems have properties that vary continuously and smoothly with position, usually as a result of natural tailoring of their mechanical response to the surrounding environment. Examples of such systems are found in animal and human bone material as well as plant structures such as wood and cellulose [1]. When specifically tailored for a man-made engineering application, materials with a continuous spatial property variation are often termed functionally graded materials (FGMs). For the purpose of this study, an FGM is a material in which the mechanical properties are continuous functions of position. Common forms of FGMs are metal/ceramic systems, which combine the beneficial properties of a metal (toughness, ductility, conductivity) with those of a ceramic (hardness, stiffness, heat resistance). In both naturally occurring and manmade FGMs, the material gradation can be either in a through-thickness or a radial fashion. A through-thickness metal/ceramic FGM has potential applications in areas such as resistance to ballistic penetration and wear resistance and heat shielding, where the ceramic side would be brought into contact with the external agent and the metal side would provide integral structural support. A radial property variation is commonly generated in many plant and bone structures and in manmade situations, such as the graded polymeric matrix material surrounding a fiber in a polymer matrix fiber reinforced composite.

Progress in implementing FGM designs has been slower than initially expected, not only because of the difficulty in manufacturing such materials [2], but also because of a lack of a fundamental understanding of their mechanical response and, in the context of the present work, their fracture response. In comparison with homogenous materials, only a limited number of studies have addressed the structure of the near-tip field in graded materials. Delale and Erdogan [3] solved the mode I problem for a crack parallel to the direction of material property gradient in an unbounded elastic FGM to extract values of the stress intensity factor. Although several studies of this type exist in the literature, detailed description of the *spatial* structure of the near tip stress field has been limited. A notable contribution is the work of Eischen [4] who studied the asymptotic nature of the stress and displacement fields around a crack tip in a graded material whose elastic moduli were specified by continuous and generally differentiable functions. Eischen [4] showed that the asymptotic stress singularity at the crack tip in an FGM is of exactly the same form, i.e., square root singular and with the same angular variation, as that present at a crack tip in a homogeneous material [5]. Eischen also indicated that higher-order terms are influenced by the details of the material gradient distribution, although in the interest of generality these terms were not studied in detail there. Recently, Parameswaran and Shukla [6] computed these higher-order terms (up to six) for an elastic FGM exhibiting a linear property gradient.

Parameswaran and Shukla [7] studied the asymptotic nature of the stress and displacement fields in a dynamically growing crack in an elastic FGM having either an exponential or linear shear modulus variation. Much like Eischen [4], they also saw that the leading term for the case of the FGM is identical to that for the homogeneous material and only higher-order terms are affected by the material gradient. In general, it is physically reasonable to assume that, locally, the material properties of an FGM can be considered homogeneous. Therefore, it can be expected that the leading term of the spatial distribution of stresses and displacements for the FGM will always be the same as the homogeneous case, although the extent of validity of this term may be minimal and has to be investigated on a case-by-case basis [8].

To our knowledge, apart from the work of Eischen [4] and Parameswaran and Shukla [6,7], there has not been any investigation of the detailed spatial structure of deformation fields in the vicinity of crack tips in FGMs, especially in regards to the precise nature of the influence of material gradient variation on higher-order terms. Since FGMs possess an *intrinsic* length scale associated with the material gradient, higher-order terms may be affected by this length scale. In addition, all theoretical studies thus far have dealt with elastic FGMs, despite the fact that the metallic FGM constituent is ductile. In the metal-rich region of a metal/

---

ceramic FGM, the local response is effectively that of a graded plastic material for which an additional length scale corresponding to the *plastic zone size* exists.

The goal of the present work is to investigate in detail the spatial distribution of the displacements and stresses near the tip of a crack embedded in a plastic (ductile) FGM. Of particular interest are the development of higher-order terms, their dependence on material gradient and the interaction between the material gradient length scale and the plastic zone size. Section 2 of this paper presents the problem formulation. For mathematical simplicity, a crack in an unbounded radially varying plastic FGM under mode 3 loading is considered. Despite the limited applicability of a radial gradient and mode 3 loading, this particular problem allows us to study the competition between the plasticity and intrinsic gradient length scales in a tractable way. The leading term of the near-tip solution is derived in Sec. 3. Section 4 presents a solution for the second-order term for both the homogeneous and graded plastic cases. A comparison of these two solutions allows us to investigate in Sec. 5 the relationship between the two length scales involved in this problem.

2 Problem Formulation

A semi-infinite, stationary crack embedded in a functionally graded material subjected to far field mode 3 loading is considered. Let (r, θ) denote a polar coordinate system with its origin positioned at the crack tip. The sole surviving equilibrium equation for the mode 3 problem then is

$$\sigma_{r,r} + \sigma_r/r + \sigma_{\theta,\theta}/r = 0, \qquad (1)$$

where $\sigma_r = \sigma_{rz}$ and $\sigma_\theta = \sigma_{\theta z}$ are the radial and tangential shear stress components, respectively, and comma denotes partial differentiation. We assume hereafter that the material response is described by the infinitesimal J_2-deformation theory of plasticity and that the elastic contribution to the near-tip strains is negligible compared to the plastic one. The Ramberg–Osgood power stress–strain law

$$\frac{\varepsilon_{ij}}{\varepsilon_o} = \frac{3}{2}\alpha\left(\frac{\sigma_e}{\sigma_o}\right)^{n-1}\frac{S_{ij}}{\sigma_o} \qquad (2)$$

is used, where α is a nondimensional amplification factor, n is the hardening exponent, ε_{ij} are the components of strain, and ε_o is the strain at the initial yield strength σ_o. In Eq. (2), σ_e denotes the von Mises equivalent stress given by

$$\sigma_e = \sqrt{\frac{3}{2}S_{ij}S_{ij}}, \qquad (3)$$

where S_{ij} are the components of the deviatoric stress tensor. In the above equations, repeated indices imply summation and the indices i and j range from 1 to 3. Gradation of the material into a plastic FGM can be achieved by allowing a spatial variation in any or all of α, n, ε_o, and σ_o.

In the simpler antiplane shear setting, the stress–strain relation takes the form

$$\frac{\varepsilon_r}{\varepsilon_o} = \frac{\sigma_r}{\sigma_o}\frac{3}{2}\alpha\left[\frac{3(\sigma_r^2 + \sigma_\theta^2)}{\sigma_o^2}\right]^{(n-1)/2},$$

$$\frac{\varepsilon_\theta}{\varepsilon_o} = \frac{\sigma_\theta}{\sigma_o}\frac{3}{2}\alpha\left[\frac{3(\sigma_r^2 + \sigma_\theta^2)}{\sigma_o^2}\right]^{(n-1)/2}, \qquad (4)$$

where $\varepsilon_r = \varepsilon_{rz}$ and $\varepsilon_\theta = \varepsilon_{\theta z}$ are the only nonzero components of the strain tensor and are related to the out-of-plane displacement w by

$$\varepsilon_r = \frac{1}{2}\frac{\partial w}{\partial r}, \quad \varepsilon_\theta = \frac{1}{2r}\frac{\partial w}{\partial \theta}. \qquad (5)$$

Equations (4) can be inverted as

$$\sigma_r = G\frac{\varepsilon_r}{[\varepsilon_e]^{(n-1)/n}}, \quad \sigma_\theta = G\frac{\varepsilon_\theta}{[\varepsilon_e]^{(n-1)/n}}, \qquad (6)$$

where $\varepsilon_e = \sqrt{2\varepsilon_{ij}\varepsilon_{ij}/3}$ is the equivalent strain and the "stiffness" parameter

$$G = \frac{2}{3}\frac{\sigma_o}{(\alpha\varepsilon_o)^{1/n}} \qquad (7)$$

is a combination of the various material parameters entering the constitutive relation Eq. (2). In the present work, in order to keep the problem mathematically tractable, we allow the material properties to vary radially except for the hardening exponent n which is taken to be constant. Therefore, the radially varying FGM can be generated by variations of the material parameters α, ε_o, and σ_o, and can be expressed collectively as a radial variation of the modulus G defined in Eq. (7), i.e., $G = G(r)$. Using the constitutive equation (6), the equilibrium relation Eq. (1) can be written in terms of the strains as

$$[n\varepsilon_r^2 + \varepsilon_\theta^2]\frac{1}{r}\frac{\partial\varepsilon_\theta}{\partial\theta} + (1-n)\varepsilon_r\varepsilon_\theta\left[\frac{1}{r}\frac{\partial\varepsilon_r}{\partial\theta} + \frac{\partial\varepsilon_\theta}{\partial r}\right] + [\varepsilon_r^2 + n\varepsilon_\theta^2]\frac{\partial\varepsilon_r}{\partial r}$$

$$+ n[\varepsilon_r^2 + \varepsilon_\theta^2]\frac{\varepsilon_r}{r} + n[\varepsilon_r^2 + \varepsilon_\theta^2]\frac{\varepsilon_r}{r}\left\{\frac{rG'}{G}\right\} = 0, \qquad (8)$$

where $G' = dG/dr$. The boundary conditions involve mode 3 symmetry ahead of the crack

$$w(r,0) = 0, \qquad (9)$$

and traction-free conditions along the crack faces

$$\sigma_\theta(r, \pi) = 0. \qquad (10)$$

3 Leading Asymptotic Term

In this section, we first consider the effect of material property variation on the leading (most singular) term of the near-tip asymptotic solution. The asymptotic near-tip fields for a mode 3 crack in a homogeneous ductile material have been derived by Amazigo [9] using a hodograph transformation technique. We use here a different approach based on the following assumed separable form of the displacement field near the crack tip:

$$\frac{w(r,\theta)}{r_P} \approx A\left(\frac{r}{r_P}\right)^{p+1}f(\theta) \quad \text{as } r \to 0, \qquad (11)$$

where the power p denotes the strain singularity and must be greater than -1 to ensure a bounded displacement at the crack tip, while A is a nondimensional amplification factor. In Eq. (11), r_p denotes the load-induced length scale (plastic zone size) and is used in this study to normalize all length parameters. An expression for r_p based on the leading terms of the elastic and plastic asymptotic solutions is provided at the end of this section.

Combining Eqs. (11), (5), and (8) leads to

$$\phi_1 + \frac{r_P}{r_F}\frac{r}{r_P}\frac{G'(r/r_F)}{G(r/r_F)}\phi_2 = 0, \qquad (12)$$

where r_F denotes the intrinsic material length scale associated with the material gradient (see Sec. 4.1), and ϕ_1 and ϕ_2 are given by

$$\phi_1 = [n(p+1)^2f^2 + (f')^2]f'' + (p+1)[(p+n) + (p+1)(1-n)]$$

$$\times (f')^2f + (p+n)(p+1)^3f^3,$$

$$\phi_2 = [n(p+1)^3f^3 + n(p+1)(f')^2f], \qquad (13)$$

with $f' = df/d\theta$ and $f'' = d^2f/d\theta^2$. In Eq. (12), the term ϕ_1 would be present for the homogeneous material, while the term associated with ϕ_2 is the contribution due to the material property gradient. If no gradient exists, Eq. (12) reduces to $\phi_1 = 0$ and the solution for the homogeneous material is recovered [9]. Note that

the term involving ϕ_2 contains a ratio of the two inherent length scales involved in this problem, namely the plastic zone size (a function of external loading and plastic properties) and the material property gradient.

Since at this stage we are only interested in the most singular term of the stress field, it can readily be observed from Eq. (12) that, as long as $G(r)$ is finite and strictly positive at the origin (as physically required), the leading order is found by solving $\phi_1 = 0$, which is identical to the equation found in the homogenous case. The leading term for the displacement and stress fields is thus the same for the plastic FGM material as for the plastic homogeneous material [9] with the material properties evaluated at the tip ($G(r=0)$) and $p = -n/(n+1)$.

The fully plastic asymptotic solution is expected to be valid in a region very close to the crack tip [10,11]. Under small scale yielding conditions, the stress state beyond this region decays to the linearly elastic asymptotic. An estimate of the plastic zone size r_P can therefore be obtained by equating the elastic and plastic leading-order asymptotic approximations for the stresses *under equivalent far-field loading* [12]. For simplicity, we compare here the estimates of the equivalent stress σ^e along the crack line ($\theta = 0$). The elastic and plastic asymptotic expressions are, respectively,

$$(\sigma^e)_{\text{elastic}} \approx \frac{K_3}{\sqrt{2\pi r}} = \sqrt{\frac{\mu J}{\pi r}}, \quad \text{as } r \to 0, \qquad (14)$$

$$(\sigma^e)_{\text{plastic}} \approx \frac{\sigma_o}{\sqrt{3}} \left(\frac{J}{\alpha \varepsilon_o \sigma_o I_n} \right)^{-p/n} r^{p/n}, \quad \text{as } r \to 0, \qquad (15)$$

with l_n denoting a known function of the exponent n (Eq. (3.25) and Fig. 2 in [13]). In Eq. (14) and Eq. (15), J denotes the value of the J-integral [14]

$$J = \int_\Gamma \left(W n_1 - \sigma_{ij} \frac{\partial u_j}{\partial x_1} n_i \right) ds \qquad (16)$$

as the contour Γ is shrunk to the crack tip in the FGM case [15]. Equating the asymptotic expressions of Eqs. (14) and (15) yields the following estimation of the plastic zone size:

$$r_P = \frac{J}{\sigma_o} \left(\frac{3\mu}{\pi \sigma_o} \right)^{(n+1)/(n-1)} (\alpha \varepsilon_o I_n)^{2/(n-1)}. \qquad (17)$$

4 Two-Term Asymptotic Approximation

We now investigate whether and how higher-order terms are affected by the material property gradient. In Sec. 4.1, we investigate the conditions under which a separable higher-order solution is possible for the asymptotic displacement field. In Sec. 4.2, we compare our results to those of the two-term asymptotic field around the tip of a crack in a plastic *homogeneous* material subjected to mode 3 loading, obtained by Aravas and Blazo [13] and Anheuser and Gross [16] using different solution techniques. The technique employed here is then extended in Sec. 4.3 to the mode 3 fracture problem in a *graded* plastic material. As will be seen in the following discussion, the exponent of the second term in the homogenous case plays a key role since it provides an upper bound for the region of validity of the FGM solution.

4.1 Existence of a Separable Solution. In Sec. 2, all the material properties that possess a spatial dependence were combined into one single functional form described by Eq. (7). A specific radial variation of $G(r)$ can be arrived at by choosing different variations of the underlying material properties (α, ε_o, and σ_o). Note once again that in the present analysis the hardening exponent n is not allowed to vary with position. The equilibrium equation (8) can be rewritten as
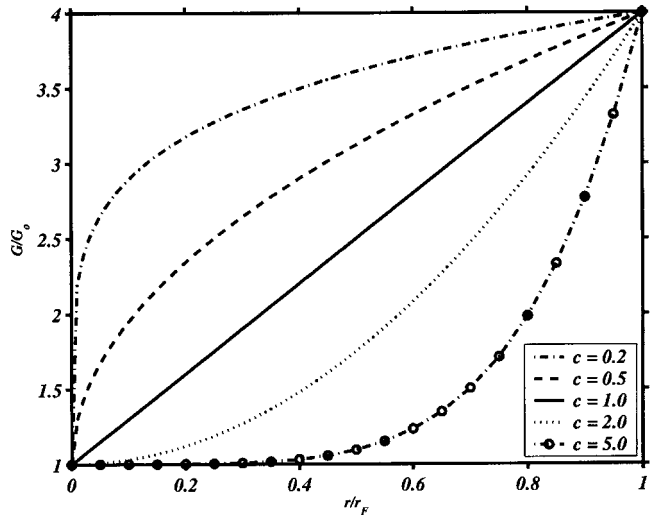
Fig. 1 Influence of property gradient exponent c on the variation of plastic material properties surrounding the crack tip located at $r/r_F = 0$

$$G\Lambda_1(\varepsilon_r, \varepsilon_\theta, n) + \Lambda_2(\varepsilon_r, \varepsilon_\theta, n) r G' = 0, \qquad (18)$$

where Λ_1 and Λ_2 are differential operators that depend only on the angle θ (Eqs. (12) and (13)). Expressing G as

$$G = G_0 + \psi(r), \qquad (19)$$

where G_0 is the material property at the crack tip (with $G_0 > 0$) and $\psi(0) = 0$, the leading orders of the equilibrium equation are recast as

$$G_0 \Lambda_1 = 0,$$

$$\Lambda_1 + r \frac{\psi'(r)}{\psi(r)} \Lambda_2 = 0. \qquad (20)$$

A variable separable solution is possible in the graded case only when

$$\frac{r \psi'(r)}{\psi(r)} = c \Rightarrow \psi(r) = G_1 \left(\frac{r}{r_F} \right)^c, \qquad (21)$$

where G_1 is a constant and r_F denotes the aforementioned intrinsic length scale associated with the material property gradient. The leading term derived in the previous section is recovered if the gradient exponent $c = 0$ and this case is not considered hereafter. To keep the modulus G positive and bounded in the vicinity of the crack tip, c and G_0 must be strictly positive ($c > 0, G_0 > 0$). The physical significance of c can be visualized in Fig. 1, which presents the variation of G over the graded region for $G_1 = 3G_0$. Decreasing the value of c makes the property variation near the crack tip more severe. Thus the value of c is expected to control whether higher-order terms near the crack tip region will be affected by the material property variation or whether they will remain the same as in the homogeneous case. For example, it is clear from Fig. 1 that, for a value of $c = 5$, when viewed at the scale of the entire gradient, the material appears essentially homogeneous in a large region surrounding the crack tip (up to approximately $r = 0.3 r_F$). The opposite is true for a small value of c ($c < 1$), where the material gradient is very strong in the vicinity of the crack tip and is therefore expected to affect the asymptotic fields. Note finally that G_1 can be negative, leading to a "softening" radially varying FGM for which the "stiffness" parameter $G(r)$ decreases away from the crack tip. As will be shown later, the combined effect of the sign and amplitude of the material gradient can be captured by a single parameter.

A variable separable solution to the crack problem does not

appear to be possible in the FGM case for a property variation other than that described by Eqs. (19) and (21). However, it is important to recognize that we are only interested in higher-order contributions to the asymptotic solution and that Eqs. (19) and (21) with $c=1$ can be considered as the first two terms of a Taylor series expansion for G in the vicinity of the crack tip

$$G(r) = G_0 + G_1\left(\frac{r}{r_F}\right) + \cdots. \tag{22}$$

To conclude this section, let us mention that the introduction of a material length scale in a fracture problem may change the conditions for existence of a separable solution. In a recent example it was found that no separable solution was possible within the framework of the gradient theory of plasticity [17].

4.2 Homogeneous Case. This homogeneous case has been solved by Aravas and Blazo [13] using a stress-based formulation and Anheuser and Gross [16] using a perturbation technique. In the displacement-based approach adopted here, we start from the following form of the near-tip displacement field w:

$$\frac{w(r,\theta)}{r_P} \approx A\left(\frac{r}{r_P}\right)^{p+1} f(\theta) + B\left(\frac{r}{r_P}\right)^{q+1} g(\theta) \quad \text{as } r \to 0, \tag{23}$$

where A and B are nondimensional amplitude factors, p and q are the exponents for the leading and second terms, respectively (with $q>p$). Substituting Eq. (23) into Eq. (8) yields, for the two leading orders,

$$A r^{3p-1}\phi_1 = 0,$$

$$B r^{2p+q-1}\phi_3 = 0, \tag{24}$$

where the first equation is identical to that used to derive p and f, and

$$\phi_3 = [n(p+1)^2 f^2 + f'^2]g'' + 2[n(p+1)(q+1)fg + f'g']f'' + (1-n)$$
$$\times (p+1)ff'(2q+1)g' + (1-n)[(p+1)fg' + (q+1)gf'](2p$$
$$+1)f' + \{(p+1)^2 f^2 + nf'^2\}q(q+1)g + 2\{(p+1)(q+1)fg$$
$$+nf'g'\}p(p+1)f + n\{(p+1)^2 f^2 + f'^2\}(q+1)g + 2n\{(p+1)$$
$$\times(q+1)fg + f'g'\}(p+1)f. \tag{25}$$

The eigenvalue problem for q and $g(\theta)$ is completed with the boundary conditions $[g(0)=g'(\pi)=0]$ and a normalization condition [chosen here to be $g'(0)=1$] and can be solved numerically, leaving the amplitude B undetermined by this asymptotic analysis. The variation of q with respect to n (for $1 \leq n \leq 10$) is shown as open circles in Fig. 2, which also contains the elastic-plastic and purely plastic solutions obtained by Aravas and Blazo [13]. As can be seen in Fig. 2, the second asymptotic term is singular for $n > 3.5$. The value of q obtained for the homogeneous case is denoted by q_h in the remainder of the paper. As discussed in the next section, q_h plays a key role in the determination of the region of validity of the asymptotic solution for the FGM case.

4.3 FGM Case. Let us now turn our attention to the asymptotic solution for the mode 3 fracture problem in a radially varying FGM with properties described by Eq. (19) and Eq. (21). We again seek solutions for the displacement field of the form of Eq. (23) with $q>p$. Substituting Eq. (23) into Eq. (8) and rewriting the boundary and normalization conditions, we obtain the following boundary value problem for g:

$$\phi_3 + m\left(\frac{r}{r_P}\right)^\beta \phi_2 = 0,$$

$$g(0) = g'(\pi) = 0, \quad g'(0) = 1, \tag{26}$$

where $\beta = c - (q-p)$, ϕ_2 and ϕ_3 have been defined in Eq. (13) and Eq. (25), and $m = \gamma A/B$ with
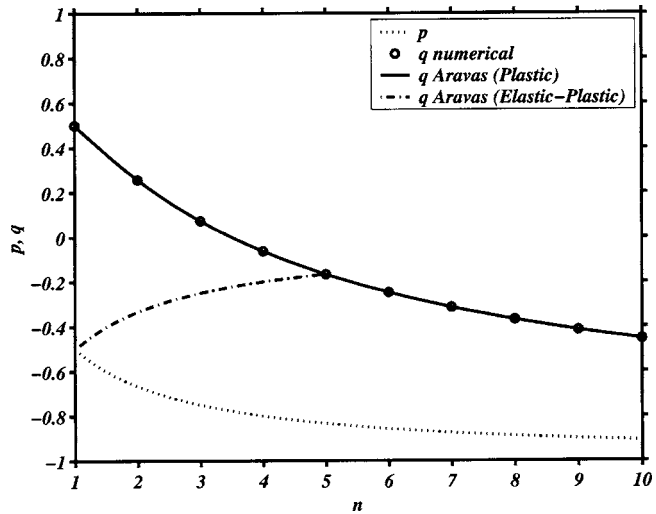
Fig. 2 Variation with respect to *n* of the exponent of the first (*p*, dotted curve) and second (*q*, symbols) asymptotic terms, including the purely plastic and elastic-plastic solutions obtained by [13]

$$\gamma = c\left(\frac{G_1}{G_0}\right)\left(\frac{r_P}{r_F}\right)^c. \tag{27}$$

The parameter γ plays a critical role in this asymptotic study, as it allows us to quantify in a simple unified fashion the combined effect of the nature, amplitude and extent of the material gradient (through the exponent c, the ratio G_1/G_0, and the ratio r_F/r_P, respectively).

Three separate cases must be considered for the solution of Eq. (26) based on the sign of β.

Case (i) $\beta<0$ ($q>p+c$):

Asymptotically, Eq. (26) reduces to $\phi_2=0$. Since ϕ_2 depends only on f [Eq. (13)], this would require f to make ϕ_1 and ϕ_2 vanish simultaneously. Such a solution is not possible and no separable solution can exist for the second term in this case.

Case (ii) $\beta>0$ ($q<p+c$):

In this case, the coefficient of ϕ_2 does not contribute asymptotically to the solution and hence the function g must satisfy $\phi_3=0$. The solution to the FGM case is thus identical to the homogeneous case described in Sec. 4.2 up to the second term (i.e., $q=q_h$). It is possible that the third-order term may be affected, but this is beyond the scope of this work.

Case (iii) $\beta=0$ ($q=c+p$):

In this particular case, the FGM solution for the second-order term is different from the homogeneous one. Unlike in the homogeneous case where q was the solution of an eigenvalue problem, q is now fully determined by the values of the "material" exponents n and c, as

$$q = c + p, \tag{28}$$

with $p=-n/(n+1)$. This FGM case presents another distinguishing feature: the amplitude parameter B appearing in the second term of the near-tip expansion Eq. (23), which was left undetermined in the homogeneous case, is now *fully determined*. Only the amplitude of the first term [denoted by A in Eq. (23)] is left undetermined by the analysis and can be related to the value of the J-integral. As indicated earlier, this separable solution is only valid for a limited range of c values: $0<c<c_{\max}=q_h-p$. Indeed, when c exceeds $c_{\max}$, we revert to case (ii) discussed above, and the second term of the asymptotic expansion in the FGM case is given by that of the homogeneous problem. In that case, the gradient may affect only the third or higher-order terms in the expansion. This result is consistent with the comments made earlier in
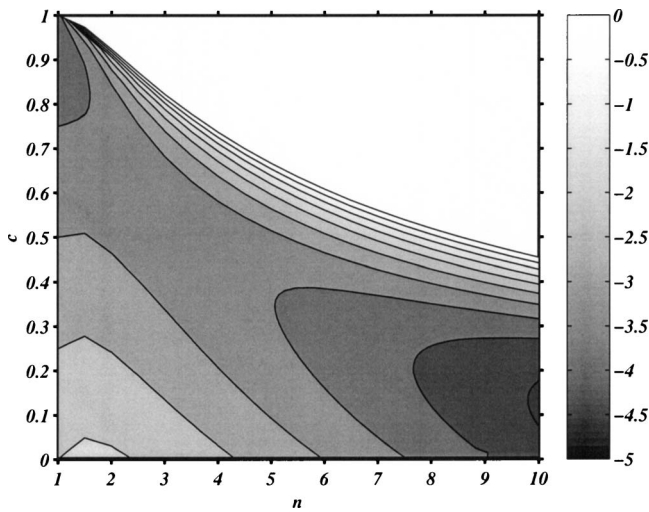
Fig. 3 Variation of *m* with respect to the material exponents *n* ($1 \leq n \leq 10$) and *c* [$0 < c \leq c_{max}(n)$]

Fig. 1. Physically, increasing c makes the property change less steep around the crack tip, thus extending the range over which the material can be considered "homogeneous." The homogeneous exponent q_h, i.e., the strength of the second term in the homogeneous case, therefore determines, through $c_{max} = q_h - p$, the material gradient variation beyond which the second term in the asymptotic series will be not affected by the heterogeneity of the material.

For the case of $0 < c \leq c_{max}$, the governing equation for g becomes

$$\phi_3 + m\phi_2 = 0, \tag{29}$$

with $m = \gamma A/B$ and γ defined by Eq. (27). The eigenvalue problem can again be solved numerically. The variation of m with respect to the material parameters c and n is shown as a contour plot in Fig. 3 for $1 < n \leq 10$ and $0 < c \leq c_{max}(n)$. As expected, m tends to zero (i.e., for a given value of γ, the second term becomes increasingly dominant) as c approaches c_{max}. When $c = c_{max}$, the exponent of the second asymptotic term of the FGM case equals that of the homogeneous case and B becomes indeterminate. The case $n = 1$ requires special attention. For this value of n, which yields a linear relation between stress and strain components, no solution can be found for the eigenvalue problem described by Eq. (29). The material gradient does not therefore affect the first two terms of the near-tip asymptotic expansion, in agreement with earlier results obtained by Parameswaran and Shukla [6].

The angular function $g(\theta)$ of the second asymptotic term for w is presented in Fig. 4 for two values of n ($n = 8$ and $n = 20$) and for four values of the exponent c for each n. As $c \to 0$, the second-order term approaches the first one ($q \to p$) and g approaches f. In the other extreme case ($c \to c_{max}$), $q \to q_h$ and the corresponding angular function of the FGM problem approaches that of the homogeneous one.

5 Discussion

It is clear from the above results that in the FGM problem a competition takes place between the crack tip stress field setting up a second term that is the same as the homogeneous case (of strength q_h) and one that is affected by material gradient (of strength $q = c + p$). This competition occurs in the shadow of the most singular term (of strength p) that is the same for the homogeneous and graded cases. To visualize the combined effect of the first two terms of the asymptotic solution, we present contour plots of the radial (σ_r) and tangential (σ_θ) shear stresses (normalized by G_0) obtained by combining Eqs. (5)–(7), (15), (17), and
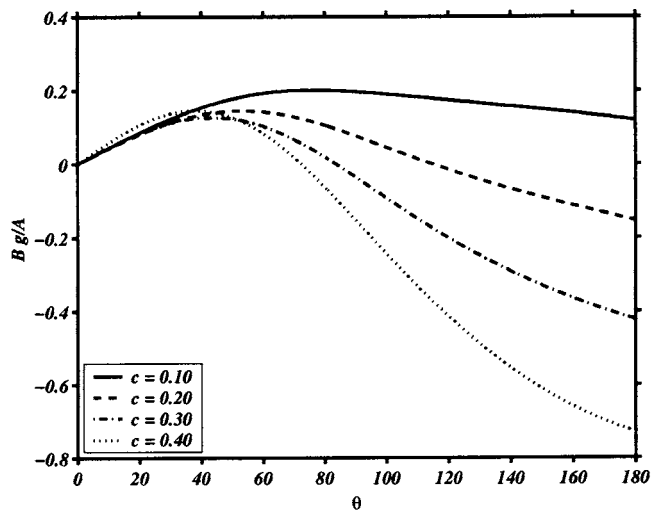
(a)
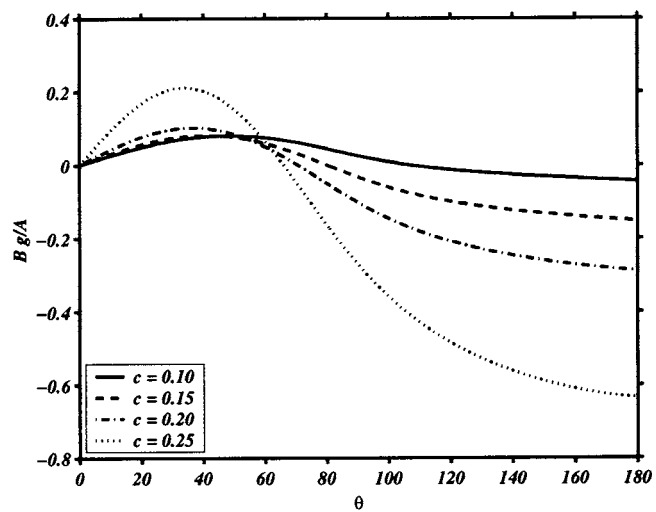
(b)

Fig. 4 Angular variation of the second asymptotic term for the graded case for (a) *n*=8 ($c_{max}=0.520$) and (b) *n*=20 ($c_{max}=0.283$), and for various values of *c*

(23). For reference, stress contours corresponding to the one- and two-term homogeneous solutions are shown in Figs. 5(a) and 5(b), respectively, for $n = 8$ ($p = -0.889$ and $q_h = -0.369$) and up to a radius $r = 0.8 \, r_P$. In the immediate vicinity of the crack tip, the solution is completely characterized by the first term. The influence of the second term is only felt outside of the "core region" and depends on the adopted B/A ratio (chosen to be 0.3 in these contour plots). Recall that for the homogeneous case, the ratio B/A is left undetermined by the asymptotic analysis. Figure 5(c) presents the one-term approximation of the near-tip stress field for a large value of n ($n = 1000$). In that case, which can be considered as approaching the "perfectly plastic situation," the first two asymptotic exponents p and q_h are very close to each other ($p = -0.999$ and $q_h = -0.991$). The corresponding angular functions are then almost identical and there is little difference between the one- and two-term approximations. Due to the more singular nature of the near-tip solution, the region of stress concentration is smaller for $n = 1000$ than for $n = 8$, and the near-tip stress fields are characterized by the appearance of sectors, characteristic of perfectly plastic solutions.
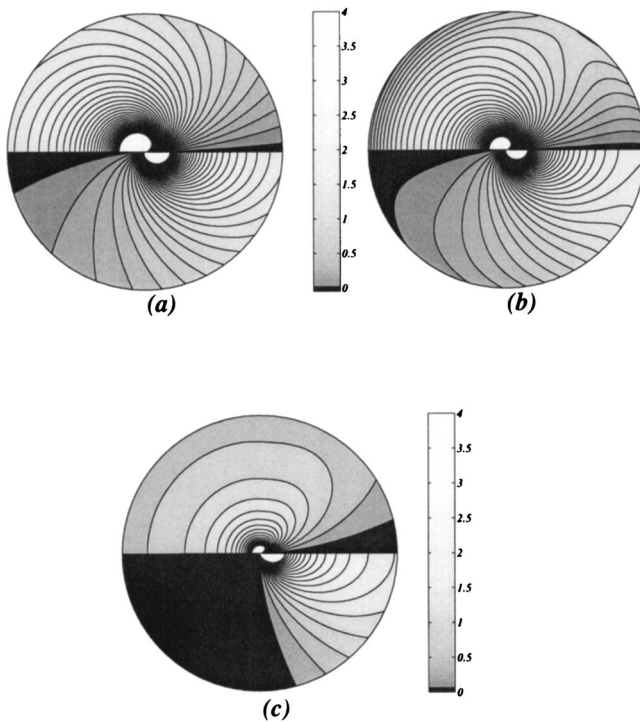
Fig. 5 Contour plot of the two near-tip stress components (normalized by G_0) in a circular domain of radius equal to $0.8r_P$ and centered at the crack tip. The top half of each circle corresponds to σ_r (which is odd in θ) and the bottom half to σ_θ (even in θ). (a) and (b), respectively, correspond to the one- and two-term approximations for the homogeneous case with $n=8$, while (c) shows the one-term solution for $n=1000$.

Stress contours obtained from the FGM two-term solution are shown in Fig. 6 for $n=8$, $c=0.4$ and for three values of γ ($\gamma = -1$, 0.1, and 1). Note that a negative value of γ corresponds to "radial softening," i.e., to the case where the modulus $G(r)$ is maximum at the crack tip and decreases radially. As mentioned earlier, unlike the homogenous case, the solution for the second term is completely determined in the FGM case, as both the exponent q and the ratio of amplification factor B/A are known. The only undetermined quantity in the FGM case is thus the amplitude A, which is related to the J-integral. This is the sole quantity controlling the extent of plasticity, since r_P is proportional to J/σ_o. Therefore, the competition between the "load-induced length scale" r_P and "material gradient length scale" r_F, quantified by the parameter γ defined by Eq. (27), is clearly visible in Fig. 6. For small values of γ (i.e., when the material gradient is small), the near-tip solution is very similar to that obtained with the one-term approximation (and shown with the same gray scale scheme as in Fig. 5(a)). For steeper positive (Fig. 6(c)) and negative (Fig. 6(a)) gradients, the region of dominance of the one-term approximation is reduced.

This competition is further illustrated in Fig. 7, which presents on a log–log plot of the radial variation of σ_θ directly ahead of the crack (i.e., for $\theta=0$) for the one- and two-term homogeneous solutions, and for five values of γ for the graded material. All curves have been obtained for $n=8$ and the FGM solution use $c=0.4$ (i.e., all FGM solutions share the same values of the asymptotic exponents p and q). As expected, very little deviation from the one-term approximation is observed in the homogeneous case (for which B/A has been chosen as 0.3), but also in the FGM case with $\gamma=0.1$. As γ increases, the material heterogeneity reduces the region of dominance of the one-term asymptotic solution. It is interesting to note that, for FGMs with a positive radial gradient (i.e., for $\gamma>0$), the stress distribution described by the first two

asymptotic terms does not decay monotonically from the crack tip, but reaches a local minimum. The situation is of course very different for radially varying FGMs with a negative material gradient ($\gamma<0$): the material heterogeneity reinforces the rapid decay of the stress field away from the crack tip. Physically, this can be rationalized by recognizing the natural tendency of a crack tip to produce a stress field amplifying stresses in the near-tip region but

Fig. 6 Stress contour plots (similar to those shown in Figs. 5) obtained with the two-term approximation for the FGM case with $n=8$, $c=0.4$, and $\gamma=-1$ (a), $\gamma=0.1$ (b), and $\gamma=1$ (c).
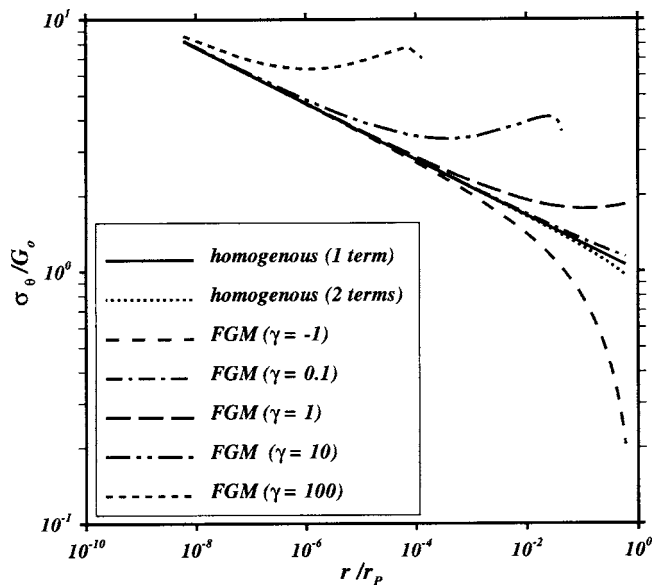
Fig. 7 Radial variation of σ_θ ahead of the crack, showing the effect of γ on the region of dominance of the most singular term (denoted by the solid curve). For comparison, the two-term approximation is also shown in the homogeneous case (obtained for $B/A=0.3$). The curves have been obtained for $n=8$ and, in the FGM case, $c=0.4$.

decaying at larger distances away from the tip. The local stresses, however, are also controlled by the local stiffness of the material. Therefore, in a FGM with a positive radial gradient ($\gamma > 0$), although the crack tip stresses tend to decay as we move away from the crack, the material is getting stiffer, which in turn raises the stress level. In the immediate vicinity of the crack tip, the leading order is dominant and the stresses decay as we move away. However, when the second-order term becomes important, the material stiffening effect is increasingly felt and the stress increases as we continue moving away from the tip. For an FGM with a negative radial gradient ($\gamma < 0$), the radially decreasing material stiffness accelerates the natural decay rate of the stresses induced by the presence of the crack. Both these effects are clearly seen in Fig. 7.

6 Conclusions

The first two terms of an asymptotic approximation of the near-tip displacement and stress fields have been obtained for a mode 3 crack embedded in a radially varying plastic FGM described by the Ramberg–Osgood power law. Taking advantage of the mathematical simplicity of the solution we draw the following conclusions:

(1) As is the case in the elastic problem, the material property gradient does not affect the form of the leading term of the near-tip approximation, but may affect the higher-order terms.

(2) A separable solution up to the second term in the FGM case is possible when the material stiffness parameter G defined by Eq. (7) follows a power law variation of the form $G = G_0 + G_1(r/r_F)^c$, with $G_0 > 0$ and $c > 0$. In this case, the exponent q of the second asymptotic term is simply given by $q = p + c$, where p is the exponent of the first (most singular) term.

(3) The existence of a second asymptotic term specific to the FGM case depends on the value of the material exponent c: if the resulting value of q is less than that obtained for the second term in the homogeneous case (q_h), the material gradient will affect the two-term asymptotic solution. Otherwise, it might only affect terms of third order or higher.

(4) Unlike in the homogeneous situation for which the amplitude of *all* asymptotic expansion terms is left undetermined by the asymptotic analysis, the amplitude of the second term in the FGM case is fully determined by the value of the first one.

(5) The competition between the two length scales characterizing the fracture problem (the plastic zone size and the material gradient length scale) has been quantified in terms of a single parameter γ introduced in Eq. (27) that incorporates the combined effect of the nature, amplitude, and extent of the material gradient. The value of this parameter strongly affects the size of the region of dominance associated with the most singular asymptotic term, and the nature of the development of the stresses and displacements outside this region of dominance when the second term becomes significant.

How these conclusions transfer to more complex material property gradients and/or loading modes is currently under investigation.

Acknowledgment

References

[1] Krassig, K., 1993, "Cellulose: Structure, Accessibility and Reactivity," *Polymer Monographs*, **11**, Gordon and Breach, New York.
[2] Rabin, B. H., and Shiota, I., 1995, "Functionally Graded Materials." Mater. Res. Bull., **20**(1), pp. 14–15.
[3] Delale, F., and Erdogan, F., 1983, "The Crack Problem for a Nonhomogeneous Plane." J. Appl. Mech., **50**, pp. 609–614.
[4] Eischen, J. W., 1987, "Fracture of Non-Homogeneous Materials." Int. J. Fract., **34**, pp. 3–22.
[5] Williams, M. L., 1957, "On the Stress Distribution at the Base of a Stationary Crack." J. Appl. Mech., **24**, pp. 109–114.
[6] Parameswaran, V., and Shukla, A., 2002, "Asymptotic Stress Fields for Stationary Cracks Along the Gradient in Functionally Graded Materials." J. Appl. Mech., **69**, pp. 240–243.
[7] Parameswaran, V., and Shukla, A., 1999, "Crack-Tip Fields for Dynamic Fracture in Functionally Graded Materials." Mech. Mater., **31**, pp. 579–596.
[8] Anlas, G., Lambros, J., and Santare, M. H., 2002, "Dominance of Asymptotic Crack Tip Fields in Elastic Functionally Graded Materials." Int. J. Fract., **115**, pp. 193–204.
[9] Amazigo, J. C., 1974, "Fully Plastic Crack in an Infinite Body Under Anti-Plane Shear." Int. J. Solids Struct., **10**, pp. 1003–1015.
[10] Hutchinson, J. W., 1968, "Singular Behavior at the End of a Tensile Crack in a Hardening Material." J. Mech. Phys. Solids, **16**, pp. 13–31.
[11] Rice, J. R., and Rosengren, G. F., 1968, "Plane Strain Deformation Near a Crack Tip in a Power-Law Hardening Material." J. Mech. Phys. Solids, **16**, pp. 1–12.
[12] Geubelle, P. H., and Knauss, W. G., 1994, "Finite Strains at the Tip of a Crack in a Sheet of Hyperelastic Material: 1. Homogeneous Case." J. Elast., **35**, pp. 31–98.
[13] Aravas, N., and Blazo, A., 1991, "Higher-Order Terms in Asymptotic Elastoplastic Mode-III Crack Tip Solutions." Acta Mech., **90**(1–4), pp. 139–153.
[14] Rice, J. R., 1968, "A Path Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks." J. Appl. Mech., **35**, pp. 379–386.
[15] Anlas, G., Santare, M. H., and Lambros, J., 2000, "Numerical Calculation of Stress Intensity Factors in Functionally Graded Materials." Int. J. Fract., **104**, pp. 131–143.
[16] Anheuser, M., and Gross, D., 1994, "Higher Order Fields at Crack and Notch Tips in Power-Law Materials Under Longitudinal Shear." Arch. Appl. Mech., **64**, pp. 509–518.
[17] Shi, M. X., Huang, Y., Gao, H., and Hwang, K. C., 2000, "Non-Existence of Separable Crack Tip Field in Mechanism-Based Strain Gradient Plasticity." Int. J. Solids Struct., **37**, pp. 5995–6010.

Asymptotics for the Characteristic Roots of Delayed Dynamic Systems

Pankaj Wahi[1]
e-mail: pankaj@mecheng.iisc.ernet.in

Anindya Chatterjee
e-mail: anindya@mecheng.iisc.ernet.in

Mechanical Engineering,
Indian Institute of Science,
Bangalore 560012, India

Delayed dynamical systems appear in many areas of science and engineering. Analysis of general nonlinear delayed systems often begins with the linearized delay differential equation (DDE). The study of these linearized constant coefficient DDEs involves transcendental characteristic equations, which have infinitely many complex roots not obtainable in closed form. Here, after motivating our study with a well-known delayed dynamical system model for tool vibrations in metal cutting, we obtain asymptotic expressions for the large characteristic roots of several delayed systems. These include first- and second-order DDEs with single delays, and a first-order DDE with distributed as well as multiple incommensurate delays. For reasonable magnitudes of the coefficients of the DDEs, the approximations in each case are very good. Subsequently, a fourth delayed system involving coefficients of disparate magnitude is analyzed using an alternative asymptotic strategy. Finally, the large root asymptotics are complemented with calculations using Padé approximants to find all the roots of these systems.
[DOI: 10.1115/1.1875492]

1 Introduction

Delay differential equations (DDEs) are infinite-dimensional systems which find application in manufacturing processes, control systems, biology, economics, chemical kinetics, and other areas [1–9]. The simplest linear DDEs have constant coefficients, as in

$$\dot{x}(t) = \alpha x(t) + \beta x(t - \Delta), \tag{1}$$

where $\Delta > 0$. The solution of Eq. (1) is a sum of terms of the form $e^{\lambda t}$ [10–14], where λ satisfies

$$\lambda = \alpha + \beta e^{-\lambda \Delta}.$$

This equation determines the infinitely many *characteristic roots* of the DDE, impossible to find in closed form. If all these roots have negative real parts, then all solutions decay to zero and the system is stable. A root with a positive real part implies an exponentially growing solution (system unstable).

In this paper, we find *all* the characteristic roots of some linear constant coefficient DDEs.

The study of linear constant coefficient DDEs is important since analysis of general nonlinear DDEs often begins with a preliminary study of the *linearized* DDE (e.g., [7]). Analysts often seek conditions under which the system is stable. This search is nontrivial because there are infinitely many characteristic roots not obtainable in closed form. Significant stability results have, nevertheless, been obtained (e.g., chap. 10 in [10], Theorems 4.1–4.3 in Sec. 11.4, Theorems 5.1–5.3 in Sec. 11.5 in [12], Sec. 1.2 and 1.4 in [13], and chap. 2 in [14], etc.; see also [15,16]). The above references contain results pertaining to either all roots having negative real parts, or specific parameter values for which a pure imaginary pair exists. It is also possible to count the number of roots in the right half of the complex plane [14,17].

Whether the system is stable or not, it is of further interest to find out *where* the characteristic roots lie. The geometrical distribution of these characteristic roots is important in proving theorems on series expansion and asymptotic behavior of solutions; see Bellman and Cooke [10]. Results in similar directions have also been obtained for specific equations by others (e.g., Sec. 1.4 of [12], Sec. 11.3 of [18], etc.). These theorems show for certain DDEs, e.g., that there are finitely many roots in any vertical or horizontal strip in the complex plane.

Numerical algorithms for finding the characteristic roots of linear constant coefficient DDEs have been given in [19,20]. However, they are computationaly efficient for finding the first few roots only. Sandquist and Rogers [21] have sought the characteristic roots for scalar linear first-order DDEs; they consider a single delay, and graphically determine the roots of a transcendental equation in one variable.

As mentioned above, in this work we aim to determine *all* the roots of some linear constant coefficient DDEs. Our approach is based on asymptotic calculations for the *large* roots, a Padé approximant for a small number of roots that are not large, and numerics (the Newton-Raphson method) to refine these roots. Alternative asymptotics are also used for the not very large roots of a DDE involving coefficients of disparate magnitude.

It may be noted that Bellman and Cooke [10] have obtained leading order asymptotics for characteristic roots of a class of DDEs with multiple commensurate delays. In contrast, here we find correction terms in the expansions, giving very accurate estimates. Moreover, incommensurate delays are included in our study. Finally, DDEs with coefficients of disparate magnitude are considered as well.

2 Mechanical System With Delay

A schematic of the turning process in three-dimensional (3D) space, and a 2D projection of the same on the *x-y* plane, are shown in Fig. 1. The derivation below is that of Stépán [1]. Introduce a "long" discrete time delay $\tau_1 = 2\pi/\Omega$, where Ω is the speed of workpiece rotation in rad/s. This is the time period of one revolution. Also introduce a "short" continuous delay $h = 2L/\Omega D$,

Fig. 1 Simple model for tool vibrations

where L is the length of the tool-chip contact region and D is the diameter of the workpiece. This is the time during which the tool and the chip are in contact.

The linear model of regenerative machine tool vibration in the x direction, considering both the long discrete and the short continuous delays, is given by Stépán [1] as

$$\ddot{x}(t) + 2\psi\omega_n\dot{x}(t) + \omega_n^2 x(t) + \frac{k_1}{m}\int_{-h}^{0} W(\theta)x(t+\theta)d\theta$$

$$- \frac{k_1}{m}\int_{-\tau_1-h}^{-\tau_1} W(\tau_1+\theta)x(t+\theta)d\theta = 0, \qquad (2)$$

where $\omega_n = \sqrt{k/m}$ is the natural angular frequency of the tool, $\psi = c/2\sqrt{mk}$ is the damping ratio, k_1 is the slope of the F_x versus f curve[2] at $f = f_0$, and $W(\theta)$ with $\theta \in [-h, 0]$, is the stress distribution function over the tool-chip contact region.

When $W(\theta) = \delta(\theta)$, there is only the long discrete delay τ_1, and on suitable rescaling of time, Stépán obtains

$$\ddot{x}(t) + 2\psi\dot{x}(t) + (1+p)x(t) - px(t-\tau) = 0, \qquad (3)$$

where $p = k_1/m\omega_n^2$, and $\tau = \omega_n\tau_1$. Equation (3) is also derived in [7]. If $\tau = 1$, for simplicity, we have

$$\ddot{x} + 2\psi\dot{x} + (1+p)x - px(t-1) = 0. \qquad (4)$$

We will study the characteristic roots of Eq. (4) in detail in Sec. 4.

When $W(\theta) \neq \delta(\theta)$ in Eq. (2), we have an integro-DDE which represents a distributed delay effect.

If $W(\theta)$ is approximated by more than one Dirac-delta function, e.g., $W(\theta) = \delta(\theta) + \delta(\theta + h)$, we get a delayed system with multiple delays which could be incommensurate. Note that systems with incommensurate delays are not amenable to analysis using Pontryagin's criteria [22] (see also [12]). But we will consider such systems below. Note that incommensurate delays can also arise in other applications, e.g., a system where a microphone picks up signals from two unequally distant speakers and gives it as a feedback to the amplifier driving them.

3 Preliminary Example

We begin with the equation (e.g., [12,13])

$$\dot{x} + ax(t-1) = 0. \qquad (5)$$

Assuming $x(t) = C e^{\lambda t}$ as usual, we obtain

$$\lambda + a e^{-\lambda} = 0. \qquad (6)$$

Letting $\lambda = \alpha + i\beta$ and separating real and imaginary parts, we obtain

$$\alpha + a e^{-\alpha}\cos\beta = 0, \qquad (7)$$

$$\beta - a e^{-\alpha}\sin\beta = 0. \qquad (8)$$

We assume a is nonzero and $\mathcal{O}(1)$ compared to $|\lambda|$, and that $\beta \geq 0$ ($-\beta$ gives another solution).

---

[2]Here f is the chip thickness, and F_x is the x component of the cutting force.
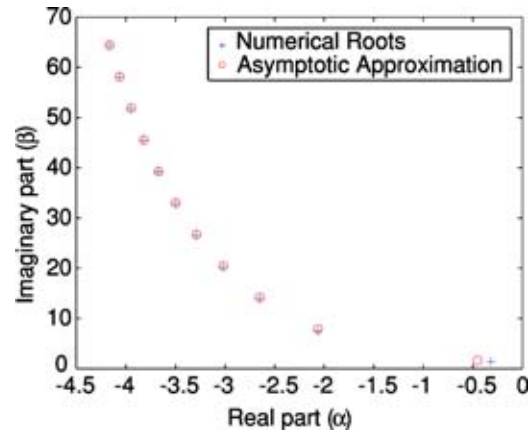
Fig. 2 Roots of Eq. (6) for $a=1$

PROPOSITION 1. *In the large roots of Eq. (6), $\alpha < 0$ and $\alpha = \mathcal{O}(\ln\beta)$.*

Proof. If, in the large roots, α is not large [i.e., is bounded or $\mathcal{O}(1)$] then Eq. (8) gives the contradiction $\beta = \mathcal{O}(1)$. If α is large and positive, then Eq. (7) gives the contradictory $\alpha = o(1)$. Therefore, α is large and *negative*. Then Eq. (7) implies, in the asymptotic limit, that $\cos\beta = 0$, which gives $\sin\beta = \pm 1$. By Eq. (8) we can show that, for integer $N \gg 1$

$$\alpha \sim -\ln\left(\frac{\beta}{|a|}\right) \text{ and } \beta \sim \left(2N + \frac{\text{sgn}(a)}{2}\right)\pi. \qquad \blacksquare$$

Roots found numerically by the Newton-Raphson method, along with the asymptotic approximations themselves, agree well in Fig. 2.

4 Second-Order DDE

We now consider the equation derived in Sec. 2,

$$\ddot{x}(t) + 2\psi\dot{x}(t) + (1+p)x(t) - p\,x(t-1) = 0, \qquad (9)$$

For this equation, for small p, all characteristic roots lie in the left half plane [7]. Here, we develop large-root asymptotics for p nonzero and finite. Note that second-order delayed systems have long been of interest due to mechanical applications [15,16,23].

In Eq. (9), we let $x = C e^{\lambda t}$, set $\lambda = \alpha + i\beta$, and separate real and imaginary parts to get

$$\alpha^2 - \beta^2 + 2\psi\alpha + 1 + p - p\,e^{-\alpha}\cos\beta = 0, \qquad (10)$$

$$2\alpha\beta + 2\psi\beta + p\,e^{-\alpha}\sin\beta = 0. \qquad (11)$$

We now eliminate large regions of the complex plane from our consideration. We divide the upper half plane qualitatively into the regions shown in Fig. 3, and investigate them one by one. The lower half plane is symmetrical, hence excluded. We will elimi-
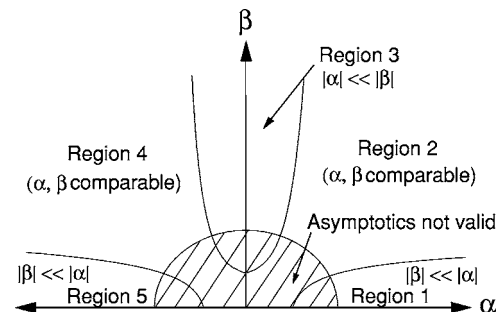
Fig. 3 Regions considered in the complex plane

nate regions *not* containing any roots. In the shaded area roots are not large, and our asymptotics do not hold.

PROPOSITION 2. *For Eq.* (4), *large roots have large* α *and occur only in the left half of region* 3, *Fig.* 3.

Proof. If α is $\mathcal{O}(1)$ in the large roots, it leaves an unbalanced β^2 in Eq. (10). So α is large.

First consider regions 1 and 2, as well as the right half of region 3. Since $\alpha > 0$ here, $0 < e^{-\alpha} < 1$. In Eq. (10), since large α means $\alpha^2 \gg \alpha$, we balance the two largest terms to get $\alpha^2 \sim \beta^2$. Using this in Eq. (11), we find nothing can balance the large term $2\alpha\beta$.

Next, consider regions 4 and 5. Here, $e^{-\alpha}$ is exponentially large compared to both α and β as well as algebraic powers thereof. Equations (10) and (11) then lead to the contradictory

$$\cos^2\beta + \sin^2\beta \ll 1.$$

By elimination, only the left half of region 3 contains large roots. ∎

PROPOSITION 3. *Equations* (10) *and* (11) *are satisfied by* $\alpha = -2\ln\beta + \ln|p| + o(1)$ *and* $\beta = [2N + (1 + \mathrm{sgn}(p))/2]\pi + o(1)$, *for integer* $N \gg 1$.

Proof. By proposition 2, $\beta \gg |\alpha| \gg 1$ for the large roots. We substitute $\alpha = \mu \ln\beta + z$ in Eqs. (10) and (11), with the assumption that $|z| \ll \ln\beta$, to get

$$\mu^2(\ln\beta)^2 + 2\mu(z + \psi)\ln\beta + z^2 - \beta^2 + 2\psi z + 1 + p - p\,\beta^{-\mu}e^{-z}\cos\beta$$
$$= 0, \tag{12}$$

$$2\mu\beta\ln\beta + 2z\beta + 2\psi\beta + p\,\beta^{-\mu}e^{-z}\sin\beta = 0. \tag{13}$$

In Eq. (12), β^2 is asymptotically bigger than all other terms [including $\mu^2(\ln\beta)^2 \sim \alpha^2$] except possibly $p\,\beta^{-\mu}e^{-z}\cos\beta$, so these two terms must balance each other. Note, $p \ne 0$. This gives

$$\beta^{-\mu}e^{-z}\cos(\beta) = \mathcal{O}(\beta^2). \tag{14}$$

Taking absolute values and then logarithms

$$-\mu\ln\beta + \ln|\cos\beta| \sim 2\ln\beta,$$

where we have dropped z since it is smaller than $\ln\beta$.

In the above, we could conceivably have $0 < |\cos\beta| \ll 1$, and in fact small enough that $\ln|\cos\beta| = \mathcal{O}(\ln\beta)$. This, however, requires $\beta^{-\mu}e^{-z} \gg \beta^2$ which leaves a large term $\beta^{-\mu}e^{-z}\sin\beta$ (with $\sin\beta \sim \pm 1$) unbalanced in Eq. (13). Therefore, to balance terms, we must let $\mu = -2$. Then Eqs. (12) and (13), on dividing by β^2 and dropping smaller terms, give

$$p\,e^{-z}\cos\beta \sim -1, \tag{15}$$

$$p\,e^{-z}\sin\beta = o(1). \tag{16}$$

Since $pe^{-z} = 0$ violates Eq. (15), we must have $\sin\beta = o(1)$ from Eq. (16), whence $\cos\beta \sim \pm 1$. Consequently, $pe^{-z} \sim \mp 1$. The result follows. ∎

We can now develop formal series as follows (assuming $p > 0$; the other case is analogous). Having balanced β^2 terms, we still retain $\mathcal{O}(\beta)$ terms, which are relatively smaller by a factor of $\mathcal{O}(1/N)$. Accordingly, we write (note the slightly different form for α, now dependent explicitly on N)

$$\beta = (2N + 1)\pi + \frac{\beta_1}{N} + \frac{\beta_2}{N^2} + \cdots, \tag{17}$$

$$\alpha = -2\ln[(2N+1)\pi] + \ln p + \frac{\alpha_1}{N} + \frac{\alpha_2}{N^2} + \cdots. \tag{18}$$

Substituting into Eqs. (12) and (13), expanding in series, collecting terms (using Maple 6 (Windows)), and solving for the unknown coefficients, we obtain

$$\beta_1 = -\frac{2\ln(2N\pi) - \psi - \ln p}{\pi},$$
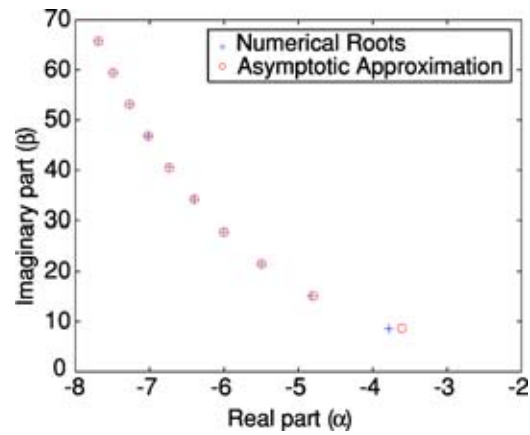
Fig. 4 Roots of Eq. (4) with $\psi = 0.1$ and $p = 2$

$$\beta_2 = \frac{2\ln(2N\pi) - \psi - \ln p - 2}{2\pi},$$

$$\alpha_1 = 0,$$

$$\alpha_2 = -\frac{1}{4\pi^2}(-1 + 8\ln\pi\ln N + 8\ln\pi\ln 2 + 8\ln 2\ln N$$
$$- 4\ln(2N\pi)\ln p - 4\psi\ln(2N\pi) + 2\ln p + 2\psi\ln p + 2\psi^2 - p$$
$$+ 4\psi).$$

While collecting terms above, we have treated $\ln N$ as $\mathcal{O}(1)$ compared to N.

The above approximations agree well with roots obtained using Newton-Raphson; see Fig. 4. There are just three relatively small roots (one real, one complex pair) not captured by the asymptotics; those are not plotted here.

5 DDE With Multiple and Distributed Delays

Consider

$$\dot{x}(t) + a_1 x(t) + a_2 x(t - 1/\sqrt{2}) + a_3 x(t - 1) + a_4 \int_0^1 x(t - s)\cos s\,ds$$
$$= 0, \tag{19}$$

with $a_3 \ne 0$. The characteristic equation is (we multiply by $\lambda^2 + 1$ to simplify the expression, but introduce spurious roots at $\lambda = \pm i$ which we ignore)

$$-a_4 e^{(-\lambda)}\lambda\cos(1) + a_4 e^{(-\lambda)}\sin(1) + a_4\lambda + a_3 e^{(-\lambda)}\lambda^2 + a_2 e^{(-\lambda)}$$
$$+ a_2 e^{(-\lambda/\sqrt{2})}\lambda^2 + a_2 e^{(-\lambda/\sqrt{2})} + a_1\lambda^2 + a_1 + \lambda^3 + \lambda = 0.$$

Substituting $\lambda = \alpha + i\beta$ and separating real and imaginary parts, we get

$$-a_4 e^{-\alpha}\alpha\cos 1\cos\beta - a_4 e^{-\alpha}\beta\cos 1\sin\beta + a_4 e^{-\alpha}\sin 1\cos\beta$$
$$+ a_4\alpha - a_3 e^{-\alpha}\beta^2\cos\beta + 2a_3 e^{-\alpha}\alpha\,\beta\sin\beta + a_3 e^{-\alpha}\alpha^2\cos\beta$$
$$+ a_3 e^{-\alpha}\cos\beta + 2a_2 e^{-\alpha/\sqrt{2}}\alpha\,\beta\sin(\beta/\sqrt{2})$$
$$+ a_2 e^{-\alpha/\sqrt{2}}\alpha^2\cos(\beta/\sqrt{2}) - a_2 e^{-\alpha/\sqrt{2}}\beta^2\cos(\beta/\sqrt{2})$$
$$+ a_2 e^{-\alpha/\sqrt{2}}\cos(\beta/\sqrt{2}) + a_1\alpha^2 - a_1\beta^2 + a_1 + \alpha^3 - 3\alpha\,\beta^2 + \alpha$$
$$= 0, \tag{20}$$

$$a_4 e^{-\alpha}\alpha\cos 1\sin\beta - a_4 e^{-\alpha}\beta\cos 1\cos\beta - a_4 e^{-\alpha}\sin 1\sin\beta + a_4\beta$$
$$- a_3 e^{-\alpha}\alpha^2\sin\beta + 2a_3 e^{-\alpha}\alpha\,\beta\cos\beta + a_3 e^{-\alpha}\beta^2\sin\beta$$
$$- a_3 e^{-\alpha}\sin\beta + 2a_2 e^{-\alpha/\sqrt{2}}\alpha\,\beta\cos(\beta/\sqrt{2})$$

$$-a_2 e^{-\alpha/\sqrt{2}}\alpha^2 \sin(\beta/\sqrt{2}) + a_2 e^{-\alpha/\sqrt{2}}\beta^2 \sin(\beta/\sqrt{2})$$

$$-a_2 e^{-\alpha/\sqrt{2}}\sin(\beta/\sqrt{2}) + 2a_1\alpha + 3\alpha^2\beta - \beta^3 + \beta = 0. \quad (21)$$

We again eliminate regions of the complex plane from consideration.

LEMMA 1. *In that portion of region 3 where $\beta \gg |\alpha| \gg 1$, Eqs. (20) and (21) simplify to*

$$2a_3 e^{-\alpha}\alpha\beta \sin\beta - a_3 e^{-\alpha}\beta^2 \cos\beta + 2a_2 e^{-\alpha/\sqrt{2}}\alpha\beta \sin(\beta/\sqrt{2})$$

$$- a_2 e^{-\alpha/\sqrt{2}}\beta^2 \cos(\beta/\sqrt{2}) - 3\alpha\beta^2 = \mathrm{DST}, \quad (22)$$

and

$$2a_3 e^{-\alpha}\alpha\beta \cos\beta + a_3 e^{-\alpha}\beta^2 \sin\beta + 2a_2 e^{-\alpha/\sqrt{2}}\alpha\beta \cos(\beta/\sqrt{2})$$

$$+ a_2 e^{-\alpha/\sqrt{2}}\beta^2 \sin(\beta/\sqrt{2}) - \beta^3 = \mathrm{DST}, \quad (23)$$

where DST stands for "demonstrably smaller terms."

Proof. Consider first the coefficients of $e^{-\alpha}\cos\beta$ among the terms appearing in Eq. (20), i.e.,

$$-\alpha_4\alpha\cos 1, \quad a_4\sin 1, \quad -\alpha_3\beta^2, \quad \alpha_3\alpha^2, \quad \text{and } a_3.$$

The largest of these is $-a_3\beta^2$, so we drop the other four. Similarly, of the two terms containing $e^{-\alpha}\sin\beta$, one is dropped; among terms containing $e^{-\alpha/\sqrt{2}}\cos(\beta/\sqrt{2})$, all but one are dropped; among the rest excluding the term with $e^{-\alpha/\sqrt{2}}\sin(\beta/\sqrt{2})$, only $3\alpha\beta^2$ needs to be retained. Similar simplifications are made for Eq. (21) (details omitted). ∎

PROPOSITION 4. *The large roots of Eqs. (20) and (21) have large α and lie only in the left half of region 3, Fig. 3.*

Proof. For large roots, $\alpha=\mathcal{O}(1)$ leaves β^3 unbalanced in Eq. (21). So α is large.

We first drop regions 1 and 2 from our consideration, as follows. In these regions, terms containing $e^{-\alpha}$ can be dropped, being smaller than other terms, giving

$$a_4\alpha + a_1\alpha^2 - a_1\beta^2 + a_1 + \alpha^3 - 3\alpha\beta^2 + \alpha = o(1), \quad (24)$$

$$a_4\beta + 2a_1\alpha\beta + 3\alpha^2\beta - \beta^3 + \beta = o(1). \quad (25)$$

In region 1, the largest term α^3 is unbalanced in Eq. (24). In region 2, the leading terms in Eqs. (24) and (25) give $\alpha^3 \sim 3\alpha\beta^2$ and $3\alpha^2\beta \sim \beta^3$, which have no nonzero solutions.

Now we drop the right half of region 3, as follows. By Lemma 1, Eqs. (22) and (23) hold in this region. Since $\alpha > 0$ in the right half, we have $0 < e^{-\alpha} < e^{-\alpha/\sqrt{2}} < 1$. This leaves the largest term β^3 unbalanced in Eq. (23).

We next consider regions 4 and 5. Here, $\alpha < 0$ (and large), so $e^{-\alpha} \gg e^{-\alpha/\sqrt{2}}$, which in turn is much greater than both α and β as well as algebraic powers thereof. Hence, retaining only the terms containing $e^{-\alpha}$ in Eqs. (20) and (21), we get

$$a_4 \sin 1\cos\beta - a_4\alpha\cos 1\cos\beta - a_4\beta\cos 1\sin\beta + a_3\alpha^2\cos\beta$$

$$+ 2a_3\alpha\beta\sin\beta - a_3\beta^2\cos\beta + a_3\cos\beta = \mathrm{DST}, \quad (26)$$

and

$$a_4\alpha\cos 1\sin\beta - a_4\beta\cos 1\cos\beta - a_4\sin 1\sin\beta - a_3\alpha^2\sin\beta$$

$$+ 2a_3\alpha\beta\cos\beta + a_3\beta^2\sin\beta - a_3\sin\beta = \mathrm{DST}. \quad (27)$$

In region 4, the quadratic terms in α and β dominate in Eqs. (26) and (27), giving (since $a_3 \neq 0$)

$$-\alpha^2 \sin\beta + 2\alpha\beta\cos\beta + \beta^2 \sin\beta = \mathrm{DST}, \quad (28)$$

$$\alpha^2 \cos\beta + 2\alpha\beta\sin\beta - \beta^2\cos\beta = \mathrm{DST}. \quad (29)$$

Multiplying Eq. (28) with $\sin\beta$, Eq. (29) with $\cos\beta$, and subtracting, we find

$$\beta^2 - \alpha^2 = \mathrm{DST},$$

which means $\beta \sim \alpha$. However, that in turn leads to the contradictory

$$\cos\beta = o(1) \quad \text{and} \quad \sin\beta = o(1), \quad \text{i.e.,} \quad \cos^2\beta + \sin^2\beta \ll 1. \quad (30)$$

In region 5, α^2 dominates in Eqs. (28) and (29), giving the same contradiction as Eq. (30). Hence, we conclude that asymptotically large roots lie only in the left half of region 3. ∎

PROPOSITION 5. *Equations (20) and (21) are satisfied by $\alpha = -\ln\beta + \ln|a_3| + o(1)$ and $\beta = (2N + \mathrm{sgn}(a_3)/2)\pi + o(1)$.*

Proof. By Proposition 4 and Lemma 1, we have the simplified equations, Eqs. (22) and (23) (reproduced below).

$$2a_3 e^{-\alpha}\alpha\beta \sin\beta - a_3 e^{-\alpha}\beta^2\cos\beta + 2a_2 e^{-\alpha/\sqrt{2}}\alpha\beta \sin(\beta/\sqrt{2})$$

$$- a_2 e^{-\alpha/\sqrt{2}}\beta^2\cos(\beta/\sqrt{2}) - 3\alpha\beta^2 = \mathrm{DST}, \quad (31)$$

and

$$2a_3 e^{-\alpha}\alpha\beta \cos\beta + a_3 e^{-\alpha}\beta^2\sin\beta + 2a_2 e^{-\alpha/\sqrt{2}}\alpha\beta \cos(\beta/\sqrt{2})$$

$$+ a_2 e^{-\alpha/\sqrt{2}}\beta^2\sin(\beta/\sqrt{2}) - \beta^3 = \mathrm{DST}. \quad (32)$$

Here $\alpha < 0$ and $|\alpha| \gg 1$ so that $e^{-\alpha} \gg e^{-\alpha/\sqrt{2}} \gg 1$. In Eq. (31), $e^{-\alpha}\beta^2\cos\beta$ is asymptotically larger than any other term unless $\cos\beta = o(1)$, whence $\sin\beta \sim \pm 1$. Using this in Eq. (32), the two largest terms have magnitudes $e^{-\alpha}\beta^2$ and β^3. This gives

$$a_3 e^{-\alpha}\beta^2\sin\beta - \beta^3 = \mathrm{DST}.$$

For $a_3 > 0$, we require $\sin\beta \sim 1$ which leads to

$$\beta = (2N + 1/2)\pi + o(1) \quad \text{and} \quad \alpha = -\ln\beta + \ln a_3 + o(1), \quad (33)$$

while for $a_3 < 0$, we require $\sin\beta \sim -1$ and hence

$$\beta = (2N - 1/2)\pi + o(1) \quad \text{and} \quad \alpha = -\ln\beta + \ln|a_3| + o(1),$$

for large integer N. ∎

We now find two correction terms in a formal series. The procedure is somewhat more complicated than before.

The second largest terms in Eqs. (22) and (23) are of $\mathcal{O}(\beta^{-(1-1/\sqrt{2})})$ compared to β^3. This suggests a formal series in powers of $N^{-(1-1/\sqrt{2})}$. However, there are terms of $\mathcal{O}(\beta^{-1}), \mathcal{O}(\beta^{-2})$, etc., whose powers are not integer multiples of $1 - 1/\sqrt{2}$, and so the formal series should have mixed powers of $N^{-(1-1/\sqrt{2})}$ and N^{-1}. We therefore anticipate a series of the form

$$\beta = (2N + 1/2)\pi + \frac{\beta_1}{N^{(1-1/\sqrt{2})}} + \frac{\beta_2}{N^{2(1-1/\sqrt{2})}} + \cdots + \frac{\beta_{11}}{N} + \frac{\beta_{12}}{N^2} + \cdots$$

$$+ \text{mixed powers.}$$

Note, however, that the first mixed power is of the form

$$\frac{1}{N^{1-1/\sqrt{2}}} \times \frac{1}{N} \ll \frac{1}{N},$$

and so if we retain only the first two small corrections, then we have the somewhat simpler expression

$$\beta = (2N + 1/2)\pi + \frac{\beta_1}{N^{(1-1/\sqrt{2})}} + \frac{\beta_2}{N^{2(1-1/\sqrt{2})}} + \text{h.o.t.} \quad (34)$$

where "h.o.t." stands for higher order terms. Note that $1/\sqrt{2} \approx 0.7$ and so (roughly) $N^{-(1-1/\sqrt{2})} \approx N^{-0.3}$. Similarly, $N^{-2(1-1/\sqrt{2})} \approx N^{-0.6} \gg N^{-1} \gg$ "mixed powers."

For analytical convenience, we take the leading order solution for α as $\alpha = -\ln\beta + z$, where $z \ll \ln\beta$ and express the correction z in a series as
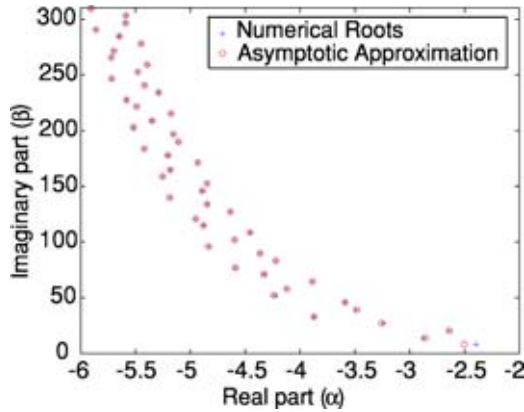
Fig. 5 Characteristic roots of Eq. (19) for $a_1=a_2=a_3=a_4=1$

$$z = \ln(a_3) + \frac{z_1}{N^{(1-1/\sqrt{2})}} + \frac{z_2}{N^{2(1-1/\sqrt{2})}} + \text{h.o.t.} \tag{35}$$

Substituting α in Eqs. (22) and (23) and dividing throughout by β^3, we get

$$-1 + a_3 e^{-z}\sin(\beta) + a_2\sin\left(\frac{\beta}{\sqrt{2}}\right)e^{-z/\sqrt{2}}\beta^{-(1-1/\sqrt{2})} + \mathcal{O}\left(\frac{\ln\beta}{\beta}\right) = 0, \tag{36}$$

$$a_3 e^{-z}\cos(\beta) + a_2\cos\left(\frac{\beta}{\sqrt{2}}\right)e^{-z/\sqrt{2}}\beta^{-(1-1/\sqrt{2})} + \mathcal{O}\left(\frac{\ln\beta}{\beta}\right) = 0. \tag{37}$$

Substituting Eqs. (34) and (35) in Eqs. (36) and (37), followed by expanding and collecting terms finally gives[3]

$$\beta_1 = 2^{(-2+\sqrt{2})/2}\pi^{(-2+\sqrt{2})/2}a_2 a_3^{-1/\sqrt{2}}\cos\left(\frac{(4N+1)}{2\sqrt{2}}\pi\right),$$

$$\beta_2 = \frac{1}{8\pi^2}\left[2^{\sqrt{2}}\pi^{\sqrt{2}}a_2^2 a_3^{-\sqrt{2}}(1-\sqrt{2})\sin\left(\frac{(4N+1)}{\sqrt{2}}\pi\right)\right],$$

$$z_1 = 2^{(-2+\sqrt{2})/2}\pi^{(-2+\sqrt{2})/2}a_2 a_3^{-1/\sqrt{2}}\sin\left(\frac{(4N+1)}{2\sqrt{2}}\pi\right),$$

$$z_2 = -\frac{1}{8\pi^2}\left(2^{\sqrt{2}}\pi^{\sqrt{2}}a_2^2 a_3^{-\sqrt{2}}(1-\sqrt{2})\cos\left(\frac{(4N+1)}{\sqrt{2}}\pi\right)\right).$$

Figure 5 shows some of the characteristic roots of Eq. (19) obtained numerically using the Newton-Raphson method along with the asymptotic approximations, which are in good agreement.

Remark 1. The smaller characteristic roots of DDEs can be found using Padé approximants ([24,25]). These, along with the asymptotics above, can give *all* the roots. The procedure requires arbitrary precision arithmetic (in, e.g., MAPLE). See the Appendix. These can be found using the Galerkin projection technique [26].

6 Asymptotics on Coefficients

In the above DDEs the term with the largest delay essentially determined the large roots. For DDEs with multiple delays, however, if the coefficients of delayed terms have disparate magni-

---

[3]Our calculations were done using MAPLE 6 (Windows), which, for these irrational powers, needs a little patience. We found it useful to do the expansion one term at a time. For each term, we divided by the (known) largest surviving power of N, and then asked for the limit as $N\rightarrow\infty$.

tudes, then the above asymptotics may only begin to hold for extremely large roots. In such cases, we could develop alternative asymptotic expansions. We present one example

$$\dot{x}(t) + \frac{x(t - 1/\sqrt{2})}{\epsilon} + ax(t - 1) = 0, \tag{38}$$

where $0 < \epsilon \ll 1$. Note that this is a special case of Eq. (19) with $a_2 = 1/\epsilon, a_3 = a$ and $a_1 = a_4 = 0$. Substituting $\lambda = \alpha + i\beta$ and separating real and imaginary parts, we get

$$\alpha + \frac{e^{-\alpha/\sqrt{2}}\cos(\beta/\sqrt{2})}{\epsilon} + ae^{-\alpha}\cos\beta = 0, \tag{39}$$

$$\beta - \frac{e^{-\alpha/\sqrt{2}}\sin(\beta/\sqrt{2})}{\epsilon} - ae^{-\alpha}\sin\beta = 0. \tag{40}$$

Remark 2. The above two equations each have three terms. However, there are regimes of root magnitudes for which, in each equation, two of the terms are much bigger than the third. Our analytical search for roots will focus on these regimes. In what follows, we adopt the following convention. If the mth and nth terms, with $m=1,2,3$ and $n=1,2,3$, are negligible compared to the other terms in Eqs. (39) and (40), respectively, then we refer to it as "Case (m,n)." The large root asymptotics developed in the previous section correspond to Case (2,2).

Remark 3. The leading order solution of Eq. (19) given by Eq. (33) also represents roots of Eq. (38), under Case (2,2). Substituting Eq. (33) in Eq. (40), the first and third terms turn out to be of $\mathcal{O}(N)$ while the second term is of $\mathcal{O}(N^{1/\sqrt{2}}/\epsilon)$. Hence, the second term is negligible for

$$N \gg N^{1/\sqrt{2}}/\epsilon, \text{ i.e., } N \gg \epsilon^{-(2+\sqrt{2})}.$$

Alternatively, the asymptotic expressions of Eqs. (34) and (35), for the roots of Eq. (19), are based on the correction term being $o(1)$. For Eq. (38), viewed as a special case of Eq. (19) (with $a_2 = 1/\epsilon$), substitution of β_1 found in Sec. 5 in the series solution for β, i.e., Eq. (34) gives the first correction term to be

$$\frac{2^{(-2+\sqrt{2})/2}\pi^{(-2+\sqrt{2})/2}a^{-1/\sqrt{2}}}{\epsilon N^{(1-1/\sqrt{2})}}\cos\left(\frac{(4N+1)}{2\sqrt{2}}\pi\right).$$

For the above to be $o(1)$, we require $N \gg \epsilon^{-(2+\sqrt{2})}$, matching the above.

The asymptotics developed below are therefore for N smaller than the above estimate. How much smaller is a somewhat tricky issue as discussed later.

By remark 2, our analytical search for the roots will concentrate on regimes where two of the three terms in each equation, i.e., Eqs. (39) and (40), are much larger than the third one. Accordingly we have nine different possibilities, one of which, Case (2, 2), has been dealt with in the previous section. Here, we check the remaining possibilities one by one for the existence of roots and obtain expressions for them.

PROPOSITION 6. *Case* (1,1) *yields a set of solutions for Eqs.* (39) *and* (40).

Proof. Equations (39) and (40) simplify to

$$\frac{e^{-\alpha/\sqrt{2}}\cos(\beta/\sqrt{2})}{\epsilon} = -ae^{-\alpha}\cos\beta + \text{ST}, \tag{41}$$

$$\frac{e^{-\alpha/\sqrt{2}}\sin(\beta/\sqrt{2})}{\epsilon} = -ae^{-\alpha}\sin\beta + \text{ST}, \tag{42}$$

where "ST" stands for an asymptotically smaller term. Squaring Eqs. (41) and (42) and adding, in the asymptotic limit we get

$$\frac{e^{-\sqrt{2}\alpha}}{\epsilon^2} = a^2 e^{-2\alpha}. \tag{43}$$

Hence, to leading order we have

$$\alpha = (2 + \sqrt{2})\ln(\epsilon a). \qquad (44)$$

Substituting α in Eqs. (41) and (42) and rearranging, we get in the asymptotic limit

$$a^{-(1+\sqrt{2})}\epsilon^{-(2+\sqrt{2})}\left[\cos(\beta/\sqrt{2}) + \cos\beta\right] = 0,$$

$$a^{-(1+\sqrt{2})}\epsilon^{-(2+\sqrt{2})}\left[\sin(\beta/\sqrt{2}) + \sin\beta\right] = 0.$$

Again squaring and adding we get

$$2a^{-(1+\sqrt{2})}\epsilon^{-(2+\sqrt{2})}\left[1 + \cos\left(\frac{\beta}{2+\sqrt{2}}\right)\right] = 0.$$

which requires $\cos(\beta/2 + \sqrt{2}) = -1$, giving

$$\beta = (2N+1)(2+\sqrt{2})\pi. \qquad (45)$$

Equations (44) and (45) give solutions for Eqs. (39) and (40). ∎

We further investigate these roots for small ϵ and large N as follows. Substituting Eqs. (44) and (45) in Eq. (40), we note that the first term is of $\mathcal{O}(N)$ while the second and the third term are of $\mathcal{O}(\epsilon^{-(2+\sqrt{2})})$. Hence, the assumption that the first term is negligible for Case (1,1) is valid as long as $N \ll \epsilon^{-(2+\sqrt{2})}$. Note that this is complementary to Case (2, 2) where $N \gg \epsilon^{-(2+\sqrt{2})}$ (by remark 3).

We proceed to a correction term. We add corrections β_1 and α_1 to the leading order solutions for β and α, i.e., in Eqs. (44) and (45), respectively. We substitute α and β in Eqs. (39) and (40), expand in a Taylor series about $\alpha_1 = \beta_1 = 0$, drop quadratic and higher order terms, and solve for β_1 and α_1 to get

$$\beta_1 = \frac{\epsilon^{2+\sqrt{2}}a^{1+\sqrt{2}}\left\{(4N+2)\pi\cos\left[(2+\sqrt{2})(2N+1)\pi\right] + 2\ln(a\epsilon)\sin\left[(2+\sqrt{2})(2N+1)\pi\right]\right\}}{-3 + 2\sqrt{2}},$$

$$\alpha_1 = \frac{\epsilon^{2+\sqrt{2}}a^{1+\sqrt{2}}\left\{(4N+2)\pi\sin\left[(2+\sqrt{2})(2N+1)\pi\right] - 2\ln(a\epsilon)\cos\left[(2+\sqrt{2})(2N+1)\pi\right]\right\}}{-3 + 2\sqrt{2}}.$$

Since we are interested in $N \ll \epsilon^{-(2+\sqrt{2})}$ for this case we find, self-consistently, that $\beta_1 = o(1)$ and $\alpha_1 = o(1)$. This concludes Case (1,1).

Remark 4. In the above, we established the range of β (or equivalently N) over which the asymptotic expressions hold, i.e., $\beta \ll \epsilon^{-(2+\sqrt{2})}$. In what follows, such sharp estimates on β may not always be easily obtainable. Sometimes, for simplicity, we will fix β at some ϵ-independent range of magnitudes and take the limit as $\epsilon \to 0$. Eventually, numerics will bear out the final analytical approximations.

LEMMA 2. *If r is an irrational number and either $\sin\theta = 0$ or $\cos\theta = 0$, then $\sin(r\theta) \neq 0$ and $\cos(r\theta) \neq 0$.*

Proof. $\sin\theta = 0$ requires $\theta = n\pi$, for some integer n. However, $\sin(r\theta) = 0$ requires $r\theta = m\pi$, for some integer m. Simultaneous satisfaction of both the above requirements imply, contradictorily, that $r = m/n$. Similar contradictions arise for the other cases. ∎

PROPOSITION 7. *Case (2,1) can be eliminated.*

Proof. From Eqs. (39) and (40), we have

$$\alpha = -ae^{-\alpha}\cos\beta + \text{ST}, \qquad (46)$$

$$\frac{e^{-\alpha/\sqrt{2}}\sin(\beta/\sqrt{2})}{\epsilon} = -ae^{-\alpha}\sin\beta + \text{ST}. \qquad (47)$$

Squaring and adding, we get in the asymptotic limit

$$\frac{e^{-\sqrt{2}\alpha}\sin^2(\beta/\sqrt{2})}{\epsilon^2} + \alpha^2 = a^2e^{-2\alpha}. \qquad (48)$$

There are three possibilities now: $\alpha \gg 1$, $\alpha = \mathcal{O}(1)$, and $\alpha \ll -1$.

1. If $\alpha \gg 1$, the right-hand side (RHS) of Eq. (48), i.e., $a^2e^{-2\alpha} = o(1)$. However, in the LHS, $\alpha^2 \gg 1$ while the first term is positive, giving a contradiction.
2. If $\alpha = \mathcal{O}(1)$, the first term remains unbalanced in Eq. (48) unless $\sin^2(\beta/\sqrt{2}) \ll 1$. However, Case (2,1) implies for Eq. (39) that $\cos^2(\beta/\sqrt{2}) \ll 1$ as well, giving a contradiction because $\sin^2(.) + \cos^2(.) = 1$.
3. If $\alpha \ll -1$, then $\alpha^2 \ll e^{-2\alpha}$ and Eq. (48) simplifies to

$$\frac{e^{-\sqrt{2}\alpha}\sin^2(\beta/\sqrt{2})}{\epsilon^2} = a^2e^{-2\alpha} + \text{ST}. \qquad (49)$$

Also, Case (2,1) implies for Eq. (39) that $\cos(\beta/\sqrt{2}) = o(1)$, when $\sin^2(\beta/\sqrt{2}) \sim 1$. In that case, Eq. (49) simplifies to Eq. (43) which has Eq. (44) as the leading order solution for α. Substituting Eq. (44) for α in Eq. (47), we require in the asymptotic limit

$$\sin(\beta/\sqrt{2}) + \sin\beta = 2\sin\left(\frac{2+\sqrt{2}}{4}\beta\right)\cos\left(\frac{2-\sqrt{2}}{4}\beta\right) = 0,$$

which contradicts our prior conclusion that $\cos(\beta/\sqrt{2}) = 0$ in the asymptotic limit[4] (by Lemma 2).

Thus, Case (2,1) gives no roots. ∎

PROPOSITION 8. *Cases (3,1), (1,2), and (3,2) can also be eliminated.*

Proof. The proof resembles that of Proposition 7 and is not given here to save space. ∎

PROPOSITION 9. *The three Cases $(m,3)$ $m = 1$, 2, 3 lead to one set of solutions for Eqs. (39) and (40).*

Proof.

1. For Case (1,3), from Eqs. (39) and (40) we get

$$\frac{e^{-\alpha/\sqrt{2}}\cos(\beta/\sqrt{2})}{\epsilon} = -ae^{-\alpha}\cos\beta + \text{ST}, \qquad (50)$$

$$\beta = \frac{e^{-\alpha/\sqrt{2}}\sin(\beta/\sqrt{2})}{\epsilon} + \text{ST}. \qquad (51)$$

Note that Case $(m,3)$ implies that $\beta \gg ae^{-\alpha}|\sin\beta|$. Accordingly, we have two subcases, in the asymptotic limit, viz., $\sin\beta = 0$ and $\sin\beta \neq 0$.

---

[4]Note that we have considered β fixed as $\epsilon \to 0$. An argument allowing β to grow as $\epsilon \to 0$ can be developed, but is trickier and avoided here.
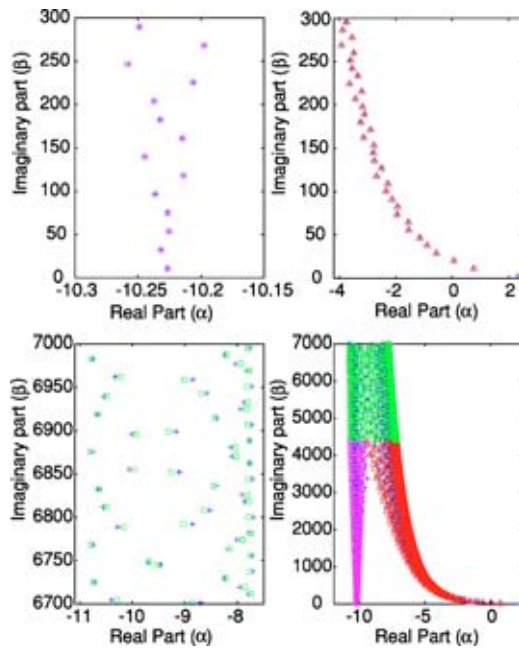
Fig. 6 Roots of Eqs. (39) and (40) for a=1 and ϵ=0.05. Plus signs: Newton-Raphson. Circles: asymptotic, Case (1,1). Triangles: asymptotic, Case $(m,3)$. Rectangles: asymptotic, Case (2,2).

- If $\sin\beta=0$ in the asymptotic limit, we have $\cos\beta=\pm1$. In that case, solving Eq. (50) for α, we get

$$\alpha=(2+\sqrt{2})\ln\left(\frac{\epsilon a}{|\cos(\beta/\sqrt{2})|}\right).$$

Substituting the above in Eq. (51), we will require

$$\epsilon^{(2+\sqrt{2})}a^{(1+\sqrt{2})}\beta=\sin(\beta/\sqrt{2})\cos^{(1+\sqrt{2})}(\beta/\sqrt{2}).$$

In the asymptotic limit as $\epsilon\to0$, we will require either $\sin(\beta/\sqrt{2})=0$ or $\cos(\beta/\sqrt{2})=0$ which contradicts $\sin\beta=0$ (by lemma 2).

- If $\sin\beta\neq0$ in the asymptotic limit, we have $\beta\gg ae^{-\alpha}$. Also, from Eq. (51), we have $e^{-\alpha/\sqrt{2}}/\epsilon\gg\beta\gg ae^{-\alpha}$. Hence, for balance in Eq. (50), we require $\cos(\beta/\sqrt{2})=0$ (in the asymptotic limit), giving

$$\beta=(2N+1/2)\sqrt{2}\pi+o(1). \qquad (52)$$

Substituting the above for β in Eq. (51) and solving for α, we get

$$\alpha=-\sqrt{2}\ln[\epsilon(2N+1/2)\sqrt{2}\pi]+o(1). \qquad (53)$$

This concludes Case (1, 3).
2. For Case (2,3), we have

$$\alpha=-ae^{-\alpha}\cos\beta+\text{ST}, \qquad (54)$$

$$\beta=\frac{e^{-\alpha/\sqrt{2}}\sin(\beta/\sqrt{2})}{\epsilon}+\text{ST}. \qquad (55)$$

Again $\beta\gg ae^{-\alpha}|\sin\beta|$, giving two subcases as before.

- If $\sin\beta=0$, we have $\cos\beta=\pm1$ and Eq. (54) reduces, in the

Digits used	6	8	10
First root	1.0452	1.044 639 8	1.044 6435 3

asymptotic limit, to

$$\alpha\pm ae^{-\alpha}=0.$$

The above has at most one solution for α which, in turn, is $\mathcal{O}(1)$. In that case, $e^{-\alpha/\sqrt{2}}=\mathcal{O}(1)$ and Case (2,3) requires $\cos(\beta/\sqrt{2})=0$ contradicting $\sin\beta=0$.

- If $\sin\beta\neq0$ in the asymptotic limit, we have $\beta\gg ae^{-\alpha}$. Also from Eq. (55), we have $e^{-\alpha/\sqrt{2}}/\epsilon\gg\beta$. Now, Case $(2,n)$ requires $[e^{-\alpha/\sqrt{2}}\cos(\beta/\sqrt{2})]/\epsilon\ll ae^{-\alpha}\ll\beta$. This is possible only when $\cos(\beta/\sqrt{2})=0$ giving Eqs. (52) and (53) as the solutions again.

3. For Case (3,3), we get from Eqs. (39) and (40)

$$\alpha=-\frac{e^{-\alpha/\sqrt{2}}\cos(\beta/\sqrt{2})}{\epsilon}+\text{ST}, \qquad (56)$$

$$\beta=\frac{e^{-\alpha/\sqrt{2}}\sin(\beta/\sqrt{2})}{\epsilon}+\text{ST}. \qquad (57)$$

In the asymptotic limit, Eqs. (56) and (57) are the same as Eqs. (7) and (8) with $a=1/\epsilon$ and the delay being $1/\sqrt{2}$ instead of 1. From Proposition 1, we get the same expressions for α and β as given in Eqs. (52) and (53).

Thus, all cases $(m,3),m=1$, 2, 3 give the same solutions given by Eqs. (52) and (53). ∎

Remark 5. All three cases $(m,3)$ lead to the same solution set. A finer analysis of small terms might tell them apart. Note also that these solutions rest on $\beta\gg1$, which implies "large" roots (though not larger than that allowed by Remark 4). As will be seen in numerics, all but a small number of roots are in fact captured accurately.

Remark 6. In the above, roots in different magnitude regimes were captured by different scalings of the nominally small/large expansion terms. Equations (52) and (53) apply for N somewhat large, but not larger than $\epsilon^{-(2+\sqrt{2})}$. Moreover, there ϵ itself is small. Elsewhere, we have held N fixed as $\epsilon\to0$. Below, to obtain correction terms to the leading order solutions, we set $\epsilon=\Delta/N^{1-1/\sqrt{2}}$, where $\Delta=o(1)$ since $N\ll\epsilon^{-(2+\sqrt{2})}$. However, now Δ will be held constant while we consider asymptotics for large N. These different scalings are motivated by mathematical convenience alone, always keeping in mind that we are finally interested in some finite nonzero values of ϵ and N, and it matters little how we get there. In final justification of these ideas, numerics and approximations will match below.

For finding a correction term, we first scale ϵ as mentioned in Remark 6, then substitute $\beta=(2N+1/2)\pi\sqrt{2}+\beta_1$ and $\alpha=-\sqrt{2}\ln[\epsilon(2N+1/2)\pi\sqrt{2}]+\alpha_1$ in Eqs. (39) and (40), expand in a Taylor series about $\alpha_1=\beta_1=0$ until first order, solve for α_1 and β_1, retain the largest power of N (a key step in simplifying very long expressions), and reinsert $\Delta=\epsilon N^{1-1/\sqrt{2}}$ to finally get

$$\beta_1 = \frac{2^{1+3/\sqrt{2}}\pi^{1+\sqrt{2}}\epsilon^{\sqrt{2}}a\cos[(2N+1/2)\pi\sqrt{2}]}{4\pi^2 N^{(1-\sqrt{2})} + 2^{3\sqrt{2}}\epsilon^{2\sqrt{2}}\pi^{2\sqrt{2}}a^2 N^{(\sqrt{2}-1)} + 2^{2+3/\sqrt{2}}\pi^{1+\sqrt{2}}\epsilon^{\sqrt{2}}a\sin[(2N+1/2)\pi\sqrt{2}]}, \tag{58}$$

$$\alpha_1 = \frac{2^{1+3/\sqrt{2}}\pi^{1+\sqrt{2}}\epsilon^{\sqrt{2}}a\sin((2N+1/2)\pi\sqrt{2}) + 2^{3\sqrt{2}}\pi^{2\sqrt{2}}\epsilon^{2\sqrt{2}}a^2 N^{\sqrt{2}-1}}{4\pi^2 N^{(1-\sqrt{2})} + 2^{3\sqrt{2}}\epsilon^{2\sqrt{2}}\pi^{2\sqrt{2}}a^2 N^{(\sqrt{2}-1)} + 2^{2+3/\sqrt{2}}\pi^{1+\sqrt{2}}\epsilon^{\sqrt{2}}a\sin[(2N+1/2)\pi\sqrt{2}]}. \tag{59}$$

Figure 6 shows the converged roots obtained from Newton-Raphson iterations, each starting from a different initial point on a large, uniform grid (exhaustive search); as well as the above analytical estimates. Agreement is good except for a single root near the origin.

7 Conclusions

We have obtained asymptotic approximations for the roots of the characteristic equations of some linear DDEs with constant coefficients. The term with the largest delay dominates in the asymptotic expansions for the large roots. A few smaller roots obtained using a Padé approximant can complement these asymptotic expressions to give all the characteristic roots of a DDE. However, if a very large coefficient is associated with a term with a smaller delay (equivalently, the term with the largest delay has a small coefficient), the large root asymptotics are useful only for extremely large roots. For such cases, determining the remaining roots using the Padé approximant is impractical and alternative asymptotic expansions have been developed. This study provides practical insight into the location of characteristic roots of DDEs on the complex plane, and may be useful for further theoretical studies as well.

Table 2 First six roots of Eq. (5) for $a=1$

No.	Padé approximant	Newton-Raphson
1	$-0.3181+1.3372i$	$-0.3181+1.3372i$
2	$-2.0623+7.5886i$	$-2.0623+7.5886i$
3	$-2.6532+13.9492i$	$-2.6532+13.9492i$
4	$-3.0202+20.2465i$	$-3.0202+20.2725i$
5	$-3.2878+26.5805i$	$-3.2878+26.5805i$
6	$-3.4997+32.8805i$	$-3.4985+32.8807i$

Table 3 First six roots of Eq. (4) for $p=2, \psi=0.1$

No.	Padé approximant	Newton-Raphson
1	-0.44008	-0.44008
2	$-0.5762+2.4326i$	$-0.5762+2.4326i$
3	$-3.7516+8.5961i$	$-3.7516+8.5961i$
4	$-4.8189+15.095i$	$-4.8190+15.095i$
5	$-5.4984+21.4961i$	$-5.4984+21.4962i$
6	$-6.0014+27.8552i$	$-6.0014+27.8553i$

Table 4 First six roots of Eq. (19) for $a_1=a_2=a_3=a_4=1$

No.	Padé approximant	Newton-Raphson
1	$-0.1639+2.4749i$	$-0.1639+2.4749i$
2	$-2.3946+8.2369i$	$-2.3946+8.2369i$
3	$-2.8736+13.6232i$	$-2.8736+13.6232i$
4	$-2.6442+20.2466i$	$-2.6442+20.2466i$
5	$-3.2404+27.0141i$	$-3.2404+27.0141i$
6	$-3.8715+32.8327i$	$-3.8703+32.8338i$

Appendix: Smaller Roots

We use MAPLE, to find Padé approximants (as in [24,25]) to obtain the smaller roots of the characteristic equations of the DDEs. MAPLE does both symbolic algebra and arbitrary-precision floating point arithmetic. To see the issues involved, consider

$$p - \sqrt{2}e^{-p} - e^{-p/\sqrt{3}} = 0. \tag{A1}$$

Expanding in a Taylor series, we obtain

$$-1 - \sqrt{2} + \left(\sqrt{2} + 1 + \frac{1}{\sqrt{3}}\right)p - \left(\frac{1}{\sqrt{2}} + \frac{1}{6}\right)p^2 + \cdots$$

Retaining terms up to p^{10}, we obtain the Padé approximant of order (5,5). We seek the zeroes of this approximant. The numerator is of the form $\Sigma_{k=0}^5 c_k p^k$, where c_0 is

$$6\,688\,184\,704\,014\,240 - 1\,634\,855\,556\,025\,440\sqrt{6}$$

$$+ 2\,840\,749\,923\,049\,920\sqrt{2} - 3\,855\,543\,089\,257\,440\sqrt{3}, \tag{A2}$$

and the other coefficients are comparably lengthy. The first few roots of the numerator polynomial give good approximations to the first few roots of Eq. (A1). However, the accuracy of the Padé roots increase with the digits of precision used in the floating point arithmetic. Results for the smallest root of Eq. (A1), which is 1.044 643 69, are given in Table 1.

In practice, to find several roots of a DDE, high order Padé approximants need to be used. The numerator polynomial then involves long/large coefficients. We therefore numerically evaluate the coefficients of the Taylor series at the start, before Padé approximants are calculated. Moreover, many floating point digits are needed for accurate results.

For the three examples studied in this paper, we used 650 digits of precision (numerical inaccuracies were observed with our previous choice of 540 digits; no optimization was done on number of digits) and took a Padé approximant of order either (21, 21) or (22, 22) depending on the number of real roots obtained. Results, in Tables 2–4, show good agreement.

References

[1] Stépán, G., 1997, "Delay Differential Equation Models for Machine Tool Chatter," *Dynamics and Chaos in Manufacturing Processes*, F. C. Moon, ed., Wiley, New York, pp. 165–191.
[2] Olgac, N., Elmali, H., Hosek, M., and Renzulli, M., 1997, "Active Vibration Control of Distributed Systems Using Delayed Resonator With Acceleration Feedback," ASME J. Dyn. Syst., Meas., Control, **119**, pp. 380–389.
[3] Santos, O., and Mondié, S., 2000, "Control Laws Involving Distributed Time Delays: Robustness of Implementation," *Proc. Amer. Control Conf*, Chicago, IL, IEEE, Piscataway, NJ, pp. 2479–2480.
[4] Insperger, T., and Stépán, G., 2000, "Remote Control of Periodic Robot Motion," *Proc. Thirteenth Sympos. on Theory and Practice of Robots and Manipulators*, Zakopane, pp. 197–203.
[5] Batzel, J. J., and Tran, H. T., 2000, "Stability of The Human Respiratory Control System. I: Analysis of a Two-Dimensional Delay State-Space Model," J. Math. Biol. **41**, pp. 45–79.
[6] Szydlowski, M., and Krawiec, A., 2001, "The Kaldor-Kalecki Model of Busi-

ness Cycle as a Two-Dimensional Dynamical System," J. Nonlinear Math. Phys. **8**, pp. 266–271.

[7] Kalmár-Nagy, T., Stépán, G., and Moon, F. C., 2001, "Subcritical Hopf Bifurcation in the Delay Equation Model for Machine Tool Vibrations," Nonlinear Dyn., **26**, pp. 121–142.

[8] Epstein, I. R., 1992, "Delay Effects and Differential Delay Equations in Chemical Kinetics," Int. Rev. Phys. Chem., **11**(1), pp. 135–160.

[9] Roussel, M. R., 1998, "Approximate State-Space Manifolds Which Attract Solutions of Systems of Delay-Differential Equations," J. Chem. Phys., **109**(19), pp. 8154–8160.

[10] Bellman, R., and Cooke, K. L., 1963, *Differential Equations*, Academic, New York.

[11] Driver, R. D., 1977, *Ordinary and Delay Differential Equations*, Springer-Verlag, New York.

[12] Hale, J. K., and Lunel, S. V., 1993, *Introduction to Functional Differential Equations*, Springer-Verlag, New York.

[13] Gopalsamy, K., 1992, *Stability and Oscillations in Delay Differential Equations of Population Dynamics*, Kluwer Academic Publishers, Dordrecht.

[14] Stépán, G., 1989, *Retarded Dynamical Systems*, Longman Group, UK.

[15] Bhatt, S. J., and Hsu, C. S., 1966, "Stability Criteria for Second-Order Dynamical Systems with Time Lag," ASME J. Appl. Mech., **33**(1), pp. 113–118.

[16] Bhatt, S. J., and Hsu, C. S., 1966, "Stability Charts for Second-Order Dynamical Systems with Time Lag," ASME J. Appl. Mech., **33**(1), pp. 119–124.

[17] Hassard, B. D., 1997, "Counting Roots of the Characteristic Equation for Linear Delay-Differential Systems," J. Diff. Eqns., **136**, pp. 222–235.

[18] Diekmann, O., Gils, S. V., Lunel, S. V., and Walther, H., 1995, *Delay Equations: Functional-, Complex-, and Nonlinear Analysis*, Springer-Verlag, New York.

[19] Breda, D., Maset, S. and Vermiglio, R., 2004, "Computing the Characteristic Roots for Delay Differential Equations," IMA J. Numer. Anal., **24**, pp. 1–19.

[20] Engelborghs, K., and Roose, D., 2002, "On Stability of LMS Methods and Characteristic Roots of Delay Differential Equations," SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal., **40**(2), pp. 629–650.

[21] Sandquist, G. M., and Rogers, V. C., 1979, "Graphical Solutions for the Characteristic Roots of the First Order Linear Differential-Difference Equation," ASME J. Dyn. Syst., Meas., Control, **101**, pp. 37–43.

[22] Pontryagin, L. S., 1955, "On the Zeros of Some Elementary Transcendental Functions," Am. Math. Soc. Transl. II, **1**, pp. 95–110.

[23] Minorsky, N., 1942, "Self-Excited Oscillations in Dynamical Systems Possessing Retarded Actions," ASME J. Appl. Mech., **9**, pp. A65–A71.

[24] Lam, J., 1993, "Model Reduction of Delay Systems Using Padé Approximants," Int. J. Control, **57**(2), pp. 377–391.

[25] Wang, Z., and Hu, H., 1999, "Robust Stability Test for Dynamic Systems with Short Delays by Using Padé Approximation," Nonlinear Dyn. **18**, pp. 275–287.

[26] Wahi, P., and Chatterjee, A., 2005, "Galerkin Projections for Delay Differential Equations," ASME J. Dyn. Syst., Meas., Control, **127**(1), pp. 80–87.

A Representation of Anisotropic Creep Damage in Fiber Reinforced Composites

D. N. Robinson
Professor Emeritus
e-mail: drobins7@tampabay.rr.com

W. K. Binienda[1]
Professor
e-mail: wbinienda@uakron.edu

Civil Engineering Department,
The University of Akron,
Akron, OH 44325-3905

A creep damage model is presented that allows for anisotropic distributions of damage in composite materials. An earlier model by the writers allowed for anisotropic damage growth rate but, based on a scalar state variable, failed to account for anisotropic distributions of damage. A vectorial state variable is introduced that allows a representation of anisotropic damage distribution. As in earlier work, a fundamental assumption is that the principally damaging stress components are tensile traction and longitudinal shear at the fiber/matrix interface. Application of the creep damage model is made to calculations involving homogenously stressed composite elements under transverse tensile and longitudinal shear stress and to cross plied thin-walled tubes under tension/torsion. Although the emphasis is phenomenological, with focus on a mathematical structure for representing anisotropic distributions of damage, a meaningful creep damage model must rest on fundamental material science and microstructural examination. Verification experiments involving tension/torsion testing of thin-walled composite tubes together with detailed microstructural examination are discussed and outlined. [DOI: 10.1115/1.1875512]

1 Introduction

Grobstein [1] observes anisotropic damage growth at the fiber/matrix interface of a W/Nb composite (MMC) under low stress creep conditions, cf. Fig. 1(*a*). Grobstein's results show a pronounced directionality of damage that is evidently influenced by the presence of transverse tensile stress at the fiber/matrix interface.

The overall microstructural observations of Grobstein [1,2] are summarized schematically in Fig. 1(*b*). With the passage of time under transverse stress, interfacial defects (voids) appear to grow and eventually coalesce, leading to increasing degradation of the interface. The damage distribution at any stage is highly directional, the major damage accumulating on interfacial tangent planes that lie essentially normal to the transverse stress direction. It is conjectured that the physical mechanisms involved may include phenomena such as Kirkendall porosity [3] and the evolution of an interphase [1,2].

This description of Grobstein's observations on the scale of fibers (e.g., on the mesoscale) could similarly be made regarding intergranular void growth in monolithic metallic alloys on the microscale (cf. Hull and Rimmer [4], Chuang and Rice [5], Cocks and Ashby [6]). Although the physical mechanisms on the two scales are very different, the schematic picture, Fig. 1(*b*), may still apply. The damage distribution is highly directional and the major damage accumulates on interfacial planes (grain boundaries) that lie essentially normal to the applied stress direction.

Intuition suggests that an analogous description may also hold for polymer matrix composites (PMC), now on a molecular scale, with Fig. 1(*b*) again applying qualitatively. Again, the scale and physical mechanisms may be vastly different but mesoscopic features such as a highly anisotropic damage distribution and the strong influence of transverse tensile stress (and longitudinal shear stress) at the fiber/matrix interface may remain essentially the same.

The objective of this research is to develop a phenomenological, anisotropic creep damage model for fiber reinforced composites (MMC, PMC, etc.) that is consistent with the general features of Fig. 1. The resulting model is an extension of Robinson et al. [7,8] and Binienda et al. [9]. The earlier model accounts for an anisotropic rate of damage growth but, resting on a *scalar state variable* for damage, fails to account for an anisotropic distribution as characterized in Fig. 1. The present model allows for an anisotropic distribution of damage through introduction of a *vectorial state variable*, i.e., a quantity that associates a scalar with each direction in space.

Although the present emphasis is phenomenological, with focus on a mathematical structure for representing anisotropic distributions of damage, it is recognized that a generally applicable damage model for a specific composite (MMC, PMC, etc.) must also rest on accepted material science and detailed microstructural examination. Macroscopic experiments supporting the present modeling must ultimately be accompanied by microscopic evidence of damage as Fig. 1(*a*).

Application of the creep damage model is made to homogenously stressed composite elements under transverse tensile and longitudinal shear stress and to cross plied thin-walled tubes under tension/torsion.

2 Anisotropic Creep Damage Distribution

Figure 2(*a*) is a schematic illustration of a (transversely isotropic) composite element presumed to contain a large number of fibers with orientation defined by the unit vector $\bar{d} = (1,0,0)$-out of page (cf. Fig. 2(*b*)). The element is stressed by a transverse tensile stress σ applied at time $t=0$, at which time the element is undamaged. Figure 2(*b*) introduces a unit circle c defined by the unit vector $\bar{n}$ with $\bar{n} \cdot \bar{d} = 0$. The direction $\bar{n}$ (or $-\bar{n}$) designates the normal to a tangent plane at the interface of a generic fiber in the composite element (cf. Fig. 2(*a*)).

$\Psi(\bar{n},t)$ represents the Kachanov [10] continuity at time t asso-

(a)

Fig. 1 (*a*) **Microphotograph of the fiber/matrix interface of W/Nb composite [2]; shows creep damage and nominal stress direction.** (*b*) **Schematic diagram of creep damage progression observed in [2].**

ciated with the interfacial tangent plane whose normal is $\bar{n}$. Kachanov introduces the concept of actual stress in the context of isotropic damage, i.e., $\sigma_A = \sigma/\Psi$, and asserts that the rate of damage growth is determined principally by the level of the actual stress. Here, the actual stress is directional $\sigma_A(\bar{n},t) = \sigma/\Psi(\bar{n},t)$
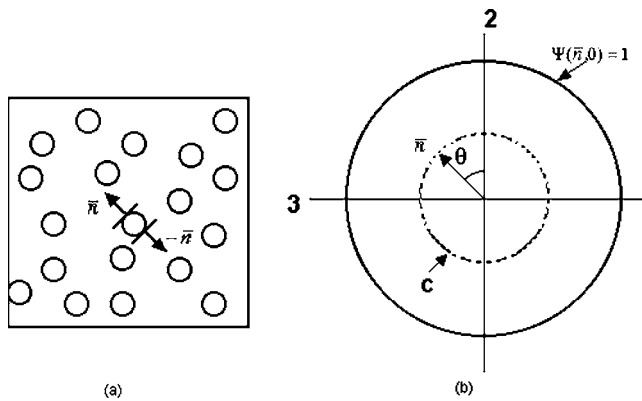
Fig. 2 (*a*) **Transversely isotropic composite element; fiber direction out of page. Unit vector $\bar{n}$ denotes normal to interfacial tangent planes at fiber/matrix interface.** (*b*) **Definition of (1,2,3) coordinates. Unit circle c (dashed) defined by $\bar{n}$. Kachanov continuity distribution $\Psi(\bar{n},0)$ mapped onto unit circle c. Undamaged $\Psi = 1$.**
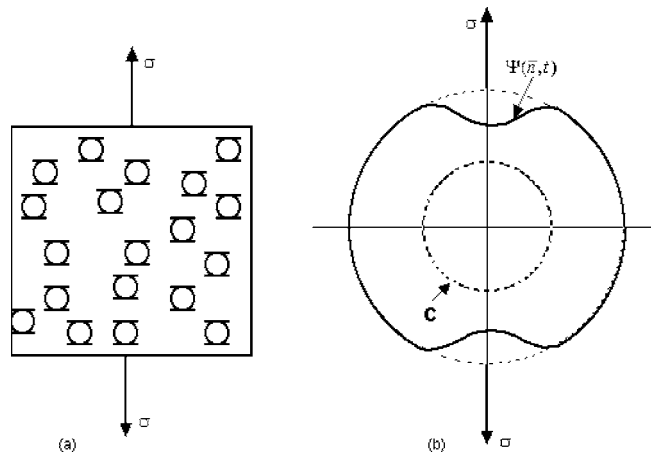
Fig. 3 (*a*) **Composite element showing directional damage—normal to stress direction.** (*b*) **Continuity distribution (solid curve) in damaged state $\Psi(\bar{n},t)$ mapped onto unit circle.**

differing on interfacial tangent planes. $\Psi(\bar{n},t)=1$ implies full continuity (no damage); $\Psi(\bar{n},t)=0$ indicates total loss of load carrying capacity across the plane with normal $\bar{n}$. The distribution of $\Psi(\bar{n},t)$ can be mapped onto the unit circle c. The distribution in Fig. 2(*b*) represents the undamaged state ($t=0$) while that in Fig. 3(*b*) represents a damaged state at $t>0$ biased by the presence of the transverse stress σ.

Onat [11,12] and Onat and Leckie [13] represent anisotropic damage distributions using expansions of irreducible tensors. They map the distribution at a material element onto the unit sphere (cf. spherical harmonics). Their representation is coordinate frame indifferent and compatible with continuum theory.

Guided by Onat and Leckie, we formulate an analogous tensorial expansion of $\Psi(\bar{n},t)$ that reflects the inherent transverse isotropy of a composite element (Fig. 2(*a*)). Our model amounts to a reduction from spherical to cylindrical coordinates, i.e., from mapping the damage distribution on the unit sphere as in [13] to mapping on the unit circle c (Fig. 2(*b*)). The formulation remains frame independent in the plane of transverse isotropy with the local fiber direction aligned in the 1 coordinate direction.

We propose the expansion:

$$\Psi(\bar{n},t) = \psi_o(t) + \psi_{ij}(t)f_{ij}(\bar{n}) + \psi_{ijkl}(t)f_{ijkl}(\bar{n}) + \cdots \quad i,j,k,l=1,2,3 \tag{1}$$

in which the usual summation convention is used. The set of functions $1, f_{ij}(\bar{n}), f_{ijkl}(\bar{n}), \ldots$ are orthogonal basis functions and $\psi_o(t), \psi_{ij}(t), \psi_{ijkl}(t), \ldots$ are the Fourier coefficients. As $\Psi(\bar{n},t) = \Psi(-\bar{n},t)$, the basis functions involve only even rank tensors, they are defined as:

$$f_{ij}(\bar{n}) = n_i n_j - \tfrac{1}{2}\delta_{ij}$$

$$f_{ijkl}(\bar{n}) = n_i n_j n_k n_l - \tfrac{1}{6}(\delta_{ij}n_k n_l + \delta_{ik}n_j n_l + \delta_{il}n_j n_k + \delta_{jk}n_i n_l + \delta_{jl}n_i n_k + \delta_{kl}n_i n_j) + \tfrac{1}{24}(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{2}$$

The basis functions are traceless, i.e., $f_{ii}=0$, $f_{ijji}=0 \ldots$ and have the symmetry properties $f_{ij}=f_{ji}$; $f_{ijkl}=f_{klij}=f_{ijlk}=f_{jikl}$; Also, $\delta_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

The Fourier coefficients are given by

$$\psi_o(t) = \frac{1}{2\pi}\int_C \Psi(\bar{n},t)d\bar{n}$$

$$\psi_{ij}(t) = \frac{1}{\pi} \int_C \Psi(\bar{n},t) f_{ij}(\bar{n}) d\bar{n} \tag{3}$$

$$\psi_{ijkl}(t) = \frac{1}{\pi} \int_C \Psi(\bar{n},t) f_{ijkl}(\bar{n}) d\bar{n}$$

Now we adopt a definite coordinate system 1,2,3 as in Fig. 2(b) with the fiber direction $\bar{d} = (1,0,0)$ along the 1 direction. The coordinate directions 2,3 define the plane of transverse isotropy; we choose $\bar{n}(\theta) = (0, \cos \theta, \sin \theta)$. With these definitions, the expansion (1) and the basis functions (2) reduce to linear combinations of trigonometric functions involving even multiples of θ. Moreover, it is readily shown that (1)–(3) reduce to a standard Fourier series representation of $\Psi(\bar{n}(\theta),t)$ over $-\pi \leqslant \theta \leqslant \pi$. Thus

$$\Psi(\theta,t) = \psi_0(t) + \psi_{2C}(t) \cos 2\theta + \psi_{2S}(t) \sin 2\theta + \psi_{4C}(t) \cos 4\theta$$
$$+ \psi_{4S}(t) \sin 4\theta + \cdots \tag{4}$$

with the Fourier coefficients given by

$$\psi_0(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Psi(\theta,t) d\theta$$

$$\psi_{2C}(t) = \frac{1}{\pi} \int_{-\pi}^{\pi} \Psi(\theta,t) \cos 2\theta d\theta \quad \psi_{2S}(t) = \frac{1}{\pi} \int_{-\pi}^{\pi} \Psi(\theta,t) \sin 2\theta d\theta \tag{5}$$

$$\psi_{4C}(t) = \frac{1}{\pi} \int_{-\pi}^{\pi} \Psi(\theta,t) \cos 4\theta d\theta \quad \psi_{4S}(t) = \frac{1}{\pi} \int_{-\pi}^{\pi} \Psi(\theta,t) \sin 4\theta d\theta$$

Although the frame indifferent representation (1)–(3) is generally applicable, we adopt the simple Fourier series form (4) and (5) in the remaining development and in subsequent applications.

3 Flow and Damage Evolutionary Laws

The governing equations are taken from Robinson et al. [7,8,14], Binienda et al. [9] and [15] with appropriate modifications. Thus,

$$\frac{\dot{\varepsilon}_{ij}}{\dot{\varepsilon}_0} = \frac{3}{2} \Phi^{p-1} \frac{\Gamma_{ij}}{\sigma_0} \frac{1}{\psi_0^p} \tag{6}$$

$$\dot{\Psi}(\theta,t) = -A \frac{\Delta^\nu}{\Psi^m(\theta,t)} \tag{7}$$

$\dot{\varepsilon}_{ij}$ is the creep deformation rate, Φ and Γ_{ij} are as defined in [9]. σ_0 is a reference stress, $\dot{\varepsilon}_0$, p, A, ν, and m are material constants and Δ is the isochronous damage function (defined below).

Note that the coupling in (6) and (7) is through a scalar ψ_0, which from the first of (5) is seen to represent the average of $\Psi(\bar{n}(\theta),t)$ over the unit circle c. This is consistent with the experimental results of Trampczynski et al. [17], directed to anisotropic deformation and failure of metals, that suggest creep rate may be more dependent on a measure of total damage (or, for example, its average ψ_0 over c), than on its detailed directional distribution. Of course, justification for this choice regarding the composites of interest here ultimately requires additional experimentation.

If only the first term ψ_0 is retained in the expansion (4) then (6) and (7) are identical to Eqs. (2) and (3) in [9]. Reduction to full isotropy reduces (6) and (7) to the "$\Phi-\Delta$" forms introduced by Leckie and Hayhurst [18] as a multiaxial generalization of the Rabotnov/Kachanov equations of continuum damage mechanics (CDM), cf. Rabotnov [19].

The isochronous damage function Δ is taken as

$$\Delta(N,S), \quad \text{where } N(\theta) = \langle \sigma_{ij} n_j(\theta) n_i(\theta) \rangle \text{ and } S(\theta) = |\sigma_{ij} n_j(\theta) d_i| \tag{8}$$

N in (8) represents the tensile traction acting on an interfacial plane with normal $\bar{n}$; S represents the longitudinal shear acting on the plane of normal $\bar{n}$ and in the direction of $\bar{d}$ (or $-\bar{d}$). As in earlier work, in the absence of a complete experimental definition of $\Delta(N,S)$ we adopt, for simplicity, the linear form

$$\Delta(N,S) = \frac{1}{\sigma_0}(N + \alpha S) \tag{9}$$

where α is an additional material constant.

Thermodynamic constraints ensuring dissipativity in a class of evolutionary laws including (6) and (7) are discussed in by Onat and Leckie in [11,13]. Equations (6) and (7) are identifiable as "Ω-forms" in the thermodynamic formalism of [11,13] (Φ relates to a dissipation potential Ω). Conditions for thermodynamic dissipativity are that Φ (or Ω) is non-negative and $\Phi = \text{const.}$ ($\Omega = \text{const.}$) are convex, nested surfaces enclosing the origin of state space. These conditions are seen to be met in Eq. (5) of [9], defining Φ. In addition to an accompanying discussion in [9], a related discussion on thermodynamic dissipativity is given in [14].

The distribution of continuity $\Psi(\bar{n}(\theta),t)$ of a composite element evolves in time. Its rate of change is obtained by differentiating (4) as,

$$\dot{\Psi}(\theta,t) = \dot{\psi}_0(t) + \dot{\psi}_{2C}(t) \cos 2\theta + \dot{\psi}_{2S}(t) \sin 2\theta + \dot{\psi}_{4C}(t) \cos 4\theta$$
$$+ \dot{\psi}_{4S}(t) \sin 4\theta + \dot{\psi}_{6C}(t) \cos 6\theta + \dot{\psi}_{6S}(t) \sin 6\theta + \cdots \tag{10}$$

where, looking ahead to applications, we have extended the expansion to include terms in 6θ.

Differentiating (5) and making use of the damage evolutionary law (7), we write

$$\dot{\psi}_0(t) = -\frac{A}{2\pi} \int_{-\pi}^{\pi} \frac{\Delta^\nu}{\Psi^m(\theta,t)} d\theta$$

$$\dot{\psi}_{2C}(t) = -\frac{A}{\pi} \int_{-\pi}^{\pi} \frac{\Delta^\nu}{\Psi^m(\theta,t)} \cos 2\theta d\theta$$

$$\dot{\psi}_{2S}(t) = -\frac{A}{\pi} \int_{-\pi}^{\pi} \frac{\Delta^\nu}{\Psi^m(\theta,t)} \sin 2\theta d\theta$$

$$\dot{\psi}_{4C}(t) = -\frac{A}{\pi} \int_{-\pi}^{\pi} \frac{\Delta^\nu}{\Psi^m(\theta,t)} \cos 4\theta d\theta \tag{11}$$

$$\dot{\psi}_{4S}(t) = -\frac{A}{\pi} \int_{-\pi}^{\pi} \frac{\Delta^\nu}{\Psi^m(\theta,t)} \sin 4\theta d\theta$$

$$\dot{\psi}_{6C}(t) = -\frac{A}{\pi} \int_{-\pi}^{\pi} \frac{\Delta^\nu}{\Psi^m(\theta,t)} \cos 6\theta d\theta$$

$$\dot{\psi}_{6S}(t) = -\frac{A}{\pi} \int_{-\pi}^{\pi} \frac{\Delta^\nu}{\Psi^m(\theta,t)} \sin 6\theta d\theta$$

Integration of (10) and (11) over the unit circle c and time t yields the current continuity distribution $\Psi(\bar{n}(\theta),t)$. In many structural applications the stress components are referred to a global coordinate system; these can be transformed readily to the local fiber coordinates as illustrated in the following applications.
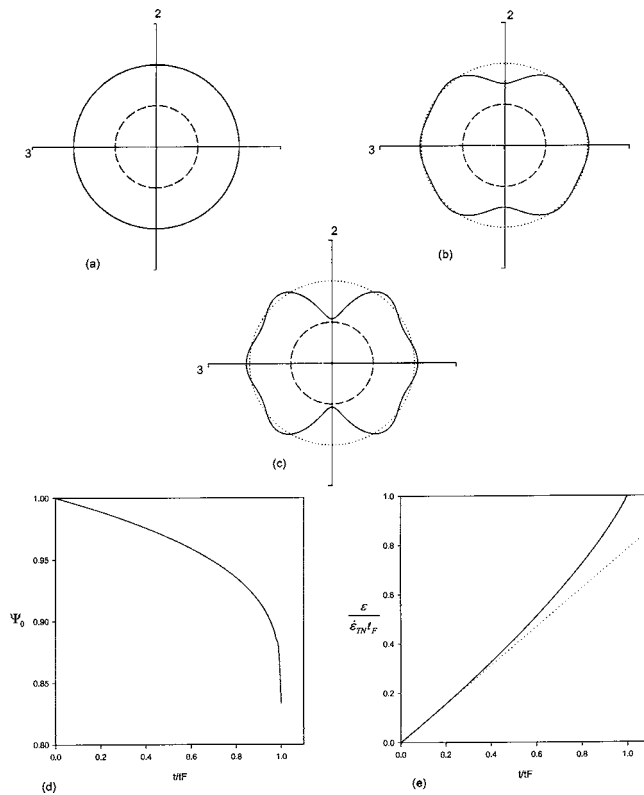
Fig. 4 Damage of composite element under transverse tension. (*a*) Undamaged state (solid curve) $\Psi(\bar{n},0)=1$ at $t=0$. (*b*) Partial damage (solid curve) $t>0$. (*c*) Continuity distribution (solid curve) $\Psi(\bar{n},t)$ at failure $t=t_F$. (*d*) ψ_o vs t/t_F. (*e*) $\varepsilon/\dot{\varepsilon}_{TN}t_F$ vs t/t_F showing acceleration of creep rate (solid curve).

Fig. 5 Isotropic damage evolution. (*a*) Undamaged state (solid curve) $\Psi=1$ at $t=0$. (*b*) Partial damage (solid curve) $t>0$. (*c*) Isotropic failure (solid curve) $\Psi=0$ at $t=t_F$.

4 Applications

We now apply (4)–(11) in some simple problems. Integration is accomplished using simple numerical techniques. The quadratures involved in (11) make use of standard procedures built into Math-CAD 2001; the time integration uses a fourth order Runge–Kutta method also available in MathCAD 2001. The expansions (4) and (10) include terms through 6θ. Obviously, additional terms and more sophisticated numerical techniques can be used to achieve increased accuracy.

4.1 Composite Element Under Transverse Tension. First, we consider damage evolution under constant transverse stress. In the coordinates of Fig. 2(*b*) the stress components are $\sigma_{22}=\sigma_o$ (the reference stress) with $\sigma_{11}=\sigma_{33}=\sigma_{12}=\sigma_{13}=\sigma_{23}=0$. Using (8) we have

$$N(\theta) = \langle \sigma_{22}n_2n_2 \rangle = \sigma_o \cos^2(\theta) \quad \text{and} \quad S(\theta)=0 \quad (12)$$

From (9) the isochronous damage function Δ is

$$\Delta(\theta) = \cos^2(\theta) \quad (13)$$

For the sake of this and subsequent calculations we chose the following values of the material parameters in the flow law (6) and damage evolutionary law (7): $A=1$, $\nu=6$, $m=4$, and $p=6.5$. These values are chosen for convenience and to match those determined experimentally for a model PMC in [8,9].

The results are shown in Fig. 4. Figure 4(*a*) depicts the undamaged state at $t=0$ with the continuity distribution $\Psi=1$; Fig. 4(*b*) shows partial damage at a later time $t>0$ and Fig. 4(*c*) is the distribution at $t=t_F$ when the interfacial tangent planes normal to the loading direction 2 have totally lost load carrying capacity (for reference in subsequent calculations t_F will generally designate the time under constant tensile stress σ_o at which failure occurs on

interfacial planes normal to the loading direction). Evidently from Fig. 4(*c*), other interfacial tangent planes within the element retain some carrying capacity.

The adoption of a general "failure" criterion for a composite element damaged as in Fig. 4(*c*) is not straightforward. Correlation with microstructural examination (cf. Fig. 1(*a*)) is critical and needs to be made using appropriate experimental techniques, however, a nondestructive, macroscopic failure criterion is needed in structural applications. Such a criterion may be specific to a given application, e.g., it may uniquely depend on the particular loading history.

A common experimental failure criterion relies on the observed acceleration of creep rate with damage. As indicated earlier, the coupling of damage and creep rate in (6) is taken through ψ_o, which from the first of (5) is seen to represent the average of $\Psi(\bar{n}(\theta),t)$ over the unit circle c. The time variation of $\psi_o(t)$ in this application is plotted in Fig. 4(*d*); $\psi_o \approx 0.83$ at the failure time t_F. Note that $\psi_o(t)$ decreases rapidly as $t \to t_F$.

The creep rate $\dot{\varepsilon}=\dot{\varepsilon}_{22}$ under a transverse tensile stress $\sigma=\sigma_{22}$ is calculated from the flow law (6) which reduces to a version of the simple Norton/Bailey form

$$\frac{\dot{\varepsilon}}{\dot{\varepsilon}_{TN}} = \left| \frac{\sigma}{\psi_o(t)\sigma_o} \right|^p \frac{\sigma}{|\sigma|} \quad (14)$$

as shown in [8]. $\dot{\varepsilon}_{TN}$ is the creep rate for the undamaged material ($\psi_o=1$) under the reference stress σ_o. Taking $\sigma=\sigma_o$ and $p=6.5$ in (14) and integrating, we calculate the creep strain $\varepsilon(t)$ in $0 \leq t \leq t_F$. This is plotted nondimensionally as $\varepsilon/\dot{\varepsilon}_{TN}t_F$ versus t/t_F in Fig. 4(*e*). We observe measurable acceleration of creep rate as indicated by the ratio $\dot{\varepsilon}(t_F)/\dot{\varepsilon}(0) \approx 3.3$. Thus, there is a theoretical correlation relating an increase in creep rate (by about $\times 3$), a corresponding decrease in ψ_o (to ≈ 0.8), and a loss of load carrying capacity normal to the loading direction (viz., Fig. 4(*c*)). Whether this correlation exists in reality must be shown by experiment (both macroscopic and microscopic). A precise definition of failure is left for future study. Presently, we shall view elemental failure as the calculated loss of load carrying capacity across any interfacial tangent plane.

In contrast to Fig. 4, Fig. 5 shows isotropic damage evolution as modeled in Binienda et al. [9] or, equivalently, the present model reduced to a single term ψ_o in the expansion (4). Again, the stress
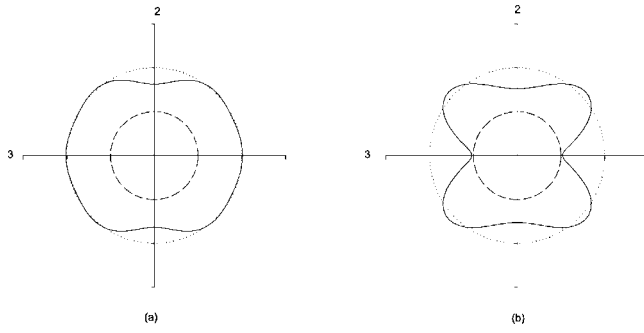
Fig. 6 Damage under stepwise stress history. (*a*) Partially damaged state (solid curve) at *t*=0.9*t_F* just before change in stress direction. (*b*) Fully damaged state (solid curve) at *t* ≈1.87*t_F* after final stage of stress $\sigma_{33}=\sigma_o$.

Fig. 7 Damage of composite element under longitudinal shear. Continuity distribution $\Psi(\bar{n}, t)$ at failure *t*≈60*t_F* (solid curve).

σ_o is applied along the 2 direction. Figure 5(*a*) is the undamaged state at *t*=0; Fig. 5(*b*) shows partial damage at *t*<0 and Fig. 5(*c*) is the fully damaged state at *t*=*t_F* where all interfacial planes have (isotropically) lost load carrying capacity, viz., $\Psi=0$.

The response to stress histories that involve changes in direction are of particular interest. To address this we consider an element first stressed along the 2 direction as in Fig. 4. At some later time the stress direction is changed to the 3 direction. The latter segment of the stress history requires a redefinition of the isochronous damage function Δ. Here, the stress components are $\sigma_{33}=\sigma_o$ with $\sigma_{11}=\sigma_{22}=\sigma_{12}=\sigma_{13}=\sigma_{23}=0$. Again, using (8) we have

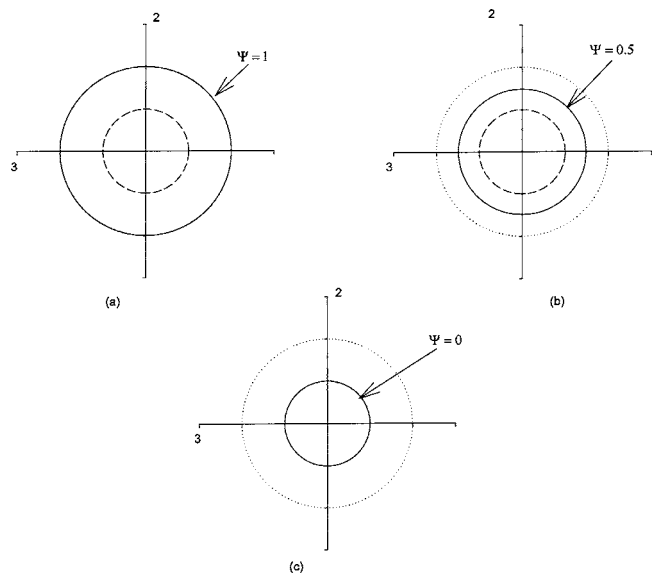$$N(\theta) = \langle \sigma_{33} n_3 n_3 \rangle = \sigma_o \sin^2(\theta) \text{ and } S(\theta) = 0 \qquad (15)$$

From (9) Δ is

$$\Delta(\theta) = \sin^2(\theta) \qquad (16)$$

The damage evolution for the stepwise stress history is shown in Fig. 6. Figure 6(*a*) shows the partially damaged state at *t* =0.9*t_F*, i.e., the selected time when the directional change in stress is made. Figure 6(*b*) shows the continuity distribution at *t* ≈1.87*t_F*, the calculated time (from *t*=0) when the interfacial planes normal to the final loading direction 3 are exhausted of carrying capacity. Recall that a constant stress σ_o applied mono-directionally produces failure at *t*=*t_F*. Thus, the total time to failure in the stepped history is about 87% greater than that under constant σ_o. Of course, this is because changing the stress direction allows material on lesser damaged interfacial planes to support the load following the stress change.

Moreover, referring back to Fig. 5, it is evident that if the damage distribution remained isotropic as in the Binienda et al. [9] model, the failure time in this application would be *t*=*t_F* regardless of any directional change in stress. Obviously, this is a conservative lifetime estimate as argued in Binienda et al. [9].

4.2 Composite Element Under Longitudinal Shear. A fundamental assumption in this research is that the damaging stress components are the tensile traction and the longitudinal shear acting at the fiber/matrix interface. The damage evolution under transverse tensile stress was examined in the previous section. Here, we examine the response under longitudinal shear stress.

Again, referring to the coordinate system of Fig. 2(*b*), we consider the stress state: $\sigma_{12}=\sigma_o$ with $\sigma_{11}=\sigma_{22}=\sigma_{33}=\sigma_{13}=\sigma_{23}=0$. From (8) we have

$$N(\theta) = 0 \text{ and } S(\theta) = |\sigma_{12} n_2 d_1| = \sigma_o |\cos \theta| \qquad (17)$$

and from (9) Δ is

$$\Delta(\theta) = \alpha |\cos \theta| \qquad (18)$$

In these calculations we take $\alpha=0.5$.

Figure 7 shows the final continuity distribution under constant longitudinal shear stress $\sigma_{12}=\sigma_o$. The failure time, i.e., the time
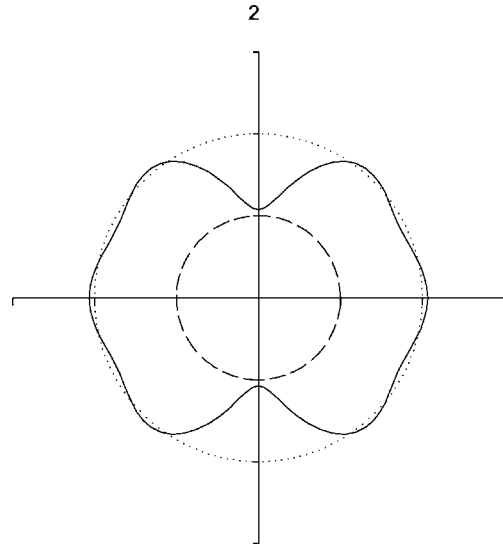
for shear failure on interfacial planes normal to the 2 direction, is calculated to be *t*≈60*t_F* (recall that *t_F* is the failure time under transverse tensile stress σ_o). A comparison of Fig. 7 and Fig. 4(*c*) shows only a slight difference in the shapes of the final continuity distributions associated with longitudinal shear and transverse tension. The shapes reflect the respective damage functions (18) and (13), but also depend on the choice of material parameters used in the calculations.

Whereas the predicted shapes of the continuity distributions at failure show only minor differences, the underlying physical features of damage relating to longitudinal shear and transverse tensile failure (cf. Fig. 1(*a*)) are likely to show major microstructural differences. Again, it is essential that microstructural examination be included as part of the experimental verification of the present damage model.

Once again, it is of interest to consider a history of changing stress. Figure 8 shows the results of first applying a shear stress $\sigma_{12}=\sigma_o$ for *t*=50*t_F*, then changing to a transverse tensile stress $\sigma_{33}=\sigma_o$. Figure 8(*a*) shows the continuity distribution at *t*=50*t_F*, following the application of the longitudinal shear stress. Figure 8(*b*) is the distribution after application of the tensile stress $\sigma_{33}=\sigma_o$ to failure, i.e., to the loss of carrying capacity on interfacial planes normal to the 3 axis. This occurs in a time period *t*≈*t_F* (following the change in stress), i.e., nearly unaffected by the
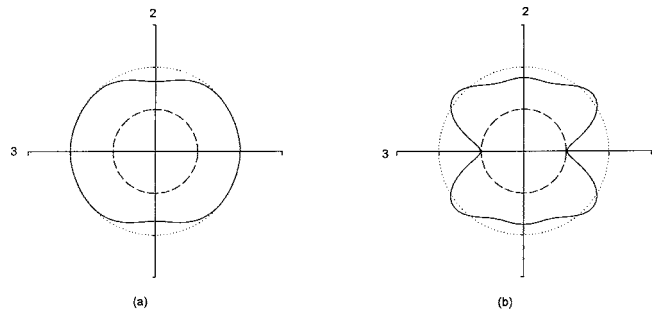
Fig. 8 Damage under stepwise longitudinal shear/transverse tension. (*a*) Partial damage under longitudinal shear for *t* =50*t_F* (solid curve). Distribution just before stress change to $\sigma_{33}=\sigma_o$. (*b*) Continuity distribution $\Psi(\bar{n}, t)$ at failure (solid curve) after $\sigma_{33}=\sigma_o$ is applied for *t*≈*t_F*.
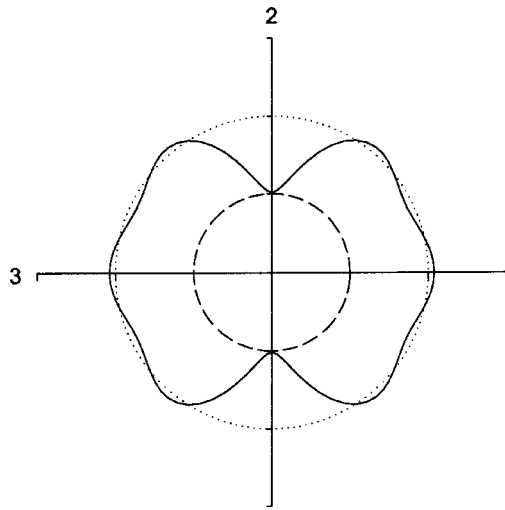
Fig. 9 Damage under stepwise longitudinal shear / transverse tension. Continuity distribution $\Psi(\bar{n}, t)$ at failure (solid curve) after application of longitudinal shear for $t=50t_F$, then $\sigma_{22}=\sigma_o$ applied for $t\approx0.15t_F$.

Fig. 10 Thin tube under tension/torsion reinforced with two fiber families $\bar{a}$ and $\bar{b}$ at $\pm\phi$. (a) Definition of coordinate axes (1,2,3). (b) Definition of fiber coordinate axes (1′,2′,3′). Fiber direction 1′. (c) Unit circle c viewing back along 1′ direction.

initial damage caused by longitudinal shear.

Figure 9 shows the result of a variation on the previous loading. Here, after applying longitudinal shear for $t=50t_F$, a tensile stress $\sigma_{22}=\sigma_o$ is applied. The continuity distribution is identical to Fig. 8(a) at $t=50t_F$. The final distribution after exposure to $\sigma_{22}=\sigma_o$ is shown in Fig. 9. In this case failure occurs in a time period $t \approx 0.15t_F$ following the stress change, only 15% of that corresponding to Fig. 8(b). This illustrates the anisotropy induced, in this case, by the initial longitudinal shear loading. What was originally a plane of isotropy (i.e., the 2–3 plane) becomes anisotropic.

4.3 Tension/Torsion of Thin Tubes With Two Families of Fibers at $\pm\phi$. Here we consider thin-walled tubes reinforced by two distinct families of fibers designated by the unit vectors $\bar{a}$ and $\bar{b}$ in Fig. 10(a); the fibers make angles $\pm\phi$ with the tube axis. We conjecture that each fiber family incurs damage as in the case of a single fiber family, i.e., by the presence of tensile stress and longitudinal shear stress normal and parallel to their respective fiber-matrix interfaces.

The tube is loaded by an axial force F and a torque T generating the state of stress

$$\sigma_{ij} = \begin{bmatrix} \sigma & \tau & 0 \\ \tau & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (19)$$

at a typical wall element. The components (19) refer to the coordinate system (1,2,3) in Fig. 10(a).

We adopt an additional fiber related coordinate system (1′,2′,3′) as in Fig. 10(b) that is formed by rotation of $\pm\phi$ about the 3 axis. The 1′ axis coincides with the fiber direction; it aligns with the $\bar{a}$ fibers under rotation ϕ and with the $\bar{b}$ fibers under rotation $-\phi$. The stress components referred to the (1′,2′,3′) system are

$$\sigma_{ij}' = \begin{bmatrix} \sigma\cos^2\phi + \tau\sin 2\phi & -\dfrac{\sigma}{2}\sin 2\phi + \tau\cos 2\phi & 0 \\ -\dfrac{\sigma}{2}\sin 2\phi + \tau\cos 2\phi & \sigma\sin^2\phi - \tau\sin 2\phi & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (20)$$

Viewing back along the 1′ axis (the fiber direction) we observe the unit circle c in Fig. 10(c). Its defining unit vector $\bar{n}$ is

$$\bar{n}(\theta) = (0, \cos\theta, \sin\theta) \quad (21)$$

in (1′,2′,3′) components.

Returning to (8), we have

$$N(\theta,\phi) = \langle \sigma_{2'2'} n_{2'} n_{2'} \rangle = \langle \sigma\sin^2\phi - \tau\sin 2\phi \rangle \cos^2\theta \quad (22)$$

and

$$S(\theta,\phi) = |\sigma_{1'2'} n_{2'} d_{1'}| = \left| -\dfrac{\sigma}{2}\sin 2\phi + \tau\cos 2\phi \right| |\cos\theta| \quad (23)$$

where, again, all components are referred to the (1′,2′,3′) system. Note that in (23), $d_{1'} \equiv a_{1'} = 1$ for $+\phi$ and $d_{1'} \equiv b_{1'} = 1$ for $-\phi$. The damage function $\Delta(N,S)$ is obtained by introducing (22) and (23) into (9), thus

$$\Delta(\theta,\phi) = \dfrac{1}{\sigma_o}\left[\langle \sigma\sin^2\phi - \tau\sin 2\phi \rangle \cos^2\theta + \alpha \left| -\dfrac{\sigma}{2}\sin 2\phi \right. \right.$$
$$\left. \left. + \tau\cos 2\phi \right| |\cos\theta| \right] \quad (24)$$

(24) holds independently for each fiber family; $\bar{a}$ with $+\phi$ and $\bar{b}$ with $-\phi$.

In applications, the independent continuity distributions
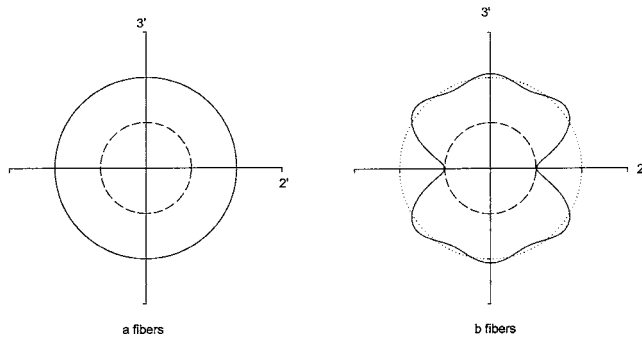
Fig. 11 Continuity distributions for both fiber families at $t=t_F$ (solid curves). Positive twist with $\tau=\sigma_o$.

$\Psi_A(\theta,t)$ and $\Psi_B(\theta,t)$ evolve according to (4)–(11) with (24). Each distribution has a damage evolutionary equation identical to (7). Referring to (4), the lead terms in each of the expansions $\Psi_A(\theta,t)$ and $\Psi_B(\theta,t)$ are $\psi_{oA}(t)$ and $\psi_{oB}(t)$. From (5) these represent the current averages of the distributions over their respective unit circles c. For an arbitrary composite volume V, we consider the volume average

$$\bar{\Psi}_o(t) = \frac{1}{V}\left(\int_{V_A}\psi_{oA}(t)dV + \int_{V_B}\psi_{oB}(t)dV\right) \tag{25}$$

Applying (25) to a homogeneous (continuum) element in which $V_A=V_B=V/2$, we have

$$\bar{\Psi}_o(t) = \frac{1}{2}(\psi_{oA}(t) + \psi_{oB}(t)) \tag{26}$$

The average (26) replaces $\psi_o(t)$ as the coupling variable in a flow law analogous to (6) for the two-fiber-family representation, cf., Robinson [16].

We are now in a position to calculate the damage response of thin-walled tubes with $\pm\phi$ fibers to a general tension/torsion loading using (4)–(11) with (24). However, we shall consider only a special case that relates to experiments proposed in [9]. We consider a tube with fiber directions $\phi=\pm\pi/4$ ($\pm45°$) under forward and reverse torsional (shear) loading, viz., $(\sigma,\tau)=(0,\pm\tau)$.

First, we consider a tube subjected to positive torque T producing a state of shear stress $\tau\equiv\sigma_{12}=\sigma_o$; with $\sigma_{11}=\sigma_{22}=\sigma_{33}=\sigma_{13}=\sigma_{23}=0$. These are components relative to the (1,2,3) axes in Fig. 10(a). The damage distributions at failure ($t\approx t_F$) are calculated using (4)–(11) with (24) and are shown in Fig. 11 for each of the fiber families. The $\bar{a}$ fibers remain undamaged while the $\bar{b}$ fibers show a loss of load carrying capacity on interfacial planes normal to their associated 2′ direction. The failure time t_F is consistent with earlier results in that the $\bar{b}$ fibers experience a transverse tensile traction σ_o; the $\bar{a}$ fibers have zero transverse traction. As identified earlier, the coupling term in the flow law for two families of fibers (cf. Robinson [16]) is $\bar{\Psi}_o(t)$ in (26). The time variation of $\bar{\Psi}_o(t)$ in this application is shown as the dotted curve in Fig. 12; it is seen that $\bar{\Psi}_o\approx0.91$ at $t=t_F$.

Next we consider a stress history where a tube is subjected to a negative torque $-T$ prior to the application of positive torque T as considered above. The state of shear stress under the negative twist is $\tau=-\sigma_o$. This is held for a selected time $t=0.9t_F$ after which the shear stress is changed to $\tau=\sigma_o$. The partial damage distributions at the stress change ($t=0.9t_F$) for both fiber families are shown in Fig. 13; the $\bar{b}$ fibers are yet undamaged. Following the stress change, failure is calculated to occur at $t\approx1.9t_F$ when the load carrying capacity of the $\bar{b}$ fiber family becomes exhausted across planes normal to their 2′ direction. The final continuity
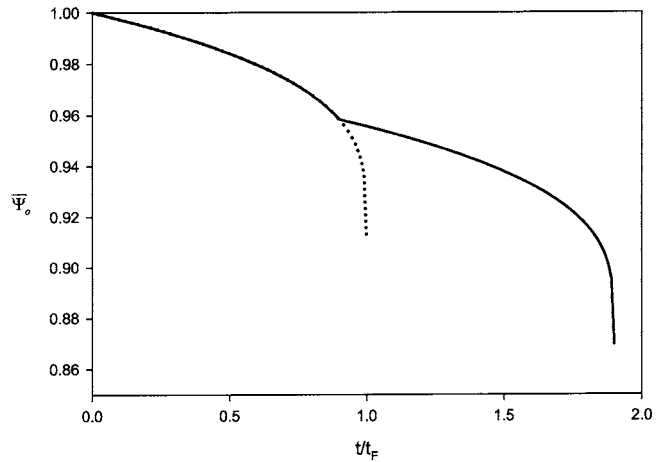
Fig. 12 $\bar{\Psi}_o$ versus t/t_F. Dotted curve relates to continuity distributions of Fig. 11; solid curve relates to continuity distributions of Fig. 14.

distributions at failure ($t\approx1.9t_F$) are given in Fig. 14. Whereas the $\bar{b}$ family has failed, the $\bar{a}$ fibers remain partially damaged as in Fig. 13.

The time variation of the deformation/damage coupling quantity $\bar{\Psi}_o(t)$ is shown in Fig. 12 (solid curve). Its value is $\bar{\Psi}_o\approx0.95$ at the stress change ($t=0.9t_F$) and $\bar{\Psi}_o\approx0.86$ at failure ($t\approx1.9t_F$).

We are interested in the shear strain in the tube over $0\leq t\leq1.9t_F$. In particular, our interest is in the calculated acceleration
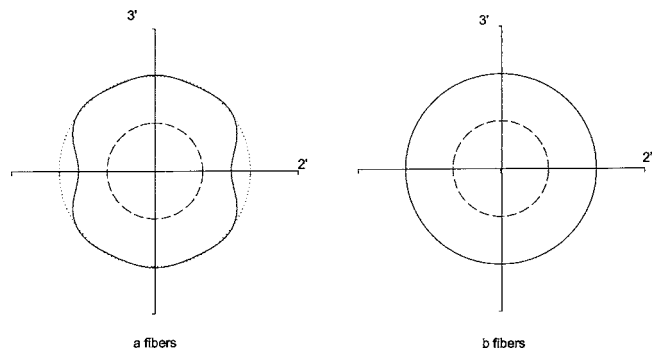
Fig. 13 Damage of tube under reversed twist history. Continuity distributions for both fiber families (solid curves) at $t=0.9t_F$, under negative twist ($\tau=-\sigma_o$) just prior to twist reversal.
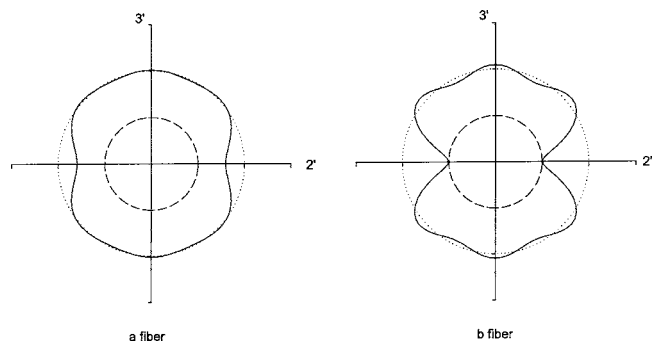
Fig. 14 Damage of tube under reversed twist history. Continuity distributions for both fiber families (solid curves) at failure, after final positive twist ($\tau=\sigma_o$). Failure time is $t\approx1.9t_F$.
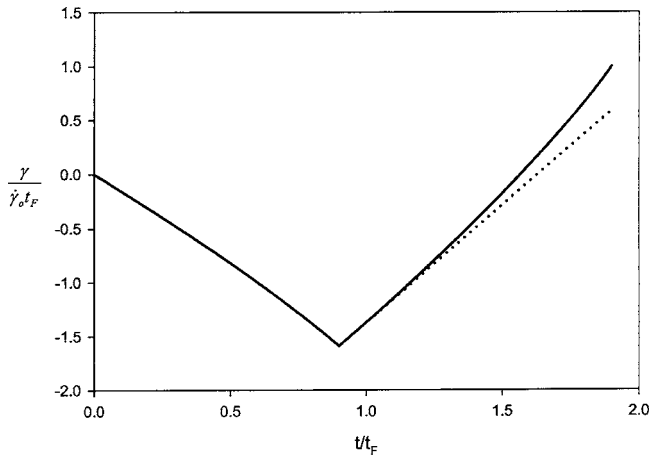
Fig. 15 Damage of tube under reversed twist history. Creep strain history $\gamma/\dot{\gamma}_o t_F$ versus t/t_F (solid curve) showing creep rate acceleration approaching $t \approx 1.9 t_F$.

of creep rate accompanying the approach of failure at $t \approx 1.9 t_F$. An equation analogous to (14) is derived in [16] for calculating the shear strain rate $\dot{\gamma}(t)$ in a composite with two fiber families. Thus,

$$\frac{\dot{\gamma}}{\dot{\gamma}_o} = \left| \frac{\tau}{\bar{\Psi}_o(t)\sigma_o} \right|^p \frac{\tau}{|\tau|} \qquad (27)$$

where $\dot{\gamma}_o$ denotes the shear rate for the undamaged (± 45 deg two-family) composite element with $\bar{\Psi}_o = 1$ and $\tau = \pm \sigma_o$.

Taking $\tau = \pm \sigma_o$ and $p = 6.5$, (27) is integrated to determine $\gamma(t)$ in $0 \leq t \leq 1.9 t_F$. This is plotted nondimensionally in Fig. 15 as $\gamma/\dot{\gamma}_o t_F$ versus t/t_F. Negative shear strain accumulates prior to the stress change at $t = 0.9 t_F$, it then reverses and accelerates as failure approaches. The extent of acceleration is characterized by the ratio $\dot{\gamma}(1.9 t_F)/\dot{\gamma}(0.9 t_F) \approx 2.4$.

Here again, a hypothetical correlation exists relating the acceleration of creep rate (by a factor of 2–3), a decrease of $\bar{\Psi}_o$ (to ≈ 0.86) and a total loss of carrying capacity on interfacial planes normal to the loading direction. As indicated earlier, verification of this correlation in a real composite material needs to be established by experiment. The application considered here involving reversed twisting of a tubular specimen having fibers at ± 45 deg provides a useful vehicle for experimental verification.

Consider experiments of this kind generating a shear strain-time response as in Fig. 15. Loading continues until the creep rate is observed to increase by ≈ 2–3 times. The "failure" time is noted and the specimen is unloaded, dissected and examined microstructurally. The proposed model infers that the $\bar{b}$ fibers should appear critically damaged (cf. Fig. 1(a)) on interfacial planes normal to their associated $2'$ (in plane) direction. The $\bar{a}$ fibers are expected to show lesser damage, cf. Fig. 14. A sequence of tests of this kind, involving various degrees of creep rate acceleration, other stress histories, etc., serve as a comprehensive assessment of the present model.

5 Summary and Conclusions

A phenomenological creep damage model is formulated that allows for an anisotropic distribution of damage in composite materials having long or continuous fiber reinforcement. Faithful representation of anisotropic damage distribution is important for accurately predicting stress redistribution in damaging structures and thereby accurately estimating lifetime. An earlier constitutive model by the writers allowed for anisotropic damage growth rate but, resting on a scalar damage state variable, failed to account for anisotropic distributions of damage. Here, in effect, a vectorial

state variable is introduced that permits a representation of anisotropic damage distribution. As in earlier work, a fundamental assumption is that the principally damaging stress components are tensile traction and longitudinal shear at the fiber/matrix interface.

This research is guided by the work of Onat [11,12] and Onat and Leckie [13] in which expansions of irreducible tensors are used to map anisotropic damage distributions on a unit sphere. Here, accounting for material symmetry of the composite and the assumed critical stress dependence, we map anisotropic damage onto a unit circle normal to the local fiber direction (or directions). The reduced tensorial expansion is shown to be equivalent to an ordinary Fourier series expansion over the unit circle.

Although the present emphasis is phenomenological, with focus on a mathematical structure for representing anisotropic distributions of damage, any creep damage model must rest on fundamental material science and microstructural examination. Macroscopic experiments supporting the present modeling need to be accompanied by microscopic evidence of damage, e.g., as in Fig. 1(a). Combined micro/macroscopic verification experiments are described in the previous section; they are based on tension/torsion testing of thin-walled composite tubes with fiber reinforcement at ± 45 deg to the tube axis. Related experiments were supported by the National Science Foundation under Grant No. 0001634 and published in [8,9]. These include tests for defining the form and functional dependency of the isochronous damage function (8). Additional experiments involving nonproportional stressing as in [17] may be required to justify coupling in (6) and (7) being based solely on the leading scalar term ψ_o in the expansion (1) (or (4)). Experiments of this kind are under definition by the authors. If experimental verification of a scalar coupling in (6) and (7) were not established, the same general representation holds but at the expense of considerable complexity.

The specification of a general "failure" criterion for an anisotropically damaged composite element is not straightforward. In terms of model development and verification, correlation with micro-structural examination is critical, however, a macroscopic failure criterion is needed in structural applications, e.g., should structural failure be based on total loss of load carrying capacity in any orientation, on some average measure over the unit circle c, on detection or calculation of accelerated creep rate, etc.? This topic is left for future study.

References

[1] Grobstein, T. L., 1989, "The interface in tungsten fiber reinforced niobium metal matrix composites," NASA TM-102122, National Aeronautics and Space Administration, Lewis Research Center, Cleveland, OH.
[2] Grobstein, T. L., 1989, "Creep behavior of tungsten fiber reinforced niobium metal matrix composites," NASA TM-102307, National Aeronautics and Space Administration, Lewis Research Center, Cleveland, OH.
[3] Shewmon, P. G., 1963, *Diffusion in Solids*, Material Science and Engineering Series, McGraw–Hill, New York.
[4] Hull, D. and Rimmer, D. E., 1959, "The growth of grain boundary voids under stress," Philos. Mag., **4**, pp. 673–687.
[5] Chuang, T-J. and Rice, J. R., 1973, "The shape of intergranular creep cracks growing by surface diffusion," Acta Metall., **21**(12), p. 1625.
[6] Cocks, A. C. F. and Ashby, M. F., 1982, "On creep fracture by void growth," Prog. Mater. Sci., **27**, pp. 189–244.
[7] Robinson, D. N., Binienda, W. K., and Miti-Kavuma, M., 1992, "Creep and creep rupture of metallic composites," J. Eng. Mech., **118**(8), pp. 1646–1660.
[8] Robinson, D. N., Binienda, W. K., and Ruggles, M. B., 2003, "Creep of polymer matrix composites. I: Norton/Bailey creep law for transverse isotropy," J. Eng. Mech., **129**, pp. 310–317.
[9] Binienda, W. K., Robinson, D. N., and Ruggles, M. B., 2003, "Creep of polymer matrix composites. II: Monkman-Grant failure relationship for transverse isotropy," J. Eng. Mech., **129**, pp. 318–323.

[10] Kachanov, L. M., 1986, *Introduction to Continuum Damage Mechanics*, Martinus-Nijhoff, Dordrecht.

[11] Onat, E. T., 1986, "Representation of mechanical behavior in the presence of internal damage," Eng. Fract. Mech., **25**, pp. 605–614.

[12] Onat, E. T., 1991, "Group theory and representation of microstructure and mechanical behavior of materials," *Modeling the Deformation of Crystalline Solids*, edited by T. C. Lowe, A. D. Rollett, P. S. Follensbee, and G. S. Daehn, The Minerals, Metals and Materials Society.

[13] Onat, E. T., and Leckie, F. A., 1988, "Representation of mechanical behavior in the presence of changing internal structure," J. Appl. Mech., **55**, pp. 1–10.

[14] Robinson, D. N., Kim, K-J. and White, J. L., 2002, "Constitutive model of a transversely isotropic Bingham fluid," J. Appl. Mech., **69**, pp. 641–648.

[15] Robinson, D. N., and Binienda, W. K., 2001, "Optimal fiber orientation in creeping composite structures," J. Appl. Mech., **68**(2), pp. 213–217.

[16] Robinson, D. N., 2004, "Creep/damage model of PMC with two families of fibers" (in preparation).

[17] Trampczynski, W. A., Hayhurst, D. R., and Leckie, F. A., 1981, "Creep rupture of copper and aluminium under nonproportional loading," J. Mech. Phys. Solids, **29**, pp. 353–374.

[18] Leckie, F. A., and Hayhurst, D. R., 1974, "Creep rupture of structures," Proc. R. Soc. London, Ser. A, **240**, p. 323.

[19] Rabotnov, Y. N., 1968, *Creep Problems in Structural Mechanics*, North–Holland, Amsterdam.

F. S. Ibrahem[1]
e-mail: fibrahim@acc.aun.edu.eg

I. A. Hassanien

Department of Mathematics,
Faculty of Science,
Assiut University, Assiut, Egypt

A. A. Bakr
Department of Mathematics,
Faculty of Science,
Al-Azhar University,
Assiut, Egypt

Nonclassical Thermal Effects in Stokes' Second Problem for Micropolar Fluids

The MCF model is used to study the nonclassical heat conduction effects in Stokes' second problem of a micropolar fluid. The effects of the thermal relaxation time and the structure wave on angular velocity, velocity field, and temperature are investigated. The skin friction, the displacement thickness, and the rate of the heat transfer at the plate are determined. [DOI: 10.1115/1.1875412]

1 Introduction

Fourier's law of heat conduction, which has been used to drive the heat conduction equation in classical unsteady heat transfer problems, results in a parabolic equation for temperature filed and an infinite speed of heat propagation, thus violating the principle of causality. Maxwell [1] derived the generalization of Fourier's heat law for the dynamical theory of gases. Maxwell's heat flux equation contains a term proportional to the time derivative of the heat flux vector multiplied by a constant relaxation time τ. Since τ had a very small magnitude in Maxwell's work, he took it to be zero. In justification he remarked, "The first term of this equation may be neglected, as the rate of conduction will rapidly establish itself. Ackerman et al. [2] established the second sound in solid helium, which gave a finite speed of propagation of thermal waves. Puri and Kythe [3] have studied the influence of generalized law of heat conduction, using the Maxwell-Cattaneo-Fox (MCF) model, on Stokes' first and second problems for Rivlin-Ericksen fluids with nonclassical heat conduction. Kythe and Puri [4] studied the unsteady magnetohydrodynamics (MHD) free-convection flows on a porous plate with time-dependent heating in a rotating medium. Puri and Kythe [5] have studied an unsteady flow problem which deals with nonclassical heat conduction effects and the structure of waves in the Stokes' second problem. In the MCF model as developed by McTaggart and Lindsay [7], the nonclassical constitutive equation for the heat-flux vector q is given by the Maxwell-Cattaneo equation in the form

$$\tau(\dot{q}_i - \omega_{ij}q_j) = -q_i - \chi\theta^*_{,i}, \tag{1}$$

where ω_{ij} is the vorticity, χ the thermal conductivity, $\dot{q}_i = \partial q/\partial t$, θ the temperature, and τ the thermal relaxation time. Equation (1) reduces to that of the Cattaneo model at $\omega_{ij}=0$ and it becomes Fourier's law for $\tau=0$ (see Joseph and Preziosi [8,9]). While there are other good models to choose from, the Cattaneo law, as stated in Joseph and Preziosi [8,9], has many desirable properties. For example, the steady heat flow is induced by temperature gradients and gives rise to finite speeds of propagation. The dimensionless thermal relaxation time, defined as $\lambda=Cp$, where C and p are the Cattaneo and Prandtl numbers, respectively, exhibits a definite

influence on the structure of waves. It significantly modifies their behavior. The number λ also appears in generalized thermoelasticity [10] where it is defined as m and is shown to be of the order 10^{-2}. Again as noted in Puri and Jordan [11], the Cattaneo number C may not be very small in astrophysical applications. For example, C is of order 10^{-2} in a low temperature hydrogen gas. However, the thermal relaxation time λ does not appreciably change the magnitude of the temperature and velocity fields. Puri and Jordan [6] studied some recent developments in the unsteady flow of dipolar fluids with hyperbolic heat conduction. Eringen [12] introduced the theory of micropolar fluids in which he explained some fundamental problems for future theoretical and experimental studies. These problems include: well-posedness basic initial boundary value problems for micropolar fluid flows, right choice of boundary data for microrotation field, comparison of micropolar fluid flows with Navier-Stokes flows, new questions in the theory of turbulence, and range of applicability of the model. A lot has been done since in the frame of this broad project, however, many important problems remain open. Peddieson and McNitt [13] and Willson [14] have introduced the boundary layer concept in such fluid. The study of flow and heat transfer for an electrically conducting micropolar fluid past a porous plate under the influence of a magnetic field has attracted the interest of many investigators in view of its applications in many engineering problems such as magnetohydrodynamic, generator, plasma studies, geothermal energy extractions, and the boundary layer control in the field of aerodynamics Soundalgekar and Takhar [15]. Micropolar fluids are fluids with microstructure belonging to a class of fluids with non-symmetrical stress tensor. Physically, they represent fluids consisting of randomly oriented particles suspended in a viscous medium Lukaszewicz [16]. Eringin [12] formulated the theory of micropolar fluids and derived constitutive laws for fluids with microstructure. This theory included the effects of local rotary inertia and couple stresses and expected to provide a mathematical model for non-Newtonian behavior observed in certain manmade liquids and theologically complex fluids such as liquid crystals, polymeric suspensions, and naturally occurring liquids such as animal blood. Kim and Fedorov [17,18] studied unsteady MHD micropolar flow and heat transfer over a vertical porous moving plate with variable suction.

One can consider this article as both a generalization of research of Puri and Kythe [5] and Puri and Jordan [19] to Stokes' second problem for micropolar fluids and as an extension and refinement of the work of Puri and Jordan [11]. Our motivation in doing this work stems from the ever growing number of flow temperature and or high heat flux applications of non-Newtonian fluids in areas such as medical research, space exploration, and low-temperature physics. Last, we must note that in the general

---

case of thermoviscous fluids, particularly monoatomic gases, a complicated mutual interaction between temperature and velocity fields, particularly monoatomic (see Müller and Ruggeri [20], pp. 1–61). Thus, because of the linear nature of the problem presented here, this work should be considered as only a first approximation to a more complex problem.

2 Mathematical Analysis

Consider the one-dimensional unsteady flow of a laminar, incompressible, micropolar fluid past a vertical flat plate in the xy plane and occupy the space $z > 0$, with the z axis in the vertical direction. The plate initially at rest and at constant temperature θ_∞ which is the free stream temperature is moved with a velocity $U_0 e^{i\omega t}$ in its own plane along the z axis, and its temperature is subjected to a periodic heating of the form $(\theta_w - \theta_\infty)e^{i\omega t}$, where $\theta_w \neq \theta_\infty$ is some constant. The basic equations of continuity, momentum, angular momentum and energy governing such a flow, subject to the Boussinesq approximation, are

$$v^*_{i,i} = 0, \tag{2}$$

$$\rho \dot{v}^*_i = -P^*_{,i} + (\mu + \mu_r)\nabla^2 v^*_i - \rho[1 - \alpha(\theta^* - \theta^*_\infty)]g\,\delta_{i1} + 2\mu_r N^*_{,i} + t_{ki,k}, \tag{3}$$

$$\rho j^* \dot{N}^* = \gamma \nabla^2 N^*, \tag{4}$$

$$\rho \varepsilon = -q_{i,i} + t_{ik}d_{ik}. \tag{5}$$

Where the vector $v = (u, 0, 0)$ represents the velocity, ρ the density, j^* the microinertia density, N^* the component of angular velocity vector, γ the spin-gradient viscosity, μ the dynamic viscosity, P the pressure, ε the specific internal energy, α the coefficient of thermal expansion, g the acceleration due to gravity, t_{ik}, the non-Newtonian stress tensor, and d_{ik} the strain tensor.

The effect of microstructure is negligible in the neighborhood of a rigid boundary since the suspended particles cannot get closer than their radius to boundary (Ahmadi [21]). Thus in our study we consider the only rotation is due to fluid shear as pointed in Eq. (4).

Taking into account the geometry of the problem which results in the disappearance of the dissipative terms and noting that $t_{ik} = 0$, Eqs. (2)–(4) reduce to the following equation of motion:

$$u^*_t = (v + v_r)u^*_{zz} + g\alpha(\theta^* - \theta^*_\infty) + 2v_r N^*_z, \tag{6}$$

$$N^*_t = \frac{1}{\eta}N^*_{zz}. \tag{7}$$

Equation (1), after substitution into Eq. (5), gives

$$\rho c_p \theta^* = -q_{i,i}. \tag{8}$$

Where $\varepsilon = c_p \theta$ for the MCF model. If we drop the nonlinear terms $\tau\omega_{ij}q_j$ in Eq. (1) because τ and ω are small quantities, we get

$$\tau \dot{q}_{i,i} = -q_{i,i} - \chi \theta_{,ii}. \tag{9}$$

Which in one-dimensional form, after dropping the convective terms (because these terms become automatically zero), leads to

$$\tau \theta^*_{tt} + \theta^*_t = \frac{\chi}{\rho c_p}\theta^*_{zz} \tag{10}$$

Note that the term $\tau\theta^*_{tt}$ in Eq. (10) is necessary to ensure finite speed of propagation. We shall use the nondimensional quantities.

$$z^* = \frac{v}{U_0}z, \quad u^* = U_0 u, \quad t^* = \frac{v}{U_0^2}t, \quad \theta = \frac{\theta^* - \theta^*_0}{\theta^*_w - \theta^*_0}, \quad N^* = \frac{U_0^2}{v}N. \tag{11}$$

Then the governing Eqs. (6), (7), and (10) for the flow, angular velocity and heat conduction, after suppressing the primes, become

$$u_t = (1 + \beta)u_{zz} + G\theta + 2\beta N_z, \tag{12}$$

$$N_t = \frac{1}{\eta}N_{zz}, \tag{13}$$

$$\lambda p\theta_{tt} + p\theta_t = \theta_{zz}. \tag{14}$$

Where G is the Grashof number and β viscosity ratio

$$G = \frac{vg\alpha(\theta^*_w - \theta^*_0)}{U_0^3}, \quad p = \frac{v\rho c_p}{\chi}, \quad C = \frac{\tau\chi U_0^2}{v^2\rho c_p},$$

$$\lambda = \frac{\tau U_0^2}{v} = C_p, \quad \eta = \frac{2}{2 + \beta}. \tag{15}$$

The boundary conditions are

$$u(0,t) = e^{i\omega t}, \quad \theta(0,t) = e^{i\omega t}, \quad \frac{\partial N(0,t)}{\partial z} = \frac{-\partial^2 u(0,t)}{\partial z^2},$$

$$u(\infty,t) = 0, \quad \theta(\infty,t) = 0, \quad N(\infty,t) = 0. \tag{16}$$

To solve the nonlinear systems, (12)–(14) with the boundary condition (16), We assume that

$$u(z,t) = U(z)e^{i\omega t},$$

$$\theta(z,t) = \Theta(z)e^{i\omega t},$$

$$N(z,t) = N(z)e^{i\omega t}. \tag{17}$$

If we substitute by Eq. (17) in Eqs. (12)–(14) and the boundary conditions (16) we get.

$$(1 + \beta)U'' - i\omega U = -G\Theta - 2\beta N', \tag{18}$$

$$N'' - i\omega\eta N = 0, \tag{19}$$

$$\Theta'' + (\lambda p\omega^2 - i\omega p)\Theta = 0. \tag{20}$$

The boundary conditions are

$$U(0) = 1, \quad \Theta(0) = 1, \quad N'(0) = -U''(0),$$

$$U(\infty) = 0, \quad \Theta(\infty) = 0, \quad N(\infty) = 0. \tag{21}$$

We get the solution under modified boundary conditions as

$$U(z) = e^{-mz} - (G_1 + iG_2)(e^{-mz} - e^{-(r_1+ir_2)z}) - (S_1 + iS_2)(e^{-mz} - e^{-m_1 z}), \tag{22}$$

$$N(z) = \frac{\omega + i\,G}{2(1 - \beta)\sqrt{\omega/(1 + \beta)}}e^{-m_1 z}, \tag{23}$$

$$\Theta(z) = e^{-(r_1+ir_2)z}. \tag{24}$$

where

$$m = (1 + i)\sqrt{\frac{\omega}{2(1 + \beta)}}, \quad m_1 = (1 + i)\sqrt{\frac{\omega}{2 + \beta}},$$

$$r_{1,2} = \sqrt{\omega P\left(\frac{\sqrt{(1 + \lambda^2\omega^2)} \mp \lambda\omega}{2}\right)},$$

$$G_1 + iG_2 = \frac{G\left(\lambda P\omega^2 - i\omega\left(\frac{1}{1+\beta} - P\right)\right)}{(1 + \beta)\left[\lambda^2 p^2\omega^4 + \omega^2\left(\frac{1}{1+\beta} - p\right)^2\right]},$$

$$S_1 + iS_2 = \frac{2(2 + \beta)(\omega + iG)}{(1 - \beta)\omega}. \tag{25}$$

From Eq. (23), we find that

$$\text{Re } u(z,t) = e^{-z\sqrt{\omega/2(1+\beta)}}\cos[\omega t - z\sqrt{\omega/2(1+\beta)}] - e^{-z\sqrt{\omega/2(1+\beta)}}\{(G_1$$

$$+ S_1)\cos[\omega t - z\sqrt{\omega/2(1+\beta)}] - (G_2 + S_2)\sin[\omega t$$

$$- z\sqrt{\omega/2(1+\beta)}]\} + e^{-r_1 z}[G_1\cos(\omega t - r_2 z) - G_2\sin(\omega t$$

$$- r_2 z)] + e^{-z\sqrt{\omega/(2+\beta)}}\{S_1\cos[\omega t - z\sqrt{\omega/(2+\beta)}]$$

$$- S_2\sin[\omega t - z\sqrt{\omega/(2+\beta)}]\}. \tag{26}$$

In the classical case, for micropolar fluid ($\beta \neq 0$) the solutions for both temperature and velocity fields can be obtained by taking $\lambda = 0$ in Eqs. (24) and (26), respectively. Then

$$\theta(z,t) = e^{-(1+i)z\sqrt{\omega p/2}+i\omega t}, \tag{27}$$

$$u(z,t) = [e^{-mz} - iG_2(e^{-mz} - e^{-m_2 z}) - (S1 + iS2)(e^{-mz} - e^{-m_1 z})]e^{iwt}, \tag{28}$$

for $G_1 = 0$ and $G_2 = G/\beta\omega$. Also in the classical case for Newtonian fluid ($\beta = 0$, $p = 1$, and $\lambda = 0$) becomes a singular case and the solution for the velocity in the this case for $p = 1$ has to be obtained directly in the form

$$u(z,t) = \left(1 + \frac{(1-i)Gz}{2\sqrt{2\omega}}\right)e^{i\omega t - (1+i)z\sqrt{\omega/2}}. \tag{29}$$

We notice that this is the same result as obtained by Puri and Kythe [5]. The velocity field based on Fourier's heat law is obtained by setting $\beta = 1$ in the above expressions. However, taking $\beta = 1$, is a singular case. The directly obtained singular solution for the velocity field is given by

$$u = (z,t) = \{e^{-m_2 z} - (G_3 + iG_4)(e^{-m_2 z} - e^{-(r_1+ir_2)z}) - (S_3 + iS_4)z(e^{-m_2 z}$$

$$- e^{-m_3 z})\}e^{-iwt}. \tag{30}$$

Where

$$m_2 = \frac{1}{2}(1+i)\sqrt{\omega}, \quad m_3 = (1+i)\sqrt{\omega/3},$$

$$G_3 + iG_4 = \frac{G\lambda P\omega^2 + iG\omega(p - 0.5)}{2(\lambda^2 p^2\omega^4 + \omega^2(p - 0.5)^2)}$$

$$S_3 + iS_4 = \frac{(\omega - G/2) + i(\omega + G/2)}{4\sqrt{\omega/3}}$$

For the singular case of $p = 1$, $\lambda = 0$, $\beta = 1$ Re u it is given by

$$\text{Re } u = \exp\left(-\frac{z}{2}\sqrt{\omega}\right)\left\{(1 - G_3 - zS_3)\cos\left(\omega t - \frac{z}{2}\sqrt{\omega}\right) + (G_4$$

$$+ zS_4)\sin\left(\omega t - \frac{z}{2}\sqrt{\omega}\right)\right\} + \exp(-zr_1)(G_3\cos(\omega t - zr_2)$$

$$- G_4\sin(\omega t - zr_2) + \exp(-z\sqrt{\omega/3})\{zS_3\cos(\omega t - z\sqrt{\omega/3})$$

$$- zS_4\sin(\omega t - z\sqrt{\omega/3})\}. \tag{31}$$

The standard definition of the displacement thickness δ^* is

$$\delta^* = \int\left(1 - \frac{u}{U_f}\right)dz, \tag{32}$$

where U_f is the free stream velocity. In our case, the plate is moving while the free stream is stationary. Therefore, the formula (32) is modified, and in the nondimensional form is given by

$$|\delta^*| = \int_0^\infty u\, dz, \tag{33}$$

this formula can also be obtained by imparting to the entire system a negative velocity equal to the velocity of the plate. Using the formula (33) the absolute value of the displacement thickness is given by

$$|\delta^*| = \left|\frac{1}{m} - (G_1 + iG_2)\left(\frac{1}{m} - \frac{1}{(r_1 + ir_2)}\right) - (S_1 + iS_2)\left(\frac{1}{m}\right.\right.$$

$$\left.\left. - \frac{1}{m_1}\right)\right| \text{ for all } p, \lambda \neq 0, \quad \text{and } \beta \neq 0. \tag{34}$$

This thickness, in classical case, is $|\delta^*| = |1/m - (i\omega/(1-\beta))(1/m - 1/m_1)|$ in the absence of heat transfer. But in the singular case of $\lambda = 0$, $p = 1$, and $\beta = 0$ it is given by

$$|\delta^*| = \frac{1}{\sqrt{\omega}}\sqrt{1 + G^2/4\omega^2}, \tag{35}$$

and for the singular case of $\lambda = 0$, $p = 1$, $\beta = 1$ is given by

$$|\delta^*| = \left|\frac{1}{m_2} - (G_3 + G_4)\left(\frac{1}{m_2} - \frac{1}{(r_1 + ir_2)}\right) - (S_3 + iS_4)\left(\frac{1}{m_2^2} - \frac{1}{m_3^2}\right)\right|. \tag{36}$$

We can now calculate the skin-friction and heat transfer coefficients in terms of the Nusselt number at the wall of the plate:

$$\tau_w = \left.\frac{\partial u}{\partial z}\right|_{z=0} = \{-m + (G_1 + iG_2)[m - (r_1 + ir_2)] + (s_1 + is_2)(m$$

$$- m1)\}\exp(i\omega t), \tag{37}$$

$$\text{Nu}\chi^{-1} = \left.-\frac{\partial\theta}{\partial z}\right|_{z=0} = (r_1 + ir_2)\exp(i\omega t). \tag{38}$$

3 Discussion

The oscillatory nature of the flow generates waves in the both temperatures, velocity fields and the angular velocity. Although these waves decay rapidly, it is of some interest to understand their structure.

3.1 Wave Structure and Thermal Waves. In the temperature field the behavior of the thermal waves can be obtained from the solution (24). From this solution we can conclude that the thermal waves exhibit one type of dispersive wave trains with a wave front at $z = \omega t/r_2$, ω is the angular frequency, r_1 the attenuation coefficient, and r_2 is the wave number, then phase velocity V_ϕ is defined as

$$V_\phi = \omega/r_2 = \sqrt{\frac{\omega}{2p(\sqrt{1 + \lambda^2\omega^2} + \lambda\omega)}}, \tag{39}$$

from Eq. (25), one can solve for ω in terms of r_2 and get

$$\omega = \frac{\sqrt{2}r_2^2}{\sqrt{p(p + 4\lambda r_2^2)}}. \tag{40}$$

The group velocity is the velocity with which energy propagates and is defined by

$$V_g = \frac{d\omega}{dr_2} = \frac{2\sqrt{2\omega(1 + \lambda^2\omega^2)}}{\sqrt{p}(\lambda\omega + \sqrt{1 + \lambda^2\omega^2})^{3/2}},$$

$$V_g < V_\phi \Rightarrow \text{Normal dispersion}, \quad V_g$$

$$> V_\phi \Rightarrow \text{Anomalous dispersion},$$

$$V_g = V_\phi \Rightarrow \text{There is no dispersion} \tag{41}$$

Normal dispersion means that waves appear to emanate from the front and disappear in the rear. Anomalous dispersion means that waves appear to emanate from the rear and disappear in the front.

No dispersion means that there is no relative motion of the group and the carrier occurs and the group travels without distorsion of the wave shape. The above Eq. (41) for $\lambda \rightarrow 0$ reduces $V_g = 2\sqrt{2\omega/p}$. In the case when the product $\lambda\omega$ is large, we find that $r_1 \approx \sqrt{p/\omega/2\lambda}$, $r_2 \approx \omega\sqrt{\lambda p}$. In this case, then, the velocity phase is $V_\phi = 1/\sqrt{\lambda p}$ at λ is very small then, the speed phase of thermal waves is huge and this wave decay very fast.

$$\tan\Psi_1 = \frac{G_2 + S_2}{G_1 + S_1 - 1} = \frac{\frac{(\beta-1)G\omega^2[p(1+\beta)-1]}{1+\beta} + 2(1+\beta)(2+\beta)G\{p^2\lambda^2\omega^4 - \omega^2[1/(1+\beta)-p]^2\}}{(\beta-1)G\omega^3 p\lambda + 2(1+\beta)(2+\beta)\omega(p^2\lambda^2\omega^4 - \omega^2(1/(1+\beta)-p)^2)}. \quad (42)$$

For $p = 1$, this reduces to the classical case [$\Psi_1 = 0$ at $\beta^2 + (G + 4\omega/2\omega)\beta - G/2\omega = 0$]. The second wave train at the phase lag $zr_2 - \Psi_2$, where

$$\tan\Psi_2 = \frac{G_2}{G_1} = \frac{p(1+\beta)-1}{\lambda p\omega(1+\beta)}. \quad (43)$$

In this case we note $\Psi_2 = 0$ when $P(1+\beta) = 1$, $\beta \neq -1$. The third wave train at $z = t\sqrt{\omega(2+\beta)}$ and the phase lag of $(z\sqrt{\omega/2+\beta} - \Psi_3)$, where

$$\tan\Psi_3 = \frac{S_2}{S_1} = \frac{G}{\omega}. \quad (44)$$

We note that from Eq (22), when $\beta = 0$, the third wave train is

omitted because in this case $m_1 = m_2 = \sqrt{\omega/2}$.

Equation (29) which describes the classical case in Newtonian fluid ($\beta = 0$, $P = 1$, $\lambda = 0$), the above three waves trains coalesce into one with a wave front at $z = t\sqrt{2\omega}$, and a phase lag of $[(z/\sqrt{2} - \phi)/\omega]$ for a fluid layer at a distance z from the plate, where

$$\tan\phi = \frac{Gz}{Gz + 2\sqrt{2\omega}}, \quad (45)$$

which is the same result as obtained in [5].

For the special case, $\beta = 1$ from Eq. (30) exhibits three types of dispersive progressive wave trains, one wave train progresses with the wave front at $z = 2t\sqrt{\omega}$ and phase lag $(z/2\sqrt{\omega} - \phi_1)$, where

$$\tan\phi_1 = \frac{G_4 + S_4}{G_4 + S_3 - 1} = \frac{4G(p-0.5)\omega^{3/2}/\sqrt{3} + 2(\omega + G/2)[(p-0.5)^2\omega^2 + \lambda^2 p^2\omega^4]}{4Gp\ \omega^{5/2}\lambda/\sqrt{3} + 2(\omega + G/2)[(p-0.5)^2\omega^2 + \lambda^2 p^2\omega^4] - 8\omega^{5/2}[0.25 - p - p^2(1+\lambda^2\omega^2)]} \quad (46)$$

The second wave-train at $z = \omega t/r_2$ and the phase lag of $[(zr_2 - \phi_2)]$, where

$$\tan\phi_2 = \frac{G_4}{G_3} = \frac{\left(p - \frac{1}{2}\right)}{\lambda P\omega}. \quad (47)$$

The third wave at $z = t\sqrt{3\omega}$ and the phase lag $(z\sqrt{\omega/3} - \phi_3)$, where
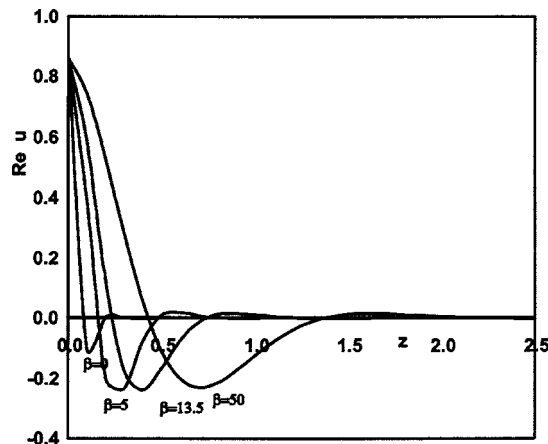
Fig. 1 Behavior of Re u versus z for $p=1$, $\omega=1000$, $t=0.1$, $\lambda =0.005$, $G=5.0$ and $\beta=0$, 5, 13.5 and 50

$$\tan\phi_3 = \frac{2\omega + G}{2\omega - G}. \quad (48)$$

3.3 Wave Structure of Angular Velocity Field. The angular velocity fields consist of the one damped harmonic component. The component has a wave front at $z = t\sqrt{(2+\beta)\omega}$, and phase lag of $(z\sqrt{2\omega/(2+\beta)} - \varphi)$ for a fluid layer at distance z from the plate, where

$$\tan\varphi = \frac{\omega(1-\beta) + 2G}{G(1-\beta) - 2\beta\omega}. \quad (49)$$

The group velocity of this wave is given by

$$V_g = \frac{1}{2\sqrt{(2+\beta)\omega}}, \quad (50)$$

ω is very small, then the speed phase of the thermal waves is huge, and this wave decays very fast.

3.4 Velocity Field. From Fig. 1 we can observe that as β increases the momentum boundary layer thickens/increases. Also, as β increases the Re u increases.

The same trend can be observed in Fig. 2 for $|u|$. On the other hand from this figure notice that as β increases, the peak velocity decreases.

The effect of λ on both Re u and $|u|$ is found to be very small,
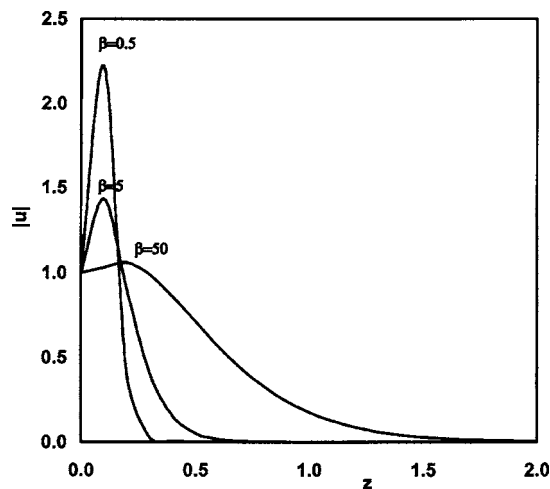
Fig. 2 Behavior of $|u|$ versus z for $p=1$, $\omega=1000$, $t=0.1$, $\lambda=0.005$, $G=5.0$ and $\beta=0.5$, 5, and 50

although it changes the character of the velocity field. Since λ takes very small values, we have presented the graphs of Re u and $|u|$ in Figs. 3 and 4 for $\lambda=0.005$, $\omega=10$, $G=\pm5$, and $t=0.1$. From these figures we can notice that as G increases $|u|$ decreases and there is no change in the fluid boundary layer. Also, the peak velocity decreases as G increases. For large ω, the influence of both λ and G is negligible; as seen in Figs. 5 and 6.

The velocity field in Stokes' second problem consists of three components: one, corresponding to e^{-mz}, defines the Stokes-Rayleigh layer which is of the order $O(1/\sqrt{\omega})$, the second, corresponding to $e^{-(r_1+ir_2)}$ represents the thermal layer which is of the order $O(1/r_1)$, and the third wave train, which is similar to e^{-m_1z}, represents the angular momentum layer which is of the order $O(1/\sqrt{\omega/\beta})$.

In order to observe the effect of λ on Re u we presented some data in Tables 1 and 2, below. Thus, in Table 1, which is for $G=-5$, it can be seen that an increase in λ tends to decrease Re u up to $z=0.24$ for an increase in λ from 0.0 to 0.005, and increase it thereafter. Again, Re u increases with respect to λ up to $z=0.2$ for $\lambda \in [0.005, 0.01]$. This behavior is reversed for $G=5$, as is obvious from Table 2. The underlined data in Tables 1 and 2 are the

Fig. 4 Behavior of $|u|$ versus z for $p=1$, $\omega=10.0$, $t=0.1$, $\lambda=0.005$, $\beta=0.2$, and $G=\pm5$
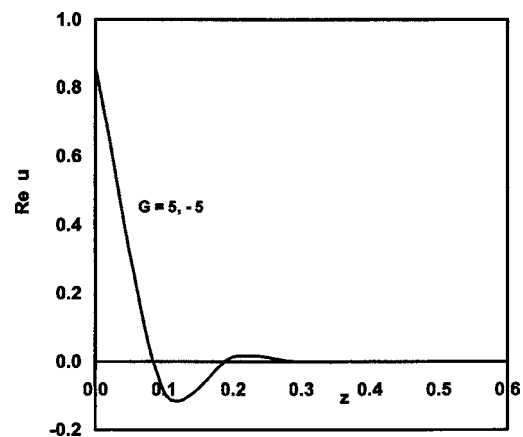
Fig. 5 Behavior of Re u versus z for $p=1.0$, $\omega=1000.0$, $t=0.1$, $\lambda=0.005$, $\beta=0.2$, and $G=\pm5$

critical values of Re u with respect to λ; at these values there is reversal in response to an increase in λ. By comparing our results for micropolar fluid ($\beta=0.2$) in Tables 1 and 2 by Tables 1 and 2 of Puri and Kythe [5], which are for Newtonian fluid ($\beta=0$), we

Fig. 3 Behavior of Re u versus z for $p=1.0$, $\omega=10.0$, $t=0.1$, $\lambda=0.005$, $\beta=0.2$, and $G=\pm5$
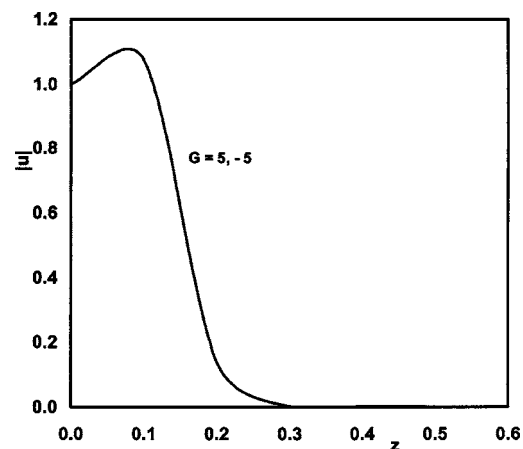
Fig. 6 Behavior of $|u|$ versus z for $p=1$, $\omega=1000$, $t=0.10$, $\lambda=0.005$, $\beta=0.2$ and $G=\pm5$

Table 1 Re *u* at *t*=0.1, *β*=0.2, *ω*=10, *G*=−5

Λ	Z=0.0	Z=0.2	Z=0.21	Z=0.22	Z=0.23	Z=0.24	Z=0.26	Z=0.27	Z=0.3
0.0	0.540 302 31	0.399 696 78	0.390 290 47	0.380 839 93	0.371 363 39	0.361 878 05	0.342 945 01	0.333 527 00	0.305 626 80
0.001	0.540 302 31	0.399 672 58	0.390 269 61	0.380 822 95	0.371 350 82	0.361 870 43	0.342 948 88	0.333 537 40	0.305 659 88
0.005	0.540 302 31	0.399 612 28	0.390 224 52	0.380 795 21	0.371 342 56	0.361 883 78	0.343 011 81	0.333 628 28	0.305 846 85
0.01	0.540 302 31	0.399 619 37	0.390 254 75	0.380 851 23	0.371 427 03	0.361 999 36	0.343 197 53	0.333 852 98	0.306 203 66

Table 2 Re *u* at *t*=0.1, *β*=0.2, *ω*=10, *G*=5

λ	Z=0.0	Z=0.2	Z=0.21	Z=0.22	Z=0.23	Z=0.24	Z=0.26	Z=0.27	Z=0.3
0.0	0.540 302 31	0.666 394 57	0.663 149 05	0.659 236 60	0.654 687 94	0.649 533 27	0.637 524 03	0.630 727 22	0.607 502 80
0.001	0.540 302 31	0.666 418 77	0.663 169 90	0.659 253 58	0.654 700 51	0.649 540 89	0.637 520 17	0.630 716 82	0.607 469 73
0.005	0.540 302 31	0.666 479 07	0.663 214 99	0.659 281 33	0.654 708 78	0.649 527 54	0.637 457 23	0.630 625 94	0.607 282 76
0.01	0.540 302 31	0.666 471 99	0.663 184 77	0.659 225 31	0.654 624 31	0.649 411 96	0.637 271 51	0.630 401 24	0.606 925 94

Fig. 7 Behavior of Re *N* versus *z* for *p*=1, *ω*=1000, *t*=0.1, λ =0.005, *G*=5 and *β*=2.5, *5, 13.5,* and *50*
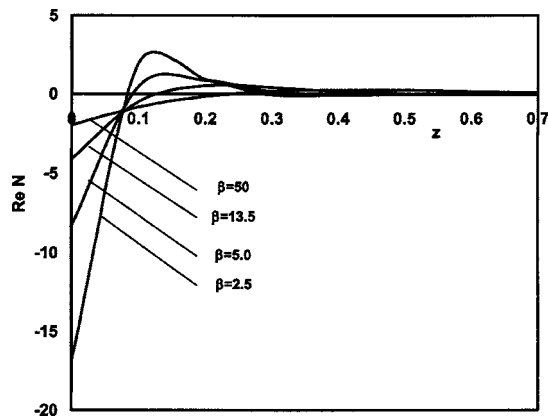
Fig. 9 Behavior of Re *N* versus *z* for *p*=1, *ω*=10, *t*=0.1, λ =0.005, *β*=0.2, and *G*=±5

find that with increasing *β* Re *u* decreases for *G*=−5 and increases for *G*=5. On the other hand, at *β*=0 our results are agree with those of Puri and Kythe [5].

3.5 Angular Velocity. In Figs. 7–12 we have prepared some graphs of the angular velocity (Re *N* and |*N*|) profiles for various values of the parameters *G,p,*λ*,ω,* and *β* which listed in figure captions.

The variation of the Re *N* with *β* is displayed in Fig. 7 for *ω*
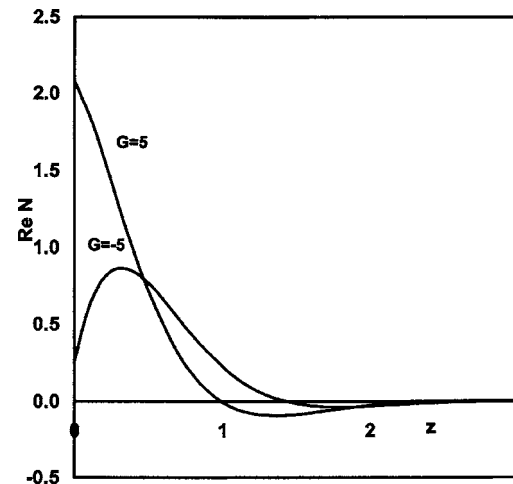
=10. From this figure it can be seen that Re *N* increases with *β*.

Figure 8 displays the buoyancy-assisted flow (*G*=5) results for the distribution of angular velocity within the boundary layer for various values of *β*. It is obvious that as the viscosity ratio *β*
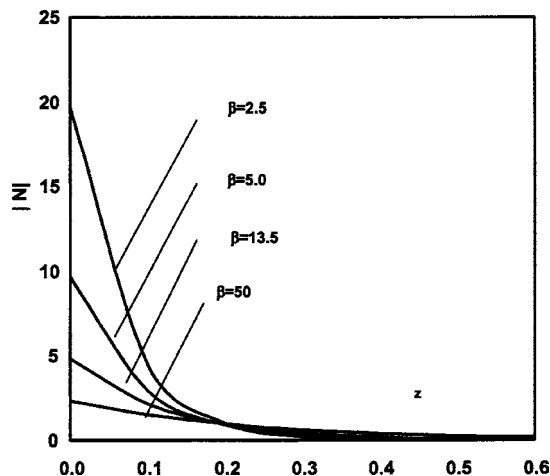
Fig. 8 Behavior of |*N*| versus *z* for *p*=1, *ω*=1000, *t*=0.1, λ =0.005, *G*=5 and *β*=2.5, *5, 13.5,* and *50*
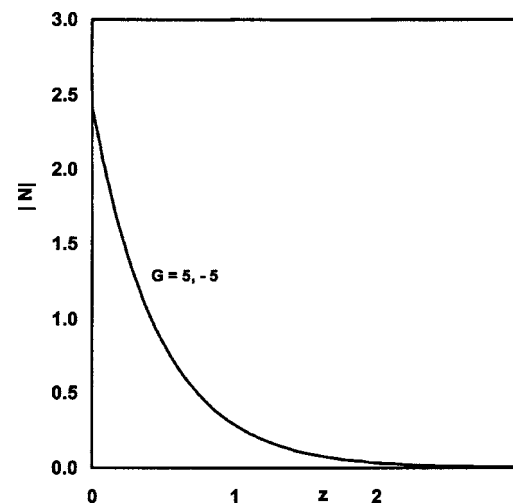
Fig. 10 Behavior of |*N*| versus *z* for *p*=1, *ω*=10, *t*=0.1, λ =0.005, *β*=0.2, and *G*=±5
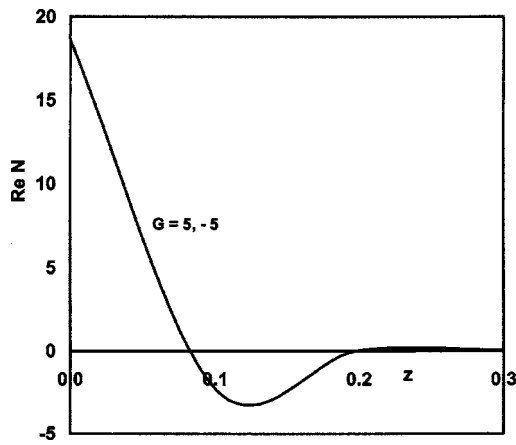
Fig. 11 Behavior of Re *N* versus *z* for *p*=1, *ω*=1000, *t*=0.1, λ =0.005, *β*=0.2, and *G*=±5
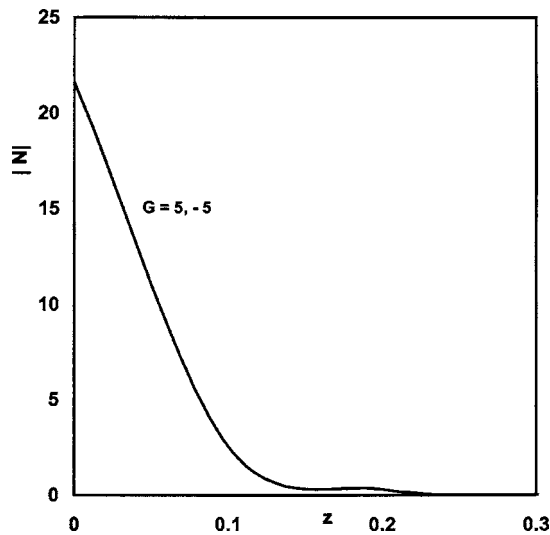
Fig. 12 Behavior of |*N*| versus *z* for *p*=1, *ω*=1000, *t*=0.1, λ =0.005, *β*=0.2, and *G*=±5

increases, the amplitude of the angular velocity profiles decreases.

It is obvious that the effect of increasing values of *G* results in a decreasing angular velocity distribution across the boundary layer as seen in Fig. 9. On the other hand for small (ω=10) the influence of *G* on |*N*| is negligible as shown in Fig. 10. For large ω=1000, the influence of *G* is negligible in both Re *N* and |*N*| (see Figs. 11 and 12).

3.6 Temperature Fields. Typical variations of the temperature profiles along the spanwise coordinate are the same that was presented by Puri and Kythe [5], therefor it is omitted here.

4 Conclusions

Based on the analysis given above, we now state the following:

(1) There are three cases of unique solution $u(x,t)$. These correspond to the nonsingular case of $\beta \neq 0$, *1*, the singular case of β=0, and the case β=1.

(2) The solution for the velocity field exhibits three types of wave motion, one corresponds to the angular velocity, the second wave corresponds to the thermal wave, and the third to the classical Stokes wave.

(3) The momentum boundary layer thickens and the real part of the velocity Re *u* increases with increasing β.

(4) The peak velocity |*u*| decreases with increasing β.

(5) The angular velocity profiles |*N*| decreases but the real part of the angular velocity Re *N* increases with increasing β.

Acknowledgment

References

[1] Maxwell, J. C., 1867, "On the Dynamical Theory of Gases," Philos. Trans. R. Soc. London, **157**, pp. 49–88.

[2] Ackerman, C. C., Bertman, B., Fairbank, H. A., and Guyer, R. A., 1966, "Second Solid Helium," Phys. Rev. Lett., **16**, pp. 789–791.

[3] Puri, P., and Kythe, P. K., 1998, "Stokes' First and Second Problems for Rivlin—Ericksen Fluids With Nonclassical Heat Conduction," ASME J. Heat Transfer, **120**, pp. 44–50.

[4] Kythe, P. K., and Puri, P., 1988, "Unsteady MHD Free-Convection Flows on a Porous Plate With Time-Dependent Heating in a Rotating Medium," Astrophys. Space Sci., **143**, pp. 51–62.

[5] Puri, P., and Kythe, P. K., 1995, "Nonclassical Thermal Effects in Stokes' Second Problem," Acta Mech., **112**, pp. 1–9.

[6] Puri, P., and Jordan, P. M., 2002, "Some Recent Developments in the Unsteady Flow of Dipolar Fluids," *Developments in Theoretical and Applied Mechanics*, XXI, pp. 499–508.

[7] McTaggart, C. L., and Lindsay, K. A., 1985, "Nonclassical Effects in the Benard Problem," SIAM J. Appl. Math., **45**, pp. 70–92.

[8] Joseph, D. D., and Preziosi, L., 1989, "Heat Waves," Rev. Mod. Phys., **61**, pp. 41–73.

[9] Joseph, D. D., and Preziosi, L., 1990, "Addendum to the Paper Heat Waves," Rev. Mod. Phys. **62**, pp 375–391.

[10] Puri, P., 1973, "Plane waves in Generalized Thermoelasticity," Int. J. Eng. Sci., **11**, pp. 735–744.

[11] Puri, P., and Jordan, P. M., 1999, "Stokes First Problem for a Dipolar With Nonclassical Heat Conduction," J. Eng. Math., **36**, pp. 219–240.

[12] Eringen, A. C. J., 1966, "Theory of Micropolar Fluids," J. Math. Mech., **16**, pp. 1–18.

[13] Peddieson, J., and McNitt, R. P., 1970, "Boundary Layer Theory for Micropolar Fluid," Recent Adv. Engng. Sci., **5**, pp 405–426.

[14] Willson, A., 1970, "Boundary Layers in Micropolar Fluid," Proc. Cambridge Philos. Soc., **67**, pp. 469–481.

[15] Soundalgekar, V. M., and, Takhar, H. S., 1977, "MHD Forced and Free Convective Flow Past a Semi-Infinite Plate," AIAA J., **15**, pp 457–485.

[16] Lukaszewicz, G., 1999, *Micropolar Fluids-Theory and Applications*, Birkhauser, Boston.

[17] Kim, Y. J., and Fedorov, A. G., 2003, "Transient Mixed Radiative Convection Flow of a Micropolar Fluid Past a Moving, Semi-Infinite Vertical Porous Plate," Int. J. Heat Mass Transfer, **46**, pp. 1751–1758.

[18] Kim, Y. J., 2001, "Unsteady MHD Micropolar Flow and Heat Transfer Over a Vertical Porous Moving Plate With Variable Suction," *Proceedings 2nd International Conference on Computational Heat and Mass Transfer*, COPPA/ UFRJ- Federal University of Rio de Janerio, Brazil, October, 22–26.

[19] Puri, P., and Jordan, P. M., 1999, "Wave Structure in Stokes' Second Problem for a Dipolar Fluid With Nonclasical Heat Conduction," Acta Mech., **133**, pp. 145–160.

[20] Müller, I., and Ruggeri, T., 1993, "Extended Thermodynamics," *Springer Tracts in Natural Philosophy*, Vol. 37, C. Truesdell, ed., Springer-Verlag, New York, p. 230.

[21] Ahmadi, G., 1976, "Self-Similar Solution of Incompressible Micropolar Boundary Layer Flow Over a Semi-Infinite Plate," Int. J. Eng. Sci., **14**, pp. 639–646.

Buckling of Long Sandwich Cylindrical Shells Under External Pressure

G. A. Kardomateas
Professor of Aerospace Engineering
Fellow ASME
e-mail: george.kardomateas@aerospace.gatech.edu

G. J. Simitses[1]
Professor Emeritus of Aerospace Engineering
Fellow ASME

Georgia Institute of Technology,
Atlanta, GA 30332-0150

An elasticity solution to the problem of buckling of sandwich long cylindrical shells subjected to external pressure is presented. In this context, the structure is considered a three-dimensional body. All constituent phases of the sandwich structure, i.e., the facings and the core, are assumed to be orthotropic. The loading is a uniform hydrostatic pressure, which means that the loading remains normal to the deflected surface during the buckling process. Results are produced for laminated facings, namely, boron/epoxy, graphite/epoxy and kevlar/epoxy laminates with 0 deg orientation with respect to the hoop direction, and for alloy-foam core. Shell theory results are generated with and without accounting for the transverse shear effect. Two transverse shear correction approaches are compared, one based only on the core, and the other based on an effective shear modulus that includes the face sheets. The results show that the shell theory predictions without transverse shear can produce highly non-conservative results on the critical pressure, but the shell theory formulas with transverse shear correction produce reasonable results with the shear correction based on the core only being in general conservative (i.e., critical load below the elasticity value). The results are presented for four mean radius over shell thickness ratios, namely 15, 30, 60, and 120 in order to assess the effect of shell thickness (and hence that of transverse shear). For the same thickness, the differences between elasticity and shell theory predictions become larger as the mean radius over thickness ratio is decreased. A comparison is also provided for the same shell with homogeneous composite construction. It is shown that the sandwich construction shows much larger differences between elasticity and shell theory predictions than the homogeneous composite construction. The solution presented herein provides a means of a benchmark for accurately assessing the limitations of shell theories in predicting stability loss in sandwich shells. [DOI: 10.1115/1.1934513]

Introduction

The need for lightweight, yet stiff and durable structures has made the sandwich composite configuration a leading edge technology with promise for innovative high performance structural designs. A typical sandwich structure is composed of two thin metallic or composite laminated faces and a thick soft core made of foam or low strength honeycomb. This lightweight sandwich construction is of great interest in the design and manufacture of aircraft, spacecraft, and marine vehicles. In addition to the high specific stiffness and strength, sandwich construction offers enhanced corrosion resistance, noise suppression, and reduction in life-cycle costs.

There are several issues and questions related to the use of sandwich construction that require attention and answers. In applications involving compressive loading, loss of stability and the accurate prediction of buckling loads is of major concern. This is particularly important in sandwich construction because of the existence of the low-modulus core, which would be expected to make transverse shear effects even more significant than in homogeneous composites. In addition, composite sandwich structures are normally envisioned in applications involving relatively thick construction, therefore thickness effects need to be properly accounted for.

Shell theory solutions for buckling and even initial postbuckling behavior have been produced by many authors (e.g., from the 60's, Hutchinson [1]; Budiansky and Amazigo [2], many of these works with elegant variational formulations). Indeed, the existence of different shell theories underscores the need for benchmark elasticity solutions, in order to compare the accuracy of predictions from the classical and the improved shell theories. Several elasticity solutions for monolithic homogeneous composite shell buckling have become available. In particular, Kardomateas [3] formulated and solved the problem for the case of uniform external pressure and orthotropic homogeneous material; in this study, just as in the present one, a long shell was studied ("ring" assumption). This simplifies the problem considerably, in that the pre-buckling stress and displacement field is axisymmetric, and the buckling modes are two dimensional, i.e., no axial component of the displacement field, and no axial dependence of the radial and hoop displacement components. The ring assumption was relaxed in a further study [4], in which a nonzero axial displacement and a full dependence of the buckling modes on the three coordinates was assumed. Other three-dimensional elasticity buckling studies are the buckling of a transversely isotropic homogeneous thick cylindrical shell under axial compression [5] and a generally cylindrically orthotropic homogeneous shell under axial compression [6]. In addition, three-dimensional elasticity results, again for homogeneous hollow cylinders subjected to combined axial compression and uniform external pressure, were provided by Soldatos and Ye [7] based on a successive approximation method.

The geometry of a circular cylindrical shell is particularly at-

---

---

tractive for constructing elasticity solutions due to the axisymmetry which simplifies the analysis. Furthermore, a prerequisite to obtaining elasticity solutions for shell buckling such as the one by Kardomateas [3], is the existence of three-dimensional elasticity solutions to the pre-buckling problem. Elasticity solutions for monolithic homogeneous orthotropic cylindrical shells have been provided by Lekhnitskii [8]. Recently, elasticity solutions for sandwich shells were obtained by properly extending the solutions for monolithic structures [9]. The latter is the pre-buckling solution needed to formulate the bifurcation problem in the elasticity context.

As far as shell theory, there are but few studies reported in the literature that deal with sandwich shell analyses [10–12]. The comparison to shell theory predictions will be based on the formulas presented in Smith and Simitses [13] and Simitses and Aswani [14] and specialized to an infinite length cylinder, whose behavior is similar to that of a sandwich ring.

Formulation

By considering the equations of equilibrium in terms of the second Piola-Kirchhoff stress tensor, subtracting these at the perturbed and initial conditions, and making order of magnitude assumptions on the products of stresses and strains/rotations, based on the fact that a characteristic feature of stability problems is the shift from positions with small rotations to positions with rotations substantially exceeding the strains, Kardomateas [3] obtained the following buckling equations:

$$\frac{\partial}{\partial r}(\sigma_{rr} - \tau_{r\theta}^0\omega_z + \tau_{rz}^0\omega_\theta) + \frac{1}{r}\frac{\partial}{\partial\theta}(\tau_{r\theta} - \sigma_{\theta\theta}^0\omega_z + \tau_{\theta z}^0\omega_\theta)$$

$$+ \frac{\partial}{\partial z}(\tau_{rz} - \tau_{\theta z}^0\omega_z + \sigma_{zz}^0\omega_\theta)$$

$$+ \frac{1}{r}(\sigma_{rr} - \sigma_{\theta\theta} + \tau_{rz}^0\omega_\theta + \tau_{\theta z}^0\omega_r - 2\tau_{r\theta}^0\omega_z) = 0, \quad (1a)$$

$$\frac{\partial}{\partial r}(\tau_{r\theta} + \sigma_{rr}^0\omega_z - \tau_{rz}^0\omega_r) + \frac{1}{r}\frac{\partial}{\partial\theta}(\sigma_{\theta\theta} + \tau_{r\theta}^0\omega_z - \tau_{\theta z}^0\omega_r)$$

$$+ \frac{\partial}{\partial z}(\tau_{\theta z} + \tau_{rz}^0\omega_z - \sigma_{zz}^0\omega_r)$$

$$+ \frac{1}{r}(2\tau_{r\theta} + \sigma_{rr}^0\omega_z - \sigma_{\theta\theta}^0\omega_z + \tau_{\theta z}^0\omega_\theta - \tau_{rz}^0\omega_r) = 0, \quad (1b)$$

$$\frac{\partial}{\partial r}(\tau_{rz} - \sigma_{rr}^0\omega_\theta + \tau_{r\theta}^0\omega_r) + \frac{1}{r}\frac{\partial}{\partial\theta}(\tau_{\theta z} - \tau_{r\theta}^0\omega_\theta + \sigma_{\theta\theta}^0\omega_r)$$

$$+ \frac{\partial}{\partial z}(\sigma_{zz} - \tau_{rz}^0\omega_\theta + \tau_{\theta z}^0\omega_r)$$

$$+ \frac{1}{r}(\tau_{rz} - \sigma_{rr}^0\omega_\theta + \tau_{r\theta}^0\omega_r) = 0. \quad (1c)$$

In the previous equations, σ_{ij}^0 and ω_j^0 are the values of stresses and rotations, respectively, at the initial equilibrium position (pre-buckling state), and σ_{ij} and ω_j are the corresponding values at the perturbed position (buckled state).

The boundary conditions associated with Eq. (1) were obtained from the traction (stress resultant) relationships in terms of the second Piola-Kirchhoff stress tensor, and by further considering the fact that because of the hydrostatic pressure loading, the magnitude of the surface load remains invariant under deformation, but its direction changes (since hydrostatic pressure is always directed along the normal to the surface on which it acts). By writing these equations for the initial and the perturbed equilibrium position and then subtracting them and using the previous argu-
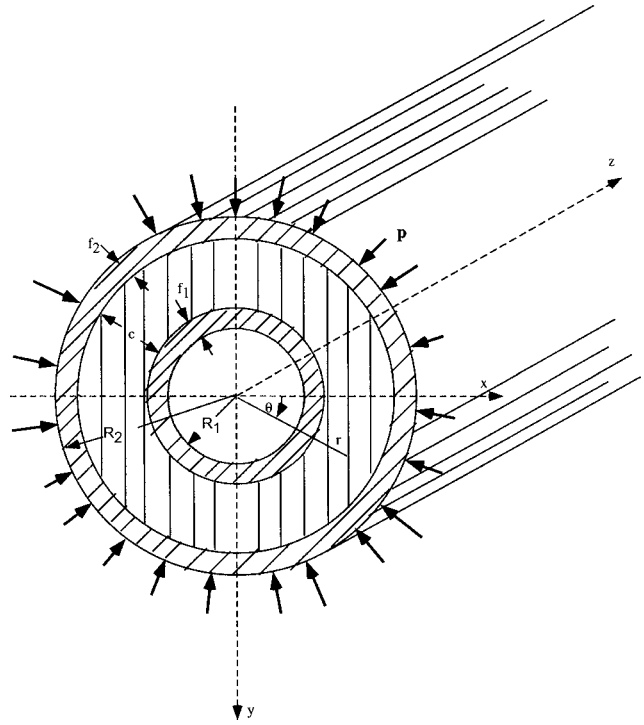
Fig. 1   Definition of the geometry and the loading

ments on the relative magnitudes of the rotations, Kardomateas [3] obtained the following boundary conditions on a surface which has outward unit normal $(\hat{\ell}, \hat{m}, \hat{n})$:

$$(\sigma_{rr} - \tau_{r\theta}^0\omega_z + \tau_{rz}^0\omega_\theta)\hat{\ell} + (\tau_{r\theta} - \sigma_{\theta\theta}^0\omega_z + \tau_{\theta z}^0\omega_\theta)\hat{m}$$

$$+ (\tau_{rz} - \tau_{\theta z}^0\omega_z + \sigma_{zz}^0\omega_\theta)\hat{n} = p(\omega_z\hat{m} - \omega_\theta\hat{n}), \quad (2a)$$

$$(\tau_{r\theta} + \sigma_{rr}^0\omega_z - \tau_{rz}^0\omega_r)\hat{\ell} + (\sigma_{\theta\theta} + \tau_{r\theta}^0\omega_z - \tau_{\theta z}^0\omega_r)\hat{m}$$

$$+ (\tau_{\theta z} + \tau_{rz}^0\omega_z - \sigma_{zz}^0\omega_r)\hat{n} = -p(\omega_z\hat{\ell} - \omega_r\hat{n}), \quad (2b)$$

$$(\tau_{rz} + \tau_{r\theta}^0\omega_r - \sigma_{rr}^0\omega_\theta)\hat{\ell} + (\tau_{\theta z} + \sigma_{\theta\theta}^0\omega_r - \tau_{r\theta}^0\omega_\theta)\hat{m}$$

$$+ (\sigma_{zz} + \tau_{\theta z}^0\omega_r - \tau_{rz}^0\omega_\theta)\hat{n} = p(\omega_\theta\hat{\ell} - \omega_r\hat{m}). \quad (2c)$$

For the lateral bounding surfaces, $\hat{m} = \hat{n} = 0$ and $\hat{\ell} = 1$. These conditions will also be used when we impose traction continuity at the core/face sheet interfaces.

Pre-buckling State. The problem under consideration is that of a sandwich hollow cylinder deformed by uniformly distributed external pressure, p (Fig. 1) and of infinite length (generalized plane deformation assumption). Then, not only the stresses, but also the displacements do not depend on the axial coordinate. Alternatively, this is the assumption we would make if the cylinder were securely fixed at the ends. An elasticity solution to this problem was provided by Kardomateas [9]. The solution is an extension of the classical one by Lekhnitskii [8] for a homogeneous, orthotropic shell and was provided in closed form. All three phases, i.e., the two face sheets and the core were assumed to be orthotropic. Moreover, there were no restrictions as far as the individual thicknesses of the face sheets and the sandwich construction could be asymmetric.

In this configuration, the axially symmetric distribution of external forces produces stresses identical at all cross sections and dependent only on the radial coordinate r. We take the axis of the body as the z axis of the cylindrical coordinate system and we denote by R_1 and R_2 the inner and outer radii. Let us also denote

each phase by i where $i=f_2$ for the outer face sheet, $i=c$ for the core, and $i=f_1$ for the inner face sheet. Then, for each phase, the orthotropic strain-stress relations are

$$
\begin{bmatrix} \epsilon_{rr}^{(i)} \\ \epsilon_{\theta\theta}^{(i)} \\ \epsilon_{zz}^{(i)} \\ \gamma_{\theta z}^{(i)} \\ \gamma_{rz}^{(i)} \\ \gamma_{r\theta}^{(i)} \end{bmatrix} = \begin{bmatrix} a_{11}^i & a_{12}^i & a_{13}^i & 0 & 0 & 0 \\ a_{12}^i & a_{22}^i & a_{23}^i & 0 & 0 & 0 \\ a_{13}^i & a_{23}^i & a_{33}^i & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{44}^i & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{55}^i & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{66}^i \end{bmatrix} \begin{bmatrix} \sigma_{rr}^{(i)} \\ \sigma_{\theta\theta}^{(i)} \\ \sigma_{zz}^{(i)} \\ \tau_{\theta z}^{(i)} \\ \tau_{rz}^{(i)} \\ \tau_{r\theta}^{(i)} \end{bmatrix}, \quad (i=f_1,c,f_2)
$$

(3)

where a_{ij}^i are the compliance constants (we have used the notation $1 \equiv r$, $2 \equiv \theta$, $3 \equiv z$).

Let us introduce the following notation for constants which enter into the stress formulas and depend on the elastic properties:

$$
\beta_{11}^i = a_{11}^i - \frac{a_{13}^{i\,2}}{a_{33}^i}; \quad \beta_{22}^i = a_{22}^i - \frac{a_{23}^{i\,2}}{a_{33}^i} \quad (i=f_1,c,f_2), \quad (4a)
$$

$$
\beta_{12}^i = a_{12}^i - \frac{a_{13}^i a_{23}^i}{a_{33}^i} \quad (i=f_1,c,f_2); \quad k_i = \sqrt{\frac{\beta_{11}^i}{\beta_{22}^i}}; \quad (i=f_1,c,f_2).
$$

(4b)

Then, the pre-buckling stresses in each of the phases, i.e., for $i=f_1, c, f_2$, are

$$
\sigma_{rr}^{0(i)}(r) = p(C_1^{(i)} r^{k_i-1} + C_2^{(i)} r^{-k_i-1}), \quad (5a)
$$

$$
\sigma_{\theta\theta}^{0(i)}(r) = p(C_1^{(i)} k_i r^{k_i-1} - C_2^{(i)} k_i r^{-k_i-1}), \quad (5b)
$$

$$
\tau_{\theta z}^{0(i)}(r) = \tau_{rz}^{(i)}(r) = \tau_{r\theta}^{0(i)}(r) = 0, \quad (5c)
$$

$$
\sigma_{zz}^{0(i)}(r) = p \left\{ -C_1^{(i)} \frac{(a_{13}^i + a_{23}^i k_i)}{a_{33}^i} r^{k_i-1} - C_2^{(i)} \frac{(a_{13}^i - a_{23}^i k_i)}{a_{33}^i} r^{-k_i-1} \right\}.
$$

(5d)

Furthermore, the pre-buckling radial displacement is found to be

$$
u^{0(i)}(r) = p \left[C_1^{(i)} \frac{(\beta_{11}^i + k_i \beta_{12}^i)}{k_i} r^{k_i} - C_2^{(i)} \frac{(\beta_{11}^i - k_i \beta_{12}^i)}{k_i} r^{-k_i} \right], \quad (5e)
$$

the other displacements being zero, i.e., $v^{0(i)}(r) = w^{0(i)}(r) = 0$.

The constants $C_1^{(i)}$, $C_2^{(i)}$ are found from the conditions on the cylindrical lateral surfaces (traction free) and the conditions at the interfaces between the phases of the sandwich structure. Specifically, the traction conditions at the face-sheet/core interfaces give two equations [9]

$$
C_1^{(f_1)}(R_1+f_1)^{k_{f1}-1} + C_2^{(f_1)}(R_1+f_1)^{-k_{f1}-1}
$$
$$
= C_1^{(c)}(R_1+f_1)^{k_c-1} + C_2^{(c)}(R_1+f_1)^{-k_c-1} \quad (6a)
$$

and

$$
C_1^{(c)}(R_2-f_2)^{k_c-1} + C_2^{(c)}(R_2-f_2)^{-k_c-1}
$$
$$
= C_1^{(f_2)}(R_2-f_2)^{k_{f2}-1} + C_2^{(f_2)}(R_2-f_2)^{-k_{f2}-1}. \quad (6b)
$$

The displacement continuity at the face-sheet/core interfaces gives another two equations

$$
C_1^{(f_1)} \frac{(\beta_{11}^{f_1} + k_{f1}\beta_{12}^{f_1})}{k_{f1}} (R_1+f_1)^{k_{f1}} - C_2^{(f_1)} \frac{(\beta_{11}^{f_1} - k_{f1}\beta_{12}^{f_1})}{k_{f1}} (R_1+f_1)^{-k_{f1}}
$$
$$
= C_1^{(c)} \frac{(\beta_{11}^c + k_c\beta_{12}^c)}{k_c} (R_1+f_1)^{k_c} - C_2^{(c)} \frac{(\beta_{11}^c - k_c\beta_{12}^c)}{k_c} (R_1+f_1)^{-k_c}
$$

(7a)

and

$$
C_1^{(c)} \frac{(\beta_{11}^c + k_c\beta_{12}^c)}{k_c} (R_2-f_2)^{k_c} - C_2^{(c)} \frac{(\beta_{11}^c - k_c\beta_{12}^c)}{k_c} (R_2-f_2)^{-k_c}
$$
$$
= C_1^{(f_2)} \frac{(\beta_{11}^{f_2} + k_{f2}\beta_{12}^{f_2})}{k_{f2}} (R_2-f_2)^{k_{f2}}
$$
$$
- C_2^{(f_2)} \frac{(\beta_{11}^{f_2} - k_{f2}\beta_{12}^{f_2})}{k_{f2}} (R_2-f_2)^{-k_{f2}}. \quad (7b)
$$

Finally, the conditions of tractions at the lateral surfaces (traction-free inner surface and pressure, p, at the outer) give

$$
C_1^{(f_1)} R_1^{k_{f1}-1} + C_2^{(f_1)} R_1^{-k_{f1}-1} = 0, \quad (8a)
$$

$$
C_1^{(f_2)} R_2^{k_{f2}-1} + C_2^{(f_2)} R_2^{-k_{f2}-1} = -1. \quad (8b)
$$

The six linear Eqs. (6)–(8) can be solved for the six constants $C_1^{(i)}$, $C_2^{(i)}$, $(i=f_1,c,f_2)$. Other details of the solution can be found in Ref. [9].

Perturbed State. In the perturbed position we seek plane equilibrium modes as follows:

$$
u_i(r,\theta) = U_i(r)\cos n\theta; \quad v_i(r,\theta) = V_i(r)\sin n\theta;
$$

$$
w_i(r,\theta) = 0, \quad i=f_1,c,f_2. \quad (9)
$$

Substituting these in the strain versus displacement

$$
\epsilon_{rr}^{(i)} = \frac{\partial u_i}{\partial r}, \quad \epsilon_{\theta\theta}^{(i)} = \frac{1}{r}\frac{\partial v_i}{\partial \theta} + \frac{u_i}{r}, \quad \epsilon_{zz}^{(i)} = \frac{\partial w_i}{\partial z}, \quad (10a)
$$

$$
\gamma_{r\theta}^{(i)} = \frac{1}{r}\frac{\partial u_i}{\partial \theta} + \frac{\partial v_i}{\partial r} - \frac{v_i}{r}, \quad \gamma_{rz}^{(i)} = \frac{\partial u_i}{\partial z} + \frac{\partial w_i}{\partial r}, \quad \gamma_{\theta z}^{(i)} = \frac{\partial v_i}{\partial z} + \frac{1}{r}\frac{\partial w_i}{\partial \theta}
$$

(10b)

and rotation versus displacement relations

$$
2\omega_r^{(i)} = \frac{1}{r}\frac{\partial w_i}{\partial \theta} - \frac{\partial v_i}{\partial z}, \quad 2\omega_\theta^{(i)} = \frac{\partial u_i}{\partial z} - \frac{\partial w_i}{\partial r},
$$

$$
2\omega_z^{(i)} = \frac{\partial v_i}{\partial r} + \frac{v_i}{r} - \frac{1}{r}\frac{\partial u_i}{\partial \theta} \quad (10c)
$$

and then using the stress-strain relations in terms of the stiffness constants, c_{ij}^i

$$
\begin{bmatrix} \sigma_{rr}^{(i)} \\ \sigma_{\theta\theta}^{(i)} \\ \sigma_{zz}^{(i)} \\ \tau_{\theta z}^{(i)} \\ \tau_{rz}^{(i)} \\ \tau_{r\theta}^{(i)} \end{bmatrix} = \begin{bmatrix} c_{11}^i & c_{12}^i & c_{13}^i & 0 & 0 & 0 \\ c_{12}^i & c_{22}^i & c_{23}^i & 0 & 0 & 0 \\ c_{13}^i & c_{23}^i & c_{33}^i & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44}^i & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{55}^i & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66}^i \end{bmatrix} \begin{bmatrix} \epsilon_{rr}^{(i)} \\ \epsilon_{\theta\theta}^{(i)} \\ \epsilon_{zz}^{(i)} \\ \gamma_{\theta z}^{(i)} \\ \gamma_{rz}^{(i)} \\ \gamma_{r\theta}^{(i)} \end{bmatrix}, \quad i=f_1,c,f_2
$$

(10d)

the buckling Eqs. (1) result in the following two linear homogeneous ordinary differential equations of the second order for $U_i(r)$, $V_i(r)$, where $i=f_1$ for $R_1 \leqslant r \leqslant R_1+f_1$; $i=c$ for $R_1+f_1 \leqslant r \leqslant R_2-f_2$ and $i=f_2$ for $R_2-f_2 \leqslant r \leqslant R_2$.

$$
c_{11}^{(i)} U_i'' + \frac{c_{11}^{(i)}}{r} U_i' - \left[\left(c_{66}^{(i)} + \frac{\sigma_{\theta\theta}^{0(i)}}{2} \right) n^2 + c_{22}^{(i)} \right] \frac{U_i}{r^2}
$$

$$
+ \left(c_{12}^{(i)} + c_{66}^{(i)} - \frac{\sigma_{\theta\theta}^{0(i)}}{2} \right) \frac{nV_i'}{r}
$$

$$
- \left(c_{22}^{(i)} + c_{66}^{(i)} + \frac{\sigma_{\theta\theta}^{0(i)}}{2} \right) \frac{nV_i}{r^2} = 0 \quad (11a)
$$

and

Material	E_2 GPa	$E_1 = E_3$ GPa	G_{31} GPa	$G_{12} = G_{23}$ GPa	ν_{31}	$\nu_{21} = \nu_{23}$
FACE SHEETS						
Boron/epoxy	221.0	20.7	3.29	5.79	0.45	0.23
Graphite/epoxy	181.0	10.3	5.96	7.17	0.49	0.28
Kevlar/epoxy	75.9	5.52	1.89	2.28	0.47	0.34
CORE						
Alloy foam (isotropic)	0.0459	0.0459	0.0173	0.0173	0.33	0.33

$$\left(c_{66}^{(i)} + \frac{\sigma_{rr}^{0(i)}}{2}\right)V_i'' + \left(c_{66}^{(i)} + \sigma_{rr}^{0(i)} - \frac{\sigma_{\theta\theta}^{0(i)}}{2} + \frac{r\sigma_{rr}^{0(i)\prime}}{2}\right)\frac{V_i'}{r}$$

$$+ \left[-\left(c_{66}^{(i)} + c_{22}^{(i)}n^2 + \frac{\sigma_{\theta\theta}^{0(i)}}{2}\right) + \frac{r\sigma_{rr}^{0(i)\prime}}{2}\right]\frac{V_i}{r^2}$$

$$- \left(c_{12}^{(i)} + c_{66}^{(i)} - \frac{\sigma_{rr}^{0(i)}}{2}\right)\frac{nU_i'}{r} + \left[-\left(c_{66}^{(i)} + c_{22}^{(i)} + \frac{\sigma_{\theta\theta}^{0(i)}}{2}\right)\right.$$

$$\left. + \frac{r\sigma_{rr}^{0(i)\prime}}{2}\right]\frac{nU_i}{r^2} = 0. \tag{11b}$$

The associated boundary conditions are as follows:

(a) At the inner and outer bounding surfaces, we have the following two traction conditions at each of the surfaces:

$$c_{11}^{(j)}U_j' + \frac{c_{12}^{(j)}}{r}(U_j + nV_j) = 0 \tag{12a}$$

and

$$\left(c_{66}^{(j)} + \frac{\sigma_{rr}^{0(j)} + p_j}{2}\right)V_j' - \left(c_{66}^{(j)} - \frac{\sigma_{rr}^{0(j)} + p_j}{2}\right)\frac{V_j + nU_j}{r} = 0, \tag{12b}$$

where $j = f_1$ and $p_j = 0$ at $r = R_1$ (inner bounding surface) and $j = f_2$ and $p_j = p$ at $r = R_2$ (outer bounding surface).

(b) At the face-sheet/core interfaces, we have the following four conditions at each of the interfaces:

Traction Continuity:

$$c_{11}^{(j)}U_j' + \frac{c_{12}^{(j)}}{r}(U_j + nV_j) = c_{11}^{(c)}U_c' + \frac{c_{12}^{(c)}}{r}(U_c + nV_c) \tag{13a}$$

and

$$\left(c_{66}^{(j)} + \frac{\sigma_{rr}^{0(j)}}{2}\right)V_j' - \left(c_{66}^{(j)} - \frac{\sigma_{rr}^{0(j)}}{2}\right)\frac{V_j + nU_j}{r}$$

$$= \left(c_{66}^{(c)} + \frac{\sigma_{rr}^{0(c)}}{2}\right)V_c' - \left(c_{66}^{(c)} - \frac{\sigma_{rr}^{0(c)}}{2}\right)\frac{V_c + nU_c}{r}. \tag{13b}$$

Displacement Continuity:

$$U_j = U_c; \quad V_j = V_c, \tag{13c}$$

where $j = f_1$ at $r = R_1 + f_1$ (inner face-sheet/core interface) and $j = f_2$ at $r = R_2 - f_2$ (outer face-sheet/core interface).

Solution of the Eigen-Boundary-Value Problem for Differential Equations. Equations (11)–(13) constitute an eigenvalue problem for differential equations, with p the parameter (two point boundary value problem). An important point is that $\sigma_{rr}^{0(i)}(r)$, $\sigma_{\theta\theta}^{0(i)}(r)$ and $\sigma_{rr}^{0(i)\prime}(r)$ depend *linearly* on the external pressure, p (the parameter) through expressions in the form of Eqs. (8) and this makes possible the direct application of standard solution techniques.

With respect to the method used there is a difference between the present problem and the one for the homogeneous orthotropic body solved by Kardomateas [3]. The complication in the present problem is due to the fact that the displacement field is continuous but has a slope discontinuity at the face-sheet/core interfaces. This is the reason that the displacement field was not defined as one

Table 2 Critical pressure in N/m². Geometry: $f = 0.1$ in., $c = 1.0$ in. and $B = 3$ in.

R_0/h	Elasticity	Classical shell[a] no shear (% versus elast)	Shell w/shear[b] based on core only (% versus elast)	Shell w/shear[c] based on $\bar{G}$ (% versus elast)
BORON/EPOXY faces w/ALLOY-FOAM core				
15	741,773	6,898,740 (+930.0%)	651,125 (−12.2%)	899,768 (+21.3%)
30	277,305	862,343 (+310.9%)	253,721 (−8.5%)	323,361 (+16.6%)
60	70,416	107,793 (+53.0%)	67,383 (−4.3%)	76,087 (+8.0%)
120	11,817	13,474 (+14.0%)	11,717 (−0.85%)	12,203 (+3.3%)
GRAPHITE/EPOXY faces w/ALLOY-FOAM core				
15	720,842	5,650,460 (+783.9%)	637,826 (−11.5%)	874,654 (+21.3%)
30	258,549	706,307 (+273.2%)	238,236 (−7.9%)	298,643 (+15.5%)
60	61,528	88,288 (+43.5%)	59,207 (−3.8%)	65,825 (+7.0%)
120	9,918	11,036 (+11.3%)	9,829 (−0.9%)	10,168 (+2.5%)
KEVLAR/EPOXY faces w/ALLOY-FOAM core				
15	605,472	2,370,590 (+391.5%)	551,668 (−8.9%)	719,856 (+18.9%)
30	171,351	296,324 (+72.9%)	162,433 (−5.2%)	188,347 (+9.9%)
60	31,418	37,040 (+17.9%)	30,712 (−2.2%)	32,397 (+3.1%)
120	4,476	4,630 (+3.4%)	4,403 (−1.6%)	4,470 (−0.13%)

[a]Equation (14).
[b]Equation (17).
[c]Equation (18b).

Table 3 Critical pressure in N/m². Effect of increased face thickness: *f*=0.3 in., *c*=0.6 in., and *B*=3 in.

R_0/h	Elasticity	Classical shell no shear (% versus elast)	Shell w/shear based on core only (% versus elast)	Shell w/shear based on $\bar{G}$ (% versus elast)
GRAPHITE/EPOXY faces w/ALLOY-FOAM core				
15	1,244,010	11,731,900 (+943.1%)	416,091 (−66.6%)	1,501,160 (+20.7%)
30	393,573	1,466,490 (+372.6%)	188,038 (−52.2%)	542,378 (+37.8%)
60	105,699	183,311 (+73.4%)	67,900 (−35.8%)	128,553 (+21.6%)
120	19,297	22,914 (+18.7%)	16,081 (−16.7%)	20,709 (+7.3%)

function but as three distinct functions for $i=f_1$, c, and f_2, i.e., the two face sheets and the core. Our formulation of the problem employs, hence, "internal" boundary conditions at the face-sheet/core interfaces, as outlined above. Due to this complication, the shooting method [15] was deemed to be the best way to solve this eigen-boundary-value problem for differential equations. A special version of the shooting method was formulated and programmed for this problem. In fact, for each of the three constituent phases of the sandwich structure, we have five variables: $y_1=U_i$, $y_2=U_i'$, $y_3=V_i$, $y_4=V_i'$, and $y_5=p$. The five differential equations are: $y_1'=y_2$, the first equilibrium Eq. (11a), $y_3'=y_4$, the second equilibrium Eq. (11b) and $y_5'=0$.

The method starts from the inner boundary $r=R_1$ and integrates the five first order differential equations from R_1 to the inner face-sheet/core interface R_1+f_1 (i.e., through the inner face sheet). At the inner bounding surface, R_1, we have three conditions, the two traction boundary conditions, Eqs. (12), and a third condition of (abritrarily) setting $U_{f1}=1.0$, therefore we have two freely specifiable variables. The freely specifiable starting values at R_1 are taken as the y_5 (pressure), and the y_3 (V_{f1}) and these are taken as the values from the shell theory (described later). Then, the three boundary conditions at r_1 allow finding the starting values for y_1, y_2 and y_4. Once we reach the inner face-sheet/core interface, R_1+f_1, the tractions from the inner face-sheet side are calculated; these should equal the tractions from the core side, according to the boundary conditions on the face-sheet/core interface, Eqs. (13a) and (13b). This allows finding the slopes of the displacements, $y_2=U_c'$ and $y_4=V_c'$, for starting the shooting into the core (notice that the other three functions, $y_1=U_c$, $y_3=V_c$, and $y_5=p$ are continuous according to Eq. (13c), and their values at R_1+f_1 have already been found at the end of the integration step through the inner face sheet). The next step is integrating the five differential equations from R_1+f_1 to R_2-f_2, i.e., through the core. In a similar manner, once we reach the outer face-sheet/core interface, R_2-f_2, the tractions from the core side are calculated; these should equal the tractions from the outer face-sheet side, per Eq. (13a) and (13b), and this allows finding the slopes of the displacements, $y_2=U_{f2}'$ and $y_4=V_{f2}'$, for starting the shooting into the outer-face sheet (again, the other three functions are continuous and their values at R_2-f_2 have already been found at the end of the integration step through the core). The third step is the integration through the outer-face sheet. Once the outer bounding surface, R_2,

is reached, the traction boundary conditions, Eqs. (12), which ought to be zero, are calculated. Multi-dimensional Newton–Raphson is then used to develop a linear matrix equation for the two increments to the adjustable parameters, y_5 and y_3, at R_1. These increments are solved for and added and the shooting repeats until convergence. For the integration phase, we used a Runge–Kutta driver with adaptive step size control. The method produced results very fast and without any numerical complication.

Results, Comparison with Shell Theory and Discussion

As an illustrative example, consider a sandwich ring with the following geometry: core, $c=25.4$ mm (1 in.), face sheets $f_1=f_2=f=2.54$ mm (0.1 in.) and width $B=76.2$ mm (3 in.). This value for B was chosen in order to assume that buckling is in the plane of the ring and not out of the plane. Note that the sandwich is symmetric about its midsurface. The total thickness of the ring is, thus, $h=2f+c=30.48$ mm (1.2 in.), and is kept constant. The mean radius, R_0, is chosen in such a manner that the ratio R_0/h ranges from 15 to 120.

Material properties for the face sheets and the core are given in Table 1. The core is isotropic alloy foam and the face sheets are boron/epoxy or graphite/epoxy or kevlar epoxy unidirectional with 0 deg. orientation with respect to the hoop direction. Note again that 1 is the radial (r), 2 is the circumferential (θ), and 3 the axial (z) direction.

Notice also that by referring to Eq. (1), the compliance constants for each orthotropic phase are

$$a_{11}=\frac{1}{E_1}; \quad a_{22}=\frac{1}{E_2}; \quad a_{33}=\frac{1}{E_3}; \quad a_{44}=\frac{1}{G_{23}}; \quad a_{55}=\frac{1}{G_{31}};$$

$$a_{66}=\frac{1}{G_{12}},$$

$$a_{12}=-\frac{\nu_{21}}{E_2}; \quad a_{13}=-\frac{\nu_{31}}{E_3}; \quad a_{23}=-\frac{\nu_{32}}{E_3}.$$

Since the shell is considered to be very long, the buckling analysis reduces to that for a ring [12]. If the transverse shear effect is neglected, the expression for the pressure becomes (classical)

Table 4 Critical pressure in N/m². Comparison with homogeneous: *f*=0.1 in., *c*=1.0 in., and *B*=3 in.

R_0/h	Elasticity	Clasical shell no shear (% versus elast)	Shell w/shear based on core only (% versus elast)	Shell w/shear based on $\bar{G}$ (% versus elast)
GRAPHITE/EPOXY homogeneous (no sandwich)				
15	12,594,400	13,407,400 (+6.5%)	12,831,500 (+1.9%)	12,924,100 (+2.6%)
30	1,641,360	1,675,930 (+2.1%)	1,657,330 (+0.97%)	1,660,400 (+1.2%)
60	208,228	209,491 (+0.61%)	208,905 (+0.33%)	209,002 (+0.37%)
120	26,180	26,186 (+0.03%)	26,168 (−0.05%)	26,171 (−0.03%)

$$p_{c\ell} = 3 \frac{(EI)_{eq}}{BR_0^3}, \qquad (14)$$

where $(EI)_{eq}$ is the equivalent bending rigidity, given in terms of the extensional moduli of the face sheets E_f and the core E_c by [16]

$$(EI)_{eq} = w\left[E_f \frac{f^3}{6} + 2E_f f\left(\frac{f}{2} + \frac{c}{2}\right)^2 + E_c \frac{c^3}{12}\right]. \qquad (15a)$$

If the transverse shear effect is accounted for, then

$$p_{w/shear} = 3 \frac{(EI)_{eq}}{BR_0^3(1 + 4k_s)}; \quad k_s = \frac{(EI)_{eq}}{CR_0^2}, \qquad (16a)$$

where

$$C = \int_A KGdA, \qquad (16b)$$

K being a shear correction factor taken as equal to one and G is the transverse shear stiffness of the sandwich cross section.

Two different expressions for C are employed herein. In the first case, it is assumed that only the core contributes, in which case, $C = BcG_{12}^c$ and

$$k_{s1} = \frac{(EI)_{eq}}{BcG_{12}^c R_0^2}, \qquad (17)$$

where G_c is the shear modulus of the core.

In the second case, an effective shear modulus for the sandwich section, $\bar{G}$, which includes the contribution of the facings, is derived based on the compliances of the constituent phases [16]. The expression for $\bar{G}$ is given by

$$\frac{2f + c}{\bar{G}} = \frac{2f}{G_{12}^f} + \frac{c}{G_{12}^c}, \qquad (18a)$$

where G_{12}^f is the shear modulus of the facings. Therefore, in this case

$$k_{s2} = \frac{(EI)_{eq}}{B(2f + c)\bar{G}R_0^2}. \qquad (18b)$$

Table 2 gives the critical pressure from the elasticity formulation for a range of mean radius over total thickness ratios, in comparison with the classical shell and the two shear deformable shell formulas.

In all cases, $n = 2$ was used in the buckling modes, Eq. (9). This has been well established for isotropic cylindrical shells under external pressure; however, since we are dealing with a sandwich structure whose core has elastic properties that are orders of magnitude different from those of the face sheets, verification of this postulate was needed. Indeed, in all cases examined, an exhaustive search was made for the n that results in the minimum eigenvalue, and it was indeed found that $n = 2$ corresponds to the lowest eigenvalue. For example, for the case of graphite/epoxy faces with alloy-foam core and $R_0/h = 30$, the eigenvalues found from the elasticity solution were (in N/m²) as follows: $(n = 2; 258,549)$, $(n = 3; 397,355)$, $(n = 4; 469,798)$, $(n = 5; 512,410)$.

Now coming to the results in Table 2, it is seen that the classical (no shear) formula can yield results highly nonconservative, even approaching ten times the elasticity value for the lower ratio of R_0/h and boron/epoxy case. Both shear correction formulas yield reasonable results with the shear correction formula based on the core only being in general conservative as opposed to the shear correction formula based on an "effective shear modulus," $\bar{G}$, which is nonconservative.

An illustration of the results in Table 2 is provided in Fig. 2, which shows the critical pressure, p_{cr}, normalized with the simple formula from classical shell theory, $p_{c\ell}$, Eq. (14), as a function of
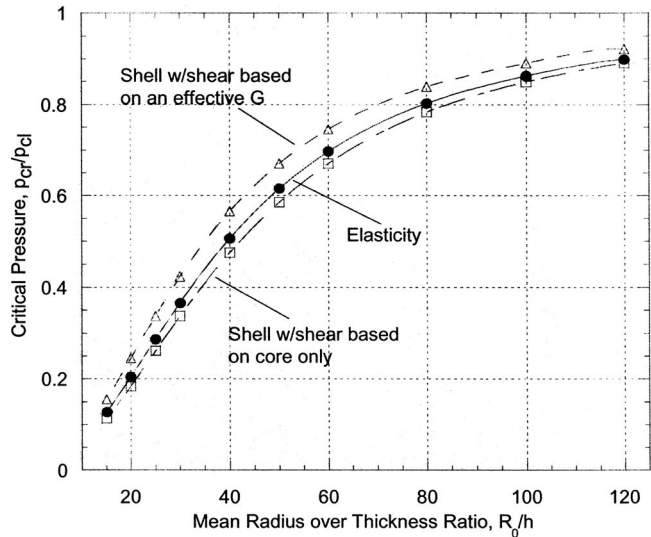
Fig. 2 Critical pressure, p_{cr}, normalized with the classical shell formula, $p_{c\ell}$, Eq. (14)

the mean radius over thickness ratio, R_0/h. The results are derived from the elasticity formulation and the two shell theory formulas with transverse shear, for the case of graphite/epoxy faces with alloy-foam core. The results show clearly the very significant effect of transverse shear as the ratio R_0/h becomes smaller (thicker shell), in the sense that p_{cr} is only about 12% of the $p_{c\ell}$ (which ignores transverse shear) for $R_0/h = 15$. It is also seen that the elasticity results are between the two shells with shear correction formulas, as already discussed in the previous paragraph. For thinner shells, the transverse shear effects get diminished; for example, for $R_0/h = 120$, the p_{cr} is about 90% of the $p_{c\ell}$.

In the results presented in Table 2, the face sheets were quite thin and the shear correction formula based on the core only, Eq. (17), seemed to be more accurate. In order to further examine this premise, the critical load was calculated for a construction in which the total thickness remains the same but the face sheet thickness is increased at the expense of the core. The results, listed in Table 3, show that the shear correction formula based on an effective modulus (which includes the core), Eq. (18b), is now more accurate.

In order to compare with the homogeneous, monolithic, Table 4 gives the critical pressure for a construction made of graphite/epoxy homogeneous, i.e., no sandwich. It is seen that the differences from the elasticity values are modest, even with the classical shell formula. This illustrates the nature of sandwich construction, in which buckling is a more demanding issue.

Acknowledgment

References

[1] Hutchinson, J. W., 1968, "Buckling and Initial Postbuckling Behavior of Oval Cylindrical Shells Under Axial Compression," J. Appl. Mech., **35**, pp. 66–72.
[2] Budiansky, B., and Amazigo, J. C., 1968, "Initial Post-Buckling Behavior of Cylindrical Shells Under External Pressure," J. Math. Phys. (Cambridge, Mass.), **47**(3), pp. 223–235.
[3] Kardomateas, G. A., 1993a, "Buckling of Thick Orthotropic Cylindrical Shells Under External Pressure," ASME J. Appl. Mech., **60**, pp. 195–202.
[4] Kardomateas, G. A., and Chung, C. B., 1994, "Buckling of Thick Orthotropic Cylindrical Shells Under External Pressure Based on Non-Planar Equilibrium Modes," Int. J. Solids Struct., **31**(16), pp. 2195–2210.
[5] Kardomateas, G. A., 1993b, "Stability Loss in Thick Transversely Isotropic

Cylindrical Shells Under Axial Compression," ASME J. Appl. Mech., **60**, pp. 506–513.

[6] Kardomateas, G. A., 1995, "Bifurcation of Equilibrium in Thick Orthotropic Cylindrical Shells Under Axial Compression," ASME J. Appl. Mech., **62**, pp. 43–52.

[7] Soldatos, K. P., and Ye, J.-Q., 1994, "Three-Dimensional Static, Dynamic, Thermoelastic, and Buckling Analysis of Homogeneous and Laminated Composite Cylinders," Compos. Struct., **29**, pp. 131–143.

[8] Lekhnitskii, S. G., 1963, *Theory of Elasticity of an Anisotropic Elastic Body*, Holden Day, San Francisco, also Mir, Moscow, 1981.

[9] Kardomateas, G. A., 2001, "Elasticity Solutions for a Sandwich Orthotropic Cylindrical Shell Under External Pressure, Internal Pressure and Axial Force," AIAA J., **39**(4), pp. 713–719.

[10] Birman, V., and Simitses, G. J., 2000, "Theory of Cylindrical Sandwich Shells with Dissimilar Facings Subjected to Thermo-mechanical Loads," AIAA J., **37**(12), pp. 362–367.

[11] Birman, V., Simitses, G. J., and Shen, L., 2000, "Stability of Short Sandwich Cylindrical Shells with Rib-Reinforced Facings," *Recent Advances in Applied Mechanics*, J. T. Katsikadelis, D. E. Beskos, and E. E. Gdoutos, eds., National Technical University of Athens, Greece, pp. 11–21.

[12] Birman, V., and Simitses, G. J., 1999, "Stability of Long Cylindrical Sandwich Shells with Dissimilar Facings Subjected to Lateral Pressure," *Advances in Aerospace Materials and Structures*, G. Newaz, ed., ASME AD-58, ASME, New York, pp. 41–51.

[13] Smith, C. V., Jr., and Simitses, G. J., 1969, "Effect of Shear and Load Behavior on Ring Stability," ASCE, J. of EM Division, **95**(3), pp. 559–569.

[14] Simitses, G. J., and Aswani, M., 1974, "Buckling of Thin Cylinders Under Uniform Lateral Loading," ASME J. Appl. Mech., **41**(3), pp. 827–829.

[15] Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., 1989, *Numerical Recipes*, Cambridge University Press, Cambridge, UK.

[16] Huang, H., and Kardomateas, G. A., 2002, "Buckling and Initial Postbuckling Behavior of Sandwich Beams Including Transverse Shear," AIAA J., **40**(11), pp. 2331–2335.

Anisotropic Elastic Tubes of Arbitrary Cross Section Under Arbitrary End Loads: Separation of Beamlike and Decaying Solutions

P. Ladevèze

Laboratoire de Mecanique et Technologie,
ENS de Cachan, CNRS-Paris VI,
Paris, France

J. G. Simmonds[1]

Fellow ASME
Department of Civil Engineering,
University of Virginia,
Charlottesville, VA 22904-4742

*First approximation analytical solutions are constructed for finite and semi-infinite, fully anisotropic elastic tubes of constant thickness h and arbitrary cross section, subject to purely kinetic or purely kinematic boundary conditions. Final results contain relative errors of $O(\sqrt{h/R})$, where R is some equivalent cross sectional radius. Solutions are decomposed into the sum of an exact beamlike or Saint-Venant solution, treated in Ladevèze et al. (Int. J. Solids Struct., **41**, pp. 1925–1944, 2004) and extended in an appendix; a rapidly decaying edge-zone solution; and a slowly decaying semi-membrane-inextensional-bending (MB) solution. Explicit conditions on the boundary data are given that guarantee decaying solutions. The MB solutions are expressed as an infinite series of complex-valued exponential functions times real-valued one-dimensional eigenfunctions which satisfy a fourth-order differential equation in the circumferential coordinate and depend on the pointwise cross sectional curvature only.* [DOI: 10.1115/1.1934532]

1 Introduction

Consider a straight tube constructed of elastic, anisotropic layers of constant but possibly different thicknesses. If the tube is under surface loads and any combination of end loads and displacements (compatible with overall equilibrium and no rigid-body displacement), then, as Ladevèze and Simmonds [1,2] have shown (within the framework of linear elasticity), the solution of the governing equations may be decomposed *exactly* into a beamlike part and a decaying part. We shall refer to these two components as a *Saint-Venant* (SV) part and a *decaying* (D) part, respectively.

As $\varepsilon \equiv h/R$, the constant thickness of the tube divided by some typical radius of the cross section, approaches zero, the D part of the solution displays *three* characteristic decay lengths: (1) a very short length, $O(h)$, associated with a three-dimensional edge effect; (2) a moderately short length, $O(\sqrt{hR})$, associated with the bending edge effect of classical (first-approximation) shell theory; and (3) a very long length, $O(R\sqrt{R/h})$, associated with the *semimembrane-inextensional bending* (MB) behavior of the shell. (We note that each of these decay lengths also depends strongly on the ratios of various anisotropic elastic coefficients.)

Ladevèze et al. [3], using the linear first-approximation shell theory of Sanders [4] and Koiter [5], have analyzed the beamlike behavior of an elastic tube of arbitrary anisotropy and cross section. Some corrections and extensions are presented in the Appendix. In the present paper, we examine the complementary solu-tions associated with the decay lengths (2) and (3) mentioned above in the absence of surface loads; an accurate resolution of the three-dimensional zone (1) lies outside the range of *any* shell theory (despite many claims to the contrary in the literature). (For references to work on this latter problem, see, for example, the book edited by Ladevèze [6] or an earlier review paper by Simmonds [7] where citations of relevant papers by Friedrichs and Dressler, Goldenveiser, Green, Gregory and Wan, Ladevèze, and others are given.) In particular, for the two extreme cases of (A) end loads only or (B) kinematic end conditions only, we develop conditions on the data that insure decaying solutions. Moreover, we show that these data may be decomposed into a set that, to lowest order, i.e., to within a relative error of $O(\varepsilon^{1/2})$, determines the MB solution and a complementary set that subsequently determines the EZ solution. We note that, except in relatively long tubes, the MB solution may well be as important in the interior of the tube as the beamlike solution.

2 Geometry

In a fixed Euclidean frame $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, let (r, θ, x) denote a set of circular cylindrical coordinates with associated orthonormal base vectors $[\mathbf{e}_r(\theta), \mathbf{e}_\theta(\theta), \mathbf{k}]$. We take the vector representation of the reference surface of the tube to be

$$\mathcal{T}: \mathbf{x} = R[x\mathbf{k} + \mathbf{r}(y)], \quad \mathbf{r} \in \mathcal{S}, \quad 0 \leqslant x \leqslant l, \quad 0 \leqslant y \leqslant 2\pi. \quad (1)$$

Here, $2\pi R$ is the circumference of $\mathcal{T}$, and x and y are, respectively, dimensionless distances along and around $\mathcal{T}$. Differentiation with respect to x and y will be denoted by a prime ($'$) and a dot ($\bullet$), respectively. Thus, with $r \equiv |\mathbf{r}|$ denoting dimensional radial distance, we have

$$\mathbf{r} = r(y)\mathbf{e}_r(\theta) \text{ and } \theta = \pm \int_0^y \frac{\sqrt{1 - r^{\cdot 2}}(\eta) d\eta}{r(\eta)}, \quad (2)$$

where the $\pm$ sign allows for the possibility that $\mathcal{S}$ might not be star shaped with respect to the chosen axis of $\mathcal{T}$. (If $\mathcal{T}$ is a circular cylinder, $r = 1$ and $y = \theta$.) Finally, we let

$$\kappa = (\mathbf{r}^\bullet \times \mathbf{r}^{\bullet\bullet}) \bullet \mathbf{k} \qquad (3)$$

denote the dimensionless curvature of $\mathcal{S}$.

3 Governing Equations

When convenient, we use Cartesian tensor notation, with $x = x_1$ and $y = x_2$. Thus, let $\sigma h N_{\alpha\beta}$, $\sigma h^2 M_{\alpha\beta}$, $(\sigma/\bar{E})E_{\alpha\beta}$, and $(\sigma/h\bar{E})K_{\alpha\beta}$, $\alpha = 1, 2$, denote, respectively, the (modified, symmetric) stress resultants, stress couples, extensional strains, and bending strains of the linear shell theory of Sanders [4] and Koiter [5], where σ is some measure of the stress level in the tube and $\bar{E}$ is some nominal Young's modulus. In component form, with the notation

$$T_{\alpha\beta} = \{T_x, T, T_y\}, \qquad (4)$$

the equilibrium and compatibility conditions of the Sanders-Koiter theory are

$$N_x' + [N - (1/2)\varepsilon\kappa M]^\bullet = 0, \qquad (5)$$

$$N' + N_y^\bullet + \varepsilon\kappa[3/2M' + M_y^\bullet] = 0, \qquad (6)$$

$$-\varepsilon(M_x'' + 2M'^\bullet + M_y^{\bullet\bullet}) + \kappa N_y = 0, \qquad (7)$$

$$K_y' - [K + (1/2)\varepsilon\kappa E]^\bullet = 0, \qquad (8)$$

$$-K' + K_x^\bullet + \varepsilon\kappa[(3/2)E' - E_x^\bullet] = 0, \qquad (9)$$

$$\varepsilon(E_y'' - 2E'^\bullet + E_x^{\bullet\bullet}) + \kappa K_x = 0. \qquad (10)$$

These equations have been written in a form that displays the *static-geometric duality* of Goldenveiser [8] and Lure [9]. This duality implies that the equilibrium Eqs. (5)–(7) go over into the compatibility conditions (8)–(10) if the variables below on the left are replaced by those on the right

$$\hat{N}_{\alpha\beta} : K_{\alpha\beta}, \qquad M_{\alpha\beta} : -\hat{E}_{\alpha\beta}. \qquad (11)$$

Here, we have introduced the "hat" notation

$$\hat{T}_{\alpha\beta} = e_{\alpha\lambda}e_{\beta\mu}T_{\lambda\mu} = \{T_y, -T, T_x\}, \qquad (12)$$

where $e_{\alpha\beta}$ is the two-dimensional alternator.

To complete the set of field equations, we must add constitutive relations. To exploit fully the economy offered by the static-geometric duality, we follow McDevitt and Simmonds [10] and write these in the form

$$-\hat{E}_{\alpha\beta} = \psi_{,\hat{N}_{\alpha\beta}} = -A_{\alpha\beta\lambda\mu}\hat{N}_{\lambda\mu} + C_{\alpha\beta\lambda\mu}K_{\lambda\mu} \qquad (13)$$

$$M_{\alpha\beta} = \psi_{,K_{\alpha\beta}} = A^*_{\alpha\beta\lambda\mu}K_{\lambda\mu} + C^*_{\alpha\beta\lambda\mu}\hat{N}_{\lambda\mu}, \qquad (14)$$

where $A^*_{\alpha\beta\lambda\mu}$ is the dual of $-A_{\alpha\beta\lambda\mu}$, $C^*_{\alpha\beta\lambda\mu} \equiv C_{\lambda\mu\alpha\beta}$ is the dual of $C_{\alpha\beta\lambda\mu}$, and

$$\psi = (1/2)(A^*_{\alpha\beta\lambda\mu}K_{\alpha\beta}K_{\lambda\mu} + C_{\alpha\beta\lambda\mu}\hat{N}_{\alpha\beta}K_{\lambda\mu} + C^*_{\alpha\beta\lambda\mu}K_{\alpha\beta}\hat{N}_{\lambda\mu}$$
$$- A_{\alpha\beta\lambda\mu}\hat{N}_{\alpha\beta}\hat{N}_{\lambda\mu}) \qquad (15)$$

is the dimensionless mixed-energy density. The quadratic form (15) implies that, in general, there are 21 dimensionless elastic coefficients.

Finally, we add the auxiliary strain-displacement relations of the Sanders-Koiter theory (which must be integrated to obtain displacements for use in boundary conditions). If $(R^2\sigma/\bar{E}h)\mathbf{U}$ denotes the displacement field, where

$$\mathbf{U} = U(x,y;\varepsilon)\mathbf{k} + V(x,y;\varepsilon)\mathbf{t}(y) - W(x,y;\varepsilon)\mathbf{n}(y) \qquad (16)$$

and $\mathbf{t} = \mathbf{r}^\bullet(y)$, $\mathbf{n} = \mathbf{k} \times \mathbf{t}(y)$, then these take the form

$$\varepsilon E_x = U', \quad 2\varepsilon E = U^\bullet + V', \quad \varepsilon E_y = V^\bullet + \kappa W \qquad (17)$$

and

$$K_x = -W'', \quad K = -W'^\bullet + (3/4)\kappa V' - (1/4)\kappa U^\bullet,$$
$$K_y = -(W^\bullet - \kappa V)^\bullet. \qquad (18)$$

(Analogous stress-stress function relations exist, but are not needed in what follows.)

Let $\sigma h R \mathbf{T}$ denote the net traction acting on the material to the left of any cross section of the tube and let $\sigma h R^2 \mathbf{M}$ denote the net moment with respect to the centroid ($\mathbf{r} = \mathbf{0}$) of the cross section. Then

$$\mathbf{T} = \int_0^{2\pi} \mathbf{N}_x(x,y;\varepsilon)dy$$

and

$$\mathbf{M} = \int_0^{2\pi} [\mathbf{r}(y) \times \mathbf{N}_x(x,y;\varepsilon) - \varepsilon M_x(x,y;\varepsilon)\mathbf{t}(y)]dy, \qquad (19)$$

where

$$\mathbf{N}_x = N_x\mathbf{k} + [N + (3/2)\varepsilon\kappa M]\mathbf{t} - \varepsilon(\mathbf{M}_x' + 2M^\bullet)\mathbf{n} \qquad (20)$$

is the *effective* dimensionless axial stress resultant in the Sanders-Koiter theory. (See Budiansky and Sanders [11], Eq. 15.) The beamlike (SV) solutions developed by Ladevèze et al. [3], in the absence of surface loads, satisfy the global beam equations

$$\mathbf{T} = \mathbf{T}(0) \quad \text{and} \quad \mathbf{M} = \mathbf{M}(0) - x\mathbf{k} \times \mathbf{T}(0) \qquad (21)$$

as well as all the local field equations [providing certain negligible terms of $O(\varepsilon)$ are added to the stress-strain relations]. The decaying EZ and MB solutions we now develop satisfy all the local field equations to $O(\varepsilon^{1/2})$ plus the global conditions $\mathbf{T} = \mathbf{M} = \mathbf{0}$.

4 Edge-zone (EZ) Solutions

To extract these from the field equations, we scale certain of the variables as follows:

$$x = \varepsilon^{1/2}\bar{x}, \quad (N,K) = \varepsilon^{1/2}(\bar{N},\bar{K}), \quad (N_x,K_y) = \varepsilon(\bar{N}_x,\bar{K}_y)$$
$$(U,V) = \varepsilon^{3/2}(\bar{U},\bar{V}), \quad W = \varepsilon\bar{W}, \qquad (22)$$

where an overbar indicates that a variable is a function of $\bar{x}$, y, and ε only. Then, we set $\partial()/\partial\bar{x} = ()\#$ and assume that differentiation with respect to $\bar{x}$ and y does not change orders of magnitude. Thus, Eqs. (5)–(10), (13), (14), (17), and (18) assume the forms

$$\bar{N}_x^\# + \bar{N}^\bullet = O(\varepsilon^{1/2}), \qquad (23)$$

$$\bar{N}^\# + \bar{N}_y^\bullet + O(\varepsilon^{1/2}), \qquad (24)$$

$$-\bar{M}_x^{\#\#} + \kappa\bar{N}_y = O(\varepsilon^{1/2}), \qquad (25)$$

$$\bar{K}_y^\# - \bar{K}^\bullet = O(\varepsilon^{1/2}), \qquad (26)$$

$$-\bar{K}^\# + \bar{K}_x^\bullet = O(\varepsilon^{1/2}), \qquad (27)$$

$$\bar{E}_y^{\#\#} + \kappa\bar{K}_x = O(\varepsilon^{1/2}), \qquad (28)$$

$$\bar{E}_{\alpha\beta} = \hat{A}_{\alpha\beta 11}\bar{N}_y - \hat{C}_{\alpha\beta 11}\bar{K}_x + O(\varepsilon^{1/2}), \qquad (29)$$

$$\bar{M}_{\alpha\beta} = A^*_{\alpha\beta 11}\bar{K}_x + C_{11\alpha\beta}\bar{N}_y + O(\varepsilon^{1/2}), \qquad (30)$$

$$\bar{E}_x = \bar{U}^\#, \quad 2\bar{E} = \bar{V}^\# + O(\varepsilon^{1/2}), \quad \bar{E}_y = \kappa\bar{W} + O(\varepsilon^{1/2}), \qquad (31)$$

$$\bar{K}_x = -\bar{W}^{\#\#}, \quad \bar{K} = -\bar{W}^{\#\bullet} + O(\varepsilon^{1/2}), \quad \bar{K}_y = -\bar{W}^{\bullet\bullet} + O(\varepsilon^{1/2}). \qquad (32)$$

From (25) and (30), we have

$$-(A^*_{1111}\bar{K}_x + C^*_{1111}\bar{N}_y)^{\#\#} + \kappa\bar{N}_y = O(\varepsilon^{1/2}), \qquad (33)$$

and from their duals, (28) and (29)

$$(A_{1111}\bar{N}_y - C_{1111}\bar{K}_x)^{\#\#} + \kappa\bar{K}_x = O(\varepsilon^{1/2}). \qquad (34)$$

Noting that $C_{1111} = C^*_{1111}$, we introduce the complex-valued unknown

$$\bar{\mathcal{N}}_y \equiv \bar{N}_y + i\sqrt{A^*_{1111}/A_{1111}}\bar{K}_x \qquad (35)$$

and constant

$$\bar{\mu}^2 = \frac{i}{\sqrt{A_{1111}A^*_{1111}} + iC_{1111}}. \qquad (36)$$

Then Eqs. (33) and (34) may be combined as

$$\bar{\mathcal{N}}_y^{\#\#} - \bar{\mu}^2\kappa(y)\bar{\mathcal{N}}_y = O(\varepsilon^{1/2}). \qquad (37)$$

Ignoring the $O(\varepsilon^{1/2})$ term, we write the general solution of Eq. (37) as

$$\bar{\mathcal{N}}_y = \bar{C}_\pm(y)\{\exp[\pm\bar{\mu}\kappa^{1/2}(y)\bar{x}]\}^\#, \qquad (38)$$

where the complex-valued function $\bar{C}_\pm$ will be determined from the boundary conditions by combining them with the MB solutions that we determine next. (If $\kappa < 0$, $\kappa^{1/2}$ is imaginary.)

5 Semimembrane-Inextensional-Bending (MB) Solutions

To obtain equations for the MB solutions, we scale certain variables as follows:

$$\varepsilon^{1/2}x = \tilde{x}, \quad (N,K) = \varepsilon^{1/2}(\tilde{N},\tilde{K}), \quad (N_y,K_x) = \varepsilon(\tilde{N}_y,\tilde{K}_x), \quad U = \varepsilon^{1/2}\tilde{U}, \qquad (39)$$

where a tilde denotes that a variable depends on $\tilde{x}$, y, and ε only. Then, with $\partial()/\partial\tilde{x} = ()^\circ$ and the assumption that differentiation with respect to $\tilde{x}$ does not change orders of magnitude, Eqs. (5)–(10), (13), (14), (17), and (18) take the form

$$\tilde{N}_x^\circ + \tilde{N}^\bullet = O(\varepsilon^{1/2}) \qquad (40)$$

$$\tilde{N}^\circ + \tilde{N}_y^\bullet + \kappa\tilde{M}_y^\bullet = O(\varepsilon^{1/2}), \qquad (41)$$

$$-\tilde{M}_y^{\bullet\bullet} + \kappa\tilde{N}_y = O(\varepsilon^{1/2}), \qquad (42)$$

$$\tilde{K}_y^\circ - \tilde{K}^\bullet = O(\varepsilon^{1/2}), \qquad (43)$$

$$-\tilde{K}^\circ + \tilde{K}_x^\bullet - \kappa\tilde{E}_x^\bullet = O(\varepsilon^{1/2}), \qquad (44)$$

$$\tilde{E}_x^{\bullet\bullet} + \kappa\tilde{K}_x = O(\varepsilon^{1/2}), \qquad (45)$$

$$\tilde{E}_{\alpha\beta} = \hat{A}_{\alpha\beta22}\tilde{N}_x - \hat{C}_{\alpha\beta22}\tilde{K}_y + O(\varepsilon^{1/2}), \qquad (46)$$

$$\tilde{M}_{\alpha\beta} = A^*_{\alpha\beta22}\tilde{K}_y + C^*_{\alpha\beta22}\tilde{N}_x + O(\varepsilon^{1/2}), \qquad (47)$$

$$\tilde{E}_x = \tilde{U}^\circ, \quad \tilde{U}^\bullet + \tilde{V}^\circ = O(\varepsilon^{1/2}), \quad \tilde{V}^\bullet + \kappa\tilde{W} = O(\varepsilon), \qquad (48)$$

$$\tilde{K}_x = -\tilde{W}^{\circ\circ}, \quad \tilde{K} = -\tilde{W}^{\circ\bullet} + (3/4)\kappa\tilde{V}^\circ - (1/4)\kappa\tilde{U}^\bullet, \qquad (49)$$

$$\tilde{K}_y = -(\tilde{W}^\bullet - \kappa\tilde{V})^\bullet.$$

By differential elimination of $\tilde{N}$ and $\tilde{N}_y$, Eqs. (40)–(42) reduce to

$$\tilde{N}_x^{\circ\circ} = (\mathcal{M}\tilde{M}_y^\bullet)^\bullet + O(\varepsilon^{1/2}), \qquad (50)$$

where

$$\mathcal{M} \equiv \frac{d}{dy}\left[\rho(y)\frac{d}{dy}\right] + \kappa(y), \quad \rho = \kappa^{-1}(y). \qquad (51)$$

Substituting Eq. (47) into Eq. (50) and noting that $C_{2222} = C^*_{2222}$, we have

$$\tilde{N}_x^{\circ\circ} = [\mathcal{M}(C_{2222}\tilde{N}_x + A^*_{2222}\tilde{K}_y)^\bullet]^\bullet + O(\varepsilon^{1/2}). \qquad (52)$$

The dual of Eq. (52) follows readily as

$$\tilde{K}_y^{\circ\circ} = [\mathcal{M}(C_{2222}\tilde{K}_y - A_{2222}\tilde{N}_x)^\bullet]^\bullet + O(\varepsilon^{1/2}). \qquad (53)$$

If we set

$$\tilde{\mathcal{N}}_x = \tilde{N}_x + i\sqrt{A^*_{2222}/A_{2222}}\tilde{K}_y \quad \text{and} \quad \tilde{\mu}^2 = C_{2222} - i\sqrt{A_{2222}A^*_{2222}}, \qquad (54)$$

then Eqs. (52) and (53) may be combined into the single complex-valued equation

$$\tilde{\mathcal{N}}_x^{\circ\circ} = \tilde{\mu}^2(\mathcal{M}\tilde{\mathcal{N}}_x^\bullet)^\bullet + O(\varepsilon^{1/2}). \qquad (55)$$

To $O(\varepsilon^{1/2})$, Eq. (55) admits homogeneous solutions of the form

$$\tilde{\mathcal{N}}_x = \tilde{C}_\pm \exp(\pm\lambda\tilde{\mu}\tilde{x})\Omega(y), \qquad (56)$$

where the $\tilde{C}_\pm$ are complex-valued constants and Ω is a *real-valued* function satisfying the differential equation

$$\lambda^2\Omega = (\mathcal{M}\Omega^\bullet)^\bullet = [\rho(y)\Omega^{\bullet\bullet}]^{\bullet\bullet} + [\kappa(y)\Omega^\bullet]^\bullet \qquad (57)$$

and auxiliary condition

$$\Omega(y + 2\pi) = \Omega(y). \qquad (58)$$

Note from Eqs. (57) and (58) that if $\lambda \neq 0$, then $\int_0^{2\pi}\Omega dy = 0$.

Since Ω is 2π-periodic, Eqs. (57) and (58) represent a standard, self-adjoint eigenvalue problem which, if κ is suitably smooth and nonvanishing on $[0,2\pi]$ (as we shall assume), admits a countable set of real-valued eigensolutions, $\{\lambda_n^2, \Omega_n\}_{n=0}^\infty$, where $0 = \lambda_0 < \lambda_1 < \lambda_2 < \cdots$ and the Ω_n satisfy the orthonormality condition

$$\int_0^{2\pi}\Omega_m(y)\Omega_n(y)dy = \delta_{mn}, \quad m,n = 0,1,2,\dots, \qquad (59)$$

where δ_{mn} is the Kronecker delta. The two eigenfunctions corresponding to $\lambda = 0$ have the explicit forms

$$\overset{0}{\Omega}_0 = 1/\sqrt{2\pi} \quad \text{and} \quad \overset{1}{\Omega}_0 = \mathbf{c}\bullet\mathbf{r}(y). \qquad (60)$$

Here, $\mathbf{c}$ is a constant vector perpendicular to $\mathbf{k}$ satisfying $\pi\mathbf{c}\bullet\mathbf{I}\bullet\mathbf{c} = 1$, where

$$\mathbf{I} = (1/\pi)\int_0^{2\pi}\mathbf{r}(y)\mathbf{r}(y)dy = \mathbf{I}^T. \qquad (61)$$

(For a circular cross section, $\mathbf{I}$ is the two-dimensional identity tensor $\mathbf{1}$.) To make $\overset{0}{\Omega}_0$ orthogonal to $\overset{1}{\Omega}_0$, we take the tail of $\mathbf{r}$ to coincide with the centroid of the cross section of the reference surface so that

$$\int_0^{2\pi}\mathbf{r}(y)dy = \mathbf{0}. \qquad (62)$$

The nondecaying solutions associated with the eigenfunctions (60) are the Saint-Venant (beamlike) solutions discussed in [3] and the Appendix.

6 Reciprocity

As in Ladevèze and Simmonds [2], the Betti reciprocity principle may be used to ensure that end data produce decaying solutions. To present this principle in its various forms, let

$$s \equiv (\mathbf{V}, \varepsilon E_y, \varepsilon^{1/2} Z_*; \mathbf{F}, \varepsilon^{1/2} Z, \varepsilon M_x) \equiv (\underline{V}, \underline{F}) \qquad (63)$$

denote a kinematic-kinetic pair, where Z and its dual Z_* are defined in Eq. (A6), and

$$\mathbf{V} \equiv U(x,y;\varepsilon)\mathbf{k} + V(x,y;\varepsilon)\mathbf{t}(y) \text{ and } \mathbf{F} \equiv F(x,y;\varepsilon)\mathbf{k} + S(x,y;\varepsilon)\mathbf{t}. \qquad (64)$$

If we introduce the following inner product at any section $x=$ constant;

$$[s,\bar{s}]_x \equiv \int_0^{2\pi} (\underline{F} \cdot \bar{\underline{V}} - \bar{\underline{F}} \cdot \underline{V})_x dy, \qquad (65)$$

then, for any two states s and $\bar{s}$ satisfying the field Eqs. (5)–(10) the Betti reciprocity principle in the Sanders-Koiter theory implies that

$$[s,\bar{s}]_a = [s,\bar{s}]_x, \quad 0 \leqslant a < x. \qquad (66)$$

That is, $[s,\bar{s}]_x$ is a constant.

7 End-Loaded, Semi-Infinite Tube

We consider first a semi-infinite tube. Of the various combinations of kinematic-kinetic end conditions that might be imposed in accordance with Eq. (A7), the dimensionless form of the external virtual work, we confine ourselves to two extreme cases: (A) $\underline{F}$ (i.e., $\mathbf{F}$, Z, and M_x) prescribed and (B) $\underline{V}$ (i.e., $\mathbf{V}$, Z_*, and E_y) prescribed.

7.1 Case A: End Loads Prescribed. In Eq. (66) let s be a linear combination of rapidly decaying edge-zone (EZ) solutions and slowly decaying semimembrane-inextensional-bending (MB) solutions. For $\bar{s}$ we merely take the rigid-body solution

$$\bar{\mathbf{U}} = \mathbf{D} + \mathbf{R} \times (\mathbf{r} + x\mathbf{k}), \quad \bar{\mathbf{F}} = \mathbf{0}, \quad \bar{Z} = \bar{Z}_* = \bar{M}_x = \bar{E}_y = 0, \quad (67)$$

where $\mathbf{D}$ and $\mathbf{R}$ are constant displacement and rotation vectors. Since s comprises exponentially decaying solutions whereas the components of $\bar{s}$ have, at most, algebraic growth, $\lim_{x\to\infty}[s,\bar{s}]_x = 0$, so that at $a=0$, Eq. (66) reduces to

$$[s,\bar{s}]_0 = -\mathbf{D}\cdot \int_0^{2\pi} \hat{\mathbf{F}}(y)dy - \mathbf{R}\cdot \int_0^{2\pi} [\mathbf{r}(y) \times \hat{\mathbf{F}}(y)]dy = 0, \qquad (68)$$

where, here and henceforth, a hat $(\hat{\ })$ is used to denote a prescribed variable. As $\mathbf{D}$ and $\mathbf{R}$ are arbitrary, Eq. (68) yields the familiar requirement that, *modulo a rigid body movement and in the absence of surface loads, the total edge force and moment must vanish for exponential decaying solutions to exist.*

To determine the decaying solutions themselves, we introduce the scaled EZ ($^-$) and MB ($\tilde{\ }$) variables defined in Eqs. (22) and (39) and take $\hat{\mathbf{F}}$ in the form

$$\hat{\mathbf{F}} = \hat{F}(y)\mathbf{k} + \varepsilon^{1/2}\hat{S}(y)\mathbf{t}(y), \qquad (69)$$

where $\int_0^{2\pi}\hat{\mathbf{F}}dy = \int_0^{2\pi}\mathbf{r}\times\hat{\mathbf{F}}dy = \mathbf{0}$ and where the normalizing stress measure σ introduced at the beginning of Sec. 3 is chosen so that

$$\max_{0 \leqslant y \leqslant 2\pi} \{|\hat{F}(y)|, |\hat{S}(y)|, |\hat{Z}(y)|, |\hat{M}_x(y)|\} = 1. \qquad (70)$$

Note that Eq. (A14) implies that $\bar{S} = O(\varepsilon^{3/2})$ and $\tilde{S} = O(\varepsilon^{1/2})$ whereas, Eq. (A13) implies that $\bar{F} = O(\varepsilon^{3/2})$ and $\tilde{F} = O(1)$. By Eqs. (A6), the boundary conditions take the form

$$\tilde{F}(0,y;\varepsilon) = \hat{F}(y) + O(\varepsilon^{3/2}) \Rightarrow \tilde{F}(0,y;0) = \hat{F}(y), \qquad (71)$$

$$\tilde{S}(0,y;\varepsilon) = \hat{S}(y) + O(\varepsilon) \Rightarrow \tilde{S}(0,y;0) = \hat{S}(y), \qquad (72)$$

$$\rho(y)\bar{M}_x^{\#}(0,y;\varepsilon) = \hat{Z}(y) + O(\varepsilon^{1/2}) \Rightarrow \rho(y)\bar{M}_x^{\#}(0,y;0) = \hat{Z}(y), \qquad (73)$$

$$\bar{M}_x(0,y;\varepsilon) + \tilde{M}_x(0,y;\varepsilon) = \hat{M}_x(y) \Rightarrow \bar{M}_x(0,y;0) = \hat{M}_x(y) - \tilde{M}_x(0,y;0). \qquad (74)$$

Thus, the prescribed vector stress resultant (69) determines the membrane-inextensional bending (MB) solution to lowest order.

We now determine the MB solution explicitly. Setting $\varepsilon=0$ and $\tilde{C}_- = \tilde{C}$, we have, by Eqs. (54), (56), (71), and (A5)

$$\tilde{F}(0,y;0) = \tilde{N}_x(0,y;0) = \sum_1^\infty \mathcal{R}\tilde{C}_n\Omega_n(y) = \hat{F}(y), \qquad (75)$$

where from the orthogonality condition Eq. (59)

$$\mathcal{R}\tilde{C}_n = \int_0^{2\pi} \hat{F}(y)\Omega_n(y)dy \equiv \hat{F}_n. \qquad (76)$$

Next, from Eqs. (40), (57), (75), and the second part of (A5), the condition that $\tilde{N}$ must approach zero as $\tilde{x}\to\infty$ is

$$\tilde{S}(0,y;0) = \tilde{N}(0,y;0) = \sum_1^\infty \mathcal{R}(\tilde{\mu}\tilde{C}_n)\mathcal{M}[\Omega_n^\bullet(y)]/\lambda_n = \hat{S}(y). \qquad (77)$$

To obtain from Eq. (77) a second relation for $\tilde{C}_n$ to complement Eq. (76), note by Eqs. (57) and (59), and an integration by parts that

$$-\int_0^{2\pi} (\mathcal{M}\Omega_m^\bullet)\Omega_n^\bullet dy = \int_0^{2\pi} (\mathcal{M}\Omega_m^\bullet)^\bullet\Omega_n dy = \lambda_m^2\delta_{mn}. \qquad (78)$$

Thus, Eq. (77) yields

$$-\mathcal{R}(\tilde{\mu}\tilde{C}_n) = (1/\lambda_n)\int_0^{2\pi} \hat{S}(y)\Omega_n^\bullet(y)dy \equiv \hat{S}_n. \qquad (79)$$

It now follows from Eqs. (76) and (79) that, with $\tilde{\mu} = \tilde{\alpha} + i\tilde{\beta}$

$$\tilde{C}_n = \hat{F}_n + i\left[\frac{\tilde{\alpha}\hat{F}_n + \hat{S}_n}{\tilde{\beta}} \right] \qquad (80)$$

and, from the second part of Eq. (54)

$$\sqrt{2}\begin{Bmatrix} \tilde{\alpha} \\ \tilde{\beta} \end{Bmatrix} = \pm \sqrt{\sqrt{A_{2222}A_{2222}^* + C_{2222}^2} \pm C_{2222}}. \qquad (81)$$

Next, we determine the edge-zone (EZ) solution. To satisfy Eqs. (73) and (74), we insert the real part of Eq. (38) into Eq. (25) integrate either once or twice with respect to $\bar{x}$, and discard (nondecaying) functions of integration. With $\bar{C}_- \equiv \bar{C}$, the resulting expressions for $\bar{M}_x$ and $\bar{M}_x^{\#}$ evaluated at $\bar{x}=0$ and inserted into Eqs. (73) and (74), yield

$$\mathcal{R}\bar{C}(y) = \hat{Z}(y) \qquad (82)$$

and

$$-\mathcal{R}[\bar{C}(y)\kappa^{1/2}(y)/\bar{\mu}] = \hat{M}_x(y) - \tilde{M}_x(0,y;0) \equiv \Delta\hat{M}(y). \qquad (83)$$

Because the MB solution can be determined first, $\tilde{M}_x(0,y;0)$—and hence $\Delta\hat{M}(y)$—is *known*. Thus, with $1/\bar{\mu} \equiv \bar{\alpha} + i\bar{\beta}$

$$\bar{C} = \hat{Z}(y) + i\hat{D}(y), \qquad (84)$$

where

$$\hat{D} = \begin{cases} \dfrac{1}{\bar{\beta}}\left[\dfrac{\Delta\hat{M}(y)}{\sqrt{\kappa(y)}} + \bar{\alpha}\hat{Z}(y) \right] & \text{if } \kappa(y) > 0 \\[3ex] \dfrac{1}{\bar{\alpha}}\left[\dfrac{\Delta\hat{M}(y)}{\sqrt{|\kappa(y)|}} - \bar{\beta}\hat{Z}(y) \right] & \text{if } \kappa(y) < 0 \end{cases} \tag{85}$$

and, from Eq. (36)

$$\sqrt{2}\left\{ \begin{matrix} \bar{\alpha} \\ \bar{\beta} \end{matrix} \right\} = \pm \sqrt{\sqrt{A_{1111}A_{1111}^* + C_{1111}^2} \pm C_{1111}^2}. \tag{86}$$

7.2 Case B: End Displacement and Rotation Prescribed.
To guarantee decaying solutions, we need conditions on the prescribed kinematic variables. However, we may deal with these later because the forms we have assumed for the EZ and MB solutions decay automatically, regardless of the boundary conditions.

Let the prescribed kinematic boundary conditions have the form

$$\mathbf{V}(0,y;\varepsilon) = \varepsilon^{1/2}\hat{U}(y)\mathbf{k} + \hat{V}(y)\mathbf{t}(y), \quad E_y(0,y;\varepsilon) = \hat{E}_y(y), \tag{87}$$

$$Z_*(0,y;\varepsilon) = \hat{Z}_*(y)$$

where Z_* is defined in the second part of Eq. (A6) and the normalizing stress measure σ is chosen so that

$$\max_{0 \leqslant y \leqslant 2\pi} \{|\hat{U}(y)|, |\hat{V}(y)|, |\hat{E}_y(y)|, |\hat{Z}_*(y)|\} = 1. \tag{88}$$

By adding the EZ and MB solutions and noting the scalings in Eqs. (22) and (39), we may express the boundary conditions as

$$\tilde{U}(0,y;\varepsilon) = \hat{U}(y) + \mathrm{O}(\varepsilon) \Rightarrow \tilde{U}(0,y;0) = \hat{U}(y), \tag{89}$$

$$\tilde{V}(0,y;\varepsilon) = \hat{V}(y) + \mathrm{O}(\varepsilon^{3/2}) \Rightarrow \tilde{V}(0,y;0) = \hat{V}(y), \tag{90}$$

$$\rho(y)\bar{E}_y^{\#}(0,y;\varepsilon) = -\hat{Z}_*(y) + \mathrm{O}(\varepsilon^{1/2}) \Rightarrow \rho(y)\bar{E}_y^{\#}(0,y;0) = -\hat{Z}_*(y), \tag{91}$$

$$\bar{E}_y(0,y;\varepsilon) + \tilde{E}_y(0,y;\varepsilon) = \hat{E}_y(y) \Rightarrow \bar{E}_y(0,y;0) = \hat{E}_y(y) - \tilde{E}_y(0,y;0), \tag{92}$$

where the last two boundary conditions are the duals of Eqs. (73) and (74). *Note that the prescribed axial and tangential end displacements determine the membrane-inextensional bending (MB) solution to lowest order.*

Because in the second and third part of Eq. (18), and Eq. (A5), the static-geometric duality (11) imply that

$$F_* = K_y + \varepsilon(\rho E_y)^{\bullet\bullet} = [\mathcal{M}(V)]^{\bullet} \tag{93}$$

and

$$-S_* = -K + \varepsilon\{[\rho(-E_y' + 2E^{\bullet})]^{\bullet} + (3/2)\kappa E\} = \mathcal{M}(U^{\bullet}), \tag{94}$$

we may replace Eqs. (89) and (90)—assuming sufficient smoothness—by the duals of Eqs. (71) and (72) namely,

$$\tilde{F}_*(0,y;\varepsilon) = \hat{F}_*(y) + \mathrm{O}(\varepsilon^{3/2}) \Rightarrow \tilde{F}_*(0,y;0) = \hat{F}_*(y) \tag{95}$$

and

$$\tilde{S}_*(0,y;\varepsilon) = \hat{S}_*(y) + \mathrm{O}(\varepsilon) \Rightarrow \tilde{S}_*(0,y;0) = \hat{S}_*(y). \tag{96}$$

Note that the dual of Eq. (69) is

$$\hat{\mathbf{F}}_* = \hat{F}_*(y)\mathbf{k} - \varepsilon^{1/2}\hat{S}_*(y)\mathbf{t}(y). \tag{97}$$

To satisfy Eqs. (95) and (96), we use the duals of Eqs. (75) and (77) to conclude that

$$\tilde{F}_*(\tilde{x},y;0) = \sum_1^{\infty} I[\tilde{C}_n^* \exp(-\lambda_n\tilde{\mu}\tilde{x})]\Omega_n(y) \tag{98}$$

and

$$\tilde{S}_*(\tilde{x},y;0) = -\sum_1^{\infty} I[\tilde{\mu}\tilde{C}_n^* \exp(-\lambda_n\tilde{\mu}\tilde{x})]\mathcal{M}[\Omega_n^{\bullet}(y)]/\lambda_n, \tag{99}$$

where $\tilde{C}_n^* \equiv \sqrt{A_{2222}/A_{2222}^*}\,\tilde{C}_n$. The duals of Eqs. (76) and (79), namely,

$$I\tilde{C}_n^* = \int_0^{2\pi} \hat{F}_*(y)\Omega_n(y)dy \equiv \hat{F}_n^* \tag{100}$$

and

$$I(\tilde{\mu}\tilde{C}_n^*) = (1/\lambda_n)\int_0^{2\pi} \hat{S}_*(y)\Omega_n^{\bullet}(y)dy \equiv \hat{S}_n^*, \tag{101}$$

yield the dual of Eq. (80)

$$\tilde{C}_n^* = \frac{\hat{S}_n^* - \tilde{\alpha}\hat{F}_n^*}{\tilde{\beta}} + i\hat{F}_n^*, \tag{102}$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ are given in terms of the elastic constants by Eq. (81). Note that all dual equations follow from the dualities

$$F{:}F_*, \quad S{:}-S_*, \quad i\tilde{C}_n{:}\tilde{C}_n^*. \tag{103}$$

It is easy to verify that if we set $\bar{C}_* = \sqrt{A_{1111}/A_{1111}^*}\,\bar{C}$ and use the dualities

$$Z{:}Z_*, \quad M_y{:}-E_y, \quad i\bar{C}{:}\bar{C}_*, \tag{104}$$

then the kinematic boundary conditions (91) and (92) yield

$$\bar{C}_* = -\hat{D}_*(y) + i\hat{Z}_*(y), \tag{105}$$

where

$$\hat{D}_* = \begin{cases} \dfrac{1}{\bar{\beta}}\left[\bar{\alpha}\hat{Z}_*(y) - \dfrac{\Delta\hat{E}_y(y)}{\sqrt{\kappa(y)}} \right] & \text{if } \kappa(y) > 0 \\[3ex] -\dfrac{1}{\bar{\alpha}}\left[\bar{\beta}\hat{Z}_*(y) + \dfrac{\Delta\hat{E}_y(y)}{\sqrt{|\kappa(y)|}} \right] & \text{if } \kappa(y) < 0 \end{cases} \tag{106}$$

and where $\bar{\alpha}$ and $\bar{\beta}$ are given by Eq. (86). (Note that $\bar{\alpha}$ and $\bar{\beta}$ are their own duals.)

8 Decay Conditions on Kinematic End Data for a Semi-Infinite Tube

To develop these, we consider $[\hat{s} - s^{\mathrm{SV}}, s^C]_x$, following earlier work in [2,12–14], Here, $\hat{s}$ is any solution of the field equations satisfying the prescribed kinematic boundary conditions and s^C is one of a set of six *canonical* solutions such that $\underline{V}^C(0,y;\varepsilon) = \underline{0}$. We denote the associated Saint-Venant solutions by s^{SV} and s^C, respectively. Finally, we take

$$\bar{\bar{\mathbf{W}}}^C(0) = \bar{\bar{\mathbf{\Phi}}}^C(0) = \mathbf{0}. \tag{107}$$

Thus, by Eqs. (65) and (66) and because $\lim_{x\to\infty}[s - s^{\mathrm{SV}}] = 0$, we have

$$[\hat{s} - s^{\mathrm{SV}}, s^C]_0 = \int_0^{2\pi} [\underline{F}^C \cdot (\underline{V}^{\mathrm{SV}} - \hat{\underline{V}})]_0 dy = 0. \tag{108}$$

But, by (A82)

$$\int_0^{2\pi} (\underline{F}^C \cdot \underline{V}^{\mathrm{SV}})_0 dy = \mathbf{T}^C(0)\bullet[\mathbf{W}_I(0) + \varepsilon\mathbf{w}(0)] + \mathbf{M}^C(0)\bullet[\boldsymbol{\Phi}_I$$

$$+ \varepsilon\boldsymbol{\phi}(0)] + \mathbf{T}^{\mathrm{SV}}(0)\bullet\int_0^{2\pi}(\underline{A}^T \cdot \underline{F}^C)_0 dy$$

$$+ \mathbf{M}^{\mathrm{SV}}(0)\bullet\int_0^{2\pi}(\underline{B}^T \cdot \underline{F}^C)_0 dy. \tag{109}$$

By Eqs. (107), (A85), and (A86), the last line of Eq. (109) vanishes. Furthermore, linearity implies that the canonical solution for $\underline{F}$ can be put in the form

$$\underline{F}^C = \mathbf{T}^C(0) \cdot \underline{\underline{F}}^C_{\mathbf{T}}(x,y;\varepsilon) + \mathbf{M}^C(0) \cdot \underline{\underline{F}}^C_{\mathbf{M}}(x,y;\varepsilon). \tag{110}$$

Thus, because $\mathbf{T}^C(0)$ and $\mathbf{M}^C(0)$ can be prescribed independently—by Eq. (21) the overall equilibrium equations are always satisfied—Eqs. (108) and (109) yield the desired kinematic boundary conditions for the *exact* beamlike Eqs. (21), (A72), and (A76)

$$\mathbf{W}_I(0) + \varepsilon\mathbf{w}(0) = \int_0^{2\pi} \underline{\underline{F}}^C_{\mathbf{T}}(0,y;\varepsilon) \cdot \underline{\hat{V}}(y)dy \tag{111}$$

and

$$\boldsymbol{\Phi}_I + \varepsilon\boldsymbol{\phi}(0) = \int_0^{2\pi} \underline{\underline{F}}^C_{\mathbf{M}}(0,y;\varepsilon) \cdot \underline{\hat{V}}(y)dy. \tag{112}$$

In particular, the vanishing of the beamlike solutions associated with $\hat{s}$, which in the present case mean that the right sides of Eqs. (111) and (112) vanish and $\mathbf{T}^{\mathrm{SV}}(0)=\mathbf{M}^{\mathrm{SV}}(0)=\mathbf{0}$, implies that the prescribed data at $x=0$ and $x=\infty$ produce decaying solutions only.

Without going into detail, we note that s^C may be constructed by setting $s^C=s^{\mathrm{SV}}_C+s^R_C$ (where "R" stands for "residual,") subject to the boundary conditions coming from Eqs. (A6), (A7), and (A20), namely,

$$F^R_*(0,y;\varepsilon) = -\overset{0}{F}^{\mathrm{SV}}_*(y), \quad S^R_*(0,y;\varepsilon) = -\varepsilon\overset{0}{\sigma}^{\mathrm{SV}}_*(y) \tag{113}$$

and

$$Z^R_*(0,y;\varepsilon) = -Z^{\mathrm{SV}}_*(0,y), \quad E^R_y(0,y;\varepsilon) = -E^{\mathrm{SV}}_y(0,y), \tag{114}$$

where the terms on the right of Eqs. (113) and (114) are the Saint-Venant solutions associated with $\mathbf{T}^C(0)$ and $\mathbf{M}^C(x)$. One then sets $s^R_C=\bar{s}^R_C+\tilde{s}^R_C$ and proceeds as in Case B of Sec. 7.

9 End-loaded Tube of Length Rl

The exact, beamlike part of the solution is governed by Eqs. (21), (A55), (A59), (A60), (A72), and (A76). To compute the EZ and MB stress solutions, we need only modify the MB solution coming from Eq. (56) by setting

$$\tilde{N}_x = \sum_1^\infty \mathcal{R}[\tilde{C}^c_n\cosh(\lambda_n\tilde{\mu}\tilde{x}) + (\tilde{C}^s_n/\tilde{\mu})\sinh(\lambda_n\tilde{\mu}\tilde{x})]\Omega_n(y), \tag{115}$$

where the superscripts "c" and "s" identify the associated hyperbolic functions cosh and sinh. In place of Eq. (76), we now have, with $\hat{F}^0$ denoting the value of F prescribed at $x=0$, etc.,

$$\mathcal{R}\tilde{C}^c_n = \int_0^{2\pi} \hat{F}^0(y)\Omega_n(y)dy \equiv \hat{F}^0_n \tag{116}$$

and in place of Eq. (79)

$$\mathcal{R}\tilde{C}^c_n = (1/\lambda_n)\int_0^{2\pi} \hat{S}^0(y)\Omega^\bullet_n(y)dy \equiv \hat{S}^0_n, \tag{117}$$

At the right end of the tube, the corresponding relations are

$$\mathcal{R}[\tilde{C}^c_n\cosh(\lambda_n\tilde{\mu}\varepsilon^{1/2}l) + (\tilde{C}^s_{n'}/\tilde{\mu})\sinh(\lambda_n\tilde{\mu}\varepsilon^{1/2}l)]$$

$$= \int_0^{2\pi} \hat{F}^l(y)\Omega_n(y)dy \equiv \hat{F}^l_n. \tag{118}$$

and

$$\mathcal{R}[\tilde{\mu}\tilde{C}^c_n\sinh(\lambda_n\tilde{\mu}\varepsilon^{1/2}l) + \tilde{C}^s_n\cosh(\lambda_n\tilde{\mu}\varepsilon^{1/2}l)]$$

$$= (1/\lambda_n)\int_0^{2\pi} \hat{P}^l(y)\Omega^\bullet_n(y)dy \equiv \hat{P}^l_n. \tag{119}$$

To solve these equations, we set $\tilde{\mu}=\tilde{\alpha}+i\tilde{\beta}$, as before, and use the identities

$$\cosh[\lambda_n(\tilde{\alpha} + i\tilde{\beta})\varepsilon^{1/2}l] = \cosh(\lambda_n\tilde{\alpha}\varepsilon^{1/2}l)\cos(\lambda_n\tilde{\beta}\varepsilon^{1/2}l)$$

$$+ i\sinh(\lambda_n\tilde{\alpha}\varepsilon^{1/2}l)\sin(\lambda_n\tilde{\beta}\varepsilon^{1/2}l) \equiv J_n + iK_n \tag{120}$$

and

$$\sinh[\lambda_n(\tilde{\alpha} + i\tilde{\beta})\varepsilon^{1/2}l] = \sinh(\lambda_n\tilde{\alpha}\varepsilon^{1/2}l)\cos(\lambda_n\tilde{\beta}\varepsilon^{1/2}l)$$

$$+ i\cosh(\lambda_n\tilde{\alpha}\varepsilon^{1/2}l)\sin(\lambda_n\tilde{\beta}\varepsilon^{1/2}l) \equiv H_n + iI_n. \tag{121}$$

Then, with

$$\tilde{C}^c_n = \hat{F}^0_n + i\hat{A}^c_n \quad \text{and} \quad \tilde{C}^s_n = \hat{S}^0_n + i\hat{A}^s_n, \tag{122}$$

it may be verified that Eqs. (116)–(119) lead to the simultaneous equations

$$(\tilde{\alpha}I_n + \tilde{\beta}H_n)\hat{A}^c_n + K_n\hat{A}^s_n = (\tilde{\alpha}H_n - \tilde{\beta}I_n)\hat{F}^0_n + J_n\hat{S}^0_n - \hat{S}^l_n \tag{123}$$

and

$$K_n\hat{A}^c_n + \left[\frac{\tilde{\alpha}I_n - \tilde{\beta}H_n}{\tilde{\alpha}^2 + \tilde{\beta}^2}\right]\hat{A}^s_n = J_n\hat{F}^0_n - \hat{F}^l_n + \left[\frac{\tilde{\alpha}H_n - \tilde{\beta}I_n}{\tilde{\alpha}^2 + \tilde{\beta}^2}\right]\hat{S}^0_n. \tag{124}$$

These may be solved in specific cases with no difficulty.

10 Conclusions

We have shown how to decompose the linear Sanders-Koiter equations applied to a fully anisotropic (21 constant) tube of arbitrary cross section into an *exact* beamlike (or Saint-Venant) solution plus edge-zone and membrane-bending solutions given explicitly to within a relative error of $O(\varepsilon^{1/2})$. For the two extreme case of (A) edge loads only and (B) edge kinematic constraints only, we give conditions on the end data that insure decaying solutions (i.e., beamlike solutions vanish identically).

Aside from a rigid body movement, the beamlike solution turns out to have a relatively large *inextensional* component, unless the combined dimensionless material constant Γ^* given in Eq. (A37) vanishes.

Appendix: Elaboration of Saint-Venant (Beamlike) Solutions

In the Sanders-Koiter theory, the virtual work identity for a tube of arbitrary cross section extending from $x=a$ to $x=b$, under edge loads only, has the dimensionless form

$$\mathrm{EVW}|^b_a = \varepsilon\int_a^b \mathrm{IVW}dx, \tag{A1}$$

where

$$\text{EVW} = \int_0^{2\pi} \{N_x U + [N + (3/2)\varepsilon\kappa M]V$$

$$+ \varepsilon[(M_x' + 2M^\bullet)W - M_x W']\}dy \qquad (A2)$$

and

$$\text{IVW} = \int_0^{2\pi} (N_x E_x + 2NE + N_y E_y + M_x K_x + 2MK + M_y K_y)dy. \qquad (A3)$$

Here, the stress resultants and couples satisfy the equilibrium Eqs. (5)–(7) whereas the displacements are arbitrary (but sufficiently smooth); EVW and IVW are mnemonics for external and internal virtual work.

To put these equations in a form that better reflects the complementary aspects of the Saint-Venant (SV), the edge zone (EZ), and the membrane-inextensional (MB) solutions, we first use the strain-displacement relations (17) to write

$$W = \rho(\varepsilon E_y - V^\bullet) \quad \text{and} \quad W' = \rho[\varepsilon(E_y' - 2E^\bullet) + U^{\bullet\bullet}], \qquad (A4)$$

then introduce these expressions into Eq. (A2), integrate by parts, and finally set

$$F \equiv N_x - \varepsilon(\rho M_x)^{\bullet\bullet}, \quad S \equiv N + \varepsilon\{[\rho(M_x' + 2M^\bullet)]^\bullet + (3/2)\kappa M\} \qquad (A5)$$

$$Z \equiv \varepsilon^{1/2}\rho(M_x' + 2M^\bullet), \quad Z_* \equiv \varepsilon^{1/2}\rho(2E^\bullet - E_y'). \qquad (A6)$$

In Eq. (A6), Z is the *effective* Kirchhoff-transverse shear stress resultant at a section $x=$ constant (the true shear stress resultant being $M_x' + M^\bullet$), and Z_* is the kinematic dual of Z. Thus, Eq. (A2) can be rewritten as

$$\text{EVW} = \int_0^{2\pi} \underline{F} \cdot \underline{V}dy, \qquad (A7)$$

where $\underline{F}$ and $\underline{V}$ are defined in Eq. (63).

On any cross section of the tube, the dimensionless force $\mathbf{T}$ and moment $\mathbf{M}$ (acting about the centroid $\mathbf{r}=\mathbf{0}$) are given by Eq. (19). Using $\mathbf{r}^\bullet = \mathbf{t} = \mathbf{n} \times \mathbf{k}$, $\mathbf{t}^\bullet = \kappa\mathbf{n}$, the periodicity of all functions defined over the tube's cross section, and integration by parts, we have

$$\int_0^{2\pi} (\rho M_x)^{\bullet\bullet}\mathbf{k}dy = \mathbf{0}, \qquad (A8)$$

$$\int_0^{2\pi} \mathbf{r} \times (\rho M_x)^{\bullet\bullet}\mathbf{k}dy = \int_0^{2\pi} \rho M_x \mathbf{r}^{\bullet\bullet} \times \mathbf{k}dy = \int_0^{2\pi} M_x\mathbf{t}dy, \qquad (A9)$$

and

$$\int_0^{2\pi} (\rho Z\mathbf{t})^\bullet dy = \mathbf{0}, \quad \int_0^{2\pi} \mathbf{r} \times (\rho Z\mathbf{t})^\bullet dy = \mathbf{0}. \qquad (A10)$$

Thus, we may write Eq. (19) in the formally simpler form

$$\mathbf{T}(0) = \int_0^{2\pi} \mathbf{F}(0,y;\varepsilon)dy \quad \text{and} \quad \mathbf{M}(x) = \int_0^{2\pi} \mathbf{r} \times \mathbf{F}(x,y;\varepsilon)dy. \qquad (A11)$$

The differential operator $\mathcal{M}$, defined in Eq. (51), appears in what follows and has the following properties:

$$\mathcal{M}(\mathbf{t}) = \mathbf{0}, \quad \mathcal{M}(\mathbf{r}\bullet\mathbf{n}) = -1, \quad \int_0^{2\pi} \mathcal{M}(f)\mathbf{t}dy = \int_0^{2\pi} f\mathcal{M}(\mathbf{t})dy = \mathbf{0}. \qquad (A12)$$

Note that if we introduce the definitions (A5) into the equilibrium Eqs. (5) and (6) and use Eq. (7) to eliminate N_y, the two resulting equations take the forms

$$F' + S^\bullet - 2\varepsilon[\mathcal{M}(M)]^\bullet = 0 \qquad (A13)$$

and

$$S' + \varepsilon\mathcal{M}(M_y^\bullet) = 0. \qquad (A14)$$

Analogous expressions can be written down for the compatibility equations.

We now modify, extend, and correct parts of the analysis in [3] concerning SV solutions by assuming that all stress resultants and couples are, at most, *linear in x*. In fact, as we shall see, *all unknowns are also, at most, linear in ε*. Indeed, ignoring terms in the stress-strain relations that are of the same order as the inherent, unavoidable errors implies that *the extensional strains and the stress couples are independent of ε*. Thus, we assume that

$$F = \overset{0}{F}(y) + x\overset{1}{F}(y), \quad S = \overset{0}{S}(y) + \varepsilon[\overset{0}{\sigma}(y) + x\overset{1}{\sigma}(y)],$$
$$(A15)$$
$$M_{\alpha\beta} = \overset{0}{M}_{\alpha\beta}(y) + x\overset{1}{M}_{\alpha\beta}(y).$$

(Here, and until further notice, we drop the tag SV to avoid a cluttered notation.) From Eqs. (A13) and (A14), these assumptions require that

$$\overset{1}{F} = -\overset{0}{S}^\bullet, \quad \overset{0}{\sigma}^\bullet = 2[\mathcal{M}(\overset{0}{M})]^\bullet \Rightarrow \overset{0}{\sigma} = c + 2\mathcal{M}(\overset{0}{M}), \quad c = \text{constant} \qquad (A16)$$

$$\overset{1}{\sigma}^\bullet = 2[\mathcal{M}(\overset{1}{M})]^\bullet \Rightarrow \overset{1}{\sigma} = c + 2\mathcal{M}(\overset{1}{M}), \quad c = \text{constant} \qquad (A17)$$

$$\overset{1}{\sigma} = -\mathcal{M}(\overset{0}{M_y}^\bullet), \qquad (A18)$$

and

$$\mathcal{M}(\overset{1}{M_y}^\bullet) = 0. \qquad (A19)$$

We further assume that the kinematic duals of F and S have the form

$$F_* = \overset{0}{F_*}(y), \quad S_* = \varepsilon[\overset{0}{\sigma_*}(y) + x\overset{1}{\sigma_*}(y)]. \qquad (A20)$$

These assumptions imply the following duals of Eqs. (A16)–(A19)

$$-\overset{0}{\sigma_*}^\bullet = 2[\mathcal{M}(\overset{0}{E})]^* \Rightarrow -\overset{0}{\sigma_*} = \overset{0}{c_*} + 2\mathcal{M}(\overset{0}{E}) \qquad (A21)$$

$$-\overset{1}{\sigma_*}^\bullet = 2[\mathcal{M}(\overset{1}{E})]^\bullet \Rightarrow -\overset{1}{\sigma_*} = \overset{1}{c_*} + 2\mathcal{M}(\overset{1}{E}) \qquad (A22)$$

$$-\overset{1}{\sigma_*} = \mathcal{M}(\overset{0}{E_x}^\bullet), \qquad (A23)$$

and

$$\mathcal{M}(\overset{1}{E_x}^\bullet) = 0. \qquad (A24)$$

The general solution of Eq. (A19) is

$$\overset{1}{M_y} = \overset{1}{a} + \overset{1}{\mathbf{a}}\bullet\mathbf{r}, \qquad (A25)$$

where $\overset{1}{a}$ and $\overset{1}{\mathbf{a}}$ are arbitrary scalar and vector constants. On the other hand, because we can neglect terms of relative $O(\varepsilon)$ in the

stress-strain relations (14), Eqs. (A5), (A16), and (A20) allow us to write

$$\overset{1}{M_y} = -C^*_{2222}\overset{\circ}{S}^\bullet. \tag{A26}$$

Thus, Eqs. (A25) and (A26) are consistent if and only if $a=0$ and $\overset{0}{S}$ has the periodic form

$$\overset{0}{S} = b - \mathbf{b}\cdot\mathbf{s}, \tag{A27}$$

where b and $\mathbf{b}$ are unknown constants and, as in Eq. (48) of [3]

$$\mathbf{s} = \int_0^y \mathbf{r}(\eta)d\eta + (1/2\pi)\int_0^{2\pi} y\mathbf{r}(y)dy. \tag{A28}$$

(Note that $\int_0^{2\pi} \mathbf{s}dy = \mathbf{0}$.) Because $\overset{1}{E_x} = -A_{2222}\overset{0}{S}^\bullet = A_{2222}\mathbf{b}\cdot\mathbf{r}$, to within a permissible relative error of $O(\varepsilon)$, Eq. (A24) is also satisfied.

Next, the consistency of Eqs. (A17) and (A18) requires that

$$\mathcal{M}(2\overset{1}{M} + \overset{0}{M}_y^\bullet) = -\overset{1}{c} \Rightarrow 2\overset{1}{M} + \overset{0}{M}_y^\bullet = \overset{1}{c}(\mathbf{r}\cdot\mathbf{n}) + \mathbf{c}\cdot\mathbf{t}, \tag{A29}$$

where $\mathbf{c}$ is a constant vector and where we have used it in the second part of Eq. (A12). But, by Eqs. (14), (A16), (A15), (A20), and (A27)

$$2\overset{1}{M} + \overset{0}{M}_y^\bullet = 4\bar{C}^*_{1222}\mathbf{b}\cdot\mathbf{r} + (A^*_{2222}\overset{0}{F}_* + C^*_{2222}\overset{0}{F})^\bullet, \tag{A30}$$

where, as in [3], $2\bar{C}^*_{1222} \equiv C^*_{1222} + C^*_{2212}$ and, again, $O(\varepsilon)$ terms have been neglected. Because the integral of the right side of Eq. (A30) from $y=0$ to 2π vanishes, $\overset{1}{c}=0$. Moreover, because the right side of Eq. (A30) must be periodic and of the form $\mathbf{c}\cdot\mathbf{t}$, $\overset{0}{F}$ and its dual must have the forms

$$\overset{0}{F} = d + \mathbf{d}\cdot\mathbf{r} + \mathbf{e}\cdot\mathbf{s}, \tag{A31}$$

$$\overset{0}{F}_* = d^* + \mathbf{d}^*\cdot\mathbf{r} - \mathbf{e}^*\cdot\mathbf{s}, \tag{A32}$$

where

$$C^*_{2222}\mathbf{e} - A^*_{2222}\mathbf{e}^* = -4\bar{C}^*_{1222}\mathbf{b} \tag{A33}$$

and d, $\mathbf{d}$, d^*, and $\mathbf{d}^*$ are unknown constants. However, as the tube can undergo no dislocations,

$$\int_0^{2\pi} \overset{0}{F}_*(y)\mathbf{k}dy = \int_0^{2\pi} \mathbf{r} \times \overset{0}{F}_*(y)\mathbf{k}dy = \mathbf{0}. \tag{A34}$$

See Eqs. (20*) and (22*) of [3]. Because $\int_0^{2\pi}\mathbf{rs}dy = B\mathbf{k}\times\mathbf{1}$, where B is the area swept out by $\mathbf{s}(y)$ as y goes from 0 to 2π and $\mathbf{1}$ is the two-dimensional identity tensor in the cross-sectional plane, Eq. (A34) implies that

$$d^* = 0, \quad \mathbf{d}^* = (B/\pi)\mathbf{H}\cdot\mathbf{e}^* \text{ where } \mathbf{H} \equiv \mathbf{I}^{-1} \times \mathbf{k},$$
$$\mathbf{H}^{-1} = -\mathbf{k} \times \mathbf{I}, \quad \mathbf{H}^T = -\mathbf{k} \times \mathbf{I}^{-1}, \tag{A35}$$

and $\mathbf{I} = \mathbf{I}^T$ is given by Eq. (61) with $\mathbf{I}^{-1}$ its inverse.

The dual of Eq. (A29) is $2\overset{1}{E} - \overset{0}{E}_x = \mathbf{c}^*\cdot\mathbf{t}$ which, with terms of relative $O(\varepsilon)$ ignored, leads to a second linear relation between $\mathbf{e}$ and $\mathbf{e}^*$

$$A_{2222}\mathbf{e} + C_{2222}\mathbf{e}^* = -4A_{1222}\mathbf{b}, \tag{A36}$$

where we have used the symmetry $A_{2212} = A_{1222}$. The solution of Eqs. (A33) and (A36) is

$$\{\mathbf{e}, \mathbf{e}^*\} = -\frac{4\{A_{1222}A^*_{2222} + C_{2222}\bar{C}^*_{1222}, A_{1222}C^*_{2222} - A_{2222}\bar{C}^*_{1222}\}}{A_{2222}A^*_{2222} + C_{2222}C^*_{2222}}\mathbf{b}$$
$$\equiv \{\Gamma, \Gamma_*\}\mathbf{b}, \tag{A37}$$

in the notation of Eq. (53) of [3]. (Things obviously simplify tremendously if $\Gamma = \Gamma_* = 0$.)

If we substitute the second part of Eq. (64), the first and second parts of (A15), (A27), and Eqs. (A31) and (A37) into Eq. (A11), note that $2A\bullet\mathbf{k} = \mathbf{r}\times\mathbf{t} = -(\mathbf{r}\cdot\mathbf{n})\mathbf{k}$, where A is the area enclosed by the cross section of the tube, and take $\overset{0}{c} = -(1/A)\int_0^{2\pi}\overset{0}{M}dy$ in Eq. (A16) to avoid any ε dependence in Eq. (A38) and (A39) below, we find that

$$\mathbf{T}(0) = 2\pi d\mathbf{k} + \pi\mathbf{I}\cdot\mathbf{b} \tag{A38}$$

and

$$\mathbf{M}(x) = (2Ab + \pi\mathbf{m}\cdot\mathbf{b})\mathbf{k} + B\Gamma\mathbf{b} - \pi\mathbf{k} \times \mathbf{I}\cdot(\mathbf{d} + x\mathbf{b}). \tag{A39}$$

Thus,

$$d = (1/2\pi)T, \quad b = (1/2A)[\bar{M} - \mathbf{m}\cdot\mathbf{I}^{-1}\cdot\mathbf{T}(0)], \quad \mathbf{b} = (1/\pi)\mathbf{I}^{-1}\cdot\mathbf{T}(0), \tag{A40}$$

and

$$\mathbf{d} + x\mathbf{b} = (1/\pi)\mathbf{H}\cdot[\mathbf{M}(x) - (B\Gamma/\pi)\mathbf{I}^{-1}\cdot\mathbf{T}(0)]. \tag{A41}$$

where $T \equiv \mathbf{k}\cdot\mathbf{T}(0)$, $\bar{M} \equiv \mathbf{k}\cdot\mathbf{M}(0) = \mathbf{k}\cdot\mathbf{M}(x)$, and, as in Eq. (63) of [3],

$$\mathbf{m} = (1/\pi)\int_0^{2\pi} \mathbf{s}(y)\mathbf{r}(y)\cdot\mathbf{n}(y)dy. \tag{A42}$$

From Eq. (A5), $N_x = F + O(\varepsilon)$ and $N = S + O(\varepsilon)$ for SV solutions, so that in view of Eqs. (A15), (A16), and (A20), the strain-stress relations (13) for SV-solutions may be given the simplified, explicit forms

$$\begin{bmatrix} E_x \\ E \\ E_y \end{bmatrix} = \begin{bmatrix} A_{2222} & -2A_{2212} & -C_{2222} \\ -A_{1222} & 2A_{1212} & C_{1222} \\ A_{1122} & -2A_{1112} & -C_{1122} \end{bmatrix} \begin{bmatrix} \overset{0}{F}(y) - x\overset{0}{S}^\bullet(y) \\ \overset{0}{S}(y) \\ \overset{0}{F}_*(y) \end{bmatrix}. \tag{A43}$$

By Eqs. (A27), (A31), (A32), (A35), and (A36), these expressions become

$$E_x = A_{2222}[d + \mathbf{r}\cdot(\mathbf{d} + x\mathbf{b})] - 2A_{1222}(b + \mathbf{s}\cdot\mathbf{b}) - (B\Gamma^*/\pi)C_{2222}\mathbf{r}\cdot\mathbf{H}\cdot\mathbf{b} \tag{A44}$$

$$E = -A_{1222}[d + \mathbf{r}\cdot(\mathbf{d} + x\mathbf{b}) + \Gamma\mathbf{s}\cdot\mathbf{b}] + 2A_{1212}(b - \mathbf{s}\cdot\mathbf{b}) - \pi\Gamma^*C_{1222}\mathbf{g}\cdot\mathbf{I}\cdot\mathbf{b} \tag{A45}$$

and

$$E_y = A_{1122}[d + \mathbf{r}\cdot(\mathbf{d} + x\mathbf{b}) + \Gamma\mathbf{s}\cdot\mathbf{b}] - 2A_{1211}(b - \mathbf{s}\cdot\mathbf{b}) - \pi\Gamma^*C_{1122}\mathbf{g}\cdot\mathbf{I}\cdot\mathbf{b}$$
$$= [(1/2\pi)A_{1122}\mathbf{k} + (\Gamma^*C_{1122} + \Gamma A_{1122})\mathbf{g} + 2A_{1112}\mathbf{h}]\cdot\mathbf{T}(0)$$
$$+ [-(1/A)A_{1112}\mathbf{k} + (1/\pi)A_{1112}\mathbf{r}\cdot\mathbf{H}]\cdot\mathbf{M}(x)$$
$$\equiv \mathbf{A}^*(y)\cdot\mathbf{T}(0) + \mathbf{B}^*(y)\cdot\mathbf{M}(x), \tag{A46}$$

where

$$\mathbf{g} \equiv (1/\pi)[\mathbf{s} - (B/\pi)\mathbf{r}\cdot\mathbf{H}]\cdot\mathbf{I}^{-1} \tag{A47}$$

and

$$\mathbf{h} \equiv (\mathbf{s}/\pi + \mathbf{m}/2A)\cdot\mathbf{I}^{-1}. \tag{A48}$$

Three of the other variables that appear in the external virtual work (A7) have the forms

$$M_x = [(1/2\pi)C_{1122}^*\mathbf{k} + (\Gamma C_{1122}^* - \Gamma^* A_{1122}^*)\mathbf{g} + 2C_{1122}^*\mathbf{h}]\bullet\mathbf{T}(0)$$
$$+ [-(1/A)C_{1122}^*\mathbf{k} + (1/\pi)C_{1122}^*\mathbf{r}\bullet\mathbf{H}]\bullet\mathbf{M}(x)$$
$$\equiv \mathbf{A}(y)\bullet\mathbf{T}(0) + \mathbf{B}(y)\bullet\mathbf{M}(x), \tag{A49}$$

$$2E^\bullet - E_y' = -[2(\Gamma A_{1222} + \Gamma^* C_{1222})\mathbf{g}^\bullet + (4A_{1212} + A_{1122})\mathbf{h}^\bullet]\bullet\mathbf{T}(0)$$
$$- (2/\pi)A_{1222}\mathbf{t}\bullet\mathbf{H}\bullet\mathbf{M}(x)$$
$$\equiv \widetilde{\mathbf{A}}^*(y)\bullet\mathbf{T}(0) + \widetilde{\mathbf{B}}^*(y)\bullet\mathbf{M}(x), \tag{A50}$$

$$M_x' + 2M^\bullet = [2(\Gamma C_{1222} - \Gamma^* A_{1222})\mathbf{g}^\bullet + (4C_{1212}^* + C_{1122}^*)\mathbf{h}^\bullet]\bullet\mathbf{T}(0)$$
$$+ (2/\pi)C_{1222}^*\mathbf{t}\bullet\mathbf{H}\bullet\mathbf{M}(x)$$
$$\equiv \widetilde{\mathbf{A}}(y)\bullet\mathbf{T}(0) + \widetilde{\mathbf{B}}(y)\bullet\mathbf{M}(x). \tag{A51}$$

Likewise, note from Eqs. (A15), (A27), (A31), and (A38)–(A41), that the second part of Eq. (64) may be given the form

$$\overset{0}{\mathbf{F}}(x,y;\varepsilon) = [\overset{0}{\mathbf{A}}(y) + \varepsilon\mathbf{A}_1(y)]\bullet\mathbf{T}(0) + [\overset{0}{\mathbf{B}}(y) + \varepsilon\mathbf{B}_1(y)]\bullet\mathbf{M}(x), \tag{A52}$$

where

$$\overset{0}{\mathbf{A}} = (1/2\pi)\mathbf{k}\mathbf{k} + \Gamma\mathbf{k}\mathbf{g} - \mathbf{t}\mathbf{h}, \quad \overset{0}{\mathbf{B}} = (1/\pi)\mathbf{k}\mathbf{r}\bullet\mathbf{H} + (1/2A)\mathbf{t}\mathbf{k} \tag{A53}$$

and

$$\mathbf{A}_1 = \mathbf{t}[C_{1222}^*(\kappa/\pi - 1/A)\mathbf{k} + 2(\Gamma C_{1222}^* - \Gamma^* A_{1222}^*)\mathcal{M}(\mathbf{g})$$
$$- (2\pi/A^2)C_{1222}^*\mathbf{m}\bullet\mathbf{I}^{-1} + 4C_{1212}^*\mathcal{M}(\mathbf{h})],$$
$$\mathbf{B}_1 = \mathbf{t}[(2\pi/A)C_{1212}^*(1/A - \kappa/\pi)\mathbf{k} + (2/\pi)C_{1222}^*\mathbf{H}^T\bullet\mathcal{M}(\mathbf{r})]. \tag{A54}$$

The final ingredient in the expression for the external virtual work (A7) is the dimensionless displacement. To put this into a useful form, we first set

$$\overset{0}{\mathbf{U}} = \overset{0}{U}(y)\mathbf{k} + [\overset{0}{V}(y) - x\overset{0}{U^\bullet}(y)]\mathbf{t}(y) + \rho(y)[\overset{0}{V^\bullet}(y) - x\overset{0}{U^{\bullet\bullet}}(y)]\mathbf{n}(y)$$
$$+ \varepsilon\mathbf{u}(x,y), \tag{A55}$$

where $\overset{0}{U} = c_1 - \mathbf{c}_1\bullet\mathbf{r}$ and $\overset{0}{V}$ are inextensional displacements with $\overset{0}{V}$ chosen so that, by the third part of Eq. (18) and Eq. (A32)

$$[\mathcal{M}(\overset{0}{V})]^\bullet = \mathbf{d}^*\bullet\mathbf{r} - \mathbf{e}^*\bullet\mathbf{s}. \tag{A56}$$

With the aid of Eqs. (A35) and (A40) the 2π-periodic solution of this equation may be verified to be

$$\overset{0}{V} = -c_2\mathbf{r}(y)\bullet\mathbf{n}(y) + \mathbf{c}_2\bullet\mathbf{t}(y) - \Gamma^*\mathbf{t}(y)\bullet\int_0^y \mathbf{n}(\eta)\mathbf{G}(\eta)d\eta\bullet\mathbf{T}(0), \tag{A57}$$

where c_2 and $\mathbf{c}_2$ are constants associated with a rigid body movement, $\mathbf{G}^\bullet = \mathbf{g}$, and, as in Eq. (54) of [3],

$$\mathbf{v} = \int_0^y \mathbf{s}(\eta)d\eta + (1/2\pi)\int_0^{2\pi} y\mathbf{s}(y)dy, \tag{A58}$$

the constant term in the above equation being chosen so that $\int_0^{2\pi}\mathbf{v}dy = \mathbf{0}$.

Thus, the inextensional component of $\mathbf{U}$ has the form

$$\overset{0}{\mathbf{U}} \equiv c_1 + \mathbf{c}_2 + x\mathbf{c}_1 + (c_2\mathbf{k} + \mathbf{k}\times\mathbf{c}_1)\times\mathbf{r} - \Gamma^*\int_0^y \mathbf{n}(\eta)\mathbf{G}(\eta)d\eta\bullet\mathbf{T}(0)$$
$$\equiv \mathbf{W}_I(x) + \mathbf{\Phi}_I\times\mathbf{r} + \mathbf{A}_I\bullet\mathbf{T}(0). \tag{A59}$$

In Eq. (A59), the constants vectors $\mathbf{c}_1$ and $\mathbf{c}_2$ lie in the $\mathbf{t}$-$\mathbf{n}$ plane. and the first two terms in the last line are rigid body movements.

More generally, Ladevèze has shown [13] that it is possible to write $\mathbf{U} - \overset{0}{\mathbf{U}}$ as

$$\mathbf{u} = \mathbf{w}(x) + \boldsymbol{\phi}(x)\times\mathbf{r} + \mathbf{A}(y)\bullet\mathbf{T}(0) + \mathbf{B}(y)\bullet\mathbf{M}(x), \tag{A60}$$

where $\mathbf{w}$ and $\boldsymbol{\phi}$ are a beamlike displacement and rotation, respectively, and $\mathbf{A}$ and $\mathbf{B}$ are material tensors. To find the latter, we first use Eqs. (17) and (A60), and the overall moment equilibrium condition $\mathbf{M}'(x) = -\mathbf{k}\times\mathbf{T}(0)$, to obtain

$$u' = w'(x) + [\mathbf{k}\times\boldsymbol{\phi}'(x)]\bullet\mathbf{r} - [\mathbf{k}\bullet\mathbf{B}(y)\times\mathbf{k}]\bullet\mathbf{T}(0) = E_x, \tag{A61}$$

where $w \equiv \mathbf{k}\bullet\mathbf{w}$. A comparison of the middle expression in Eq. (A61) with the right side of Eq. (A44) yields

$$w' = A_{2222}d - 2A_{1222}b + \mathbf{C}\bullet\mathbf{b} = (1/2\pi)A_{2222}T + [(1/A)A_{1222}\mathbf{m}$$
$$+ (1/\pi)\mathbf{C}]\bullet\mathbf{I}^{-1}\bullet\mathbf{T}(0) - (1/A)A_{1222}M \tag{A62}$$

and, by Eq. (A35),

$$(\mathbf{k}\times\boldsymbol{\phi}')\times\mathbf{k} = \boldsymbol{\phi}' - \phi'\mathbf{k}$$
$$= [A_{2222}(\mathbf{d} + x\mathbf{b}) - C_{2222}\mathbf{d}^* + (1/\pi)\mathbf{C}\bullet\mathbf{b}]\times\mathbf{k}$$
$$= -[(4B/\pi^2)A_{1212}\mathbf{H}^{2T} + \mathbf{k}\times\mathbf{C}]\bullet\mathbf{I}^{-1}\bullet\mathbf{T}(0)$$
$$- (1/\pi)A_{2222}\mathbf{k}\times\mathbf{H}\bullet\mathbf{M}(x), \tag{A63}$$

where $\phi \equiv \mathbf{k}\bullet\boldsymbol{\phi}$ and $\mathbf{C}$ and $\mathbf{C}$ are, respectively, a constant vector and tensor. By the third part of (A40),

$$\mathbf{k}\bullet\mathbf{B} = B_{11}(y)\mathbf{k} - (1/\pi)(2A_{1222}\mathbf{s} + \mathbf{r}\bullet\mathbf{C} + \mathbf{C})\bullet\mathbf{H}. \tag{A64}$$

Still to be determined are $\mathbf{C}$, $\mathbf{C}$, $\mathbf{A}$, $B_{11} \equiv \mathbf{k}\bullet\mathbf{B}\bullet\mathbf{k}$, and $\mathbf{t}\bullet\mathbf{B}$.

Next, inserting Eqs. (A35), (A38), (A39), and (A47) into the third strain-displacement relation $E_y = v^\bullet + kw$, we have

$$A_{1122}[d + (\mathbf{d} + x\mathbf{b})\bullet\mathbf{r} + \Gamma\mathbf{b}\bullet\mathbf{s}] - 2A_{1112}(b - \mathbf{b}\bullet\mathbf{s}) + \pi\Gamma^*C_{1122}\mathbf{b}\bullet\mathbf{I}\bullet\mathbf{g}$$
$$= \mathbf{t}\bullet\{\mathbf{A}^\bullet(y)\bullet(2\pi d\mathbf{k} + \pi\mathbf{I}\bullet\mathbf{b}) + \mathbf{B}^\bullet(y)\bullet[(2Ab + \pi\mathbf{m}\bullet\mathbf{b})\mathbf{k} + B\Gamma\mathbf{b}$$
$$- \pi\mathbf{k}\times\mathbf{I}\bullet(\mathbf{d} + x\mathbf{b})]\}. \tag{A65}$$

Thus,

$$\mathbf{A} = \mathbf{k}\mathbf{k}\bullet\mathbf{A} + (1/2\pi)A_{1122}\mathbf{r}\mathbf{k} + (1/A)A_{1112}\mathbf{r}\mathbf{m}\bullet\mathbf{I}^{-1} - (B/\pi^2)(\Gamma A_{1122}$$
$$+ \Gamma^*C_{1122})\mathbf{t}\mathbf{s}\bullet\mathbf{H}\bullet\mathbf{I}^{-1} + (1/\pi)(A_{1122}\Gamma + 2A_{1112} + \Gamma^*C_{1122})\mathbf{t}\mathbf{v}\bullet\mathbf{I}^{-1}$$
$$+ \mathbf{t}A \tag{A66}$$

and

$$\mathbf{B} = \mathbf{k}[B_{11}(y)\mathbf{k} - (1/\pi)(2A_{1222}\mathbf{s} + \mathbf{r}\bullet\mathbf{C} + \mathbf{C})\bullet\mathbf{H}] - (1/A)A_{1112}\mathbf{r}\mathbf{k}$$
$$+ (1/\pi)A_{1122}\mathbf{t}\mathbf{s}\bullet\mathbf{H} + \mathbf{t}\mathbf{B}, \tag{A67}$$

where $\mathbf{k}\bullet\mathbf{A}$, B_{11}, the constant vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, and the constant tensor $\mathbf{C}$ are yet to be determined. However, the terms arising from $\mathbf{A}$ and $\mathbf{B}$, representing rigid body movements, may be absorbed into the constant $\mathbf{c}_2$ in Eq. (A57) and so ignored.

Finally, we use the second part of Eq. (17), Eq. (A35), the third part of (A40), and Eqs. (A45), (A55), (A59), and (A67) to write

$$v' = \mathbf{t}(y)\bullet\mathbf{w}'(x) + 2A^\bullet(y)\phi'(x) + A_{1122}\mathbf{b}\bullet\mathbf{s} = 2E - u^\bullet, \tag{A68}$$

where, as in Eq. (62) of [3],

$$A(y) = (1/2)\mathbf{k} \cdot \int_0^y \mathbf{r}(\eta) \times \mathbf{t}(\eta) d\eta, \quad A = A(2\pi) \quad \text{(A69)}$$

Substituting Eqs. (A45), (A59), (A66), and (A67) into the right side of Eq. (A68), using Eq. (A41) and the equilibrium condition $\mathbf{M}'(x) = -\mathbf{k} \times \mathbf{T}(0)$, and cancelling some terms, we obtain

$$\mathbf{t} \cdot \mathbf{w}'(x) + 2A^\bullet(y)\phi'(x) + A_{1122}\mathbf{b} \cdot \mathbf{s}$$
$$= -2A_{1222}d - 2\pi(\Gamma A_{1222} + \Gamma^* C_{1222})\mathbf{g} \cdot \mathbf{I} \cdot \mathbf{b}$$
$$+ 4A_{1212}(b - \mathbf{b} \cdot \mathbf{s}) + \mathbf{t} \cdot [(1/\pi)\mathbf{C} \cdot \mathbf{H} \cdot \mathbf{M}(x) - \mathbf{k} \times \phi]$$
$$- \mathbf{k} \cdot \mathbf{A}^\bullet \cdot T(0) - B_{11}^\bullet(y)\bar{M}. \quad \text{(A70)}$$

To determine ϕ, we integrate both sides of Eq. (A70) from $y = 0$ to 2π. Noting Eqs. (A40), and (A47), we find that

$$\phi' = (2\pi/A)(2A_{1212}b - A_{1222}d)$$
$$= -(1/A)A_{1222}T - (2\pi/A^2)A_{1212}\mathbf{m} \cdot \mathbf{I}^{-1} \cdot \mathbf{T}(0)$$
$$+ (2\pi/A^2)A_{1212}\bar{M}. \quad \text{(A71)}$$

This relation combined with Eq. (A63) gives an expression of the following form for the beamlike bending strain:

$$\Omega \equiv \phi' = \mathbf{\Lambda_{MT}} \cdot \mathbf{T}(0) + \mathbf{\Lambda_{MM}} \cdot \mathbf{M}(x), \quad \text{(A72)}$$

where

$$\mathbf{\Lambda_{MM}} = (2\pi/A^2)A_{1212}\mathbf{kk} - (1/\pi)A_{2222}\mathbf{k} \times \mathbf{I}^{-1} \times \mathbf{k}. \quad \text{(A73)}$$

An expression for $\mathbf{\Lambda_{MT}}$ will be given shortly.

Next, to determine $\mathbf{\Gamma}$, the extensional-shear strain in the overall beamlike stress-strain relations, we multiply both sides of Eq. (A70) by s and integrate from $y = 0$ to 2π. Introducing the 2π-periodic function

$$P \equiv A(y) - (y/2\pi)A \quad \text{(A74)}$$

and noting the following integrals:

$$\int_0^{2\pi} \mathbf{st}dy = -\pi\mathbf{I}, \quad \int_0^{2\pi} \mathbf{s}P^\bullet dy = -(\pi/2)\mathbf{m}, \quad \int_0^{2\pi} \mathbf{ss}dy \equiv \pi\mathbf{J},$$

$$\int_0^{2\pi} \mathbf{sg}dy = [\mathbf{J} + (B/\pi)^2\mathbf{k} \times \mathbf{H}] \cdot \mathbf{I}^{-1}, \quad \text{(A75)}$$

$$\int_0^{2\pi} \mathbf{s}(\mathbf{A}^\bullet)^T dy \cdot \mathbf{k} \equiv \mathbf{D}, \quad \int_0^{2\pi} \mathbf{s}B_{11}^\bullet dy \equiv \mathbf{D},$$

we obtain an expression of the form

$$\mathbf{\Gamma} \equiv \mathbf{w}' + \mathbf{k} \times \phi = \mathbf{\Lambda_{TT}} \cdot \mathbf{T}(0) + \mathbf{\Lambda_{TM}} \cdot \mathbf{M}(x). \quad \text{(A76)}$$

When we impose the conditions that $\mathbf{\Lambda_{TT}} = \mathbf{\Lambda_{TT}^T}$ and $\mathbf{\Lambda_{TM}} = \mathbf{\Lambda_{MT}^T}$, we find that $\mathbf{C} = -(2B/\pi)A_{2222}\mathbf{H}$, $\mathbf{D} = \mathbf{0}$, and $\mathbf{C} = \mathbf{D} = \mathbf{0}$. Thus, in view of Eq. (A62)

$$\mathbf{\Lambda_{TT}} = (1/2\pi)A_{2222}\mathbf{kk} + (1/A)A_{1222}(\mathbf{km} \cdot \mathbf{I}^{-1} + \mathbf{I}^{-1} \cdot \mathbf{mk})$$
$$+ (4/\pi)A_{1212}\mathbf{I}^{-1} \cdot [\mathbf{J} + (\pi^2 + 2A^2)\mathbf{mm}] \cdot \mathbf{I}^{-1}$$
$$+ (1/\pi)A_{1122}\mathbf{I}^{-1} \cdot \mathbf{J} \cdot \mathbf{I}^{-1} + (2/\pi)(\Gamma A_{1222} + \Gamma^* C_{1222})\mathbf{I}^{-1} \cdot [\mathbf{J}$$
$$+ (B/\pi)^2\mathbf{k} \times \mathbf{I}^{-1} \times \mathbf{k}] \cdot \mathbf{I}^{-1} \quad \text{(A77)}$$

and

$$\mathbf{\Lambda_{TM}} = -(2/\pi)[(\pi/A)^2 A_{1212}\mathbf{m} \cdot \mathbf{I}^{-1}\mathbf{k} + (B/\pi)A_{1212}\mathbf{H}^2]. \quad \text{(A78)}$$

Note that the last three lines of Eq. (A77) define a *shear coefficient tensor* as opposed to a shear coefficient scalar that one often sees introduced in theories for isotropic shear deformable beams.

Returning to Eq. (A70) and equating similar terms, we obtain

$$\mathbf{k} \cdot \mathbf{A} = (2/A)A_{1222}P(y)\mathbf{k} + (4\pi/A^2)A_{1212}\mathbf{m} \cdot \mathbf{I}^{-1}P(y) - (1/\pi)(4A_{1212}$$
$$+ A_{1122})\mathbf{v} \cdot \mathbf{I}^{-1} + (2/\pi)(\Gamma A_{1222} + \Gamma^* C_{1222})[(B/\pi)\mathbf{s} \cdot \mathbf{H} - \mathbf{v}] \cdot \mathbf{I}^{-1}$$
$$- \mathbf{r} \cdot \mathbf{\Lambda_{TT}} \quad \text{(A79)}$$

and

$$B_{11} = (2\pi/A^2)A_{1212}[\mathbf{r} \cdot \mathbf{I}^{-1} \cdot \mathbf{m} - 2P(y)]. \quad \text{(A80)}$$

The *exact* beamlike strain-stress relations (A72) and (A76) have now been determined *completely*.

A.1 The Generalized Displacement-rotation

Following the definition given by Ladevéze [13], we first use Eqs. (65) and (A8) to set

$$[s, s^{SV}]_x = \int_0^{2\pi} (F \cdot \underline{V}^{SV} - \underline{F}^{SV} \cdot \underline{V})_x dy. \quad \text{(A81)}$$

To define a *generalized displacement and rotation* $\bar{\bar{\mathbf{W}}}$ and $\bar{\bar{\mathbf{\Phi}}}$ we use Eqs. (A8), (A46), (A49)–(A52), (A59), and (A60) to write, in an obvious notation,

$$\underline{V}^{SV} = \mathbf{W}_I(x) + \varepsilon\mathbf{w}(x) + [\mathbf{\Phi}_I + \varepsilon\phi(x)] \times \mathbf{r} + \underline{A}(y) \cdot \mathbf{T}(0)$$
$$+ \underline{B}(y) \cdot \mathbf{M}(x) \quad \text{(A82)}$$

and

$$\underline{F}^{SV} = \underline{\overset{0}{A}}(y) \cdot \mathbf{T}(0) + \underline{\overset{0}{B}}(y) \cdot \mathbf{M}(x), \quad \text{(A83)}$$

where $\underline{A}$, $\underline{\overset{0}{A}}$, $\underline{B}$, and $\underline{\overset{0}{B}}$ are 4×2 matrices. Then,

$$[s, s^{SV}]_x = \mathbf{T}(0) \cdot [\mathbf{W}_I(x) + \varepsilon\mathbf{w}(x) - \bar{\bar{\mathbf{W}}}(x; \varepsilon)]$$
$$+ \mathbf{M}(x) \cdot [\mathbf{\Phi}_I + \varepsilon\phi(x) - \bar{\bar{\mathbf{\Phi}}}(x; \varepsilon)], \quad \text{(A84)}$$

where

$$\bar{\bar{\mathbf{W}}} \equiv \int_0^{2\pi} [\underline{A}^T(y; \varepsilon) \cdot \underline{V}(x, y; \varepsilon) - \underline{\overset{0}{A}}^T(y; \varepsilon) \cdot \underline{F}(x, y; \varepsilon)]dy$$
$$= \int_0^{2\pi} [(\mathbf{A} + \varepsilon\mathbf{A}_1)^T \cdot \mathbf{V} - \mathbf{F} \cdot (\mathbf{A}_I + \varepsilon\mathbf{A}) + \varepsilon^{3/2}(Z_*\mathbf{A} - Z\mathbf{A}^*)$$
$$+ \varepsilon^2\rho(E_y\tilde{\mathbf{A}} - M_x\tilde{\mathbf{A}}^*)]dy \quad \text{(A85)}$$

and

$$\bar{\bar{\mathbf{\Phi}}} \equiv \int_0^{2\pi} [\underline{B}^T(y; \varepsilon) \cdot \underline{V}(x, y; \varepsilon) - \underline{\overset{0}{B}}^T(y; \varepsilon) \cdot \underline{F}(x, y; \varepsilon)]dy$$
$$= \int_0^{2\pi} [(\mathbf{B} + \varepsilon\mathbf{B}_1)^T \cdot \mathbf{V} - \varepsilon\mathbf{F} \cdot \mathbf{B} + \varepsilon^{3/2}(Z_*\mathbf{B} - Z\mathbf{B}^*)$$
$$+ \varepsilon^2\rho(E_y\tilde{\mathbf{B}} - M_x\tilde{\mathbf{B}}^*)]dy. \quad \text{(A86)}$$

Noting that $[s_{SV}, s_{SV}]_x = 0$ and following the same arguments as in [14], we can show that

$$\bar{\bar{\mathbf{W}}} = \mathbf{W}_I + \varepsilon\mathbf{w} = \bar{\bar{\mathbf{W}}}^{SV} \quad \text{and} \quad \bar{\bar{\mathbf{\Phi}}} = \mathbf{\Phi}_I + \varepsilon\phi = \bar{\bar{\mathbf{\Phi}}}^{SV}. \quad \text{(A87)}$$

Thus, by Eqs. (A55) and (A59), we also have

$$\bar{\bar{\mathbf{\Gamma}}} = \varepsilon^{-1}(\bar{\bar{\mathbf{W}}}' + \mathbf{k} \times \bar{\bar{\mathbf{\Phi}}}) = \mathbf{\Gamma}^{SV} \quad \text{and} \quad \bar{\bar{\mathbf{\Omega}}} = \varepsilon^{-1}\bar{\bar{\mathbf{\Phi}}}' = \mathbf{\Omega}^{SV}, \quad \text{(A88)}$$

so that the beamlike strain-stress relations Eqs. (A72) and (A76) developed for the special class of Saint-Venant solutions are exact for any solution of the shell equations.

References

[1] Ladevèze, P., and Simmonds, J. G., 1996, "De Nouveaux Concepts en Théorie des Poutres pour des Charges et Géométries Quelconques," Compt. Rend. **322**(II) pp. 455–462.

[2] Ladevèze, P., and Simmonds, J. G., 1998, "New Concepts for Linear Beam Theory with Arbitrary Geometry and Loading," Eur. J. Mech. A/Solids **17**, pp. 377–402.

[3] Ladevèze, P., Sanchez, Ph., and Simmonds, J. G., 2004, "Beamlike (Saint-Venant) Solutions for Fully Anisotropic Elastic Tubes of Arbitrary Cross Section," Int. J. Solids Struct. **41**, pp. 1925–1944.

[4] Sanders, J. L. Jr., 1959, "An Improved First-Approximation Theory for Thin Shells", NASA Rep. 24.

[5] Koiter, W. T., 1959, "A Consistent First Approximation in the General Theory of Thin Elastic Shells," The Theory of Thin Elastic Shells, *Proc IUTAM Symposium*, Delft, 1959, Koiter W. T., ed., North-Holland, Amsterdam.

[6] Ladevèze, P. ed, 1985, *Local Effects in the Analysis of Structures*, Elsevier, Amsterdam.

[7] Simmonds, J. G., 1976, "Recent Advances in Shell Theory," Advances in Engineering Science, *13th Annual Meeting Society of Engineering Science*, NASA CP-2001, pp. 617–626.

[8] Goldenveiser, A. L., 1940, "Equations of the Theory of Thin Shells" (in Russian), Prikl. Mat. Mekh., **4**, pp. 35–42.

[9] Lure A. I., 1940, "General Theory of Thin Elastic Shells" (in Russian), Prikl. Mat. Mekh., **4**, pp. 7–34.

[10] McDevitt, T. J., and Simmonds, J. G., 1999, "Reduction of the Sanders-Koiter Equations for Fully Anisotropic Circular Cylindrical Shells to Two Coupled Equations for a Stress and a Curvature Function," J. Appl. Mech., **66**, pp. 593–597.

[11] Budiansky, B., and Sanders, J. L., Jr., 1963, "On the 'Best' First-Order Linear Shell theory," *Progress in Applied Mechanics* (Prager Anniversary Volume), Macmillan, New York, pp. 129–140.

[12] Gregory, R. D., and Wan, F. Y. M., 1984, "Decaying States of Plane Strain in a Semi-Infinite Strip and Boundary Conditions for Plate Theory," J. Elast. **14**, pp. 27–64.

[13] Ladevèze, P., 1983, "Sur le Principe de Saint-Venant en élasticité," J. Mec. Theor. Appl. **2**, pp. 161–184.

[14] Ladevèze, P., and Simmonds, J. G., 2001, "The Exact One-Dimensional Theory for End-Loaded Fully Anisotropic Beams of Narrow Cross Section," J. Appl. Mech. **68**, pp. 865–868.

Towards Chaos in Vibrating Damaged Structures—Part I: Theory and Period Doubling Cascade

Alberto Carpinteri

Nicola Pugno

Department of Structural Engineering
and Geotechnics,
Politecnico di Torino,
Corso Duca degli Abruzzi 24,
10129 Torino, Italy

The aim of the present paper is to evaluate the complex oscillatory behavior, i.e., the transition toward deterministic chaos, in damaged nonlinear structures under excitation. In the present paper (Part I), we show the developed theoretical approach and how it allows us to capture not only the super-harmonic and offset components (predominant for moderate nonlinear systems) but also the subharmonics of the structural dynamic response, describing complex and highly nonlinear phenomena, like the experimentally observed period doubling. Moreover, a period doubling cascade with a route to chaos seems to emerge from our considerations. [DOI: 10.1115/1.1934582]

1 Introduction

Vibration-based inspection of structural behavior offers an effective tool of non-destructive testing. The analysis of the dynamic response of a structure to excitation forces and the monitoring of alterations which may occur during its lifetime can be employed as a global integrity-assessment technique to detect, for example, play in joints or the presence of a crack. Indeed it is well known that, in the case of simple structures, crack position and depth can be determined from changes in natural frequencies, modes of vibration or the amplitude of the forced response.

The damage assessment problem in cracked structures has been extensively studied in the last decade, highlighting that the vibration based inspection is a valid method to detect, localize, and quantify cracks especially in beam structures. Dealing with the presence of a crack in the structure, previous studies have demonstrated that a transverse crack can change its state (from open to closed and vice versa) when the structure, subjected to an external load, vibrates. As a consequence, a nonlinear dynamic behavior is introduced.

In the past many studies have illustrated that a crack in a structure such as a beam, may exhibit nonlinear behavior if it is open during part of the response and closed in the remaining intervals. This phenomenon has been detected during experimental testing performed by Gudmundson [1] in which the influence of a transverse breathing crack upon the natural frequencies of a cantilever beam was investigated. The main result obtained was that the experimentally observed decrease in the natural frequencies of the beam due to the presence of the crack was not sufficient to be described by a model of crack which is always open. Therefore, it must be concluded that the crack alternately opened and closed giving rise to natural frequencies falling between those corresponding to the always-open and always-closed (e.g., integral) cases. In fact, if an always-open crack is assumed in the analysis of a beam with a so-called breathing crack, which both opens and closes during the time interval considered, the reduced decrease in

the experimental natural frequencies will lead to an underestimate of the crack depth if determined via a test-model correlation approach.

Friswell and Penny [2] have simulated the nonlinear behavior of a beam with a breathing crack vibrating in its first mode of vibration through a simple one-degree-of-freedom model with bilinear stiffness. The analysis of the fast Fourier transform and the response to harmonic loading, obtained by numerical integration, demonstrates the occurrence of harmonics in the response spectra which are integer multiples of the exciting frequency.

Krawczuk and Ostachowicz [3,4] presented an analysis of the forced vibrations of a cantilevered beam with a breathing crack, in which the equations of motion are solved using the harmonic balance technique. The periodically time-variant beam stiffness is simulated by a square wave function with a fundamental frequency equal to the forcing term frequency. According to [2] this research has shown that, when a breathing crack is present in a beam, higher harmonic components in the frequency spectrum of the response are generated if excited by a sinusoidal forcing function, indicating that the structure behaves nonlinearly.

Crespo, Ruotolo, and Surace [5] have solved and discretized the nonlinear equation of motion of a beam with breathing crack using the finite element method. According to the cited papers, the main assumption has been that the crack can be either fully open or fully closed during the vibration.

In [6] the vibrational response of a cantilevered beam with closing crack to harmonic forcing has been analyzed and its dynamic behavior characterized by using the so-called higher order frequency response functions.

Carpinteri and Carpinteri [7] highlight how in reality the crack opening and closing are continuous phenomena, i.e., the crack can be even partially open (or closed) as a function of the cracked element curvature.

The aim of this paper is to develop a coupled theoretical and numerical approach to evaluate the complex oscillatory behavior in damaged nonlinear structures under excitation. In particular, we have focused our attention on a cantilever beam with several breathing transverse cracks and subjected to harmonic excitation perpendicular to its axis. The method, that is an extension of the super-harmonic analysis carried out by Pugno, Surace, and Ruotolo [8] to subharmonic and zero frequency components, has allowed us to capture the complex behavior of the nonlinear structure, e.g., the occurrence of *period doubling*, as experimentally observed by Brandon and Sudraud [9] in cracked beams. The first results of this approach have been presented by Carpinteri and

Pugno in [10]. The method described assumes, as previously done in [10] and according to [7], that the cracks open and close continuously instead of instantaneously, as suggested by the experiments. As a consequence, the cracks are not considered to be either fully open or fully closed, but the intermediate configurations with partial opening have also been taken into account.

The period of the response is not assumed a priori equal to the period of the harmonic excitation, as classically supposed (absence of subharmonic components). This has allowed us to capture the complex behavior of the highly nonlinear structure, e.g., the occurrence of period doubling.

A pioneer work on period doubling was written in 1978, when Mitchell Feigenbaum [11] developed a theory to treat its route from the ordered to chaotic states. Even if oscillators showing the period doubling can be of different nature, as in mechanical, electrical, or chemical systems, these systems all share the characteristic of recursiveness. He provided a relationship in which the details of the recursiveness become irrelevant, through a kind of universal value, measuring the ratio of the distance between successive period doublings, the so called *Feigenbaum's delta* [12]. His understanding of the phenomenon was later experimentally confirmed [13], so that today we refer to the so-called *Feigenbaum's period doubling cascade*. However, even if the period doubling has a long history, only recently it has been experimentally observed in dynamics of cracked structures [9]. The aim of our study is the understanding of such phenomenon that, according to our model, seems to be ruled by the breathing of the cracks during the oscillation of the structure.

In addition to the super-harmonics, the analysis has systematically emphasized a presence of an offset (zero frequency component) in the structural response also for weak nonlinearities. Furthermore, subharmonic components appear in the response of the structure for stronger nonlinearities, leading, in particular conditions, to the period doubling.

The differential nonlinear equations governing the oscillations of the continuum structure, discretized by the finite element method, have been analyzed by means of the Fourier transforms or Fourier trigonometric series coupled with the harmonic balance approach. This allows us to obtain a nonlinear system of algebraic equations, easy to be solved numerically. In the numerical examples, the phenomenon of the period doubling is discussed, with an eye to the phase space trajectories.

2 Theoretical Continuum Approach

Let us consider a multicracked cantilever beam, clamped at one end and subjected to a dynamic distributed force p (with rotating frequency ω). Modeling the breathing cracks as concentrated nonlinear compliances (or stiffnesses) (in [14] linear stiffnesses are assumed), the equation of the motion of each integer beam segment, is the classical equation of the beam dynamics. Furthermore, the boundary conditions between two adjacent segments are represented by the continuity of the transversal displacement and of its second and third spatial derivatives (proportional to the bending moment and to the shearing force respectively), as well as by the compatibility with the crack. This implies that the difference in the rotations between the two adjacent sections must be equal to the rotation of the connecting concentrated nonlinear stiffness. The problem formally can be written as

$$\rho A \frac{\partial^2 q(z,t)}{\partial t^2} + EI \frac{\partial^4 q(z,t)}{\partial z^4} = p(z,t) \text{ for } z_i < z < z_{i+1};$$

whereas for $z = z_i$:

$$q(z_i^-) = q(z_i^+), \ q''(z_i^-) = q''(z_i^+), \ q'''(z_i^-) = q'''(z_i^+), \ q'(z_i^+) - q'(z_i^+)$$
$$= \frac{EI q''(z_i^\pm)}{k_i[q'(z_i^\pm)]} \tag{1}$$

where ρ is the density, A the cross-section area, q the transversal

displacement, E the Young's modulus, and I the moment of inertia of the beam. k_i is the nonlinear concentrated rotational stiffness (a function of the rotations $q'(z_i^\pm)$) of the crack placed at z_i (the symbol prime denotes the derivation with respect to z). We will discuss the form of the nonlinear stiffness in the next section, according to the experimental evidence.

Equation (1) can be formally solved by applying the Fourier trigonometric series searching a solution in the form

$$q(z,t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{+\infty} c_k(z) e^{ik\omega/\Theta t}, \tag{2}$$

where $c_k(z)$ are unknown functions and $\widetilde{P} = \Theta 2\pi/\omega = \Theta P$ is the period of the response, assumed a priori to be different from the period P of the excitation (describing a so-called complex behavior, thus we call Θ complexity index). On the other hand, if the period of the response tends to infinite, i.e., $\Theta \rightarrow \infty$ (nonperiodic response, i.e., chaotic deterministic behavior), it is well-known that Eq. (2) becomes formally a Fourier transform, i.e.

$$q(z,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega t} Q(z,\omega) d\omega, \tag{3}$$

where $Q(z,\omega)$ is the new unknown function. Thus, an approximation of the continuum spectrum in the response $q(z,t)$ of Eq. (3) is represented by the discrete spectrum of Eq. (2) if a sufficiently large complexity index Θ is considered. In addition, instead of the continuum approach of Eq. (1), a discretization of the system could also be considered. These discretizations allow an easier solution of the problem, as we will point out in the next section.

3 Theoretical Discrete Approach

By discretizing the structure with the finite element method [8], Eq. (1) can be rewritten as

$$[M]\{\ddot{q}\} + [D]\{\dot{q}\} + [K]\{q\} + \sum_m [\Delta K^{(m)}] f^{(m)}(\{q\})\{q\} = \{F\}, \tag{4}$$

where $[M]$ is the mass matrix, $[D]$ the damping matrix, $[K] + \Sigma_m[\Delta K^{(m)}]$ the stiffness matrix of the undamaged beam, and $[\Delta K^{(m)}]$ is half of the variation in stiffness introduced when the mth crack is fully open (see the Appendix and [8]). $\{F\}$ is the vector of the applied forces (with angular frequency ω) and $\{q\}$ is the vector containing the generalized displacements of the nodes (translations and rotations). According to this notation, $f^{(m)}(\{q\})$ is between -1 and $+1$ and models the transition between the conditions of mth crack fully open and fully closed. Assuming that this transition is instantaneous and hence takes place discontinuously, $f^{(m)}(\{q\})$ is a step function and has the sign of the curvature of the corresponding cracked element. With this simple model of crack opening and closing, $f^{(m)}(\{q\})$ can thus only be equal to -1 or $+1$. On the other hand, in the present investigation as in the previous [8], $f^{(m)}(\{q\})$ is assumed to be a linear function of the curvature of the corresponding cracked element. In other words, the cracks are not considered fully open or fully closed, as the intermediate configurations with partial opening are also taken into account. Thus, the stiffness varies continuously between the two extremes of undamaged or totally damaged beam (fully open cracks), rather than stepwise. The solution for the elements of the $\{q\}$ vector $\in L^2$ (i.e., $|q_i|^2$ can be integrated according to Lebesgue) can be found by means of Eq. (3), or by the approximation of Eq. (2), that for our discrete system can be rewritten as

$$q_i = \sum_{j=0}^{N} \left(A_{ij} \sin j\frac{\omega}{\Theta} t + B_{ij} \cos j\frac{\omega}{\Theta} t \right), \tag{5}$$

in which the complexity integer Θ must be a positive integer, to take into account not only the super-harmonics (and offset) but

also the subharmonic components of the dynamic response and theoretically $N=\infty$. As previously discussed, this means that the response could have a period $\tilde{P}=\Theta P$ that is not a priori coincident with the period P of the excitation. A value for Θ tending to infinite (Fourier trigonometric series become Fourier transforms) describes a transition toward a chaotic (nonperiodic) response.

It is interesting to note that, even if the trigonometric series (2) converges, it could not be a trigonometric Fourier series. In fact, the Fischer-Riesz theorem affirms that it is a fourier series if and only if $\Sigma_{j=0}^{\infty}(|A_{ij}|^2+|B_{ij}|^2)$ converges. In this case, the Parseval equation

$$\sum_{j=0}^{\infty}(|A_{ij}|^2+|B_{ij}|^2)=\frac{2}{\tilde{P}}\int_{-\tilde{P}/2}^{+\tilde{P}/2}|q_i(t)|^2\mathrm{d}t, \qquad (6)$$

obviously implies

$$\lim_{j\to\infty}A_{ij}=\lim_{j\to\infty}B_{ij}=0. \qquad (7)$$

The last relationships allow us to consider a finite number N in Eq. (5), large enough to provide a good approximation. The function $f^{(m)}(\{q\})$ is considered linear versus the curvature of the corresponding cracked element, i.e.,

$$f^{(m)}(\{q\})=\frac{q_{m_k}-q_{m_h}}{|q_{m_k}-q_{m_h}|}\max=\Lambda_m(q_{m_k}-q_{m_h}), \qquad (8)$$

where the numerator represents the difference in the rotations at the ends of the corresponding cracked element and the denominator is the maximum absolute value that can be reached by this difference: consequently, the generic component of function $\{g^{(m)}\}=f^{(m)}(\{q\})\{q\}$ (that appears in Eq. (4)) can be expressed as

$$g_i^{(m)}=\Lambda_m(q_{m_k}-q_{m_h})q_i. \qquad (9)$$

The same concepts argued for the q_i components can be now applied to the $g_i^{(m)}$, ensuring that they can be developed in a trigonometric Fourier series and can thus be put in the approximate form

$$g_i^{(m)}=\sum_{j=0}^{N}\left(C_{ij}^{(m)}\sin j\frac{\omega}{\Theta}t+D_{ij}^{(m)}\cos j\frac{\omega}{\Theta}t\right), \qquad (10)$$

with

$$C_{ij}^{(m)}=\frac{2}{\tilde{P}}\int_{0}^{\tilde{P}}g_i^{(m)}(t)\sin\left(j\frac{\omega}{\Theta}t\right)\mathrm{d}t, \qquad (11)$$

$$D_{ij}^{(m)}=\frac{2}{\tilde{P}}\int_{0}^{\tilde{P}}g_i^{(m)}(t)\cos\left(j\frac{\omega}{\Theta}t\right)\mathrm{d}t. \qquad (12)$$

Inserting relation (9), in its explicit form according to Eq. (5) for the degrees of freedom i, m_h, and m_k, into Eqs. (11) and (12) and developing the integrals, gives the following expressions:

$$C_{ij}^{(m)}=\frac{\Lambda_m}{2}\sum_{j_1,j_2:j_1+j_2=j}\{(A_{m_kj_1}-A_{m_hj_1})B_{ij_2}+(B_{m_kj_1}-B_{m_hj_1})A_{ij_2}\}$$
$$+\frac{\Lambda_m}{2}\sum_{j_1,j_2:j_1-j_2=\pm j}\pm\{(A_{m_kj_1}-A_{m_hj_1})B_{ij_2}-(B_{m_kj_1}-B_{m_hj_1})A_{ij_2}\}, \qquad (13)$$

$$D_{ij}^{(m)}=\frac{\Lambda_m}{2}\sum_{j_1,j_2:j_1+j_2=j}\{-(A_{m_kj_1}-A_{m_hj_1})A_{ij_2}+(B_{m_kj_1}-B_{m_hj_1})B_{ij_2}\}$$
$$+\frac{\Lambda_m}{2}\sum_{j_1,j_2:j_1-j_2=\pm j}\pm\{(A_{m_kj_1}-A_{m_hj_1})A_{ij_2}-(B_{m_kj_1}-B_{m_hj_1})B_{ij_2}\}. \qquad (14)$$

As the nonlinearity of the components of $\{g^{(m)}\}=f^{(m)}(\{q\})\{q\}$ were expressed in a form analogous to that of the components of $\{q\}$, as indicated by Eq. (10), it is possible at this stage to apply the harmonic balance method (Eqs. (5) and (10)—with Eqs. (13) and (14)—must be introduced in Eq. (4) and the terms of the harmonics of the same order must be "balanced"), which leads to N different systems of nonlinear algebraic equations. We find the following solution:

$$\begin{bmatrix}[K]-\dfrac{j^2\omega^2}{\Theta^2}[M] & -\dfrac{j\omega}{\Theta}[D] \\ \dfrac{j\omega}{\Theta}[D] & [K]-\dfrac{j^2\omega^2}{\Theta^2}[M]\end{bmatrix}\begin{Bmatrix}\{A_j\} \\ \{B_j\}\end{Bmatrix}$$
$$=\{F_j\}-\sum_m\begin{bmatrix}[\Delta K^{(m)}] & [0] \\ [0] & [\Delta K^{(m)}]\end{bmatrix}\begin{Bmatrix}\{C_j^{(m)}\} \\ \{D_j^{(m)}\}\end{Bmatrix}, \qquad (15)$$

where $j=0,1,\ldots,N$ and for each vector we have $\{V_j\}^T=\{V_{1j},V_{2j},\ldots,\}^T$. In addition

$$F_{ij}=F\delta_{j\Theta}\delta_{ip}, \qquad (16)$$

p being the node position corresponding to the point where the sinusoidal force is applied.

Each system can be solved numerically using an iterative procedure interrupted by an appropriate convergence test when the relative j error for the $\{A_j\}$ and $\{B_j\}$ vectors becomes lower than a specified value; it is a function of the kth iteration and has been defined as

$$e_{jk}=\left\|\begin{Bmatrix}\{A_j\} \\ \{B_j\}\end{Bmatrix}_k-\begin{Bmatrix}\{A_j\} \\ \{B_j\}\end{Bmatrix}_{k-1}\right\|\Bigg/\left\|\begin{Bmatrix}\{A_j\} \\ \{B_j\}\end{Bmatrix}_{k-1}\right\|. \qquad (17)$$

The procedure consists in determining the unknowns A_{ij} and B_{ij}. It is very interesting to note that, assuming the coefficients $C_{ij}^{(m)}$, $D_{ij}^{(m)}$ to be zero at the first step, implies to force also the subharmonic components to be zero (see Eq. (15)). So, differently from the super-harmonic analysis [8], we have to start with non-zero values for the coefficients $C_{ij}^{(m)}$, $D_{ij}^{(m)}$. To obtain good initial values for these coefficients, we have considered as a zero step a super-harmonic analysis ($\Theta=1$); in this case, we can determine the unknowns A_{ij} and B_{ij} starting with zero coefficients $C_{ij}^{(m)}$, $D_{ij}^{(m)}$ and, by Eqs. (13) and (14), we have their initial values for the subharmonic analysis. The solution thus obtained is used to determine the known vector of the right hand-side of Eq. (15). The procedure is repeated until the desired precision is achieved and coefficients A_{ij} and B_{ij} are found, while Eq. (5) is used to determine the components of the approximate vector, which satisfies the nonlinear Eq. (4), giving an approximated solution of the continuum system described in Eq. (1).

4 Period Doubling Cascade

We can consider two different numerical examples: a weakly nonlinear structure and a strongly nonlinear one. Only in the latter case, the so-called period doubling phenomenon, experimentally observed by Brandon and Sudraud [9], clearly appears. The beam here considered is the same as that described in the mentioned experimental analysis. It is 270 mm long and has a transversal rectangular cross section of base and height, respectively, of 13 and 5 mm. The material is (UHMW)-ethylene, with a Young's modulus of 8.61×10^8 N/m^2 and a density of 935 kg/m^3. We have assumed a modal damping of 0.002. It is discretized with 20
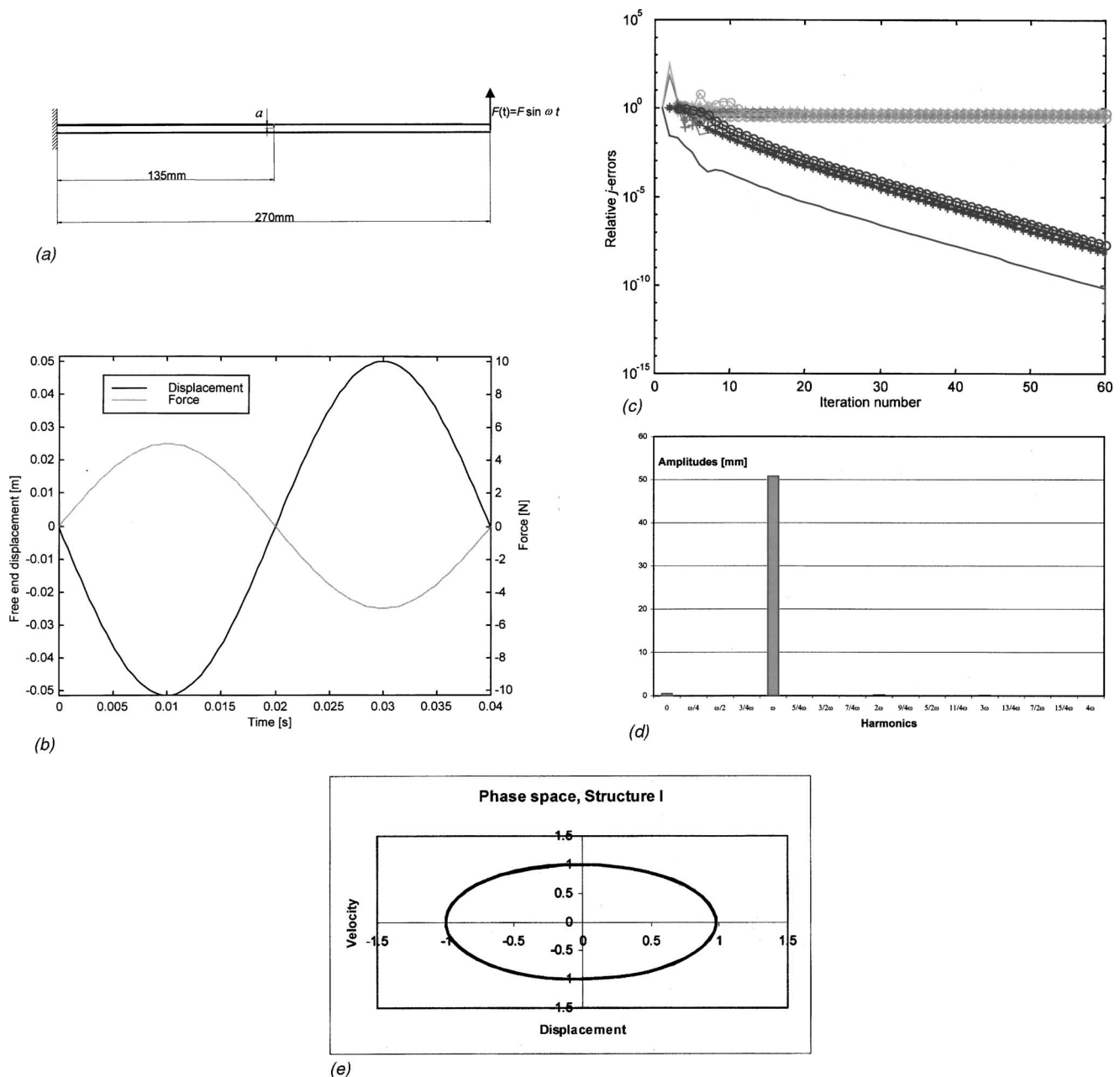
Fig. 1 (*a*) Structure I—Damaged structure and characteristics of the excitation ($a=2.4$ mm, $F=5$N, $f=\omega/2\pi=25$ Hz). (*b*) Structure I—Time history of the free end displacement and of the applied force. (*c*) Structure I—Relative error as a function of the iteration number, for each *j*th harmonic ($j=0,1,\ldots,16$). (*d*) Structure I—Zero- (offset), sub- and super-harmonic components for the free end displacement (i.e., $\sqrt{A_{20j}^2+B_{20j}^2}$ for $j=0,1,\ldots,16$). (*e*) Structure I—Dimensionless phase diagram of the response (free-end displacement).

finite elements. We have found that a complexity index $\Theta=4$ and $N=16$ give a good approximation. For larger values of Θ and N, substantially coincident solutions are obtained. The first natural frequency of the undamaged structure is $f_u=10.6$ Hz.

For each of the two considered structures (Figs. 1(*a*) and 2(*a*)) it is shown the time history of the applied force and of the free-end displacement (Figs. 1(*b*) and 2(*b*)), the relative errors as functions of the iteration number (Figs. 1(*c*) and 2(*c*)) and the zero-, sub-, and super-harmonic components for the free-end displacement (Figs. 1(*d*) and 2(*d*)). In Tables 1 and 2, the frequency components are considered separately as sin and cos components.

In a hypothetical linear structure, the structural response is linear by definition with obviously only one harmonic component at the same frequency of the excitation.

In the weakly nonlinear structure of Fig. 1(*a*), the response converges and it appears only weakly nonlinear, as depicted in Fig. 1(*b*). The relative errors, shown in Fig. 1(*c*), tend to zero or by definition are equal to 1 if related to the harmonic components identically equal to zero. The harmonic components in the structural response are the zero-one (presence of a negative offset in the displacement of the free-end, downward in Fig. 1(*a*)) and the super-harmonic ones (Fig. 1(*d*) and Table 1). No subharmonic components can be observed. The corresponding phase diagram of the response is shown in Fig. 1(*e*). Due to the weak nonlinearity the trajectory in the phase diagram is close to an ellipse. The diagram is nonsymmetric as the spatial positions of the cracks (placed in the upper part of the beam). The trajectory is an unique closed curve since here the period of the response is equal to the
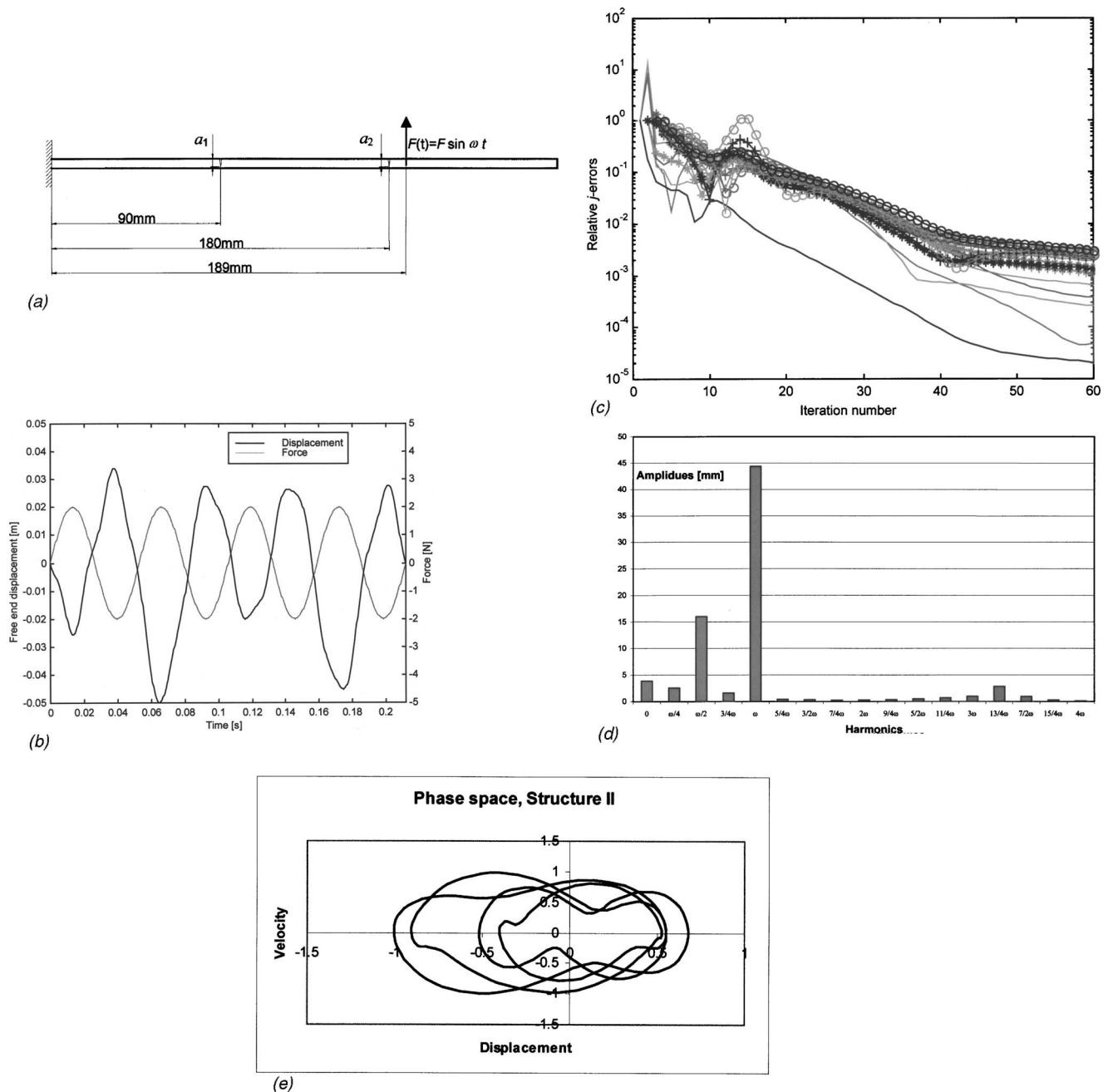
Fig. 2 (*a*) Structure II—Damaged structure and characteristics of the excitation (a_1=4.25 mm, a_2=4.25 mm, *F*=2N, *f*=$\omega/2\pi$ =18.9 Hz). (*b*) Structure II—Time history of the free end displacement and of the applied force. (*c*) Structure II—Relative error as a function of the iteration number, for each *j*th harmonic (*j*=0,1,...,16). (*d*) Structure II—Zero- (offset), sub- and super-harmonic components for the free end displacement (i.e., $\sqrt{A_{20j}^2+B_{20j}^2}$ for *j*=0,1,...,16). (*e*) Structure II—Dimensionless phase diagram of the response (free-end displacement).

period of the excitation. An increasing of the nonlinearity will distort the trajectory as we will show in the next example.

In the strongly nonlinear structure of Fig. 2(*a*) the response converges (Fig. 2(*c*)) and the nonlinearity increases, as depicted in Fig. 2(*b*). The harmonic components in the structural response are the zero one, the super harmonic as well as the subharmonic ones (Fig. 2(*d*) and Table 2). It should be emphasized that a strong presence of the component causes the period doubling of the response, i.e., the $\omega/2$ component. The free-end vibrates practically with a period doubled with respect to the excitation. A non-negligible component at $\omega/4$ is observed too, representing a route to chaos through a period doubling cascade.

The corresponding phase diagram of the response is shown in Fig. 2(*e*). The trajectory is a unique closed curve since the response is still periodic; it is composed by multiple cycles since here the period of the response is not equal to the period of the excitation. The distortions in the trajectory are consequences of the presence of the super- or subharmonics, as well as the multiple cycles emphasize the presence of the subharmonics (four cycles are due to the component $\omega/4$), i.e., the presence of a complexity with associated route to chaos. Also in this case, the diagram is nonsymmetric as the spatial positions of the cracks.

An extensive parametrical investigation can be found in the companion paper.

Table 1 Structure I—Zero- (offset), sub- and super-harmonic sin- and cos- components [mm], for the free end displacement (i.e., A_{20j}, B_{20j}, for j=0, 1,…, 16)

Harmonic	Sin	Cos	Amplitude
0	0	−0.5472	0.5472
$\omega/4$	0	0	0
$\omega/2$	0	0	0
$3/4\omega$	0	0	0
ω	−50.7971	−0.1414	50.7973
$5/4\omega$	0	0	0
$3/2\omega$	0	0	0
$7/4\omega$	0	0	0
2ω	0.0039	0.1688	0.1689
$9/4\omega$	0	0	0
$5/2\omega$	0	0	0
$11/4\omega$	0	0	0
3ω	0.0745	−0.0008	0.0745
$13/4\omega$	0	0	0
$7/2\omega$	0	0	0
$15/4\omega$	0	0	0
4ω	0.0001	0.0035	0.0035

5 Conclusions

The proposed approach extends the theory proposed by Pugno et al. [8] to (offset and) subharmonic components. We have demonstrated that our approach corresponds to an approximated solution of the continuum spectrum of the response of the continuum system. The method has allowed us to catch complex phenomena, i.e., transition toward deterministic chaos, like the occurrence of a period doubling, as shown in the numerical examples and experimentally observed in the context of cracked beam by Brandon and Sudraud [9]. In this analysis, of crucial importance appears the complexity index Θ. For higher values of Θ we have to increase also N (e.g., $N \approx \Theta^2$), so that the complexity of the numerical simulations considerably increases. On the other hand, larger values of Θ allow us to catch higher structural complexity, as emphasized by multiple cycles of the trajectory in the phase space diagrams.

From the reported numerical examples (for an extensive numerical parametrical investigation see the companion paper), we can affirm that if the nonlinearity is zero, the structural response (i.e., Eq. (5)) can be obviously caught with $N = \Theta = 1$. If a weak nonlinearity is considered, only offset and super-harmonic components can be observed in the structural response. As a conse-quence, for this case, it can be easily caught using classical Fourier series ($\Theta = 1$) with $N > 1$ (and large enough, in the sense that $\text{Response}(N) \cong \text{Response}(N' > N)$). If the nonlinearity becomes stronger, offset, and super-harmonic components, as well as sub-harmonic ones, can be observed in the structural response. As a consequence, in this case, it can be caught using a complex index Θ larger than 1 ($\Theta > 1$ and $N \gg 1$ large enough, in the sense that $\text{response}(\Theta, N) \cong \text{response}(\Theta' > \Theta, N' > N)$). Theoretically, values of Θ tending to infinity (Fourier series become Fourier transforms, with theoretically N tending to infinity too) allow us to catch a route to chaos through a period doubling cascade, that here would imply a nonperiodic dynamic response. These considerations are summarized in Table 3.

Table 2 Structure II—Zero- (offset), sub- and super-harmonic sin- and cos-components [mm], for the free end displacement (i.e., A_{20j}, B_{20j}, for j=0, 1,…, 16)

Harmonic	Sin	Cos	Amplitude
0	0	−3.2415	3.2415
$\omega/4$	0.5278	−0.5708	0.7774
$\omega/2$	9.9741	7.6932	12.5964
$3/4\omega$	−0.0201	−1.0667	1.0669
ω	−30.2576	−0.3138	30.2592
$5/4\omega$	0.151	0.3043	0.3397
$3/2\omega$	0.2172	−0.0064	0.2173
$7/4\omega$	−0.1981	−0.2078	0.2871
2ω	−0.1384	0.3512	0.3775
$9/4\omega$	0.421	0.1979	0.4652
$5/2\omega$	−0.1315	−0.3162	0.3425
$11/4\omega$	−0.5769	0.1225	0.5898
3ω	0.5754	0.1795	0.6027
$13/4\omega$	1.7804	−2.4618	3.0382
$7/2\omega$	0.6258	−0.296	0.6923
$15/4\omega$	−0.038	−0.7144	0.7154
4ω	−0.0078	0.0539	0.0545

Nonlinearity	Zero	Weak	Strong	Very strong (chaos)
Θ	1	1	>1	∞
N	1	>1	$\gg 1$	∞

Appendix

The mathematical model used for the considered beam of Leonhard Euler (1707–1783) with several transverse one-side non-propagating breathing cracks is based on the finite element model. According to the principle of Ademare Jean-Claude Barré de Saint-Vénant (1797–1886) the stress field is influenced only in the region adjacent to the crack. The element stiffness matrix, with the exception of the terms which represent the cracked element, may be regarded as unchanged under a certain limitation of the element size. The additional stress energy of a crack leads to a flexibility coefficient expressed by a stress intensity factor derived by means of the theorem of Carlo Alberto Castigliano (1847–1884) in the linear elastic regime.

The cracked element is shown in Fig. 3.

Neglecting shear action (Euler beam), the strain energy of an element without a crack can be obtained as

$$W^{(0)} = \frac{1}{2EI} \int_0^l (M + Pz)^2 dz = \frac{1}{2EI}(M^2 + P^2 l^3/3 + MPl^2),$$

(A1)

where E is the Young's modulus of the material constituting the finite element, $I = bh^3/12$ is the moment of inertia of its cross section, having base b and height h, and M and P are the generalized forces acting at the ends of the finite element of length l. The additional energy due to the crack is

$$W^{(1)} = b \int_0^a [[K_I^2(x) + K_{II}^2(x)]/E' + (1 + \nu)K_{III}^2(x)/E]dx, \quad \text{(A2)}$$

where $E' = E$ for plane stress, $E' = E/(1 + \nu)$ for plane strain and ν is the Poisson's ratio. $K_{I,II,III}$ are the stress intensity factors for opening, sliding and tearing-type crack, of depth a, respectively.

Taking into account only bending

$$W^{(1)} = b \int_0^a [[K_{IM}(x) + K_{IP}(x)]^2 + K_{IIP}^2(x)]/E' dx, \quad \text{(A3)}$$

with

$$K_{IM} = (6M/bh^2)\sqrt{\pi a}F_I(s)$$

$$K_{IP} = (3Pl/bh^2)\sqrt{\pi a}F_I(s),$$

$$K_{IIP} = (P/bh)\sqrt{\pi a}F_{II}(s) \quad \text{(A4)}$$

where $s = a/h$ and

$$F_I(s) = \sqrt{2/(\pi s)\tan(\pi s/2)}$$
$$\times \{0.923 + 0.199[1 - \sin(\pi s/2)^4]\}/\cos(\pi s/2)$$
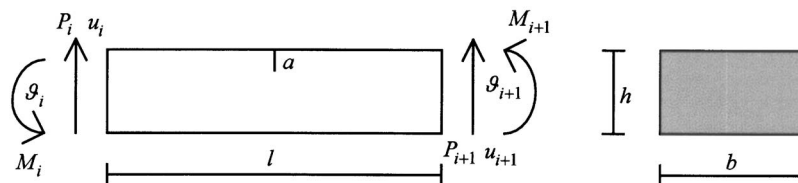
$$F_{II}(s) = (3s - 2s^2)(1.122 - 0.561s + 0.085s^2 + 0.18s^3)/\sqrt{1 - s}.$$

(A5)

The term $c_{ik}^{(0)}$ of the flexibility matrix $[C_e^{(0)}]$ for an element without crack can be written as

$$c_{ik}^{(0)} = \frac{\partial^2 W^{(0)}}{\partial P_i \, \partial P_k} \quad i,k = 1,2 \quad P_1 = P, P_2 = M. \quad \text{(A6)}$$

The term $c_{ik}^{(1)}$ of the additional flexibility matrix $[C_e^{(1)}]$ due to the crack can be obtained as

$$c_{ik}^{(1)} = \frac{\partial^2 W^{(1)}}{\partial P_i \, \partial P_k} \quad i,k = 1,2 \quad P_1 = P, P_2 = M. \quad \text{(A7)}$$

The term c_{ik} of the total flexibility matrix $[C_e]$ for the damaged element is

$$c_{ik} = c_{ik}^{(0)} + c_{ik}^{(1)}. \quad \text{(A8)}$$

From the equilibrium condition (Fig. 3)

$$(P_i \; M_i \; P_{i+1} \; M_{i+1})^T = [T](P_{i+1} \; M_{i+1})^T, \quad \text{(A9)}$$

where

$$[T] = \begin{bmatrix} -1 & -l & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}^T. \quad \text{(A10)}$$

Applying the theorem of Enrico Betti (1823–1892), the stiffness matrix of the undamaged element can be written as

$$[K_e] = [T][C_e^{(0)}]^{-1}[T]^T, \quad \text{(A11)}$$

or

$$[K_e] = \frac{EI}{l^3} \begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix}, \quad \text{(A12)}$$

while the stiffness matrix of the cracked element may be derived as

$$[K_{de}] = [T][C_e]^{-1}[T]^T, \quad \text{(A13)}$$

In order to evaluate the dynamic response of the cracked beam when acted upon by an applied force, it is supposed that the crack does not affect the mass matrix. Therefore, for a single element, the mass matrix can be formulated directly

$$[M_e] = [M_{de}] = \frac{ml}{420} \begin{bmatrix} 156 & 22l & 54 & -13l \\ 22l & 4l^2 & 13l & -3l^2 \\ 54 & 13l & 156 & -22l \\ -13l & -3l^2 & -22l & 4l^2 \end{bmatrix}, \quad \text{(A14)}$$

where m is the mass for unity length of the beam.

Assuming that the damping matrix $[D]$ is not affected by the crack, it can be calculated through the inversion of the modeshape matrix $[\phi]$ relative to the undamaged structure

$$[D] = ([\phi]^T)^{-1}[d][\phi]^{-1}, \quad \text{(A15)}$$

where $[d]$ is the following matrix:

Fig. 3 Cracked element

$$[d] = 2 \begin{bmatrix} \zeta_1 \omega_1 M_1 & 0 & 0 & \ldots & 0 \\ 0 & \zeta_2 \omega_2 M_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & 0 & \zeta_n \omega_n M_n \end{bmatrix}, \quad (A16)$$

in which ζ_i is the modal damping ratio, ω_i is the ith natural frequency, and M_i is the ith modal mass relative to the undamaged beam.

References

[1] Gudmundson, P., 1983, "The Dynamic Behaviour of Slender Structures with Cross-Sectional Cracks," J. Mech. Phys. Solids, **31**, pp. 329–345.

[2] Friswell, M. I., and Penny, J. E. T., 1992, "A Simple Nonlinear Model of a Cracked Beam," *Proc. X Int. Modal Analysis Conf.*, pp. 516–521.

[3] Krawczuk, M., and Ostachowicz, W., 1994, "Forced Vibrations of a Cantilever Timoshenko Beam With a Closing Crack," ISMA **19**, pp. 1067–1078.

[4] Ostachowicz, W., and Krawczuk, M., 1990, "Vibration Analysis of a Cracked Beam," Comput. Struct., **36**, pp. 245–250.

[5] Crespo, P., Ruotolo, R., and Surace, C., 1996, "Non-Linear Modelling of Cracked Beam," *Proc. of XIV Int. Modal Analysis Conf.*, pp. 1017–1022.

[6] Ruotolo, R., Surace, C., Crespo, P., and Storer, D., 1996, "Harmonic Analysis of the Vibrations of a Cantilevered Beam with a Closing Crack," Comput. Struct., **61**, pp. 1057–1074.

[7] Carpinteri, Al., and Carpinteri, An., 1984, "Softening and Fracturing Process in Masonry Arches," *Proc. of the Sixth Int. Brick Masonry Conf.*, pp. 502–510.

[8] Pugno, N., Surace, C., and Ruotolo, R., 2000, "Evaluation of the Non-Linear Dynamic Response to Harmonic Excitation of a Beam with Several Breathing Cracks," J. Sound Vib., **235**, pp. 749–762.

[9] Brandon, J. A., and Sudraud, C., 1998, "An Experimental Investigation into the Topological Stability of a Cracked Cantilever Beam," J. Sound Vib., **211**, pp. 555–569.

[10] Carpinteri, A., and Pugno, N., 2002, "Complexity of the Nonlinear Forced Vibrations in Multicracked Structures," *Proc. of the IX Int. Con. On Sound & Vibration*, July 8–11, Orlando, FL, (CD-ROM N. P114-1).

[11] Feigenbaum, M. J., 1978, "Quantitative Universality for a Class of Nonlinear Transformations," J. Stat. Phys., **19** (1), 25–52.

[12] Feigenbaum, M. J., 1983, "Universal Behavior in Nonlinear Systems," Physica D, **7**, pp. 16–39.

[13] Linsay, P. S., 1981, "Period Doubling and Chaotic Behavior in a Driven Anharmonic Oscillator," Phys. Rev. Lett., **47**, pp. 1349–1352.

[14] Li, Q. S., 2003, "Vibratory Characteristics of Timoshenko Beams with Arbitrary Number of Cracks," J. Eng. Mech., **129**, pp. 1355–1359.

Towards Chaos in Vibrating Damaged Structures—Part II: Parametrical Investigation

Alberto Carpinteri
e-mail: carpinteri@polito.it

Nicola Pugno
e-mail: nicola.pugno@polito.it

Department of Structural Engineering and
Geotechnics,
Politecnico di Torino,
Corso Duca degli Abruzzi 24,
10129 Torino, Italy

The aim of the present paper is to evaluate the complex oscillatory behavior, i.e., the transition to chaos, in damaged nonlinear structures under excitation. In the present paper, Part II, we apply the theoretical approach described in Part I to perform an extensive parametrical investigation. We focus our attention on a cantilevered beam with several breathing cracks subjected to sinusoidal excitation. The numerical simulations have been performed by varying the number of cracks, their depth and position, as well as the amplitude, frequency and position of the excitation, for a total of 83 different cases. [DOI: 10.1115/1.1934631]

1 Introduction

As shown in Part I, the proposed theoretical and numerical approach can be successfully applied to the study of damaged structures. The aim of the present paper, Part II, is to perform an extensive parametrical investigation to describe the influence of the main parameters on the dynamic behavior of the considered system. For a given model, the system complexity is a function of the complexity of the structure, as well as of the complexity of the excitation. Focusing our attention to the excited cracked cantilevered beam introduced in Part I, we have performed a parametrical investigation by varying the main parameters influencing the structural complexity, i.e., the cracks' number, depth and position, as well as the force amplitude, frequency and position, for a total of 83 different cases.

Several researchers have studied the problem of a beam with a breathing crack from analytical, numerical and experimental viewpoints [1–7]. In particular, relevant numerical investigations have been presented in Ref. [1], by using the Finite Element Method and in Ref. [2], applying directly numerical integration. In spite of this, an extensive parametrical investigation on the topic is entirely absent in the literature and is the object of the present paper (Part II).

The method, described in detail in Part I, has permitted to capture the influence of the different parameters on the complex behavior for the nonlinear structure, as well as the transition towards deterministic chaos, i.e., towards a nonperiodic response of the structure subjected to periodic excitation.

In particular, we have found that, if a weak nonlinearity is considered, only offset and super-harmonic components can be observed in the structural response. On the other hand, if the nonlinearity becomes stronger, also sub-harmonic components can be observed in the structural response, providing the so-called *complex* behavior.

Furthermore, the influence of each parameter on the structural behavior will be discussed on the basis of the presented extensive parametrical investigation. A new methodology for vibration-based inspections will also be presented.

2 Parametrical Simulations

As an example, we focus our attention onto a clamped beam. It is 270 mm long and has a transversal rectangular cross section of base and high, respectively, equal to 13 and 5 mm. The material is UHMW-ethylene, with a Young's modulus of 8.61×10^8 N/m^2 and a density of 935 kg/m^3. We have assumed a modal damping of 0.002. The beam has been discretized with 20 finite elements. We have found that a Complexity Index $\Theta = 4$ and a number of terms $N = 16$ give a good approximation (i.e., for larger values of Θ and N substantially identical solutions are obtained). The first natural frequency of the undamaged structure is $f_u = 10.6$ Hz.

The extensive parametrical investigation has been performed by varying the main parameters quoted in Fig. 1. These parameters affect the behavior of the system, as summarized in the following:

(A) By varying the depth of a crack localized at one-half of the total length of the beam;
(B) by varying the depth of a crack localized at one-third of the total length of the beam;
(C) by varying the crack position;
(D) by varying the excitation frequency;
(E) by varying the excitation amplitude;
(F) by varying the depth of one crack (in a beam containing two cracks);
(G) by varying the position of the excitation (in a beam containing two cracks).

Each of these families of parametrical simulations is separately treated in a specific section. The outputs from each simulation are the same as presented in the examples of Part I. As structural response we present only the normalized amplitude, corresponding to a given frequency component. It is defined as the ratio of the amplitude of the considered frequency component to the amplitude of the linear one (the component of the response with the same frequency of the excitation) related to the displacement of the free-end, i.e.,

$$\text{Normalized Amplitude}|_{j \neq \Theta} = \frac{\sqrt{A_{20j}^2 + B_{20j}^2}}{\sqrt{A_{20\Theta}^2 + B_{20\Theta}^2}} \tag{1}$$

with reference to the variables introduced in Part I.

2.1 Parametrical Simulations by Varying the Depth of a Crack Localized at One-Half of the Total Length of the Beam (A). These numerical simulations consider one crack with a variable depth of a_1. They are indicated by the letter A. Referring to Fig. 1, the coordinate of the crack is $d_1 = 135$ mm (at one-half of the total length of the beam), the force amplitude is $F = 5$N with a

Fig. 1 The considered nonlinear system and the main parameters numerically investigated

frequency f of 25 Hz (to compare with the first natural frequency of the undamaged structure, $f_u = 10.6$ Hz). The parameters of the simulations are summarized in Table 1. The numerical responses in terms of the normalized amplitude of Eq. (1) as a function of the crack depth are presented in Fig. 2. Figure 2(a) considers the frequency components $j = 0, 1, 2, 3$, as well as Figs. 2(b)–2(d), respectively, the components groups $j = 5–8; j = 9–12–j = 13–16$. It is very interesting to note that the first symptom of the presence of a crack, i.e., of the nonlinearity, is the offset ($j = 0$, 0-frequency) in Fig. 2(a), as well as the super-harmonic components, i.e., $j = 8(2\omega)$, $j = 12(3\omega)$, and $j = 16(4\omega)$, which are also present for small crack depths. As a consequence, we can affirm that the nonlinearity implies a natural rupture of the symmetry of the problem (i.e., an offset). A rather considerable presence of sub-harmonic components arises after a threshold value of crack depth, which is around one-half of the total height of the beam. For this excitation frequency (around twice the first natural frequency), the component of *period doubling* ($\omega/2$) and its multiples ($3\omega/2$, $5\omega/2$, $7\omega/2$) are clearly prevailing.

2.2 Parametrical Simulations by Varying the Depth of a Crack Localized at One-Third of the Total Length of the Beam (B). These numerical simulations consider one crack with a variable depth of a_1. They are indicated by the letter B. Referring to Fig. 1, the coordinate of the crack is $d_1 = 90$ mm (at one-third of the total length of the beam), the force amplitude is $F = 2$N with a frequency f of 19 Hz. The parameters of the simulations are summarized in Table 2. The numerical responses in terms of the normalized amplitude of Eq. (1) as a function of the crack depth, are presented in Fig. 3. The results are similar to those of the previous

Table 1 One crack (localized at one-half of the total length of the beam)—Numerical simulations by varying the crack depth (A)

One crack—a_1 variable $d_1 = 135$ mm; $F = 5$N; $f = 25$ Hz	
a_1 [mm]	*Case*
1.0	A1
2.0	A2
2.2	A3
2.4	A4
2.6	A5
2.8	A6
3.0	A7
3.2	A8

(a)

(b)

(c)

(d)
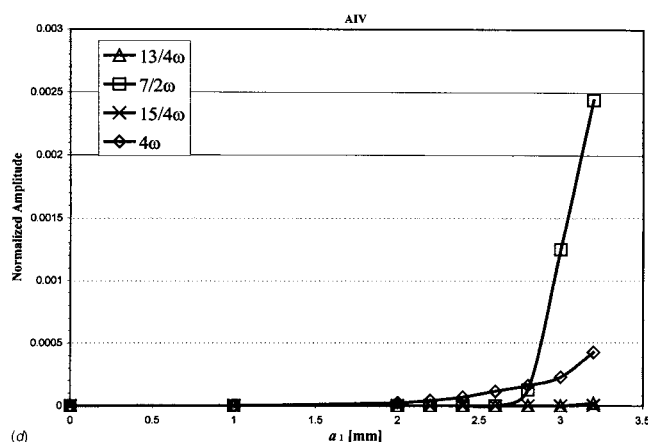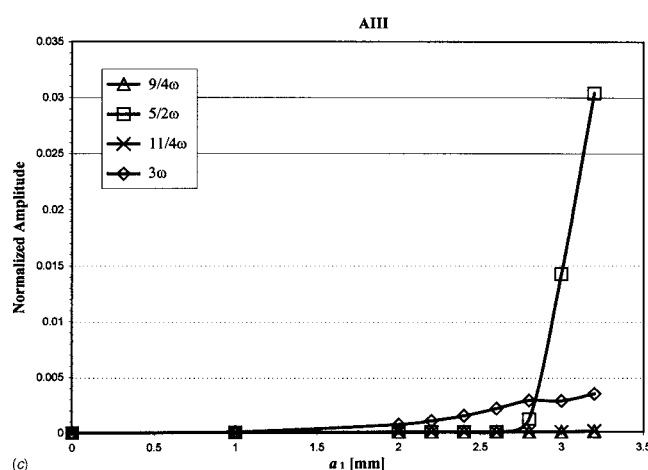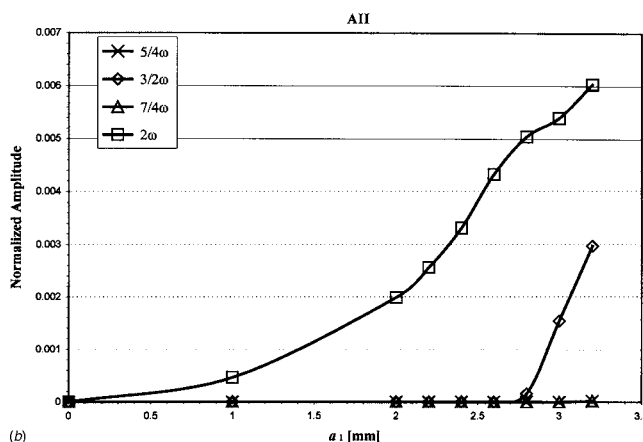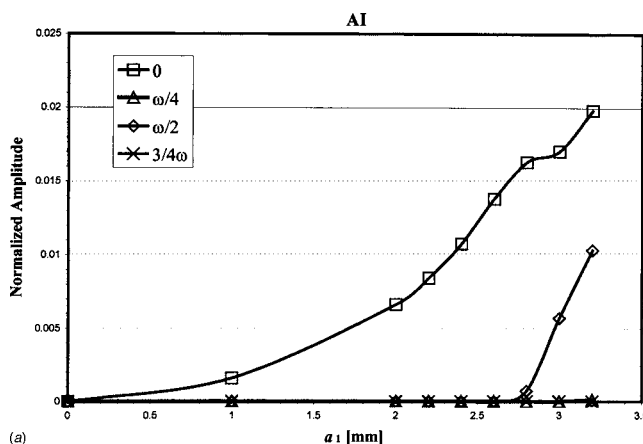
Fig. 2 One crack (localized at one-half of the total length of the beam)—Numerical simulations by varying the crack depth (a) (AI). (b) (AII). (c) (AIII). (d) (AIV).

Table 2 One crack (localized at one-third of the total length of the beam)—Numerical simulations by varying the crack depth (B)

One crack—a_1 variable $d_1=90$ mm; $F=2$N; $f=19$ Hz	
a_1 [mm]	Case
1.0	B1
2.0	B2
2.2	B3
2.4	B4
2.6	B5
2.8	B6
3.0	B7
3.2	B8
3.4	B9
3.6	B10
3.8	B11
4.0	B12
4.2	B13
4.4	B14
4.6	B15
4.8	B16

Table 3 One crack—Numerical simulations by varying the crack position (C)

One crack—d_1 variable $a_1=4.25$ mm; $F=5$N; $f=25$ Hz	
d_1 [mm]	Case
260	C1
240	C2
220	C3
200	C4
180	C5
160	C6
140	C7
120	C8
100	C9
80	C10
60	C11
40	C12

doubling, arise in the dynamic response. In this case, as in the previous one, the frequency of the excitation is around twice the first natural one.

2.3 Parametrical Simulations by Varying the Crack Position (C). These numerical simulations consider one crack with a depth of $a_1=4.25$ mm. They are indicated by the letter C. Referring to Fig. 1, the coordinate d_1 of the crack is assumed variable; the force amplitude is $F=5$N with a frequency f of 25 Hz. The pa-

case. The main difference herein is that we have a higher nonlinearity due to the reduction of the distance between crack and clamp. In addition, in this case we have also considered a crack with a higher depth. If the nonlinearity increases (larger crack depth or lower distance between crack and clamp), other subharmonic components, not necessarily a multiple of that of period

Fig. 3 One crack (localized at one-third of the total length of the beam)—Numerical simulations by varying the crack depth (a) (BI). (b) (BII). (c) (BIII). (d) (BIV).
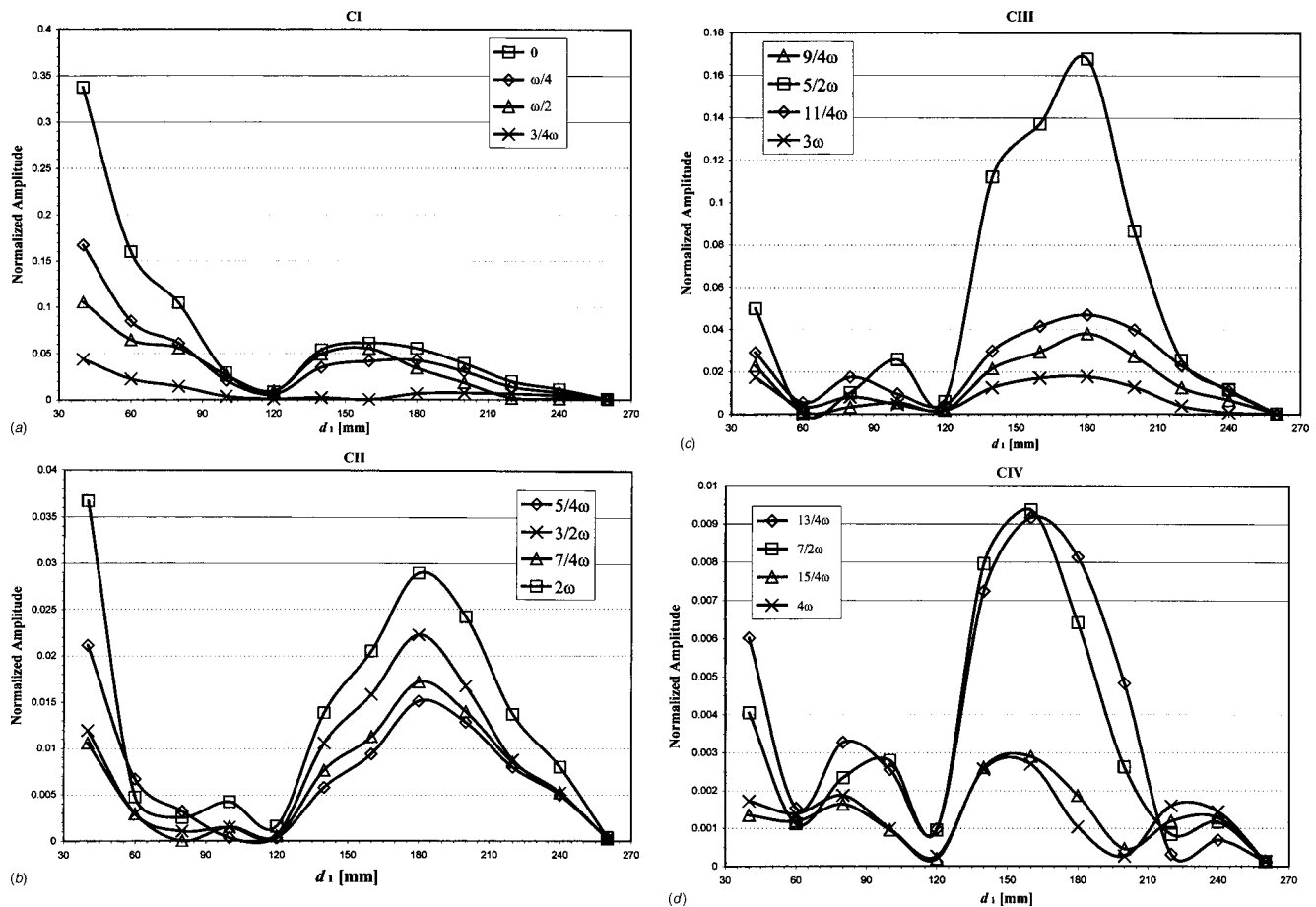
Fig. 4 One crack—Numerical simulations by varying the crack position (*a*) (CI). (*b*) (CII). (*c*) (CIII). (*d*) (CIV).

rameters of the simulations are summarized in Table 3. The numerical responses in terms of the normalized amplitude of Eq. (1) as a function of the crack position, are presented in Fig. 4. These diagrams clearly show that some particular crack positions, corresponding to a linear behavior, can be identified along the beam. These positions correspond to inflexion points in the beam elastic line, where the curvature is zero. In these positions the crack does not breath, so that it does not introduce a nonlinear behavior. For our cases, the inflexion point is between one-half and one-third of the beam length, starting from the clamp. Another inflexion point is clearly shown at the free-end of the beam: A crack placed in the extreme finite element does not change the linear behavior of the structure. This phenomenon can be used to detect the crack position. A real structure can be, in fact, monitored by varying the excitation (typically in terms of frequency). A linear behavior, corresponding to a particular value of the excitation frequency, implies a crack in the inflection point of the elastic line corresponding to that frequency. In the case considered in Fig. 4, the nonlinearity vanishes around the inflexion point corresponding to the second modal shape (consider that the first natural frequency of the undamaged structure is around one-half of that of excitation). In addition, Fig. 4 clearly shows that the nonlinearity increases if the distance between crack and clamp decreases, as previously observed combining simulations A and B. As a matter of fact, the sub-harmonic components can become predominant with respect to the super-harmonic ones.

2.4 Parametrical Simulations by Varying the Excitation Frequency (D). These numerical simulations consider one crack with a depth of $a_1 = 4.25$ mm. They are indicated by the letter D. Referring to Fig. 1, the coordinate of the crack is $d_1 = 90$ mm, the force amplitude is $F = 2$N with a variable frequency f. The param-

eters of the simulations are summarized in Table 4. The numerical responses in terms of the normalized amplitude of Eq. (1) as a function of the excitation frequency are presented in Fig. 5. The most interesting result is that a particular harmonic component

Table 4 One crack—Numerical simulations by varying the amplitude of the excitation (D)

One crack—f variable $a_1 = 4.25$ mm; $d_1 = 90$ mm; $F = 2$N	
f [Hz]	*Case*
2.0	D1
2.5	D2
3.3	D3
4.0	D4
5.0	D5
5.5	D6
6.5	D7
8.5	D8
11.0	D9
12.0	D10
13.0	D11
14.5	D12
15.0	D13
17.0	D14
18.0	D15
19.0	D16
19.5	D17
23.0	D18
25.0	D19
30.0	D20
34.0	D21
38.0	D22
49.8	D23
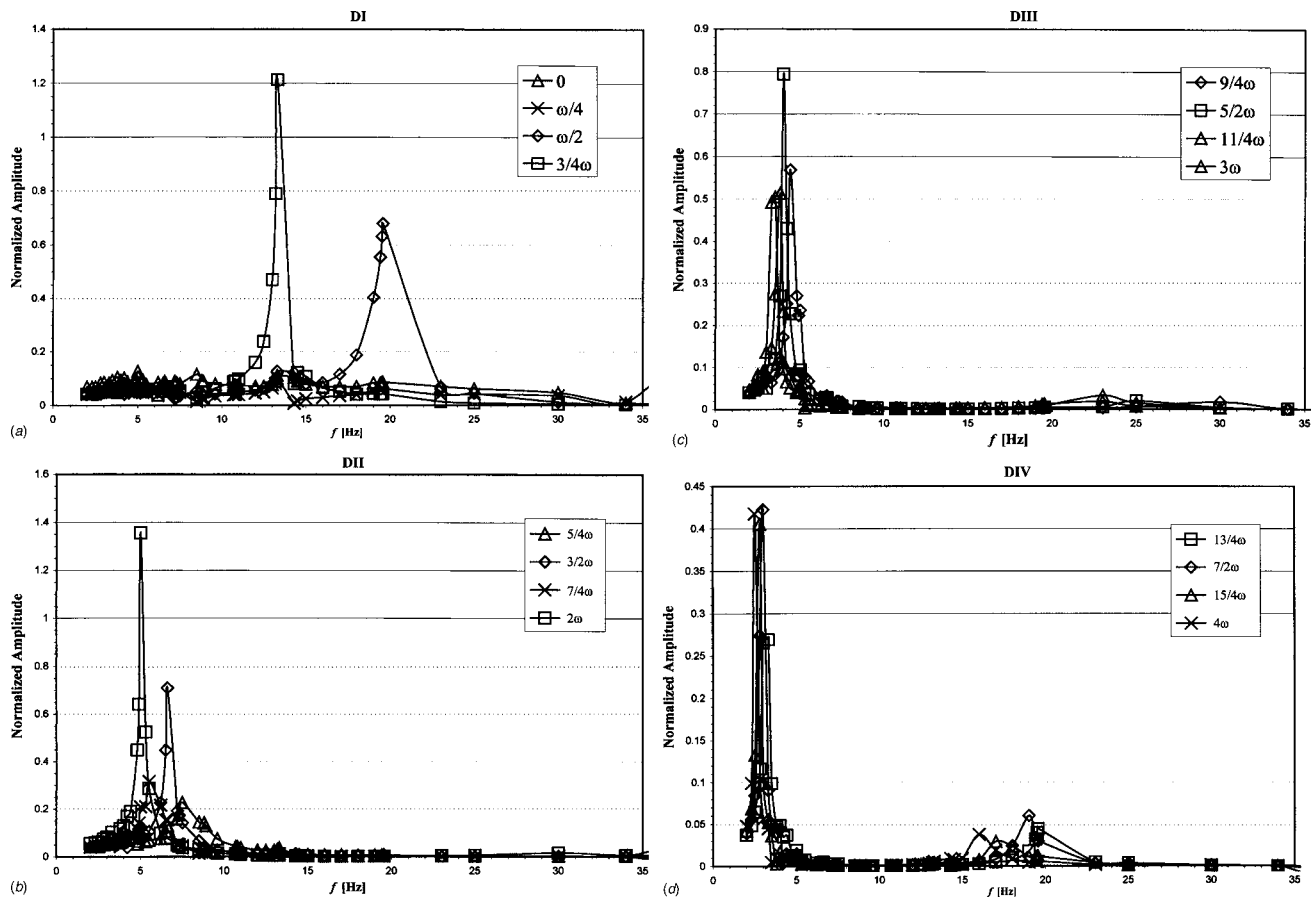
Fig. 5 One crack—Numerical simulations by varying the frequency of the excitation (*a*) **(DI).** (*b*) **(DII).** (*c*) **(DIII).** (*d*) **(DIV).**

becomes predominant in relation to its own resonance. This means that the component of frequency mf ($m=j/\Theta$) becomes predominant when the frequency of the excitation satisfies:

$$f \approx \frac{f_0}{m} \qquad (2)$$

f_0 being the first natural frequency of the damaged structure (in the present case it is around 9.5 Hz). Therefore, the $\frac{3}{4}\omega$ component, for example, goes into resonance around $f \approx 4/3 \times 9.5$ Hz ≈ 13 Hz, as well as the $\frac{1}{2}\omega$ component goes into resonance around $f \approx 2 \times 9.5$ Hz ≈ 19 Hz, according to the numerical results of Fig. 5(*a*). This phenomenon has been observed for all the frequency components.

2.5 Parametrical Simulations by Varying the Excitation Amplitude (E). These numerical simulations consider one crack with a depth of $a_1=4.25$ mm. They are indicated by the letter E. Referring to Fig. 1, the coordinate of the crack is $d_1=90$ mm, the excitation frequency is $f=12$ Hz and the force amplitude F is variable. The parameters of the simulations are summarized in Table 5. The numerical responses in terms of the normalized amplitude of Eq. (1) as a function of the frequency of the excitation are presented in Fig. 6. The results show that the stable solution is the trivial one of linearity with respect to the force amplitude. The existence of this linear solution appears rather obvious, as suggested by the motion equation reported in the companion paper (Part I).

2.6 Parametrical Simulations Considering Two Cracks and Varying the Depth of One of them (F). These numerical simulations consider two cracks, one of depth $a_1=4.25$ mm and the other of variable depth a_2. The simulations are indicated by

the letter F. Referring to Fig. 1, the coordinates of the cracks are $d_1=90$ mm and $d_2=180$ mm, the force amplitude is $F=2$N with a frequency f of 19 Hz. The parameters of the simulations are summarized in Table 6. The numerical responses in terms of the normalized amplitude of Eq. (1) as a function of the crack depth are presented in Fig. 7. According to these diagrams, the nonlinearity seems to be less sensitive with respect to the crack depth a_2. This simply means that the predominant crack is the first one, since it is closer to the clamp. The trend changes only for very high depths a_2.

2.7 Parametrical Simulations Considering Two Cracks and Varying the Position of the Excitation (G). These numerical simulations consider two cracks, both of depth $a_1=a_2=4.25$ mm. The simulations are indicated by the letter G. Referring to Fig. 1, the coordinates of the cracks are $d_1=90$ mm and $d_2=180$ mm, the

Table 5 One crack—Numerical simulations by varying the amplitude of the excitation (E)

One crack—F variable $a_1=4.25$ mm; $d_1=90$ mm; $f=12$ Hz	
F [N]	*Case*
2.0	E1
1.0	E2
0.5	E3
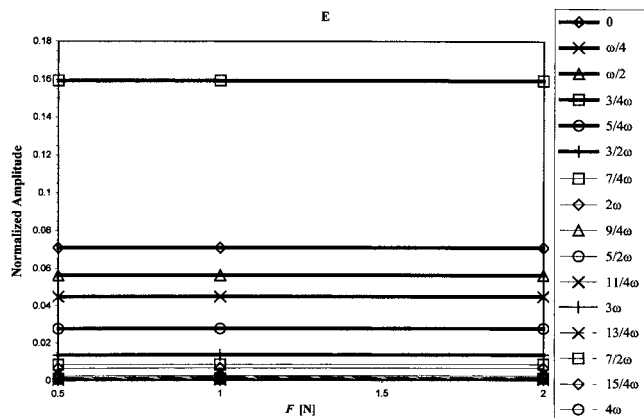
Fig. 6 One crack—Numerical simulations by varying the amplitude of the excitation (E)

force amplitude is $F=2$N, with a frequency f of 19 Hz and the position of the excitation d_F is variable. The parameters of the simulations are summarized in Table 7. The numerical responses in terms of the normalized amplitude of Eq. (1) as a function of the position of the excitation are presented in Fig. 8. According to these diagrams, the nonlinearity presents a clear transition between the two cracks, larger near the first than near the second one. The largest nonlinearities arise for values of the force position between the first crack and the clamp. A very interesting

result is that the stronger nonlinearity appears for excitations near the position of the predominant crack. It is important to note that the trend does not change substantially when the force position is closer to the clamp (consider that these are contributions normalized with respect to the linear one, see Eq. (1), so that they do not vanish near the clamp).
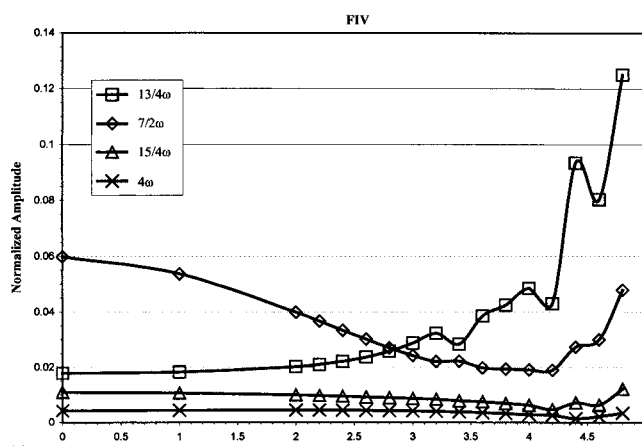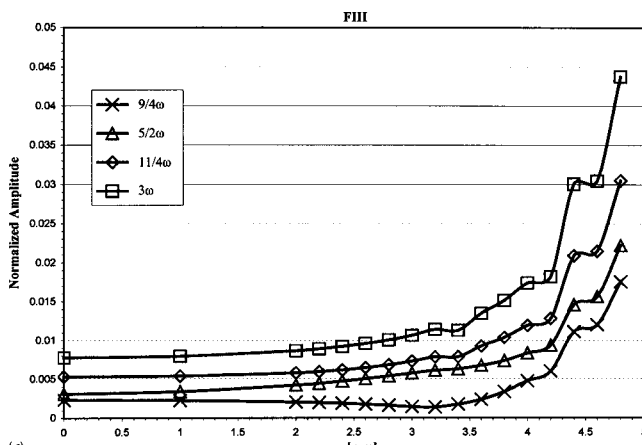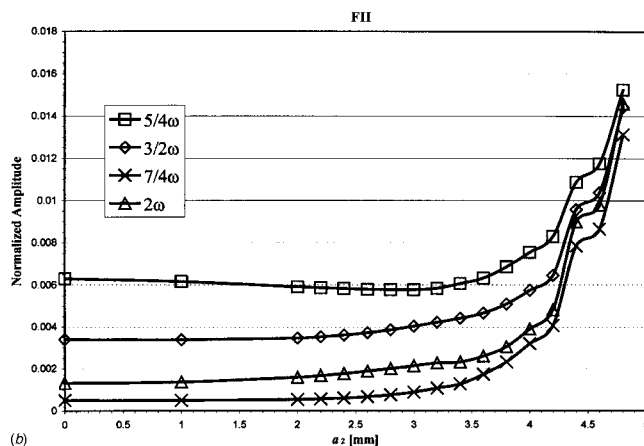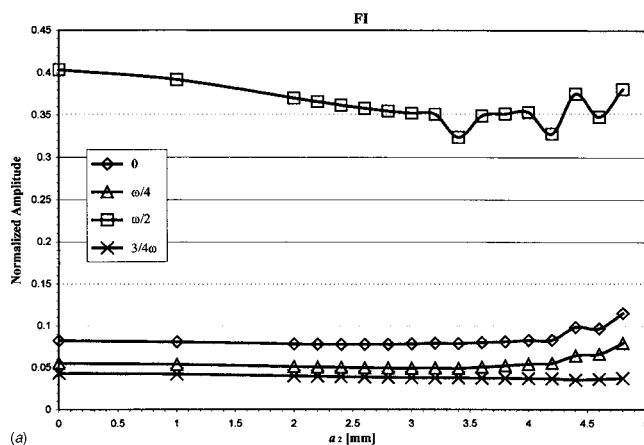
Fig. 7 Two cracks—Numerical simulations by varying the depth of one crack (*a*) (FI). (*b*) (FII). (*c*) (FIII). (*d*) (FIV).

Table 7 Two cracks—Numerical simulations by varying the position of the excitation (G)

d_F [mm]	Case
Two cracks—d_F variable $a_1=a_2=4.25$ mm; $d_1=90$ mm; $d_2=180$ mm; $F=2$N; $f=19$ Hz	
270.0	G1
189.0	G2
135.0	G3
94.5	G4
13.5	G5

3 General Discussion

The theoretical and numerical approach presented in the companion paper (Part I) appears very useful in the study of highly nonlinear forced vibrations for damaged structures. It permits us to take into account the interaction of several breathing cracks. In the case of high nonlinearity, the super-harmonic frequency components become insufficient to catch the real behavior of the structure. As a consequence (offset and) sub-harmonic components must be taken into account. One example is given by the period doubling phenomenon, recently experimentally observed and discussed in both Parts I and II. The extensive parametrical simulations, presented in Part II, have been performed by varying all the main parameters influencing the dynamic behavior of the structure: The number of cracks, their depth and position, as well as the amplitude, frequency, and position of the excitation.

The results can be summarized as follows:

(1) For a weak nonlinearity, we have to take into account in the structural response not only the super-harmonic frequency components but also an offset (zero-frequency);

(2) for a stronger nonlinearity, we have to take into account not only the super-harmonic frequency and offset components but also the sub-harmonic ones (complexity and transition to deterministic chaos);

(3) if a crack implies, as a particular case, a linear behavior of the structure, we can conclude that the crack position is close to an inflection point of the elastic line corresponding to the excitation frequency. This result could be useful to improve the techniques of vibration-based inspection;

(4) the nonlinearity increases if the position of the excitation becomes closer to the crack (or, obviously, if the crack position becomes closer to the clamp, or if the crack depth becomes larger);

(5) the component of frequency mf becomes predominant in correspondence of its own resonance, when the frequency of the excitation satisfies $f \approx f_0/m$;

(6) the behavior becomes linear with respect to the amplitude of the excitation.
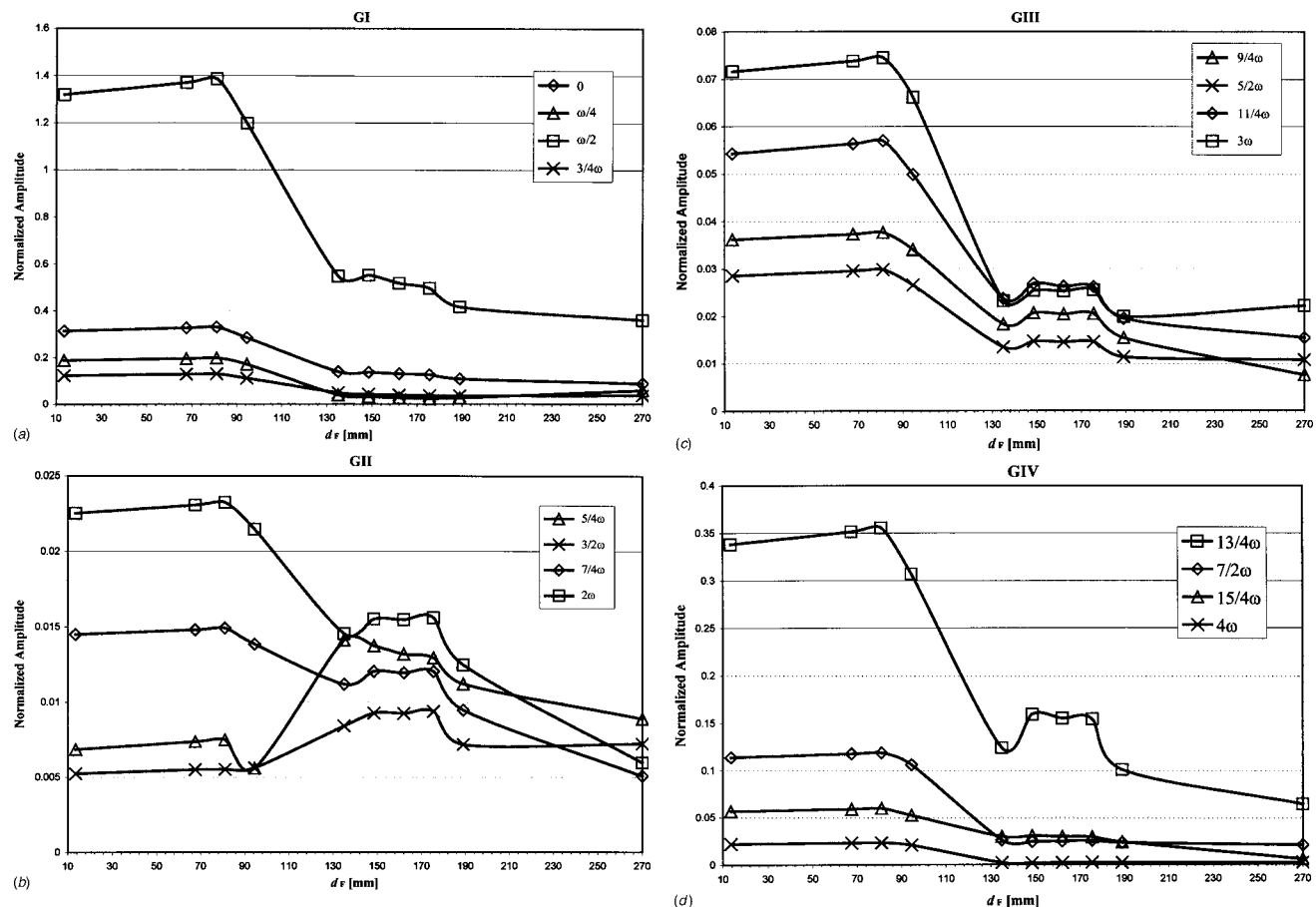
Fig. 8 Two cracks—Numerical simulations by varying the position of the excitation (a) (GI). (b) (GII). (c) (GIII). (d) (GIV).

References

[1] Zastrau, B., 1985, "Vibrations of Cracked Structures," Arch. Mech., **37**, pp. 731–743.

[2] Collins, K. R., Plaut, P. H., and Wauer, P. H., 1992, "Free and Forced Longitudinal Vibrations of a Cantilevered Bar With a Crack," J. Vibr. Acoust., **114**, pp. 171–177.

[3] Ibrahim, A., Ismail, F., and Martin, H. R., 1987, "Modeling of Dynamics of Continuous Beam Including Nonlinear Fatigue Crack," Int. J. of Experimental Modal Analysis, **2**, pp. 76–82.

[4] Ismail, F., Ibrahim, A., and Martin, H. R., 1990, "Identification of Fatigue Cracks From Vibration Testing," J. Sound Vib., **140**, pp. 305–317.

[5] Qian, G. L., Gu, S. N., and Jiang, J. S., 1990, "The Dynamic Behavior and Crack Detection of a Beam With a Crack," J. Sound Vib., **138**, pp. 233–243.

[6] Shen, M.-H. H., and Chu, Y. C., 1992, "Vibrations of Beams With a Fatigue Crack," Comput. Struct., **45**, pp. 79–93.

[7] Chu, Y. C., and Shen, M.-H. H., 1992, "Analysis of Forced Bilinear Oscillators and the Application to Cracked Beam Dynamics," AIAA J., **30**, pp. 2512–2519.

Newmark's Time Integration Method From the Discretization of Extended Functionals

Lorenzo Bardella

Francesco Genna
e-mail: genna@bscivgen.ing.unibs.it

Department of Civil Engineering,
University of Brescia,
Via Branze, 38,
25123 Brescia, Italy

In this note we illustrate how to obtain the full family of Newmark's time integration algorithms within a rigorous variational framework, i.e., by discretizing suitably defined extended functionals, rather than by starting from a weak form (for instance, of the Galerkin type), as done in the past. The availability of functionals as a starting point is useful both as a tool to obtain new families of time integration methods, and as a theoretical basis for error estimates. To illustrate the first issue, here we provide some examples of how to obtain modified algorithms, in some cases significantly more accurate than the basic Newmark one despite having a comparable computational cost. [DOI: 10.1115/1.1934648]

1 Introduction

This work aims at giving an application to some ideas, presented for the first time in [1], concerning the variational basis of time integration algorithms for the initial-value structural dynamics problem.

In recent times, the term "variational approach" has assumed a somewhat slack meaning, including techniques, most notably the Galerkin or Hamilton/Ritz ones, in which the basic equations (usually in weak form) are written as the variation of something which is never defined. The use of these strategies is a forced consequence of the lack of a potential function for the problem to be solved; nevertheless, it has been very successful, to the point of extending the meaning of the term "variational" to something that is not truly variational.

Since the work described in [2–4] has shown that every mathematical problem can be given a rigorous variational formulation, we believe that such a terminology would be better maintained in its proper sense. Therefore, here we use the expression "variational formulation" solely to denote the statement of a problem in terms of the stationarity of an explicitly defined functional.

The variational framework adopted in [1] is described by the theory developed in [2–4], whose main results can be summarized as follows. Let us write a generic problem (linear or nonlinear) in the following form:

$$\mathcal{N}(x) - P = 0, \tag{1.1}$$

where $\mathcal{N}(\cdot)$ denotes an operator, in general nonpotential (i.e., which cannot be seen as the gradient of a functional), containing all the equations governing the problem (i.e., all field equations and both boundary and initial conditions), whereas x and P indicate, respectively, an unknown function and a given known term, both possibly vector or tensor valued. Starting from the theory presented for the first time in [2], Ref. [3] discusses a general way to construct variational formulations of problem (1.1), by splitting operator $\mathcal{N}(\cdot)$ into a linear, symmetric, positive definite part $\mathcal{S}(\cdot)$, hereafter called *kernel*, and a residual part $\mathcal{R}(\cdot)$

$$\mathcal{N}(\cdot) = \mathcal{S}(\cdot) + \mathcal{R}(\cdot) \tag{1.2}$$

The kernel $\mathcal{S}(\cdot)$ must be symmetric with respect to a nondegenerated bilinear functional [2] indicated by the symbol $\langle . , . \rangle$, i.e., such that the following identity holds:

$$\langle x, \mathcal{S}y \rangle = \langle \mathcal{S}x, y \rangle \quad \forall x, y \tag{1.3}$$

Problem (1.1) always admits an *extended* [2,5] variational formulation, governed by the following functional:

$$F[x,y] = \langle x - y, \mathcal{N}(x) - P \rangle - \frac{1}{2} \langle x - y, \mathcal{S}(x - y) \rangle \tag{1.4}$$

It has been proved [4] that the critical points of functional (1.4) are solutions of Eq. (1.1) (and vice versa) and that $y = x$ at any critical point of Eq. (1.4).

The general theory developed in [2,3] requires the definition of a linear, invertible kernel $\mathcal{S}(\cdot)$. The range of admissible kernels has subsequently been widened, first in [4] and, very recently, in [5]. In particular, Brun et al. [4] have shown that the choice $\mathcal{S}(\cdot) = 0$ leads to the construction of a meaningful extended functional associated with problem (1.1), that reads

$$G[x,y] = \langle x - y, \mathcal{N}(x) - P \rangle \tag{1.5}$$

In [1], it has been shown that all the Galerkin or Hamilton/Ritz-based algorithms for numerical time integration can be derived from the stationarity of a functional of the type (1.5) with respect to the auxiliary unknown y; such a stationarity equation, in fact, formally reads

$$\delta_y G[x,y] = -\langle \delta y, \mathcal{N}(x) - P \rangle = 0 \tag{1.6}$$

which can be seen as the Petrov–Galerkin weak form of problem (1.1) [5], where no restriction is placed on the choice of the bilinear functional, and the variation δy assumes the rôle of a weight function.

All the traditional weak form approaches do indeed start from an equation of the form (1.6), and proceed by discretizing the main unknown function x, and by defining the *weight* function δy. We wish to remark that, in general, different results would be obtained by approaching the problem at the level of the functional (1.5), i.e., by discretizing directly the auxiliary unknown y by means of *test* functions, instead of defining its variation δy. One of the main features of the theory discussed in the present work is indeed this more "consistent" treatment of the discretization of main and auxiliary unknowns, possible only if a functional is available.

Previous work done on the application of this theory to the equations of structural dynamics [1,6,7] has never been concerned with the discretization of functionals F and G to obtain known time integration methods; here, we focus our attention on Newmark's method [8]. This family of time integration algorithms was originally derived in a sort of "empirical" way, but has subsequently been given a variety of theoretical fundaments, both as a Taylor-series expansion and as a recurrence equation deriving from the discretization of a Galerkin-type weak form (see [9] for a review of this field). Some special cases of the Newmark family have also been obtained and discussed in [10] by starting from a Hamilton approach and without using shape functions. Owing to the absence of classical extremum principles governing the initial-value structural dynamics problem, Newmark's method (as well as all the existing others) has never been derived from the stationarity of a functional.

Here, we wish to show how the theory developed in [1], briefly summarized here above, does indeed allow one to recast Newmark's method into a rigorous variational framework, which might prove useful from both the theoretical and the practical viewpoints. For instance, the availability of functionals might allow the derivation of new algorithms in which some desirable feature—energy conservation in the nonlinear range is the first which comes to the mind—is prescribed a priori as a constraint in the stationarity process.

We first discuss how to derive Newmark's method from the discretization of extended functionals of the type (1.5). Next, we will illustrate, by means of examples, some ways to exploit the new theoretical framework available, in order to obtain new time integration algorithms, some of which appear to have an accuracy far superior to that of the basic Newmark method.

2 Discretization of Extended Functionals to Obtain the Newmark Method

The equations describing the Newmark method are [8]

$$u_{n+1} = u_n + \Delta t v_n + \frac{\Delta t^2}{2}[(1 - 2\beta)a_n + 2\beta a_{n+1}] \qquad (2.1a)$$

$$v_{n+1} = v_n + \Delta t[(1 - \gamma)a_n + \gamma a_{n+1}] \qquad (2.1b)$$

u, v, and a indicating displacement, velocity, and acceleration fields, respectively, and $\Delta t = t_{n+1} - t_n$ denoting the considered time step. From here on, for simplicity and without loss of generality, we will always set $t_n = 0$ and $t_{n+1} = \Delta t$.

A first important observation is that, owing to the mutual independence of the algorithmic parameters β and γ, in order to obtain these equations variationally one must take a two-field approach in writing the equations of motion, in which the velocity is treated as an independent variable. Therefore, we write the equations of motion, for simplicity in terms of a single degree of freedom, for $0 \leq t \leq \Delta t$, as

$$m\dot{v}(t) + cv(t) + ku(t) - f(t) = 0; \qquad (2.2a)$$

$$v(t) - \dot{u}(t) = 0; \qquad (2.2b)$$

$$v(0) - v_n = 0; \qquad (2.2c)$$

$$-u(0) + u_n = 0 \qquad (2.2d)$$

where m indicates the mass, c the damping, and k the elastic stiffness. In the classical Newmark approach, Eq. (2.2a), evaluated at time $t = t_{n+1}$ and in which $\dot{v}(t)$ is replaced with $a(t)$, is combined with Newmark's update Eqs. (2.1), in order to obtain displacements, velocities, and accelerations at the end of the time step.

Turning now to the issue of choosing a functional to be discretized, in order to obtain Newmark's method, observe that formulas (2.1) define time continuous displacements and velocities.

Carini and Genna [1] provide several functionals of the type (1.5) for the structural dynamics equations, based on different choices of the bilinear functional. In view of our purpose and of the preceding observations, here we will make reference only to a particularization of one of them (called G_1^a in [1]), namely to a *two-field, time continuous* functional, i.e., one in which displacement and velocity are independent variables, and the initial conditions (2.2c) and (2.2d) are prescribed in strong form. By setting $x = \{u(t), v(t)\}$ as the real unknown vector, whereas $y = \{u_s(t), v_s(t)\}$ contains the corresponding auxiliary unknowns, functional G of Eq. (1.5) then reads

$$G[x, y] = G[u(t), v(t), u_s(t), v_s(t)]$$
$$= \int_0^{\Delta t} [u(t) - u_s(t)][m\dot{v}(t) + cv(t) + ku(t) - f(t)]dt$$
$$+ \int_0^{\Delta t} m[v(t) - v_s(t)][v(t) - \dot{u}(t)]dt \qquad (2.3)$$

where, without altering the starting problem, we have inserted the mass coefficient m into the second integral, for dimensional consistency.

The possible discretization leading to Newmark's Eqs. (2.1) is suggested by the inspection of the stationarity conditions of functional (2.3) with respect to the auxiliary unknown functions u_s and v_s, and by their comparison with the target Eqs. (2.1). The following requirements are obtained, with specific reference to the discretization of functional (2.3):

1. the shape functions, approximating the real displacement and velocity fields, can be taken as linear functions of time, each starting from the respective initial values;
2. the test functions [the discretization of the auxiliary displacement and velocity fields $u_s(t)$ and $v_s(t)$] *must exhibit coupling between the unknown coefficients of u_s and those of v_s*, and there is no need for their continuity over time steps;
3. β and γ must not appear in the definition either of the shape or of the test functions; the calculations are simpler, and more similar to others of analogous approaches [9], if one lets them appear in the test functions only; and
4. the forcing term must be discretized as a linear function of time.

Thus, we are led to writing the following sets of approximating functions (where a superimposed hat indicates unknown constant coefficients).

- Main unknowns (shape functions):

$$u(t) = u_n + \frac{\hat{u}_{n+1} - u_n}{\Delta t}t; \qquad (2.4a)$$

$$v(t) = v_n + \frac{\hat{v}_{n+1} - v_n}{\Delta t}t \qquad (2.4b)$$

- Auxiliary unknowns (test functions):

$$u_s(t) = U_u(t, \beta, \gamma)\hat{u}_s + U_v(t, \beta, \gamma)\hat{v}_s;$$
$$v_s(t) = V_u(t, \beta, \gamma)\hat{u}_s + V_v(t, \beta, \gamma)\hat{v}_s \qquad (2.5)$$

where U_u, U_v, V_u, and V_v are functions of time to be identified.

- Forcing term:

$$f(t) = f_n + \frac{f_{n+1} - f_n}{\Delta t}t \qquad (2.6)$$

The primary unknown coefficients $\hat{u}_{n+1}$ and $\hat{v}_{n+1}$ are computed by plugging Eqs. (2.4)–(2.6) into functional (2.3) and solving its sta-

tionarity equations with respect to the auxiliary unknown coefficients $\hat{u}_s$ and $\hat{v}_s$. By matching these stationarity equations with Eqs. (2.1) we can identify the functions U_u, U_v, V_u, and V_v, which, finally, have to satisfy the following relationships:

$$\int_0^{\Delta t} U_u dt = \Delta t; \tag{2.7a}$$

$$\int_0^{\Delta t} U_u t dt = \gamma \Delta t^2; \tag{2.7b}$$

$$\int_0^{\Delta t} U_v dt = \frac{\Delta t^2}{2}; \tag{2.8a}$$

$$\int_0^{\Delta t} U_v t dt = \beta \Delta t^3; \tag{2.8b}$$

$$\int_0^{\Delta t} V_u dt = 0; \tag{2.9a}$$

$$\int_0^{\Delta t} V_u t dt = 0; \tag{2.9b}$$

$$\int_0^{\Delta t} V_v dt = -\Delta t; \tag{2.10a}$$

$$\int_0^{\Delta t} V_v t dt = -\frac{\Delta t^2}{2} \tag{2.10b}$$

Since we have eight conditions to identify the unknown functions U_u, U_v, V_u, and V_v, the simplest possible choice, among the infinite ones, is to define them as all linear in time:

$$U_u = k_1 + k_2 t; \quad U_v = k_3 + k_4 t; \quad V_u = k_5 + k_6 t; \quad V_v = k_7 + k_8 t \tag{2.11}$$

Thereafter, one can compute all the coefficients k_i, thus coming to the following test functions to be plugged into functional (2.3):

$$u_s(t) = \left[4 - 6\gamma + 6(2\gamma - 1)\frac{t}{\Delta t}\right]\hat{u}_s$$
$$+ [2(3\beta - 1)\Delta t - 3(4\beta - 1)t]\hat{v}_s \tag{2.12a}$$

$$v_s(t) = \hat{v}_s \tag{2.12b}$$

The discretization defined by Eqs. (2.4), (2.6), and (2.12), inserted into functional (2.3), leads to Eqs. (2.1).

It is worth pointing out that the time derivative of the discretized velocity $v(t)$ in Eq. (2.4b), i.e., the constant expression $(\hat{v}_{n+1} - v_n)/\Delta t$, turns out to be precisely Newmark's result $(1 - \gamma)a_n + \gamma a_{n+1}$, and therefore it is *not* the final acceleration a_{n+1}, as computed by inserting u_{n+1} and v_{n+1} into Eq. (2.2a) and inverting it. This is not a contradiction, but just a consequence of the fact that the variational procedure enforces the validity of both equations of motion (2.2) in *weak* form.

In conclusion, essential requisites to obtain the Newmark method from a variational formulation are (i) the two-field and time continuous approach, and, at least when starting from functional (2.3), (ii) the linear approximations for both the main unknown fields and the forcing term, and (iii) the coupling between the two test functions in terms of unknown parameters. Even though we have been able to obtain several particular Newmark algorithms (with fixed values of β and γ, or, for instance, for an undamped motion) by means of different, often simpler choices,

we have never been able to generalize them, without satisfying all the afore said requisites (i–iii), to the full Newmark family.

Finally, we wish to point out that, owing to the existence of multiple solutions to Eqs. (2.7)–(2.10), but also owing to the possibility of using slightly different strategies altogether, there are multiple ways to obtain time integration schemes sharing exactly the same features and the same algorithmic properties with Newmark's method. The path indicated in this section, albeit possibly the simplest one, is not unique. We have indeed found other ways to derive Newmark's method from the discretization of extended functionals, but here we do not expand further on this aspect for the sake of brevity.

3 Examples of Application of the New Theoretical Framework

Here, we wish to illustrate, by means of some examples, how the availability of a functional governing the problem under study can help in obtaining numerical methods with possibly improved features. Indeed, the possibility of working with test and shape functions which have the same "dignity" (as hinted in the Introduction) allows one to start from the already excellent Newmark family of algorithms and try to improve on them in a rather simple way, reasoning by analogy.

We will present three methods. First, we will just increase the number of unknown parameters in the basic discretization defined by Eqs. (2.4), (2.6), and (2.12). Then, in the second method we will maintain the same number of unknowns, but we will modify the discretization of the main unknowns according to the suggestion provided by Eqs. (2.12). In both these methods we will search for maximum accuracy. In the third one, we will illustrate the effects of preconditioning kernels $\mathcal{S}(\cdot)$ on the stability of the resulting algorithms. Both linear and nonlinear examples will be shown.

3.1 A Modification of Newmark's Method by Increasing the Number of Unknowns. The simplest idea which comes to mind, in order to improve the accuracy of Newmark's method starting from the theory proposed, is to modify the discretization defined by Eqs. (2.4), (2.6), and (2.12) by increasing the number of unknown parameters. This could be done in several ways, as long as the main unknown parameters are as many as the auxiliary ones, in such a way as to maintain the uncoupling of the stationarity equations (see also [1,4,5]) and being thus able to solve for the main unknowns only. An improved algorithm, at least in terms of accuracy, is defined by the following approximation:

- main unknowns (shape functions)

$$u(t) = u_n + \frac{\hat{u}_{n+1} - u_n}{\Delta t}t + \hat{b}t^2; \quad v(t) = v_n + \frac{\hat{v}_{n+1} - v_n}{\Delta t}t \tag{3.1}$$

- auxiliary unknowns (test functions)

$$u_s(t) = \left[4 - 6\gamma + 6(2\gamma - 1)\frac{t}{\Delta t}\right]\hat{u}_s$$
$$+ [2(3\beta - 1)\Delta t - 3(4\beta - 1)t]\hat{v}_s \tag{3.2a}$$

$$v_s(t) = \hat{v}_s + \hat{b}_s t \tag{3.2b}$$

Here, $\hat{b}$ and $\hat{b}_s$ represent two new unknown parameters; the values of the main unknowns $\hat{u}_{n+1}$, $\hat{v}_{n+1}$, and $\hat{b}$ are obtained by solving the stationarity equations of functional (2.3) with respect to $\hat{u}_s$, $\hat{v}_s$, and $\hat{b}_s$. The cost, with respect to what is required by the standard

Algorithm comparison for undamped motion ($\beta = 1/6$, $\gamma = 1/2$)

$m = 1$; $c = 0$; $k = 1$; $u_0 = 1$; $v_0 = 0$; $\Delta t = 1/2$

Legend:
— Exact
—●— Newmark "augmented"
—△— Newmark "improved"
—□— Standard Newmark

Fig. 1 Integration of undamped, unforced linear motion with $m=1$, $k=1$. Comparison of standard Newmark, "augmented" Newmark (Sec. 3.1), and "improved" Newmark (Sec. 3.2) solutions, all with $\beta=1/6$, $\gamma=1/2$, $\Delta t=0.5$.

Newmark method, is of course significantly higher.

Appendix A reports the amplification matrix[1] of this family of methods, as well as of those defined in the next two sections.

The family of algorithms generated by means of Eqs. (3.1) and (3.2) requires $\gamma=1/2$ for second-order accuracy. If $\beta=1/6$ it is possible to get third-order accuracy only in the undamped case, which is, however, an improvement with respect to Newmark's method, which never does go beyond second-order accuracy. We have also found that, by adding only one unknown to the basic Newmark discretization, the resulting algorithms tend to become unstable, or to acquire poor characteristics in terms of energy dissipation. The special case defined by Eqs. (3.1) and (3.2) is indeed the best we could obtain, and some results of its application are shown in Figs. 1 and 2, together with those given by the algorithm illustrated in the next section.

Only switching to the even more expensive choice of four unknown parameters for both shape and test functions it is possible to significantly improve in all respects on the basic scheme. In any event, the availability of the framework developed in this work allows one to easily generate new families of algorithms which could also be tailored to suit several specifications (accuracy, stability, energy dissipation), by following standard techniques.

3.2 A Modification of Newmark's Method With the Same Unknowns. The observation of the most characteristic feature of Eqs. (2.12)—the coupling between the discretizations of the aux-

---

[1]The amplification matrix **A** of a time integration algorithm relates quantities at the end of one step to quantities at the beginning of the next one, under initial conditions only (i.e., without considering forcing terms), as follows:

$$\begin{bmatrix} u_{n+1} \\ v_{n+1} \end{bmatrix} = \mathbf{A} \begin{bmatrix} u_n \\ v_n \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} u_n \\ v_n \end{bmatrix}$$

iliary velocity and displacement—suggests a way to increase the polynomial order of the main unknowns, from what is required in Newmark's method (Eq. (2.4)), *without increasing the number of unknowns*. Among various ways of proceeding, we have found the following one to furnish interesting results:

• main unknowns (shape functions)

$$u(t) = u_n + \frac{\hat{u}_{n+1} - u_n}{\Delta t}t + (h_1\gamma + h_2\beta)\frac{\hat{v}_{n+1} - v_n}{\Delta t}t^2;$$

(3.3a)

$$v(t) = v_n + \frac{\hat{v}_{n+1} - v_n}{\Delta t}t \qquad (3.3b)$$

• auxiliary unknowns (test functions) as in Eqs. (2.12).

Here, we have introduced a coupling between the unknown velocity and displacement similar to that already necessary *in the test functions only* to obtain the standard Newmark method, and have also introduced Newmark's algorithmic parameters β and γ into the shape functions; h_1 and h_2 are further algorithmic parameters.

In Eqs. (3.3) the number of main unknown coefficients is the same—two—as in the basic Newmark method, which leaves this modification roughly as expensive as the basic Newmark scheme. Even though the approach illustrated in this paper seems to imply the doubling of the unknown variables, with respect to what is required by a standard approach (i.e., one simply defined by Eqs. (2.1)), it is possible to show that, both for the standard Newmark case (as obvious) and for the "improved" one illustrated in this section, even in the multi-degrees of freedom (DOF) case, the main computational cost entailed at each integration step corresponds to the solution of a linear system of equations whose size is equal to the number of degrees of freedom of the system. Ap-

Algorithm comparison for damped motion ($\beta = 1/6$, $\gamma = 1/2$)

m = 1; c = 0.04; k = 1; u_0 = 1; v_0 = 0; Δt = 0.02
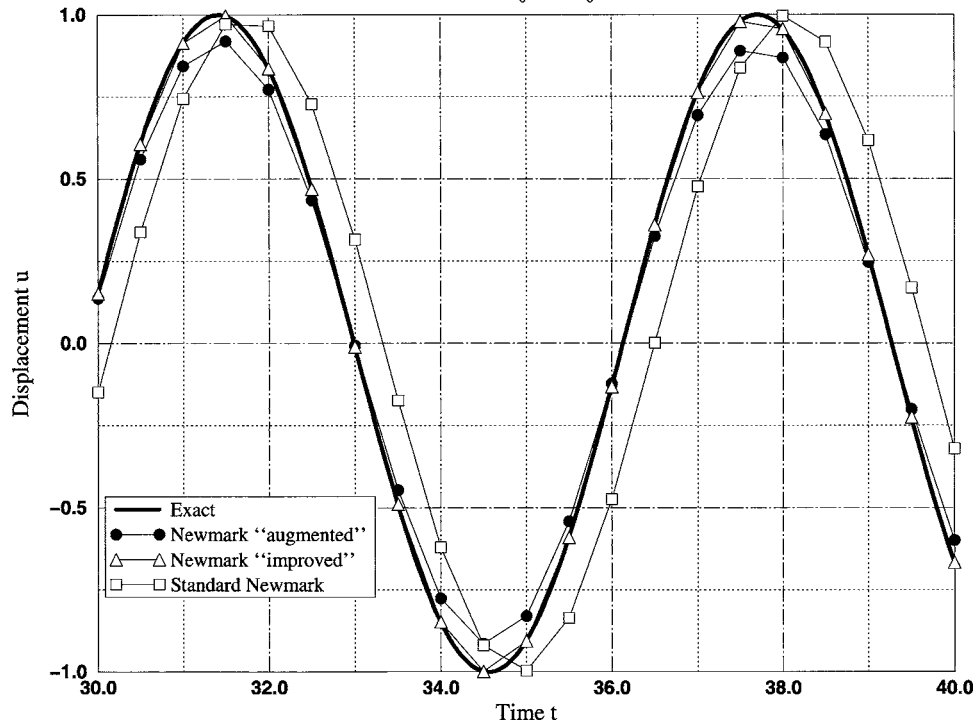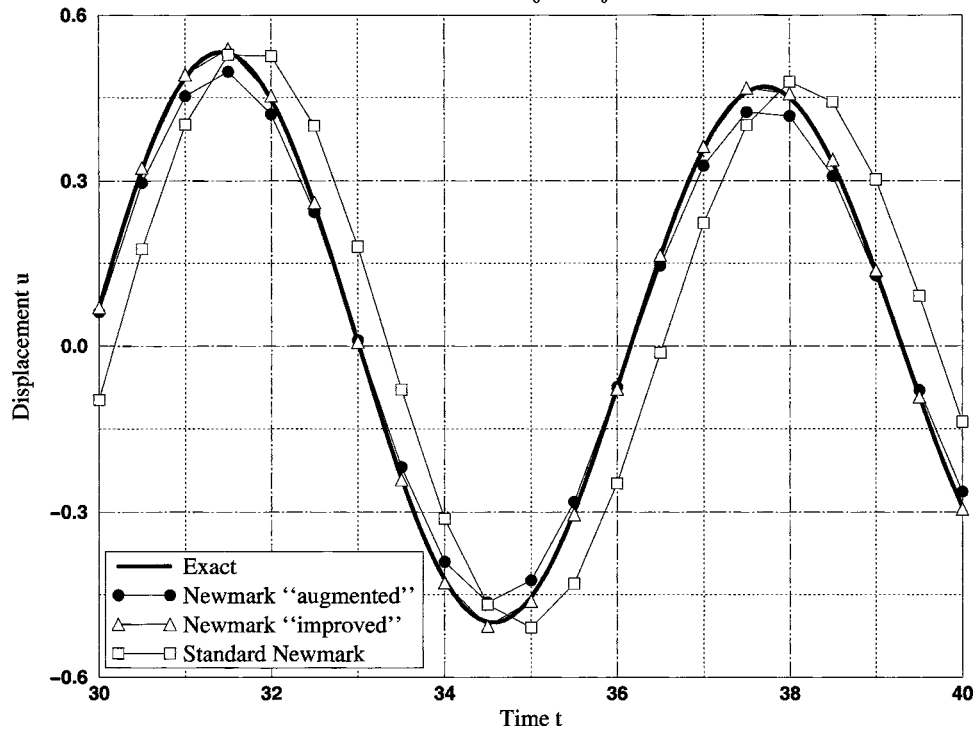
Fig. 2 Integration of damped, unforced linear motion with $m=1$, $k=1$, $c=0.04$. Comparison of standard Newmark, "augmented" Newmark (Sec. 3.1), and "improved" Newmark (Sec. 3.2) solutions, all with $\beta=1/6$, $\gamma=1/2$, $\Delta t=0.5$.

pendix B, in which, for the sake of clarity and completeness, we report the calculations required by our approach for a general, multi-DOF system, will also sketch one possible way to illustrate this assertion.

We have optimized the resulting family of algorithms in terms of accuracy only, with the following results: $\gamma=1/2$ is required for second-order accuracy regardless of the value of all the other parameters; if $\beta=1/6$ and $h_2=3(1-h_1)$ then, in the undamped case only, the algorithm is fourth-order accurate. With these parameters its accuracy is anyway far superior to that of the basic Newmark method, as shown by the numerical results of Figs. 1 and 2, which refer precisely to this set of parameter values, for all the three methods considered (basic Newmark, "augmented" Newmark as in the previous section, and "improved" Newmark as defined here).

Figure 1 refers to the undamped case and Fig. 2 to a damped one; in both cases it is apparent that the "improved" algorithm is quite effective; the "augmented" one, more expensive, is affected by a significant numerical damping, even though its phase error is smaller than that of the basic Newmark method. With this set of parameters, the "improved" method is unconditionally stable, whereas, interestingly, the basic Newmark is only conditionally stable. We finally observe that, with the chosen set of parameters, the added quadratic term in the discretization of the displacements [Eq. (3.3a)] turns out to be exactly the time integral of the linear term in the velocity.

3.3 Effects of Preconditioning Kernels: An Example for the Linear Case. We now consider the effect of the inclusion of a kernel $\mathcal{S}(\cdot)$ into an extended functional governing problem (2.2). Our purpose is to see how the inclusion of a kernel, in the basic extended functional, *this time keeping otherwise fixed the adopted*

discretization, modifies Newmark's method, as obtained in Sec. 2.

For the sake of simplicity, we set $m=1$, $c=0$, and $k=\omega^2$ in Eq. (2.2a), and rewrite Eqs. (2.2a) and (2.2b) in the following, handier matrix form:

$$\begin{bmatrix} \omega^2 & \dfrac{d}{dt} \\[2mm] -\dfrac{d}{dt} & 1 \end{bmatrix} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \begin{bmatrix} f(t) \\ 0 \end{bmatrix} \tag{3.4}$$

and, as done before, we assume to satisfy a priori the initial conditions (2.2c) and (2.2d). With the aim of writing down an extended functional of the type (1.4), the simplest possible symmetric kernel $\mathcal{S}(\cdot)$ is the following linear one:

$$\mathcal{S}(y) = \mathcal{S}(u_s, v_s) = \begin{bmatrix} \omega^2 \ell_1 & \ell_2 \\ \ell_2 & \ell_3 \end{bmatrix} \begin{bmatrix} u_s(t) \\ v_s(t) \end{bmatrix} \tag{3.5}$$

where the constants ℓ_1, ℓ_2, and ℓ_3 are free algorithmic parameters.

The extended functional of the type (1.4), with the choice (3.5) for the kernel, reads:

$$F[x,y] = F[u(t),v(t),u_s(t),v_s(t)] = \int_0^{\Delta t} (u - u_s)(\dot{v} + \omega^2 u - f)dt$$

$$+ \int_0^{\Delta t} (v - v_s)(v - \dot{u})dt - \frac{1}{2}\int_0^{\Delta t} (u - u_s)[\omega^2 \ell_1 (u - u_s)$$

$$+ \ell_2(v - v_s)]dt - \frac{1}{2}\int_0^{\Delta t} (v - v_s)[\ell_2(u - u_s) + \ell_3(v - v_s)]dt$$

$$\tag{3.6}$$
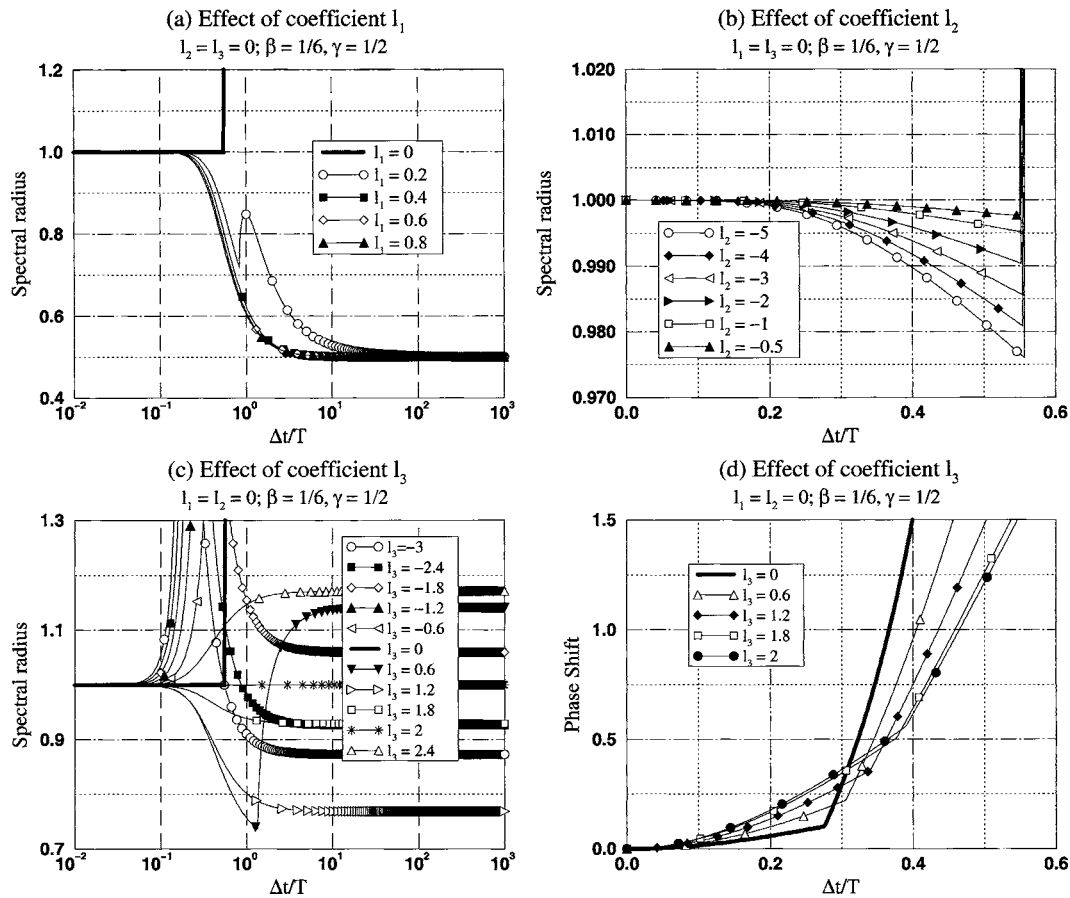
Fig. 3 Effect of the inclusion of a constant kernel on the algorithm based on functional (3.6) with discretization as in Eqs. (2.4), (2.6), and (2.12): stability and phase shift

By adopting the discretizations (2.4), (2.6), and (2.12), and computing the stationarity conditions of functional (3.6) with respect to all the unknown coefficients, we obtain a whole family of time integration schemes, which, of course, reduces to Newmark's one for the choice $\ell_1 = \ell_2 = \ell_3 = 0$. It is quite cumbersome to examine the algorithmic features of this family in terms of all its free parameters. Therefore, here we show only some results for the special case $\beta = 1/6$, $\gamma = 1/2$. In the linear case, such as that governed by Eqs. (2.2), the analysis of the algorithm can be performed in a standard way, by writing down in explicit form the amplification matrix and computing numerically (analytical calculations are too complex) the spectral radius (i.e., the maximum eigenvalue of the amplification matrix), the numerical damping, and the phase shift (i.e., the relative error between the numerical and the exact time periods of the motion, as described, for instance, in [11]).

Any choice of parameters ℓ_1, ℓ_2, and ℓ_3 yields second-order accurate algorithms; this can be checked analytically, from the coefficients of the amplification matrix, reported in Appendix A. All the other results, part of which are summarized in Fig. 3, have been obtained numerically.

Coefficient ℓ_1 must be non-negative in order to guarantee some stability. If $0 \leq \ell_1 < 0.157$, with $\ell_2 = \ell_3 = 0$, one has conditional stability, with critical time step increasing with ℓ_1; if $\ell_1 \geq 0.157$ unconditional stability with numerical damping is obtained [Fig. 3(a)].

Coefficient ℓ_2 must be nonpositive to guarantee stability; its inclusion has almost negligible effects on all the algorithmic properties (it adds an almost negligible amount of numerical damping, and alters very slightly both the phase shift and the stability limit—see Fig. 3(b)).

Finally, we have checked the influence of coefficient ℓ_3 only for $\ell_1 = \ell_2 = 0$. For $\ell_3 < 0$ and $\ell_3 > 2$ there is unconditional instability, while for $0 \leq \ell_3 < 0.746$ one has conditional stability, with critical time step increasing with ℓ_3. The choice $0.746 \leq \ell_3 < 2$, with $\ell_1 = \ell_2 = 0$, yields unconditional stability with numerical damping. Only for $\ell_3 = 2$ unconditional stability with no numerical damping is found (Fig. 3(c)).

These are indeed surprising results, since the adopted discretization, if plugged into the no-kernel functional (2.3), with $\beta = 1/6$, $\gamma = 1/2$, gives rise to a conditionally stable algorithm. In this case, unlike what was found in [1] with reference to a wholly different discretization, it turns out that the addition of a kernel can yield unconditional stability, where the no-kernel approach, keeping fixed all the other ingredients, leads to conditional stability; by selecting a careful combination of coefficients ℓ_i one can also control numerical damping.

Figure 3(d) shows the influence of coefficient ℓ_3, with $\ell_1 = \ell_2 = 0$, on the phase shift $(\bar{T} - T)/T$ (T being the exact time period of the motion, and $\bar{T}$ the computed one) of the solution, which is slightly worse than that of the basic Newmark scheme for small time steps, and much better for large ones.

The same kernel discussed here has similar effects also on the Central Difference method, another special case of Newmark's, with $\beta = 0$, $\gamma = 1/2$, therefore explicit. Here, for example, any choice $\ell_1 = \ell_2 = 0$, $0.2 \leq \ell_3 \leq 2$ makes the new algorithms, obtained by means of the inclusion of a constant kernel, unconditionally stable and, for $\ell_3 = 2$, with zero numerical damping as well. In this case, however, the modified algorithms should be implicit, even though we have not checked this analytically (we have obtained only the relevant amplification matrix).

3.4 An Example for a Nonlinear Case. Finally, we discuss some results obtained by applying the variationally based integration methods described previously to the following nonlinear (modified Duffing) equation:

$$\ddot{u}(t) + c\dot{u}(t) + u^3(t) = f(t) \equiv -\cos t - c \sin t + \cos^3 t \quad (3.7)$$

where everything is intended to be nondimensional, and which, if coupled with the initial conditions

$$u(0) = 1; \quad (3.8a)$$

$$\dot{u}(0) = 0 \quad (3.8b)$$

has the trivial exact solution

$$u(t) = \cos t \quad (3.9)$$

Such a problem has been considered in [12] for $c = 0.05$; it was observed that the solution (3.9), if computed numerically, is unstable, regardless of the adopted integration method and of the chosen time step. By unstable solution here we mean that the motion described by Eq. (3.9), if computed by numerical integration of Eq. (3.7), no matter how small is the adopted time step, is always abandoned after some cycles, with the computed solution going towards one of several others, which seem to exist for Eq. (3.7), and to powerfully attract the numerical solution; by changing details of the integration scheme one or another of them can be reached.

The methods we have compared are:

1. the basic Newmark's method as given by Eqs. (2.1), coupled with Eq. (3.7) evaluated at t_{n+1}, with $\beta = 1/6$, $\gamma = 1/2$;
2. the method arising from the procedure discussed in Sec. 2. Recall that such a procedure yields exactly Newmark's method in the linear case, but in the nonlinear one, in general, it does not; now, we must rewrite functional (2.3), according to the formal structure of Eq. (1.5), accounting for the new problem under study, i.e.:

$$G[x,y] = G[u(t), v(t), u_s(t), v_s(t)]$$

$$= \int_0^{\Delta t} [u(t) - u_s(t)][\dot{v}(t) + cv(t) + u(t)^3 - f(t)]dt$$

$$+ \int_0^{\Delta t} [v(t) - v_s(t)][v(t) - \dot{u}(t)]dt \quad (3.10)$$

to be discretized as in Eqs. (2.4), (2.6), and (2.12);
3. the same as in point 2 above, but exploiting the formal structure of functional (3.10) to integrate the forcing term analytically. By this we mean that, instead of adopting the discretization (2.6), we put into functional (3.10) the analytic expression for the forcing term $f(t)$ as defined by Eq. (3.7) and then, adopting the discretization defined by Eqs. (2.4) and (2.12), we compute the relevant integrals in closed form. Such a procedure cannot be applied in general, of course, but it may be useful to distinguish between two different sources of error: that due to the discretization, and the quadrature one introduced when approximating the forcing term. The availability of a functional allows one to mitigate this last error by adopting better quadrature formulas than that implied by Eq. (2.6);
4. the "improved" Newmark method already illustrated in Sec. 3.2 for the linear case, obtained now by starting from functional (3.10), and adopting the discretization defined by Eqs. (2.6), (2.12), and (3.3);
5. the method already illustrated in Sec. 3.3 for the linear case, which makes use of the same formulation and discretization as in method 2, but starts from a functional of the type (1.4), with the linear kernel

$$\mathcal{S}(y) = \mathcal{S}(u_s, v_s) = \begin{bmatrix} \ell_1 & \ell_2 \\ \ell_2 & \ell_3 \end{bmatrix} \begin{bmatrix} u_s(t) \\ v_s(t) \end{bmatrix} \quad (3.11)$$

The resulting functional, associated with Eq. (3.7), reads therefore:

$$F[x,y] = F[u(t), v(t), u_s(t), v_s(t)]$$

$$= \int_0^{\Delta t} (\dot{v} + cv + u^3 - f)dt + \int_0^{\Delta t} (v - v_s)(v - \dot{u})dt$$

$$- \frac{1}{2}\int_0^{\Delta t} (u - u_s)[\ell_1(u - u_s) + \ell_2(v - v_s)]dt$$

$$- \frac{1}{2}\int_0^{\Delta t} (v - v_s)[\ell_2(u - u_s) + \ell_3(v - v_s)]dt \quad (3.12)$$

and the discretization is again given by Eqs. (2.4), (2.6), and (2.12). Here, we show the results obtained for $\ell_1 = -1.352\,21$, $\ell_2 = 0$, $\ell_3 = 0$; some comments about the influence of the parameter choice for this problem will be given later.

We have adopted $\beta = 1/6$ and $\gamma = 1/2$ in all the methods, which, for this problem, are all conditionally stable (in the usual sense).

Figure 4 shows time histories, as computed numerically using the described methods, for $c = 0.05$ and $\Delta t = 0.5$ (i.e., $\Delta t/T = 0.078\,75$, T being the time period of the exact solution of Eq. (3.7)). As expected, none of the employed methods could remove the numerical instability, as each algorithm, after following the exact solution (3.9) for some cycles, goes in its own way towards different time histories. Methods 1 and 2 (Figs. 4(a) and 4(b)) do not yield the same result, even though they coincide in the linear case. The use of an exact integration of the forcing term (method 3, Fig. 4(c)) improves only marginally the stability of the solution. When using the "improved" Newmark scheme with this time step (result not illustrated in Fig. 4) an explosive instability of the solution is obtained, which could be removed only by reducing the time step.

This suggests that the accuracy of the algorithm is less of a factor, in controlling how long the numerical solution of problem (3.7) can stay on the right one (3.9), than its numerical damping. Indeed, both methods 3 and 4 are affected by a verifiably smaller error than both 1 and 2 in the linear case; nevertheless, it appears that their energy behavior, for this specific problem, is quite far from optimal (recall, however, that method 4 has been "designed," in the linear case, with accuracy only as a target).

Our choice of parameters ℓ_1, ℓ_2, and ℓ_3, for the kernel (3.11) in method 5, allows the algorithm to stay on the exact solution for a much longer time than the others (Fig. 4(d)).

It must be said that this behavior is not strictly associated with the use of this specific kernel, but that the kernel itself just acts as a further ingredient which (i) changes the conditioning of the resulting algorithm, as discussed in [1,13], and (ii) can control the numerical damping, as shown in Sec. 3.3 for the linear case. In this last respect, it is worth observing that the "good" result of Fig. 4(d) is obtained for a value of ℓ_1 which would lead to an unstable algorithm, with this time step, in the linear case. Here, the tendency of the numerical solution to abandon the exact one of Eq. (3.9) is due to numerical errors usually associated with some energy loss at each step; if the algorithm is tuned in such a way as to compensate this loss, by adding some small artificial amount of energy (therefore being intrinsically unstable), it might cause the computed solution to stay on the exact, unstable one for a longer time[2]. Therefore, for this problem the obvious conclusion is that

---

[2]This phenomenon, which we were unaware of, before finding it in our calculations, can be also obtained in the numerical solution of linear problems, when the physical damping and the spurious introduction or removal of energy by numerical
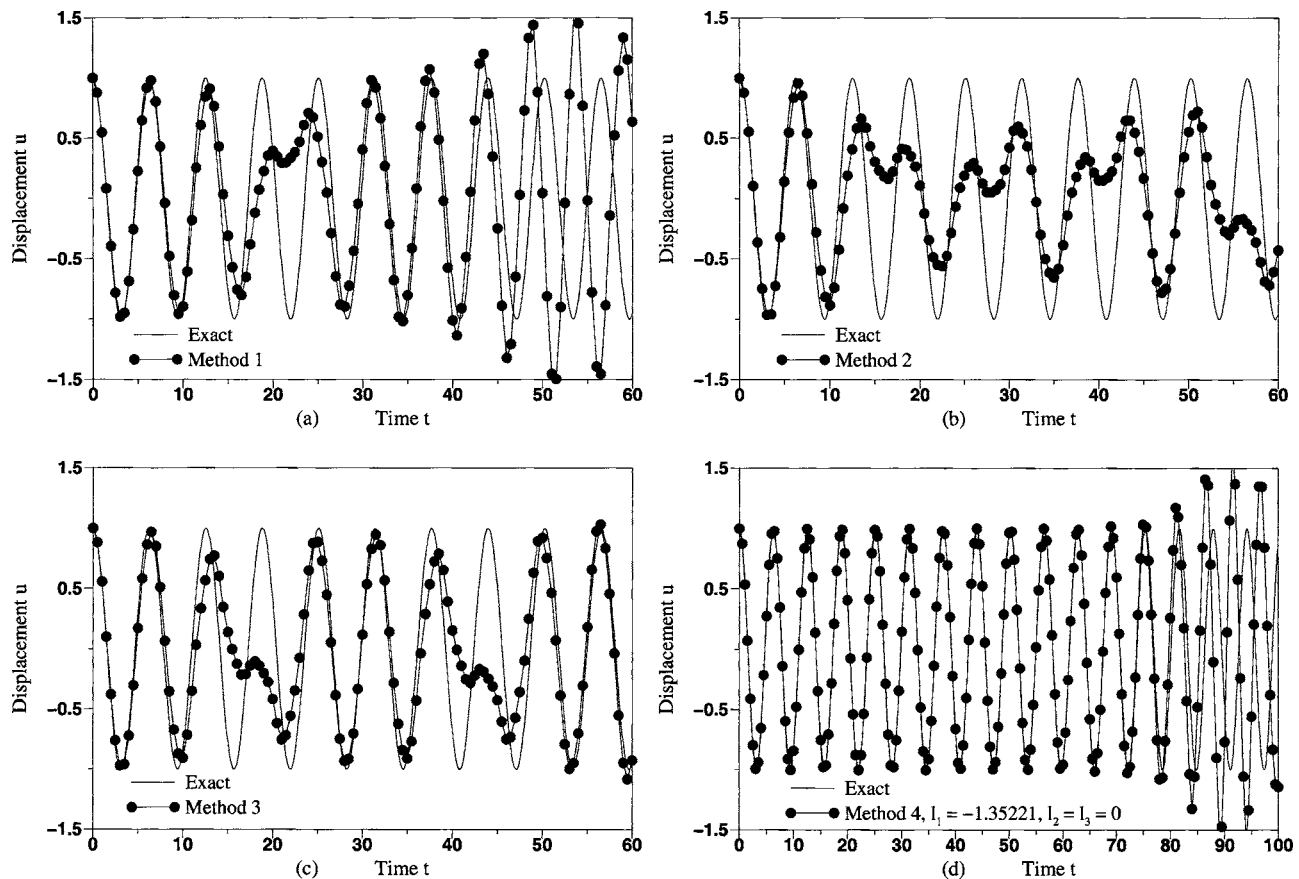
Fig. 4 Time histories computed numerically for the modified Duffing equation of Eq. (3.7). Curve (a): standard Newmark method, $\beta=1/6$, $\gamma=1/2$; curve (b): variationally derived Newmark method, $\beta=1/6$, $\gamma=1/2$; curve (c): same as in (b), with analytical integration of the forcing term; curve (d): kernel-modified variational method (functional (3.12)), discretization as in Eqs. (2.4), (2.6), and (2.12), $\beta=1/6$, $\gamma=1/2$; $l_1=-1.35221$; $l_2=l_3=0$. $\Delta t=0.5$ in all curves.

this specific choice of a kernel can do nothing good; nevertheless, it is quite reasonable to expect that other methods, specifically tailored with respect to stability taking advantage of the available theoretical framework, might perform much better.

4 Conclusions

We have illustrated a technique to obtain the Newmark method by discretizing an extended functional. The required formulation is a time-continuous, two-field one, and the discretization is linear in time for both the shape and the test functions, this latter being not trivial at all.

We have provided examples that illustrate some of the advantages coming from the availability of functionals for the structural dynamics problem with given initial conditions. For instance, the method of Sec. 3.2 exhibits an accuracy far superior to that of the basic Newmark method, at a comparable cost, whereas the inclusion of kernels, exemplified in the methods of Secs. 3.3 and 3.4, has proved to be an important factor in controlling the stability and conditioning properties of the numerical algorithms.

The availability of a functional allows one to explore new ways for deriving time integration algorithms, and to better understand their features. In particular, the knowledge of discretizations leading to the whole family of Newmark algorithms in the linear case appears to be an important factor which facilitates the development of new, more efficient, algorithms for many particular problems, which might meet different requirements, for instance in-

volving accuracy order, stability range, dissipation/energy conservation properties. Specific constraints, deriving from the physical problem studied, may be included in the functional and thus be automatically incorporated into the final numerical algorithm derived from the stationarity process. All these items are the subject of investigations under way.

Appendix A

We report the coefficients of the amplification matrix $\mathbf{A}$ for the algorithms described in Secs. 3.1–3.3. Note that, for the sake of brevity, here we have computed and written these results only for the special case with mass $m=1$, i.e., only in terms of the damping ratio $\xi=c/m$ and of the angular frequency $\omega=\sqrt{k/m}$ of the undamped motion. The more general results holding for the case of a non-unitary mass should be recomputed on the basis of what is illustrated in the main text; nevertheless, the equations hereafter reported allow one to easily check all the properties of the corresponding algorithms.

For the family of algorithms derived in Sec. 3.1, the coefficients of matrix $\mathbf{A}$ have the following expressions:

schemes can influence each other in such a way as to produce stable but meaningless results.

$$A_{11} = \frac{M_1}{Q_1}; \quad A_{12} = \frac{M_2}{Q_2}; \quad A_{21} = \frac{M_3}{Q_2}; \quad A_{22} = \frac{M_4}{Q_3}$$

where

$$M_1 = 12 + (-7 + 12\beta)\Delta t^2\omega^2 + 2\Delta t^4(-2\beta + \gamma)\omega^4 \\ + \xi[12\Delta t\gamma + \Delta t^3(6\beta\omega^2 - 3\gamma\omega^2)]$$

$$M_2 = 6\Delta t\{-2 + \xi\Delta t[1 - 2\gamma + \xi\Delta t(-2\beta + \gamma)]\} \\ + \Delta t^3(1 + 6\beta\xi\Delta t - 3\xi\Delta t\gamma)\omega^2$$

$$M_3 = 6\Delta t\omega^2(2 + \Delta t^2(2\beta - \gamma)\omega^2)$$

$$M_4 = 12 - \Delta t(-12\xi(-1 + \gamma) + \Delta t(1 + 6\beta(-2 + 3\xi\Delta t) \\ + 3(4 - 3\xi\Delta t)\gamma)\omega^2 + \Delta t^3(-2\beta + \gamma)\omega^4)$$

$$Q_1 = 12 + (-1 + 12\beta)\Delta t^2\omega^2 + 2\Delta t^4(-2\beta + \gamma)\omega^4 \\ + 3\xi(4\Delta t\gamma + \Delta t^3(-2\beta + \gamma)\omega^2)$$

$$Q_2 = -12(1 + \xi\Delta t\gamma) + \Delta t^2(1 + 6\beta(-2 + \xi\Delta t) - 3\xi\Delta t\gamma)\omega^2 \\ - 2\Delta t^4(-2\beta + \gamma)\omega^4$$

$$Q_3 = 12 + 3\xi\Delta t(4\gamma + \Delta t^2(-2\beta + \gamma)\omega^2) \\ + \Delta t^2\omega^2(-1 + 12\beta + 2\Delta t^2(-2\beta + \gamma)\omega^2)$$

For the family of algorithms derived in Sec. 3.2, the coefficients of matrix **A** turn out to have the following expressions:

$$A_{11} = \frac{N_1}{D}; \quad A_{12} = \frac{N_2}{D}; \quad A_{21} = \frac{N_3}{D}; \quad A_{22} = \frac{N_4}{D}$$

where

$$D = 12 + \Delta t(12\xi\gamma + 12\beta\Delta t\omega^2 + \Delta t(\gamma h_1 + \beta h_2)\omega^2 \\ \times(-2 + \Delta t^2(-2\beta + \gamma)\omega^2))$$

$$N_1 = 12 + \Delta t\{12\xi\gamma - 2\Delta t[3 + 3\xi\Delta t\gamma + \gamma h_1 \\ + \beta(-6 - 6\xi\Delta t + h_2)]\omega^2 + \Delta t^3(-2\beta + \gamma)(\gamma h_1 + \beta h_2)\omega^4\}$$

$$N_2 = -2\Delta t(-6 + \Delta t(3\xi(1 - 2\gamma + \xi\Delta t(-2\beta + \gamma)) \\ + \Delta t(\gamma h_1 + \beta h_2)\omega^2))$$

$$N_3 = -6\Delta t\omega^2(2 + \Delta t^2(2\beta - \gamma)\omega^2)$$

$$N_4 = 12 + \Delta t(12\xi(-1 + \gamma) - 2\Delta t(\gamma(6 - 3\xi\Delta t + h_1) \\ + \beta(-6 + 6\xi\Delta t + h_2))\omega^2 + \Delta t^3(-2\beta + \gamma)(\gamma h_1 + \beta h_2)\omega^4)$$

For the family of algorithms derived in Sec. 3.3, the coefficients of matrix **A** turn out to have the following expressions:

$$A_{11} = \frac{C_1}{D_1}; \quad A_{12} = \frac{C_2}{D_2}; \quad A_{21} = \frac{C_3}{D_2}; \quad A_{22} = \frac{C_4}{D_1}$$

where

$$C_1 = 6(12 + \ell_2^2\Delta t^2) - 3\Delta t^2(2 + 2\ell_3(-1 + \ell_1) \\ - 4\ell_1 + \ell_2\Delta t)(12 + \ell_2^2\Delta t^2)\omega^2 + 3\Delta t^4(4(-2 + \ell_1)(1 + \ell_1) \\ + \ell_3^2(3 + (-2 + \ell_1)\ell_1) - 2\ell_2(-1 + \ell_1)\Delta t + \ell_3[-3(2 \\ + \ell_2\Delta t) + \ell_1(6 - 4\ell_1 + \ell_2\Delta t)])\omega^4 + 2(-2 + \ell_3)\ell_1\Delta t^6\omega^6$$

$$C_2 = -3\Delta t(12 + \ell_2^2\Delta t^2)^2 + 3(-2 - \ell_3 + 2(-2 + \ell_3)\ell_1)\Delta t^3 \\ \times(12 + \ell_2^2\Delta t^2)\omega^2 - 3(-2 + \ell_3)\ell_1[\ell_3(-1 + \ell_1) - 2(1 \\ + \ell_1)]\Delta t^5\omega^4$$

$$C_3 = 36\Delta t(12 + \ell_2^2\Delta t^2)\omega^2 + 3\Delta t^3(12 + 24\ell_1 + \ell_2\Delta t(-12 + \ell_2\Delta t) \\ + 6\ell_3[4 - 2\ell_1 + \ell_2\Delta t])\omega^4 - 3(2 + 3\ell_3(-1 + \ell_1) - 6\ell_1)\Delta t^5\omega^6$$

$$C_4 = 6(12 + \ell_2^2\Delta t^2) + 3\Delta t^2(-2 - 2\ell_3(-1 + \ell_1) + 4\ell_1 + \ell_2\Delta t) \\ \times(12 + \ell_2^2\Delta t^2)\omega^2 - 3\Delta t^4[-(\ell_3^2(-3 + \ell_1)(1 + \ell_1)) - 2(1 + \ell_1) \\ \times(-4 + 2\ell_1 + \ell_2\Delta t) + \ell_3[-6 - \ell_2\Delta t + \ell_1(-6 + 4\ell_1 \\ + \ell_2\Delta t)]]\omega^4 \\ + 2(-2 + \ell_3)\ell_1\Delta t^6\omega^6$$

$$D_1 = 6(12 + \ell_2^2\Delta t^2) - 6[\ell_3(-1 + \ell_1) - 2(1 + \ell_1)]\Delta t^2 \\ \times(12 + \ell_2^2\Delta t^2)\omega^2 + 3\Delta t^4(\ell_3^2(-3 + \ell_1)(1 + \ell_1) + 4[(1 + \ell_1)^2 \\ - 2\ell_2\Delta t] - 4\ell_3[-3 + \ell_1^2 - \ell_2\Delta t])\omega^4 - 4(-2 + \ell_3)\ell_1\Delta t^6\omega^6$$

$$D_2 = -3(12 + \ell_2^2\Delta t^2)^2 + 6[\ell_3(-1 + \ell_1) - 2(1 + \ell_1)]\Delta t^2 \\ \times(12 + \ell_2^2\Delta t^2)\omega^2 - 3\Delta t^4(\ell_3^2(-3 + \ell_1)(1 + \ell_1) \\ + 4[(1 + \ell_1)^2 - 2\ell_2\Delta t] - 4\ell_3[-3 + \ell_1^2 - \ell_2\Delta t])\omega^4 \\ + 4(-2 + \ell_3)\ell_1\Delta t^6\omega^6$$

Appendix B

Here, we show how the variational technique discussed in this work can be applied to the case of a multi-DOF linear system. First, we go through the main steps needed to obtain the standard Newmark scheme, and, eventually, we will show that the "improved" scheme proposed in Sec. 3.2 does not require a substantial increase of the numerical burden.

Let us write the equations of motion for a multi-DOF linear elastic system by generalizing Eqs. (2.2) as follows (with obvious meaning of symbols):

$$\mathbf{M}\dot{\mathbf{v}}(t) + \mathbf{C}\mathbf{v}(t) + \mathbf{K}\mathbf{u}(t) - \mathbf{f}(t) = \mathbf{0} \qquad (A1)$$

$$\mathbf{v}(t - \mathbf{u}(t) = \mathbf{0} \qquad (A2)$$

plus initial conditions in terms of vectors $\mathbf{u}_n$ and $\mathbf{v}_n$. Consider now the variational procedure to obtain the standard Newmark method illustrated in Sec. 2 (Eqs. (2.4), (2.6), and (2.12), now intended to hold for every component of the unknown vectors), and define, for convenience, the following constant coefficients:

$$A = 4 - 6\gamma; \quad B = \frac{6(2\gamma - 1)}{\Delta t}; \quad D = 2(3\beta - 1)\Delta t;$$

$$E = -3(4\beta - 1) \qquad (A3)$$

such that the discretization (2.12) for the auxiliary unknown vectors can be written as

$$\mathbf{u}_s(t) = (A + Bt)\hat{\mathbf{u}}_s + (D + Et)\hat{\mathbf{v}}_s; \quad \mathbf{v}_s(t) = \hat{\mathbf{v}}_s \qquad (A4)$$

Writing now the extended functional (2.3), in which, for the time being, only the discretization (A4) of the auxiliary unknowns is explicitly written out, but in which it is intended that also the main unknowns are discretized according to Eqs. (2.4), we obtain:

$$G[\hat{\mathbf{u}}_{n+1}, \hat{\mathbf{v}}_{n+1}, \hat{\mathbf{u}}_s, \hat{\mathbf{v}}_s] = \int_0^{\Delta t} [\mathbf{u}(t) - (A + Bt)\hat{\mathbf{u}}_s - (D + Et)\hat{\mathbf{v}}_s]$$

$$\cdot [\mathbf{M}\dot{\mathbf{v}}(t) + \mathbf{C}\mathbf{v}(t) + \mathbf{K}\mathbf{u}(t) - \mathbf{f}(t)]dt$$

$$+ \int_0^{\Delta t} [\mathbf{v}(t) - \hat{\mathbf{v}}_s] \cdot \mathbf{P}[\mathbf{v}(t) - \dot{\mathbf{u}}(t)]dt \quad (A5)$$

where the symbol · indicates a vector dot product and, in the second integral, we have introduced a non-singular mass-type matrix **P**, similarly to what done with reference to Eq. (2.3). Here too, this corresponds simply to writing Eq. (A2) premultiplied by

matrix $\mathbf{P}$, and then applying the formalism of Eq. (1.5). It is convenient, for the time being, to keep matrix $\mathbf{P}$ different from the actual mass matrix $\mathbf{M}$ in order to understand, later on, the differences in cost between the basic Newmark method and the improved one of Sec. 2.3. It is intended, however, that matrix $\mathbf{P}$ is fully known.

The stationarity conditions of function (A5) with respect to the auxiliary unknown vectors $\hat{\mathbf{u}}_s$ and $\hat{\mathbf{v}}_s$ give us the equations to compute the main unknowns $\hat{\mathbf{u}}_{n+1}$ and $\hat{\mathbf{v}}_{n+1}$. We obtain what follows:

$$\frac{\partial G}{\partial \hat{\mathbf{u}}_s} = \int_0^{\Delta t} A[\mathbf{M}\dot{\mathbf{v}}(t) + \mathbf{C}\mathbf{v}(t) + \mathbf{K}\mathbf{u}(t) - \mathbf{f}(t)]dt$$

$$+ \int_0^{\Delta t} B[\mathbf{M}\dot{\mathbf{v}}(t) + \mathbf{C}\mathbf{v}(t) + \mathbf{K}\mathbf{u}(t) - \mathbf{f}(t)]tdt = \mathbf{0} \quad (A6)$$

$$\frac{\partial G}{\partial \hat{\mathbf{v}}_s} = \int_0^{\Delta t} D[\mathbf{M}\dot{\mathbf{v}}(t) + \mathbf{C}\mathbf{v}(t) + \mathbf{K}\mathbf{u}(t) - \mathbf{f}(t)]dt + \int_0^{\Delta t} E[\mathbf{M}\dot{\mathbf{v}}(t)$$

$$+ \mathbf{C}\mathbf{v}(t) + \mathbf{K}\mathbf{u}(t) - \mathbf{f}(t)]tdt + \int_0^{\Delta t} \mathbf{P}[\mathbf{v}(t) - \dot{\mathbf{u}}(t)]dt = \mathbf{0}$$

$$(A7)$$

i.e., a linear system of equations in the two unknown vectors $\hat{\mathbf{u}}_{n+1}$ and $\hat{\mathbf{v}}_{n+1}$. Now, we need to reduce this system to one in a single unknown vector (instead of two), similarly to what is required by Newmark's method. One way to do so starts by explicitly inserting the discretization of the main unknowns into the stationarity Eqs. (A6) and (A7) and evaluating the integrals, thus obtaining

$$\left(\frac{A\Delta t}{2} + \frac{B\Delta t^2}{3}\right)\mathbf{K}\hat{\mathbf{u}}_{n+1} + \mathbf{g}[\hat{\mathbf{v}}_{n+1}, \mathbf{u}_n, \mathbf{v}_n, \mathbf{f}(t)] = \mathbf{0} \quad (A8)$$

$$\left[\left(\frac{D\Delta t}{2} + \frac{E\Delta t^2}{3}\right)\mathbf{K} - \mathbf{P}\right]\hat{\mathbf{u}}_{n+1} + \mathbf{h}[\hat{\mathbf{v}}_{n+1}, \mathbf{u}_n, \mathbf{v}_n, \mathbf{f}(t)] = \mathbf{0} \quad (A9)$$

where

$$\mathbf{g}[\hat{\mathbf{v}}_{n+1}, \mathbf{u}_n, \mathbf{v}_n, \mathbf{f}(t)] = (\mathbf{M} + \gamma\mathbf{C}\Delta t)\hat{\mathbf{v}}_{n+1} + (1 - \gamma)\Delta t\mathbf{K}\mathbf{u}_n$$

$$+ [(1 - \gamma)\mathbf{C}\Delta t - \mathbf{M}]\mathbf{v}_n + \Delta t(1 - \gamma)\mathbf{f}_n - \Delta t\mathbf{f}_{n+1}$$

$$(A10)$$

$$\mathbf{h}[\hat{\mathbf{v}}_{n+1}, \mathbf{u}_n, \mathbf{v}_n, \mathbf{f}(t)] = \left((\mathbf{P} - \mathbf{M})\frac{\Delta t}{2} - \beta\mathbf{C}\Delta t^2\right)\hat{\mathbf{v}}_{n+1}$$

$$+ \left[\mathbf{P} + \left(\beta - \frac{1}{2}\right)\mathbf{K}\Delta t^2\right]\mathbf{u}_n$$

$$+ \left[(\mathbf{P} + \mathbf{M})\frac{\Delta t}{2} + \left(\beta - \frac{1}{2}\right)\mathbf{C}\Delta t^2\right]\mathbf{v}_n$$

$$- \left(\beta - \frac{1}{2}\right)\Delta t^2\mathbf{f}_n + \beta\Delta t^2\mathbf{f}_{n+1} \quad (A11)$$

Recalling the definitions (A3), we can rewrite Eqs. (A8) and (A9) as follows:

$$\gamma\Delta t\mathbf{K}\hat{\mathbf{u}}_{n+1} + \mathbf{g}[\hat{\mathbf{v}}_{n+1}, \mathbf{u}_n, \mathbf{v}_n, \mathbf{f}(t)] = \mathbf{0} \quad (A12)$$

$$(\mathbf{P} + \beta\Delta t^2\mathbf{K})\hat{\mathbf{u}}_{n+1} - \mathbf{h}[\hat{\mathbf{v}}_{n+1}, \mathbf{u}_n, \mathbf{v}_n, \mathbf{f}(t)] = \mathbf{0} \quad (A13)$$

Multiplying Eq. (A12) by $\beta\Delta t$ and Eq. (A13) by γ, and taking the difference between the two equations, one obtains the following expression:

$$\mathbf{P}\hat{\mathbf{u}}_{n+1} = \frac{\beta\Delta t}{\gamma}\mathbf{g}[\hat{\mathbf{v}}_{n+1}, \mathbf{u}_n, \mathbf{v}_n, \mathbf{f}(t)] + \mathbf{h}[\hat{\mathbf{v}}_{n+1}, \mathbf{u}_n, \mathbf{v}_n, \mathbf{f}(t)] \quad (A14)$$

which shows us that, as long as the coefficient of $\hat{\mathbf{u}}_{n+1}$ in the discretization of $\mathbf{u}(t)$ is that described by Eq. (2.4), it is always possible to express $\hat{\mathbf{u}}_{n+1}$ as function of $\hat{\mathbf{v}}_{n+1}$ by inverting only the matrix $\mathbf{P}$. We will come back later to this point, crucial to understand the computational cost required by methods thus formulated, but it is obvious that any method which allows one to choose a diagonal matrix $\mathbf{P}$ would be much less expensive than what appears from the aspect of the general Eqs. (A6) and (A7).

Exploiting results (A10) and (A11), making use of the equations of motion (A1) written at time t_n, setting $\mathbf{a}_n = \dot{\mathbf{v}}(t_n)$, and going through all the calculations, one can rewrite Eq. (A14) as follows:

$$\mathbf{P}\hat{\mathbf{u}}_{n+1} = \mathbf{P}\mathbf{u}_n + \frac{\Delta t}{2}\mathbf{P}\mathbf{v}_n + \frac{\Delta t}{2}\mathbf{P}\hat{\mathbf{v}}_{n+1} - \left(\frac{\beta}{\gamma} - \frac{1}{2}\right)\Delta t^2\mathbf{M}\mathbf{a}_n$$

$$+ \left(\frac{\beta}{\gamma} - \frac{1}{2}\right)\Delta t^2\mathbf{M}\left(\frac{\hat{\mathbf{v}}_{n+1} - \mathbf{v}_n}{\Delta t}\right) \quad (A15)$$

If one now sets $\mathbf{P} = \mathbf{M}$ into Eq. (A15), as required to obtain Newmark's equations, one arrives easily at the following:

$$\hat{\mathbf{u}}_{n+1} = \mathbf{u}_n + \Delta t\left(1 - \frac{\beta}{\gamma}\right)\mathbf{v}_n + \frac{\beta\Delta t}{\gamma}\hat{\mathbf{v}}_{n+1} + \left(\frac{1}{2} - \frac{\beta}{\gamma}\right)\Delta t^2\mathbf{a}_n$$

$$(A16)$$

an explicit relationship between $\hat{\mathbf{u}}_{n+1}$ and $\hat{\mathbf{v}}_{n+1}$ which requires no matrix inversion at all, and which, if reinserted into Eq. (A12), provides the solution of the problem according to Newmark's equations, through the solution of a linear system of equations in $\hat{\mathbf{v}}_{n+1}$ only.

If we now turn to the discretization (3.3) of the main unknowns, that, coupled with the same discretization (2.12) of the auxiliary ones, produces a modified Newmark scheme with superior properties with respect to the standard one (at least in some cases), we observe that the coefficient of vector $\hat{\mathbf{u}}_{n+1}$ remains exactly the same as in the case of the basic Newmark scheme (Eq. (2.4)). Therefore, even though the functions $\mathbf{g}[\hat{\mathbf{v}}_{n+1}, \mathbf{u}_n, \mathbf{v}_n, \mathbf{f}(t)]$ and $\mathbf{h}[\hat{\mathbf{v}}_{n+1}, \mathbf{u}_n, \mathbf{v}_n, \mathbf{f}(t)]$ in Eqs. (A8) and (A9) are no more defined as in Eqs. (A10) and (A11), the matrices of coefficients of vector $\hat{\mathbf{u}}_{n+1}$ in Eqs. (A8) and (A9) remain unaltered, and the result (A14) still holds.

If we wish to recover our improved Newmark method of Sec. 3.2, we must still set $\mathbf{P} = \mathbf{M}$ in function (A5), but now all the simplifications allowed by the structure of Eqs. (A8) and (A9) hold no more, and the simple result (A16) cannot be written, because, in Eq. (A15), it is very unlikely (we have not checked it, though) that only matrices $\mathbf{P}$ and $\mathbf{M}$ would appear.

Nevertheless, a result analogous to Eq. (A16) can always be obtained, by just inverting matrix $\mathbf{P}$, as already suggested, and as apparent from the examination of Eq. (A14). In our case, since we must set $\mathbf{P} = \mathbf{M}$, and since quite often matrix $\mathbf{M}$ is diagonal, we can consider this step comparatively inexpensive, and the whole procedure, even in the case of a non-diagonal mass matrix, is certainly much less expensive than solving a coupled system in both unknowns $\hat{\mathbf{u}}_{n+1}$ and $\hat{\mathbf{v}}_{n+1}$. In this sense we have said, earlier in this paper, that this new scheme has the same computational cost as the basic Newmark method.

References

[1] Carini, A., and Genna, F., 2000, "Saddle-Point Principles and Numerical Integration Methods for Second-Order Hyperbolic Equations," Comput. Methods Appl. Mech. Eng., **190**, pp. 1663–1678.
[2] Tonti, E., 1984, "Variational Formulations for Every Nonlinear Problem," Int. J. Eng. Sci., **22**(11–12), pp. 1343–1371.
[3] Auchmuty, G., 1988, "Variational Principles for Operator Equations and Initial

Value Problems," Nonlinear Anal. Theory, Methods Appl., **12**(5), pp. 531–564.

[4] Brun, M., Carini, A., and Genna, F., 2001, "On the Construction of Extended Problems and Related Functionals for General Nonlinear Equations," J. Mech. Phys. Solids, **49**(4), pp. 839–856.

[5] Salvadori, A., 2002, "Extended Functionals and Approximation Schemes for Non-Potential operators. Part I: A Unified View and Some New Results," Quaderni del Seminario Matematico di Brescia, **42/02**, pp. 1–11.

[6] Carini, A., and Genna, F., 1998, "Some Variational Formulations for Continuum Nonlinear Dynamics," J. Mech. Phys. Solids, **46**(7), pp. 1253–1277.

[7] Carini, A., Castiglioni, L., and Genna, F., 1996, "Extremal Formulations in Nonlinear Dynamics," *Proc. ASME International Mechanical Engineering Congress and Exposition*, Atlanta, Georgia, 17–22 November.

[8] Newmark, N. M., 1959, "A Method for Computation of Structural Dynamics," *Proc. of the ASCE*, 85 EM3, pp. 67–94.

[9] Zienkiewicz, O. C., and Taylor, R. L., 2000, *The Finite Element Method*, Third, fourth and fifth editions, McGraw–Hill, London, 1977 and 1989/91; Butterworth and Heinemann, Oxford, UK; Boston, US.

[10] Kane, C., Marsden, J. E., Ortiz, M., and West, M., 2000, "Variational Integrators and the Newmark Algorithm for Conservative and Dissipative Mechanical Systems," Int. J. Numer. Methods Eng., **49**, pp. 1295–1325.

[11] Hughes, T. J. R., 1983, "Analysis of Transient Algorithms with Particular Reference to Stability Behavior," *Computational Methods for Transient Analysis*, T. Belytschko and T. J. R. Hughes, eds., Elsevier, Amsterdam, pp. 67–155.

[12] Addison, P. S., Chan, A. H. C., Ervine, D. A., and Williams, K. J., 1992, "Observations of Numerical Method Dependent Solutions of a Modified Duffing Oscillator," Commun. Appl. Numer. Methods, **8**, pp. 519–528.

[13] Carini, A., Gelfi, P., and Marchina, E., 1995, "An Energetic Formulation for the Linear Viscoelastic Problem. Part I: Theoretical Results and First Calculations," Int. J. Numer. Methods Eng., **38**, pp. 37–62.

Multiscale Shock Heating Analysis of a Granular Explosive

Keith A. Gonthier
Assistant Professor
Mem. ASME
e-mail: gonthier@me.lsu.edu

Venugopal Jogi
Graduate Research Assistant

Mechanical Engineering Department,
Louisiana State University,
Baton Rouge, LA 70803

A multiscale model is formulated and used to characterize the duration and amplitude of temperature peaks (i.e., hot spots) at intergranular contact surfaces generated by shock compaction of the granular high explosive HMX (octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine). The model tracks the evolution of both bulk variables and localized temperature subject to a consistent thermal energy localization strategy that accounts for inelastic and compressive heating, phase change, and thermal conduction at the grain scale (grain size ~50 μm). Steady subsonic compaction waves having a dispersed two-wave structure are predicted for mild impact of dense HMX (porosity ~19%), and steady supersonic compaction waves having a discontinuous solid shock followed by a thin compaction zone are predicted for stronger impact. Short duration hot spots having peak temperatures in excess of 900 K are predicted near intergranular contact surfaces for impact speeds as low as 100 m/s; these hot spots are sufficient to induce sustained combustion as determined by a two-phase thermal explosion theory. Thermal conduction and phase change significantly affect hot-spot formation for low impact speeds (~100 m/s), whereas bulk inelastic heating dominates the thermal response at higher speeds resulting in longer duration hot spots. Compressive grain heating is shown to be largely inconsequential for the range of impact speeds considered in this work ($100 \leqslant u_p \leqslant 1000$ m/s). Predictions for the variation in inelastic strain, pressure, and porosity through the compaction zone are also shown to qualitatively agree with the results of detailed mesoscale simulations. [DOI: 10.1115/1.1934666]

1 Introduction

Localized heating due to the rapid deformation of heterogeneous reactive solids is an important ignition source that may lead to their detonation. Because thermal energy localization occurs at the sub-grain scale (grain size ~50 μm) due to plastic deformation, fracture, and friction, it is difficult to experimentally characterize. Multiscale modeling is, therefore, required to better understand the interplay of localized heating and ignition at the grain scale and the bulk system response. Such modeling necessarily involves the coupling of physical phenomena occurring over disparate length and time scales as illustrated in Fig. 1 for the shock compaction of a granular explosive of engineering dimension (~10 cm). Here, a shock wave is propagating to the right through the unstressed ambient material (porosity ~20%) leaving it in a stressed, compacted state (porosity <5%). The compaction process is largely inelastic resulting in significant deformation and heating near intergranular contact surfaces. Depending on the loading conditions, the mass of material locally heated to a high temperature may be significantly less than the total grain mass yet sufficient to induce bulk combustion. The bulk material response is the integrated manifestation of the grain scale response, and is commonly described using principles of continuum mixture theory [1–3]. Continuum-based models, while important because they can be easily correlated with experiments and applied to engineering scale systems, lack the degrees of freedom needed to accurately describe phenomena occurring at scales smaller than a representative volume element for the bulk material. Indeed, continuum-based models used to describe mechanically induced

transition to detonation for both pressed and granular explosives are often only predictive over a narrow range of impact conditions due to insufficient descriptions of grain scale phenomena (e.g., hot-spot formation). This shortcoming has motivated the development of more robust models for describing hot-spot formation in energetic solids [4–6].

Recently, Gonthier [7] described a modeling approach that can resolve key features of hot-spot formation in a manner compatible with both grain contact mechanics and bulk compaction energetics. The approach requires (1) a model for the bulk material response, (2) a model for the material structure (e.g., grain size and packing), (3) a localization strategy for depositing bulk dissipated energy at the grain scale, and (4) a model for the grain scale response. It assumes that the bulk material response is experimentally well-characterized and can be accurately predicted by the bulk model. Bulk dissipated energy is thermalized at localization sites centered at intergranular contact surfaces. The grain scale response tracks the evolution of hot-spot temperature subject to the localization strategy. The integrity of bulk model predictions is maintained by requiring that the integrated mass, momentum, and energy at the grain scale equals that given by the bulk model.

In this paper, we modify the localization strategy of Ref. [7] to account for compressive heating and phase change at the grain scale, and use the modified model to characterize the duration and magnitude of hot spots induced by mild and strong impact. While compression is unimportant for weak deformation waves having peak solid pressures that are much less than the solid bulk modulus ($P_s \ll K \approx 13$ GPa), it is important for supersonic waves as it gives rise to structures consisting of a lead solid shock followed by a thin compaction zone similar to those first reported by Powers et al. [8]; we give an example of such a structure including its grain scale thermal response. However, even for supersonic waves, we show that compressive heating plays a smaller role than localized inelastic heating. Solid–liquid phase change may also significantly affect hot-spot energetics, particularly at low impact speeds. The latent heat of fusion of HMX at standard conditions corresponds to an equivalent temperature change of approximately $\Delta T = q_m^0 / c_v \approx 147$ K, where $q_m^0 = 0.22$ MJ/kg and c_v

Fig. 1   An illustration of multiscale features for the shock compaction of a granular solid

$=1.5$ kJ/(kg K). It is the relative rates of localized inelastic heating, thermal diffusion, and chemical reaction that determine if sustained combustion will occur. Prior modeling has focused on identifying hot-spot temperatures and sizes needed for ignition and sustained combustion of high-explosives independent of the dynamic bulk material response [9,10]. While useful, such analysis does not provide a fully coupled description of hot-spot creation, growth, and decay due to bulk deformation within the context of an engineering scale model. The primary objective of this work is to predict and resolve the temperature field in the vicinity of intergranular contact surfaces for shock compaction of HMX as part of an ongoing effort to develop coupled bulk-localization models that accurately describe transition to detonation in energetic solids. The model accounts for key mechanisms that influence hot-spot evolution, in an albeit simple manner, with the exception of combustion which will be addressed by subsequent work.

A secondary objective of this work is to report preliminary comparisons between our model predictions and those given by detailed mesoscale simulations [11]. Mesoscale simulations track the discrete dynamical interactions of a small number of identifiable grains (≈ 500), attempting to resolve the stress and temperature fields within grains; they can provide useful information about grain scale hot-spot fluctuations, especially in the absence of experimental data, but are computationally expensive, with fine scale structure difficult to numerically resolve, and are impractical to apply to engineering scale systems. The comparisons given in this paper are a first step in assessing the merits of our localization strategy based on such simulations.

This work is also a preliminary step in the development and characterization of a comprehensive multidimensional model that can be used to predict the combined (hydrodynamic and deviatoric) loading response of granular solids. Though it is justifiable to ignore deviatoric stresses for strong compaction waves because $P \gg \|\tau\|$, where τ is the stress deviator, there exists ample experimental data indicating that bulk shear can significantly affect material compaction for weaker waves by reducing the minimum pressure needed for the onset of inelastic volumetric deformation (i.e., material crush-up); this process is commonly referred to as shear enhanced compaction. Shear enhanced compaction is observed in many types of porous solids including powdered metals [12,13] and geological materials [14,15]. There also exists experimental evidence that shear enhanced compaction may play a prominent role in the ignition of energetic solids [16–18]; a detailed discussion is given in Ref. [19]. Though an analysis of shear enhanced compaction is beyond the scope of this study, the analy-

sis given here will enable the effects of hydrodynamic and deviatoric stresses on the dynamic loading response of granular solids to be systematically isolated and studied.

Finally, we note that the multiscale modeling approach formulated in this paper can also be used to address other problems in engineering science for which localized stress and thermal fluctuations induced by the dynamic loading of heterogeneous materials is important. One obvious example is the synthesis of high strength materials (e.g., intermetallic alloys and ceramics) having small grain crystal structures by shock compaction of nano and micro-size grains [20,21]. For this application, localized heating and phase change occurring near intergranular contact surfaces is essential for obtaining densely consolidated product material. Our modeling approach can be readily applied to such problems with slight modifications to its constitutive theory.

The outline of this paper is as follows. We first briefly describe both the bulk compaction and grain scale heating models, highlighting modifications that account for grain scale compressive heating and phase change. We then give numerical predictions for both subsonic and supersonic steady compaction waves, characterizing the duration and amplitude of hot-spot temperatures induced by these waves. Lastly, we give predictions showing the influence of key model parameters on compaction wave structure, and compare our predictions to those obtained by mesoscale simulations.

2   Bulk Scale Compaction Model

The unsteady, one-dimensional, bulk hydrodynamic compaction model used in this work is a limiting form of the multidimensional model recently formulated by the lead author to investigate the combined effects of hydrodynamic and deviatoric loading [19]. The hydrodynamic model equations are given in conservative form by

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{f}(\mathbf{w})}{\partial x} = \mathbf{g}(\mathbf{w}), \qquad (1)$$

where

$$\mathbf{w} = \left[\rho, \ \rho u, \ \rho\left(e + \frac{u^2}{2}\right), \ \rho\phi, \ \rho\tilde{\phi}, \ n \right]^{\mathrm{T}}, \qquad (2)$$

$$\mathbf{f}(\mathbf{w}) = \left[\rho u, \ \rho u^2 + P, \ \rho u\left(e + \frac{u^2}{2} + \frac{P}{\rho}\right), \ \rho\phi u, \ \rho\tilde{\phi}u, \ un \right]^{\mathrm{T}}, \quad (3)$$

$$\mathbf{g}(\mathbf{w}) = \left[0, \ 0, \ 0, \ \frac{\rho\phi(1-\phi)}{\mu_c}(P_s - \beta), \ \rho\Lambda, \ 0 \right]^{\mathrm{T}} \qquad (4)$$

and

$$\Lambda = \begin{cases} \dfrac{1}{\tilde{\mu}}(f - \tilde{\phi}) & \text{if } f > \tilde{\phi}, \\[2mm] 0 & \text{otherwise.} \end{cases} \qquad (5)$$

These equations track the evolution of granular solid mass, linear momentum, total energy, volume fraction (ϕ), no-load volume fraction ($\tilde{\phi}$), and grain number density (n), respectively. The no-load solid volume fraction is the equilibrium value of the solid volume fraction in the absence of an applied load (i.e., $\beta \rightarrow 0$ as $\phi \rightarrow \tilde{\phi}$, where β is an intergranular stress), and is a measure of inelastic volumetric strain. The grain number density is related to the solid volume fraction and grain radius, r, by $n = \phi/(\frac{4}{3}\pi r^3)$, assuming spherical grains. The assumption of spherical grains, while simplistic, results in a tractable model that allows us to estimate the number of intergranular contacts per unit volume which is needed by the localization strategy outlined in Sec. 3. Other granular solid variables in these equations have their usual meaning: ρ is density, u is particle velocity, P is pressure, and e is internal energy. Quantities associated with the pure phase solid,

denoted by subscript "s," are related to the granular solid variables by $\rho_s = \rho/\phi$, $P_s = P/\phi$, and $e_s = e - B$, where $B = \int_0^\phi \beta/\rho \, d\phi'$ is recoverable compaction energy. It is shown in Ref. [26] that this thermodynamic description is compatible with the Helmholtz free energy function $\Psi(\rho, T, \phi, \tilde{\phi}) = \Psi_s(\rho_s, T) + B(\phi - \tilde{\phi})$ and the Gibbs equation $T d\eta = de - P/\rho^2 d\rho + (P_s - \beta)/\rho d\phi + \beta/\rho d\tilde{\phi}$, where $\Psi_s(\rho_s, T)$ and $\eta = \eta_s(\rho_s, T)$ are the pure phase solid free energy and entropy, respectively, and $T = T_s$ is the temperature. Here, the bulk entropy is equal to the pure phase solid entropy because excess surface entropy is ignored. The model is similar to that of Baer [22] and Powers et al. [8], but better accounts for compaction energetics due to the inclusion of $\tilde{\phi}$ in the theory. A comprehensive discussion of the thermodynamic foundation of the model is given in Ref. [7].

The internal variables ϕ and $\tilde{\phi}$ are bulk (average) measures of the material's microstructural configuration and are needed to specify its thermodynamic state. Rate equations for these internal variables are constructed so that compaction does not violate the strong form of the entropy inequality, and so that $\phi_f \leq \phi \leq 1$, where ϕ_f is the volume fraction of the loose, virgin material. The parameters μ_c and $\tilde{\mu}$ appearing in Eqs. (4) and (5) govern the relaxation rates to the equilibria $P_s = \beta$ and $f = \tilde{\phi}$, respectively. It is worth noting the similarity between the rate equation for $\tilde{\phi}$, which is proportional to $(f - \tilde{\phi})/\tilde{\mu}$, and those for inelastic strain in classical viscoplasticity theory [23]. Here, f can be interpreted as a yield surface for inelastic volumetric strain and $\tilde{\mu}$ controls the relaxation rate to the yield surface. Importantly, the inclusion of $\tilde{\phi}$ enables history-dependent loading to be described that is a direct consequence of inelastic compaction, as illustrated later, and as discussed in detail in Ref. [33]. An appropriate value for μ_c can be easily determined from experimental compaction wave thickness data, whereas the value for $\tilde{\mu}$ is more difficult to determine, but can usually be estimated from the stress relaxation time measured by quasistatic compaction experiments.

Constitutive relations needed to mathematically close Eqs. (1)–(5) include expressions for $P_s(\rho_s, T)$, $e_s(\rho_s, T)$, $\beta(\rho_s, \phi, \tilde{\phi})$, $f(\phi)$, μ_c, and $\tilde{\mu}$. To this end, we use the Hayes equation of state to describe the thermodynamics of solid HMX [1]. The Helmholtz free energy for an initially stress-free solid is given by

$$\Psi_s(\rho_s, T) = c_v \left[(T - T_0)\left(1 + \frac{g}{\rho_{s0}}\left(1 - \frac{\rho_{s0}}{\rho_s}\right)\right) + T \ln\left(\frac{T_0}{T}\right) \right]$$
$$+ t_1 \left[\left(\frac{\rho_s}{\rho_{s0}}\right)^{N-1} - (N-1)\left(1 - \frac{\rho_{s0}}{\rho_s}\right) - 1 \right], \quad (6)$$

from which the thermal and caloric equations of state can be obtained:

$$P_s(\rho_s, T) \equiv \rho_s^2 \left.\frac{\partial \Psi_s}{\partial \rho_s}\right|_T = c_v g (T - T_0) + \frac{A}{N}\left[\left(\frac{\rho_s}{\rho_{s0}}\right)^N - 1 \right], \quad (7)$$

$$e_s(\rho_s, T) \equiv \Psi_s - T \left.\frac{\partial \Psi_s}{\partial T}\right|_{\rho_s} = c_v(T - T_0) - c_v T_0 \frac{g}{\rho_{s0}}\left(1 - \frac{\rho_{s0}}{\rho_s}\right)$$
$$+ t_1 \left[\left(\frac{\rho_s}{\rho_{s0}}\right)^{N-1} - (N-1)\left(1 - \frac{\rho_{s0}}{\rho_s}\right) - 1 \right]. \quad (8)$$

Here, $A = t_1 N(N-1)\rho_{s0}$ and N are the first and second Hayes parameters, respectively. Values for the parameters contained in Eqs. (7) and (8) are listed in Table 1. Expressions for β and f are based on the quasi-static compaction data for granular HMX reported by Coyne et al. [24], and Elban and Chiarito [25], and are given by

Table 1 Parameters values used in the Hayes equation of state for solid HMX [1]

Parameter	Value	Units
A	1.35×10^{10}	Pa
c_v	1.50×10^3	J/(kg K)
g	2.10×10^3	kg/m^3
N	9.8	$\cdots$
T_0	300	K
ρ_{s0}	1.90×10^3	kg/m^3

$$\beta(\rho_s, \phi, \tilde{\phi}) = -\beta_0 \frac{\rho_s}{\rho_{s0}} \phi(\phi - \tilde{\phi}) \frac{\ln(\kappa - (\phi - \tilde{\phi}))}{\kappa - (\phi - \tilde{\phi})} \quad (9)$$

and

$$f(\phi) = \phi_f + c(\phi - \phi_f), \quad (10)$$

where $\beta_0 = 6.0$ MPa, $c = 0.913$, $\kappa = 0.03$, and $\phi_f = 0.655$. This expression for β contains a weak (linear) dependence on solid density as required by thermodynamic restrictions [26]. This constraint may be relaxed with minimal consequences on the predicted solution, but is retained here for completeness. Further, we take $\mu_c = 100$ kg/(s m) so that the model predicts compaction zone lengths and wave speeds commensurate with experiments [27]. A baseline value of $\tilde{\mu} = 9.39 \times 10^{-12}$ s is used in this work as suggested by the fast relaxation response of $\tilde{\phi}$ observed in quasi-static compaction experiments. Because the value of $\tilde{\mu}$ is difficult to determine from dynamic compaction experiments, it will later be varied to characterize its effect on compaction wave structure and localized heating. Equations (1)–(5) and (7)–(10) constitute a nonstrictly hyperbolic system of nine equations and nine unknowns that can be numerically solved for the bulk material response provided that suitable initial and boundary conditions are supplied. Because the equations are hyperbolic, compaction wave structures having discontinuous shocks are admitted.

An important aspect of this work is the localization of bulk thermal energy at the grain scale to form hot-spots. It is, therefore, convenient to obtain an expression for the evolution of granular solid internal energy in terms of Lagrangian derivatives. Using the Eulerian form of mass and momentum conservation to eliminate kinetic energy from the total energy equation, and introducing the Lagrangian derivative $d/dt \equiv \partial/\partial t + u \partial/\partial x$, we obtain

$$\frac{de}{dt} = \frac{de_s}{dt} + \frac{dB}{dt}, \quad (11)$$

where

$$\frac{de_s}{dt} = \underbrace{\frac{(P_s - \beta)}{\rho_s \phi}\frac{d\phi}{dt} + \frac{\beta}{\rho_s \phi}\frac{d\tilde{\phi}}{dt}}_{\text{Compaction}} + \underbrace{\frac{P_s}{\rho_s^2}\frac{d\rho_s}{dt}}_{\text{Compression}} = \frac{de_\phi}{dt} + \frac{de_\rho}{dt} \quad (12)$$

and

$$\frac{dB}{dt} = \frac{\beta}{\rho}\frac{d}{dt}(\phi - \tilde{\phi}). \quad (13)$$

Equations (12) and (13) govern the evolution of thermal energy for the pure phase solid and compaction potential energy for the granular solid, respectively. Thermal energy evolution is affected by both compaction and compression, denoted by de_ϕ/dt and de_ρ/dt, whereas compaction potential energy is affected only by changes in the elastic component of solid volume fraction given by $d/dt(\phi - \tilde{\phi})$. It is shown in Ref. [26] that de_ϕ/dt is the result of compaction induced dissipative processes that lead to an increase in granular solid entropy, whereas the compression energy is recoverable. It is important to note that while the bulk model does not attribute dissipation to specific processes, it can be easily cor-
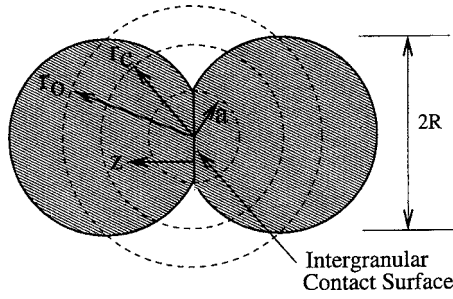
Fig. 2 An illustration of the intergranular contact geometry used for the thermal energy localization strategy

related with loading–unloading compaction experiments to accurately reflect the measured net dissipated work that will constrain the thermal energy localization strategy. Various energetically consistent localization strategies can be formulated in a manner compatible with grain scale heating mechanisms that are believed important for a given problem. One such strategy is outlined in the following section.

3 Grain Scale Heating Model

Because the bulk material response gives no information about the magnitude and duration of temperature fluctuations that are important for combustion initiation, it is necessary to rationally localize bulk thermal energy to describe hot-spot formation at the grain scale. Here, we modify the localization strategy of Gonthier [7] to account for compressive heating and phase change induced by shock compaction. The localization strategy is consistent with mesoscale simulations and grain contact mechanics which indicate that the applied bulk load is transmitted through the material by intergranular contact [28]. This contact results in inelastic deformation and friction near contact surfaces. As illustrated in Fig. 2, we track the evolution of thermal energy within solid regions surrounding intergranular contact surfaces referred to as *localization spheres*. The number of localization spheres per unit volume, $n_c(x,t)$, is related to the number of contact points per grain, γ, and the grain number density by $n_c = \frac{1}{2}\gamma n$; the prefactor $1/2$ appears because each localization sphere involves contact between two grains. The localization spheres are assumed to be evenly distributed, and have radii r_0, where $r_0 = R(\frac{1}{2}\gamma)^{-1/3}$. The expression for n_c can be combined with the expression for r_0 to obtain $n_c = \phi/(\frac{4}{3}\pi r_0^3)$; thus, all solid mass is encompassed by the localization spheres.

Many phenomenological strategies can be established to prescribe the partitioning of dissipated energy within a localization sphere based on the underlying physics. Because there exists uncertainty about stress and temperature distributions within the neighborhood of intergranular contact surfaces for real systems, and their dependence-on loading rate, microstructure, etc., we choose a simple, but plausible, strategy that uniformly deposits bulk dissipated compaction energy (given by de_ϕ/dt) within expanding material volumes of radius $r_c(x,t) \leq r_0$ centered at the contact surfaces; this radius defines a *localization center*. The localization center should encompass all material involved in dissipative heating due to friction and plastic deformation; thus, its volume should increase due to plastic flow. The initial value for r_c is taken as the radius of the intergranular contact surface, a, at the onset of plastic deformation within the grain. This assumption is reasonable in that, prior to the onset of plastic deformation, most dissipated energy will be due to intergranular friction and will, thus, be localized near the contact surface within the region $r < r_c(x,0)$. We equate the volumetric rate of work done by the plastic flow stress P_Y to the bulk volumetric compaction induced dissipated energy given by Eq. (12); the following evolution equation for r_c results:

$$\frac{dr_c}{dt} = \frac{\rho_s \phi}{4\pi n_c r_c^2 P_Y} \frac{de_\phi}{dt}. \tag{14}$$

Though we uniformly deposit dissipated energy within the localization centers, and describe the evolution of r_c solely due to plastic deformation, it is possible to develop more refined localization strategies that better describe both frictional and plastic heating, and viscous dissipation with a liquid phase, using this modeling framework. Such descriptions may require the introduction and evolution of multiple localization volumes within a single localization sphere. We also assume here that compressive heating is a bulk process and, thus, uniformly affects all material within a localization sphere; details are given below.

For consistency, we require that the evolution of mass, linear momentum, and thermal energy at the grain scale locally equals that given by the bulk model. It can be easily shown that the mass and linear momentum constraints are trivially satisfied by requiring that the velocity of each grain is given by the bulk velocity u; i.e.,

$$\frac{d}{dt}\left(\int_0^L \rho_s \phi dx\right) \equiv \frac{d}{dt}\left(\int_0^L \rho_s n_c 4\pi \int_0^{r_0} r^2 dr dx\right),$$

$$\frac{d}{dt}\left(\int_0^L \rho_s \phi u dx\right) \equiv \frac{d}{dt}\left(\int_0^L \rho_s n_c u 4\pi \int_0^{r_0} r^2 dr dx\right),$$

where $n_c = \phi/(\frac{4}{3}\pi r_0^3)$, r is radial position within the localization sphere, and x is position within the granular system of total length L. The thermal energy constraint is given by

$$\frac{d}{dt}\left(\int_0^L \rho_s \phi e_s dx\right) = \frac{d}{dt}\left(\int_0^L \rho_s n_c 4\pi \int_0^{r_0} r^2 \hat{e} dr dx\right), \tag{15}$$

where $\hat{e}$ is the specific internal energy of the solid within a localization sphere. Throughout this paper, variables labeled with a "hat" $(\hat{\cdot})$ are associated with the localization sphere and vary not only with x and t but also with r. The left-hand side of this equation is the evolution of bulk thermal energy. The right-hand side of this equation is the evolution of integrated thermal energy at the grain scale. It is important to note that bulk compression and expansion will cause the radius of both the spherical grains, R, and the localization spheres, r_0, to change. In this study, we assume that the rate of change of localization radius is small compared to the compaction wave speed (i.e., $(dr_0/dt)/D \ll 1.0$, where D is compaction wave speed) and, thus, we take r_0 to be constant in our localization model. In particular, we take $r_0 = R_0(\frac{1}{2}\gamma)^{-1/3}$ for subsonic compaction waves and $r_0 = R_{sh}(\frac{1}{2}\gamma)^{-1/3}$ for supersonic waves, where R_0 is the ambient particle radius and R_{sh} is the shocked particle radius whose value is determined by Rankine–Hugoniot shock relations. We believe this assumption to be reasonable for all compaction waves studied in this work. With this assumption, and using the definition of n_c, Eq. (15) reduces to

$$\frac{de_s}{dt}(x,t) = \frac{3}{r_0^3}\int_0^{r_0} r^2 \frac{d\hat{e}}{dt}(x,r,t)dr. \tag{16}$$

The evolution of thermal energy within a localization sphere is given by

$$\frac{d\hat{e}}{dt} = -\frac{1}{\rho_{s0} r^2}\frac{\partial}{\partial r}(r^2 q) + \frac{\hat{S}_\phi}{\rho_{s0}} + \frac{\hat{S}_\rho}{\rho_{s0}}, \tag{17}$$

where $q = -k\partial \hat{T}/\partial r$ is the conductive heat flux, $\hat{S}_\phi$ and $\hat{S}_\rho$ are the volumetric heating rates due to inelastic compaction and compression, respectively, and ρ_{s0} is the initial solid density for subsonic waves and the shocked solid density for supersonic waves. Though we assume that the localization sphere is incompressible throughout the compaction zone (i.e., r_0=constant), bulk com-

pressive heating is still accounted for at the grain scale by the source term $\hat{S}_\rho$. Energetically consistent expressions for $\hat{S}_\phi$ and $\hat{S}_\rho$ are obtained by substituting Eq. (17) into Eq. (16), integrating, imposing the adiabatic conditions $q(x,0,t)=q(x,r_0,t)=0$ and equating the result to Eq. (12). The adiabatic condition at $r=r_0$ is a consequence of assuming a symmetric grain packing arrangement.

$$\hat{S}_\phi(x,r,t) = \begin{cases} \left[\dfrac{r_0}{r_c(x,t)}\right]^3 \rho_{s0} \dfrac{de_\phi}{dt}(x,t) & \text{for } 0 \leq r \leq r_c(x,t), \\ 0 & \text{for } r_c(x,t) < r \leq r_0, \end{cases}$$
$$(18)$$

$$\hat{S}_\rho(x,t) = \rho_{s0} \frac{de_\rho}{dt}(x,t). \qquad (19)$$

Here, functional dependencies are explicitly shown to indicate the coupling of bulk and grain scale variables. The bulk and grain scale models are coupled in an energetically consistent manner through the source terms $\hat{S}_\phi$ and $\hat{S}_\rho$. Though this energy partitioning strategy is not unique, we believe that it reasonably describes grain scale localized heating based on the success of a similar strategy used by the lead author to predict experimentally measured sustained combustion thresholds for granular HMX due to mild impact [7]. Nonetheless, this partitioning should be more carefully scrutinized in the future based on numerically resolved mesoscale simulations.

We also account for the energetics of solid–liquid phase change as it may significantly affect hot-spot temperature. Further, HMX is known to undergo phase change prior to combustion. HMX melts near $T_m^0 = 520$ K at atmospheric pressure $P_m^0 = 100$ kPa; the latent heat of fusion is $q_m^0 = 220$ kJ/kg. An estimate for the variation in melt temperature with pressure can be determined by the Kraut–Kennedy relation (a detailed discussion of the application of this relation to HMX melting is given in Ref. [29]). For the highest pressures considered in this work, $P_m \approx 500$ MPa, the melt temperature increases to only $T_m \approx 600$ K. Thus, we assume isothermal phase change, and take

$$\frac{d\hat{e}}{dt} = \begin{cases} c_v \dfrac{d\hat{T}}{dt} & \text{for } T \neq T_m^0 \\ q_m^0 \dfrac{d\hat{\chi}}{dt} & \text{for } T = T_m^0 \end{cases} \qquad (20)$$

where $0 \leq \hat{\chi} \leq 1$ is the liquid mass fraction. For simplicity, we assume that the value of the specific heat c_v is constant and the same for both the solid and liquid, and that the solid and liquid densities are equal; as such, we only describe leading-order effects of phase change energetics.

4 Analysis and Discussion

Predictions are given for the bulk and grain scale response of inert HMX ($\phi_0 = 0.81$) due to steady compaction waves. Though the analysis of steady waves is incapable of describing time-dependent phenomena associated with variable speed impact or wave propagation through materials having spatially nonuniform porosity, it does provide a simple means of analyzing the more fundamental problem of dynamic compaction of spatially homogeneous materials by constant speed impact. To this end, the model equations are expressed as a coupled system of partial differential equations (PDEs) in a steady wave frame using the transformation $\xi = x - Dt$ and $v = u - D$, where D is the wave speed, and ξ and v are position and velocity measured relative to the wave. Equilibrium solutions of the steady equations are first analyzed to identify compaction wave end states, and these end states are compared to experimental data. Next, the initial boundary volume problem (IBVP) defined by the steady equations, the wave speed,

and the ambient material state, is numerically solved to predict bulk compaction wave structure and localized heating within the compaction zone. Predicted structures for both subsonic and supersonic compaction waves are illustrated for speeds of $D = 748.2$ and 3500 m/s corresponding to impact speeds of $u_p = 106$ and 1053 m/s, respectively; these speeds are chosen because they demonstrate interesting features that are also observed in mesoscale simulations. The sensitivity of the model to variations in key parameters is then explored. Lastly, model predictions are compared to the predictions of Menikoff and Kober [11] obtained by detailed mesoscale simulations.

The steady model equations describing compaction wave structure are given by

$$\frac{d\phi}{d\xi} = \frac{\phi(1-\phi)}{v\mu_c}(P_s - \beta), \qquad (21)$$

$$\frac{d\tilde{\phi}}{d\xi} = \frac{1}{v\tilde{\mu}}(f - \tilde{\phi}), \qquad (22)$$

$$\frac{\partial \hat{T}}{\partial \xi} = \frac{\alpha}{v}\left[\frac{\partial^2 \hat{T}}{\partial r^2} + \frac{2}{r}\frac{\partial \hat{T}}{\partial r}\right] + \frac{\hat{S}_\phi}{\rho_{s0}c_v v} + \frac{\hat{S}_\rho}{\rho_{s0}c_v v}, \qquad (23)$$

$$\frac{\partial \hat{\chi}}{\partial \xi} = \frac{c_v \alpha}{q_m v}\left[\frac{\partial^2 \hat{T}}{\partial r^2} + \frac{2}{r}\frac{\partial \hat{T}}{\partial r}\right] + \frac{\hat{S}_\phi}{\rho_{s0}q_m v} + \frac{\hat{S}}{\rho_{s0}q_m v}, \qquad (24)$$

$$\frac{dr_c}{d\xi} = \frac{\rho_s \phi}{4\pi n_c r_c^2 P_Y}\frac{de_\phi}{d\xi}, \qquad (25)$$

where $\alpha = k/(\rho_{s0}c_v)$ is the thermal diffusivity of the pure phase solid. All remaining variables can be expressed as a function of the integration variables ϕ, $\tilde{\phi}$, $\hat{T}$, $\hat{\chi}$, and r_c. To this end, the conservation equations (i.e., the first three components of Eqs. (1)–(4)) expressed in the steady frame are homogeneous ODEs that can be directly integrated, and the initial conditions $\phi(0) = \tilde{\phi}(0) = \phi_0$, $T_0 = 300$ K, and $P_0 = 0$ Pa applied, to obtain the following algebraic relations:

$$\rho v = -\rho_0 D, \quad \rho v^2 + P = \rho_0 D^2, \quad e + \frac{v^2}{2} + \frac{P}{\rho} = \frac{D^2}{2}. \qquad (26)$$

Here, the notation subscript "0" refers to properties of the ambient state. These relations, together with the equations of state for the pure phase solid, given by Eqs. (7) and (8), are sufficient to express the bulk solid variables $\rho_s = \rho/\phi$, v_s, and $P_s = P/\phi$ in terms of ϕ and $\tilde{\phi}$. The required mathematical operations are straightforward, and are omitted for brevity. Initial conditions for the remaining integration variables are $\hat{T}(0,r) = T_0$, $\hat{\chi}(0,r) = 0$, and $r_c(0) = 1.6\pi R^* Y/(2E^*)$ for an unshocked solid, where $R^* = R/2$, and $E^* = E/(2(1-\nu^2))$. Here, E, Y, and ν are the elastic modulus, yield strength, and Poisson's ratio for the pure solid; we take $R = 25$ μm, $E = 24$ GPa, $Y = 0.37$ GPa, and $\nu = 0.2$ which are representative of HMX. Initial conditions for a supersonic wave are based on the shocked state rather than the ambient state. The shocked state is determined from the Rankine–Hugoniot relations given in vector form by $[\mathbf{w}]_-^+ D = [\mathbf{f}(\mathbf{w})]_-^+$, where $\mathbf{w}$ and $\mathbf{f}(\mathbf{w})$ are given by Eqs. (2) and (3), and the notation $[\cdot]_-^+$ denotes a jump in the enclosed quantity across the shock. Because $[\phi]_-^+ = [\tilde{\phi}]_-^+ = 0$ across the discontinuity, these relations reduce to the classical ones for a pure phase solid. These classical relations are combined with Eqs. (7) and (8) to obtain a single transcendental expression for ρ_{sh} which is solved numerically using a Newton–Raphson technique. With ρ_{sh} known, all remaining shocked solid variables can be determined. Equation (23) is integrated for $\hat{T} \neq T_m^0$. Equation (24) is integrated through the melt region ($0 < \hat{\chi} < 1$) for which $\hat{T} = T_m^0$.

Equations (21)–(25) are numerically solved using a Method of Lines (MOL) technique. Central differencing is used to approximate the radial derivatives on a $N_r = 100$ node grid, and the resulting system of ODEs in ξ are integrated using the implicit routine ODE15s contained in the MATLAB software package. The numerical algorithm is nominally fourth-order accurate in ξ and second-order accurate in r. Numerical convergence is demonstrated in Sec. 4.2 based on the variation in grain scale heating predictions with increasing grid resolution. A typical simulation requires approximately 2 min of CPU time on a Linux, 2.0 GHz, Pentium IV workstation.

4.1 Compaction Wave End States. Compaction wave end states are obtained by an algebraic analysis that is independent of wave structure; as such, this analysis provides no information about hot-spot formation but does give a simple framework for comparing bulk predictions to experimental data. From Eqs. (21)–(25), it is seen that equilibrium states, formally corresponding to $\xi \to -\infty$, are defined by $P_s = \beta$ and $\tilde{\phi} = f$, with radially uniform temperature $\hat{T}$ and solid mass fraction $\hat{\chi}$ within the localization sphere. The equilibrium value of $\hat{T}$ approaches the bulk equilibrium temperature T_s as a consequence of having energetically consistent bulk and grain scale models. It is conventional to analyze potential end states geometrically in the bulk pressure-specific volume (P-ν) phase plane. An expression for the Rayleigh line in this plane is obtained by combining the integrated form of mass and momentum conservation, given by the first and second expressions of Eq. (26), to obtain

$$P = \left(\frac{D}{\nu_0}\right)^2 (\nu_0 - \nu). \tag{27}$$

For a given ambient state, the slope of this line in the P-ν plane is dependent only on compaction wave speed D. The expression can be combined with the integrated form of the energy equation to obtain the Hugoniot curve for the granular solid:

$$e - e_0 = \tfrac{1}{2} P(\nu_0 - \nu). \tag{28}$$

This expression gives the mass specific work done in compressing the granular material from its initial specific volume, ν_0, to its final values P and ν. Again, not all of this work is recoverable due to internal dissipation. Using Eqs. (7) and (8), and the equilibrium conditions $P_s = \beta(\rho_s, \phi, \tilde{\phi})$ and $\tilde{\phi} = f(\phi)$, compaction wave end states are given by the intersection of the Rayleigh line and Hugoniot curve. The piston speed is then computed from the wave speed and the equilibrium value of ν by $u_p = (1 - \nu/\nu_0)D$.

Figure 3 summarizes predicted compaction wave end states in the P-ν and D-u_p phase planes for granular HMX having $\phi_0 = \phi_f = 0.655$, and compares these predictions to the experimental values reported by Sheffield [30] and Sandusky and Liddiard [31] for granular HMX having similar initial densities. Also shown are the predicted shock Hugoniots for solid HMX and experimental data for PBX 9404, a pressed plastic bonded explosive consisting of 94% granular HMX and 6% binder. The PBX data [32] is included to illustrate that the shock Hugoniot for solid HMX is approached by materials possessing a high granular HMX loading density. Several features are noteworthy. First, as seen in Fig. 3(a), the model reasonably describes the P-ν response of the initially stress free material. Low pressure compaction of loose HMX ($\nu_0 = 0.79$ cm³/g) is largely inelastic resulting in a significant decrease in material volume due to grain fracture and rearrangement, and plastic deformation. The material stiffens as the grains consolidate at higher pressure, and the granular material response tracks that of the pure phase solid. Second, the model accurately describes the D-u_p response as shown in Fig. 3(b). Weak impact results in low pressure compaction waves that propagate at speeds much less than the ambient solid acoustic speed ($D < c_{s0} = 2767$ m/s) due to internal dissipation. There is predicted a

(a)

(b)
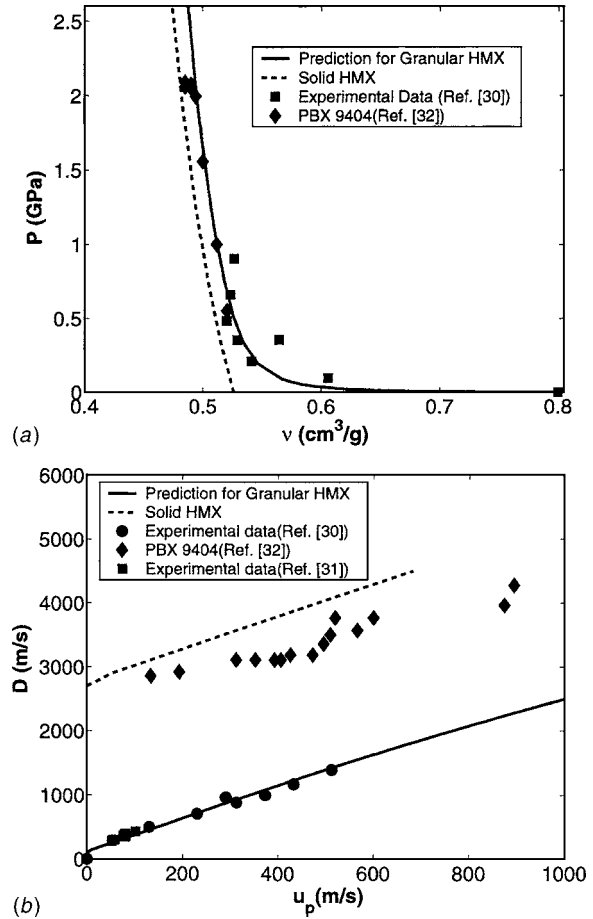
Fig. 3 Predicted and measured Hugoniot curves for granular and solid HMX in (a) P-ν and (b) D-u_p planes. The PBX 9404 data is for a pressed mixture of 94% granular HMX and 6% plastic binder.

nearly linear increase in wave speed over the range of impact conditions considered in this work. The minimum predicted wave speed for the existence of a steady compaction wave is $D = 290$ m/s. This analysis demonstrates that the model reasonably captures experimentally measured results.

Finally, we remark that the Hugoniot curves depend on initial material density which is characteristic of materials with memory. Dynamic loading of precompacted materials (i.e., $\phi_0 > \phi_f$) exhibit time-dependent, two-wave structures for a range of impact speeds (typically $10 < u_p < 100$ m/s) that depend on ϕ_0 which are similar to those in dynamically loaded solids possessing a Hugoniot elastic limit. Such compaction waves consist of a lead viscoelastic precursor followed by a slower viscoplastic wave. The inelastic component of solid volume fraction is constant through the precursor ($\tilde{\phi} = \phi_0$), whereas it irreversibly increases through the viscoplastic wave. Here, the meaning of the terms viscoelastic and viscoplastic are conventional [23] in that the granular solid stress depends on both volumetric strain and strain rate (loading history) by its dependence on ϕ and $\tilde{\phi}$, respectively. A consequence of this modeling attribute is that experimentally observed features such as strain hardening, stress relaxation, and hysteresis can be predicted; a recent discussion and analysis is given in Ref. [33]. We restrict our analysis in this paper to steady waves, though some of these waves still retain features of unsteady two-wave structures.

4.2 Compaction Wave Structure. Results for a typical subsonic and supersonic wave structure are presented in this section. For both cases, the ambient stress free material has a solid volume
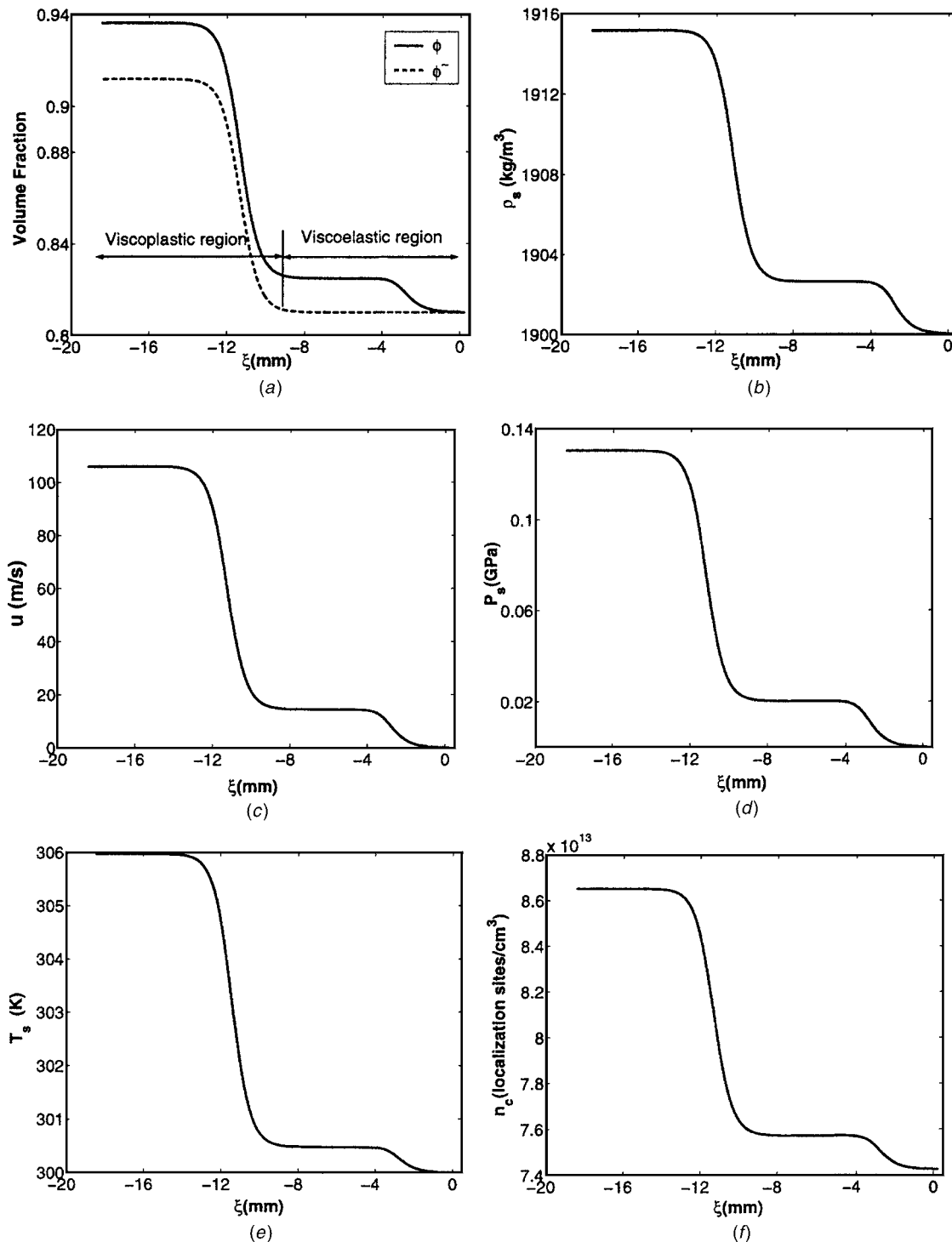
Fig. 4 Predicted variation in bulk quantities through the compaction zone for $\phi_0 = 0.81$, $u_p = 106$ m/s, and $D = 748.2$ m/s: (a) Solid volume fraction, (b) solid density, (c) velocity, (d) solid pressure, (e) solid temperature, and (f) grain number density.

fraction of $\phi_0 = 0.81$; thus, the material has been pre-compacted. This value is chosen so that we can directly compare our model predictions to results of the mesoscale simulations performed by Menikoff and Kober [11] as discussed in Sec. 4.4.

Figure 4 gives predictions for the variation in bulk quantities through the compaction zone for $u_p = 106$ m/s and $D = 748.2$ m/s. Here, the head of the compaction wave is located at $\xi = 0$ mm and its structure lies in the region $\xi < 0$ mm. This piston speed is only slightly higher than that needed for a steady wave to exists; at lower speeds, unsteady two-wave structures are pre-

dicted. Remnants of the viscoelastic precursor are evident in the figure. To our knowledge, experimental evidence of such precursors resulting from mild impact is lacking, possibly because of inadequate resolution, though mesoscale simulations clearly indicate the presence of such waves [11,34]. A continuous, dispersed structure is predicted. All variables shown monotonically increase from their initial to final values through both the viscoelastic and viscoplastic regions of the compaction wave. As seen in Fig. 4(e), a peak bulk temperature of approximately $T_s = 306$ K is predicted, which is well below the ignition temperature of HMX (T_{ig}
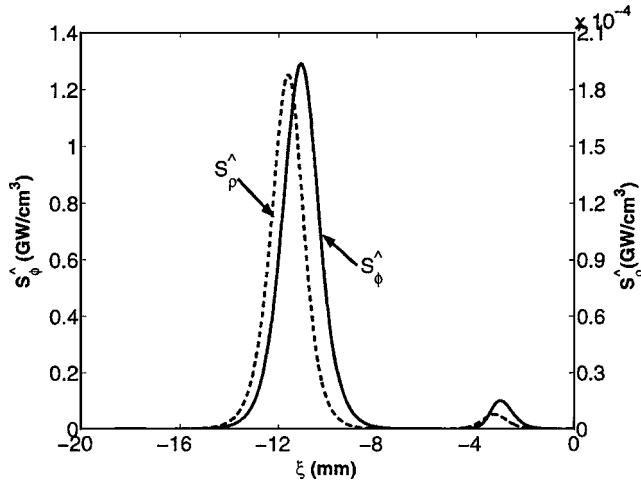
Fig. 5 Predicted variation in the compressive ($\hat{S}_\rho$) and inelastic ($\hat{S}_\phi$) heating rates through the compaction zone for ϕ_0 =0.81, u_p=106 m/s, and D=748.2 m/s

≈ 600 K), yet confined DDT experiments clearly indicate the onset of sustained combustion for piston speeds of this magnitude. Thus, ignition models based on bulk temperature are too insensitive to accurately predict combustion thresholds.

We now describe the predicted localized heating response of the material. The number density of localization sites surrounding intergranular contact surfaces, n_c, increases as the material is compacted (see Fig. 4(f)). The volumetric compressive and inelastic heating rates within the vicinity of these localization sites, $\hat{S}_\rho$ and $\hat{S}_\phi$, respectively, are shown in Fig. 5. For this impact speed, compressive heating is several orders of magnitude less than the maximum inelastic heating rate ($\hat{S}_\phi \approx 1.28$ GW/cm^3) and is thus inconsequential. The corresponding grain and localization radii, and grain scale temperature, are shown in Fig. 6. Inelastic heating occurs within the spherical volume defined by the localization center radius r_c; as mentioned, this volume characterizes the mass of plastically deformed material due to integranular contact near the surface located at $r=0$ μm. A measure of plastic strain for our localization model is the fraction of solid volume affected by plastic work, $\epsilon_{pl}=(r_c/r_0)^3$. For this simulation, $\epsilon_{pl}=0.0089$; thus, little material is plastically deformed. The viscoelastic region of the wave (for which $\tilde{\phi}=\phi_0$) induces a weak hot spot near the integranular contact surface which is rapidly quenched by thermal conduction prior to the onset of viscoplastic heating. More intense precursor hot spots may be induced under different loading conditions. Plastic work occurring within the viscoplastic region is highly localized near the contact surface resulting in a peak hot-spot temperature of 970.3 K. Temperatures of this magnitude may trigger prompt combustion initiation consistent with DDT experiments. However, the hot-spot duration is short (relative to the compaction zone length) as it contains insufficient thermal inertia to overcome conductive losses. A fully coupled thermochemical analysis would be required to determine whether such a hot spot will result in sustained combustion; nonetheless, the likelihood of a thermal explosion occurring within the compaction wave triggered by our predicted hot-spot temperature is analyzed below based on a asymptotic analysis. Importantly, melting is shown to reduce the predicted hot-spot temperature by 100.4 K based on an equivalent simulation without phase change. The melt region is identified in Fig. 6(b). An estimate for the volume fraction of liquid formed is given by $\phi_l=(r_l/r_0)^3\phi$, where r_l is the liquid core radius. For this simulation, $\phi_l \approx 0.0038 \ll \phi$; thus, we are justified in ignoring the presence of a bulk liquid phase.

To estimate the explosion length induced by the predicted hot-
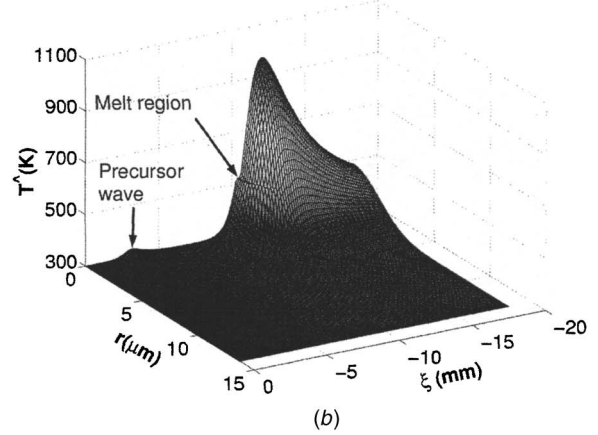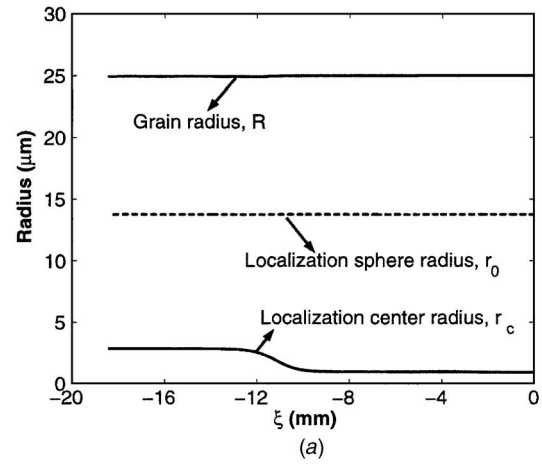

Fig. 6 Predicted variation in (a) grain radius, localization radii, and (b) grain temperature through the compaction zone for ϕ_0=0.81, u_p=106 m/s, and D=748.2 m/s

spot temperature within the compaction zone, we perform a thermal explosion analysis for a two-phase reacting system (HMX $\rightarrow$ product gas). To this end, we use the rate expression for HMX combustion given in Ref. [35], and assume (1) a well-stirred reactor, (2) an incompressible solid, (3) there is initially no gas present, and (4) the solid and product gas are in thermal equilibrium. With these assumptions, a two-phase thermal explosion model can be expressed by

$$\rho_{s0}\phi + \rho_g\phi_g = \rho_{s0}\phi_0, \tag{29}$$

$$\rho_{s0}\phi e_s + \rho_g\phi_g e_g = \rho_{s0}\phi e_{s0}, \tag{30}$$

$$\frac{d\lambda}{dt} = Z(1-\lambda)\exp\left(-\frac{T^\ddagger}{T}\right), \tag{31}$$

$$\lambda = \frac{\rho_g\phi_g}{\rho_{s0}\phi + \rho_g\phi_g} = 1 - \frac{\phi}{\phi_0}, \tag{32}$$

$$\phi + \phi_g = 1, \quad e_s = c_vT + q, \quad e_g = c_{vg}T. \tag{33}$$

Equations (29) and (30) are statements of mass and energy conservation, respectively. Equation (31) gives the evolution of reaction progress ($0 \leqslant \lambda \leqslant 1$, where $\lambda=1$ is complete reaction) and Eq. (32) relates the reaction progress variable λ to the solid volume fraction ϕ. Equation (33) gives the saturation constraint and equations of state for the solid and gas, respectively. Other variables in these equations are the gas-phase density and volume fraction, ρ_g and ϕ_g. Constant parameters include the pre-exponential factor $Z=5.0\times10^{19}$ s^{-1}, activation temperature $T^\ddagger=2.65\times10^4$ K, gas

phase specific heat $c_{vg}=1000$ J/(kg K), and specific heat of combustion $q=5.84 \times 10^6$ J/kg. The model consists of seven equations for seven unknowns and is, thus, mathematically closed. These equations can be nondimensionalized and combined, with the energy equation [Eq. (30)] differentiated with respect to time, to obtain an evolution equation for temperature:

$$\frac{dT^*}{dt^*} = \frac{\pi_1[\pi_3 + (1-\pi_2)T^*]^2}{1-\pi_2+\pi_3} \exp\left(\frac{T^*-1}{\pi_1 T^*}\right), \quad (34)$$

where $T^*=T/T_c$, $t^*=t/[(\pi_1/Z)\exp(T^{\ddagger}/T_c)]$, $\pi_1=T_c/T^{\ddagger}$, $\pi_2=c_{vg}/c_v$, and $\pi_3=q/(c_v T_c)$; here, T_c is a characteristic hot-spot temperature. We assume $\pi_1 \ll 1$, and seek a first-order asymptotic solution of the form $T^*=1+\pi_1 T_1^*(t)$. Substituting this expression into Eq. (34) gives the following ODE to leading order in π_1:

$$\frac{dT_1^*}{dt^*} = (1-\pi_2+\pi_3)\exp(T_1^*). \quad (35)$$

This equation can be analytically solved, with the initial condition $T_1^*(0)=0$ applied, to obtain $T_1^*=-\ln[1-(1-\pi_2+\pi_3)t^*]$. This solution becomes unbounded for $t^*=(1-\pi_2+\pi_3)^{-1}$ which gives an estimate for the time to thermal explosion. An estimate for thermal explosion length with respect to the steady wave is then given by

$$l_{\text{ex}} \equiv Dt_{\text{ex}} = \frac{c_v T_c^2 D}{Z T^{\ddagger}[q + (c_v - c_{vg})T_c]} \exp\left(\frac{T^{\ddagger}}{T_c}\right). \quad (36)$$

For the subsonic compaction wave structure discussed above, we take $T_c=970.3$ K resulting in $l_{\text{ex}}=0.092$ μm. Because this value for explosion length is less than the predicted hot-spot length shown in Fig. 5, we anticipate that the hot-spot temperatures induced by this wave are sufficient to trigger sustained combustion of the material in agreement with experimental observations of DDT for granular HMX.

The numerical convergence of our algorithm is briefly demonstrated for the subsonic compaction wave structure just discussed. The plots in Fig. 7 show the predicted variation in (a) the grain temperature $\hat{T}(\xi, r=0\ \mu\text{m})$, and (b) the radial grain temperature distribution $\hat{T}(\xi=-12.07\ \text{mm}, r)$, with the number of radial grid points N_r for $10 \leq N_r \leq 500$. Because the peak grain temperature occurs at or near $r=0$ μm in all of our simulations (and at $\xi=-12.07$ mm for the present case), we believe it is an appropriate quantity for establishing convergence. For the cases shown, little difference in the predicted solutions exists for radial grids in excess of $N_r=100$ nodes, though a significant increase in computational time occurred for approximately $N_r>200$ due to overhead of the MATLAB software used to integrate the model equations. Thus, we chose $N_r=100$ for all simulations performed in this work.

Figure 8 gives predictions for the variation in bulk quantities through the compaction zone for $u_p=1053$ m/s and $D=3500$ m/s. This simulation illustrates the structure of a typical supersonic compaction wave, and is the maximum impact speed considered in this work. Because $D>c_{s0}=2767$ m/s, a discontinuous shock is predicted at the head of the compaction wave. All solid thermodynamic variables increase across the shock as does the particle velocity measured relative to the lab frame. However, the solid volume fraction is continuous across the shock because the dynamic compaction equation is a linearly degenerate characteristic field. As seen in Fig. 8(a), a thin viscoelastic region immediately follows the shock through which $\bar{\phi}=\phi_0$, and it is followed by a thick viscoplastic region through which $\tilde{\phi}$ increases. The wave strength is nearly sufficient to compact the material to its theoretical maximum density ($\phi \approx 1$). It is noted that the compaction zone length is approximately 250 μm, and is approaching the order of a single grain diameter. The bulk temperature is seen in Fig. 8(e) to reach a maximum value of $T_s=663.0$ K$>T_{ig}$ sug-
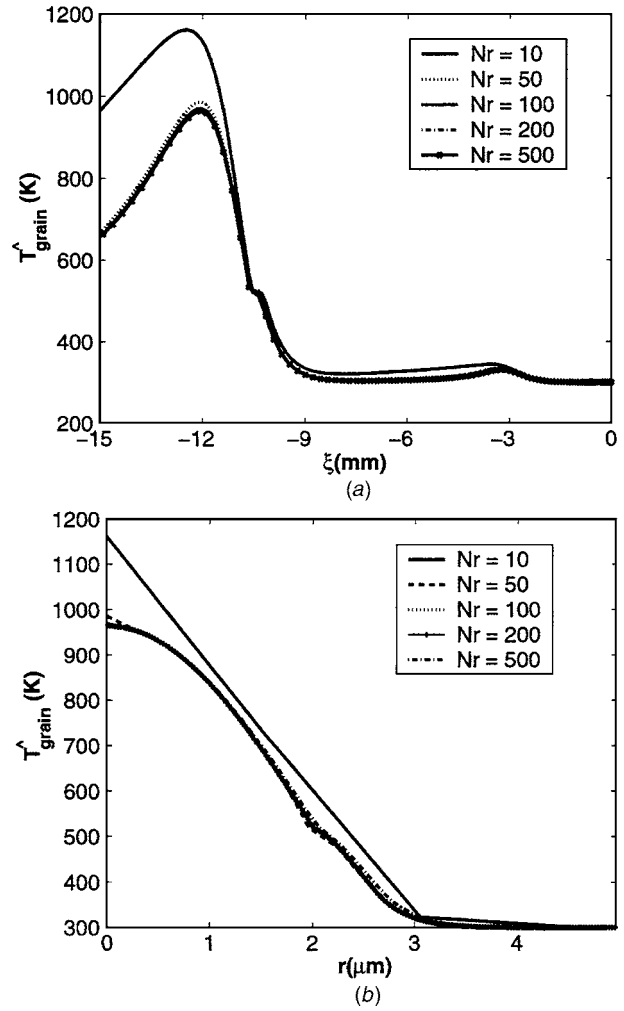
Fig. 7 Convergence of the numerical algorithm for the subsonic compaction wave structure: (a) **predicted variation in grain temperature at the center of the localization volume through the compaction wave,** $\hat{T}_{\text{grain}}=\hat{T}(\xi, r=0\ \mu\text{m})$; (b) **radial distribution of grain temperature at the location** $\xi=-12.07$ **mm,** $\hat{T}_{\text{grain}}=\hat{T}(\xi=-12.07\ \text{mm}, r)$. **Here,** N_r **is the number of radial grid points within the localization sphere.**

gesting that complete phase change and bulk combustion of the material would occur. This result is plausible for such high impact speeds, though the localized heating prediction indicates that material in the vicinity of grain contact surfaces is not uniformly heated to high temperature. The volumetric compressive and inelastic heating rates are shown in Fig. 9. The compressive heating rate is discontinuous at the shock; while significantly larger than that for the subsonic case, it remains several orders of magnitude smaller than the peak inelastic heating rate. This result does not imply that compressibility is unimportant altogether as it does give rise to the formation of the lead solid shock, but its role in localized heating remains inconsequential. The inelastic heating rate is continuous throughout the compaction wave (not apparent in the figure due to the plot scale), reaching a maximum value of $\hat{S}_\phi=2.4 \times 10^4$ GW/cm^3 immediately following the shock. The predicted evolution of grain radius, localization radii, and grain scale temperature are shown in Fig. 10. Both the grain and localization radius, R and r_0, respectively, discontinuously decrease across the shock due to the increase in solid density, and subsequently decrease slightly through the compaction zone. Most of the solid mass has plastically deformed as reflected by the esti-
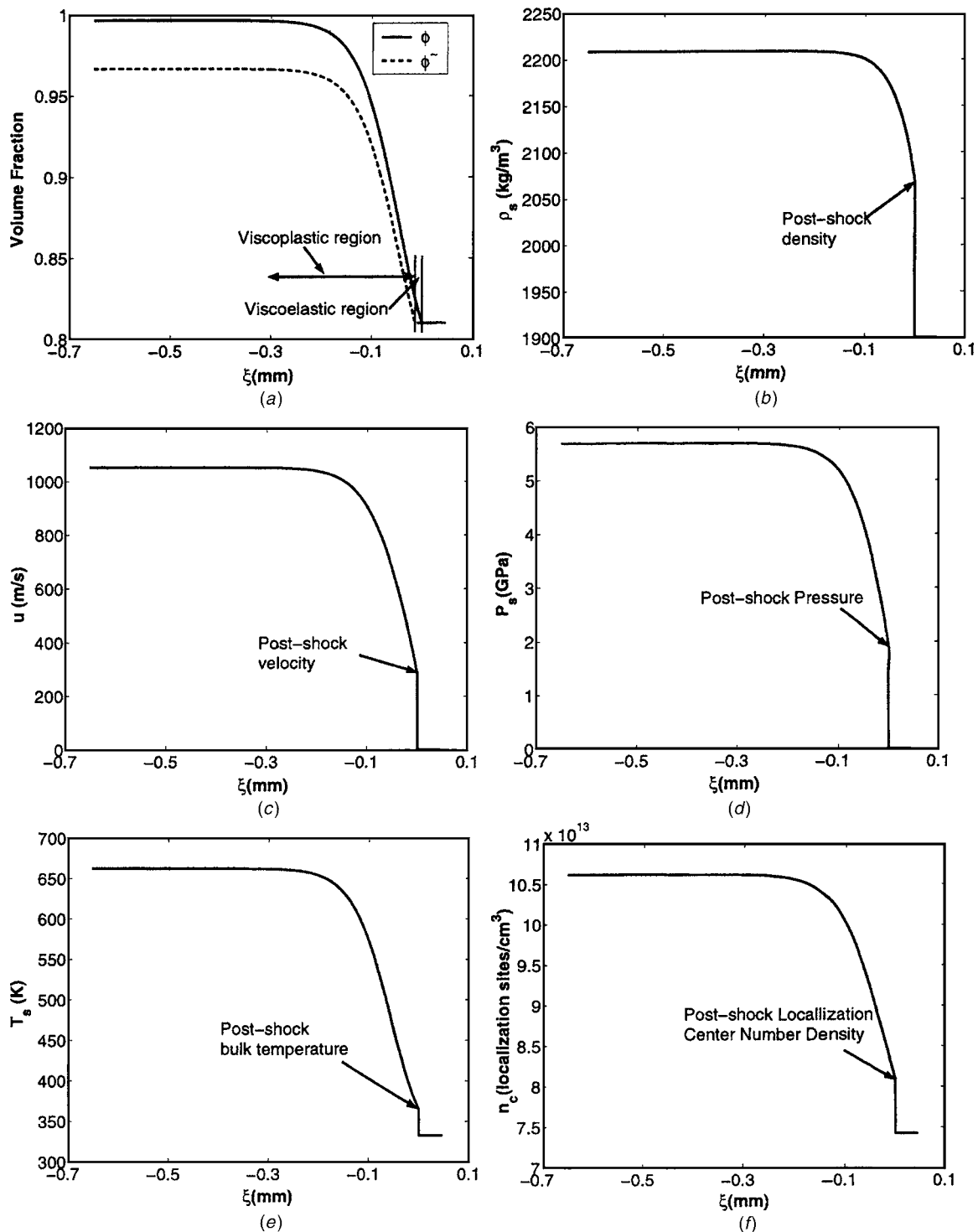
Fig. 8 Predicted variation in bulk quantities through the compaction zone for $\phi_0=0.81$, $u_p=1053$ m/s, and $D=3500$ m/s: (a) Solid volume fraction, (b) solid density, (c) velocity, (d) solid pressure, (e) solid temperature, and (f) grain number density.

mate for volumetric plastic strain $\epsilon_{pl}=(r_c/r_0)^3=0.812$. All material within the localization sphere is simultaneously heated to temperature $T_{sh}=332.6$ K across the shock. Subsequently, inelastic heating causes the temperature near the contact surface to rapidly increase to approximately 2700 K, while the temperature in the grain interior gradually increases due to progressive spreading of plastic deformation. Thermal conduction plays an insignificant role at this impact speed resulting in long duration hot spots (relative to the wave thickness) that would rapidly induce combustion.

An explosion length of $l_{ex} \ll 1$ nm is predicted by Eq. (36) indicating prompt initiation of sustained combustion. The phase change front is evident near the exterior of the localization sphere. The volume fraction of liquid formed is $\phi_l=(r_l/r_0)^3\phi \approx 0.47$. Also, gas phase products likely exist at elevated temperature due to sublimation and vaporization. Though ignored here, these results suggest that multiphase fluid mechanics may play an important role in hot-spot formation for strong impact.
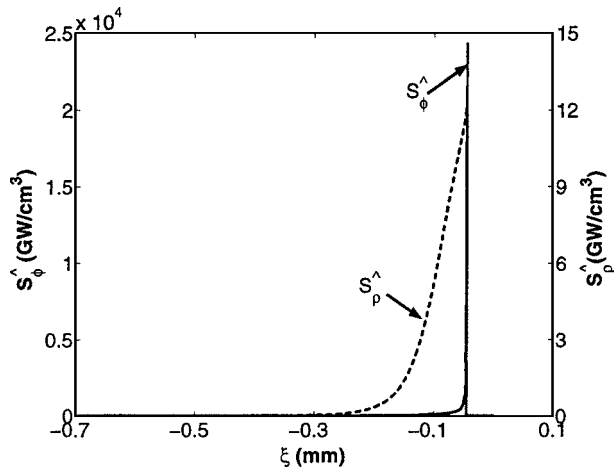
Fig. 9 Predicted variation in the compressive ($\hat{S}_\rho$) and inelastic ($\hat{S}_\phi$) heating rates through the compaction zone for ϕ_0 =0.81, u_p=1053 m/s, and D=3500 m/s

4.3 Parametric Response. Simulations are performed to determine the model's sensitivity to variations in key parameters such as impact speed u_p, initial solid volume fraction ϕ_0, and dimensionless ratio $\Omega \equiv \mu_c/(\rho_{s0}D^2\tilde{\mu})$. Here, Ω is the ratio of relaxation rates at which the no-load volume fraction $\tilde{\phi}$ and the total solid volume fraction ϕ approach their respective equilibrium
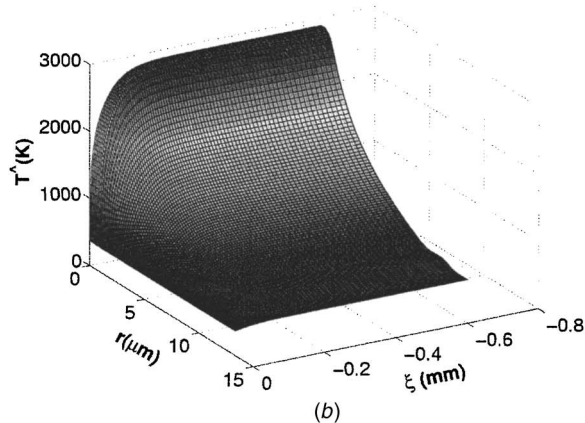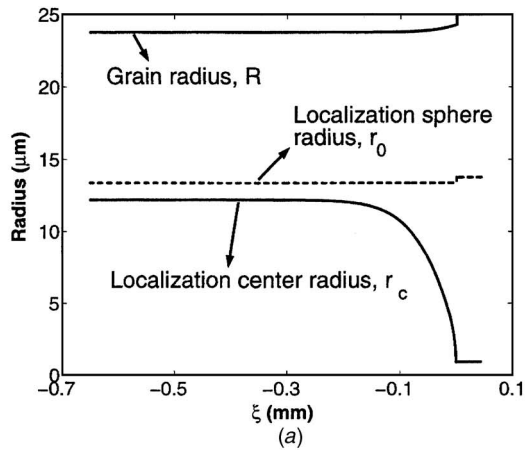
Fig. 10 Predicted variation in (a) grain radius, localization radii, and (b) grain temperature through the compaction zone for ϕ_0=0.81, u_p=1053 m/s, and D=3500 m/s

condition given by $\tilde{\phi}=f$ and $P_s=\beta$. These parameters are chosen because both u_p and ϕ_0 are easily controllable in experiments, and because Ω is difficult to experimentally characterize, particularly for dynamic compaction. Baseline values chosen for this study are u_p=150 m/s, ϕ_0=0.81, and Ω=1.0×10^4 [with μ_c =100 kg/(m s)]. This value of u_p is sufficiently large to guarantee that steady compaction waves are predicted for the entire range of ϕ_0 (0.655 ≤ ϕ_0 ≤ 0.93) considered.

We first vary piston impact speed over the range 100 ≤ u_p ≤ 1100 m/s to determine its influence on bulk compaction and grain scale heating; the predictions are summarized in Fig. 11. Subsonic wave structures are predicted for approximately u_p < 746 m/s, and supersonic structures are predicted for piston speeds in excess of this value. Compaction wave speed increases from a value of D=748.2 m/s for u_p=106 m/s to D =3500.0 m/s for u_p=1053.0 m/s; these piston speeds give rise to the structures shown in Figs. 4 and 8, respectively. Final solid volume fraction, ϕ, and no-load volume fraction, $\tilde{\phi}$, increase with impact speed as seen in Fig. 11(b). For impact speeds within the range 106 ≤ u_p ≤ 300 m/s, there is predicted a relatively steep rise in final solid volume fraction due to material crush-up, whereas a more gradual increase in final solid volume fraction is predicted for higher speeds due to material consolidation and stiffening. Stronger impact results in shorter compaction zone length [Fig. 11(c)], greater volumetric plastic strain [Fig. 11(d)], and more significant liquid formation due to phase change [Fig. 11(e)]. The predicted data for compaction zone length and liquid volume fraction in Figs. 11(c) and 11(e) are curve fit due to small, nonsmooth variations in the computed values. Compaction zone length was numerically defined by the ξ-location for which $\phi(\xi_i)-\phi(\xi_{i-1})$ ≤ 1.0×10^{-5} mm, where $\phi(\xi_i)$ is the value of solid volume fraction at nodal point ξ_i. Liquid volume fraction was estimated based on the location where $\hat{\chi} \approx 0.5$ within the phase change front. These results suggest the following. First, the compaction zone length is slowly approaching the order of a grain diameter (e.g., 50 μm) with increasing impact speed; because the compaction process is driven by interaction between grains, it is reasonable to expect that the compaction zone length should approach this value in the limit of strong waves. As compaction zone length data becomes available for high impact speeds, it is possible to match this data by varying the value of the relaxation parameter μ_c. Second, a significant increase in equivalent volumetric plastic strain occurs with increasing impact speed. While detailed mesoscale simulations of granular systems (~500 grains) may reasonably capture the mechanics of plastic deformation for strong impact, they probably lack the resolution needed to accurately capture the plastic zone for weak impact, particularly without the use of dynamically adaptive grids. The phenomenological model described in this paper, though simplistic, is capable of predicting the average manifestation of localized plastic deformation for a wide range of impact conditions in a manner consistent with contact mechanics. Third, the amount of liquid formed also increases significantly with impact speed suggesting that multiphase physics may be important for strong impact, as previously mentioned. Figure 11(f) gives the predicted ratio of the maximum grain scale temperature to bulk temperature, τ_T, through the compaction zone. The value of this ratio is in excess of 3.2 for the entire range of impact speeds considered here, and reaches a maximum value of τ_T=6.6 for u_p=450 m/s. Because the peak grain scale temperature is considerably larger than the peak bulk temperature, suitable localization strategies are needed to accurately predict hot-spot induced combustion of granular HMX (and other energetic solids) for a wide range of impact conditions. It is possible, however, that ignition models based on bulk temperature can be used for strong impact as the combustion rate becomes less sensitive to hot-spot fluctuations and can be suitably correlated with bulk temperature.

Figure 12 summarizes predictions for the impact of initially stress free granular HMX having initial solid volume fraction

Fig. 11 Predicted variation in (*a*) compaction wave speed, (*b*) solid volume fraction, (*c*) compaction zone length, (*d*) volumetric plastic strain, (*e*) liquid volume fraction, and (*f*) ratio of maximum grain temperature to bulk solid temperature with piston impact speed for $\phi_0=0.81$. Plots (*c*) and (*e*) are curve fits to the predicted data.

within the range $\phi_f=0.655 \leqslant \phi_0 \leqslant 0.93$. Here, it is understood that all material for which $\phi_0 > \phi_f$ has been pre-compacted and unloaded resulting in a strain hardened granular solid of lower porosity than the uncompacted, virgin material. We take u_p =150 m/s for these simulations. There is predicted a nonlinear increase in both compaction wave speed and solid pressure with ϕ_0 as seen in Figs. 12(*a*) and 12(*c*); the corresponding increase in solid volume fraction is shown in Fig. 12(*b*). The material crush-up pressure increases with strain hardening resulting in greater elastic compaction, wave speeds, and input power for fixed u_p. The input power per unit cross-sectional area, $\mathcal{P}_i$, is the

time rate of change of integrated total energy within the material and is given by $\mathcal{P}_i=P_s\phi u_p$. Though not plotted here, the input power varies from approximately $\mathcal{P}_i=1.38$ MW/cm^2 for ϕ_0 =0.81 to $\mathcal{P}_i=7.77$ MW/cm^2 for $\phi_0=0.93$. Interestingly, compaction zone length reaches a maximum value near 6.4 mm for ϕ_0 =0.66. A similar trend was predicted by Powers et al. [8], for the dynamic compaction of granular HMX. While the origin of this trend is unclear, it cannot be attributed to recoverable changes in volume fraction induced by strain hardening because, as opposed to our more physically realistic model, the model of Ref. [8] does not account for it; thus, this result is independent of key differ-
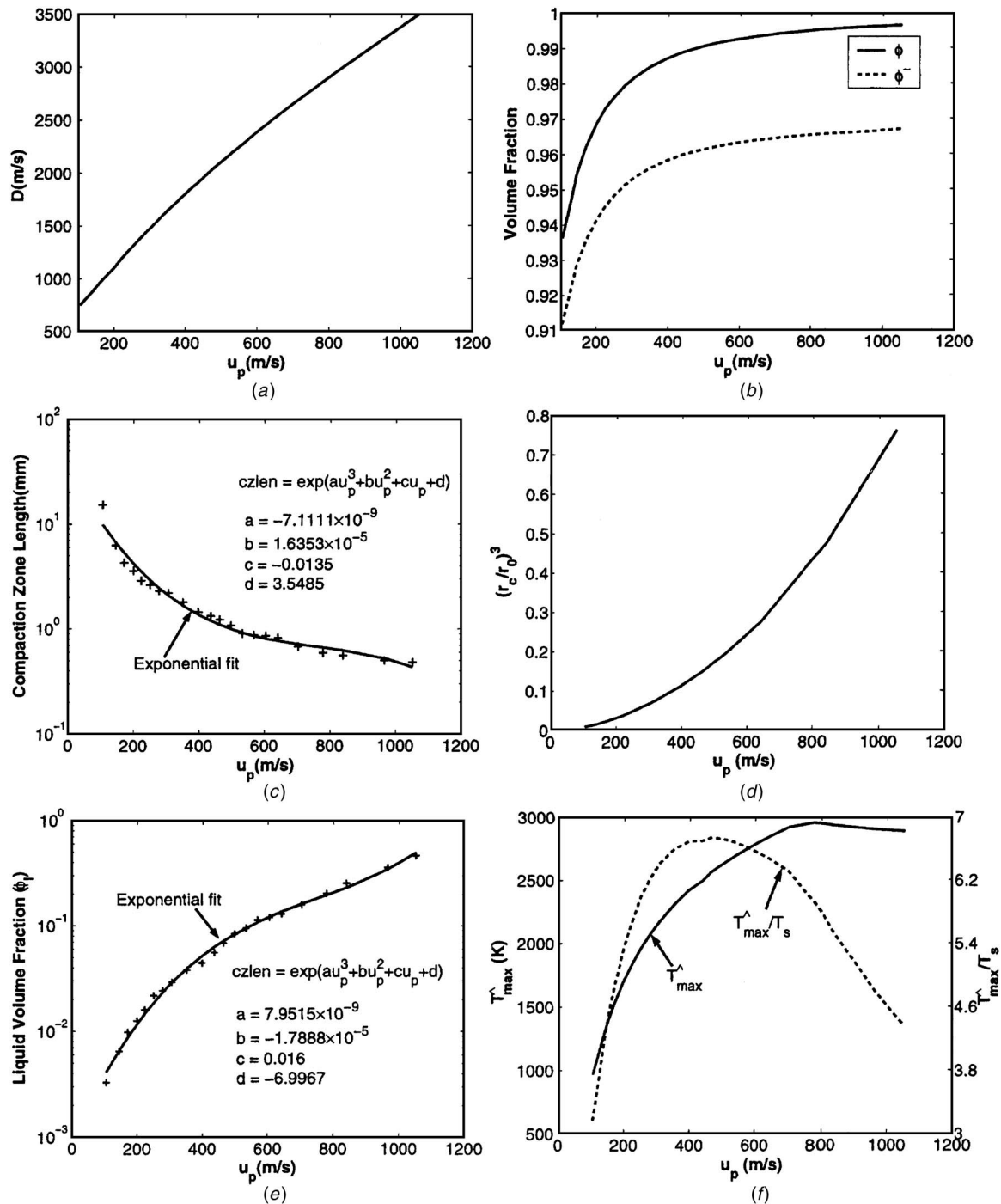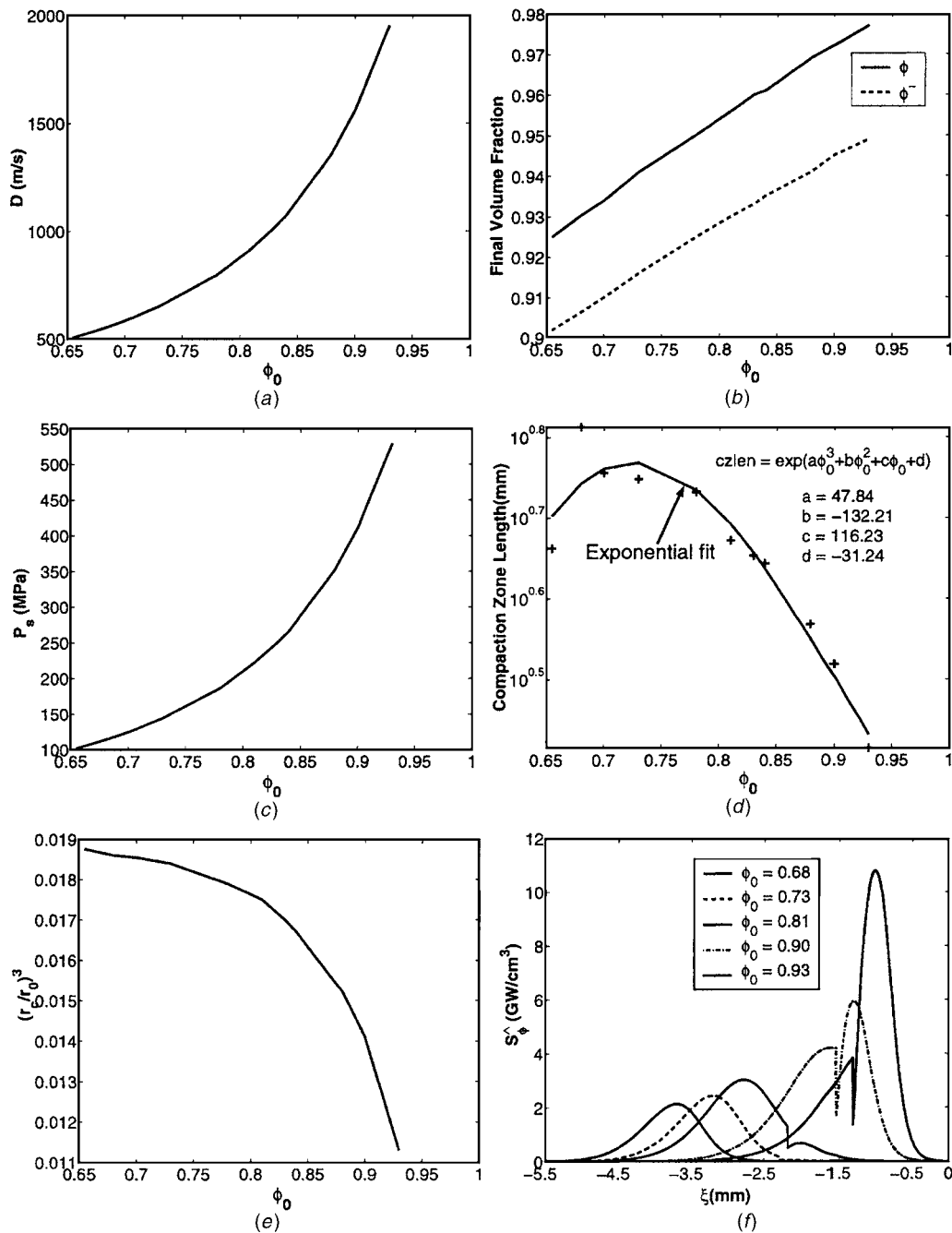
Fig. 12 Predicted variation in (a) compaction wave speed, (b) solid volume fraction, (c) solid pressure, (d) compaction zone length, (e) volumetric plastic strain, and (f) inelastic heating rate with initial solid volume fraction ϕ_0 for u_p=150 m/s

ences in our constitutive theories. The variation in volumetric plastic strain ϵ_{pl} with ϕ_0 is shown in Fig. 12(e). Here, ϵ_{pl} decreases with ϕ_0 though both the input power and compaction wave strength increase; a minimum value of ϵ_{pl}=0.0112 is predicted for ϕ_0=0.93. The number density of intergranular contact sites is large for densely packed material. As such, the applied load, assumed uniformly distributed between grains through their contact surfaces, results in much less localized plastic deformation than for looser material. Figure 12(f) gives the variation in $\hat{S}_\phi$ within the compaction zone for several values of ϕ_0. The magnitude of the inelastic heating rate increases with ϕ_0 to a maximum value of $\hat{S}_\phi$=10.4 GW/cm³ for ϕ_0=0.93. The discontinuity seen in these plots is due to the abrupt onset of viscoplastic heating

occurring when the pressure reaches that needed for material crush-up; a detailed discussion of dynamic compaction for strain hardened material is given in Ref. [7].

When the governing equations are suitably scaled, an important parameter that appears is the ratio of time scales associated with the relaxation processes given by Eqs. (21) and (22), $\Omega \equiv \mu_c/(\rho_{s0}D^2\tilde{\mu})$ [36]. The value of this parameter can significantly affect compaction dynamics as illustrated in Fig. 13. Here, we take ϕ_0=0.81, u_p=106 m/s, and μ_c=100 kg/(m s), and vary $\tilde{\mu}$ by several orders so that Ω=0.1, 1.0, and 1.0×10^4. For each of these cases, the final equilibrium state is the same (ϕ=0.936, P_s=130.45 MPa), as is the wave speed (D=748.2 m/s), because Ω only controls that rate at which the compaction end state is ap-
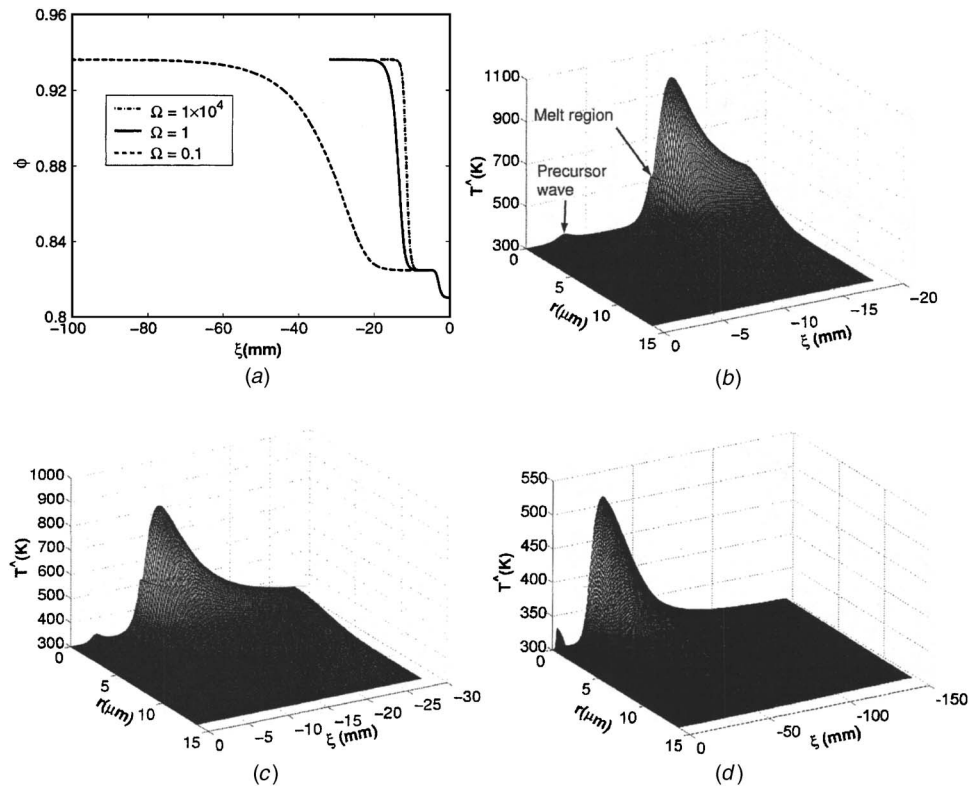
Fig. 13 Predicted variation in (a) solid volume fraction and (b)–(d) grain scale temperature through the compaction zone for Ω =1.0×10⁴, 1.0, and 0.1, respectively

proached; however, the evolution of grain scale temperature is compaction rate dependent and will thus vary with Ω. As seen in the figure, compaction zone length decreases with increasing Ω while peak grain temperature decreases. Importantly, reasonably small differences in compaction zone length ($<$4 mm) and peak grain temperature ($<$130 K) are predicted for $1.0 \leqslant \Omega \leqslant 1.0 \times 10^4$, whereas significant differences are predicted for approximately $\Omega < 0.1$. Quasistatic compaction data for HMX suggest that Ω is large [26]; thus, it is unnecessary to determine a precise value for Ω in light of these predictions.

4.4 Comparison with Mesoscale Simulations. Lastly, we compare in Fig. 14 our model predictions for the variation in plastic strain and solid pressure within the compaction zone with those reported by Menikoff and Kober [11] based on two-dimensional (2D) mesoscale simulations of inert granular HMX for u_p=200, 500, and 1000 m/s. We have used initial conditions similar to those of the simulations. The mesoscale data are running local averages of the plastic strain field within discrete grains. The predictions qualitatively agree. Both descriptions indicate that ϵ_{pl} is compaction rate dependent, though our description shows more sensitivity. Relative to the mesoscale data, we underpredict ϵ_{pl} for u_p=200 m/s; our prediction gives an equilibrium value of ϵ_{pl}=0.034 whereas the mesoscale data gives a value near 0.13. Predictions for ϵ_{pl} at u_p=500 m/s agree well, but we overpredict its value for u_p=1000 m/s. Our pressure predictions qualitatively agree with the mean axial stress predicted by the simulations. For u_p=1000 m/s, we predict that D=3378 m/s $>c_{s0}$=2767 m/s, where c_{s0} is the ambient solid sound speed; thus, as indicated in the pressure plot, there exists a solid shock at the head of the compaction wave. Discrepancies between the predictions are likely due, in part, to numerical resolution, phase change, and grain packing geometry. Because the mesoscale simulations were performed using a coarse computational grid, it is likely that thin plastic zones induced by low speed impact were
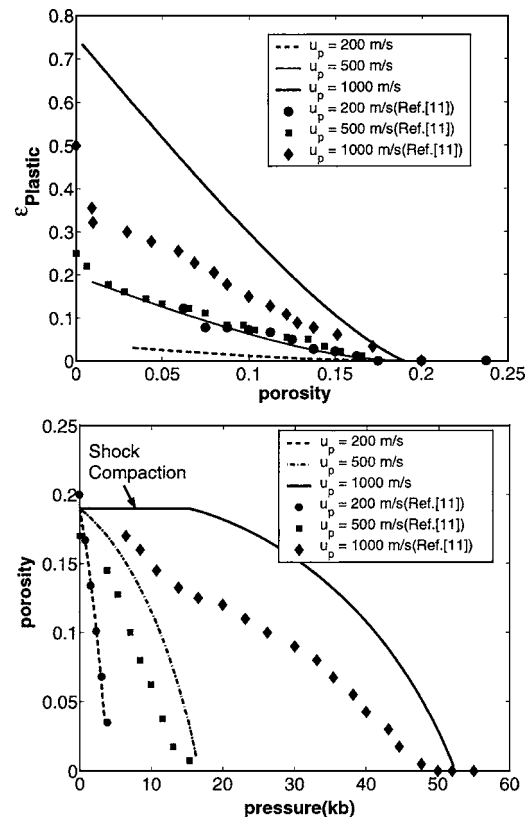
Fig. 14 Comparison of the predicted variation in plastic strain, pressure, and porosity through the compaction zone with the 2D mesoscale predictions reported in Ref. [11]

not adequately resolved. Our analysis predicts hot-spot volumes of 92 μm^3 near intergranular contact surfaces for $u_p=200$ m/s; such small volumes would require significant computational resolution. Further, phase change limits the amount of plastic strain within the solid. This observation may explain our large value of plastic strain for $u_p=1000$ m/s. Further, we assume $\gamma=12$ intergranular contacts/grain that is compatible with a three-dimensional (3D) packing arrangement, whereas the mesoscale simulations correspond to a 2D arrangement. Such mesoscale simulations are needed to provide information about hot-spot distributions that can be used to improve our localization strategy, particularly in the absence of experimental data at that scale.

5 Conclusion

An energetically consistent localization strategy for predicting hot-spot histories for HMX compaction has been presented. The strategy attributes bulk compaction induced dissipation to plastic work occurring in the vicinity of intergranular contact surfaces, and accounts for compressive heating, phase change, and thermal conduction at the grain scale. Steady subsonic compaction waves having a dispersed two-wave structure are predicted for mild impact of strain hardened HMX (porosity ~19%), and steady supersonic compaction waves having a discontinuous solid shock followed by a thin compaction zone are predicted for stronger impact. Predictions for mild impact ($u_p<110$ m/s) indicate short duration (relative to compaction zone length) hot-spot temperatures in excess of 900 K that are sufficient to induce sustained combustion based on a two-phase thermal explosion theory. Grain scale compressive heating is shown to be largely inconsequential compared to inelastic heating, and the latent heat of fusion is shown to reduce peak hot-spot temperatures by over 100 K. An analysis of the system response to variations in impact speed and initial solid volume fraction indicates an increase in hot-spot temperature, volumetric plastic strain, and liquid volume fraction with increasing impact speed at fixed initial solid volume fraction ($\phi_0=0.81$), and a decrease in these quantities with increasing initial solid volume fraction at fixed impact speed ($u_p=150$ m/s). Importantly, the ratio of peak grain scale temperature to bulk temperature within the compaction zone is predicted to increase with impact speed to a maximum value of $\tau_T\approx6.6$ for $u_p=400$ m/s, and subsequently decreases as bulk inelastic grain deformation becomes more pronounced. This large disparity in temperatures emphasizes the necessity of resolving grain scale temperature fluctuations that may lead to combustion initiation within the context of engineering scale models for energetic solids. Lastly, the model qualitatively reproduces features predicted by detailed mesoscale simulations including localized viscoelastic and viscoplastic heating; as such, it is possible to correlate our localization model with mesoscale predictions in an effort to develop improved bulk combustion models based on hot-spot formation.

Acknowledgment

References

[1] Baer, M. R., and Nunziato, J. W., 1986, "A Two-Phase Mixture Theory for the Deflagration to Detonation Transition (DDT) in Granular Reactive Materials," Int. J. Multiphase Flow, **12**, pp. 861–889.

[2] Powers, J. M., Stewart, D. S., and Krier, H., 1990, "Theory of Two-Phase Detonation-Part I: Modeling," Combust. Flame, **80**, pp. 264–279.

[3] Bdzil, J. B., Menikoff, R., Son, S. F., Kapila, A. K., and Stewart, D. S., 1999, "Two-Phase Modeling of Deflagration-to-Detonation Transition in Granular Materials: A Critical Examination of Modeling Issues," Phys. Fluids, **11**(2), pp. 378–402.

[4] Kang, J., Butler, P. B., and Baer, M. R., 1992, "A Thermomechanical Analysis of Hot Spot Formation in Condensed-Phase Energetic Materials," Combust. Flame, **89**, pp. 117–139.

[5] Bonnett, D. L., and Butler, P. B., 1996, "Hot-Spot Ignition of Condensed Phase Energetic Materials," J. Propul. Power, **12**(4), pp. 680–690.

[6] Massoni, J., Saurel, R., Baudin, G., and Demol, G., 1999, "A Mechanistic Model for Shock Initiation of Solid Explosives," Phys. Fluids, **11**, pp. 710–736.

[7] Gonthier, K. A., 2003, "Modeling and Analysis of Reactive Compaction for Granular Energetic Solids," Combust. Sci. Technol., **175**, pp. 1679–1709.

[8] Powers, J. M., Stewart, D. S., and Krier, H., 1989, "Analysis of Steady Compaction Waves in Porous Materials," J. Appl. Mech., **56**, pp. 15–24.

[9] Johnson, J. N., Tang, P. K., and Forest, C. A., 1985, "Shock-Wave Initiation of Heterogeneous Reactive Solids," J. Appl. Phys., **57**, pp. 4323–4334.

[10] Tarver, C. M., Chidester, S. K., and Nichols, A. L., III, 1996, "Critical Conditions for Impact and Shock-Induced Hot Spots in Solid Explosives," J. Phys. Chem., **100**, pp. 5794–5799.

[11] Menikoff, R., and Kober, E., 1999, "Compaction Waves in Granular HMX," LA-13546-MS, Los Alamos National Laboratory, Los Alamos, NM.

[12] Park, S. J., Han, H. N., Oh, K. H., and Lee, D. N., 1999, "Model for Compaction of Metal Powders," Int. J. Mech. Sci., **41**, pp. 121–141.

[13] Swegle, J. W., 1980, "Constitutive Equation for Porous Materials with Strength," J. Appl. Phys., **51**(5), pp. 2574–2580.

[14] Issen, K. A., 2002, "The Influence of Constitutive Models on Localization Conditions for Porous Rock," Eng. Fract. Mech., **69**, pp. 1891–1906.

[15] Issen, K. A., and Rudnicki, J. W., 2001, "Theory of Compaction Bands in Porous Rock," Phys. Chem. Earth, **26**(1–2), pp. 95–100.

[16] Cart, E. J., Lee, R. J., Gustavson, P. K., Coffey, C. S., and Sutherland, G. T., 2003, "The Role of Shear in Shock Initiation of Explosives," *Proceedings of the Conference of the American Physical Society Topical Group on Shock Compression of Condensed Matter*, Seattle, WA, in press.

[17] Tamura, S., and Horie, Y., 1998, "Discrete Meso-Dynamic Simulation of Thermal Explosion in Shear Bands," J. Appl. Phys., **84**(7), pp. 3574–3580.

[18] Wilson, W. H., Tasker, D. G., Dick, R. D., and Lee, R. J., 1998, "Initiation of Explosives Under High Deformation Loading Conditions," *Proceedings of the Eleventh (International) Detonation Symposium*, Snowmass, CO, pp. 565–572.

[19] Gonthier, K. A., 2005, "Modeling Shear Enhanced Compaction for Granular Explosive," Khim. Fiz., to appear.

[20] Cline, C. F., 1996, "Dynamic Compaction of Ceramic Powders," Ind. Ceram., **16**(3), pp. 189–194.

[21] Joshi, V. S., 1995, "Materials Processing by Shock Compaction: Status and Application to Nanocrystalline Ceramics," *Proceedings of the ASME Materials Division*, ASME, New York, MD-Vol. 69-1, pp. 633–651.

[22] Baer, M. R., 1988, "Numerical Studies of Dynamic Compaction of Inert and Energetic Granular Materials," J. Appl. Mech., **55**, pp. 36–43.

[23] Lubliner, J., 1990, *Plasticity Theory*, Macmillan Publishing Company, New York.

[24] Coyne, P. J., Elban, W. L., and Chiarito, M. A., 1985, "The Strain Rate Behavior of Coarse HMX Porous Bed Compaction," *8th International Symposium on Detonation*, Albuquerque, NM, July 15–19, pp. 645–657.

[25] Elban, W. L., and Chiarito, M. A., 1986, "Quasi-Static Compaction Study of Coarse HMX Explosive," Powder Technol., **46**, pp. 181–193.

[26] Gonthier, K. A., Menikoff, R., Son, S. F., and Asay, B. W., 1998, "Modeling Compaction Induced Energy Dissipation of Granular HMX," *11th International Symposium on Detonation*, Snowmass, CO, August 31–September 4, pp. 153–161.

[27] McAfee, J. M., Asay, B., Campbell, W., and Ramsay, J. B., 1989, "Deflagration to Detonation Transition in Granular HMX," *9th International Symposium on Detonation*, Portland, OR, August 28–September 1, pp. 265–279.

[28] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, New York.

[29] Menikoff, R., and Sewell, T. D., 2001, "Constituent Properties of HMX Needed for Mesoscale Simulations," LA-UR-00-3804-rev, Los Alamos National Laboratory, Los Alamos, NM.

[30] Sheffield, S. A., Gustavsen, R. L., and Anderson, M. U., 1997, "Shock Loading of Porous High Explosives," *High-Pressure Shock Compression of Solids IV*, Springer-Verlag New York, pp. 24–61.

[31] Sandusky, H. W., and Liddiard, T. P., 1985, "Dynamic Compaction of Porous Beds," Technical Report No. 83-246, Naval Surface Warfare Center.

[32] Marsh, S. P., 1980, *LASL Shock Hugoniot Data*, University of California Press, Berkeley, CA.

[33] Gonthier, K. A., 2004, "Predictions for Weak Mechanical Ignition of Strain Hardened Granular Explosive," J. Appl. Phys., **95**(7), pp. 3482–3494.

[34] Bardenhagen, S. G., and Brackbill, J. U., 1998, "Dynamic Stress Bridging in Granular Material," J. Appl. Phys., **83**(11), pp. 5732–5740.

[35] Gibbs, T. R., and Popalato, A., 1980, LASL Explosives Property Data.

[36] Jogi, V., 2003, Predictions for Multi-Scale Shock Heating of a Granular Energetic Material, M.S. thesis, Louisiana State University.

D. Cope
Associate Professor
Departmemt of Mathematics,
North Dakota State University,
Fargo, ND 58105
e-mail: davis.cope@ndsu.edu

S. Yazdani[1]
Professor and Chair
Department of Civil Engineering and
Construction,
North Dakota State University, Fargo, ND 58105
e-mail: frank.yazdani@ndsu.edu

J. W. Ju
Professor
Department of Civil and Environmental
Engineering,
University of California,
Los Angeles, CA 90095
e-mail: juj@ucla.edu

A Method to Generate Damage Functions for Quasi-Brittle Solids

In continuum damage mechanics theories, damage functions are identified based on experimental records. These functions also serve as strain hardening-softening functions similar to the conventional plasticity formulations. In a class of damage theories described in this paper it will be shown that if care is not taken, internal contradictions will arise as manifested by a snapback in the strain–stress space. This paper establishes a formal method by which different damage functions can consistently be developed leading to no snap-back in the solution. [DOI: 10.1115/1.1935524]

Introduction

There are classes of quasi-brittle solids where cracks tend to follow tortuous paths when the local stress or strain conditions for the crack propagation are satisfied. Examples are ceramics, where microcracks are intergranular, and concrete, where cracks emanate from weak interfacial bonds, propagate through the mortar phase, and go around aggregate particles that act as energy barriers. During the process of microcracking, material grains are severed leaving the strained solid as "damaged." Such a process alters elastic moduli and can lead to a strong material anisotropy.

In the absence of any inelastic flow, which is assumed here, the progressive damage can be modeled using damage mechanics theories (CDM) [1–4]. Different theories of CDM have been published to address diverse characteristics of material inelasticity associated with damage. One class of damage models that has received great interest among researchers was published by Ortiz [1] and Ju [2,3] where a fourth-order damage representation was employed. The basic approach by Ortiz, which was a stress-based formulation, was further refined to capture the response of brittle solids under proportional and nonproportional load paths [5,6], to model softening and localization phenomena [7,8], to model load induced damage in ceramics [9,10], and to model concrete inelasticity using a strain-based formulation [11].

Yazdani et al. [12] reported that a problem arose when a bilinear damage function was used instead of the logarithmic one that had been utilized in the original models. They reported that an apparent snapback was observed in the solution, for the uniaxial stress path, in the strain-softening regime regardless of the slope used. Although it is recognized that some physical damage processes do lead to localized deformations, the term *apparent* snapback is used in this paper to refer to the development of an internal contradiction in the damage model itself where snapback in

the deformation is predicted even though the theory was not structured to capture this aspect of material response. In this paper, we will examine the problem further, and we will provide a formal method with theorems and associated proofs by which different damage functions could consistently be developed leading to no snapback in the solution.

General Formulation

It is assumed that damage remains distributed within the representative volume element. This physically corresponds to the existence of a multitude of microcracks uniformly distributed within the material element. It is further assumed that neighboring or constraint equilibrium states exist for all irreversible processes. For small, rate-independent and isothermal deformations, and in the absence of any body couples, a thermodynamic potential can be used to construct the general formulation. Adopting a stress space formulation in which only mechanical deformations are considered, the Clausius–Duhem inequality yields:

$$\dot{G} - \dot{\boldsymbol{\sigma}}:\boldsymbol{\varepsilon} \geqslant 0 \qquad (1)$$

in which the thermodynamic state function is represented by $G(\boldsymbol{\sigma},k)$ as the Gibbs free energy (GFE) and the strain tensor is denoted by $\boldsymbol{\varepsilon}$. The Cauchy stress tensor is identified as $\boldsymbol{\sigma}$ and k is used as an internal variable accounting for the load induced damage accumulation in the material. The symbol ":" indicates the tensor contraction operation. It is assumed that damage is irreversible and that no healing takes place in the material, i.e., $\dot{k} \geqslant 0$. For elastic-damaging processes, the total strain is obtained by the following constitutive relation as:

$$\varepsilon(\boldsymbol{\sigma},k) = \mathbf{C}(k):\boldsymbol{\sigma} \qquad (2)$$

in which $\mathbf{C}$ represents the fourth-order compliance tensor for the material. The dependence of $\mathbf{C}$ on k reflects the idea that damage alters elastic properties and allows for the description of load-induced anisotropy [13,14]. As was the case in the original damage model by Ortiz [1], it is assumed that no permanent deformation will arise due to the misfit of crack faces. The inclusion of the inelastic damage strain tensor will not alter the results of this paper although appropriate relations including a corresponding evolutionary relation for the inelastic strain tensor must be added. The Clausius–Duhem inequality establishes two results. The first one is the dissipation inequality,d_s, given as

$$d_s = \frac{\partial G(\boldsymbol{\sigma}, k)}{\partial k} \dot{k} \geqslant 0 \tag{3}$$

and, the second one is the statement that GFE is a potential for the strain tensor as:

$$\varepsilon(\boldsymbol{\sigma}, k) = \frac{\partial G(\boldsymbol{\sigma}, k)}{\partial \boldsymbol{\sigma}} = \mathbf{C}(k){:}\boldsymbol{\sigma} \tag{4}$$

To progress further, assume that the rate of the added flexibility tensor is expressed as a linear form in k as $\dot{\mathbf{C}}^c(k) = \dot{k}\,\mathbf{R}$ in which $\mathbf{R}$ is a fourth-order damage response tensor. Integrating Eq. (4) with respect to $\boldsymbol{\sigma}$ and substituting back in Eq. (3) yields:

$$d_s = \left(\frac{1}{2}\boldsymbol{\sigma}{:}\mathbf{R}{:}\boldsymbol{\sigma} - \frac{\partial A^i(k)}{\partial k}\right)\dot{k} \geqslant 0 \tag{4a}$$

in which $A^i(k)$ arises as a constant of integration representing the inelastic component of the Helmholtz free energy associated with surface free energy of cracks. Since it is assumed that damage is irreversible, it then follows that the term inside the parentheses in Eq.(4a) must itself be non-negative. One could then formulate a potential, $\Psi(\boldsymbol{\sigma}, k)$ such that:

$$\Psi(\boldsymbol{\sigma}, k) = \frac{1}{2}\boldsymbol{\sigma}{:}\mathbf{R}{:}\boldsymbol{\sigma} - \frac{\partial A^i}{\partial k} - g^2(\boldsymbol{\sigma}, k) = 0 \tag{4b}$$

for some function "g." The individual identification of functions A^i or g is not needed as long as the function $t^2(\boldsymbol{\sigma}, k) = 2\{(\partial A^i/\partial k) + g^2(\boldsymbol{\sigma}, k)\}$, known as the *damage function*, could be determined. With the damage function given as t, a general form of the damage surface is established as:

$$\Psi(\boldsymbol{\sigma}, k) = \tfrac{1}{2}\boldsymbol{\sigma}{:}\mathbf{R}{:}\boldsymbol{\sigma} - \tfrac{1}{2}t^2(\boldsymbol{\sigma}, k) = 0. \tag{5}$$

Anisotropy

To bring in the anisotropy for tensile stress path, Ortiz [1] proposed the following damage response tensor for damage mode I:

$$\mathbf{R} = \frac{\boldsymbol{\sigma}^+ \otimes \boldsymbol{\sigma}^+}{\boldsymbol{\sigma}^+{:}\boldsymbol{\sigma}^+} \tag{6}$$

in which $\boldsymbol{\sigma}^+$ represents the positive cone of the stress tensor. The operational requirements to obtain $\boldsymbol{\sigma}^+$ are formally stated in references cited above and will not be repeated here. The substitution of Eq. (6) into Eq. (5) yields the following result in the uniaxial tensile stress path that:

$$(\sigma_1)^2 = t^2(\boldsymbol{\sigma}, k) \tag{7}$$

where σ_1 represents the component of the stress in the 1-1 direction. Equation (7) is usually used to obtain a form for the damage function from the uniaxial stress–strain path as was done by Ortiz. Following the experimental work of Smith and Young [15], the following logarithmic form for the damage function was used:

$$t(\sigma, k) = \frac{f_t e \ln(1 + E_0 k)}{(1 + E_0 k)} \tag{8}$$

where f_t denotes the uniaxial tensile strength of the material, E_0 is the initial value of Young's modulus, and e is 2.71828. The graph of this function is shown in Fig. 1. Using the rate form of Eq. (2) together with Eqs. (5) and (6), and integrating over the path, the closed form of the stress-strain relation for the uniaxial tensile stress path can be shown to be

$$\varepsilon = \left(\frac{1}{E_0} + k\right)\sigma_1(k) = \left(\frac{1}{E_0} + k\right)t(k) \tag{9}$$

From the structure of Eq. (9) one can see that the initial slope of the curve is given by E_0. Furthermore, the right-hand side of Eq. (9) is the product of two functions. One function, $((1/E_0) + k)$, monotonically increases with k; the other function, $t(k)$, first in-
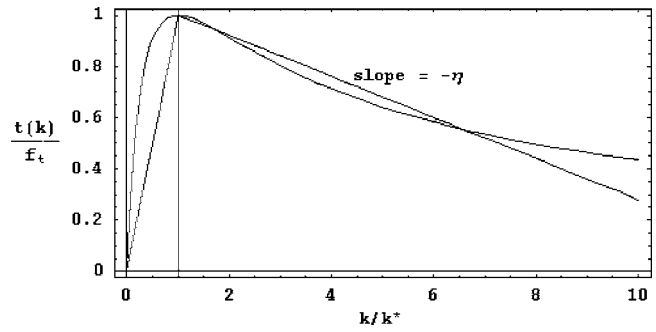
Fig. 1 Normalized damage function $t(k)/f_t$ vs normalized k/k^* for logarithmic and bilinear damage functions

creases (hardening regime) and then decreases (softening regime) with k. Therefore, if care is not taken in the appropriate specification of $t(k)$, the strain function could be an increasing and then decreasing function of k representing therefore a behavior that is termed here as the *apparent* snapback.

For example, in many cases engineers approximate nonlinear functions, such as the one represented by Eq. (8), and replace them with bilinear functions as shown in Fig. 1. It was shown by Yazdani et al. [12] that in the strain-softening regime, bilinear damage functions exhibit apparent snapback, that is, multivalued behavior, no matter how steep or mild the slope of the damage function in the post-peak regime is. This behavior is shown in Fig. 2 for several bilinear damage functions along with the response obtained by using Eq. (8). In the latter, as strain increases with increasing damage, stress increases to a maximum, f_t, and then decreases. It can also be shown that damage functions with three straight-line segments or in quadratic or sinusoidal forms will also lead to the snapback. It is therefore clear that a criterion is needed to aid in the selection of appropriate damage functions that will produce well behaved, stress–strain response with no snapback.

In the following section, we present a formal method with theorems and proofs in developing a family of damage functions that produces well behaved stress–strain curves with no snapbacks. The work here is considered an extension to the work reported by Yazdani et al. [12] in that (a) it presents a formal method with theorems and proofs to the general approach, and (b) it removes the shortcoming of the previous work by including the new capability of matching the theoretical curve with the observed initial slope and observed points of maximum uniaxial strength for all families of curves introduced in the analysis.
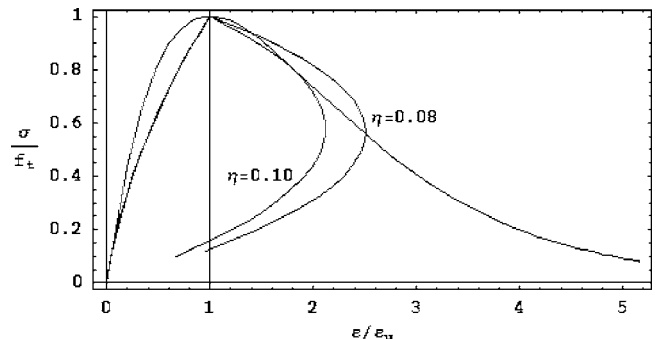
Fig. 2 Stress–strain curves for logarithmic and bilinear damage functions for $\eta = 0.08$, 0.10 ($E_0 \varepsilon_u / f_t = e$)

Model Functions and Properties

In developing a solution strategy for selecting appropriate damage functions, it proves convenient to introduce a set of functions called "model functions" with the following properties. A model function, $m(x)$, is a function such that:

(a) $m(x)$ is $C[0, +\infty)$ and is strictly increasing with $m(0)=0$ and $m(+\infty)=+\infty$,

(b) $m'(x)$ is $C(0, +\infty)$ with $xm'(x)/m(x)$ strictly decreasing and such that

$$1 \leq \lim\left(\frac{xm'(x)}{m(x)}\right) \leq +\infty \text{ as } x \to 0^+ \text{ and}$$

$$\lim\left(\frac{xm'(x)}{m(x)}\right) = 0 \text{ as } x \to +\infty \quad (10)$$

Note that the strictly decreasing condition on the quotient implies that $m'(x) > 0$.

LEMMA 1. *For each $a>0$, the function $m(x)/(a)+x$ is initially strictly increasing, then strictly decreasing on $[0, +\infty)$. The maximum occurs at $x = x_a$ given by*

$$m'(x_a) = \frac{m(x_a)}{a + x_a}. \quad (11)$$

The proof of the lemma is given in the Appendix. Also, a second lemma is given in the appendix that will be useful in generating families of damage functions.

Having defined $m(x)$ as a model function, one can then show that there exist positive constants c_1 and c_2 such that a strictly positive function $\varepsilon(k)$ can be developed in the form of $\varepsilon(k) = c_2 m(c_1 k)$. The constants c_1 and c_2 are determined from a set of experimentally identifiable points, namely the initial modulus of elasticity, E_0, the tensile strength, f_t, and the associated strain ε_u. This is shown below.

THEOREM 1. *Let $m(x)$ be a model function. Then, for each triple of positive values E_0, ε_u, f_t with $f_t < E_0 \varepsilon_u$, constants c_1 and c_2 can be found such that with:*

$$\varepsilon(k) = c_2 m(c_1 k) \quad (12)$$

the constitutive relation

$$\sigma(k) = \frac{\varepsilon(k)}{\frac{1}{E_0} + k} = \frac{c_1 c_2 m(c_1 k)}{\frac{c_1}{E_0} + c_1 k} \quad (13)$$

will display no snapback in the $\sigma - \varepsilon$ space and will have a unique maximum at the point $(\varepsilon, \sigma) = (\varepsilon_u, f_t)$.

Proof. We assume that there is a unique positive constant k^* associated with the uniaxial tensile strength, f_t, and is determined experimentally by $k^* = (\varepsilon_u/f_t) - 1/E_0$. To show that $\varepsilon(k)$ is a strictly increasing function, we recall the property (a) of the model functions defined previously. Since $m(x)$ is a model function and constants c_1 and c_2 are positive, then $\varepsilon(k)$ is a strictly increasing function of k and therefore snapback cannot occur. Now we need to show that Eq. (13) is initially strictly increasing, then strictly decreasing with maximum occurring at k^* corresponding to $\sigma = f_t$ and $\varepsilon = \varepsilon_u$. We first note that from Lemma 1, the stress function by Eq. (13) is initially strictly increasing, then strictly decreasing. To show the last part corresponding to the maximum, let r^* be a dimensionless parameter given by $r^* = E_0 k^*$. By solving the equation

$$\frac{x^* m'(x^*)}{m(x^*)} = \frac{r^*}{1 + r^*} \quad (14)$$

we obtain a positive root x^* by the definition of model function and that x^* satisfies the equation
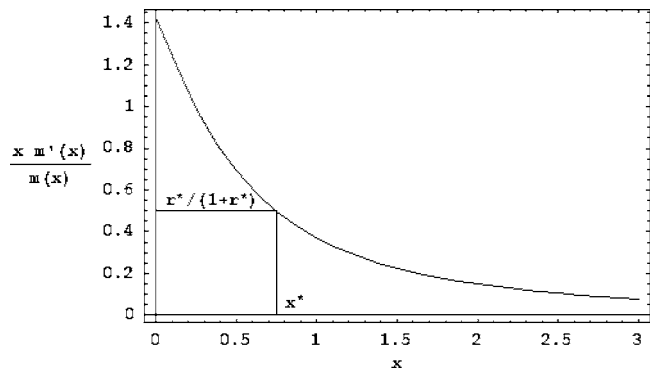
Fig. 3 Schematic representation of Eq. (14)

$$\frac{m'(x^*)}{m(x^*)} = \frac{1}{a + x^*} \quad (15)$$

with $a = x^*/r^*$ and $x^* = x_a$ signifying the location of unique maximum for $m(x)/(a) + x$. The schematic representation of the solution to Eq. (14) is provided in Fig. 3. Then, it follows that with specifying $c_1 = E_0 a$ and $c_2 = \varepsilon_u/m(x^*)$, the maximum of the function occurs at $k = k^*$ with the maximum of the function being equal to f_t.   QED

Generating Model Functions

To construct model functions, consider a function $g(s)$ that (a) is $C(-\infty, +\infty)$, (b) is strictly decreasing, (c) satisfies the limits $1 \leq \lim g(s) \leq +\infty$ for $s \to -\infty$ and $\lim g(s) = 0$ for $s \to +\infty$, and (d) satisfies the integral condition

$$\int_0^\infty g(s)ds = +\infty. \quad (16)$$

The model function $m(x)$ is then formulated and obtained as

$$m(x) = \exp\left(\int_0^{\ln(x)} g(u)du\right) \quad (17)$$

THEOREM 2. *The function obtained from Eq. (17) is a model function.*

Proof. With the change of variable $s = \ln(u)$, we get

$$m(x) = \exp\left(\int_{s=0}^{s=\ln(x)} g(s)ds\right) = \exp\left(\int_1^x \frac{g(\ln(u))}{u}du\right) \quad (18)$$

It is immediate from condition (a) above on $g(s)$ that $m(x)$ defined by Eq. (16) is $C(0, \infty)$ and that $m'(x)$ is also $C(0, \infty)$. By conditions (b) and (c) listed above, $\lim m(x) = 0$ as $x \to 0^+$ and $m(x)$ will be strictly increasing. By condition (d), $m(+\infty) = +\infty$ and part (a) of the definition of model function is therefore satisfied. By differentiating Eq. (16) one obtains

$$\frac{xm'(x)}{m(x)} = g(\ln(x)) \quad (19)$$

so that conditions (b) and (c) on $g(s)$ translate directly into part (b) of the definition for a model function.   QED

Example 1. Consider the function $m(x) = \ln(1+x)$. This function satisfies the definition of a model function as can be checked by direct computation. In particular, $xm'(x)/m(x)$ is strictly decreasing on $[0, +\infty)$ and has the limits of one and zero as x approaches zero and $+\infty$, respectively. As noted in the proof of Theorem 1, the following equation:
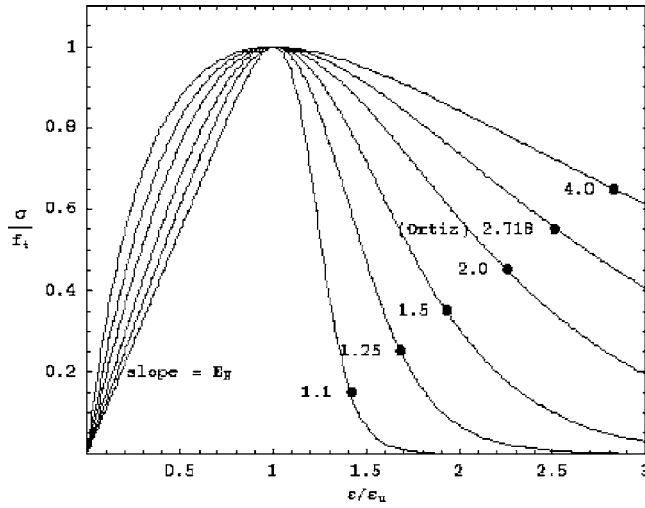
Fig. 4 Normalized stress–strain curves in uniaxial tension from Example 1. Initial slopes $E_N = E_0 \varepsilon_u / f_t = 1.1, 1.25, 1.5, 2.0, 2.718$ (Ortiz), and 4.0.

Fig. 5 Normalized stress–strain curves in uniaxial tension from Example 2. Model function $\ln(1 + x^p)$ with $p = 1$ (Ortiz), 1.5, 2, 4, 8.

$$\frac{x m'(x)}{m(x)} = \frac{x}{(1+x)\ln(1+x)} = \frac{r^*}{1+r^*} \qquad (20)$$

has a unique positive root at $x = x^*$. By letting $r^* = E_0 k^*$, $c_1 = E_0 = (x^*/r^*)$, and $c_2 = \varepsilon_u / \ln(1 + x^*)$, we obtain

$$\varepsilon(k) = \frac{\varepsilon_u}{\ln(1+x^*)} \ln(1 + x^* k / k^*) \qquad (21)$$

and the corresponding expression for stress as

$$\sigma(k) = \frac{\varepsilon_u}{\ln(1+x^*)} \frac{1}{(k + 1/E_0)} \ln(1 + x^* k / k^*) \qquad (22)$$

The stress–strain response is plotted in Fig. 4 for different values of stiffness with curves showing no snapback behavior. For one particular value of stiffness, Ortiz's model is obtained.

Example 2. Lemma 2 in the Appendix provides some simple ways of generating new model functions from any known model function. For example, starting with the model function $\ln(1 + x)$ of Example 1, it is immediate from Lemma 2 that

(a) $\ln(1 + x^p)$, $p \geq 1$
(b) $(\ln(1+x))^q$, $q \geq 1$
(c) $(\ln(1 + x^p))^q$ $p, q \geq 1$
(d) $\ln(1 + \ln(1+x))$

are all model functions, and each individual instance can be used as $m(x)$ for creating a damage function. For example, for case (a) consider the function $m(x) = \ln(1 + x^p)$ with $p \geq 1$. For a set of experimentally obtained values f_t, ε_u, and E_0, and following the steps stated in the proof of Theorem 1, we calculate dimensionless quantity $r^* = (E_0 \varepsilon_u / f_t - 1)$ and constants c_1 and c_2. The plot of this family of stress–strain curves is shown in Fig. 5 for different values of p and normalized initial slope $E_0 \varepsilon_u / f_t = e$. All curves are well behaved showing no snapback characteristics. As p increases, the curves start to merge, so that the curves are essentially identical for large p. This is characteristic of this particular model function and may not be seen if other model functions are used. Similarly, Fig. 6 shows stress–strain curves based on the model function $m(x) = (\ln(1+x))^q$ for different values of q. No snapback in the solution occurs. Other model functions could be developed from Eq. (17) resulting in well behaved stress–strain responses following the steps outlined in this paper.

Conclusion

It was shown that if care was not taken, internal contradiction within a class of continuum damage mechanics models could arise where a snapback in the stress–strain curve is observed. To avoid this problem, a formal method of specifying a family of damage functions was presented with associated theorem and proofs. It was shown that by establishing a set of functions identified as "model functions" possessing certain characteristics the formulation could lead to the construction of a well behaved response in the stress–strain space. A set of experimentally identifiable parameters is used to correlate model prediction to experimental data. The approach was illustrated with two examples for a general class of elastic-perfectly damaging behavior.

Appendix

Proof of Lemma 1. The derivative can be written as a product of two terms:
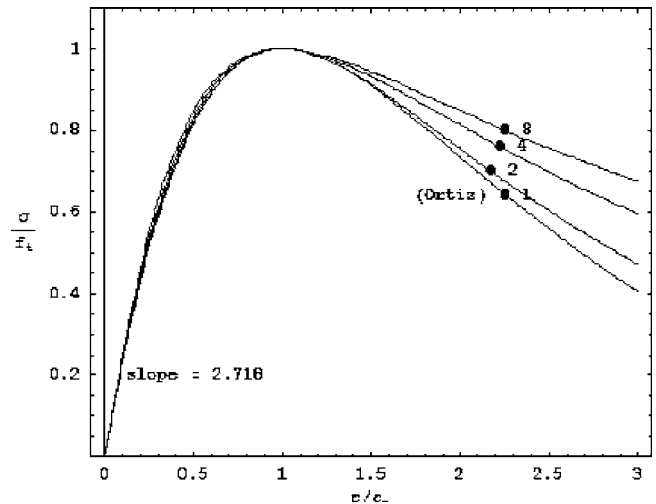
Fig. 6 Normalized stress–strain curves in uniaxial tension from Example 2. Model function $(\ln(1+x))^q$ with $q = 1$ (Ortiz), 2, 4, 8.

$$\frac{d}{dx}\left(\frac{m(x)}{a+x}\right) = \frac{m(x)}{(a+x)^2}\left(\frac{am'(x)}{m(x)} + \frac{xm'(x)}{m(x)} - 1\right) \quad \text{(A1)}$$

On $(0, +\infty)$, the first term in the parentheseis is strictly positive. The second term is a sum of two strictly decreasing functions and is therefore strictly decreasing. The first function is $\{am'(x)/m(x)\}$, which decreases from $+\infty$ to zero. The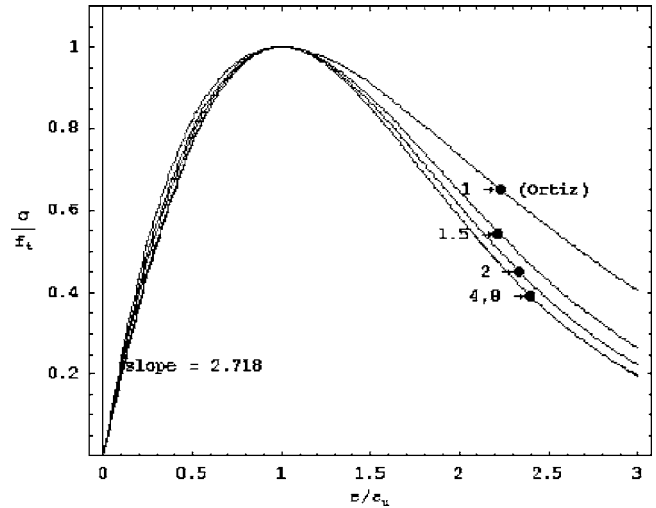 second function is $\{xm'(x)/m(x) - 1\}$, which decreases to the negative value -1. The sum therefore strictly decreases from $+\infty$ to a negative value, changing sign exactly once and implying that the original function is first strictly increasing, then strictly decreasing. QED

LEMMA 2. *Let functions $m(x)$ and $m_1(x)$ be model functions. Then:*

(1) $n(x) = bm(ax)$ *is a model function for all $a, b > 0$.*
(2) $n(x) = m(x^\alpha)$ *is a model function for each $\alpha \geq 1$*
(3) $n(x) = (m(x))^\alpha$ *is a model function for each $\alpha \geq 1$.*
(4) $n(x) = m_1(m(x))$ *is a model function.*

Proof of Lemma 2. The main step in each case is to check that $xn'(x)/n(x)$ is strictly decreasing. Here is the main step shown for (4). The quotient can be written as a product of two terms:

$$\frac{xn'(x)}{n(x)} = \frac{xm'(x)m_1'(m(x))}{m_1(m(x))} = \frac{xm'(x)}{m(x)}\frac{m(x)m_1'(m(x))}{m_1(m(x))} \quad \text{(A2)}$$

By the definition of model functions, both terms are positive functions and because, $m(x)$ is strictly increasing, both terms are strictly decreasing functions. The product is therefore strictly de-creasing. The product satisfies the two limit conditions because each term does so. QED

References

[1] Ortiz, M., 1985, "A Constitutive Theory for the Inelastic Behavior of Concrete," Mech. Mater., **4**(1), pp. 67–93.
[2] Ju, J. W., 1989a, "On Energy-Based Coupled Elastoplastic Damage Theories: Constitutive Modeling and Computational Aspects," Int. J. Solids Struct., **25**(7), pp. 803–833.
[3] Ju, J. W., 1989b, "On Energy-Based Coupled Elastoplastic Damage Models at Finite Strains," J. Eng. Mech., **115**(11), pp. 2507–2525.
[4] Ju, J. W., 1990, "Isotropic and Anisotropic Damage Variables in Continuum Damage Mechanics," J. Eng. Mech., **116**(12), pp. 2764–2770.
[5] Karnawat, S., and Yazdani, S., 2001, "Effects of Preloading on Brittle Solids," J. Eng. Mech., **127**(1), pp. 11–17.
[6] Yazdani, S., 1993, "On a Class of Continuum Damage Mechanics Theories," Int. J. Damage Mech., **2**, pp. 162–176.
[7] Chen, Z., and Schreyer, H. L., 1991, "Secant Structural Solutions under Element Constraint for Incremental Damage," Comput. Methods Appl. Mech. Eng., **90**, pp. 869–884.
[8] Schreyer, H. L., and Neilsenm, M., 1996, "Analytical and Numerical Tests for loss of Material Stability," Int. J. Numer. Methods Eng., **39**, pp. 1721–1736.
[9] Ortiz, M., and Giannakopoulos, A. E., 1990a, "Crack Propagation in Monolithic Ceramics under mixed mode Loading," Int. J. Fract., **44**, pp. 233–258.
[10] Ortiz, M., and Giannakopoulos, A. E., 1990b, "Mixed Mode Crack-tip Fields in Monolithic Ceramics," Int. J. Solids Struct., **26**(7), pp. 705–723.
[11] Stevens, D. J., and Liu, D., 1992, "Strain-Based Constitutive Model with Mixed Evolution Rules for Concrete," J. Eng. Mech., **118**(6), pp. 1184–1200.
[12] Yazdani, S., Cope, D., and Very, K., 2002, "Requirements on Damage Functions," J. Eng. Mech., **128**(1), pp. 126–129.
[13] Budiansky, B., and O'Connell, R. J., 1976, "Elastic Moduli of a Cracked Solid," Int. J. Solids Struct., **12**(2), pp. 81–97.
[14] Horii, H., and Nemat-Nasser, S., 1983, "Overall Moduli of Solids with Microcracks: Load-Induced Anisotropy," J. Mech. Phys. Solids, **31**(2), pp. 155–171.
[15] Smith, and Young, 1955, "Ultimate theory in flexure by exponential functions," Proceedings of ACI, Vol. 52(3), pp. 349–359.

Transmission of Elastic Stress Through Circular and Elliptic Cross Sections of Microstructural Elements Embedded in a Matrix Material

C. M. Kennefick

Reston, VA 20191

With the use of contact stress theory and complex variable methods in two dimensions, the transmission of a compressive stress through a circular cross section of a small material particle is calculated in the infinite plane of material below the circular cross section. The circular cross section of the particle is embedded in and completely bonded to an infinite plane of matrix material. It is shown that part of the stress is transmitted with a dependence of $1/r$, where r is a radial coordinate. Additionally, the stress is calculated in two dimensions for the interior of an ellipse that could model a cross section of a grain or particle. The boundary of the ellipse is loaded with the stress holding an elliptic kernel in place in an elastic matrix material after the kernel has undergone a small rotation under an applied tensile load. The resulting stresses are shown in contour plots for elliptic cross sections of varying shapes and orientations. [DOI: 10.1115/1.1935525]

1 Introduction

The stress distribution around particles and grains is important in the design of materials because such distributions can indicate areas of increased toughening or areas for fracture. For increased resistance to fracture, for example, small particles are often dispersed into the grain boundaries and into the interior of grains [1–4]. It is therefore the goal of this work to record what happens when a particle interacts with an applied stress coming from another part of the matrix material.

Other analytical solutions for stress distributions have involved entities that are circular [5–10], spheroidal [11], elliptic [12–15], and ellipsoidal [16–20], or are small inclusions [21–25]. Complementing these works, the study here considers a small circular cross section of a particle that is embedded in and completely bonded to an infinite plane of matrix material. The circular cross section is loaded in compression on the upper half of its circumference, as shown in Fig. 1. The stress transmitted to the lower half of the circumference is then used as a boundary condition to calculate the stress transmitted into the infinite plane of material below the particle.

This paper also calculates the stress distribution in the interior of an ellipse representing a grain or inclusion that resides in an elastic material matrix. The major axis of the ellipse is at an angle to an applied tensile load, allowing the ellipse to undergo a slight rotation denoted by the angle ε [26], as shown in Fig. 2.

2 Transmission of a Compressive Stress Through a Circular Cross Section

A compressive stress $-P$, which is the applied load in the y direction per unit length in the x direction, is being transmitted through the matrix material toward the upper half of a circular cross section, as shown in Fig. 1. Stresses resulting from the matrix material, which is being held in equilibrium with the particle, are assumed to be small in comparison with the compressive stress applied to the upper half of the surface of the particle.

The angles η_1 and η_2, also shown in Fig. 1, are measured from the line of the load to the line connecting point A to the edge of the applied load. η_1 and η_2 are positive when measured counterclockwise and are negative when measured clockwise. Stress states σ and τ will denote a loads per unit length in which the first subscript denotes the direction of the vector perpendicular to the plane upon which the load acts. The second subscript denotes the direction of the load itself.

For the Cartesian coordinates x and y shown in Fig. 1, the final stress state of a point below the surface can be found by integrating over the upper half of the circle, which can represent a cross section of unit thickness of a small particle [27]. The final result is

$$\sigma_{xx} = (-P/2\pi)[2(\eta_2 - \eta_1) - (\sin 2\eta_2 - \sin 2\eta_1)]$$

$$\sigma_{yy} = (-P/2\pi)[2(\eta_2 - \eta_1) + (\sin 2\eta_2 - \sin 2\eta_1)] \qquad (1)$$

$$\tau_{xy} = (P/2\pi)(\cos 2\eta_1 - \cos 2\eta_2)$$

θ is an angular coordinate beginning at zero at the x axis and is measured counterclockwise. The angle γ was then defined as a polar angle beginning at the left-hand side of the circle in Fig. 1 and is equal to θ-π. From Fig. 1, it was then calculated that $\eta_2 - \eta_1 = \pi/2$, $\sin 2\eta_2 - \sin 2\eta_1 = 2 \sin \gamma$ and $\cos 2\eta_1 - \cos 2\eta_2 = -\cos \gamma$. With these substitutions, along with γ being equal to θ-π, the stresses in Eq. (1) may be written

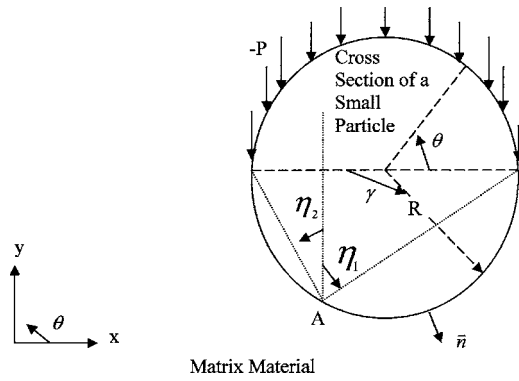$$\sigma_{xx} = -\frac{P}{2} - \frac{P}{\pi} \sin \theta$$

Fig. 1 Diagram showing the Cartesian, cylindrical, and angular coordinates for the circular cross section

$$\sigma_{yy} = -\frac{P}{2} + \frac{P}{\pi}\sin\theta \qquad (2)$$

$$\tau_{xy} = \frac{P}{2\pi}\cos\theta$$

The stresses in the infinite plane below the circular cross section may be obtained from two stress functions Φ and Ψ of the complex variable z [26]. The complex variable z is $x+iy$, where x and y have the orientation shown in Fig. 1. The variable z may also be written in polar coordinates as $re^{i\theta}$, where r is a radial coordinate. The radial coordinate r has a value of zero at the center of the circular cross section and a value of R at the perimeter of the circle.

The stresses σ_{xx}, σ_{yy}, and τ_{xy} are related to the two functions $\Phi(z)$ and $\Psi(z)$ by the equations [26]
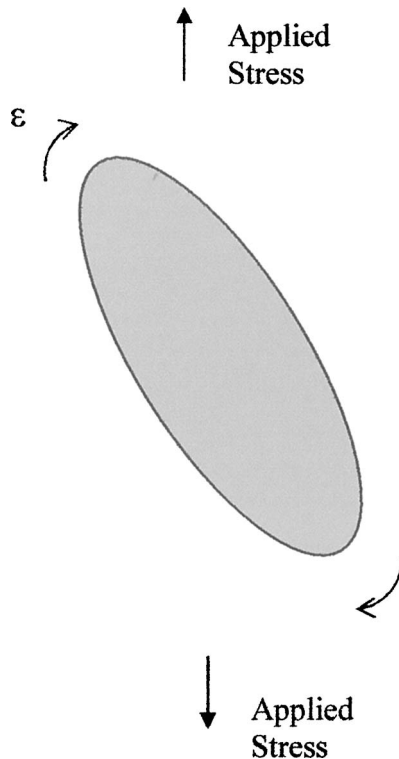
Fig. 2 Diagram showing the slight rotation through an angle ε of an elliptically shaped cross section embedded in an elastic material matrix and under a remote applied tensile stress

$$\sigma_{xx} + \sigma_{yy} = 4\,\mathrm{Re}[\Phi(z)] \qquad (3)$$

and

$$\sigma_{yy} - \sigma_{xx} + 2i\tau_{xy} = 2[\bar{z}\Phi'(z) + \Psi(z)] \qquad (4)$$

In Eq. (3), the symbol Re denotes the real part of the function $\Phi(z)$. In Eqs. (3) and (4) and in what follows in this paper, the prime next to a function denotes the derivative of that function with respect to the variable in parentheses. A bar over a variable or function denotes the complex conjugate of that variable or function.

A combination of Eqs. (2) and (3) gives $\sigma_{xx}+\sigma_{yy}=-P$ $=4\,\mathrm{Re}[\Phi(z)]$. Consequently $\mathrm{Re}[\Phi(z)]$ is equal to $-P/4$. Since Eq. (2) shows that σ_{xx}, σ_{yy}, and τ_{xy} are simple functions of the sine and cosine of θ at r equal to R, as a trial, $\Phi(z)$ was set at $-P/4$ and $\Psi(z)$ was made a function of θ. Then $\Phi'(z)$ became zero and Eq. (4) became

$$\frac{2P}{\pi}\sin\theta + \frac{iP}{\pi}\cos\theta = 2\Psi(z) \quad \text{at } r=R \qquad (5)$$

Use of the complex form of the sine function, $(e^{i\theta}-e^{-i\theta})/2i$, and of the cosine function, $(e^{i\theta}+e^{-i\theta})/2$, gives

$$\Psi(z) = \frac{-iP}{4\pi}e^{i\theta} + \frac{3iP}{4\pi}e^{-i\theta} \quad \text{at } r=R \qquad (6)$$

Since $z=re^{i\theta}$ and the complex conjugate of z is $re^{-i\theta}$, $\Psi(z)$ may be written as

$$\Psi(z) = \frac{-iPR}{4\pi\bar{z}} + \frac{3iPR}{4\pi z} \quad \text{at } r=r \qquad (7)$$

Equations (3) and (4) can be combined to give

$$\sigma_{yy} + i\tau_{xy} = 2\,\mathrm{Re}[\Phi(z)] + \bar{z}\Phi'(z) + \Psi(z) \qquad (8)$$

Combining $\Phi(z)$ equal to $-P/4$ with Eqs. (3), (7), and (8) gives

$$\sigma_{yy} = \frac{-P}{2} + \frac{PR\sin\theta}{\pi r}$$

$$\sigma_{xx} = \frac{-P}{2} - \frac{PR\sin\theta}{\pi r} \quad \text{at } r=r \qquad (9)$$

As a check on the previously constructed functions $\Phi(z)$ and $\Psi(z)$, it can be seen that at r equal to R, Eq. (9) is in agreement with Eq. (2). The stress σ_{yy} from Eq. (9) is shown in Fig. 3.

3 Stress in the Interior of an Ellipse Representing a Grain Under an Applied Tensile Load

The method introduced will again use the two functions Φ and Ψ of the complex variable z [26]. The method also uses the conformal mapping $z=\omega(\zeta)$, which maps points in the z plane onto circles in the ζ plane and points in the ζ plane back onto the z plane. ζ is a complex variable in polar coordinates in the image plane equal to $\rho e^{i\theta}$. An ellipse in the z plane can also be expressed in terms of elliptical coordinates ρ and θ. For the elliptical coordinates ρ and θ, the unit vector $\hat{\rho}$ is in a direction perpendicular to the perimeter of the ellipse. The unit vector $\hat{\theta}$ is orthogonal to $\hat{\rho}$ in an angular direction measured counterclockwise from $\hat{\rho}$.

With the two functions Φ and Ψ and the mapping $z=\omega(\zeta)$ just discussed, the final stresses in the real plane in a polar coordinate system ρ, θ are [26]

$$\sigma_{\rho\rho} - i\sigma_{\rho\theta} = \Phi(\zeta) + \overline{\Phi(\zeta)} - \frac{\zeta^2}{\rho^2\overline{\omega'(\zeta)}}\{\overline{\omega(\zeta)}\Phi'(\zeta) + \omega'(\zeta)\Psi(\zeta)\}$$

$$(10)$$

In Eq. (10), the functions $\Phi(\zeta)$ and $\Psi(\zeta)$ are related to two other stress functions φ and ψ and the mapping function $\omega(\zeta)$ by [26]

Compressive Stress in y Direction
Distance (in units of the radius R)
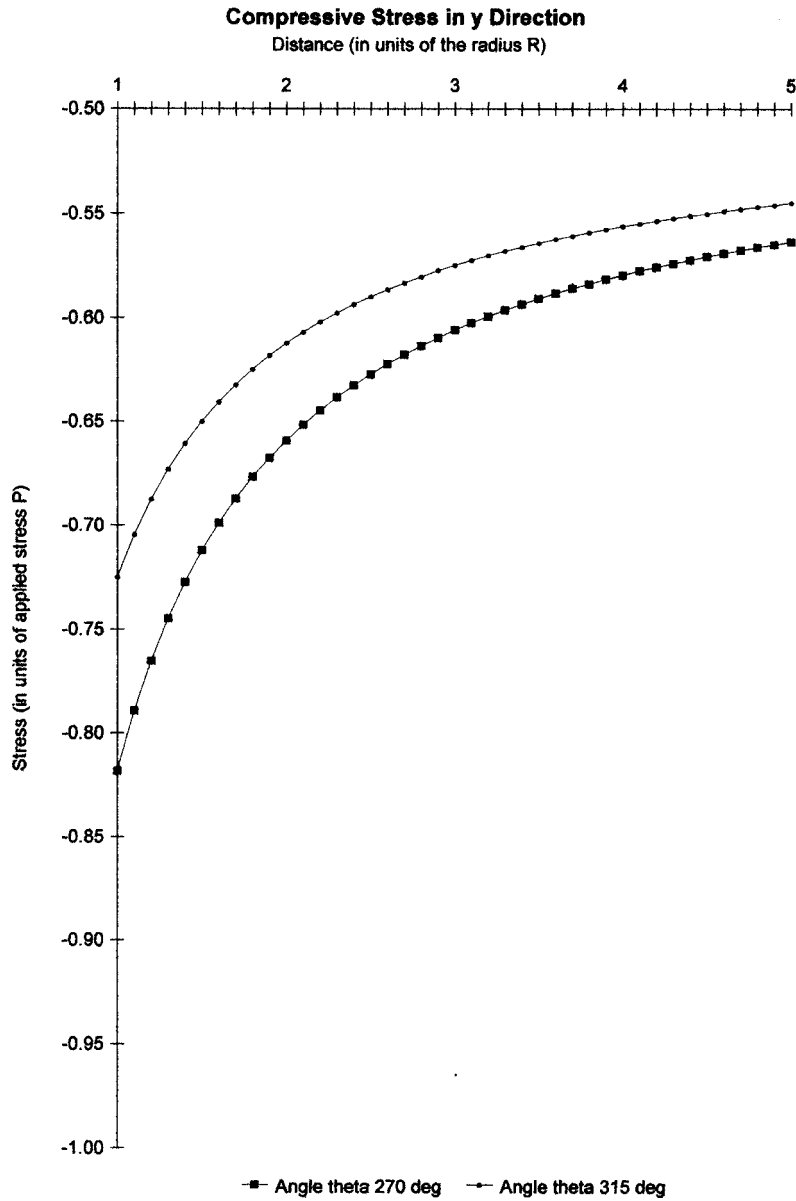
Fig. 3   Compressive stress σ_{yy} below the circular cross section as a function of distance from the center of the cross section

$$\Phi(\zeta) = \frac{\varphi'(\zeta)}{\omega'(\zeta)} \qquad (11)$$

$$\Psi(\zeta) = \frac{\psi'(\zeta)}{\omega'(\zeta)} \qquad (12)$$

The two functions φ and ψ are already known for the region exterior to the ellipse [26]. In this study, the two stress functions φ_0 and ψ_0 for the interior of the ellipse were found. The elastic stresses for both the interior and exterior of the ellipse were then calculated and plotted.

The two stress functions φ and ψ are in general related to a function $f_1 + if_2$ by the equation [26]

$$\varphi(\zeta_1) + \frac{\omega(\zeta_1)}{\omega'(\zeta_1)}\overline{\varphi'(\zeta_1)} + \overline{\psi(\zeta_1)} = f_1 + if_2 \qquad (13)$$

In Eq. (13), ζ_1 is the value of ζ along the boundary of the geo-

metric entity in the image plane from a conformal mapping. The function $f_1 + if_2$, referred to in this paper as the function f, is related to the net stresses acting on the boundary of the geometric entity in the real plane.

The interior and exterior of the ellipse use a different scale in curvilinear coordinates. What remains the same for the regions both inside and outside ellipse, however, is the scale based upon the complex variable z. With the mapping $z = \omega(\zeta)$ for the region exterior to the ellipse, the function f, calculated for the exterior of the ellipse from Eq. (13), was expressed in terms of the complex variable z. The function $z = \omega(\zeta)$, specific to the elliptic rings modeling the interior of the ellipse, was then used to express the function f in terms of the curvilinear coordinates for the ellipse interior.

The process of finding φ_0 and ψ_0 for the interior of the ellipse then involved taking the complex conjugate of Eq. (13) for the ellipse interior and finding a series in complex form for the function f:
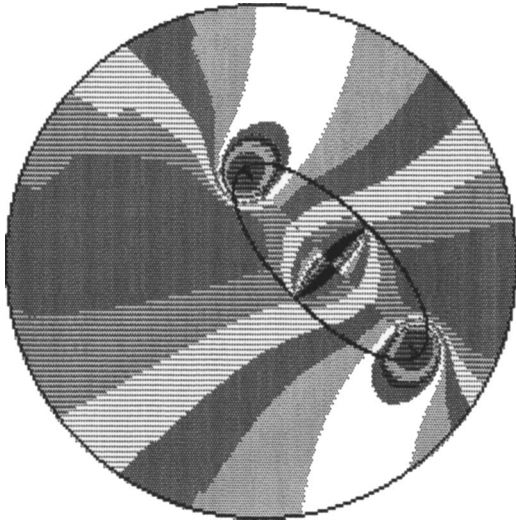
Fig. 4 Stress $\sigma_{\rho\rho}$ in the direction perpendicular to the perimeter of the ellipse. The remote applied tensile stress is aligned with the vertical direction pointing to the top and bottom of the figure. The major axis of the ellipse makes an angle of 45 deg with the applied load. The ratio of the length of the major axis of the ellipse to the length of the minor axis of the ellipse is 5 to 2.

Fig. 6 The stress σ_{yy} is plotted for an elliptical cross section that is nearly circular. In contrast with σ_{yy} shown in Fig. 3 when a circular cross section transmits a compressive stress, shown above is an almost circular cross section that rotates slightly under a tensile load that is aligned with the vertical direction pointing to the top and bottom of the figure. The major axis of the ellipse is at an angle of 45 deg from the remote applied tensile load. The ratio of the length of the major axis of the ellipse to the length of the minor axis of the ellipse is 15 to 14.

$$f_1 - if_2 = \sum_{-\infty}^{\infty} C_k e^{ik\theta} \qquad (14)$$

After expressing several functions in the form of Laurent series [28,29], a series solution [26] was used to find the two functions φ_0 and ψ_0. It was found that φ_0 and ψ_0 have dominant terms consisting of ζ in powers of 1 and -1.

To overcome the differences in form between φ, ψ, φ_0, and ψ_0, and to retain the correct angular distribution of stresses in the region exterior to the ellipse, φ and ψ were left alone and φ_0 and ψ_0 were multiplied by the constants α and β:

$$\varphi_0 = \alpha a_1 \zeta + \frac{\alpha a_{-1}}{\zeta} \qquad (15)$$

$$\psi_0 = \beta a_1' \zeta + \frac{\beta a_{-1}'}{\zeta} \qquad (16)$$

The boundary conditions used to evaluate α and β in Eqs. (15) and (16) are

$$\sigma_{\rho\rho} = \sigma_{\rho\rho}^0 \qquad (17)$$

$$\sigma_{\rho\theta} = \sigma_{\rho\theta}^0 \qquad (18)$$

In Eqs. (17) and (18) σ^0, refers to the stress for the ellipse interior at the outer perimeter of the ellipse and σ without the superscript refers to the stress for the exterior of the ellipse.

To retain the correct angular distribution of the stresses outside the ellipse, the parameters α and β were made functions of θ at each value of θ along the perimeter of the ellipse when satisfying Eqs. (17) and (18). Since the stresses outside the ellipse are single valued and continuous at the boundary, each $\alpha(\theta)$, $\beta(\theta)$ set is single valued and continuous, as well as the resulting stresses inside that are derived from φ_0 and ψ_0.

The results of calculating the stresses outside and inside ellipses of various shapes are shown in Figs. 4–7. Figure 8 shows the scale of shading used for the contours in Figs. 4–7.

4 Discussion

The result that the functions φ_0 and ψ_0 for the interior of an ellipse primarily consist of terms containing ζ and ζ^{-1} appears to be a general case for regions of a shape that allow conformal mapping onto a circular ring. For k greater than or equal to 2, the coefficients a_k in Eq. (15) are given by [26]

$$a_k = \frac{k(\rho_0^2 - \rho_0^{-2})c_k - (\rho_0^{2k} - \rho_0^{-2k})\bar{c}_k}{k^2(\rho_0^2 - \rho_0^{-2})^2 - (\rho_0^{2k} - \rho_0^{-2k})^2} \qquad (19)$$

The coefficients c_k in Eq. (19) above are related to the coefficients C_k in Eq. (14) by [26]

$$c_k = C_{-k}\rho_0^k - C_k\rho_0^{-k} \qquad (20)$$

In Eqs. (19) and (20) above, ρ_0 is the radius of the outer circle of a pair of circles in the image plane. The pair of circles, or a circular ring, is the entity upon which a pair of elliptic rings is

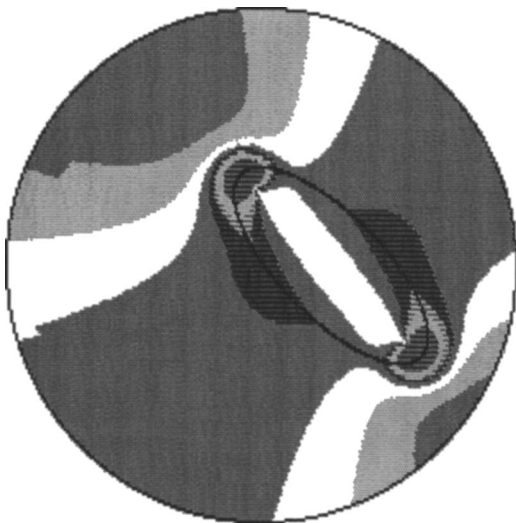Fig. 5 Shear stress $\sigma_{\rho\theta}$ in the direction parallel to the perimeter of the ellipse. The orientation of the applied load, as well as the size and orientation of the ellipse, are the same as in Fig. 4.
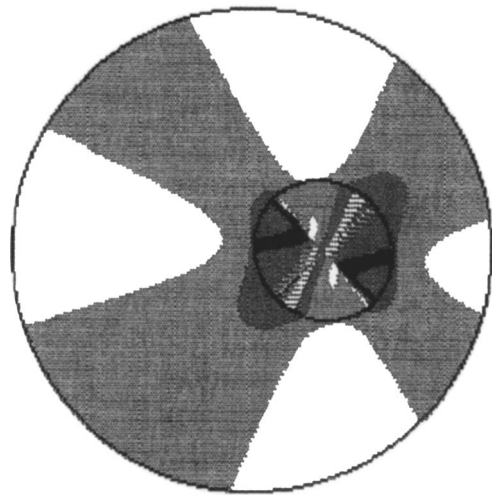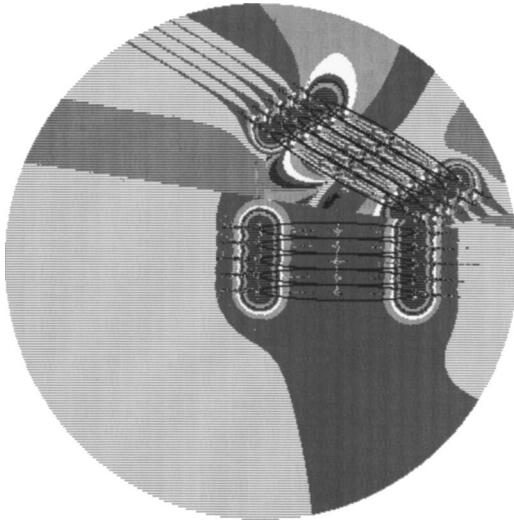
Fig. 7 Stress $\sigma_{\rho\rho}$ plotted for bundles of ellipses adjacent to each other. The stress due to the rotation of a particular ellipse is plotted up to a line midway between it and the next nearest ellipse. The remote applied tensile stress is aligned with the vertical direction pointing to the top and bottom of the figure. The major axis of the ellipses in the lower bundle make and angle of 89 deg with respect to the applied load. The major axes of the ellipses in the upper bundle make and angle of 60 deg with respect to the applied load. The ratio of the length of the major axis of each ellipse to the length of its minor axis is 10 to 1.

mapped. The elliptic rings themselves model the interior the ellipse for a series solution for elastic stresses [26]. It is readily seen from Eq. (19) that for k greater than or equal to 2, a_k becomes rapidly small, in agreement with Eq. (15) derived from the boundary function f.

Close inspection of the derivation of Eq. (19) [26] shows that the factor ρ_0^{2k} appearing in the denominator of Eq. (19) is independent of the mathematical form of the mapping function $z = \omega(\zeta)$ that appears in Eq. (13). The factor ρ_0^{2k} eventually appears
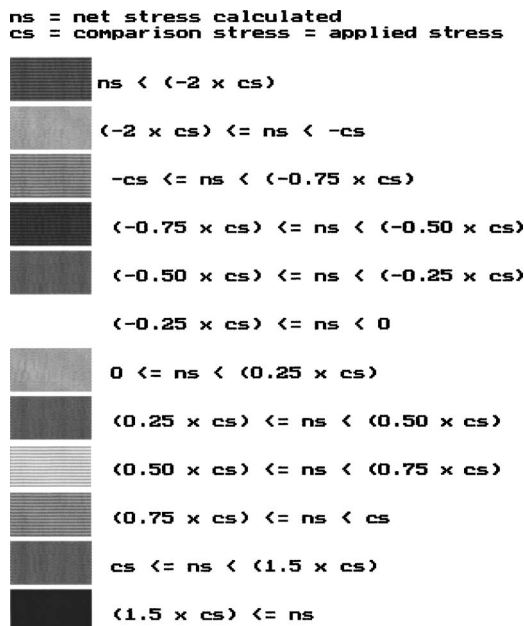
Fig. 8 Shading scale used for the stress contours in Figs. 4–7

in the derivation from being able to express the stress functions in Eq. (13) in both positive and negative powers of ζ when a mapping is done onto a circular ring.

The coefficients a_k' for the series form of ψ_0 in Eq. (16) will likewise become very small as k becomes greater than or equal to 2. In the derivation of the coefficients a_k' [26], it is shown that they can be related to the coefficients a_k themselves, which from Eq. (19) can be seen to become small for k greater than 1. The coefficients a_k' are also related to the coefficients C_k in Eq. (14) multiplied by a factor of ρ_0^{-k}, which again will make the coefficients a_k' small for k greater than 1.

5 Conclusions

(1) Contact stress theory predicts that a constant compressive stress $-P$ will have about 80% of its initial value at the bottom of the lower part of the perimeter of a circular cross section when the stress is transmitted though the circular cross section.

(2) Complex variable methods predict that the compressive stress $-P$ applied to the upper side of a circular cross section will transmit into the infinite plane from the lower side partly as $-P/2$. The stress also transmits as a term that attenuates as $1/r$ and that has an angular dependence of $\sin \theta$.

(3) When the scale of the curvilinear coordinates for the real and image planes is different for two regions on opposite sides of a boundary, the function f may be expressed in terms of z using the first mapping function and then expressed in curvilinear coordinates for the second region using the second mapping function.

(4) For the interior of an ellipse being held in place after rotating under a remote applied tensile load, and in general for a geometric entity that can be mapped onto a circular ring, the dominant terms in the stress functions for the interior have exponents of 1 and -1.

(5) Continuity of the stresses along the boundary makes it possible for the multiplicative constants α and β multiplying the terms of φ_0 and ψ_0 to be functions of the angular coordinate moving around the perimeter. Allowing α and β to be functions of the angular coordinate ensures continuity of stresses across the boundary when the functions φ_0 and ψ_0 on one side contain a disparate number of terms in comparison with the functions φ and ψ on the other side.

References

[1] Ohji, T., Kusunose, T., and Niihara, K., 1998, "Threshold Stress in Creep of Alumina-Silicon Carbide Nanocomposites," J. Am. Ceram. Soc., **81**, pp. 2713–2716.

[2] Oh, S.-T., Sando, M., and Niihara, K., 1998, "Preparation and Properties of Alumina/Nickel-Cobalt Alloy Nanocomposites," J. Am. Ceram. Soc., **81**, pp. 3013–3015.

[3] Davis, L. C., and Allison, J. E., 1993, "Residual Stresses and Their Effects on Deformation in Particle-Reinforced Metal-Matrix Composites," Metall. Trans. A, **24**, pp. 2487–2496.

[4] Kovalev, S., Ohji, T., Yamauchi, Y., and Sakai, M., 2000, "Grain Boundary Strength in Non-Cubic Polycrystals with Misfitting Intragranular Inclusions," J. Mater. Sci., **35**, pp. 1405–1412.

[5] Mizushima, I., Hamada, M., and Shakudo, T., 1978, "Tensile and Compressive Stress Problems for a Rigid Circular Disk in an Infinite Plate," Bull. JSME, **21**, pp. 1325–1333.

[6] Wang, J., Andreasen, J. H., and Karihaloo, B. L., 2000, "The Solution of an Inhomogeneity in a Finite Plane Region and its Application to Composite

Materials," Compos. Sci. Technol., **60**, pp. 75–82.

[7] Mizushima, I., Hamada, M., and Kusano, N., 1979, "Tensile and Compressive Stress Problems for a Circular Disk in an Infinite Plate," Bull. JSME, **22**, pp. 1175–1181.

[8] Gladwell, G. M. L., and Iyer, K. R. P., 1974, "Unbonded Contact Between a Circular Plate and an Elastic Half-Space," J. Elast., **4**, pp. 115–130.

[9] Miller, G. R., and Keer, L. M., 1983, "Interaction Between a Rigid Indenter and a Near-Surface Void or Inclusion," J. Appl. Mech., **50**, pp. 615–620.

[10] Wang, J., Andreasen, J. H., and Karihaloo, B. L., 2000, "The Solution of an Inhomogeneity in a Finite Plane Region and its Application to Composite Materials," Compos. Sci. Technol., **60**, pp. 75–82.

[11] Onaka, S., and Kato, M., 1999, "Effects of Elastic Modulus, Shape and Volume Fraction of an Elastically Inhomogeneous Second Phase on Stress States in a Loaded Composite," Mater. Trans., JIM, **40**, pp. 1102–1107.

[12] Batista, M., 1999, "Stresses in a Confocal Elliptic Ring Subject to Uniform Pressure," J. Strain Anal. Eng. Des., **34**, pp. 217–221.

[13] Chen, D.-H., 1996, "Green's Functions for a Point Force and Dislocation Outside an Elliptic Inclusion in Plane Elasticity," Z. Angew. Math. Mech., **47**, pp. 894–905.

[14] Gross, R. S., Goree, J. G., 1991, "Torsion of a Rigid Smooth Elliptic Insert in an Infinite Elastic Plane," J. Appl. Mech., **58**, pp. 370–375.

[15] Sendeckyj, G. P., 1970, "Elastic Inclusion Problems in Plane Elastostatics," Int. J. Solids Struct., **6**, pp. 1535–1543.

[16] Karihaloo, B. L., and Viswanathan, K., 1988, "A Partially Debonded Ellipsoidal Inclusion in an Elastic Medium. Part I: Stress and Displacement Fields," Mech. Mater., **7**, pp. 191–197.

[17] Noda, N.-A., Tomari, K., and Matsuo, T., 1999, "Interaction Effect Between Ellipsoidal Inclusions in an Infinite Body Under Asymmetric Uniaxial Tension," JSME Int. J., Ser. A, **42**, pp. 372–380.

[18] Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion and Related Problems," Proc. R. Soc. London, Ser. A, **241**, pp. 376–396.

[19] Eshelby, J. D., 1959, "The Elastic Field Outside an Ellipsoidal Inclusion," Proc. R. Soc. London, Ser. A, **252**, pp. 561–569.

[20] Edwardes, D., 1893, "Steady Motion of a Viscous Liquid in Which an Ellipsoid is Constrained to Rotate About a Principal Axis," Q. Appl. Math., **26**, pp. 70–78.

[21] Huang, Y., Hu, K. X., and Chandra, A., 1995, "Stiffness Evaluation for Solids Containing Dilute Distributions of Inclusions and Microcracks," ASME J. Appl. Mech., **62**, pp. 71–77.

[22] Erdogan, F., Gupta, G. D., and Ratwani, M., 1974, "Interaction Between an Inclusion and an Arbitrarily Oriented Crack," ASME J. Appl. Mech., **41**, pp. 1007–1013.

[23] Dundurs, J. and Mura, T., 1964, "Interaction Between an Edge Dislocation and a Circular Inclusion," J. Mech. Phys. Solids, **12**, pp. 177–189.

[24] Boniface, V., and Hasebe, N., 1998, "Solution of the Displacement Boundary Value Problem of an Interface Between Two Dissimilar Half-Planes and a Rigid Elliptic Inclusion at the Interface," ASME J. Appl. Mech., **65**, pp. 880–888.

[25] Ballarini, R., 1990, "A Rigid Line Inclusion at a Bimaterial Interface," Eng. Fract. Mech., **37**, pp. 1–5.

[26] Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, P. Noordhoff, Ltd.

[27] Frocht, M. M., 1948, *Photoelasticity*, Wiley, New York, Vol. II.

[28] Wylie, C. R., and Barrett, L. C., 1982, *Advanced Engineering Mathematics*, McGraw–Hill, New York.

[29] Churchhill, R. V., Brown, J. W., and Verhey, R. F., 1976, *Complex Variables and Applications*, McGraw–Hill, New York.

Leone Corradi

Lelio Luzzi
e-mail: lelio.luzzi@polimi.it

Fulvio Trudi

Department of Nucelar Engineering,
Politecnico di Milano,
Via Ponzio 34/3,
20133 Milan, Italy

Collapse of Thick Cylinders Under Radial Pressure and Axial Load

This paper provides the theoretical collapse loads of thick, long cylindrical shells subject to pressure and axial forces. Tubes are made of isotropic, perfectly plastic von Mises' material. Axial strains are assumed to be constant but possibly different from zero, so that elongation is permitted. This assumption, together with axial symmetry and the isochoric nature of plastic flow, unambiguously defines the set of possible collapse mechanisms, and collapse loads are computed on this basis. Results are contrasted to those presently available, based on thin-shell assumptions. Comparison shows that differences are of engineering significance, well worth considering for thick tubes, such as those envisaged in some nuclear power plant applications. [DOI: 10.1115/1.1938204]

1 Introduction

The assessment of load-bearing capacity of shells can be considered an issue satisfactorily settled when shells are thin enough to collapse because of elastic buckling, as typical of aeronautic or aerospace applications. Outside this context, however, shells of higher thickness often are required. Medium-thick shells are employed, for instance, in the oil industry as pipes or casings (with thickness increasing as the depth of the water in which the pipes operate), and recent proposals for innovative nuclear power plant design consider steam generator tubes of significant thickness pressurized from outside [1].

When thick tubes are subject to external pressure, collapse is initiated (and often dominated) by yielding, but interaction with instability is meaningful, in that imperfections reduce the load bearing capacity by an amount of engineering significance also when thickness is considerable. At present, such an effect is accounted for by means of more or less empirical formulas, defining the reduction with respect to the plastic collapse load induced by coupling with instability [2–8].

Independent of the adequacy of such formulas, often borrowed from problems, such as beam columns, only partially similar to thick tubes, the very definition of the reference value demands discussion. In general, the plastic collapse pressure is computed by exploiting thin-shell assumptions, which consider stresses constant throughout. Under uniform pressure the tube becomes statically determinate, with the consequence that the elastic limit is overestimated and the collapse pressure underestimated. Discrepancies are negligible as long as the ratio between the radius of the cylinder and its wall thickness is large, but get more and more significant as this ratio decreases.

The pressure values at the onset of yielding (elastic limit) are easily computed from the well-known elastic solutions [9], and the correct thick shell values are used by most codes (see, e.g., [10]). The analogous results for plastic collapse, on the contrary, are available only for tubes in plane strain [11,12], a situation of interest but by no means the only significance.

To clarify this point, the kinematics of deformation of long cylinders is examined. The tube being subject to uniform pressures (for completeness, internal pressure is also included) and, possibly, to constant axial force, each cross section undergoes the same loading, and its response is essentially "plane," in that stresses and strains are independent of the axial coordinate, say z. Classical plane solutions, however, are not adequate. A slice of the cylinder in plane stress conditions would experience (in the elastic-plastic range) nonuniform transverse strains ε_z, in general, conflicting with those of adjacent slices. A plane-strain assumption solves the conflict by imposing that ε_z be zero throughout, but this constraint appears excessively severe; in fact, continuity between adjacent slices is merely expected to make axial strains uniform without preventing possible elongation. For long tubes the most realistic model seems that of *generalized plane strain*, which assumes that ε_z is constant, but not necessarily zero.

In this paper long, thick cylinders subject to (external and/or internal) pressure and axial load are considered and the values of such loads bringing, individually or together, the cylinder to collapse are determined. To this purpose the *kinematic theorem* of limit analysis is employed in conjunction with the von Mises yield criterion. In spite of the upper-bound nature of the kinematic theorem, the result is exact; in fact, the assumptions of axial symmetry, generalized plane strain, and isochoric plastic flow unambiguously define the set of possible collapse mechanisms, governed by the ratio among two parameters, namely, the radius variation and axial elongation.

Results are contrasted to the corresponding elastic limits and to predictions stemming from thin-shell assumptions. Comparison permits the assessment of some points, such as the resources with respect to the elastic limit provided by stress redistribution and the range of validity of thin-shell approximation. In particular, it appears that the latter assumption is too restrictive for really thick tubes, as those required by some fourth-generation nuclear plant applications.

Limit analysis is based inherently on the small strain hypothesis and results are unable to assess the influence of the plasticity-instability interaction on the collapse level—influence which is significant in medium-thick cylinders and plays some role even in the definitely thick cylinders. This aspect is presently investigated and some preliminary results are presented in [13]. As was already mentioned, however, formulas evaluating such effects use the collapse load as a reference value and its correct definition seems to be a preliminary, but important, starting point toward a rational assessment of the load-bearing capacity of this structural typology.

2 General Relations

The cylinder in Fig. 1 is considered. Loads consist of external pressure q, internal pressure p, and axial force F, all constant throughout. Pressures are supposed to be always positive, while F can assume either sign, with $F > 0$ corresponding to tension. The
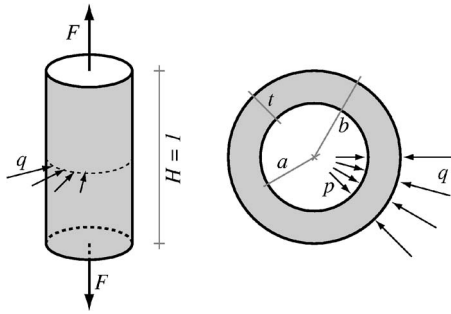
Fig. 1   Geometry and load conditions

material is elastically isotropic and perfectly plastic, with a yield limit governed by the von Mises criterion. The cylinder is in a generalized plane-strain situation.

2.1   Elastic Solution and Elastic Limit. The elastic solution is constructed by considering that a slice of the cylinder in plane stress conditions would undergo uniform displacements in the axial direction [9]; hence, when a number of slices piles up to build a long cylinder, no conflict among them arises and the plane stress solution maintains its validity. On this basis, one can write

$$\sigma_r = -q\frac{b^2}{b^2-a^2}\frac{r^2-a^2}{r^2} - p\frac{a^2}{b^2-a^2}\frac{b^2-r^2}{r^2} \tag{1a}$$

$$\sigma_\vartheta = -q\frac{b^2}{b^2-a^2}\frac{r^2+a^2}{r^2} + p\frac{a^2}{b^2-a^2}\frac{b^2+r^2}{r^2} \tag{1b}$$

$$\sigma_z = n \tag{1c}$$

where $(1a)$ and $(1b)$ are the plane stress components and

$$n = \frac{F}{\pi(b^2-a^2)} \tag{2}$$

is the axial force per unit cross section, defining the axial stress due to F.

Since shearing stresses vanish throughout, the von Mises yield criterion reads

$$\sigma_e = \frac{1}{\sqrt{2}}\sqrt{(\sigma_r-\sigma_\vartheta)^2 + (\sigma_\vartheta-\sigma_z)^2 + (\sigma_z-\sigma_r)^2} \leqslant \sigma_0 \tag{3}$$

σ_e being the *effective stress* and σ_0 the tensile yield strength. The elastic limit is attained when Eqs. $(1a)$–$(1c)$ fulfill condition (3) as an equality for some r in the interval $a \leqslant r \leqslant b$.

2.2   Plastic Collapse. The kinematic theorem of limit analysis [14] states that the ultimate load can be computed by equating the work of external loads to the work plastically dissipated in the motion corresponding to the *collapse mechanism* (strictly speaking, the equality involves powers rather than works, the mechanism motion being defined in terms of velocities, but the terminology used often is preferred). For arbitrary mechanisms the procedure only provides upper bounds; however, the exact value is obtained when the actual collapse mechanism is employed. In the present case, this can be defined unambiguously because of the conditions of polar symmetry, generalized plane strain, and isochoric plastic flow, the latter imposed by the normality rule when the von Mises criterion is used. These conditions imply
Polar symmetry:

$$v_r = v(r) \quad v_\vartheta = 0, \quad \dot{\varepsilon}_r = \frac{dv}{dr} \quad \dot{\varepsilon}_\vartheta = \frac{v}{r} \tag{4a}$$

Generalized plane strain:

$$\dot{\varepsilon}_z = C \quad \text{(const)} \tag{4b}$$

Isochoric plastic flow:

$$\dot{\varepsilon}_r + \dot{\varepsilon}_\vartheta + \dot{\varepsilon}_z = 0 \tag{4c}$$

Equations $(4a)$–$(4c)$ produce the following differential equation:

$$\frac{dv}{dr} + \frac{v}{r} + C = 0$$

which defines the radial velocity field

$$v(r) = -V\frac{b}{r} + \frac{1}{2}C\left(\frac{b^2-r^2}{r}\right) \quad \text{(positive outward)} \tag{5}$$

and the consequent strain rates

$$\dot{\varepsilon}_r = V\frac{b}{r^2} - \frac{1}{2}C\left(1 + \frac{b^2}{r^2}\right) \tag{6a}$$

$$\dot{\varepsilon}_\vartheta = -V\frac{b}{r^2} - \frac{1}{2}C\left(1 - \frac{b^2}{r^2}\right) \tag{6b}$$

$$\dot{\varepsilon}_z = C \tag{6c}$$

The integration constant V is chosen so as to represent the outer radius velocity, positive if directed inward $[V = -v(b)]$. The velocity U of the inner radius (positive outward) is

$$U = v(a) = -V\frac{b}{a} + \frac{1}{2}C\left(\frac{b^2-a^2}{a}\right)$$

Equations (5) and $(6a)$–$(6c)$ govern the motion of the only mechanisms consistent with the assumptions, which are governed by the two parameters V and C. For von Mises' materials the *dissipation* (per unit volume), when expressed as an explicit function of strain rates, reads

$$\hat{D} = \frac{\sqrt{2}}{3}\sigma_0\sqrt{(\dot{\varepsilon}_r-\dot{\varepsilon}_\vartheta)^2 + (\dot{\varepsilon}_\vartheta-\dot{\varepsilon}_z)^2 + (\dot{\varepsilon}_z-\dot{\varepsilon}_r)^2} = \sqrt{\frac{2}{3}}\sigma_0\sqrt{\dot{\varepsilon}_r^2 + \dot{\varepsilon}_\vartheta^2 + \dot{\varepsilon}_z^2}$$

where equality holds because of constraint $(4c)$. Then, from Eqs. $(6a)$–$(6c)$ one obtains

$$\hat{D} = \frac{2\sigma_0}{\sqrt{3}}\sqrt{V^2\frac{b^2}{r^4} - VC\frac{b^3}{r^4} + \frac{1}{4}C^2\left(3 + \frac{b^4}{r^4}\right)} \tag{7}$$

The kinematic theorem of limit analysis establishes

$$\Pi = D \tag{8}$$

where

$$\Pi = 2\pi bqV + 2\pi apU + FC = 2\pi b(q-p)V + \pi(b^2-a^2)(n+p)C \tag{9a}$$

is the external work and

$$D = 2\pi\int_a^b \hat{D}r\,dr = \frac{4\pi}{\sqrt{3}}\sigma_0\int_a^b\sqrt{V^2\frac{b^2}{r^4} - VC\frac{b^3}{r^4} + \frac{1}{4}C^2\left(3 + \frac{b^4}{r^4}\right)}\,r\,dr \tag{9b}$$

is the amount dissipated by the material. In writing Eqs. $(9a)$ and $(9b)$ the length H of the cylinder was assumed as unitary. The two parameters V and C must be selected so that, for the load condition considered, the external work $(9a)$ is positive.

The dissipation D is a positive homogeneous of degree one function of V, C, and one can write

$$D = \frac{\partial D}{\partial V}V + \frac{\partial D}{\partial C}C$$

so that Eq. (8) becomes

$$\left(2\pi b(q-p) - \frac{\partial D}{\partial V}\right)V + \left(\pi(b^2-a^2)(n+p) - \frac{\partial D}{\partial C}\right)C = 0 \tag{10}$$

This relation applies independently of the relative values of V and C. Hence

$$2\pi b(q-p)=\frac{\partial D}{\partial V}=\frac{2\pi}{\sqrt{3}}\sigma_0\int_a^b\frac{2V\dfrac{b^2}{r^4}-C\dfrac{b^3}{r^4}}{\sqrt{V^2\dfrac{b^2}{r^4}-VC\dfrac{b^3}{r^4}+\dfrac{1}{4}C^2\left(3+\dfrac{b^4}{r^4}\right)}}rdr$$

$$(11a)$$

$$\pi(b^2-a^2)(n+p)=\frac{\partial D}{\partial C}$$

$$=\frac{2\pi}{\sqrt{3}}\sigma_0\int_a^b\frac{-V\dfrac{b^3}{r^4}+\dfrac{1}{2}C\left(3+\dfrac{b^4}{r^4}\right)}{\sqrt{V^2\dfrac{b^2}{r^4}-VC\dfrac{b^3}{r^4}+\dfrac{1}{4}C^2\left(3+\dfrac{b^4}{r^4}\right)}}rdr$$

$$(11b)$$

The mechanism amplitude being arbitrary, different mechanisms merely correspond to different ratios C/V. Let this ratio be represented by a dimensionless parameter κ, defined by the relation

$$C=\kappa\frac{V}{b} \qquad (12)$$

then one can write

$$V^2\frac{b^2}{r^4}-VC\frac{b^3}{r^4}+\frac{1}{4}C^2\left(3+\frac{b^4}{r^4}\right)=V^2\left(\frac{b^2}{r^4}\left(1-\frac{1}{2}\kappa\right)^2+\frac{3}{4}\kappa^2\frac{1}{b^2}\right)$$

$$2V\frac{b^2}{r^4}-C\frac{b^3}{r^4}=2V\frac{b^2}{r^4}\left(1-\frac{1}{2}\kappa\right)$$

$$-V\frac{b^3}{r^4}+\frac{1}{2}C\left(3+\frac{b^4}{r^4}\right)=\frac{V}{b}\left(\frac{3}{2}\kappa-\frac{b^4}{r^4}\left(1-\frac{1}{2}\kappa\right)\right)$$

and Eqs. (11a) and (11b) become

$$q-p=\frac{V}{|V|}\frac{2}{\sqrt{3}}\sigma_0\int_a^b\frac{1-\dfrac{1}{2}\kappa}{\sqrt{\left(1-\dfrac{1}{2}\kappa\right)^2+\dfrac{3}{4}\kappa^2\dfrac{r^4}{b^4}}}\frac{dr}{r} \qquad (13a)$$

$$n+p=\frac{V}{|V|}\frac{2}{\sqrt{3}}\sigma_0\frac{b^2}{b^2-a^2}\int_a^b\frac{\dfrac{3}{2}\kappa\dfrac{r^4}{b^4}-\left(1-\dfrac{1}{2}\kappa\right)}{\sqrt{\left(1-\dfrac{1}{2}\kappa\right)^2+\dfrac{3}{4}\kappa^2\dfrac{r^4}{b^4}}}\frac{dr}{r}$$

$$(13b)$$

where the sign of V is dictated by the condition

$$\Pi=\frac{\pi}{b}[2b^2(q-p)+\kappa(b^2-a^2)(n+p)]V>0 \qquad (14)$$

For any assigned value of κ, Eqs. (13a) and (13b) establish the relations among q,p, and F at collapse. Closed-form expressions for the integrals on their right-hand sides are available, but do not provide significant advantages with respect to numerical integration, which is carried out easily by exploiting standard software, such as MATLAB.

2.3 Thin-Shell Approximation. For comparative small ratios t/b, $t=b-a$ being the wall thickness, stresses are usually assumed to be constant. This makes the cylinder *statically determinate*, and global equilibrium establishes

$$\sigma_\vartheta=-q\frac{b}{t}+p\frac{a}{t},\quad \sigma_z=n,\quad \sigma_r=0 \qquad (15)$$

If tubes are actually thin, often the further assumption $a\simeq b\simeq R$, $R=\frac{1}{2}(a+b)$ being the mean radius, is introduced when writing the
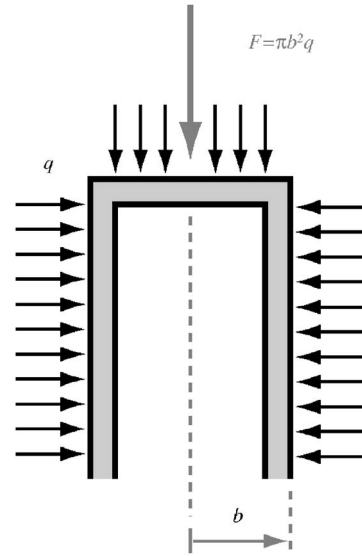
Fig. 2 Hydrostatic pressure

first of Eqs. (15). However, the distinction among internal and external radii is of importance when the assumption is used for moderately thick cylinders, as frequently done in the literature.

The cylinder being statically determinate, its elastic and collapse limits coincide and are obtained from condition (3) imposed as an equality for the stresses above.

3 No Internal Pressure

3.1 Elastic Limit. The meaning of the results in the preceding sections is better understood if the special case $p=0$ is first considered. As for the elastic limit, observe that the von Mises effective stress is maximum at the inner radius, where one has

$$\sigma_r=0,\quad \sigma_\vartheta=-2q\frac{b^2}{b^2-a^2},\quad \sigma_z=n$$

Then condition (3) reduces to a quadratic equation that is easily solved to give

$$q=\sigma_0\frac{1}{2}\frac{b^2-a^2}{b^2}\left(\sqrt{1-\frac{3}{4}\left(\frac{n}{\sigma_0}\right)^2}-\frac{1}{2}\frac{n}{\sigma_0}\right) \qquad (16)$$

Equation (16) expresses the pressure at the elastic limit as a function of the axial load. Particular cases worth mentioning are *Pure pressure*: When $n=0$ the solution reads

$$n=0,\quad q=q_e=\sigma_0\frac{t}{b}\left(1-\frac{1}{2}\frac{t}{b}\right) \qquad (17a)$$

Pure axial load: For $q=0$ one obtains the obvious result

$$q=0\quad n=n_e=\pm\sigma_0 \qquad (17b)$$

Hydrostatic pressure: This condition implies (see Fig. 2)

$$F=-\pi b^2q\Rightarrow n=-\frac{b^2}{b^2-a^2}q$$

As a consequence, at the elastic limit one has

$$q_{He}=\frac{2}{\sqrt{3}}\sigma_0\frac{t}{b}\left(1-\frac{1}{2}\frac{t}{b}\right),\quad n_{He}=-\frac{1}{\sqrt{3}}\sigma_0 \qquad (17c)$$

In writing the results above, use was made of the geometric relation

$$\frac{1}{2}\frac{b^2-a^2}{b^2}=\frac{t}{b}\left(1-\frac{1}{2}\frac{t}{b}\right) \qquad (18)$$

3.2 Collapse. For $p=0$, Eqs. (13a) and (13b) become

$$q = \pm \frac{2}{\sqrt{3}}\sigma_0 \int_a^b \frac{1 - \frac{1}{2}\kappa}{\sqrt{\left(1 - \frac{1}{2}\kappa\right)^2 + \frac{3}{4}\kappa^2 \frac{r^4}{b^4}}}\frac{dr}{r} \qquad (19a)$$

$$n = \pm \frac{2}{\sqrt{3}}\sigma_0 \frac{b^2}{b^2 - a^2}\int_a^b \frac{\frac{3}{2}\kappa\frac{r^4}{b^4} - \left(1 - \frac{1}{2}\kappa\right)}{\sqrt{\left(1 - \frac{1}{2}\kappa\right)^2 + \frac{3}{4}\kappa^2 \frac{r^4}{b^4}}}\frac{dr}{r} \qquad (19b)$$

the sign being decided by condition (14).

For particular values of κ the solution is easy. As it is immediatly verified, $\kappa=2$ implies

$$\kappa = 2: \Rightarrow q = 0 \quad n = n_0 = \pm \sigma_0 \qquad (20)$$

and corresponds to collapse under (tensile or compressive) axial load alone. For the case $\kappa=0$, which implies a plane strain situation (no elongation in the axial direction), one obtains

$$\kappa = 0: \Rightarrow q = q_H = \frac{2}{\sqrt{3}}\sigma_0 \lg\frac{b}{a}$$

$$n = n_H = -\frac{2}{\sqrt{3}}\sigma_0\frac{b^2}{b^2 - a^2}\lg\frac{b}{a} = -\frac{b^2}{b^2 - a^2}q_H \qquad (21)$$

The first of these results is well known [11]. The second establishes that this situation corresponds to hydrostatic pressure (Fig. 2).

The case of pure external pressure is not equally straightforward. The relevant value κ_q of κ is obtained by solving Eq. (19b), written for $n=0$. Namely,

$$\int_a^b \frac{\frac{3}{2}\kappa\frac{r^4}{b^4} - \left(1 - \frac{1}{2}\kappa\right)}{\sqrt{\left(1 - \frac{1}{2}\kappa\right)^2 + \frac{3}{4}\kappa^2 \frac{r^4}{b^4}}}\frac{dr}{r} = 0 \Rightarrow \kappa_q$$

When (19a) is evaluated for κ_q, results turn out to be approximated to an excellent accuracy by the equation

$$q = \sigma_0\frac{t}{b}\left(1 + k\frac{t}{b}\right)$$

A best fit over the interval $4 \leqslant b/t \leqslant 20$ establishes $k=0.247$ and, with a further slight approximation, one can write

$$\kappa = \kappa_q: \Rightarrow q = q_0 \simeq \sigma_0\frac{t}{b}\left(1 + \frac{1}{4}\frac{t}{b}\right) \quad n = 0 \qquad (22)$$

The expression above is plotted in Fig. 3 (solid line). Comparison with numerical results (dots) shows that it is fully acceptable for engineering purposes.

It is of interest to note that the same formula with $k=0.235$ was arrived at empirically on the basis of numerical solutions for long cylinders, with no a priori enforcement of the generalized plane strain constraint [3]. The difference in coefficient k affects the result by $<0.4\%$ already for $b/t=4$, and the error diminishes rapidly with increasing slenderness.

The solutions for other values of κ are computed from Eqs. (19a) and (19b). In this way, interaction curves in the q-n plane can be constructed. For some b/t ratios they are depicted in Fig. 4. Dots indicate where the solution predicts $V=0$, when the lateral expansion caused by axial compression exactly compensates the effect of external pressure and separate zones with different signs in Eqs. (19a) and (19b).
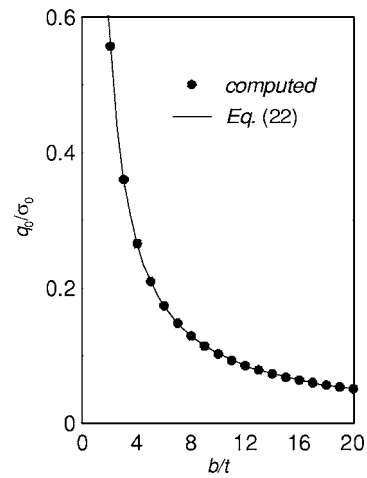
Fig. 3 Limit external pressure

3.3 Thin-Cylinder Approximation. By introducing expressions (15) written for $p=0$ in condition (3) and by enforcing it as an equality, one obtains

$$q = \sigma_0\frac{t}{b}\left(\sqrt{1 - \frac{3}{4}\left(\frac{n}{\sigma_0}\right)^2} - \frac{1}{2}\frac{n}{\sigma_0}\right) \qquad (23)$$

Eq. (23) is widely used to express the limit external pressure as function of the axial load [3,8]. One has, in particular: *Pure pressure*:

$$q_0^{TS} = \sigma_0\frac{t}{b} \quad n = 0 \qquad (24a)$$

Axial load only:

$$q = 0 \quad n_0^{TS} = \pm \sigma_0 \qquad (24b)$$

Hydrostatic pressure: The thin-cylinder approximation implies in this case

$$\frac{\sigma_z}{\sigma_\vartheta} = -\frac{n t}{q b} = \frac{1}{2}$$

and from (23) one obtains

$$q_H^{TS} = \frac{2}{\sqrt{3}}\sigma_0\frac{t}{b} \quad n_H^{TS} = -\frac{1}{\sqrt{3}}\sigma_0 \qquad (24c)$$

4 General Loading

Fully analogous considerations apply when internal pressure also acts, and results are presented and discussed, omitting computational details.

The elastic limit is evaluated straightforwardly by enforcing equality in Eq. (3) at the inner radius $r=a$, where yielding always initiates. Only, it must be noted that values for internal pressure
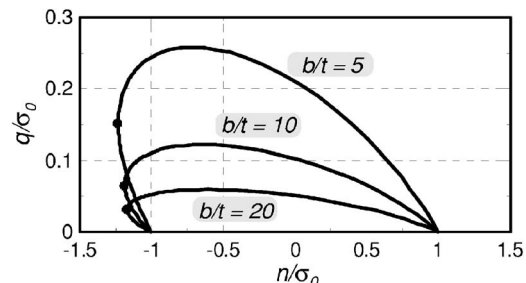
Fig. 4 External pressure versus axial load interaction curves
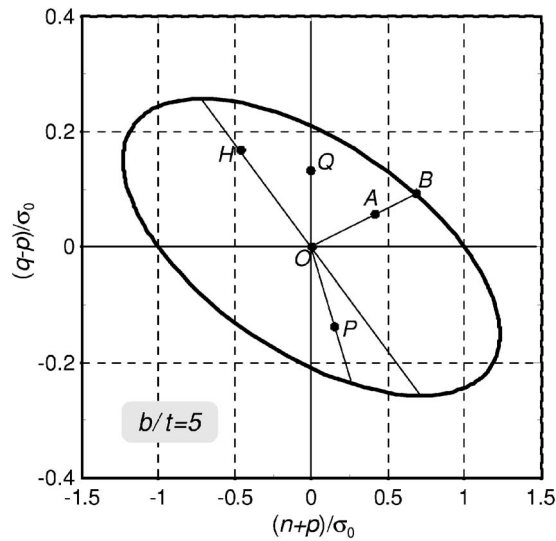
Fig. 5 Collapse limit for different q, p, n combinations

Fig. 6 Internal pressure versus axial load interaction curves

might be slightly different from those usually found in the literature, mainly based on the Tresca criterion (e.g., [12]).

As for collapse, observe that Eqs. (13a) and (13b) become identical to Eqs. (19a) and (19b) when q and n are replaced by $q-p$ and $n+p$, respectively. Hence, previous solutions maintain their validity, as well as their graphical representations, which only require changing the coordinate labels, but negative ordinates now are meaningful. Figure 5 shows the resulting collapse limit for $b/t=5$; the upper part of the curve is as in Fig. 4, and the lower part is nothing but its reflection about both axes.

Any load condition corresponds to a point in the plane, even if the same point is representative of different loading. When a point (such as A) is internal to the domain, the corresponding load conditions do not exceed the load-bearing capacity of the tube and the margin with respect to collapse (*safety factor*) is given by the ratio OB/OA. Pure external pressure conditions locate themselves on the ordinate axis (point Q), while the opposite situation of pure internal pressure corresponds to points like P, moving along a line with a slope of -45 deg if the same scales were used for both axes.

For a closed tube subject to both external and internal pressures, one has

$$ n = -q\frac{b^2}{b^2-a^2} + p\frac{a^2}{b^2-a^2} \tag{25} $$

corresponding in Fig. 5 to points, such as H, moving along the straight line

$$ (q-p) = -\frac{b^2-a^2}{b^2}(n+p) $$

Collapse always corresponds to $\kappa=0$ (plane strain).

When defining the collapse mechanisms, the radial parameter was identified with the velocity V at the outer radius. As a consequence, the interaction curves obtained, such as that in Fig. 5, essentially refer to q and n, with the internal pressure p playing the role of a "correction," which diminishes the effect of the external one and contributes to axial tension (see Eqs. (13a) and (13b)). This makes cumbersome the evaluation of the safety factor when $p > q$, i.e., when the lower curves are to be used.

The choice of V instead of U as a free parameter is, in a sense, arbitrary. It was made because the case $q > p$ was considered of greater interest, and the main purpose was to obtain interaction curves, permitting an easy study of this situation. Nevertheless, the opposite one is relevant enough to deserve some attention.
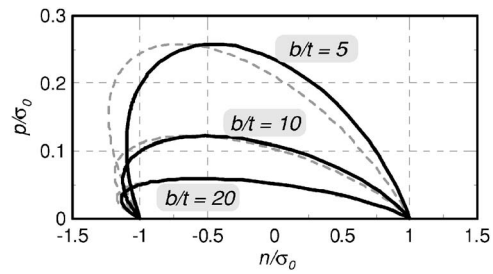
For the sake of simplicity, discussion is limited to the case $q=0$, which does not require that the procedure be restarted by replacing Eq. (5) with a velocity field expressed in terms of its inner radius value U. The collapse internal pressure p associated to any given κ can be obtained from Eq. (13a) and used to solve (13b) for n. The resulting $p-n$ interaction curves are depicted in Fig. 6 (solid lines) and compared to the $q-n$ curves from Fig. 4 (dashed), which can be considered as representative of a hypothetical negative external pressure. As cylinders get thin, the two curves approach each other, and for $b/t=20$, they are nearly undistinguishable; for stocky tubes, however, differences are significant enough to be considered.

Some comments on the results are in order. The first refers to the ratio between the collapse and the elastic limits, measuring the resources associated to stress redistribution in the plastic range. Two thicknesses are considered, namely, $b/t=5$, typical of nuclear power plant steam generators when pressurized from outside, and $b/t=10$, representative of deep water pipeline or casing applications. The relevant limit curves for $p=0$ are redrawn in Figs. 7(a) and 7(b), respectively, and supplemented by the corresponding elastic limits. Obviously, the tube does not exhibit any plastic resources when acted upon by axial load only, but resources increase with increasing pressure up to the hydrostatic situation (corresponding to the maxima of the curves), where gains with respect to the elastic limit are of 24% for the thicker tube and 11%

(a)

(b)
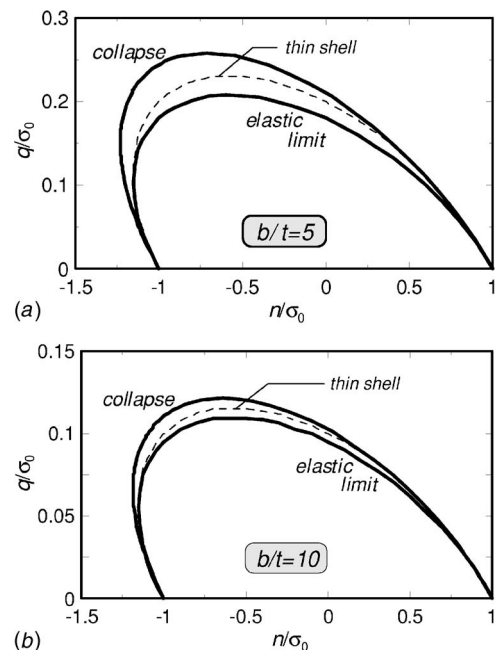
Fig. 7 Comparison of collapse levels, elastic limits, and thin-shell approximation

for the comparatively thin one. Already in the case of pure pressure, stress redistribution contributes to a significant amount (17% and 8%, respectively).

The limit loads predicted by the thin-shell assumption are also depicted in Fig. 7 (dashed lines). They are simpler to compute, but underestimate the actual collapse pressure by an amount not completely negligible for pure pressure (5% and 2.5% in the two cases) and of some significance in hydrostatic conditions, where the assumption predicts a collapse pressure 10% below the correct value for the thicker tube and about 5% in the second case.

It may be concluded that for $b/t \leqslant 10$, plastic resources are of importance and a design based on the elastic limit is strongly conservative. The thin-shell approximation appears adequate for moderately thick casing or pipelines, above all when external pressure is associated to axial tension, as often is the case. On the contrary, nuclear power plant applications, involving very thick tubes commonly operating in hydrostatic conditions, require a more precise assessment of the load-bearing capacity, and the approach developed in this paper may be useful under this respect.

5 Conclusions

This study proposes a procedure for the evaluation of the collapse load of cylindrical shells subject to pressure and axial force, accounting for possibly significant wall thickness. The result is obtained by using the kinematic theorem of limit analysis, which produces the exact result (instead of a mere upper bound) provided that the collapse mechanism is identified without ambiguity. A crucial role to this end is played by the assumption of *generalized plane strain*, imposing that axial strains are uniform without preventing possible elongation of the tube. The assumption is reasonable for long cylinders under axially symmetric loading, since any attempt at possibly nonuniform longitudinal strains is contrasted by the adjacent portions and received indirect corroboration by numerical analyses performed on tubes of moderate length, with no a priori enforcement of the constant axial strain condition [3]. For the load conditions examined, the same results as in this study were obtained.

Results are produced in parametric form, the parameter being the ratio among the radial and axial velocities in the motion associated to the collapse mechanism. Despite the unfriendly aspect of the equations, their numerical integration by means of standard software is easy (closed-form expressions for the integrals are available, but advantages with respect to numerical integration are questionable). In this way, families of limit interaction curves (each referring to different radius-to-thickness ratios) can be produced, which permit, for any of the load conditions considered, an easy assessment of the safety factor with respect to plastic collapse and may provide a useful alternative to computation. For specific, but important, situations the collapse limits are expressed by simple formulas, such as Eq. (22) for pure external pressure or Eqs. (21) for the hydrostatic condition.

Comparison to existing formulas, mostly based on simplifications that consider the tube as a thin shell, shows that the more accurate collapse loads obtained in this paper are systematically higher. For moderately thick tubes, such as those used for pipelines or casings, benefits are limited, even if not completely negligible. On the contrary, when tubes are really thick, as envisaged for steam generators of some fourth-generation nuclear power plant studies, advantages are significant and well worth exploiting.

The results obtained are fairly general in that they account for any combination of external and internal pressures and axial load. However, the situation considered of greater interest refers to tubes subject, besides axial forces, to external pressure acting alone or exceeding the internal one. In such instances, interaction with instability should be considered. This is known to reduce the load-bearing capacity of medium-thick shells by a non-negligible amount and it is expected to play some role also in tubes of significant thickness. The subject is presently under study. In any case, the theoretical collapse load is the reference value to which corrections accounting for the effects of instability are applied and its precise evaluation seems to also be important to this purpose.

References

[1] Carelli, M. D., 2003, "IRIS: A Global Approach to Nuclear Power Renaissance," Nuclear News, **46**(10), pp. 32–42.
[2] Haagsma, S. C., and Schaap, D., 1981, "Collapse Resistance of Submarine Lines Studied," Oil & Gas Journal, Feb. 2, pp. 86–91.
[3] Tamano, T., Mimaki, T., and Yanagimoto, S., 1985, "A New Empirical Formula for Collapse Resistance of Commercial Casing," Nippon Steel Technical Report No. 26, pp. 19–26.
[4] Yeh, M. K., and Kyriakides, S., 1986, "On the Collapse of Inelastic Thick-Walled Tubes under External Pressure," ASME J. Energy Resour. Technol., **108**(1), pp. 35–47.
[5] Tokimasa, K., and Tanaka, K., 1986, "FEM Analysis of the Collapse Strength of a Tube," ASME J. Pressure Vessel Technol., **108**(2), pp. 158–164.
[6] Yeh, M. K., and Kyriakides, S., 1988, "Collapse of Deepwater Pipelines," ASME J. Energy Resour. Technol., **110**(1), pp. 1–11.
[7] Assanelli, A. P., Toscano, R. G., and Dvorkin, E. N., 1998, "Analysis of the Collapse of Steel Tubes Under External Pressure," *Computational Mechanics: New Trends and Applications*, S. Idelshon et al., eds., CIMNE, Barcelona.
[8] Huang, X., and Mihsein, M., 2000, "Finite Element Prediction of the Ultimate Collapse Strength of Casings," Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci. **214**, pp. 1515–1527.
[9] Timoshenko, S., and Goodier, J. N., 1951, *Theory of Elasticity*, McGraw-Hill, New York.
[10] *2001 ASME Boiler & Pressure Vessel Code*, 2001, ASME, New York.
[11] Prager, W., and Hodge, Ph. G., Jr., 1951, *Theory of Perfectly Plastic Solids*, Wiley, New York.
[12] Mendelson, A., 1968, *Plasticity: Theory and Application*, MacMillan, New York.
[13] Corradi, L., Luzzi, L., and Trudi, F., 2004, "Plasticity-Instability Coupling in the Ultimate Behavior of Thick Tubes," Int. J. Struct. Stab. Dyn., **5**(1), pp. 1–18.
[14] Koiter, W. T., 1960, *General Theorems of Elastic Plastic Solids*, Progress in Solids Mechanics, **1**, North-Holland, Amsterdam, pp. 165–213.

On a Perturbation Method for the Analysis of Unsteady Belt-Drive Operation

Michael J. Leamy
Advanced Science and Automation,
Smithfield, VA 23430
e-mail: michael.leamy@ascience.com

A perturbation method is presented for use in analyzing unsteady belt-drive operation. The method relies on the important assumption that for operating states close to steady operation, the friction state (i.e., whether the belt is creeping or sticking at any location on the pulley) is similar to that of the well-known steady solution in which a lone stick arc precedes a lone slip arc (Johnson, K. L., 1985, Contact Mechanics, Cambridge U.P., London, Chap. 8; Smith, D. P., 1999, Tribol. Int., 31(8), pp. 465–477). This assumption, however, is not used to determine the friction force distribution, and, in fact, the friction forces in the stick zone are found to be nonzero, in direct contrast to the steady solution. The perturbation analysis is used to derive expressions for the span tensions, the pulley tension distributions, the contact forces between the belt and the pulleys, and the angular velocity of the driven pulleys. Validity criteria are developed which determine bounds on the operation state for which the assumed friction state is upheld. Verification of response quantities from the perturbation solution is accomplished through comparison to quantities predicted by an in-house dynamic finite element model and excellent agreement is found. Additionally, the finite element model is used to verify the key assumption that a lone slip arc precedes a lone stick arc. [DOI: 10.1115/1.1940660]

1 Introduction

Belt drives are widely used to transmit power between machine elements. Common applications include drives transmitting power from electric motors to rotational elements in home appliances such as washing machines, vacuum cleaners, and tape drives; from gas engines to cutting elements in lawn and garden equipment such as lawnmowers, rototillers, and snow blowers; and from the crankshaft pulley to accessory pulleys in automobiles and other transportation vehicles, where the accessories include alternators, air-conditioning compressors, and power-steering pumps. The life of the belt drive in all these applications depends critically on the tension magnitudes in the belt spans and the extent of belt creep on the pulley.

Even in a belt drive transmitting a constant torque between machine elements, the translating belt is subjected to cyclic tension variations as its tension transitions from a larger to a smaller tension on the driver pulley, and then from a smaller to a larger tension on each driven pulley, before returning again to the driver pulley. As a result, fatigue of the belt, and the subsequent permanent set and loss of compliance, is a large consideration in belt-drive design. Additionally, the belt is subjected to sliding wear as the belt creeps against the pulley during tension transitions. This wear can have a detrimental effect on the belt's friction characteristics as the belt surface deteriorates, and can lead to gross slip and noisy operation. These considerations motivate the need for a thorough understanding of belt-drive mechanics, and the need for belt-drive models which can accurately predict belt span tensions and belt creep.

The earliest studies of belt-drive mechanics include Leonard Euler's study [1] of a belt wrapped around a fixed pulley or cap-

stan, and Grashof's study [2] of the frictional mechanics of belt drives under steady operating conditions. A comprehensive review of studies on belt-drive mechanics after Grashof and up to 1981 is given by Fawcett [3]. The aforementioned studies of Euler and Grashof developed the classical creep theory of belt-drive operation. In this theory, a Coulomb law governs the belt-pulley frictional contact, and the belt is treated as a string which adheres to the pulley in an initial adhesion arc, and creeps against the pulley in a subsequent slip arc. Classical creep theory was reviewed by Johnson [4], and recently updated with new inertial effects by Bechtel et al. [5] and Rubin [6]. Smith [7] experimentally verified the existence and locations of the classically predicted stick and slip zones using a very thin, data tape cartridge. Other studies have considered the mechanics of the belt drive with belt shear effects, including Firbank [8] and Gerbert [9,10]. Gerbert [9,10] also included seating/unseating and radial compliance effects in his analysis. Townsend and Salisbury [11] derived the power loss expression and the efficiency limit of a belt drive assuming the validity of the classical creep theory.

Much recent emphasis of belt-drive studies has been on the dynamic response of automotive *serpentine* belt drives to crankshaft excitation. Serpentine belt drives include an automatic tensioner which attempts to take up belt slack in the drive system. These studies have considered both the rotational response of the pulleys and/or the transverse response of the axially moving belt, and have simplified the belt-pulley contact to linear stretching and viscous damping models. Barker [12] studied belt-drive tensions resulting from rapid engine acceleration, Hwang et al. [13] studied the periodic rotational response of the serpentine belt drive, and Beichman et al. [14–16] studied the coupled rotational and transverse response of a three-pulley prototypical serpentine belt drive. Leamy et al. [17,18] included a Coulomb dry friction damper to the tensioner arm element, and also studied the serpentine drive's rotational response. Kraver et al. [19] linearized the dry friction in the tensioner arm and developed a complex modal approach to analyze the drive's rotational response. Most recently, Kong and Parker [20,21] have included bending stiffness in their analysis of the belt spans and have found an alternate mechanism responsible for coupling rotational and transverse belt motions.

The two groups of studies reviewed above, namely belt-drive

mechanics studies and serpentine belt-drive dynamic response studies, have had little connection to each other due to the lack of dynamic excitation in the belt-drive mechanics studies, and the lack of true frictional belt-pulley modeling in the serpentine belt-drive studies. Leamy et al. [22–24] attempted to bridge this gap by studying simplified dynamic models for small [22] and large [23,24] rotational speeds. These studies considered individual pulleys only, and did not calculate the global response of the entire belt drive. Furthermore, the case of medium rotational speeds was not addressed.

Detailed modeling of the belt-pulley contact and the rotational and transverse response of a two-pulley, spring-supported, belt drive has recently been completed by Leamy and Wasfy [25,26], which does combine accurate belt-pulley contact mechanics with serpentine belt-drive system response [27]. In the latter studies, a dynamic finite element model of the belt drive was developed using truss or beam elements for the belt, rigid constraints for the pulleys, and a penalty formulation for modeling the belt-pulley contact. No restrictions were made on the steady/unsteady character of the rotational speeds or accessory torques, and the resulting model was shown to be general enough to capture coupled pulley rotational and span transverse response. Although accurate and effective, the finite element model has inherently two major disadvantages: (1) a large computational expense, which is particularly inconvenient for parametric and sensitivity studies, and (2) an inability to easily reveal underlying physical phenomena due to the availability of only time-history data.

As an alternative to the finite element model, this study presents a closed-form analysis of the frictional contact and global behavior of belt drives for the practically important case of *unsteady* belt-drive rotational response. Use of a perturbation method necessitates focusing on unsteady belt drives operating in proximity to a steady state, although validity criteria reveal that due to the large tension differential required to initiate belt sliding, the valid range of operation states is large enough to include typical automotive applications. The lone stick and sliding regions, their tension distributions, and their locations on the pulley are investigated using closed-form expressions ideally suited for parametric and sensitivity studies. The solution procedure is applied to an example two-pulley drive and global response quantities are calculated and compared to the finite element model in order to verify the analytical solutions.

2 Closed-Form Analysis: Governing Equations

The analysis of the unsteady operation of a belt drive is presented herein using a perturbation approach in which the relative belt motion (as compared to the pulley motion) is assumed to be similar to that of a steadily rotating belt-drive—i.e., where sliding is present, the two share the same direction of sliding, but do not necessarily share the same extent of sliding and/or sliding magnitude. During steady operation (i.e., constant applied torque and angular velocities), the belt slips in a single direction in a lone slip zone at the trailing edge of the pulley [4,7], as shown in Fig. 1. A stick zone occupies the remaining portion of the belt-pulley contact zone. The analytical solutions presented herein for unsteady operation are developed based on an assumption of similar contact behavior—single slip zone, single stick zone—which forms the basis for an asymptotic perturbation procedure. Following development of the perturbation solution, criteria are presented in Sec. 4 for assessing the appropriateness of this assumption based on a candidate drive's parameter space.

2.1 Element Conservation Equations. The equations governing the motion of a belt element in contact with any pulley (driven or driver) are developed using a fixed element control volume evaluated using conservation principles, resulting in a Eulerian description of the belt kinetics. Belt strain and stress, as well as material constitutive modeling, are developed using a Lagrangian description of the belt. Following development of the
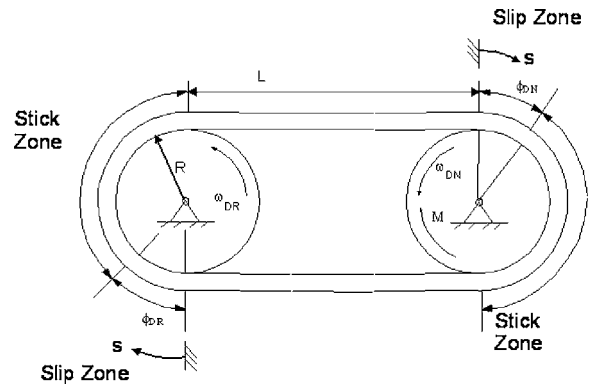
Fig. 1   Example two-pulley belt drive

governing equations, boundary conditions are specified based on the quasi-static assumption of a single span tension. In this way the spans serve to connect the belt element domains. Although a two-pulley drive is explicitly considered, generalization to a drive consisting of any number of pulleys is straightforward.

Considering a belt element in Fig. 2 for any pulley, the mass flow rate G of material is given by

$$G(s,t) = \rho vA, \tag{1}$$

where $v(s,t)$ denotes the belt velocity, $\rho(s,t)$ the belt density, and $A(s,t)$ the belt cross-sectional area at location s and time t. Note from Fig. 1 that each pulley is considered to have an arc measure s originating at the trailing edge of the pulley. Conservation of linear momentum yields tension distribution equations along the contact arc,

$$\sum \mathbf{F}_{C.V.} = \frac{\partial}{\partial t}(m\mathbf{v})_{C.V.} - (m\dot{\mathbf{v}})_{\text{in}} + (m\dot{\mathbf{v}})_{\text{out}}, \tag{2}$$

where the tangential component of (2) yields

$$\frac{\partial T(s,t)}{\partial s} + f(s,t) = \frac{\partial G(s,t)}{\partial t} + G\frac{\partial v(s,t)}{\partial s} + v\frac{\partial G(s,t)}{\partial s}, \tag{3}$$

and where f denotes the friction force per unit length between the belt and the pulley. The normal component of (2) yields

$$n(s,t) = \frac{T(s,t) - G(s,t)v(s,t)}{R}, \tag{4}$$

where n denotes the normal force per unit length. In the following, (3) and (4) are specialized to driver and driven pulleys with introduction of $G^{DR}(s,t), T^{DR}(s,t), v^{DR}(s,t)$ and $G^{DN}(s,t)$, $T^{DN}(s,t), v^{DN}(s,t)$ representing mass flow, tension, and velocity along the driver (^{DR}) and driven (^{DN}) pulleys, respectively.
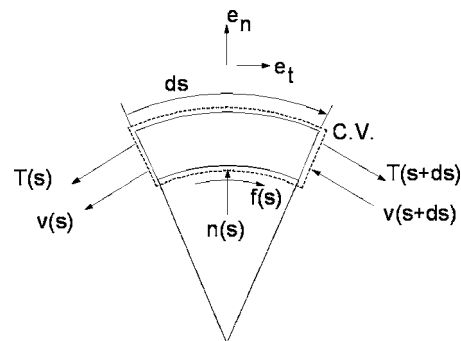
Fig. 2   Belt element used to develop the belt-drive governing equations

Belt spans connect the driver and driven belt domains. For excitation periods $T = 2\pi/\omega$ much longer than the time for longitudinal waves to travel down the belt span and reflect back (i.e., $\omega \ll (\pi\sqrt{E/\rho})/L$), it can be accurately assumed that a uniform strain exists throughout the span length at any time t. With this quasi-static assumption, the following boundary conditions can be used to relate the belt tensions at the inlet and exit of the driven and driver pulleys,

$$T_H(t) = (T^{DR}(s,t)|_{s=R\phi_{DR}} = T^{DN}(s,t)|_{s=0}), \tag{5}$$

$$T_L(t) = (T^{DR}(s,t)|_{s=0} = T^{DN}(s,t)|_{s=R\phi_{DN}}), \tag{6}$$

where $T_H(t)$ and $T_L(t)$ denote tension in the high-tension and low-tension spans and $\phi_{DR}(t)$, $\phi_{DN}(t)$ denote slip arc metrics along the driven and driver pulleys—see Fig. 1. Note: the tension boundary conditions as stated are only approximate at $s = R\phi_{DR}$ and $s = R\phi_{DN}$ since, due to the stick zone, the tensions at these points are actually equal to high and low tensions, respectively, at a previous time. This previous time is shifted from time t by an amount equal to the time it takes the belt to traverse the stick zone and reach $s = R\phi_{DR}$ or $s = R\phi_{DN}$. Similar boundary conditions relate the driven and driver belt speeds at the inlets and exits, as well as relating belt speed to pulley angular velocities. Specifically, since a stick zone is assumed to exist at the inlet, the belt speed at this location is related to the pulley speed by

$$v_H^{DR} = v^{DR}(s,t)|_{s=R\phi_{DR}} = R\omega_{DR}, \tag{7a}$$

$$v_L^{DN} = v^{DN}(s,t)|_{s=R\phi_{DN}} = R\omega_{DN}. \tag{7b}$$

At the exits the velocities must be determined and are given as

$$v_H^{DN} = v^{DN}(s,t)|_{s=0}, \tag{8a}$$

$$v_L^{DR} = v^{DR}(s,t)|_{s=0}, \tag{8b}$$

where v_H^{DR}, v_H^{DN}, v_L^{DR}, and v_L^{DN} denote the belt speeds entering and exiting the high-tension and low-tension spans. Note that unlike the steady solution, the high-tension belt speed exiting the driven pulley v_H^{DN} is *not* equal to the high-tension belt speed entering the driver pulley v_H^{DR}—similarly for the low-tension belt speeds.

2.2 Friction Law. A Coulomb friction law is adopted in this study to describe the contact friction between the belt and the pulleys,

$$f(s) = \begin{cases} -\mu n(s), & \text{slip zone(driver)}, \\ \mu n(s), & \text{slip zone(driven)}, \end{cases} \tag{9}$$

where μ denotes the sliding coefficient of friction. The friction force in the sick zones is not known a priori—see Sec. 4.2 for its calculation. Use of (9) in (3) and (4) results in the following equations governing the tension distributions along the driver and driven pulleys:

$$\frac{\partial T^{DR}(s,t)}{\partial s} - \mu \frac{T^{DR}(s,t) - G^{DR}v^{DR}(s,t)}{R} = \frac{\partial G^{DR}(s,t)}{\partial t} + G^{DR}\frac{\partial v^{DR}(s,t)}{\partial s}$$

$$+ v^{DR}\frac{\partial G^{DR}(s,t)}{\partial s}, \tag{10}$$

$$\frac{\partial T^{DN}(s,t)}{\partial s} + \mu \frac{T^{DN}(s,t) - G^{DN}v^{DN}(s,t)}{R} = \frac{\partial G^{DN}(s,t)}{\partial t}$$

$$+ G^{DN}\frac{\partial v^{DN}(s,t)}{\partial s} + v^{DN}\frac{\partial G^{DN}(s,t)}{\partial s}. \tag{11}$$

**2.3 Stress-Strain Relations.** Using a Lagrangian description, a differential belt element with undeformed length ds_{ref} has deformed length,
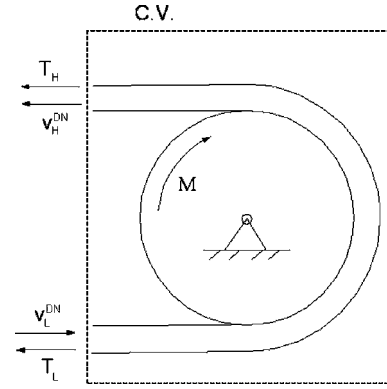
Fig. 3 Control volume used for driven pulley conservation of angular momentum

$$ds = (1 + \lambda T(s))ds_{ref}, \tag{12}$$

and deformed area (due to Poisson effects)

$$A(s) = (1 - \nu\lambda T(s))^2 A_{ref}, \tag{13}$$

where $\lambda = 1/EA$ is a measure of belt compliance. Finally, the density in the deformed state follows as

$$\rho(s) \equiv \frac{m_{ref}}{V} = \frac{\rho_{ref}A_{ref}ds_{ref}}{A(s)ds} = \left(\frac{1}{(1 - \nu\lambda T(s))^2(1 + \lambda T(s))}\right)\rho_{ref}. \tag{14}$$

A more convenient expression is $\rho(s)A(s)$,

$$\rho(s)A(s) = \left(\frac{\rho_{ref}A_{ref}}{1 + \lambda T(s,t)}\right). \tag{15}$$

Note, the definition of G can be updated to the following expression:

$$G(s,t) = \frac{\rho_{ref}A_{ref}v(s,t)}{1 + \lambda T(s,t)}. \tag{16}$$

2.4 Global Drive Relations. Global drive relations are needed to determine global quantities such as the span tensions and speeds. Application of conservation of angular momentum to the fixed control volume shown in Fig. 3,

$$\mathbf{r} \times \mathbf{F} + \mathbf{M} = \frac{\partial}{\partial t}\int_{C.V.}(\mathbf{r} \times \mathbf{v})\rho dV + \int_{C.S.}(\mathbf{r} \times \mathbf{v})\rho\mathbf{v} \cdot d\mathbf{A}, \tag{17}$$

yields

$$R(T_L - T_H) + M = \frac{\partial}{\partial t}\int_{Belt} -RG^{DN}(s,t)ds - I\dot{\omega}_{DN}$$

$$+ R(G_L^{DN}v_L^{DN} - G_H^{DN}v_H^{DN}), \tag{18}$$

where I denotes the pulley's mass moment of inertia about its axis of rotation, M is the moment resisting the driven pulley motion, and where subscripts L and H denote quantities evaluated for the low- and high-tension spans. A second relation derives from belt length compatibility which equates the unstretched belt length calculated from the geometry of the deformed (or operating) configuration to the known unstretched belt length L_{ref}^{belt},

$$\oint \frac{ds}{1 + \in} = L_{ref}^{belt}, \tag{19}$$

where the integral is taken over the entire belt length, ds is an element of length in the deformed configuration, and strain is

related to tension by $\in = \lambda T$ where $\lambda = 1/EA_{\text{ref}}$ is a measure of compliance. As will be discussed in Sec. 3, exact satisfaction of (19) should not be expected due to the quasi-static assumption invoked for the span tensions, but near satisfaction is possible as discussed in Sec. 5.

Time-dependent consideration of each span is also required for unsteady operation. The rate of change of the span's strain determines the rate of change of each span tension. Considering the top span, the strain rate can be expressed as

$$\dot{\in}_H = \frac{d}{dt}\left(\frac{L - L_0}{L_0}\right) = -\frac{(1 + \lambda T_H)}{L_0}\frac{dL_0}{dt}, \quad (20)$$

where L denotes the span's current length and $L_0 = L/(1 + \lambda T_H)$ denotes an equivalent unstretched length. Accounting for changes in L_0 due to belt entering the span $(\Delta L_0)_{\text{in}} = v_H^{DN}/(1 + \lambda T_H) \cdot \Delta t$ and leaving the span $(\Delta L_0)_{\text{out}} = R\omega^{DR}/(1 + \lambda T_H) \cdot \Delta t$, the desired expression for the time rate of change of the span tension can be stated as

$$\frac{dT_H}{dt} = EA\dot{\in}_H = \frac{(1 + \lambda T_H)}{\lambda L_0}(R\omega_{DR} - v_H^{DN}), \quad (21)$$

while a similar expression can be stated for the low-tension span,

$$\frac{dT_L}{dt} = EA\dot{\in}_L = \frac{(1 + \lambda T_L)}{\lambda L_0}(R\omega_{DN} - v_L^{DR}). \quad (22)$$

The global drive relations (18)–(22) apply to the two-pulley drive explicitly considered here, but a drive consisting of any number of pulleys is easily analyzed by including an additional angular momentum relation (18) for each driven pulley, and by including a time-dependent span relation similar to (21) for each additional span.

2.5 Perturbation Quantities. A natural small quantity arises from the prescribed expression for the driver pulley angular velocity,

$$\omega_{DR}(t) = \omega_{DR}^0 + \varepsilon\omega_{DR}^1(t) = \omega_{\text{steady}} + \omega_{osc}\sin\omega t$$
$$= \omega_{\text{steady}}(1 + \varepsilon\sin\omega t), \quad (23)$$

where for small oscillations about a steady speed, the small parameter ε can be defined as $\varepsilon = \omega_{osc}/\omega_{\text{steady}}$. Other quantities in the problem can then be expanded in a classical perturbation approach using this parameter. Following this approach, the driven pulley angular velocity is assumed to have the form

$$\omega_{DN} = \omega_{DN}^0(t) + \varepsilon\omega_{DN}^1(t) + \varepsilon^2\omega_{DN}^2(t) + O(\varepsilon^3). \quad (24)$$

Expanding all remaining response quantities results in the following perturbed expressions for the mass flow rate, tensions, velocities, span tensions, and slip arc magnitudes,

$$G^{DR} = G_0^{DR} + \varepsilon G_1^{DR}(t) + O(\varepsilon^2), \quad G^{DN} = G_0^{DN} + \varepsilon G_1^{DN}(t) + O(\varepsilon^2), \quad (25a)$$

$$T^{DR} = T_0^{DR}(s) + \varepsilon T_1^{DR}(s,t) + O(\varepsilon^2), \quad T^{DN} = T_0^{DN}(s) + \varepsilon T_1^{DN}(s,t) + O(\varepsilon^2), \quad (25b)$$

$$v^{DR} = v_0^{DR}(s) + \varepsilon v_1^{DR}(s,t) + O(\varepsilon^2), \quad v^{DN} = v_0^{DN}(s) + \varepsilon v_1^{DN}(s,t) + O(\varepsilon^2), \quad (25c)$$

$$T_L = T_L^0 + \varepsilon T_L^1(t) + O(\varepsilon^2), \quad T_H = T_H^0 + \varepsilon T_H^1(t) + O(\varepsilon^2), \quad (25d)$$

$$v_L^{DN} = v_{L,0}^{DN} + \varepsilon v_{L,1}^{DN}(t) + O(\varepsilon^2), \quad v_H^{DN} = v_{H,0}^{DN} + \varepsilon v_{H,1}^{DN}(t) + O(\varepsilon^2), \quad (25e)$$

$$v_L^{DR} = v_{L,0}^{DR} + \varepsilon v_{L,1}^{DR}(t) + O(\varepsilon^2), \quad v_H^{DR} = v_{H,0}^{DR} + \varepsilon v_{H,1}^{DR}(t) + O(\varepsilon^2), \quad (25f)$$

$$\phi^{DR} = \phi_0^{DR} + \varepsilon\phi_1^{DR}(t) + O(\varepsilon^2), \quad \phi^{DN} = \phi_0^{DN} + \varepsilon\phi_1^{DN}(t) + O(\varepsilon^2), \quad (25g)$$

where it is expected that $O(\varepsilon^0)$ quantities will recover the steady solution. Note that the spatial variation of the mass flow rate G on the pulley contact arc has been neglected in the assumed expansion (25a) in order to simplify the analysis of the tension distributions. This assumption is instead a fact for the steady analysis [5], and thus for the $O(\varepsilon^0)$ analysis presented herein. Tension terms arising due to the spatial variation of G_1 are expected to be negligible for stiff belts, where stretching is minimal—evaluating a mass conservation statement shows the spatial change in G_1 is inversely proportional to the belt stiffness. Analysis of the belt velocity is the primary concern of Sec. 4 and is computed directly from the tension distributions without a need to restrict the functional form of G.

3 Analytical Solution Approach

The perturbation quantities are now introduced into the equations governing the tension distributions and are then separated into orders of ε.

3.1 Ordered Equations. Substituting the expansions into the definition of the mass flow rate $G(s,t)$, (16) yields

$$\varepsilon^0: \quad G_0^{DR} = \frac{\rho_{\text{ref}}A_{\text{ref}}v_0^{DR}(s)}{1 + \lambda T_0^{DR}(s)}, \quad G_0^{DN} = \frac{\rho_{ref}A_{ref}v_0^{DN}(s)}{1 + \lambda T_0^{DN}(s)}, \quad (26a)$$

$$\varepsilon^1: \quad G_1^{DR}(t) = \frac{\rho_{\text{ref}}A_{\text{ref}}v_1^{DR}(s,t)}{1 + \lambda T_0^{DR}(s)} - \frac{G_0^{DR}\lambda T_1^{DR}(s,t)}{1 + \lambda T_0^{DR}(s)},$$

$$G_1^{DN}(t) = \frac{\rho_{\text{ref}}A_{\text{ref}}v_1^{DN}(s,t)}{1 + \lambda T_0^{DN}(s)} - \frac{G_0^{DN}\lambda T_1^{DN}(s,t)}{1 + \lambda T_0^{DN}(s)}, \quad (26b)$$

and the tension distribution equations (10) and (11),

$$\varepsilon^0: \quad \frac{\partial T_0^{DR}(s)}{\partial s} - \mu\frac{T_0^{DR} - G_0^{DR}v_0^{DR}}{R} = \frac{\partial}{\partial s}(G_0^{DR}v_0^{DR}(s)), \quad (27a)$$

$$\frac{\partial T_0^{DN}(s)}{\partial s} + \mu\frac{T_0^{DN} - G_0^{DN}v_0^{DN}}{R} = \frac{\partial}{\partial s}(G_0^{DN}v_0^{DN}(s)), \quad (27b)$$

$$\varepsilon^1: \quad \frac{\partial T_1^{DR}(s,t)}{\partial s} - \mu\frac{T_1^{DR} - G_0^{DR}v_1^{DR} - G_1^{DR}v_0^{DR}}{R} = G_0^{DR}\frac{\partial v_1^{DR}}{\partial s}$$

$$+ G_1^{DR}\frac{\partial v_0^{DR}}{\partial s} + \frac{\partial G_1^{DR}}{\partial t}, \quad (28a)$$

$$\frac{\partial T_1^{DN}(s,t)}{\partial s} + \mu\frac{T_1^{DN} - G_0^{DN}v_1^{DN} - G_1^{DN}v_0^{DN}}{R} = G_0^{DN}\frac{\partial v_1^{DN}}{\partial s} + G_1^{DN}\frac{\partial v_0^{DN}}{\partial s}$$

$$+ \frac{\partial G_1^{DN}}{\partial t}. \quad (28b)$$

The tension and velocity boundary conditions (5)–(8) separate into orders of ε readily at $s=0$,

$$\varepsilon^0: \quad T_H^0 = T_0^{DN}(s)|_{s=0}, \quad T_L^0 = T_0^{DR}(s)|_{s=0}, \quad (29a)$$

$$v_{H,0}^{DN} = v_0^{DN}(s)|_{s=0}, \quad v_{L,0}^{DR} = v_0^{DR}(s)|_{s=0}, \quad (29b)$$

$$\varepsilon^1: \quad T_H^1(t) = T_1^{DN}(s,t)|_{s=0}, \quad T_L^1(t) = T_1^{DR}(s,t)|_{s=0}, \quad (30a)$$

$$v_{H,1}^{DN}(t) = v_1^{DN}(s,t)|_{s=0}, \quad v_{L,1}^{DR}(t) = v_1^{DR}(s,t)|_{s=0}, \quad (30b)$$

whereas the presence of ε in both the expansions for ϕ_{DR}, ϕ_{DN} and in the expansions for tensions and velocities do not yet allow $O(\varepsilon)$ separation of the $s=R\phi_{DR}$ and $s=R\phi_{DN}$ boundary conditions. Conservation of driven pulley angular momentum (18) yields

$$\varepsilon^0: \quad R(T_L^0 - T_H^0) + M = RG_0^{DN}(\mathrm{v}_{L,0}^{DN} - \mathrm{v}_{H,0}^{DN}), \qquad (31a)$$

$$\varepsilon^1: \quad R(T_L^1(t) - T_H^1(t)) + I\frac{\partial \omega_{DN}^1}{\partial t} = RG_0^{DN}(\mathrm{v}_{L,1}^{DN}(t) - \mathrm{v}_{H,1}^{DN}(t))$$
$$+ RG_1^{DN}(\mathrm{v}_{L,0}^{DN} - \mathrm{v}_{H,0}^{DN}), \quad (31b)$$

where the belt acceleration term at $O(\varepsilon^1)$ has been assumed negligible compared to the pulley acceleration term. Exact compatibility (19) separates into ε orders as

$$\varepsilon^0: \quad \oint \frac{ds}{1 + \lambda T_0(s)} = L_{ref}^{belt}, \qquad (32a)$$

$$\varepsilon^1: \quad \oint \frac{\lambda T_1(s,t)ds}{(1 + \lambda T_0(s))^2} = 0. \qquad (32b)$$

Finally, the time-dependent span relations separate into

$$\varepsilon^0: \quad 0 = (1 + \lambda T_H^0)(R\omega_{DR}^0 - \mathrm{v}_{H,0}^{DN})/\lambda L_0, \quad 0 = (1 + \lambda T_L^0)(R\omega_{DN}^0$$
$$- \mathrm{v}_{L,0}^{DR})/\lambda L_0, \qquad (33a)$$

$$\varepsilon^1: \quad \frac{dT_H^1}{dt} = \frac{1 + \lambda T_H^0}{\lambda L_0}(R\omega_{DR}^1(t) - \mathrm{v}_{H,1}^{DN}) + \frac{\lambda T_H^1(t)}{\lambda L_0}(R\omega_{DR}^0 - \mathrm{v}_{H,0}^{DN}),$$
$$(33b)$$

$$\frac{dT_L^1}{dt} = \frac{1 + \lambda T_L^0}{\lambda L_0}(R\omega_{DN}^1(t) - \mathrm{v}_{L,1}^{DR}) + \frac{\lambda T_L^1(t)}{\lambda L_0}(R\omega_{DN}^0 - \mathrm{v}_{L,0}^{DR}).$$
$$(33c)$$

3.2 ε^0 Solution. The solution for the $O(\varepsilon^0)$ response quantities now follows sequentially from the $O(\varepsilon^0)$ equations developed above. An expression for v_0 can be obtained in terms of the mass flow rate G_0 and the tension T_0 from (26):

$$\mathrm{v}_0(s) = \frac{G_0}{\rho_{ref}A_{ref}}(1 + \lambda T_0(s)). \qquad (34)$$

Substitution of (34) into (27) yields differential equations governing the spatial change of $T_0^{DR}(s)$, $T_0^{DN}(s)$ in the slip zones. Solution of these equations and application of the $s=0$ boundary conditions (29a) yields the tension distributions,

$$T_0^{DR}(s) = \left(T_L^0 - \frac{G_0^{DR^2}}{\rho_{ref}A_{ref} - \lambda G_0^{DR^2}}\right)e^{\mu s/R} + \frac{G_0^{DR^2}}{\rho_{ref}A_{ref} - \lambda G_0^{DR^2}},$$
$$(35a)$$

$$T_0^{DN}(s) = \left(T_H^0 - \frac{G_0^{DN^2}}{\rho_{ref}A_{ref} - \lambda G_0^{DN^2}}\right)e^{-\mu s/R} + \frac{G_0^{DN^2}}{\rho_{ref}A_{ref} - \lambda G_0^{DN^2}},$$
$$(35b)$$

which are those obtained in the steady case [5]. With functional forms now known for $T_0^{DR}(s)$ and $T_0^{DN}(s)$, the boundary conditions at $s=R\phi_{DR}$ and $s=R\phi_{DN}$ can be stated at $O(\varepsilon^0)$,

$$\varepsilon^0: \quad T_H^0 = T_0^{DR}(s)|_{s=R\phi_{DR}}, \quad T_L^0 = T_0^{DN}(s)|_{s=R\phi_{DN}}, \qquad (36)$$

$$\mathrm{v}_{H,0}^{DR} = \mathrm{v}_0^{DR}(s)|_{s=R\phi_{DR}} = R\omega_{DR}^0, \quad \mathrm{v}_{L,0}^{DN} = \mathrm{v}_0^{DN}(s)|_{s=R\phi_{DN}} = R\omega_{DN}^0.$$
$$(37)$$

The two boundary conditions (37) yield the mass flow rates,

$$G_0^{DR} = \frac{R\rho_{ref}A_{ref}}{1 + \lambda T_H^0}\omega_{DR}^0, \quad G_0^{DN} = \frac{R\rho_{ref}A_{ref}}{1 + \lambda T_L^0}\omega_{DN}^0, \qquad (38)$$

while substitution of (34) and (38) into (33a) yields the fact that the $O(\varepsilon^0)$ mass flow rates are equal as in the steady solution, i.e., $G_0^{DR}=G_0^{DN}=G_0$, and yields the driven pulley's angular velocity,

$$\omega_{DN}^0 = \frac{1 + \lambda T_L^0}{1 + \lambda T_H^0}\omega_{DR}^0. \qquad (39)$$

The slip arc metrics are given by solution of (36),

$$\phi_0^{DR} = \phi_0^{DN} = \frac{1}{\mu}\ln\left(\frac{T_H^0 - G_0^2/(\rho_{ref}A_{ref} - \lambda G_0^2)}{T_L^0 - G_0^2/(\rho_{ref}A_{ref} - \lambda G_0^2)}\right). \qquad (40)$$

The only remaining quantities to find are then T_L^0 and T_H^0 using the global drive relations. This is accomplished by solving (31a) for T_L^0 and substituting this result and (35), (38), and (40) into (32a). Note that the quasi-static tension assumption invoked earlier is exact for the $O(\varepsilon^0)$ tensions since they are independent of time, and so exact satisfaction of (32a) is possible. Finally, a root solver can be used to solve the remaining equation for T_H^0 once the belt-drive parameter space has been defined. This procedure is completed for an example drive in Sec. 5.

3.3 ε^1 Solution. Following a similar procedure to that developed for $O(\varepsilon^0)$, the $O(\varepsilon^1)$ equations are solved sequentially. Rearranging (26b) for v_1,

$$\mathrm{v}_1(s,t) = \frac{(1 + \lambda T_0(s))G_1 + \lambda T_1(s,t)G_0}{\rho_{ref}A_{ref}}, \qquad (41)$$

the driver and driven tension equations (28) can be expressed as unknown functions of $T_1^{DR}(s,t)$ and $T_1^{DN}(s,t)$. Upon substitution of (41) into (28), solutions are found, and the $O(\varepsilon^1)$ $s=0$ boundary conditions (30a) are applied. The resulting tension solutions are given as

$$T_1^{DR}(s,t) = \left(T_L^1(t) - \frac{\rho_{ref}A_{ref}(2G_0^{DR}G_1^{DR} - (R/\mu)(\rho_{ref}^2 A_{ref}^2 - \lambda G_0^{DR^2})(\partial G_1^{DR}/\partial t))}{(\rho_{ref}A_{ref} - \lambda G_0^{DR^2})^2}\right)e^{(\mu/R)s}$$
$$+ \frac{\rho_{ref}A_{ref}(2G_0^{DR}G_1^{DR} - (R/\mu)(\rho_{ref}^2 A_{ref}^2 - \lambda G_0^{DR^2})(\partial G_1^{DR}/\partial t))}{(\rho_{ref}A_{ref} - \lambda G_0^{DR^2})^2}, \qquad (42a)$$

$$T_1^{DN}(s,t) = \left(T_H^1(t) - \frac{\rho_{ref}A_{ref}(2G_0^{DN}G_1^{DN} + (R/\mu)(\rho_{ref}^2 A_{ref}^2 - \lambda G_0^{DN^2})(\partial G_1^{DN}/\partial t))}{(\rho_{ref}A_{ref} - \lambda G_0^{DN^2})^2}\right)e^{-(\mu/R)s}$$
$$+ \frac{\rho_{ref}A_{ref}(2G_0^{DN}G_1^{DN} + (R/\mu)(\rho_{ref}^2 A_{ref}^2 - \lambda G_0^{DN^2})(\partial G_1^{DN}/\partial t))}{(\rho_{ref}A_{ref} - \lambda G_0^{DN^2})^2}. \qquad (42b)$$

With the functional forms for the $O(\varepsilon^0)$, $O(\varepsilon^1)$ tensions and the $O(\varepsilon^0)$ slip arc metrics known, the boundary conditions at $s=R\phi_{DR}$ and $s=R\phi_{DN}$ can be stated. Substitution of $T_0^{DR}(s)$, $T_0^{DN}(s)$, $T_1^{DR}(s,t_1)$, $T_1^{DN}(s,t)$, ϕ_0^{DR}, and ϕ_0^{DN} into (5) and (6) yields

$$\mu\phi_1^{DR}\left(T_L^0 - \frac{G_0^{DR^2}}{\rho_{\text{ref}}A_{\text{ref}} - \lambda G_0^{DR^2}}\right)e^{\mu\phi_0^{DR}} + T_1^{DR}(s=R\phi_0^{DR},t) = T_H^1,$$

(43a)

$$\mu\phi_1^{DN}\left(T_H^0 - \frac{G_0^{DN^2}}{\rho_{\text{ref}}A_{\text{ref}} - \lambda G_0^{DN^2}}\right)e^{-\mu\phi_0^{DN}} + T_1^{DN}(s=R\phi_0^{DN},t) = T_L^1,$$

(43b)

where all dependence on ϕ_1^{DR} and ϕ_1^{DN} in (43) appears outside of the exponential, and outside of the expressions for T_1. This allows for ready evaluation of the $O(\varepsilon^1)$ slip arc metrics,

$$\phi_1^{DR} = \frac{T_H^1 - T_1^{DR}(s=R\phi_0^{DR},t)}{\mu\left(T_L^0 - G_0^{DR^2}/(\rho_{\text{ref}}A_{\text{ref}} - \lambda G_0^{DR^2})\right)e^{\mu\phi_0^{DR}}},$$

$$\phi_1^{DN} = \frac{-T_L^1 + T_1^{DN}(s=R\phi_0^{DN},t)}{\mu\left(T_H^0 - G_0^{DN^2}/(\rho_{\text{ref}}A_{\text{ref}} - \lambda G_0^{DN^2})\right)e^{-\mu\phi_0^{DN}}}. \quad (44)$$

A similar procedure for the velocity boundary conditions (7) can be pursued, but yields unnecessarily lengthy expressions for $G_1^{DR}(t)$, $G_1^{DN}(t)$. Since the quantities of interest in this section of the analysis are the tensions, a simpler, approximate approach is taken to determine the expressions for $G_1^{DR}(t)$ and $G_1^{DN}(t)$ to be used in the tension distributions. Decoupling ordered belt velocities at $s=R\phi_{DR}$ and $s=R\phi_{DN}$ in a manner similar to that at $s=0$, (7) can be replaced with the approximate expressions

$$\text{v}^{DR}(s,t)\big|_{s=R\phi_{DR}} \cong \text{v}_0^{DR}(s)\big|_{s=R\phi_{DR}^0} + \varepsilon\text{v}_1^{DR}(s,t)\big|_{s=R\phi_{DR}^0}$$

$$= R\omega_{DR}^0 + \varepsilon R\omega_{DR}^1(t), \quad (45a)$$

$$\text{v}^{DN}(s,t)\big|_{s=R\phi_{DN}} \cong \text{v}_0^{DN}(s)\big|_{s=R\phi_{DN}^0} + \varepsilon\text{v}_1^{DN}(s,t)\big|_{s=R\phi_{DN}^0}$$

$$= R\omega_{DN}^0 + \varepsilon R\omega_{DN}^1(t). \quad (45b)$$

Expressions for the mass flow rates at $O(\varepsilon^1)$ then follow from (41) evaluated at $s=R\phi_{DR}$ and $s=R\phi_{DN}$ together with (45),

$$G_1^{DR}(t) = \frac{R\rho_{\text{ref}}A_{\text{ref}}\omega_{DR}^1(t) - \lambda G_0^{DR}T_H^1(t)}{1 + \lambda T_H^0},$$

$$G_1^{DN}(t) = \frac{R\rho_{\text{ref}}A_{\text{ref}}\omega_{DN}^1(t) - \lambda G_0^{DN}T_L^1(t)}{1 + \lambda T_L^0}. \quad (46)$$

Unlike in the $O(\varepsilon^0)$ analysis, three quantities remain to be determined: $T_L^1(t)$, $T_H^1(t)$, and $\omega_{DN}^1(t)$. $O(\varepsilon^1)$ conservation of angular momentum (31b) provides one relationship, while (33b) and (33c) provide the other two necessary relations. Together, these equa-

tions comprise three nonhomogenous, constant coefficient, linear, ordinary differential equations for the $O(\varepsilon^1)$ span tensions $T_L^1(t)$, $T_H^1(t)$ and the $O(\varepsilon^1)$ driven pulley angular velocity $\omega_{DN}^1(t)$, which with use of (23) can be placed in the form

$$\mathbf{A}\dot{\mathbf{x}} + \mathbf{B}\mathbf{x} = \mathbf{f}\cdot\sin(\omega t), \quad \mathbf{x}(t) = \begin{Bmatrix} T_H^1(t) \\ T_L^1(t) \\ \omega_{DN}^1(t) \end{Bmatrix}, \quad (47)$$

where $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{f}$ are given in the Appendix. Following definition of the belt-drive parameter space, particular solutions to (47) are found numerically with the resulting expressions being expressed as

$$\mathbf{x}(t) = \mathbf{c}\sin(\omega t) + \mathbf{d}\cos(\omega t), \quad (48)$$

where

$$\mathbf{d} = [-\omega\mathbf{A} - \mathbf{B}\cdot[\omega\mathbf{A}]^{-1}\cdot\mathbf{B}]^{-1}\cdot\mathbf{f}, \quad \mathbf{c} = -[\omega\mathbf{A}]^{-1}\cdot\mathbf{B}\cdot\mathbf{d}. \quad (49)$$

4 Validity Criteria

In this section, a criterion governing the validity of the solution approach is developed by requiring that in the slip zone, the belt velocity is always greater than the pulley surface velocity for the driven pulleys, and always less than the pulley surface velocity for the driver pulley. A second criterion is developed by requiring the stick zone friction forces to be less than the maximum available friction force. Together, the two criteria form the required conditions on the solutions for the valid existence of a lone slip and stick zone.

First, the belt velocity in the pulley slip is accurately analyzed using the resulting tension distributions of Sec. 3. Recall that an approximate mass flow rate at $O(\varepsilon^1)$, and thus belt velocity at $O(\varepsilon^1)$, was used in determining the tension distribution expressions. While this is expected to have very little effect on the tension distributions calculated due to the small effect of inertial tensioning, it may have a significant effect on the calculation of the difference between the belt's angular velocity and the pulley's angular velocity, this difference being an $O(\varepsilon^1)$ quantity.

4.1 Slip Zone Criterion. In what follows, the tensions (Sec. 3) are used to determine the belt velocity in the driven pulley slip zone. A reconstitution procedure is first performed in which all response quantities are returned to their unperturbed values. Reconstituting the driven pulley slip-zone tension using (25), (35), and (42), the tension takes the form

$$T^{DN}(s,t) = T_H^e(t)e^{-(\mu/R)s} + T_G^e(t), \quad (50)$$

where

$$T_L^e(t) = \left(T_L^0 - \frac{G_0^{DN^2}}{\rho_{\text{ref}}A_{\text{ref}} - \lambda G_0^{DN^2}}\right) + \varepsilon\left(T_L^1(t) - \frac{\rho_{\text{ref}}A_{\text{ref}}(2G_0^{DN}G_1^{DN} + (R/\mu)(\rho_{\text{ref}}^2 A_{\text{ref}}^2 - \lambda G_0^{DN^2})(\partial G_1^{DN}/\partial t))}{(\rho_{\text{ref}}A_{\text{ref}} - \lambda G_0^{DN^2})^2}\right),$$

$$T_H^e(t) = \left(T_H^0 - \frac{G_0^{DN^2}}{\rho_{\text{ref}}A_{\text{ref}} - \lambda G_0^{DN^2}}\right) + \varepsilon\left(T_H^1(t) - \frac{\rho_{\text{ref}}A_{\text{ref}}(2G_0^{DN}G_1^{DN} + (R/\mu)(\rho_{\text{ref}}^2 A_{\text{ref}}^2 - \lambda G_0^{DN^2})(\partial G_1^{DN}/\partial t))}{(\rho_{\text{ref}}A_{\text{ref}} - \lambda G_0^{DN^2})^2}\right), \quad (51)$$

$$T_G^e(t) = \left(\frac{G_0^{DN^2}}{\rho_{\text{ref}}A_{\text{ref}} - \lambda G_0^{DN^2}}\right) + \varepsilon\frac{\rho_{\text{ref}}A_{\text{ref}}(2G_0^{DN}G_1^{DN} + (R/\mu)(\rho_{\text{ref}}^2 A_{\text{ref}}^2 - \lambda G_0^{DN^2})(\partial G_1^{DN}/\partial t))}{(\rho_{\text{ref}}A_{\text{ref}} - \lambda G_0^{DN^2})^2}$$
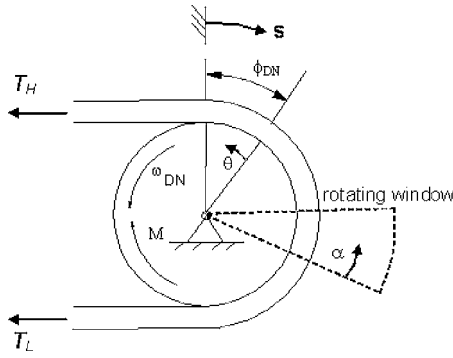
Fig. 4 Coordinates used in determining the belt angular velocity

are effective low, high, and mass flow tensions, respectively. Similar reconstitution procedures are carried out for the driven (ω_{DN}) pulley angular velocity.

The angular velocity of a belt segment in the driven pulley slip zone can now be developed using a rotating reference frame attached to the driven pulley, as shown in Fig. 4. Coordinate θ denotes an Eulerian coordinate measuring angular position of the belt segment in the slip zone—it does not convect with the rotating reference frame. Coordinate α is a Lagrangian coordinate that does convect exactly with the rotating reference frame and mea-

sures a second angular position. These two coordinates are related by the following transformation involving the driven pulley angular velocity,

$$\theta = \alpha + \int_0^t \omega_{DN}(\tau)d\tau, \tag{52}$$

where no loss in generality is incurred by assuming θ and α are aligned at time zero. In what follows, all dependent variables will be *chosen* to depend on the independent quantities α and t, but their functional dependence may be specified in terms θ, with conversion to α implied by (52).

The displacement $u_\alpha(\alpha, t)$ in the α direction at any time t follows from the tension distribution derived in Sec. 3. Specifically,

$$u_\alpha(\alpha, t) = u_0(t) + \int_{\theta'=0}^{\theta'=\theta} \lambda(T_L^e e^{\mu\theta'} + T_G^e)Rd\theta', \tag{53}$$

where $u_0(t)$ denotes $u_\alpha(\alpha, t)$ evaluated at $\theta=0$, i.e., evaluated at $\alpha = -\int_0^t \omega_{DN}(\tau)d\tau$. Evaluation of (53) yields the expression

$$u_\alpha(\alpha, t) = u_0(t) + R\left(\frac{\lambda}{\mu}T_L^e(t)(e^{\mu\theta}-1)\right) + \lambda R T_G^e(t)\theta. \tag{54}$$

The total time derivative of $u_\alpha(\alpha, t)$ *as seen by* an observer moving with the pulley is then given by a material derivative expression

$$\frac{Du_\alpha}{Dt} = \frac{\partial u_\alpha}{\partial t} + \frac{1}{R}\frac{\partial u_\alpha}{\partial \alpha}\frac{Du_\alpha}{Dt}, \tag{55}$$

such that

$$\frac{Du_\alpha}{Dt} = \frac{\partial u_\alpha/\partial t}{1-(1/R)(\partial u_\alpha/\partial \alpha)} = \frac{\partial u_\alpha/\partial t + (\partial u_\alpha/\partial \theta)(\partial \theta/\partial t)}{1-(1/R)(\partial u_\alpha/\partial \theta)(\partial \theta/\partial \alpha)} = \frac{\partial u_0/\partial t + \lambda R((1/\mu)(dT_L^e/dt)(e^{\mu\theta}-1) + (dT_G^e/dt)\theta + (T_L^e e^{\mu\theta} + T_G^e)\omega_{DN})}{1-\lambda(T_L^e e^{\mu\theta} + T_G^e)}, \tag{56}$$

where the time derivative of (52) gives $\partial\theta/\partial t = \omega_{DN}$. The final (i.e., relative to an inertial observer) angular velocity expression for a belt segment in the driven pulley slip zone follows as

$$\omega^{DN}(\theta, t) = \omega_{DN} + \frac{1}{R}\frac{Du_\alpha}{Dt}. \tag{57}$$

An expression is still required for $\partial u_0/\partial t$. Using the boundary condition that the belt is stuck to the pulley at the start of the slip zone, $\omega^{DN}(\theta=0, t) = \omega_{DN}(t)$, one finds

$$\frac{\partial u_0}{\partial t} = -\lambda R\omega_{DN}(T_L^e + T_G^e), \tag{58}$$

such that the final expression for the driven pulley belt angular velocity is

$$\omega^{DN}(\theta, t)$$
$$= \frac{(1-\lambda(T_L^e + T_G^e))\omega_{DN} + \lambda(dT_G^e/dt)\theta + (\lambda/\mu)(dT_L^e/dt)(e^{\mu\theta}-1)}{1-\lambda(T_L^e e^{\mu\theta} + T_G^e)}. \tag{59}$$

The first validity criteria on the solution can then be determined by requiring the velocity of the belt to be greater than the driven pulley velocity for all θ and t, insuring the friction force direction inherited from the steady solution is kinematically correct. This condition is met when $\omega^{DN}(\theta, t) > \omega_{DN}(t)$, or when

$$\frac{dT_G^e/dt\,\theta}{e^{\mu\theta}-1} + \frac{1}{\mu}\frac{dT_L^e}{dt} > -\omega_{DN}(t)T_L^e(t) \quad \forall\, t. \tag{60}$$

Evaluating this condition at the position of lowest expected velocity ($\theta=0$) yields the requirement

$$\frac{dT_L^e}{dt} > -\mu\omega_{DN}(t)T_L^e(t) \quad \forall\, t. \tag{61}$$

Note that this criterion and its development are more general than the analysis of Sec. 3 in that they are independent of the form of the time rate of change of T_L^e—it holds for tension changes due to driver pulley harmonic excitation *as well as* tension changes due to acceleration and deceleration of the driver pulley, etc. For example, one could propose a single slip zone and stick zone exists on either pulley during driver pulley acceleration and again find that (50) governs the tension distribution. Equation (61) would then determine the validity of this proposal in regards to the slip zone.

After a similar analysis for the driver pulley, the following criterion governs the validity of the proposed solutions in the driver pulley slip zone,

$$\frac{dT_H^e}{dt} < \mu\omega_{DR}(t)T_H^e(t) \quad \forall\, t. \tag{62}$$

4.2 Stick Zone Criterion. A second verification of the proposed solutions arises from the requirement that the friction force required to maintain the tension distribution in the stick zone *not*

Table 1 Parameter space for the example two-pulley belt-drive

ω_{steady} (rad/s)	ω_{osc} (rad/s)	$\varepsilon = \omega_{osc}/\omega_{\text{steady}}$	ω (rad/s)	EA (N)	ρA (kg/m)	R (m)	M (Nm)	L (m)	μ	T_0 (N)	I (kg-m^2)
120	1.885	0.016	100π	80,068	0.1036	0.08125	45	0.5105	1.2	827	0.2

be greater than that available. Again considering the driven pulley, this requirement dictates that

$$|f(s,t)| < \mu n(s,t) \quad \forall\, s,t. \tag{63}$$

Equation (3) can be solved for the friction force $f(s,t)$,

$$f(s,t) = \frac{\partial G}{\partial t} + \frac{\partial(\text{v}G)}{\partial s} - \frac{\partial T}{\partial s}, \tag{64}$$

which involves derivatives of the mass flow rate G. This expression can be simplified using conservation of mass,

$$\frac{\partial}{\partial t}\int_{C.V.} \rho d\text{V} + \int_{C.S.} \rho\mathbf{v}\cdot d\mathbf{A} = 0. \tag{65}$$

For the fixed, non-deforming control volume of Fig. 1 this expression evaluates to

$$\frac{\partial}{\partial t}(\rho A) + \frac{\partial}{\partial s}(G) = 0. \tag{66}$$

Multiplying (66) by velocity v and introducing v into the derivatives results in

$$\frac{\partial G}{\partial t} = \rho A\frac{\partial \text{v}}{\partial t} + G\frac{\partial \text{v}}{\partial s} - \frac{\partial(\text{v}G)}{\partial s}. \tag{67}$$

Substitution of (67) into (64) then yields the expression

$$f(s,t) = \rho A\frac{\partial \text{v}}{\partial t} + G\frac{\partial \text{v}}{\partial s} - \frac{\partial T}{\partial s}, \tag{68}$$

which in the stick zones simplifies further since the stick zone velocity $\text{v}=R\omega_{DN}(t)$ is not a function of s,

$$f(s,t) = R\rho A\frac{\partial \omega_{DN}}{\partial t} - \frac{\partial T}{\partial s}. \tag{69}$$

Substituting (4) and (69) into (63), the stick zone criterion can be restated as

$$\left| R\rho A\frac{\partial \omega_{DN}}{\partial t} - \frac{\partial T}{\partial s}\right| < \frac{\mu}{R}(T - GR\omega_{DN}(t)) \quad \forall\, s,t. \tag{70}$$

The tension distribution in the stick zone, by definition, is such that the tension at any location s is that of the low-tension span at a previous instance of time. Therefore,

$$\frac{\partial T}{\partial s} = -\frac{1}{R}\frac{1}{\omega_{DN}}\left.\frac{dT_L}{dt}\right|_{t=t'}, \tag{71}$$

where T_L is evaluated at the previous time t', i.e., the present time minus the time for a belt element to have reached s starting from $s=R\pi$. Since the right-hand side of (70) can be considered relatively constant in time due to the perturbation expansion about the steady solution, (70) must be satisfied at a point s which maximizes the left-hand side, which is a function of nonconstant rates of change. It is chosen to evaluate (71) at a location s such that $dT_L/dt|_{t=t'} = dT_L/dt|_{t=t}$, which is likely to yield the largest value of the left-hand side at all times t.[1] The stick zone criterion can then be updated to

---

[1] This statement would be exactly true if ω_{DN} was constant.

$$\left| R\rho A\frac{\partial \omega_{DN}}{\partial t} + \frac{1}{R\omega_{DN}}\frac{dT_L}{dt}\right| < \frac{\mu}{R}(T_L - GR\omega_{DN}(t)) \quad \forall\, t. \tag{72}$$

Note that when the inertia of the belt is considered to be negligible, as is the case in most belt drives, this criterion simplifies to

$$\frac{dT_L}{dt} < \mu\omega_{DN}(t)T_L(t) \quad \forall\, t, \tag{73}$$

which closely matches the slip zone criterion (71) evaluated with negligible belt inertia.

A similar analysis of the driver stick zone yields the criterion

$$\frac{dT_H}{dt} < \mu\omega_{DR}(t)T_H(t) \quad \forall\, t. \tag{74}$$

4.3 Summary of Validity Criteria. The validity criteria from Secs. 4.1 and 4.2 can be summarized into two compact criteria,

$$\frac{dT_H}{dt} < \mu\omega_{DR}(t)T_H(t), \quad \frac{dT_L}{dt} < \mu\omega_{DN}(t)T_L(t) \quad \forall\, t, \tag{75}$$

where an assumption of negligible belt inertia has been made. Note that for a given set of tensions T_L and T_H, drives operating at lower oscillation frequencies and higher steady angular speed increasingly satisfy the validity criteria.

5 Validation Results for an Example Two-Pulley Drive

The solution procedure presented in Secs. 3.1–3.3 is applied to predict response quantities for an example belt drive composed of a single drive pulley powering a single driven pulley, although as earlier noted the procedure can consider any number of driven pulleys. The analytical results are then compared to a finite element simulation of the same drive. The parameter space for the example drive defined in Table 1 is based closely on that found in [25] and is typical of an automotive application. An initial belt tension T_0, not previously defined, of 827 N is assumed. The belt spans are chosen to be of length $2\pi R$ such that the belt reference length $L_{\text{ref}}^{\text{belt}}$ in (19) is $6\pi R/(1+\lambda T_0)$.

As remarked in Sec. 3.2, the tensions T_L^0 and T_H^0 can be found from (31a) and (32a), completing the $O(\varepsilon^0)$ solution. The important $O(\varepsilon^0)$ response quantities for the example drive are given in Table 2. The tensions T_L^1 and T_H^1 follow from solution of (47), completing the $O(\varepsilon^1)$ solution, and also appear in the table. The full span tensions $T_L(t)$ and $T_H(t)$ are plotted in Fig. 5, as well as the full pulley angular velocities in Fig. 6. Use of the tension and angular velocity information in the table shows that this drive meets the validity criteria (75). It is noted from these results that the analytical solution predicts span tensions neither in-phase nor out-of-phase with the driver pulley excitation, but nearly out-of-phase with each other. Inspection of (32b) shows that this is what is essentially required to satisfy exact compatability at $O(\varepsilon^1)$. The growth and decay of the slip zones, as measured by $\phi_1^{DR}(t)$ and $\phi_1^{DN}(t)$, also show little phase correlation with the excitation source. Finally, for this parameter set, the analytical solution predicts the driven pulley angular velocity oscillations to be small relative to the driver oscillations, and nearly out-of-phase with the driver excitation source.

$O(\varepsilon^0)$						
T_L^0 (N)	T_H^0 (N)	ϕ_0^{DR} (rad)	ϕ_0^{DN} (rad)	G_0 (kg/s)	ω_{DR}^0 (rad/s)	ω_{DN}^0 (rad/s)
553.0	1107	0.586	0.586	0.996	120.0	119.18

$O(\varepsilon^1)$			
$\varepsilon T_L^1(t)$ (N)	$\varepsilon T_H^1(t)$ (N)	$\varepsilon\phi_1^{DR}(t)$ (rad)	$\varepsilon\phi_1^{DN}(t)$ (rad)
$-11.39\sin(100\pi t)$ $84.02\cos(100\pi t)$	$11.55\sin(100\pi t)$ $-85.17\cos(100\pi t)$	$0.0131\sin(100\pi t)$ $-0.0963\cos(100\pi t)$	$0.0087\sin(100\pi t)$ $0.0641\cos(100\pi t)$
$\varepsilon G_1^{DR}(t)$ (kg/s)	$\varepsilon G_1^{DN}(t)$ (kg/s)	$\varepsilon\omega_{DR}^1(t)$ (rad/s)	$\varepsilon\omega_{DN}^1(t)$ (rad/s)
$1.55\text{E}-2\sin(100\pi t)$ $+0.104\text{E}-2\cos(100\pi t)$	$-0.169\text{E}-2\sin(100\pi t)$ $-0.105\text{E}-2\cos(100\pi t)$	$1.885\sin(100\pi t)$	$-0.219\sin(100\pi t)$ $-0.030\cos(100\pi t)$

In addition to the perturbation solution of Sec. 3, an in-house finite element code [25], termed dynamic interaction simulator (DIS), was used to generate results for a similar two-pulley belt drive. This finite element solution technique is *not* limited to small oscillations and is capable of simulating belt-drive performance for a wide range of operating tensions, excitation frequencies, and excitation amplitudes, although at a large computational burden. Some insight into the behavior of the solution is also lost in the
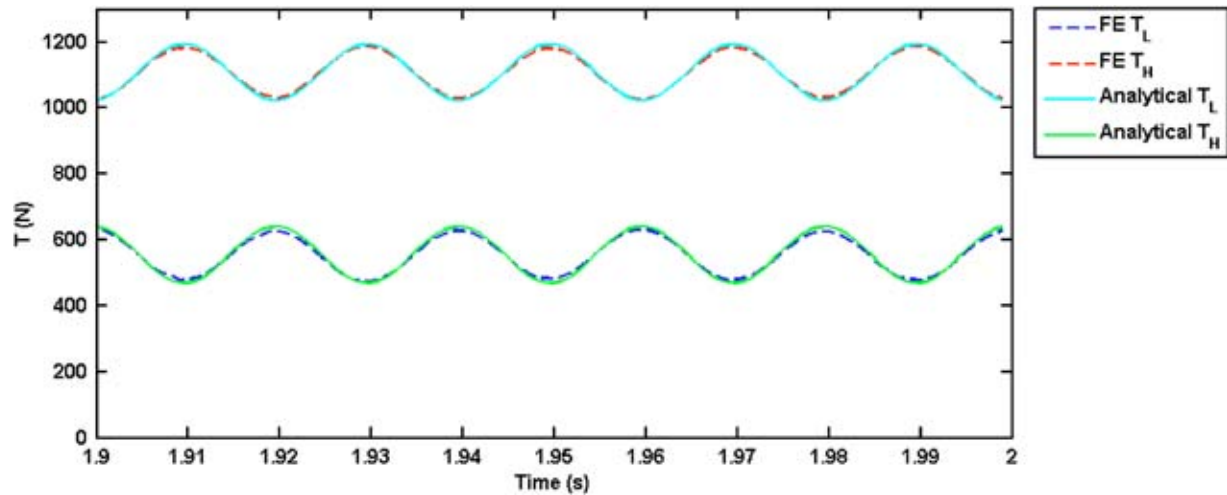
Fig. 5 Tension time histories in the low- and high-tension spans for the analytical and finite element solutions
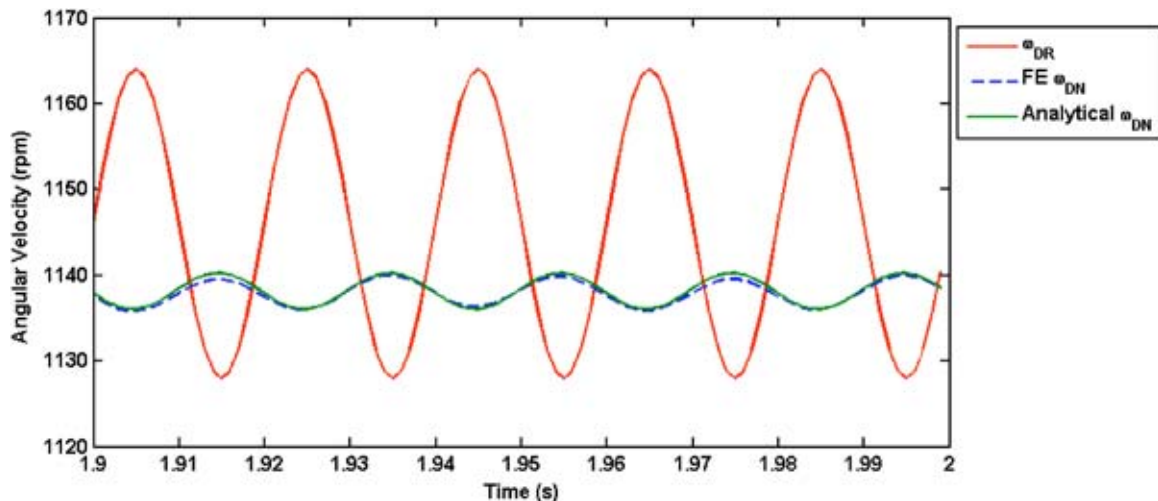
Fig. 6 Pulley angular velocity time histories for the analytical and finite element solutions

finite element solution as response information is available only as tabulated numbers and not as closed-form expressions. In this article, the finite element solution is used to verify the response quantities predicted by the analytical solution, and to verify the existence of a lone slip zone.

The finite element model consists of 450 belt truss elements and two pulley analytical rigid surfaces, and shares the parameter space of the belt drive defined by Table 1. A penalty approach is chosen to model the contact between the belt and the pulleys, as detailed in [25]. The simulation requires starting all pulleys at rest and accelerating the driver pulley to an angular speed of 120 rad/s. Superimposed on the steady angular speed is an oscillating angular velocity component of amplitude 18 rpm (1.885 rad/s) at a frequency of 50 Hz, typical of an automotive application. The simulation is continued until transients in the response decay to near zero.

The span tension and pulley angular velocity time histories for the finite element solution are also plotted in Figs. 5 and 6. As is demonstrated in the figures, the phase information predicted by the analytical solution for both the span tensions and the driven pulley angular velocity is in very good agreement with the finite element solution. The figures also document very good agreement between analytical and finite element steady-state (i.e., $O(\varepsilon^0)$) magnitudes and with oscillatory (i.e., $O(\varepsilon^1)$) magnitudes.

Evidence supporting the existence of a lone stick and slip zone is provided by the finite element solution and is documented in Fig. 7. The tension distribution along the driver pulley is plotted for a single excitation period. Location on the driver pulley is given in degrees measured counter-clockwise from the horizontal (i.e., 3 o'clock) position. As in the analytical solution, the figure documents an initial stick zone in which tension in the high-tension span is "recorded" and then convected unchanged with the pulley as time increases. All tension transitioning then occurs in a lone slip zone at the pulley exit.

6 Conclusions

A closed-form perturbation technique has been presented for the analysis of unsteady belt-drive operation. Response quantities predicted in closed form include the pulley tension distributions, belt-pulley contact forces, and the slip zone metrics. Validity criteria have also been presented which guide the usage of the de-
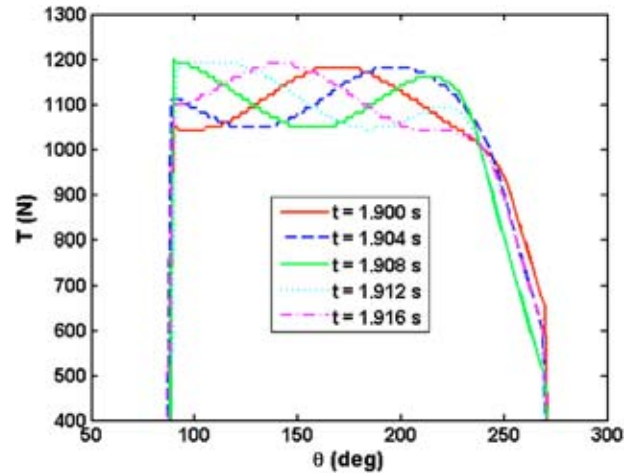
Fig. 7 Tension distribution (predicted by the finite element model) along the driver pulley contact arc for one period of excitation

veloped solutions. The perturbation solution technique was applied to an example two-pulley belt drive in order to determine the drive's global response. The resulting span tensions and the angular velocity of the driven pulley were compared to those predicted by a model analyzed using an in-house finite element simulator. Excellent phase and magnitude agreement was shown for all response quantities compared. Furthermore, the finite element simulation verified the existence of a lone stick and slip arc, an important assumption used in developing the analytical solution.

Appendix

The expressions for $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{f}$ appearing in Sec. 3.3 are as follows:

$$\mathbf{A} = \begin{bmatrix} \lambda L & 0 & 0 \\ 0 & \lambda L & 0 \\ 0 & 0 & I \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} \lambda R \omega_0^{DR} & -\dfrac{(1+\lambda T_H^0)\lambda R \omega_0^{DR}}{1+\lambda T_L^0} & \dfrac{R(1+\lambda T_H^0)^2}{1+\lambda T_L^0} \\[2ex] -\dfrac{(1+\lambda T_L^0)^2 \lambda R \omega_0^{DR}}{(1+\lambda T_H^0)^2} & \dfrac{(1+\lambda T_L^0)\lambda R \omega_0^{DR}}{1+\lambda T_H^0} & -R(1+\lambda T_L^0) \\[2ex] -R + \dfrac{\lambda R^3 \rho_{ref} A_{ref} \omega_{DR}^{0^2}}{(1+\lambda T_H^0)^2} & R - \dfrac{\lambda R^3 \rho_{ref} A_{ref} \omega_{DR}^{0^2}}{(1+\lambda T_H^0)(1+\lambda T_L^0)} - \dfrac{\lambda^2 R^3 \rho_{ref} A_{ref} \omega_{DR}^{0^2}(T_H^0 - T_L^0)}{(1+\lambda T_H^0)^2(1+\lambda T_L^0)} & \dfrac{2\lambda R^3 \rho_{ref} A_{ref} \omega_{DR}^0(T_H^0 - T_L^0)}{(1+\lambda T_H^0)(1+\lambda T_L^0)} \end{bmatrix}$$

$$\mathbf{f} = \begin{bmatrix} R(1+\lambda T_H^0)\omega_{DR}^0 \\[2ex] -\dfrac{R(1+\lambda T_L^0)^2}{1+\lambda T_H^0}\omega_{DR}^0 \\[2ex] 0 \end{bmatrix}$$

References

[1] Euler, M. Leonard, 1762, "Remarques sur l'effect du frottement dans l'equilibre," Mem. Acad. Sci., pp. 265–278.
[2] Grashof, B. G., 1883, *Theoretische Maschinenlehre*, **2**, Leopold Voss, Hamburg.
[3] Fawcett, J. N., 1981, "Chain and Belt Drives—A Review," Shock Vib. Dig., **13**(5), pp. 5–12.
[4] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, London, Chap. 8.
[5] Bechtel, S. E., Vohra, S., Jacob, K. I., and Carlson, C. D., 2000, "The Stretching and Slipping of Belts and Fibers on Pulleys," ASME J. Appl. Mech., **67**, pp. 197–206.
[6] Rubin, M. B., 2000, "An Exact Solution for Steady Motion of an Extensible

Belt in Multipulley Belt Drive Systems," ASME J. Mech. Des., **122**, pp. 311–316.

[7] Smith, D. P., 1999, "Tribology of the Belt-Driven Data Tape Cartridge," Tribol. Int., **31**(8), pp. 465–477.

[8] Firbank, T. C., 1970, "Mechanics of the Belt Drive," Int. J. Mech. Sci., **12**, pp. 1053–1063.

[9] Gerbert, G. G., 1991, "On Flat Belt Slip," *Vehicle Tribology*, Tribology Series 16, Elsevier, Amsterdam, pp. 333–339.

[10] Gerbert, G. G., 1996, "Belt Slip—A Unified Approach," ASME J. Mech. Des., **118**, pp. 432–438.

[11] Townsend, W. T., and Salisbury, J. K., 1988, "The Efficiency Limit of Belt and Cable Drives," ASME J. Mech., Transm., Autom. Des., **110**, pp. 303–307.

[12] Barker, C. R., Oliver, L. R., and Brieg, W. F., 1991, "Dynamic Analysis of Belt Drive Tension Forces During Rapid Engine Acceleration," SAE Congress, Detroit, MI, Paper No. 910687, pp. 239–254.

[13] Hwang, S. J., Perkins, N. C., Ulsoy, A. G., and Meckstroth, R. J., 1994, "Rotational Response and Slip Prediction of Serpentine Belt Drive Systems," ASME J. Vibr. Acoust., **116**(1), pp. 71–78.

[14] Beikmann, R. S., Perkins, N. C., and Ulsoy, A. G., 1996, "Free Vibration of Serpentine Belt Drive Systems," ASME J. Vibr. Acoust., **118**(3), pp. 406–413.

[15] Beikmann, R. S., Perkins, N. C., and Ulsoy, A. G., 1996, "Nonlinear Coupled Vibration Response of Serpentine Belt Drive Systems," ASME J. Vibr. Acoust., **118**(4), pp. 567–574.

[16] Beikmann, R. S., Perkins, N. C., and Ulsoy, A. G., 1997, "Design and Analysis of Automotive Serpentine Belt Drive Systems for Steady State Performance," ASME J. Mech. Des., **119**(2), pp. 162–168.

[17] Leamy, M. J., Perkins, N. C., Barber, J. R., and Meckstroth, R. J., 1997, "The Influence of Tensioner Friction on Accessory Drive Dynamics," 1997 SAE Noise & Vibration Conference and Expedition, Traverse City, MI, May 20–22, Paper No. 97NV103.

[18] Leamy, M. J., and Perkins, N. C., 1998, "Nonlinear Periodic Response of Engine Accessory Drives with Dry Friction Tensioners," ASME J. Vibr. Acoust., **120**(4), pp. 909–916.

[19] Kraver, T. C., Fan, G. W., and Shah, J. J., 1996, "Complex Modal Analysis of a Flat Belt Pulley System with Belt Damping and Coulomb-Damped Tensioner," ASME J. Mech. Des., **118**, pp. 306–311.

[20] Kong, L., and Parker, R. G., 2003, "Equilibrium and Belt-Pulley Vibration Coupling in Serpentine Belt Drives," ASME J. Appl. Mech., **70**, pp. 739–750.

[21] Kong, L., and Parker, R. G., 2004, "Coupled Belt-Pulley Vibration in Serpentine Drives With Belt Bending Stiffness," ASME J. Appl. Mech., **71**, pp. 109–114.

[22] Leamy, M. J., Barber, J. R., and Perkins, N. C., 1998, "Distortion of a Harmonic Elastic Wave Reflected From a Dry Friction Support," ASME J. Appl. Mech., **65**(4), pp. 851–857.

[23] Leamy, M. J., Barber, J. R., and Perkins, N. C., 1998, "Dynamics of Belt/Pulley Frictional Contact," IUTAM Symposium on Unilateral Multibody Contacts, Proceedings, Munich, August 3–7, Kluwer Academic Press, pp. 277–286.

[24] Leamy, M. J., 1998, "The Influence of Dry Friction in the Dynamic Response of Accessory Belt Drive Systems," doctoral dissertation, The University of Michigan.

[25] Leamy, M. J., and Wasfy, T., 2002, "Transient and Steady-State Dynamic Finite Element Modeling of Belt-Drives," ASME J. Dyn. Syst., Meas., Control, **124**(4), pp. 575–581.

[26] Wasfy, T. M., and Leamy, M. J., 2002, "Effect of Bending Stiffness on the Dynamic and Steady-State Responses of Belt-Drives," ASME 2002 Design Engineering Technical Conferences, Montreal, Canada, September 29–October 2.

[27] Meckstroth, R. J., Wasfy, T. M., and Leamy, M. J., 2004, "Finite Element Study of Dynamic Response of Serpentine Belt-Drives With Isolator Clutches," Proceedings of the 2004 SAE World Congress, Paper 04B–174, March 8–11, Detroit, MI.

Electrostatic Forces and Stored Energy for Deformable Dielectric Materials

Robert M. McMeeking
Fellow ASME
Department of Mechanical and Environmental
Engineering and Materials Department,
University of California,
Santa Barbara, CA 93106
e-mail: rmcm@engineering.ucsb.edu

Chad M. Landis
Mem. ASME
Department of Mechanical Engineering and
Materials Science,
Rice University,
Houston, TX 77251
e-mail: landis@rice.edu

An isothermal energy balance is formulated for a system consisting of deformable dielectric bodies, electrodes, and the surrounding space. The formulation in this paper is obtained in the electrostatic limit but with the possibility of arbitrarily large deformations of polarizable material. The energy balance recognizes that charges may be driven onto or off of the electrodes, a process accompanied by external electrical work; mechanical loads may be applied to the bodies, thereby doing work through displacements; energy is stored in the material by such features as elasticity of the lattice, piezoelectricity, and dielectric and electrostatic interactions; and nonlinear reversible material behavior such as electrostriction may occur. Thus the external work is balanced by (1) internal energy consisting of stress doing work on strain increments, (2) the energy associated with permeating free space with an electric field, and (3) by the electric field doing work on increments of electric displacement or, equivalently, polarization. For a conservative system, the internal work is stored reversibly in the body and in the underlying and surrounding space. The resulting work statement for a conservative system is considered in the special cases of isotropic deformable dielectrics and piezoelectric materials. We identify the electrostatic stress, which provides measurable information quantifying the electrostatic effects within the system, and find that it is intimately tied to the constitutive formulation for the material and the associated stored energy and cannot be independent of them. The Maxwell stress, which is related to the force exerted by the electric field on charges in the system, cannot be automatically identified with the electrostatic stress and is difficult to measure. Two well-known and one novel formula for the electrostatic stress are identified and related to specific but differing constitutive assumptions for isotropic materials. The electrostatic stress is then obtained for a specific set of assumptions in regard to a piezoelectric material. An exploration of the behavior of an actuator composed of a deformable, electroactive polymer is presented based on the formulation of the paper. [DOI: 10.1115/1.1940661]
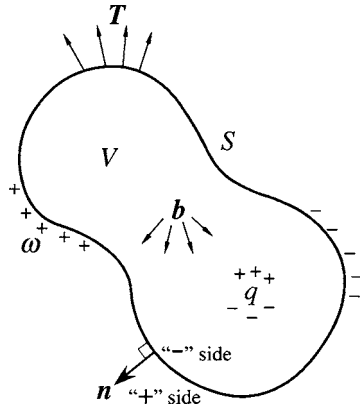
Introduction

The subject of electrostatics is the study of the effect of forces generated between charges. Therefore, when electric fields are present simultaneously with mechanical loading in a material, it seems obvious that the influence of electrostatic forces should be accounted for when stresses in the material are calculated. However, stresses due to electrostatic effects are often second order compared to those due to other purely mechanical effects [1] and therefore electrical forces are often neglected even as electric fields are analyzed. An example of this is the linear theory of piezoelectrics [2], where electric fields induce strain but stress due to electrostatic forces, being quadratic in electric field, is neglected. On the other hand, finite strains, nonlinear material behavior, the lack of mechanical loading, and other effects can lead to situations in which electrostatically induced stresses are comparable with forces from other sources. Such cases have been addressed for dielectric materials by several authors, beginning with the pioneering contributions of Toupin [1] and Eringen [3]. Beyond those examples given above, new reasons for the need to include the effect of electrostatically induced loading in the analy-

sis of stress continue to emerge. For example, the analysis of cracks including the effects of cohesive zones requires a consistent treatment of electrically induced forces across cracks [4]. Actuators composed of electrically active polymers are emerging as important devices [5] and these components function due to the electrostatic forces generated in them. Thus, the question of the electrostatic contribution to stress continues to increase in importance.

As noted above, treatments of electrostatic contributions to stress have been provided previously [1,3]. In these papers, a series of electric fields is identified and added together to form the total field. One field is that which would exist if the dielectric material was not present and a second one is due to the effect of polarization charge in the material and on its surface. These two fields together compose the classical Maxwell-Faraday electric field. To this is added what is designated as the local electric field that is considered to be a function, through a constitutive law, of the material's strain and polarization. In addition, the electrostatic forces (i.e., the Maxwell stress [6], whose divergence is the electrical body force and whose surface tractions are the electrical forces per unit area acting on interfaces) are taken to have a specific relationship to the electric field and consequently the formulation is seemingly restricted. Apparently, no allowance is made for the possibility that experiments will show that for different classes of material the electrostatic forces and the Maxwell stress will have some other relationship to the electric field and the polarization than the one assumed. It should be noted that this may not be an insurmountable restriction, since it is well known that there is some arbitrariness in how stress is divided up into electrical and material contributions [1]. Therefore, any discrep-

Fig. 1 A dielectric body with body forces, surfaces tractions, and free charges

ancy with experiment as far as the Maxwell stress is concerned may be compensated for by adjustment to the constitutive law for the material stress, presumably leading to nonlinear electrostrictive contributions. However, the structure just described is at the very least inconvenient due to the restrictions on the permitted Maxwell stress and due to the formalities used to construct the electric field.

In the current paper, a different formulation of the problem is presented, though one in which the basic physics is no different from that used in Refs. [1,3]. However, the structure of the electric field is taken to have a simpler form than used in Refs. [1,3] since we make no attempt to identify different contributions to it such as an external field, a depolarization field, or a local field. Instead, a single electric field is utilized throughout and it is assumed to be connected to material strain and polarization by a constitutive law. Furthermore, no general prejudgment is made of the relationship between the electrostatic forces (i.e., the Maxwell stress or alternatively the electrostatic stress) and the electric field and polarization. Instead, it is assumed that the electrostatic stress is measurable in experiments either directly through characterization of stress and electric fields or through measurement of the constitutive properties of the material. We note that our approach parallels that used by Landau and Lifschitz [6] but we do not restrict ourselves to infinitesimal elastic strains of isotropic materials and piezoelectric materials. It is our belief that this formulation is more versatile than what has been presented in the past and that it is more suitable for incorporation into modern methods such as finite element analysis.

Balance of Energy for a Deformable Dielectric Material

Consider a system consisting of dielectric materials, perfect conductors, and free space. In the current configuration, the system occupies the volume V as shown in Fig. 1. The system is considered to be isolated so that there is no interaction between electrical charges residing in the system and those outside. Formally, this implies that the volume of interest is shielded electrically from its exterior or that the extent of the system is infinite, since charges interact with each other over very long distances. However, when practical calculations with approximations are attempted, this formality can be ignored. The perimeter of the system plus interfaces within it are designated S in the current configuration. The internal interfaces separate the dielectric materials, the conductors, and free space from each other. In addition, sectors of dielectric with homogeneous or heterogeneous properties may be separated by surfaces included within S, as may sectors of free space.

Let the free charge per unit volume within V be $q(x_i, t)$ where x_i is the position of material points in the current configuration and t is time. Free charge may be placed in free space, in which case x_i

is used to designate the position of the points occupied by free space in the current configuration. Furthermore, let $\omega(x_i, t)$ be the free charge per unit area on the surfaces S and define $\phi(x_i, t)$ to be the electrical potential everywhere within the system such that it is continuous everywhere in space. Note that we will consider only the electrostatic limit so that t plays the role of a load parameter as far as the electrical variables are concerned and no attempt will be made to explore Maxwell's equations relevant to the electrodynamic limit. On the other hand, time may have a real meaning as far as the deformation of material is concerned, as may be the case in strain-rate-dependent response or in the acoustic limit where inertia has to be included in the balance laws. To complete the variables to be considered in the external work statement, we include the velocity $v_i(x_i, t)$ of material points, the surface traction $T_i(x_i, t)$ defined as the force per unit area acting on S, and $b_i(x_i, t)$, which is the body force per unit volume acting at points in V. Note that the surface traction T_i and the body force b_i arise from sources other than electrical effects and, therefore, do not represent the influence of charges interacting at a distance or electrical fields exerting forces on charges. The surface traction T_i and the body force b_i and any equivalent quantity defined in the current state will be designated mechanical, though we do not make any attempt to characterize how they may arise, whether they come about by gravitational effects or other sources of force in materials. To attempt to do so in too fine a detail and make distinctions between forces that arise by electrostatic effects and forces that arise from other effects would illuminate the arbitrariness in how the designations electrical and mechanical are utilized in our scheme; e.g., consider a purely ionic solid pressing against another purely ionic solid to produce supposedly mechanical traction between them. Much of the traction between the two solids in this case will in fact be electrostatic due to the repulsion of like atoms, although some of it will also develop due to quantum exclusion effects. However, in any macroscopic treatment of this problem, this interaction between the bodies when pressed together will be represented by elasticity and the tractions thereby designated to be mechanical.

Consider the physical laws governing the electromechanical fields in the material. In the quasi-static limit, Maxwell's laws state that the electric field must be curl-free and Gauss' law states that the divergence of the electric displacement must be equal to the volume density of free charge. Therefore,

$$\in_{ijk} \frac{\partial E_j}{\partial x_k} = 0 \Rightarrow E_i = -\frac{\partial \phi}{\partial x_i} \tag{1}$$

$$\frac{\partial D_i}{\partial x_i} = q \text{ in } V \tag{2}$$

$$\|D_i n_i\| = \omega \text{ on } S \tag{3}$$

Here, n_i are the Cartesian components of the unit normal to the surface S pointing from the "−" side of the surface out towards the "+" side as shown in Fig. 1, and $\in_{ijk}$ are the components of the permutation symbol. Then the notation $\| \ \|$ represents the difference or jump in the included quantity across the surface S such that

$$\|D_i\| = D_i^+ - D_i^- \tag{4}$$

Furthermore, the electric displacement can be decomposed into two parts such that

$$D_i = \kappa_0 E_i + P_i \tag{5}$$

where κ_0 is the dielectric permittivity of free space and P_i are the Cartesian components of the material polarization.

Conservation of mass implies that for a given material volume

$$\frac{d}{dt}\int_V \rho dV = 0 \Rightarrow \frac{d\rho}{dt} + \rho\frac{\partial v_i}{\partial x_i} = 0 \qquad (6)$$

where $\rho(x_i,t)$ is the mass density of the material. The principles of conservation of linear and angular momentum are stated as

$$\int_V (b_i + b_i^E)dV + \int_S (T_i + T_i^E)dS = \frac{d}{dt}\int_V \rho v_i dV \qquad (7)$$

and

$$\int_V \in_{ijk} x_j(b_k + b_k^E)dV + \int_S \in_{ijk} x_j(T_k + T_k^E)dS = \frac{d}{dt}\int_V \rho \in_{ijk} x_j v_k dV \qquad (8)$$

Here the components of the electrical body force b_i^E and surface traction T_i^E have been introduced. These forces arise directly from electric fields acting in the material and are in addition to the mechanical body force and surface traction. Furthermore, it is assumed that the electrical body force can be derived from the Maxwell stress tensor σ_{ij}^M such that

$$b_i^E = \frac{\partial \sigma_{ji}^M}{\partial x_j} \text{ in } V \qquad (9)$$

where the electrical body force is the effect of charges interacting at a distance or, equivalently, the force per unit volume arising from electric fields acting on charges. The traction relationship for Maxwell stress is then

$$T_i^E = n_j \|\sigma_{ji}^M\| \text{ on } S. \qquad (10)$$

Then, in order to satisfy the principle of conservation of linear momentum for a small surface element, the Cauchy stress in the material, σ_{ij}, must balance the total surface traction such that along with Eq. (10)

$$T_i + T_i^E = -n_j\|\sigma_{ji}\| \Rightarrow T_i = -n_j\|\sigma_{ji} + \sigma_{ji}^M\| \qquad (11)$$

where again $T_i(x_i,t)$ is the nonelectrical (i.e., mechanical) surface force per unit area acting on S. This statement has an equivalent meaning to Eq. (10). The Cauchy stress difference across a surface must balance both the electrical and mechanical surface tractions. Note, however, that the result in Eq. (11) illustrates why it is difficult, if not impossible, to separately measure the Cauchy and Maxwell stresses, since it shows that any traction measured by mechanical means (i.e., by the only method available) is related to their sum. Since there are no experiments that can separate the effects of the Cauchy and Maxwell stresses unambiguously [1,3,6], it is generally more profitable to consider their sum and not to try to identify them separately. The sum will be termed the total true stress.

Next, application of Eqs. (6), (9), and (11) within the principles of conservation of linear and angular momentum and recognition that the resultant integrals must be valid for any arbitrary volume yield

$$\frac{\partial \sigma_{ji}}{\partial x_j} + \frac{\partial \sigma_{ji}^M}{\partial x_j} + b_i = \rho\frac{dv_i}{dt} \text{ in } V \qquad (12)$$

and

$$\sigma_{ji} + \sigma_{ji}^M = \sigma_{ij} + \sigma_{ij}^M \text{ in } V \qquad (13)$$

Thus, for the balance of angular momentum to be satisfied, the total true stress must be symmetric. This requirement reflects the fact that moments due to mechanical body forces and inertia can be assumed, as usual, to be second order [7,8], but we must allow for the possibility that electric effects induce first-order moments, e.g., due to electric fields acting on dipoles in the material [1,3]. Since only the total true stress in the material must be symmetric, it is possible that both the Maxwell and Cauchy stress tensors can be nonsymmetric.

Now consider the rate of work by agencies external to the system. This external work rate is

$$\frac{dW}{dt} = \int_V b_i v_i dV + \int_S T_i v_i dS + \int_V \phi\frac{d}{dt}(qdV) + \int_S \phi\frac{d}{dt}(\omega dS) \qquad (14)$$

Note that this statement conforms to the usual definition of electrical work, i.e., an increment of electrical work is given by the electric potential ϕ multiplied by the increment of charge ΔQ. For example, $\phi\Delta(qdV)$ is the work done by external agencies to bring the charge $\Delta(qdV)$ from infinity to the point where the potential is ϕ. The interpretation of the mechanical work done by the body forces and surface tractions is obvious.

Based on the standard continuum mechanics results for the material time derivatives of field quantities and volume and surface elements [7, pp. 211–213], it can be shown that

$$\frac{d}{dt}(qdV) = \left(\frac{dq}{dt} + q\frac{\partial v_k}{\partial x_k}\right)dV = \left(\frac{\partial q}{\partial t} + v_k\frac{\partial q}{\partial x_k} + q\frac{\partial v_k}{\partial x_k}\right)dV \qquad (15)$$

and

$$\frac{d}{dt}(\|D_i n_i\|dS) = \left(\left\|\frac{dD_i}{dt} + D_i\frac{\partial v_k}{\partial x_k} - D_j\frac{\partial v_i}{\partial x_j}\right\| n_i\right)dS \qquad (16)$$

Then, applying these results along with the balance of linear momentum, Eqs. (11) and (12), we can write the work rate as

$$\frac{dW}{dt} = \int_S \phi n_i\left\|\frac{dD_i}{dt} + D_i\frac{\partial v_k}{\partial x_k} - \frac{\partial v_i}{\partial x_j}D_j\right\|dS + \int_V \phi\left(\frac{\partial q}{\partial t} + v_k\frac{\partial q}{\partial x_k}\right.$$
$$\left. + q\frac{\partial v_k}{\partial x_k}\right)dV - \int_S n_j\|\sigma_{ji} + \sigma_{ji}^M\|v_i dS - \int_V \left(\frac{\partial \sigma_{ji}}{\partial x_j} + \frac{\partial \sigma_{ji}^M}{\partial x_j}\right)v_i dV$$
$$+ \int_V \rho\frac{dv_i}{dt}v_i dt \qquad (17)$$

Use of the divergence theorem for a collection of subvolumes whose union is V and which are separated by the interfaces that collectively compose S gives

$$\frac{dW}{dt} = -\int_V \frac{\partial}{\partial x_i}\left[\phi\left(\frac{dD_i}{dt} + D_i\frac{\partial v_k}{\partial x_k} - \frac{\partial v_i}{\partial x_j}D_j\right)\right]dV$$
$$+ \int_V \phi\left(\frac{\partial^2 D_i}{\partial t\partial x_i} + v_k\frac{\partial^2 D_i}{\partial x_k\partial x_i} + \frac{\partial D_i}{\partial x_i}\frac{\partial v_k}{\partial x_k}\right)dV$$
$$+ \int_V \frac{\partial}{\partial x_j}[(\sigma_{ji} + \sigma_{ji}^M)v_i]dV - \int_V \left(\frac{\partial \sigma_{ji}}{\partial x_j} + \frac{\partial \sigma_{ji}^M}{\partial x_j}\right)v_i dV$$
$$+ \int_V \rho\frac{dv_i}{dt}v_i dV \qquad (18)$$

where Eq. (2) has been used to dispose of the free charge density q. Equation (18) simplifies to

$$\frac{dW}{dt} = \int_V \left[E_i\frac{dD_i}{dt} + (\sigma_{ji} + \sigma_{ji}^M - E_iD_j + E_kD_k\delta_{ij})\frac{\partial v_i}{\partial x_j}\right]dV$$
$$+ \int_V \rho\frac{dv_i}{dt}v_i dV \qquad (19)$$

where δ_{ij} is the Kronecker delta.

Note that in regions of space without material, the electric displacement is defined to be [6]

$$D_i = \kappa_o E_i \tag{20}$$

where κ_o is the permittivity of free space. In addition, the Maxwell stress in such regions is [6]

$$\sigma_{ij}^M = \kappa_o\left(E_i E_j - \tfrac{1}{2} E_k E_k \delta_{ij}\right) \tag{21}$$

so that from Eqs. (2), (9), and (20), we find that the electrical force per unit volume is

$$b_i^E = q E_i \tag{22}$$

which is consistent with the standard result that the force on a charge is given by the charge times the electric field.

Now we recognize the material polarization P_i and use Eq. (5) to replace D_i wherever it appears in Eq. (19) to obtain

$$\frac{dW}{dt} = \frac{d}{dt}\int_V \frac{1}{2}\kappa_o E_i E_i dV + \frac{d}{dt}\int_V \frac{1}{2}\rho v_i v_i dV + \int_V \left[E_i \frac{dP_i}{dt} \right.$$
$$\left. + (\sigma_{ji} + \sigma_{ji}^M - \hat{\sigma}_{ji}^M - E_i P_j + E_k P_k \delta_{ij}) \frac{\partial v_i}{\partial x_j} \right] dV \tag{23}$$

where

$$\hat{\sigma}_{ij}^M = \kappa_o\left(E_i E_j - \tfrac{1}{2} E_k E_k \delta_{ij}\right) \tag{24}$$

is thus the Maxwell stress at the specified electric field for free space absent any material and the second integral on the right-hand side has been obtained from the second one on the right of Eq. (19) by use of the first form of conservation of mass, Eq. (6). In the derivation of Eq. (23), the fact that

$$\kappa_o \int_V \left[E_i \frac{dE_i}{dt} - (E_i E_j - E_k E_k \delta_{ij})\frac{\partial v_i}{\partial x_j} \right] dV = \int_V \left[\kappa_o E_i \frac{\partial E_i}{\partial t} - \hat{\sigma}_{ij}^M \frac{\partial v_i}{\partial x_j} \right.$$
$$\left. + \frac{\partial}{\partial x_k}\left(\frac{1}{2}\kappa_o E_i E_i v_k\right) \right] dV \tag{25}$$

has been used and the divergence theorem then provides

$$\int_V \left[\kappa_o E_i \frac{\partial E_i}{\partial t} + \frac{\partial}{\partial x_k}\left(\frac{1}{2}\kappa_0 E_i E_i v_k\right) \right] dV = \int_V \kappa_o E_i \frac{\partial E_i}{\partial t} dV$$
$$- \int_S \frac{1}{2}\kappa_o \|E_i E_i\| n_k v_k dS \tag{26}$$

The terms on the right-hand side of Eq. (26) combine to give the first term on the right-hand side of Eq. (23). Note that in regions of free space without material, the second and third integrals on the right-hand side of Eq. (23) are both zero (given that the Cauchy stress is zero there), indicating that the first integral gives the rate of energy storage in space to permeate it with the electric field. Thus the third integral on the right-hand side of Eq. (23) is the rate at which work is stored or dissipated in the material, other than kinetic energy, which is, of course, accounted for by the first integral.

Finally, we note that Eq. (23) can be converted to a principle of virtual work; however, the details will not be emphasized here. Equation (23) is valid for the isothermal response of any electromechanical material whether the behavior is reversible or dissipative. However, the next section will focus attention on reversible material response.

Conservative Materials

In this section we formally introduce the first and second laws of thermodynamics for the electromechanical situation under consideration. Ultimately, we will specialize the results to reversible material behavior, i.e., conservative materials. In conservative materials, the work done by external agencies that is not absorbed by kinetic energy is stored in the material in the form of elastic distortion, dielectric polarization, piezoelectric response, electrostrictive behavior, electrostatic interactions, and any other recoverable

energy storage mechanism that is active. The balance of energy required by the first law of thermodynamics can be stated as

$$\frac{dU}{dt} = \frac{dW}{dt} + \frac{dQ}{dt} \tag{27}$$

where U represents the stored internal energy of the material, the electrical energy stored by free space, and the kinetic energy of the material; dW/dt is the external work rate derived in Eq. (14), and dQ/dt is the rate at which heat is transferred to the system. Note that the free space in which energy is stored is not just that occupied by the body or bodies under consideration. Material subject to electric fields couple with the surrounding aether and permeate it with an electrical field. Thus, the free space being considered includes any surrounding space affected by the electric fields associated with the body or bodies. Specifically, dU/dt and dQ/dt are written as

$$\frac{dU}{dt} = \frac{d}{dt}\int_V \rho u dV + \frac{d}{dt}\int_V \frac{1}{2}\kappa_o E_i E_i dV + \frac{d}{dt}\int_V \frac{1}{2}\rho v_i v_i dV \tag{28}$$

and

$$\frac{dQ}{dt} = \int_V \rho \dot{r} dV - \int_{S_e} \hat{q}_i n_i dS \tag{29}$$

where u is the internal energy per unit mass, $\dot{r}$ is the heat input rate per unit mass, and $\hat{q}_i$ are the components of the heat flux vector with the positive sense directed out of the external surface S_e of the system. Note that at all points interior to the system the heat flux vector is taken to be continuous, eliminating the possibility of surfaces acting as sources of heat. Equations (23) and (27)–(29) can be combined and must hold for any arbitrary volume yielding a local form for the first law as

$$\rho \frac{du}{dt} = (\sigma_{ji} + \sigma_{ji}^M - \hat{\sigma}_{ji}^M - E_i P_j + E_k P_k \delta_{ij})\frac{\partial v_i}{\partial x_j} + E_i \frac{dP_i}{dt} + \rho\dot{r} - \frac{\partial \hat{q}_i}{\partial x_i} \tag{30}$$

The second law of thermodynamics states that the entropy production rate must be equal to or exceed the rate of entropy input to a region, i.e.,

$$\frac{d}{dt}\int_V \rho s dV \geqslant \int_V \frac{\rho\dot{r}}{\theta} dV - \int_{S_e} \frac{\hat{q}_i n_i}{\theta} dS \tag{31}$$

where s is the entropy per unit mass of the material and θ is the absolute temperature. Equation (31) can be manipulated by application of the divergence theorem to the last term on the right-hand side and then required to be valid for any arbitrary volume to yield a local form of the second law as

$$\rho \frac{ds}{dt} \geqslant \frac{\rho\dot{r}}{\theta} - \frac{1}{\theta}\frac{\partial \hat{q}_i}{\partial x_i} + \frac{\hat{q}_i}{\theta^2}\frac{\partial \theta}{\partial x_i} \tag{32}$$

Then, defining ψ as the Helmholtz free energy per unit mass, we have

$$\psi = u - \theta s \tag{33}$$

and in combination with Eq. (30)–(33), Eq. (32) can be rewritten as

$$(\sigma_{ji} + \sigma_{ji}^M - \hat{\sigma}_{ji}^M - E_i P_j + E_k P_k \delta_{ij})\frac{\partial v_i}{\partial x_j} + E_i \frac{dP_i}{dt} - \rho s \frac{d\theta}{dt} - \rho \frac{d\psi}{dt}$$
$$- \frac{\hat{q}_i}{\theta}\frac{\partial \theta}{\partial x_i} \geqslant 0 \tag{34}$$

For a conservative electro-active material it is assumed that in general ψ is a function of the deformation gradient, polarization, and temperature, i.e., $\psi = \psi(F_{ij}, P_i, \theta)$. Note that for material

points, x_i designates their positions in the current configuration and a mapping

$$x_i = x_i(X_j, t) \tag{35}$$

associates these points at time t with their positions X_i in the reference configuration [7]. The components of the deformation gradient are given as

$$F_{ij} = \frac{\partial x_i}{\partial X_j} \tag{36}$$

Then, using

$$\frac{\mathrm{d}F_{ik}}{\mathrm{d}t} = \frac{\partial x_j}{\partial X_k} \frac{\partial v_i}{\partial x_j} \tag{37}$$

we can modify Eq. (34) to

$$\left[(\sigma_{ji} + \sigma_{ji}^M - \hat{\sigma}_{ji}^M - E_i P_j + E_k P_k \delta_{ij}) - \rho \frac{\partial \psi}{\partial F_{ik}} F_{jk} \right] \frac{\partial v_i}{\partial x_j}$$
$$+ \left[E_i - \rho \frac{\partial \psi}{\partial P_i} \right] \frac{\mathrm{d}P_i}{\mathrm{d}t} - \left[\rho s + \rho \frac{\partial \psi}{\partial \theta} \right] \frac{\mathrm{d}\theta}{\mathrm{d}t} - \frac{\hat{q}_i}{\theta} \frac{\partial \theta}{\partial x_i} \geq 0 \tag{38}$$

Following the methods of Coleman and Noll [9], we postulate that (38) must hold for every admissible process. First, consider processes where the deformation gradient, polarization, and temperature are spatially homogeneous but arbitrary functions of time. Such processes can be thought as being controlled by the appropriate applications of body forces, charge densities, and heat supply. For spatially homogeneous temperature distributions the last term on the left-hand side of (38) vanishes. If $\mathrm{d}P_i/\mathrm{d}t$ and $\mathrm{d}\theta/\mathrm{d}t$ are taken to be zero and $\partial v_i/\partial x_j$ is arbitrarily chosen, then (38) is satisfied only if

$$\sigma_{ji} + \sigma_{ji}^M - \hat{\sigma}_{ji}^M - E_i P_j + E_k P_k \delta_{ij} = \rho \frac{\partial \psi}{\partial F_{ik}} F_{jk} \tag{39}$$

If $\mathrm{d}\theta/\mathrm{d}t$ and $\partial v_i/\partial x_j$ are zero and $\mathrm{d}P_i/\mathrm{d}t$ is chosen arbitrarily, then (38) implies

$$E_i = \rho \frac{\partial \psi}{\partial P_i} \tag{40}$$

If $\partial v_i/\partial x_j$ and $\mathrm{d}P_i/\mathrm{d}t$ are zero and $\mathrm{d}\theta/\mathrm{d}t$ is arbitrarily chosen, then (38) yields

$$s = -\frac{\partial \psi}{\partial \theta} \tag{41}$$

Finally, consider spatially homogeneous and time-independent distributions of the deformation gradient and polarization, and spatially inhomogeneous but time-independent distributions of temperature. Such processes and Eq. (38) then imply the heat conduction inequality

$$-\hat{q}_i \frac{\partial \theta}{\partial x_i} \geq 0 \tag{42}$$

The requirements of objectivity [7] place restrictions on the forms that the constitutive laws can take. This leads us to conclude that the Helmholtz free energy per unit mass at fixed temperature must have the form

$$\psi = \psi(U_{ij}, \Pi_i) \tag{43}$$

where Π_i are the components of the rotation invariant polarization defined as

$$\Pi_i = P_j R_{ji} \tag{44}$$

where R_{ji} is the orthogonal transformation arising from polar decomposition of F_{ij} into a pure deformation and a pure rotation [7] as

$$F_{ij} = R_{ik} U_{kj} \tag{45}$$

Finally, U_{ij} are the components of the right stretch tensor.

Note that the choice we have made for the functional dependence of ψ in terms of a polarization measure is not the only one possible that would satisfy objectivity. However, we find it to be of the most convenience. In general, to satisfy the requirements of objectivity, the Helmholtz free energy density can depend on any measures of polarization and strain that in turn depend only on the rotation-invariant polarization and the right stretch tensor. Since their evaluation does not require the solution of an eigenvalue problem, it is common to define the relevant polarization measure as $\mathbb{P}_i = \Pi_j U_{ji} = P_k F_{ki}$ and the strain measure as $\varepsilon_{ij} = (U_{ik}U_{kj} - \delta_{ij})/2 = (F_{ki}F_{kj} - \delta_{ij})/2$. However, the form $\mathbb{P}_i$ for the polarization measure is somewhat convoluted since it is affected by both electrical polarization *and* the stretch of the material element. Hence, for the purposes of this work the polarization measure will be taken as the rotation invariant polarization Π_i, which is independent of the material stretch, and the strain measure will be taken as the Green-Lagrange strain ε_{ij} mentioned above and repeated below as

$$\varepsilon_{ij} = \tfrac{1}{2}(F_{ki}F_{kj} - \delta_{ij}) \tag{46}$$

Using these polarization and strain measures, we find more convenient forms of Eqs. (39) and (40) to be

$$E_i = \rho R_{ij} \frac{\partial \psi}{\partial \Pi_j} \tag{47}$$

$$\sigma_{ji} + \sigma_{ji}^M = \rho \frac{\partial \psi}{\partial \varepsilon_{lm}} F_{il} F_{jm} + \beta_{jikl} P_k E_l + P_j E_i - P_k E_k \delta_{ij} + \hat{\sigma}_{ji}^M \tag{48}$$

where

$$\beta_{jikl} = F_{jm} \frac{\partial R_{kn}}{\partial F_{im}} R_{ln} \tag{49}$$

The derivative of the rotation tensor is given by [10]

$$\frac{\partial R_{ik}}{\partial F_{jl}} = \frac{1}{\nu} \big[(V_{mm}\delta_{ij} - V_{ij})(U_{nn}\delta_{kl} - U_{kl})$$
$$- (V_{mm}R_{il} - F_{il})(V_{nn}R_{jk} - F_{jk}) \big] \tag{50}$$

with

$$\nu = \mathrm{Det}(V_{kk}\delta_{ij} - V_{ij}) = \mathrm{Det}(U_{kk}\delta_{ij} - U_{ij}) \tag{51}$$

in which V_{ij} is the left stretch tensor given by a polar decomposition in the form [7]

$$F_{ij} = V_{ik} R_{kj} \tag{52}$$

Thus

$$\beta_{jikl} = \frac{1}{\nu} V_{jm}(V_{nn}\delta_{mp} - V_{mp})(V_{qq}\delta_{ir} - V_{ir})(\delta_{rk}\delta_{pl} - \delta_{rl}\delta_{pk}) \tag{53}$$

Furthermore, it can be shown that [10]

$$\beta_{jikl} P_k E_l + P_j E_i = \beta_{ijkl} P_k E_l + P_i E_j \tag{54}$$

and this confirms that the total true stress $\sigma_{ji} + \sigma_{ji}^M$ is symmetric as required by conservation of angular momentum. In fact, as long as the free energy density is objective, then the law of conservation of angular momentum, Eqs. (8) and (13), will be satisfied automatically [1].

We note that Eqs. (47) and (48) are together equivalent to concepts developed by Landau and Lifschitz [6] who studied these issues for infinitesimally strained isotropic elastic materials and piezoelectric systems through the use of a free energy. We believe that we have therefore placed these ideas into a more general framework.

Dielectrically Linear, Isotropic, Conservative Materials

Consider now a material in which the stored energy per unit reference volume is the sum of an isotropic function of the deformation plus a term quadratric in the polarization. Thus

$$\psi = \psi_e + \frac{1}{2\rho\kappa_0}\Pi_i\chi_{ij}^{-1}\Pi_j \qquad (55)$$

where ψ_e depends only on the deformation and χ_{ij} is the electrical susceptibility of the material, which is assumed to be dependent on the strain but independent of polarization. Since the material is assumed to be isotropic, ψ_e and χ_{ij} will be functions of the invariants of the Green-Lagrange strain or, equivalently, the right or left stretch tensors [7]. Possible forms of interest for the elastic response would include neo-Hookean, Mooney-Rivlin, Blatz-Ko, and Ogden formulations [7,11–16], but allowance should be made for compressibility of the material to ensure consistency with any dilatancy that is assumed in association with the electrical behavior. The susceptibility tensor is symmetric and must be isotropic in the limit of zero strain so that the unstrained material is electrically isotropic. Due to strain χ_{ij} can become anisotropic [6].

The electric field can now be deduced from Eq. (47) to be

$$E_i = \frac{1}{\kappa_0}R_{ij}\chi_{jk}^{-1}P_lR_{lk} \qquad (56)$$

which is thus a linear relationship between the components of electric field and polarization modified by the effects of straining and rotation of the material. Note that in the limit of zero strain, this relationship is consistent with what is expected for isotropic dielectrics since then the susceptibility is

$$\chi_{ij}^{-1} = \frac{1}{\widetilde{\chi}}\delta_{ij} \qquad (57)$$

where $\widetilde{\chi}$ is the isotropic susceptibility of the unstrained material. The relationship in Eq. (56) then becomes

$$E_i = \frac{1}{\kappa_o\widetilde{\chi}}P_i \qquad (58)$$

The result for stress obtained from Eq. (48) is

$$\sigma_{ji} + \sigma_{ji}^M = F_{jk}\left(\rho\frac{\partial\psi_e}{\partial\varepsilon_{kl}} - \frac{\kappa_o}{2}R_{mn}\frac{\partial\chi_{np}}{\partial\varepsilon_{kl}}R_{qp}E_mE_q\right)F_{il}$$

$$+ \kappa_o R_{jk}\chi_{kl}R_{ml}E_mE_i - \frac{\kappa_o}{2}R_{kl}\chi_{lm}R_{nm}E_kE_n\delta_{ji}$$

$$+ \kappa_o\beta_{jikl}R_{km}\chi_{mn}R_{pn}E_lE_p + \hat{\sigma}_{ji}^M \qquad (59)$$

where

$$\frac{\partial\rho}{\partial\varepsilon_{lm}}F_{il}F_{jm} = -\rho\delta_{ij} \qquad (60)$$

has been utilized and the inverse of Eq. (56) has been used to eliminate the polarization. In Eq. (59), the term containing the derivatives of ψ_e can be considered to be the elastic stress and the remainder of the right-hand side of the expression can be taken to be the electrostatic stress as in the usage of Landau and Lifschitz [6]. However, this does not imply that the Maxwell stress and the electrostatic stress are identical, since the Cauchy stress, σ_{ij}, can have a constitutive relationship that allows it to depend quadratically on the polarization independently of whatever electrical behavior is associated with the Maxwell stress. Because of the difficulty of measuring the Cauchy and Maxwell stresses separately, it is probably impossible to resolve this question experimentally. Therefore, we will focus on the elastic stress

$$\sigma_{ji}^e = F_{jk}\rho\frac{\partial\psi_e}{\partial\varepsilon_{kl}}F_{il} \qquad (61)$$

and dielectrostatic stress

$$\sigma_{ji}^{es} = -\frac{\kappa_o}{2}F_{jk}R_{mn}\frac{\partial\chi_{np}}{\partial\varepsilon_{kl}}R_{qp}F_{il}E_mE_q + \kappa_0\beta_{jikl}R_{km}\chi_{mn}R_{pn}E_lE_p$$

$$+ \kappa_o R_{jk}\chi_{kl}R_{ml}E_iE_m - \frac{\kappa_o}{2}R_{kl}\chi_{lm}R_{nm}E_kE_n\delta_{ji} + \hat{\sigma}_{ji}^M \qquad (62)$$

without attempting to separately identify the Cauchy and Maxwell stresses. Of course, the sum of the elastic and electrostatic stresses is equal to the sum of the Cauchy and Maxwell stresses and both sums are equal to the total true stress.

Linear, Isotropic, Conservative Materials at Infinitesimal Strain

For illustration, now consider the case where the strain is infinitesimal, so that to first order in the infinitesimal strains, e_{ij}, the susceptibility can be written [6]

$$\chi_{ij} = \widetilde{\chi}\delta_{ij} + \frac{\kappa_1}{\kappa_o}e_{ij} + \frac{\kappa_2}{\kappa_o}e_{kk}\delta_{ij} \qquad (63)$$

where κ_i are constants. With terms of higher order in strain neglected, the expression for the electric field in Eq. (56) becomes identical to Eq. (58), which is the usual relationship for isotropic dielectrics. When the strains are infinitesimal and higher order terms in strain are neglected, Eq. (59) shows that the total stress is given by

$$\sigma_{ij} + \sigma_{ij}^M = 2Ge_{ij} + \left(B - \frac{2G}{3}\right)e_{kk}\delta_{ij} + \left(\widetilde{\kappa} - \frac{\kappa_1}{2}\right)E_iE_j$$

$$- \frac{1}{2}(\widetilde{\kappa} + \kappa_2)E_kE_k\delta_{ij} \qquad (64)$$

as identified for this situation by Landau and Lifschitz [6]. In Eq. (64), G is the shear modulus, B is the bulk modulus, and $\widetilde{\kappa}$ is the dielectric permittivity of the unstrained material such that

$$\widetilde{\kappa} = (1 + \widetilde{\chi})\kappa_o \qquad (65)$$

so that the relationship between the electric displacement and the electric field during infinitesimal straining is given by

$$D_i = \widetilde{\kappa}E_i \qquad (66)$$

Note that, as observed by Landau and Lifschitz [6], the resulting total stress in Eq. (64) is symmetric and therefore the conservation of angular momentum is satisfied. Furthermore, consistent with the usage of Landau and Lifschitz [6], the elastic stress is

$$\sigma_{ij}^e = 2Ge_{ij} + \left(B - \frac{2G}{3}\right)e_{kk}\delta_{ij} \qquad (67)$$

and the electrostatic stress is

$$\sigma_{ij}^{es} = \left(\widetilde{\kappa} - \frac{\kappa_1}{2}\right)E_iE_j - \frac{1}{2}(\widetilde{\kappa} + \kappa_2)E_kE_k\delta_{ij} \qquad (68)$$

However, there is no implication that the electrostatic stress and the Maxwell stress are identical.

Conservative Materials that Remain Dielectrically Isotropic During Straining

Now return to the general case of arbitrarily large deformations but assume that, upon straining, the susceptibility remains isotropic and thus objectivity is assured. The susceptibility will be permitted to depend on the material density and thus is given by

$$\chi_{ij} = \chi(\rho)\delta_{ij} \qquad (69)$$

Note that the material density obeys $\rho = \rho_o/\text{Det}(U_{ij})$, where ρ_o is the density of the material in the reference state. Thus the dependence of the susceptibility on the density represents a contribution to the dependence of the Helmholtz free energy on the right

stretch tensor, as stated in Eq. (43). Given Eq. (69), Eq. (56) then simplifies to

$$E_i = \frac{P_i}{\kappa_0 \chi(\rho)} \tag{70}$$

The total true stress from Eq. (59) is then given by

$$\sigma_{ij} + \sigma_{ij}^M = F_{ik} \frac{1}{J} \frac{\partial \psi_e}{\partial \varepsilon_{kl}} F_{jl} + \kappa \left(E_i E_j - \frac{1}{2} E_k E_k \delta_{ij} \right) + \frac{1}{2} \rho \frac{d\chi}{d\rho} \kappa_0 E_k E_k \delta_{ij} \tag{71}$$

where κ is the possibly density-dependent dielectric permittivity given by

$$\kappa = (1 + \chi) \kappa_o \tag{72}$$

so that

$$D_i = \kappa E_i \tag{73}$$

To obtain Eq. (71), use was made of the fact that $\beta_{ijkl} E_k E_l = 0$. As expected, the total true stress is symmetric, reflecting the fact that the balance of angular momentum is satisfied. From Eqs. (70)–(73) it can be seen that through measurement of the elastic response and the dielectric susceptibility, possibly a function of density, the properties of this type of material can be fully characterized. The combined true stress can then be calculated for a given strain and electric field. When the susceptibility remains isotropic during deformation, the electrostatic stress thus becomes

$$\sigma_{ij}^{es} = \kappa \left(E_i E_j - \frac{1}{2} E_k E_k \delta_{ij} \right) + \frac{1}{2} \rho \frac{d\chi}{d\rho} \kappa_0 E_k E_k \delta_{ij} \tag{74}$$

Now we may explore the implications of different assumptions of how the dielectric permittivity depends on the material density. If we take the susceptibility to be independent of the density (or equivalently that the stored polarization energy per unit current volume is proportional to the square of the magnitude of the polarization vector but insensitive to the density, see Eq. (55)), the electrostatic stress from Eq. (74) becomes

$$\sigma_{ij}^{es} = \kappa \left(E_i E_j - \frac{1}{2} E_k E_k \delta_{ij} \right) \tag{75}$$

a form commonly seen in the literature as an expression for the Maxwell stress in a dielectric [6]. The body force per unit volume associated with the electrostatic stress given in Eq. (75) is

$$\frac{\partial \sigma_{ji}^{es}}{\partial x_j} = \kappa \frac{\partial E_j}{\partial x_j} E_i = q E_i \tag{76}$$

where Eqs. (2) and (73) have been used to establish the free charge density. Thus the body force associated with this particular electrostatic stress in this particular material is the load per unit volume due to the electric field acting on the free charges, reflecting the fact that an electric field applies a unit force on a unit charge [6]. This would seem to be a reasonable choice as a postulate for the electrical body force and it is often seen as such in the literature. However, there is no implication in our approach that the electrostatic stress in Eq. (75) is the Maxwell stress, nor that the electrical body force is the expression given in Eq. (76). Indeed, there is no need to adopt any particular postulate for the electrical body force and the Maxwell stress, since the total true stress is in equilibrium with the mechanical body force whatever the form of the electrical body force and the Maxwell stress and this is all that is needed for a complete formulation of the behavior of the material in response to electric field and strain. Furthermore, the electrical body force and the Maxwell stress cannot be determined from measurements of the total true, elastic, or electrostatic stress, neither in terms of their body forces nor their interface or surface tractions.

Another possible assumption for the susceptibility is that

$$\chi = \frac{\rho \chi_1}{\rho_o} \tag{77}$$

where χ_1 is a constant. This means that the polarization energy per unit mass of the dielectric is proportional to the square of the dipole moment per unit mass, i.e.,

$$\psi_P = \frac{\rho_o}{2 \kappa_0 \chi_1} \left(\frac{P_k}{\rho} \right) \left(\frac{P_k}{\rho} \right) \tag{78}$$

where ψ_P is the polarization energy per unit mass of the material and the term in parentheses in Eq. (78) is clearly the dipole moment per unit mass. This relationship for the stored energy of polarization is often assumed to be the correct one for a linear, isotropic, dielectric material [1,3]. The result for the electrostatic stress from Eq. (74) is then

$$\sigma_{ij}^{es} = \kappa E_i E_j - \frac{1}{2} \kappa_o \chi E_k E_k \delta_{ij} = P_i E_j + \hat{\sigma}_{ij}^M \tag{79}$$

This is another form that can be found in the literature [1,3] as an expression for the Maxwell stress. The body force per unit volume arising from this electrostatic stress in this material is given by

$$\frac{\partial \sigma_{ji}^{es}}{\partial x_j} = q E_i + P_j \frac{\partial E_i}{\partial x_j} \tag{80}$$

where the free charge has been obtained from Eqs. (2), (5), and (73). The expression in Eq. (80), as a postulate for the electrical body force per unit volume, has a provenance based on molecular models for the electrical response of a lattice [3] and therefore lends credence to Eq. (79) as an expression for the Maxwell stress in this particular material. In this case, the body force is the effect of the electric field acting on free charge plus the net force acting on induced dipoles because of a gradient in the electric field. The latter effect is due to the different force magnitude on the positive and negative charges of the dipole because of the different electric field acting upon them. However, as before there is no implication in our approach that the electrostatic stress in Eq. (79) is the Maxwell stress, nor that the electrical body force is the expression given in Eq. (80). As we have emphasized already, there is no need to adopt any particular postulate for the electrical body force and the Maxwell stress.

A third possibility is that the isotropic susceptibility is inversely proportional to ρ, which leads to an electrostatic stress given by

$$\sigma_{ij}^{es} = \kappa (E_i E_j - E_k E_k \delta_{ij}) + \frac{1}{2} \kappa_o E_k E_k \delta_{ij} = P_i E_j - P_k E_k \delta_{ij} + \hat{\sigma}_{ij}^M \tag{81}$$

The body force per unit volume from this particular expression for the electrostatic stress in this particular material is given by

$$\frac{\partial \sigma_{ji}^{es}}{\partial x_j} = q E_i - \frac{\partial P_j}{\partial x_i} E_j \tag{82}$$

where, as before, the free charge has been obtained from Eqs. (2), (5), and (73). These forms, Eqs. (81) and (82), as a postulate for the Maxwell stress and the electrical body force, seem to be absent from the literature. However, they can be rationalized as giving a body force that accounts for the electric field acting on free charges but that also provides for an effect in which the gradient of dipole density produces a force opposite to the direction of the gradient. The latter action can be understood as being due to the electric field in association with a surplus of positive charges over negative charges (or the opposite) at a given point in the material when there is a gradient of polarization. However, as before, there is no implication in our approach that the electrostatic stress in Eq. (81) is the Maxwell stress, nor that the electrical body force is the expression given in Eq. (82). As we have emphasized already, there is no need to adopt any particular postulate for the electrical body force and the Maxwell stress.

Quasi-Linear Piezoelectric Materials

To illustrate results in the context of a piezoelectric material, we write down a stored energy given by

$$\psi = \frac{1}{2\rho_o}\varepsilon_{ij}c_{ijkl}\varepsilon_{kl} - \frac{1}{\rho}\Pi_i h_{ijk}\varepsilon_{jk} + \frac{\rho_o}{2\kappa_o\rho^2}\Pi_i\chi_{ij}^{-1}\Pi_j \qquad (83)$$

where c_{ijkl} is a constant linear elasticity tensor at fixed polarization, h_{ijk} is a constant tensor of piezoelectric coefficients, and χ_{ij} is a constant susceptibility tensor denoting the response at zero strain. We note that in light of Eq. (43), ψ in Eq. (83) is objective. The choice of the form in Eq. (83) is somewhat arbitrary in the sense that the piezoelectric coefficients and the susceptibility tensor are fixed. By making such choices, we have eliminated the possibility that the susceptibility and the piezoelectric coefficients can experience a change of axiality during strain. On the other hand, in devising Eq. (83) we have utilized the notion that the stored energy per unit mass should depend on the polarization through the dipole moment per unit mass, an assumption that is fairly common in the literature.

Now we use Eqs. (47) and (48) and specialize to infinitesimal strain and rotation to obtain piezoelectric relationships in the form

$$E_i = -h_{ijk}e_{jk} + \frac{1}{\kappa_o}\chi_{ij}^{-1}P_j \qquad (84)$$

and

$$\sigma_{ij} + \sigma_{ij}^M = c_{ijkl}e_{kl} - h_{kij}P_k + \frac{1}{2}(P_iE_j + E_iP_j) + \hat{\sigma}_{ij}^M \qquad (85)$$

Note that the specialization to infinitesimal strain and rotation involves neglect of terms of higher order in strain and rotation. Complete linearization of the equations would eliminate the terms in Eq. (85) that are products of the polarization and electric field with the electric field. Note also that since the elasticity and piezoelectric coefficient tensors in Eq. (85) are symmetric on interchange of the subscripts i and j, the total true stress in Eq. (85) is symmetric, confirming that the balance of angular momentum is satisfied. As before, there is little mileage in attempting to separate the Cauchy and Maxwell stresses. Instead, we will identify the first two terms on the right-hand side of Eq. (85) to comprise the piezoelectric stress, σ_{ij}^p, and the remainder to be the electrostatic stress. Thus

$$\sigma_{ij}^p = c_{ijkl}e_{kl} - h_{kij}P_k \qquad (86)$$

and

$$\sigma_{ij}^{es} = \frac{1}{2}(P_iE_j + E_iP_j) + \hat{\sigma}_{ij}^M \qquad (87)$$

The body force per unit volume associated with this electrostatic stress in this material can be readily obtained by taking the spatial divergence of the expression in Eq. (87). However, the result is not particularly revealing. In any case, as before, there is no implication in our approach that the electrostatic stress in Eq. (87) is the Maxwell stress, nor that the electrical body force is its divergence. As we have emphasized already, there is no need to adopt any particular postulate for the electrical body force and the Maxwell stress. Instead we regard the electrostatic stress tensor in Eq. (87) simply to be an illustration of the consequences of certain constitutive assumptions embedded in Eq. (83).

Compliant Isotropic Dielectrics

Polymer dielectrics that have low shear moduli and are highly deformable to stretch ratios of order 10 have been introduced recently as actuator materials [5,11,16]. It is assumed that they are isotropic before straining and Kofod [11] has had success fitting their elastic response to a large strain, isotropic constitutive law for incompressible deformations developed by Ogden [12]. Pelrine et al. [5,16] and Kofod [11] also show that the dielectric permittivities of some of the polymers are unaffected or only slightly affected by straining, though there is evidence [17] that at least some polymers have susceptibilities that change significantly
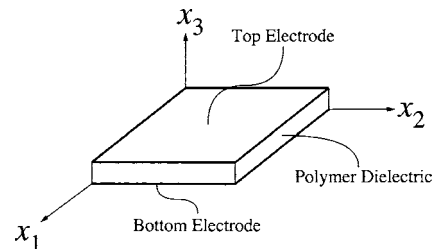
Fig. 2   A polymer dielectric actuator in the form of a slab with planar deformable electrodes

upon elastic deformation. The assumptions of incompressibility and insensitivity of the dielectric permittivity to straining are consistent with each other but not general. Therefore, we wish to study these highly deformable dielectrics within our formulation but when elastic dilatation is possible and when the permittivity depends on the strain. Ogden [14] has formulated an elasticity law for compressible materials that has a similar structure as his incompressible case; this could be used in our development and, presumably, it would model the stress-strain response of Kofod's [11] material well. However, the Ogden materials, whether incompressible or compressible, are somewhat complicated. Therefore, we will use a simpler elastic formulation, as we strive only to illustrate our approach to analyzing these materials and not to model them to a high degree of accuracy. Consequently, we will use a compressible form of the neo-Hookean material that is a variant of the Blatz-Ko [14,15] constitutive law. We will find that compressibility is essential to our ability to compute meaningful results in specific boundary value problems.

We choose also to present the results in terms of principal stresses and stretches as developed by Ogden [12,13], who also pointed out the great advantages of such an approach over formulating the results in terms of the strain tensor. In particular, it is directly applicable to the deformations of an actuator as shown in Fig. 2, which is a slab of polymer dielectric between two planar deformable electrodes [5,11,16]. For the polymer dielectric, we use a compressible generalization of the neo-Hookean formulation with an elastic strain energy density given by [14,15]

$$U_o = \frac{G}{2}(\lambda_i\lambda_i - 3J^{2/3}) + B(J - \ln J - 1) \qquad (88)$$

where λ_i are the principal stretch ratios in the three orthogonal principal directions of the deformation so that the ratio of volume in the current state to volume in the reference state $J = \lambda_1\lambda_2\lambda_3 = \rho_o/\rho$. Note that the form presented in Eq. (88) is not a Blatz-Ko material per se because the term containing the bulk modulus B differs from their form [14,15]. However, our form is consistent with Blatz-Ko usage with their parameter k chosen to be unity and is used in the spirit of simplicity. Given that the principal elastic true stresses are given by [12,13]

$$\sigma_i^e = \frac{\lambda_i}{J}\frac{\partial U_o}{\partial \lambda_i} \text{ (no sum on } i) \qquad (89)$$

they become

$$\sigma_i^e = G(\lambda_i^2 J^{-1} - J^{-1/3}) + B(1 - J^{-1}) \qquad (90)$$

Note that for polymers B/G is usually very large compared to unity and that this effect will keep the dilatational strains very small during deformations.

Now consider the condition of the actuator shown in Fig. 2 subject to an electric field $E_3 = E$ given by the voltage difference between the electrodes divided by the distance between the electrodes in the deformed state. The thickness of the actuator is very small compared to the in-plane dimensions so that a uniform electric field and stress state may be assumed in the dielectric and the effect of fringing fields at the edge can be ignored. Since we

assume that the actuator is constrained so that it does not rotate and that gravitational loads are negligible, the rotation tensor $R_{ij} = \delta_{ij}$ and the actuator remains rectilinear in shape during deformation. The susceptibility is isotropic before straining and remains at least orthotropic thereafter because of the lack of shear strain. The electrostatic stress from Eq. (62) in all situations for this actuator is then given by

$$\sigma_{11}^{es} = -\left(1 + \chi_{33} + \lambda_1^2 \frac{\partial \chi_{33}}{\partial \varepsilon_{11}}\right) \frac{\kappa_o}{2} E^2$$

$$\sigma_{22}^{es} = -\left(1 + \chi_{33} + \lambda_2^2 \frac{\partial \chi_{33}}{\partial \varepsilon_{22}}\right) \frac{\kappa_o}{2} E^2 \qquad (91)$$

$$\sigma_{33}^{es} = \left(1 + \chi_{33} - \lambda_3^2 \frac{\partial \chi_{33}}{\partial \varepsilon_{33}}\right) \frac{\kappa_o}{2} E^2$$

with the shear components zero.

The electrodes are assumed to be very thin and highly compliant, as they are in practice [5,11,16], so that they offer no constraint on the deformation of the dielectric. Thus, the only mechanical loads are the tractions T_{33} applied through the electrodes and that act parallel to the x_3 axis or, alternatively, the tractions T_{11} or T_{22} applied parallel to the plane of the electrodes. In addition, the stretching of the actuator parallel to the plane of the electrodes may also be controlled kinematically.

In the first deformation we consider, the principal stretch ratios λ_1 and λ_2 are equal to each other and, consequently, the elastic stresses from Eq. (85) are

$$\sigma_{11}^e = \sigma_{22}^e = G\left(\frac{1}{\lambda_3} - \frac{1}{\lambda_1^{2/3}\lambda_3^{1/3}}\right) + B\left(1 - \frac{1}{\lambda_1^2 \lambda_3}\right)$$

$$\sigma_{33}^e = G\left(\frac{\lambda_3}{\lambda_1^2} - \frac{1}{\lambda_1^{2/3}\lambda_3^{1/3}}\right) + B\left(1 - \frac{1}{\lambda_1^2 \lambda_3}\right) \qquad (92)$$

with the shear components all zero. The case we will address first is that where the principal stretch ratios λ_1 and λ_2 are controlled but within the constraint of being equal. The result, deduced from Eq. (7), that the traction T_{33} is balanced by the sum of the σ_{33} components of the elastic and electrostatic stresses provides

$$G\left(\frac{\lambda_3}{\lambda_1^2} - \frac{1}{\lambda_1^{2/3}\lambda_3^{1/3}}\right) + B\left(1 - \frac{1}{\lambda_1^2 \lambda_3}\right) + \frac{\kappa_o}{2} E^2\left(1 + \chi_{33} - \frac{\partial \chi_{33}}{\partial \varepsilon_{33}}\lambda_3^2\right) = T_{33} \qquad (93)$$

which, in principle, can be solved for λ_3. When χ_{33} is a function of strain, its derivative will be a function of λ_3 in a manner that can only be determined by experiment. Therefore, the order of Eq. (93), through the dependence of the susceptibility on the strain, is indeterminate without information from such experiments. Furthermore, if the potential difference between the electrodes is controlled, the electric field will be a function of λ_3, further complicating the equation order. Rather than pursue the solution of Eq. (93), we will consider the situation where the actuator, without mechanical load or electric field (i.e., in short circuit), is strained in the planar direction to a stretch ratio given by $\lambda_1 = \lambda_1^0$ and then an electric field applied along with blocking tractions sufficient to maintain the stretch ratios as they were before application of the field. The resulting through thickness strain before application of the electric field and traction conforms to a stretch $\lambda_3 = \lambda_3^o$ satisfying

$$G\left(\frac{\lambda_3^o}{(\lambda_1^o)^2} - \frac{1}{(\lambda_1^o)^{2/3}(\lambda_3^o)^{1/3}}\right) + B\left(1 - \frac{1}{(\lambda_1^o)^2 \lambda_3^o}\right) = 0 \qquad (94)$$

and the blocking tractions required to maintain these stretch ratios are

$$T_{33} = \left[1 + \chi_{33}(\lambda_1^o, \lambda_3^o) - (\lambda_3^o)^2 \frac{\partial \chi_{33}(\lambda_1^o, \lambda_3^o)}{\partial \varepsilon_{33}}\right] \frac{\kappa_o}{2} E^2 \qquad (95)$$

A special case of this result is where all the stretch ratios are fixed at unity. The blocking traction is then

$$T_{33} = \frac{1}{2}(\tilde{\kappa} - \kappa_1 - \kappa_2) E^2 \qquad (96)$$

where the relevant value of the partial derivative of the susceptibility component has been deduced from Eq. (53) and, of course, Eq. (96) agrees with the result that can be obtained from the theory of Landau and Lifschitz [6], since this case is also the limit of zero infinitesimal strains.

Note that compressibility of the material is essential to our ability to calculate a meaningful result for the problem just addressed. If the polymer were exactly incompressible, then the kinematic relationships would become

$$\lambda_3 = \frac{1}{\lambda_1^2} \qquad (97)$$

and

$$\lambda_3^o = \frac{1}{(\lambda_1^o)^2} \qquad (98)$$

Then once the stretch parallel to the electrodes of the actuator has been established, applications of electric field and mechanical stress are incapable of changing the through thickness strain. Thus, actuation in the sense of thickness change for the device becomes impossible and there is no need for a blocking traction to suppress the actuation. All that applications of electric field and tractions do is to modify the hydrostatic stress in the dielectric polymer without changing the strain.

Having demonstrated that one must be careful when using an incompressible material model, we now turn to a problem that can be successfully and accurately assessed with a volume-preserving constitutive law as an approximation to the true behavior when the bulk modulus, B, is much greater than the shear modulus, G. In this problem, the actuator is first stretched in the x_2 direction with no other tractions or constraints applied and under short circuit conditions so that E is zero. This stretch ratio is held fixed thereafter so that $\lambda_2 = \lambda_2^o$ and due to incompressibility at this stage

$$\lambda_1 = \lambda_3 = \frac{1}{\sqrt{\lambda_2^o}} \qquad (99)$$

An electric field is then applied and simultaneously a traction T_{11}. In the incompressible limit, the elastic stress is given by Eq. (90) with $J = 1$ and the term containing B is replaced by the negative of a pressure p. After the electric field is switched on, the stretch ratios must obey the relationship

$$\lambda_3 = \frac{1}{\lambda_1 \lambda_2^o} \qquad (100)$$

and the pressure is calculated from the condition that the sum of the elastic and electrostatic stress in the σ_{33} orientation is zero because there is no traction applied in that direction. Thus

$$p = G\left[\frac{1}{(\lambda_1 \lambda_2^o)^2} - 1\right] + \left[1 + \chi_{33} - \frac{1}{(\lambda_1 \lambda_2^o)^2} \frac{\partial \chi_{33}}{\partial \varepsilon_{33}}\right] \frac{\kappa_o}{2} E^2 \qquad (101)$$

and it follows that the balance of forces in the σ_{11} orientation provides

$$G\left[\lambda_1^2 - \frac{1}{(\lambda_1 \lambda_2^o)^2}\right] - \left\{1 + \chi_{33} + \frac{1}{2}\left[\lambda_1^2 \frac{\partial \chi_{33}}{\partial \varepsilon_{11}} - \frac{1}{(\lambda_1 \lambda_2^o)^2} \frac{\partial \chi_{33}}{\partial \varepsilon_{33}}\right]\right\} \kappa_o E^2$$
$$= T_{11} \qquad (102)$$

which can be solved for λ_1, in principle to determine the degree of actuation. As before, the solution is complicated by the fact that

the dependence of the dielectric permittivity on the strain (i.e., on the stretch ratios) must be determined from experiment. Furthermore, if a fixed potential difference is applied between the electrodes and a fixed load provides the traction T_{11}, both the traction and the electric field are functions of λ_1 as well, further complicating the solution. A simpler situation is where T_{11} is the blocking traction which resists the effect of the electric field and maintains the stretch ratios at the levels given in Eq. (99) that prevailed prior to the application of the field. The result from Eq. (101) for this blocking traction is then

$$T_{11} = -\left\{ 1 + \chi_{33} + \frac{1}{2\lambda_2^o}\left[\frac{\partial\chi_{33}}{\partial\varepsilon_{11}} - \frac{\partial\chi_{33}}{\partial\varepsilon_{33}} \right] \right\}\kappa_o E^2 \quad (103)$$

where the electric field and the traction are simple to establish because the thickness strain of the actuator is given by Eq. (99). The result in Eq. (103) predicts to within reasonable accuracy the results of an experiment carried out by Kofod [11]. It is notable that if the susceptibility is independent of strain, the traction predicted by Eq. (103) is compressive, reflecting the fact that application of the electric field will cause the actuator to try to expand parallel to the x_1 axis.

Discussion

The formulation presented in this paper is general and valid for materials in which the stress can be described by a local theory [7] and for which couple stresses or a director theory of materials response are not needed. In addition, the presentation is, we believe, free of unnecessary assumption in regard to the nature of the electrical body force and Maxwell stress in electrostatic systems. Instead, the expressions are presented in such a way that measurable behavior can be used to determine all the necessary functions and constants to completely describe the material constitutive behavior for reversible response. Indeed, it is not necessary to know the electrical body force or the Maxwell stress to obtain a usable constitutive law. Instead, given the assumptions of the paper regarding the constitutive law, measurements of the elastic response at zero electric field and of the material's dielectric permittivities as a function of strain will fully characterize the constitutive law. If the assumptions of the paper are incorrect in the sense that the constitutive law is more complex than that presented in Eq. (55), more extensive experiments will be needed to characterize the free energy as stated in Eq. (43), perhaps because the dielectric response involves a nonlinear dependence of the electric displacement on the electric field or that the elastic and electrostatic energy do not separate in the manner assumed in Eq. (55).

In regard to the terminology in the paper, quibbles can be raised, e.g., concerning what we call the Cauchy stress, which some workers regard to be what we have called the total true stress, i.e., the sum of what we call the elastic and electrostatic stress. However, this would be to focus on the wrong issues, because it is the formulation that is important, not the names of the terms. In any case we would not object to instructions to rename our entities, although we prefer our choice of names for what we have called the Cauchy, Maxwell, elastic, and electrostatic stresses.

Finally, we reemphasize the comments made in the Introduction that the formulation we have presented is desirable for use in a modern setting where feasibility of experiments, possible applications in finite element calculations, and the need for developments in the fracture mechanics of electromechanically loaded components have guided our thinking. We believe that our presentation is not only valid, but is attractive in the context of these needs. We have focused on conservative materials. However, all of what is developed up to and including Eq. (34) is correct for dissipative material behavior in the electrostatic limit, such as ferroelectric switching [18,19]. Thus, the formulation can and will be extended to dissipative materials in due course.

References

[1] Toupin, R. A., 1956, "The Elastic Dielectric," J. Rational Mech. Anal., **5**, pp. 849–914.
[2] Voigt, W., 1910, *Lehrbuch der Kristallphysik*, Teubner, Leipzig.
[3] Eringen, A. C., 1963, "On the Foundations of Electroelastostatics," Int. J. Eng. Sci., **1**, pp. 127–153.
[4] Landis, C. M., and McMeeking, R. M., 2000, "Modeling of Fracture in Ferroelectric Ceramics," *Smart Materials and Structures: Active Materials: Behavior and Mechanics*, Proceedings of SPIE, edited by C. S. Lynch, Vol. 3992, pp. 176–184.
[5] Pelrine, R., Kornbluh, R., Pei, Q., and Joseph, J., 2000, "High-Speed Electrically Actuated Elastomers with Over 100% Strain," Science, **287**, pp. 836–839.
[6] Landau, L. D., and Lifschitz, E. M., 1960, *Electrodynamics of Continuous Media*, Pergamon, Oxford.
[7] Malvern, L. E., 1969, *Introduction to the Mechanics of a Continuous Medium*, Prentice-Hall, Englewood Cliffs, NJ.
[8] McMeeking, R. M., 1998, "A Maxwell Stress for Material Interactions," J. Colloid Interface Sci., **199**, pp. 187–196.
[9] Coleman, B. D., and Noll, W., 1963, "The Thermodynamics of Elastic Materials with Heat Conduction and Viscosity," Arch. Ration. Mech. Anal., **13**, pp. 167–178.
[10] Carroll, M. M., 2004, "Derivatives of the Rotation and Stretch Tensors," Math. Mech. Solids, **9**, pp. 543–553.
[11] Kofod, G., 2001, "Dielectric Elastomer Actuators," Ph.D. thesis, Department of Chemistry, Technical University of Denmark, Riso Report No. Riso-R-1286(EN).
[12] Ogden, R. W., 1972, "Large Deformation Isotropic Elasticity—On the Correlation of Theory and Experiment for Incompressible Rubberlike Solids," Proc. R. Soc. London, Ser. A, **326**, pp. 565–584.
[13] Ogden, R. W., 1972, "Large Deformation Isotropic Elasticity: On the Correlation of Theory and Experiment for Compressible Rubberlike Solids," Proc. R. Soc. London, Ser. A, **328**, pp. 567–583.
[14] Blatz, P. J., and Ko, W. L., 1962, "Application of Finite Elasticity Theory to the Deformation of Rubbery Materials," Trans. Soc. Rheol., **6**, pp. 223–251.
[15] Blatz, P. J., 1963, "Application of Finite Elastic Theory to the Behavior of Rubberlike Materials," Rubber Sci. Technol., **36**, pp. 1459–1496.
[16] Pelrine, R., Kornbluh, R., Heydt, J. J. R., Pei, Q., and Chiba, S., 2000, "High-Field Deformation of Elastomeric Dielectrics for Actuators," Mater. Sci. Eng., C, **C11**, pp. 89–100.
[17] Shkel, Y. M., and Klingenberg, D. J., 1996, "Material Parameters for Electrostriction," J. Appl. Phys., **80**, pp. 4566–4572.
[18] Cocks, A. C. F., and McMeeking, R. M., 1999, "A Phenomenological Constitutive Law for the Behavior of Ferroelectric Ceramics," Ferroelectrics, **228**, pp. 219–228.
[19] Landis, C. M., 2002, "Fully Coupled, Multi-Axial, Symmetric Constitutive Laws for Polycrystalline Ferroelectric Ceramics," J. Mech. Phys. Solids, **50**, pp. 127–152.

Dynamic Fluid-Structure Interaction Analysis Using Boundary Finite Element Method–Finite Element Method

S. C. Fan[1]

e-mail: cfansc@ntu.edu.sg

S. M. Li

G. Y. Yu

Protective Technology Research Center,
School of Civil and Environmental Engineering,
Nanyang Technological University,
Nanyang Avenue,
Singapore 639798

In this paper, the boundary finite element method (BFEM) is applied to dynamic fluid-structure interaction problems. The BFEM is employed to model the infinite fluid medium, while the structure is modeled by the finite element method (FEM). The relationship between the fluid pressure and the fluid velocity corresponding to the scattered wave is derived from the acoustic modeling. The BFEM is suitable for both finite and infinite domains, and it has advantages over other numerical methods. The resulting system of equations is symmetric and has no singularity problems. Two numerical examples are presented to validate the accuracy and efficiency of BFEM-FEM coupling for fluid-structure interaction problems. [DOI: 10.1115/1.1940664]

1 Introduction

With the increasing human activities in the ocean, dynamic fluid-structure interaction (FSI) analysis has attracted more attentions in the past 50 years. Since analytical solutions are difficult to obtain for general FSI problems, numerical techniques for FSI problems are developed.

Amongst the prevailing numerical techniques, the finite element method (FEM) and the boundary element method (BEM) are commonly used. Very often, the FEM is employed to model the structure, while either the FEM or BEM is used to model the fluid domain. In those numerical simulations, FEM-to-FEM domain coupling procedures or FEM-to-BEM domain coupling procedures are necessary. Nevertheless, both FEM-FEM and FEM-BEM simulations exhibit some pitfalls.

When both the structure and fluid domain are bounded, FEM-FEM simulations have been shown to be efficient and able to yield reasonably accurate dynamic responses for FSI problems as reported in [1–9]. In those FEM-FEM analyses, various formulations were presented with regard to the acoustic fluid domain: Lagrangian fluid finite elements formulation in [1]; a generic displacement formulation in [2]; a mixed displacement-based finite element formulation in [3]; displacement and pressure mixed formulation in [4–6]; and the velocity potential formulation in [7–9]. All these formulations demonstrated their robustness in coupling with the conventional solid structural elements. According to their characteristics, those formulations can be classified into two major categories, namely, the displacement-based formulations and the potential-based formulations. In the displacement-based formulations, the fluid motion is described by the nodal displacements. The coupling responses between fluid and structure are ensured by equating the normal displacement components along the fluid-structure interface. This type of formulation is identical to the displacement formulation for continuum mechanics but with zero shear modulus at the interface. However, this pure displacement formulation for an inviscid fluid will exhibit spurious circulation modes. On the other hand, the potential-based formulations can avoid these spurious circulation modes because the fluid motion is represented by some form of scalar potential functions. In it, both the compatibility and the equilibrium conditions along the interface are explicitly enforced.

Moreover, when the FEM is employed to model an unbounded fluid domain, the size of the numerical model is limited and thereby the unbounded fluid domain has to be truncated but at a sufficiently far distance such that all major responses are not distorted. It requires the domain to be stretched beyond where the scattered waves can reach. By doing so, the effects of the scattered waves will be implicitly taken into account and no further complications will arise. However, in most practical cases, the FE mesh discretization for such a huge fluid domain will become prohibitively expensive in terms of computational cost. Practically, a compromised solution is to truncate the unbounded fluid domain at a reasonable not-too-far distance away from the structure. Consequently, some scattered waves will reach the truncated (artificial) boundary before the termination of the analysis. It results in artificial waves reflecting back into the fluid mesh and contaminating the responses. To overcome or minimize these undesired wave reflections from the artificial boundary, an alternative is to put in place a kind of nonradiating boundary that is "transparent" to the scattered waves.

In modeling a fluid domain with nonradiating boundary, many researchers have been heralding this approach without using the FEM. Mindlin and Bleich [10] are among the pioneer endeavors. They developed an early-time approximation technique, namely plane wave approximation (PWA). Successful applications of the PWA technique for the analyses of submerged spherical and spheroidal shells were reported by DiMaggio et al. [11], Hamdan and Dowling [12], and Fan et al. [13]. Very often, the early-time responses obtained from the PWA method agree well with the exact solution, but it is not so for the late-time responses. Geers [14] developed an analytical method based on virtual mass approximation (VMA) of the infinite acoustic medium. The validity of VMA was illustrated through a study of the elastic response of a cylindrical shell excited by a transient acoustic wave. Compared to the results obtained from PWA, the VMA results demonstrated its superior performance, in particular the late-time behaviors and the low-frequency response. By superimposing PWA and VMA, Ranlet et al. [15] developed the doubly asymptotic approximation (DAA), which was used to model the infinite fluid medium, while

modal analysis was employed for the structure. The DAA method was proved to be accurate to model both early- and late-time behavior. Zilliacus [16] used the DAA method to analyze the response of a submerged fluid-filled cylinder subjected to an incident plane step wave. It is worth noting that, in the DAA formulation, the mass matrix for the fluid media is fully populated. More comprehensive reviews of the nonradiation boundary can be found in [17].

In solving the general FSI problems, the exact or analytical nonradiating boundary is difficult to implement. With the emergence of BEM which has been gradually recognized to have advantages in modeling infinite domain, FEM-BEM coupling procedures were developed. Estorff and Antes [18], Czygan and Estorff [19], and Yu et al. [20] demonstrated the advantages of using FEM-BEM procedures for FSI problems. In their numerical simulations, no artificial boundaries or wave reflections were present. Moreover, the FEM-BEM procedures have computational advantages derived from its BEM formulation, which reduces the spatial dimension by one. However, it inherited the disadvantages of BEM which demands for prerequisite fundamental solutions, which can be very complicated or may not be available. Furthermore, it leads to asymmetric coefficient matrices in its formulation. Not only does it increase the requirement for memory storage, but it also compromises (if not nullifies) its other computational efficiency.

Recently, Wolf and Song developed the boundary finite element method (BFEM) [21–23], which was based on the geometry similarity and finite element cell concept [24]. Contrary to the domain-based methods, the BFEM formulation needs discretization only along the boundary. Therefore, it is suitable for both finite and infinite domains. In addition, the BFEM formulation leads to symmetric matrices and encounters no singularity problem, which often emerges in the BEM. Wolf and Song [21–23] successfully applied the BFEM to solve some infinite soil-structure interaction problems and also some bounded solid problems. They demonstrated that the BFEM is accurate for both bounded and unbounded domains.

In this paper, the BFEM is applied to model the acoustic problems in an infinite fluid medium. Through coupling of BFEM-FEM procedures, the dynamic response of a submerged cylinder is analyzed. In comparison with the PWA, VMA, and DAA, the present formulation using the BFEM does not impose any restriction on the structure and the incident wave as well. Hence, it is applicable for both early- and late-time response analyses. In comparison with the FEM-FEM and FEM-BEM coupling procedures, the present formulation has the advantages that no artificial boundaries are needed, no singularity is encountered, no asymmetric matrix appears, and no fundamental solution is required. The effectiveness of the coupling BFEM-FEM procedure will be shown through numerical examples.

2 BFEM-FEM Coupling Formulation

In the current study, the BFEM is used to model the unbounded acoustic fluid medium, while the FEM is used to model the structure.

2.1 FEM Model for the Structure. In modeling the structure, the mass matrix $\mathbf{M}$, the damping matrix $\mathbf{C}$, and the stiffness matrix $\mathbf{K}$ can be treated in the standard manner according to the FE discretization procedures. For 2D thin hollow cylindrical problems, simple two-node beam elements are used.

2.2 BFEM Model for the Unbounded Acoustic Fluid Medium

2.2.1 Acoustic Approximation. When a structure submerged in an unbounded fluid medium is subjected to incident waves, the major concern is the integrity of the structure, in particular its strength adequacy against the dynamic pressure acting on the structural surface (wet surface). To begin, the total fluid pressure

and velocity along the wet surface may be considered of composing two components: a free-field component and a scattered component [13]:

$$\mathbf{p}_{inf} = \mathbf{p}_{ff} + \mathbf{p}_{sc} \tag{1}$$

$$\mathbf{v}_{inf} = \mathbf{v}_{ff} + \mathbf{v}_{sc} \tag{2}$$

where $\mathbf{p}_{inf}$ and $\mathbf{v}_{inf}$ are the total fluid pressure and normal velocity along the wet surface, $\mathbf{p}_{ff}$ and $\mathbf{v}_{ff}$ are the corresponding component fluid pressure and normal velocity but are caused by the incident wave in the absence of the structure (usually referred as the free-field response), and $\mathbf{p}_{sc}$ and $\mathbf{v}_{sc}$ are the corresponding component fluid pressure and normal velocity corresponding to the scattered wave, which are the difference between the total and free-field solutions. The relationship between $\mathbf{p}_{sc}$ and $\mathbf{v}_{sc}$ is described in the next section. The fluid medium is assumed to have a linear constitutive relationship, which excludes the effect due to cavitation, and thereby the free-field pressure and velocities will be the same as the incident waves, i.e.,

$$\mathbf{p}_{ff} = \mathbf{p}_{inc} \tag{3a}$$

$$\mathbf{v}_{ff} = \mathbf{v}_{inc} \tag{3b}$$

where $\mathbf{p}_{inc}$ and $\mathbf{v}_{inc}$ are the fluid pressure along the wet surface and the corresponding normal velocity caused by the incident wave.

2.2.2 Pressure-Velocity Relationship for the Scattered Wave. In PWA approximation, the pressure-velocity relationship corresponding to scattered wave can be described as

$$\mathbf{p}_{sc} = \rho c \mathbf{v}_{sc} \tag{4a}$$

While in the DAA, the relationship can be expressed as

$$\frac{1}{\rho c}(\dot{\mathbf{p}}_{inf} - \dot{\mathbf{p}}_{ff}) - \mathbf{M}^{-1}\mathbf{A}(\mathbf{p}_{inf} - \mathbf{p}_{ff}) = \mathbf{a}_{inf} - \mathbf{a}_{ff} \tag{4b}$$

where an over-dot denotes differentiation with respect to time, $\mathbf{M}^{-1}$ is the inverse of the fluid-added mass matrix, $\mathbf{A}$ denotes the diagonal matrix which converts diagonal pressures to forces, and $\mathbf{a}_{inf}$, $\mathbf{a}_{ff}$ denotes the accelerations on the wet surface. ρ and c denote fluid density and wave speed in fluid, respectively. In the current study, another form of the relationship based on the BFEM and acoustic approximation is developed and described in detail in the next section.

2.2.3 BFEM Formulation for the Scattered Wave

a. Basic description. Basically, the BFEM describes the dynamic behavior of an unbounded medium through a dynamic stiffness (or mass) matrix in the frequency domain relating the displacement (or potential) amplitudes in the degrees of freedom on the boundary to the corresponding force (or velocity) amplitudes. By discretizing the domain into sectors radiating from a single center (namely scale center), the geometry of the sectors can be conveniently described in a transformed coordinate system (in which one ordinate radiates outward from the scale center, while others run along the boundary curve/surface). More precisely, the unbounded medium lies in a semi-infinite domain (see Fig. 1(a)). Along the radial direction, the near side is bounded while the far side is unbounded. The derivation of the dynamic stiffness/mass matrix is based on a "cloning" technique, in which the small differential (w) between the two similar semi-bounded sectors (see Fig. 1(b)) is taken to the analytical limit, zero. Consider the infinitesimal finite-element cell i-j-i'-j' (see Fig. 1(b)) which lies on the boundary of the semi-infinite fluid domain. Note that the near face i-j is parallel to its cloned surface i'-j', and the two side faces O_i and O_j are both originated from the scale center O. The governing equation relating the potential and the velocity can be written as
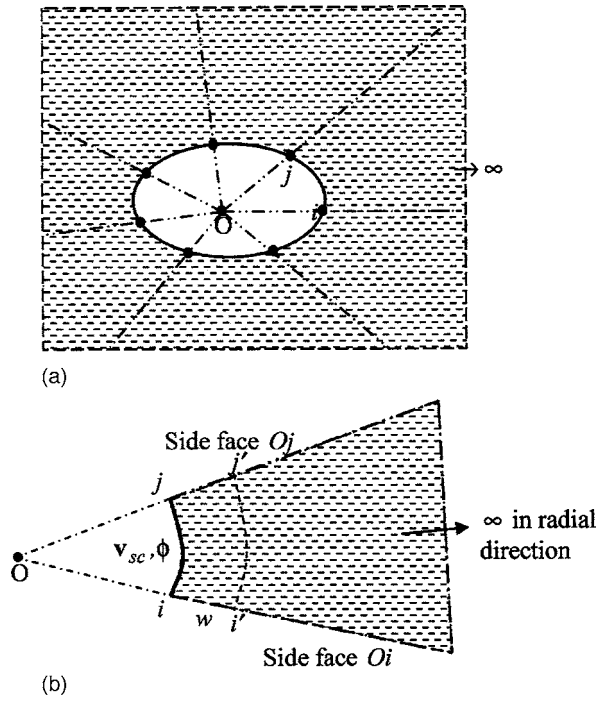
Fig. 1 (*a*) **Sectorial discretization of an unbounded domain and** (*b*) **a typical BEFM element with differential width** *w* **lying on the boundary of a semi-infinite domain**

$$\mathbf{V}_{sc}(t) = \int_0^t \mathbf{M}^\infty(t - \tau)\ddot{\boldsymbol{\varphi}}(\tau)d\tau \qquad (5)$$

where $\boldsymbol{\varphi}(t)$ denotes a velocity potential vector composed of nodal velocity potential $\phi(t)$ for scattered wave on the wet surface, $\mathbf{M}^\infty(t)$ denotes the mass matrix of the unbounded fluid medium, and $\ddot{\boldsymbol{\varphi}}(t)$ denotes the second derivative of $\boldsymbol{\varphi}(t)$ with respect to time. Note that in Eq. (5), the scattered velocity $\mathbf{V}_{sc}$ and the velocity potential $\boldsymbol{\varphi}(t)$ along the wet surface are variables. The matrix $\mathbf{M}^\infty(t)$ depends only on the geometry of the wet surface and is independent of the dynamic response of the structure and the fluid. Hence, it can be obtained before solving the dynamic response equation. Wolf and Song [21–23] showed the detail derivation of the dynamic mass matrix and no duplication is given here. In brief, the first step is to establish the integral form of the governing equation in frequency domain, and then, by taking the differential width w as the analytical limit to zero, it yields a consistent infinitesimal finite-element cell (IFEC) equation in frequency domain. Appling the inverse Fourier transformation to this frequency equation leads to an equivalent IFEC equation in time domain. Subsequently, the mass matrix $\mathbf{M}^\infty(t)$ can be obtained by solving the IFEC equation in time domain. (More recently, the same equation was successfully derived using different approaches by Song and Wolf [25] and Deeks and Wolf [26], respectively.) The scattered wave along the wet surface $\mathbf{V}_{sc}(t)$ in Eq. (5) can be discretized through conventional finite element technique, i.e.,

$$\mathbf{V}_{sc}(t) = \sum_e \int_\Gamma N_f v_{sc}(t)d\Gamma \qquad (6)$$

where v_{sc} denotes the velocity of the scattered wave normal to the wet surface; Γ denotes the wet surface; $\mathbf{N}_f$ is the shape function for the fluid element; and Σ_e denotes assemblage of all fluid elements along the wet surface.

b. Formulation for the scattered wave. Now, consider the temporal discretization of Eq. (5). Within each time step Δt, the mass $\mathbf{M}^\infty(t)$ is assumed constant. Hence, Eq. (5) can be rewritten as

$$\mathbf{V}_{sc}(n\Delta t) = \sum_{j=1}^n \mathbf{M}^\infty[(n - j + 1)\Delta t]\int_{(j-1)\Delta t}^{j\Delta t} \ddot{\boldsymbol{\varphi}}(t)dt$$

$$= \sum_{j=1}^n \mathbf{M}^\infty(n - j + 1)\dot{\boldsymbol{\varphi}}|_{(j-1)\Delta t}^{j\Delta t} \qquad (7a)$$

at the nth time step ($t=n\Delta t$). It can be written in a simpler index form as follows:

$$\mathbf{V}_{sc}^n = \sum_{j=1}^n \mathbf{M}_{n-j+1}^\infty(\dot{\boldsymbol{\varphi}}_j - \dot{\boldsymbol{\varphi}}_{j-1}) \qquad (7b)$$

where $\mathbf{V}_{sc}^n = \mathbf{V}_{sc}(n\Delta t)$, $\mathbf{M}_{n-j+1}^\infty = \mathbf{M}^\infty[(n-j+1)\Delta t]$, and $\dot{\boldsymbol{\varphi}}_j = \dot{\boldsymbol{\varphi}}|_{j\Delta t}$. Rearranging Eq. (7b) to make $\mathbf{M}_1^\infty\dot{\boldsymbol{\varphi}}_n$ appeared on the left-hand side, we have

$$\mathbf{M}_1^\infty\dot{\boldsymbol{\varphi}}_n = V_{sc}^n - \sum_{j=1}^{n-1}(\mathbf{M}_{n-j+1}^\infty - \mathbf{M}_{n-j}^\infty)\dot{\boldsymbol{\varphi}}_j + \mathbf{M}_n^\infty\dot{\boldsymbol{\varphi}}_0 \qquad (7c)$$

Given the following initial conditions,

$$\boldsymbol{\varphi}(0) = \mathbf{0} \qquad (8a)$$

$$\dot{\boldsymbol{\varphi}}(0) = \mathbf{0} \qquad (8b)$$

substituting $\mathbf{p}_{sc} = -\rho\dot{\boldsymbol{\varphi}}$ and Eq. (8b) into Eq. (7c) yields

$$\mathbf{M}_1^\infty\mathbf{p}_{sc}^n = -\rho\mathbf{V}_{sc}^n - \sum_{j=1}^{n-1}(\mathbf{M}_{n-j+1}^\infty - \mathbf{M}_{n-j}^\infty)\mathbf{p}_{sc}^j \qquad (9)$$

where $\mathbf{p}_{sc}^n = -\rho\dot{\boldsymbol{\varphi}}_n$, $\mathbf{p}_{sc}$ denotes the fluid pressure corresponding to the scattered wave, and $\mathbf{p}_{sc}^j = \mathbf{p}_{sc}(j\Delta t)$. Note that Eq. (9) describes the relationship between the pressure $\mathbf{p}_{sc}$ and the corresponding normal velocity $\mathbf{v}_{sc}$.

2.3 BFEM-FEM Coupling. The FEM formulation for the structure can be written as

$$\mathbf{Ma} + \mathbf{Cv} + \mathbf{Kd} = \mathbf{F}_{ext} + \mathbf{F}_{inf} \qquad (10)$$

where $\mathbf{F}_{ext}$ is the sum of body and traction forces; $\mathbf{F}_{inf}$ is the force derived from the unbounded fluid medium; $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are respectively the mass, damping, and stiffness matrices derived from the structure; and $\mathbf{a}$, $\mathbf{v}$, and $\mathbf{d}$ denote the structure's acceleration, velocity, and displacement vectors, respectively. The derivation of Eq. (10) can be obtained following the standard FE procedures. What follows will elaborate on the derivation of the last term in Eq. (10).

By virtue of the principle of virtual work, the nodal force $\mathbf{F}_{inf}$ on the wet surface can be expressed in terms of the shape function $\mathbf{N}$ for the structure and the total pressure ($=p_{sc}+p_{ff}$) as follows:

$$\mathbf{F}_{inf} = -\int_\Gamma \mathbf{N}^T(p_{sc} + p_{ff})d\Gamma = -\int_\Gamma \mathbf{N}^T\mathbf{N}_p\mathbf{p}_{sc}d\Gamma - \int_\Gamma \mathbf{N}^T p_{ff}d\Gamma$$

$$= -\mathbf{F}_{sc} - \mathbf{F}_{ff} \qquad (11)$$

where Γ denotes the wet surface, $\mathbf{F}_{ff} = \int_\Gamma \mathbf{N}^T p_{ff}d\Gamma$ and $\mathbf{F}_{sc} = \int_\Gamma \mathbf{N}^T\mathbf{N}_p\mathbf{p}_{sc}d\Gamma$. Note that $\mathbf{N}$ is the shape function for the structural element, while the shape function for scattered wave can be different, say $\mathbf{N}_p$, i.e.,

$$p_{sc} = \mathbf{N}_p\mathbf{p}_{sc} \qquad (12)$$

By substituting Eq. (11) into Eq. (10), the governing dynamic equation for the structure-infinite fluid system can be obtained, i.e.,

$$\mathbf{Ma} + \mathbf{Cv} + \mathbf{Kd} = -\mathbf{F}_{ff} - \mathbf{F}_{sc} + \mathbf{F}_{ext} \qquad (13)$$

In Eq. (13), the first term $\mathbf{F}_{ff}$ on the right-hand side is a function of the incident wave p_{ff}. It can be evaluated easily. What follows are the details for deriving the second term $\mathbf{F}_{sc}$ on the right-hand side.

Consider the conditions along the wet surface. Two conditions should be enforced. First, kinematic continuity across the wet surface requires the normal velocity of the structure to be identical to that of the fluid, i.e.,

$$v_{inf} = v_n \qquad (14a)$$

Second, the corresponding dynamic compatibility demands the compressive traction on the structural surface to be equal to the fluid pressure, i.e.,

$$p = -p_{inf} \qquad (14b)$$

Substituting Eq. (14a) into Eq. (2), and then the result into Eq. (6) and subsequently into Eq. (9) leads to

$$\mathbf{M}_1^{\infty}\mathbf{p}_{sc}^n = -\rho\sum_e \int_\Gamma \mathbf{N}_f[v_n(t) - v_{ff}(t)]d\Gamma - \sum_{j=1}^{n-1}(\mathbf{M}_{n-j+1}^{\infty} - \mathbf{M}_{n-j}^{\infty})\mathbf{p}_{sc}^j \qquad (15)$$

Note that once $\mathbf{p}_{sc}$ is determined, the second term $\mathbf{F}_{sc}$ on the right-hand side of Eq. (13) can be obtained.

Now, discretize Eq. (13) in time domain using Newmark's time-integration scheme. We have

$$\left(\mathbf{K} + \frac{1}{\alpha\Delta t^2}\mathbf{M} + \frac{\delta}{\alpha\Delta t}\mathbf{C}\right)\mathbf{d}^{t+\Delta t}$$

$$= {}^{t+\Delta t}\mathbf{F}_{ext} - {}^{t+\Delta t}\mathbf{F}_{ff} - {}^{t+\Delta t}\mathbf{F}_{sc} + \mathbf{M}\left[\frac{1}{\alpha\Delta t^2}\mathbf{d}^t + \frac{1}{\alpha\Delta t}\mathbf{v}^t + \left(\frac{1}{2\alpha} - 1\right)\mathbf{a}^t\right]$$

$$+ \mathbf{C}\left[\frac{\delta}{\alpha\Delta t}\mathbf{d}^t + \left(\frac{\delta}{\alpha} - 1\right)\mathbf{v}^t + \left(\frac{\delta}{2\alpha} - 1\right)\Delta t\mathbf{a}^t\right] \qquad (16)$$

where ${}^{t+\Delta t}\mathbf{F}_{sc} = \int_\Gamma \mathbf{N}^T\mathbf{N}_p\mathbf{p}_{sc}^{t+\Delta t}d\Gamma$, $\alpha = 0.25$, and $\delta = 0.5$. Assume $n\Delta t = t + \Delta t$. Hence, Eq. (15) can be rewritten as

$$\mathbf{M}_1^{\infty}\mathbf{p}_{sc}^n = -\rho\sum_e \int_\Gamma \mathbf{N}_f[v_n^t(t) - v_{ff}^{t+\Delta t}(t)]d\Gamma - \sum_{j=1}^{n-1}(\mathbf{M}_{n-j+1}^{\infty} - \mathbf{M}_{n-j}^{\infty})\mathbf{p}_{sc}^j \qquad (17)$$

in which the velocity $v_n^{t+\Delta t}$ of the current time step is approximately set equal to the velocity v_n^t of the previous time step. However, in some cases, it is desirable to include that term into the global iteration scheme, i.e.,

$$\left(\mathbf{K} + \frac{1}{\alpha\Delta t^2}\mathbf{M} + \frac{\delta}{\alpha\Delta t}\mathbf{C}\right)^{t+\Delta t}\mathbf{d}^j$$

$$= {}^{t+\Delta t}\mathbf{F}_{ext} - {}^{t+\Delta t}\mathbf{F}_{ff} - {}^{t+\Delta t}\mathbf{F}_{sc}({}^{t+\Delta t}\mathbf{v}_n^{j-1} - {}^{t+\Delta t}\mathbf{v}_{ff})$$

$$+ \mathbf{M}\left[\frac{1}{\alpha\Delta t^2}{}^t\mathbf{d} + \frac{1}{\alpha\Delta t}{}^t\mathbf{v} + \left(\frac{1}{2\alpha} - 1\right){}^t\mathbf{a}\right]$$

$$+ \mathbf{C}\left[\frac{\delta}{\alpha\Delta t}{}^t\mathbf{d} + \left(\frac{\delta}{\alpha} - 1\right){}^t\mathbf{v} + \left(\frac{\delta}{2\alpha} - 1\right)\Delta t^t\mathbf{a}\right] \qquad (18)$$

where j denotes the jth iteration within a time step. The term $v_{ff}(t)$ can be obtained via other analytical solutions. For example, in shock-wave analysis, Lamb [27] gives the following explicit expression for the incident fluid pressure along the wet surface (see Fig. 8):
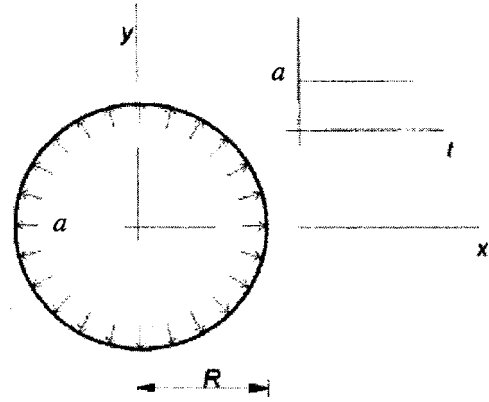

Fig. 2 A cylindrical cavity subjected to a suddenly applied acceleration

$$p_{inc}(x,t) = p_0 H\left(t - \frac{R-x}{c}\right) = p_0 H\left(t - \frac{R - r\cos\gamma}{c}\right) \qquad (19)$$

where H is a Heaviside step function, and p_0 is the magnitude of the pressure at wave front. r, γ, and R are as shown in Fig. 8. The incident fluid radial velocity is

$$v_{ff} = v_r = -\frac{p_{inc}}{\rho c}\cos\gamma \qquad (20)$$

3 Numerical Examples

To validate the present BFEM formulation and the pressure-velocity relationship corresponding to scattered wave based on the BFEM, three 2D cases for submerged structures subjected to internal pressures or external shock loading are considered.

3.1 A Cylindrical Cavity Subjected to a Suddenly Applied Acceleration. Figure 2 shows a cylindrical cavity of radius R in an infinite fluid medium. At $t=0$, boundary acceleration a is suddenly applied and then kept constant until the end of the analysis. The objective of this analysis is to check the accuracy of the "mass" matrix $\mathbf{M}^{\infty}(t)$ for the infinite fluid medium. The surrounding fluid medium is modeled by the BFEM using 8 or 32 elements. The results are plotted in Fig. 3. The time axis is normalized with respect to R/c, while the pressure axis is normalized with respect to ρaR. The same problem was considered by Yu et al. [28] using the FEM-BEM coupling procedures. The FEM-BEM and standard BEM results are also plotted in Fig. 3 and compared with the current result. From Fig. 3, one can see that the 8-element results are equal to the 32-element results, and both are almost identical to the FEM-BEM and BEM results. It shows the efficiency and accuracy of the present formulation.
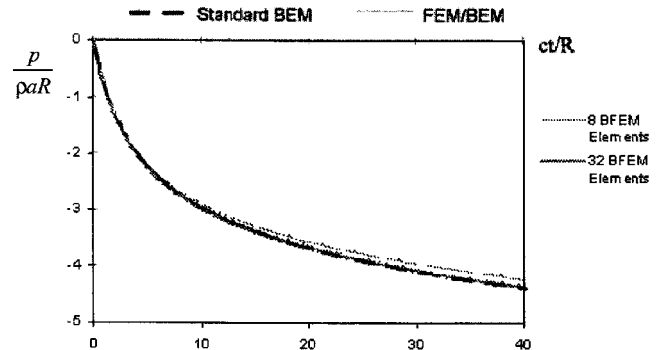

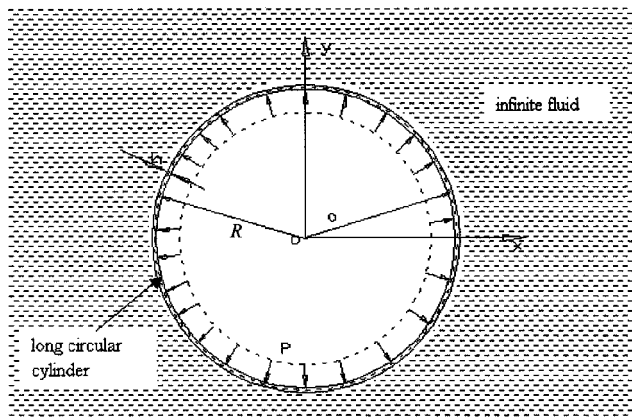Fig. 3 Pressure of the cavity boundary

Fig. 4 Geometry of a long cylindrical shell subjected to an internal pressure

Fig. 6 Matching discretization meshes for the cylinder and the fluid boundary

3.2 Submerged Long Cylindrical Shell Subjected to Internal Pressure. Consider a thin, elastic, and infinitely long cylindrical shell, submerged in water. It is subjected to a suddenly applied uniform outward pressure (see Figs. 4 and 5). A ramp for a very short period $(0.5 \times 10^{-4}$ s) is put in place in order to avoid numerical difficulties. The mean radius of the cylinder is $R=0.2$ m and the thickness of the wall is $h=0.006$ m. The material properties for the steel cylinder are Young's modulus $E_s=210 \times 10^6$ kN/m^2, density $\rho_s=7.8 \times 10^3$ kg/m^3, and Poisson ratio $\nu=0.3$. The density of fluid is $\rho=1.0 \times 10^3$ kg/m^3 and the sound speed in fluid $c=1500$ m/s.

This problem is a simple 2D axisymmetric problem. The steel cylindrical shell is discretized into 16 two-node beam elements. Along the wet surface, the same discretization mesh is applied to the fluid boundary. The 16 BFEM elements match the 16 structural beam elements side by side (see Fig. 6). In the analysis, the time step is set to 0.013333 ms. The results of the dynamic response are shown in Fig. 7, which shows the time history of the outward displacement (positive). The displacement is normalized with respect to the corresponding static displacement d_s at mid-surface $(R=0.2$ m), i.e.,

$$d_s = \frac{R_i^2 P_0}{E_s(R_e^2 - R_i^2)} \left(\frac{(1+\nu)R_e^2}{R} + (1+\nu)R \right) \quad (21)$$

where R_i, R_e are the internal and external radii of the cylindrical shell, respectively. The analytical results based on 1D plane-wave approximation (PWA) are plotted in Fig. 7 for comparison. One can see from Fig. 7 that the present results are in good accord with the PWA results, which were obtained by solving a second-order differential equation at early time. Note that the PWA approach can only yield relatively accurate results for early time.
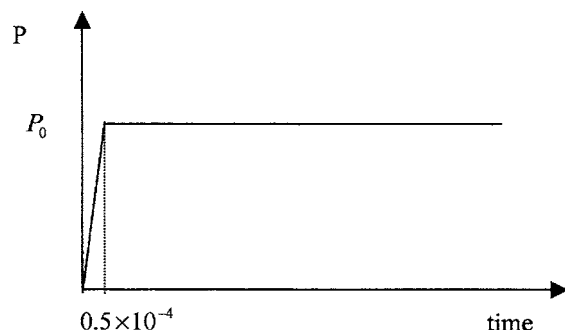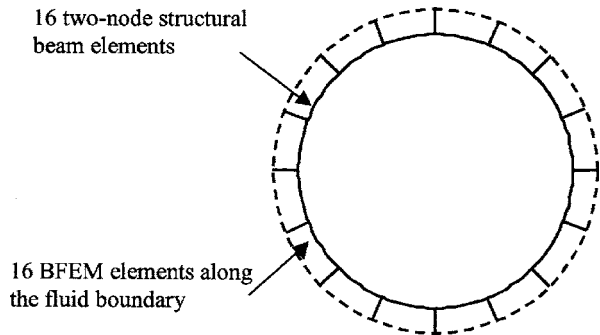
3.3 Submerged Long Cylindrical Shell Subjected to Plane Wave. Consider an elastic, infinitely long, cylindrical shell subjected to a step plane acoustic wave. The geometry of the cylindrical shell is shown in Fig. 8. This study is to check the accuracy of the present BFEM formulation against the benchmark solutions. The geometric and material properties of the cylindrical shell and the surrounding fluid medium are thickness of cylindrical wall $h=0.006$ m, mean radius of cylindrical wall $R=0.2$ m, density of steel cylindrical wall $\rho_s=7.8 \times 10^3$ kg/m^3, Young's modulus of the steel cylindrical wall $E_s=210 \times 10^6$ kN/m^2, Poisson ratio $\nu=0.3$, density of fluid $\rho=1.0 \times 10^3$ kg/m^3, and sound
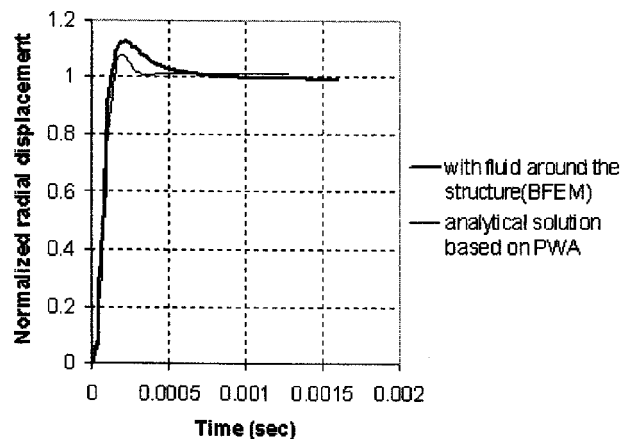

Fig. 7 Dynamic response of a cylindrical shell

Fig. 5 Loading conditions for a cylindrical shell
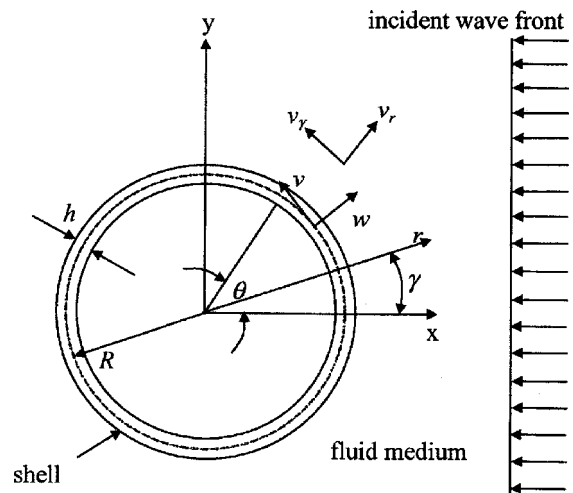

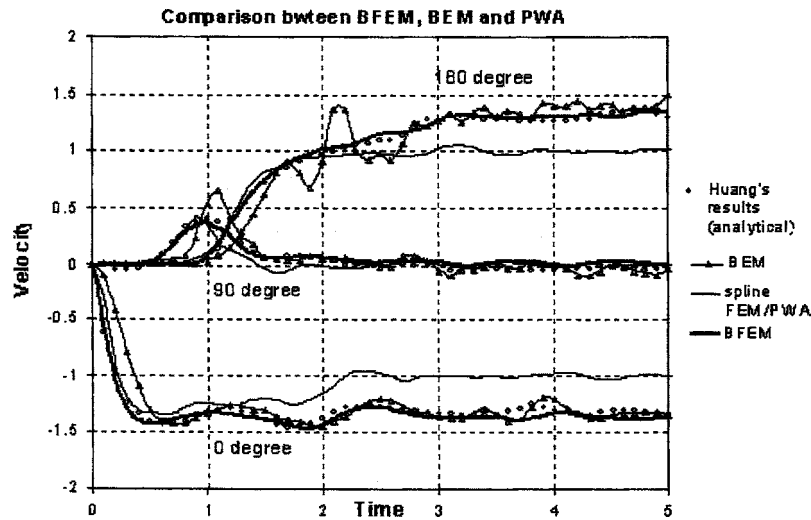Fig. 8 Geometry of an infinite cylinder

Fig. 9 Radial velocity of a cylinder

speed in fluid: $c = 1500$ m/s.

The wall of the shell is discretized into 32 two-node beam elements, while the fluid boundary is matched by 32 BFEM elements (in the similar manner as in Fig. 6). The time step chosen for the analysis is 0.002667 ms. The magnitude of the incident pressure wave impinging on the structure is taken to be equal to ρc^2. The same problem was investigated by Huang [29] and reported by Zilliacus [16], who obtained the velocity history using an analytical approach. The same problem was also analyzed by Fan et al. [13], who modeled the shell using 2×12 nine-node spline shell elements and modeled the fluid using the plane wave approximation. Plane-strain conditions were imposed along the axis of the cylinder. Recently, Yu et al. [20] also analyzed the same problem using the BEM-FEM coupling technique.

Results obtained from the present BFEM formulation are compared with solutions by others. Figure 9 shows the dimensionless radial velocity history at different locations ($\theta = 0°, 90°, 180°$). The velocity is normalized with respect to sound speed c, while the time is normalized with respect to R/c. From Fig. 9, one can see that the present results are in good accord in all time with analytical solution [28], and also in good agreement in early time with PWA solution, which is known to be relatively accurate in early time. On the other hand, the FEM-BEM results are also in fairly good agreement but exhibit random undulations around the analytical solution, particularly during late time, and the peak value is delayed. It demonstrates that the present BFEM formulation can yield more accurate results than the PWA and the BEM, in particular at late time. Note that the true velocities at $\theta = 0°, 180°$ should approach 1.377 at late time [28]. From Fig. 9, one can see that the BFEM results approach that true value of 1.377 at late time. Compared to the DAA results obtained by Zilliacus [16] (not shown in Fig. 9 for clarity), BFEM results are closer to the analytical solution and exhibit more subdued oscillations. Note that the results shown in Fig. 9 are obtained using the iterative Newmark scheme (Eq. (18)), which has iterations within each time step. The effects of using the Newmark scheme without iterations are also studied. The results shown in Fig. 10 suggest that the noniterative Newmark scheme leads to slightly inferior results. In order to study the convergence of the current BFEM formulation, three different BFEM meshes (comprising 8, 16, or 32 elements) are used to represent the fluid boundary. The time step is set to 0.01333 ms in all analyses. The results are shown in Fig. 11. One can see that except for the 8-element results, the 16- and 32-element results are nearly the same as the analytical solution. The poorer 8-element results are not unexpected because the velocity variations around the cylinder are ex-

pected to be more refined than 8 representative sectors. Nevertheless, the efficiency of the BFEM formulation is clearly demonstrated by using a fairly coarse 16-element mesh.

4 Conclusion

This paper presented the boundary finite element method (BFEM) formulation for acoustic fluid-structure interaction (FSI) problems. In the process of incorporating the acoustic effect, the authors developed the velocity-pressure relationship for the scattered wave. This relationship enables the extended applications of BFEM in solving the FSI problems. The formulation was verified through checking against benchmark solution—a submerged infinite long cylindrical shell subjected to step incident wave. The results show that the present formulation is able to yield more accurate solution than many prevailing numerical results (such as BEM, PWA solution). In a nutshell, the BFEM formulation is shown to be useful and efficient in solving unbounded fluid-structure interaction problems. It can represent accurately the unbounded fluid medium.
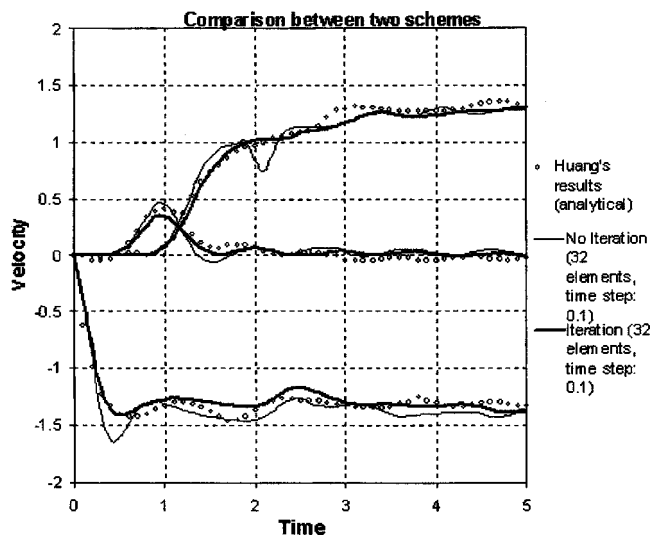

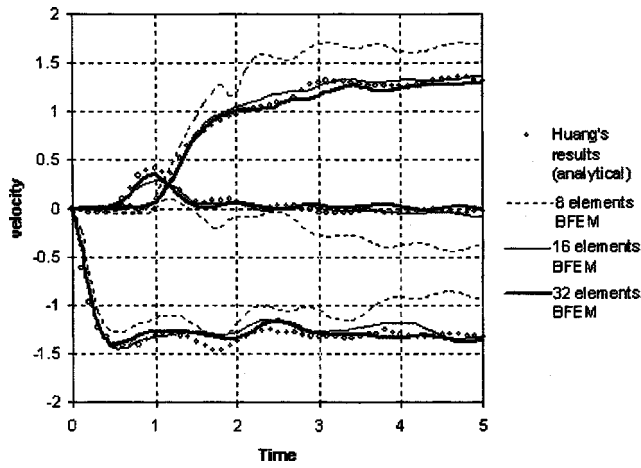Fig. 10 Comparison of results obtained from two Newmark schemes

Fig. 11 Convergence studies using 8, 16, and 32 elements

Nomenclature

$\mathbf{a}$ = structural accelerations velocity vector
$\mathbf{a}_{inf}$ = total accelerations vector on wet surface
$\mathbf{a}_{ff}$ = acceleration vector for incident wave
$\mathbf{A}$ = converting matrix
c = sound wave speed in fluid
$\mathbf{C}$ = structural damping matrix
d_s = structural displacement at mean radius
$\mathbf{d}$ = structural displacement vector
$\mathbf{F}_{ext}$ = body or traction force vector
$\mathbf{F}_{inf}$ = force vector for unbounded fluid medium
$\mathbf{F}_{ff}$ = force vector for incident wave
$\mathbf{F}_{sc}$ = force vector for scattered wave
i = nodal number
j = nodal number; jth time step; jth iteration
$\mathbf{K}$ = structural stiffness matrix
$\mathbf{M}$ = structural mass matrix
$\mathbf{M}^{\infty}$ = dynamic mass matrix for unbounded fluid medium
$\mathbf{M}^{-1}$ = inverse of fluid-added mass matrix
n = nth time step
$\mathbf{N}$ = shape function for structure element
$\mathbf{N}_f$ = shape function for fluid element
$\mathbf{N}_p$ = interpolation function for scattered wave pressure
O = scalar center
p = fluid pressure on wet surface
p_{inc} = fluid pressure on wet surface caused by incident wave
$\mathbf{p}_{inc}$ = fluid pressure vector caused by incident wave
p_{inf} = total fluid pressure on wet surface
$\mathbf{p}_{inf}$ = total fluid pressure vector
p_{ff} = fluid pressure on wet surface caused by free-field incident wave
$\mathbf{p}_{ff}$ = fluid pressure vector caused by free-field incident wave
p_{sc} = fluid pressure on wet surface caused by scattered wave
$\mathbf{p}_{sc}$ = fluid pressure vector caused by scattered wave
r = radial coordinate
R = mean radius of cylindrical shell
R_i = radius of inner surface of cylindrical shell
R_e = radius of outer surface of cylindrical shell
t = time variable
Δt = time increment
$\mathbf{v}$ = structural velocity vector

$\mathbf{v}_{inc}$ = normal velocity vector on wet surface caused by incident wave
v_{inf} = total normal velocity at one point on wet surface
$\mathbf{v}_{inf}$ = total normal velocity vector on wet surface
v_{ff} = normal velocity on wet surface caused by free-field incident wave
$\mathbf{v}_{ff}$ = normal velocity vector on wet surface caused by free-field incident wave
v_n = normal velocity of structure
v_r = normal velocity on wet surface caused by incident wave
v_{sc} = normal velocity on wet surface caused by scattered wave
$\mathbf{v}_{sc}$ = normal velocity vector on wet surface caused by scattered wave
$\mathbf{V}_{sc}$ = efficient normal velocity vector for scattered wave
w = width of scalar finite element cell
α = coefficient in Newmark scheme
δ = coefficient in Newmark scheme
ϕ = velocity potential
φ = nodal velocity potential vector
γ = measure of angle
θ = measure of angle
ρ = fluid density
Γ = wet surface
Σ_e = assemblage of finite elements

References

[1] Hamdan, F. H., 1999, "Near-Field Fluid-Structure Interaction Using Lagrangian Finite Elements," Comput. Struct., **71**, pp. 123–141.
[2] Belytschko, T., 1980, "Fluid-Structure Interaction," Comput. Struct., **12**, pp. 459–469.
[3] Bathe, K. J., Nitikitpaiboon, C., and Wang, X., 1995, "A Mixed Displacement-Based Finite Element Formulation for Acoustic Fluid-Structure Interaction," Comput. Struct., **56**, pp. 225–237.
[4] Morand, H., and Ohayon, R., 1979, "Substructure Variational Analysis of the Vibrations of Coupled Fluid-Structure Systems. Finite Element Results," Int. J. Numer. Methods Eng., **14**, pp. 741–755.
[5] Mellado, M., and Rodriguez, R., 2001, "Efficient Solution of Fluid-Structure Vibration Problems," Appl. Numer. Math., **36**, pp. 389–400.
[6] Biswal, K. C., Bhattacharyya, S. K., and Sinha, P. K., 2003, "Free-Vibration Analysis of Liquid-Filled Tank With Baffles," J. Sound Vib., **259**, pp. 177–192.
[7] Olson, L. G., and Bathe, K. J., 1985, "Analysis of Fluid-Structure Interactions. A Direct Symmetric Coupled Formulation Based on the Fluid Velocity Potential," Comput. Struct., **21**, pp. 21–32.
[8] Pal, N. C., Bhattacharyya, S. K., and Sinha, P. K., 2003, "Non-Linear Coupled Slosh Dynamics of Liquid-Filled Laminated Composite Containers: A Two Dimensional Finite Element Approach," J. Sound Vib., **261**, pp. 729–749.
[9] Nitikitpaiboon, C., and Bathe, K. J., 1993, "An Arbitrary Lagrangian-Eulerian Velocity Potential Formulation for Fluid-Structure Interaction," Comput. Struct., **47**, pp. 871–891.
[10] Mindlin, R. D., and Bleich, H. H., 1953, "Response of an Elastic Cylindrical Shell to a Transverse Step Shock Wave," ASME J. Appl. Mech., **20**, pp. 189–195.
[11] DiMaggio, F. L., Sandler, I. S., and Rubin, D., 1981, "Uncoupling Approximations in Fluid-Structure Interaction Problems with Cavitation," ASME J. Appl. Mech., **48**, pp. 753–756.
[12] Hamdan, F. H., and Dowling, P. J., 1995, "Far-Field Fluid-Structure Interaction Formulation and Validation," Comput. Struct., **56**, pp. 949–958.
[13] Fan, S. C., Wang, K., Yu, G. Y., and Lie, S. T., 2001, "Spline Shell Element and Plane-Wave Approximation for Dynamic Response of Submerged Structures," Comput. Struct., **79**, pp. 1635–1644.
[14] Geers, T. L., 1969, "Excitation of an Elastic Cylindrical Shell by a Transient Acoustic Wave," ASME J. Appl. Mech., **36**, pp. 459–469.
[15] Ranlet, D., DiMaggio, F. L., Bleich, H. H., and Baran, M. L., 1977, "Elastic Response of Submerged Shells with Internally Attached Structures to Shock Wave Loading," Comput. Struct., **7**, pp. 355–364.
[16] Zilliacus, S., 1983, "Fluid-Structure Interaction and ADINA," Comput. Struct., **17**, pp. 763–773.
[17] Givoli, D., 1991, "Non-Reflecting Boundary Conditions: A Review," J. Comput. Phys., **94**, pp. 1–29.
[18] Estorff, O. V., and Antes, H., 1991, "On FEM-BEM Coupling for Fluid-Structure Interaction Analyses in the Time Domain," Int. J. Numer. Methods Eng., **31**, pp. 1151–1168.
[19] Czygan, O., and Estorff, O. V., 2002, "Fluid-Structure Interaction by Coupling

BEM and Nonlinear FEM," Eng. Anal. Boundary Elem., **26**, pp. 773–779.

[20] Yu, G. Y., Lie, S. T., and Fan, S. C., 2002, "A Stable BEM/FEM Coupling Procedure Applied in Time Domain Fluid-Structure Interaction Problems," J. Eng. Mech., **128**, pp. 909–915.

[21] Song, C., and Wolf, J. P., 1996, "Consistent Infinitesimal Finite-Element Cell Method: Three-Dimensional Vector Wave Equation," Int. J. Numer. Methods Eng., **39**, pp. 2189–2208.

[22] Wolf, J. P., and Song, C., 1996, "Consistent Infinitesimal Finite-Element Cell Method: Three-Dimensional Scalar Wave Equation," ASME J. Appl. Mech., **63**, pp. 650–654.

[23] Wolf, J. P., and Song, C., 1996, *Finite-Element Modeling of Unbounded Media*, Wiley, Chichester.

[24] Dasupta, G., 1982, "A Finite Element Formulation for Unbounded Homogeneous Continua," ASME J. Appl. Mech., **49**, pp. 136–140.

[25] Song, C., and Wolf, J. P., 1997, "The Scaled Boundary Finite-Element Method-Alias Consistent Infinitesimal Finite-Element Cell Method-for Elastodynamics," Comput. Methods Appl. Mech. Eng., **147**, pp. 329–355.

[26] Deeks, J. A., and Wolf, J. P., 2002, "A Virtual Work Derivation of the Scaled Boundary Finite-Element Method for Elastostatics," Comput. Mech., **28**, pp. 489–504.

[27] Lamb, H., 1932, *Hydrodynamics*, 6th ed., Dover, New York, pp. 476–477.

[28] Mansur, W. J., Yu, G. Y., Carrer, J. A. M., Lie, S. T., and Siqueira, E. F. N., 2000, "The θ Scheme for Time-Domain BEM/FEM Coupling Applied to the 2-D Scalar Wave Equation," Commun. Numer. Methods Eng., **16**, pp. 439–448.

[29] Huang, H., 1970, "An Exact Analysis of the Transient Interaction of Acoustic Plane Waves with a Cylindrical Elastic Shell," ASME J. Appl. Mech., **37**, pp. 1091–1099.

Strength Analyses of Sandwich Pipes for Ultra Deepwaters

Segen Farid Estefen

Theodoro Antoun Netto

Ilson Paranhos Pasqualino
e-mail: ilson@lts.coppe.ufrj.br

Ocean Engineering Department,
COPPE-Federal University of Rio de Janeiro,
P.O. Box 68508,
Rio de Janeiro, RJ, 21945-970, Brazil

Design requirements for pipelines regarding both ultimate strength and flow assurance in ultra deepwater scenarios motivated the development of a new sandwich pipe which is able to combine high structural and thermal insulation properties. In this concept, the annulus is filled with low cost materials with adequate thermal insulation properties and good mechanical resistance. The aim of this research work is to perform small-scale laboratorial tests and to develop a finite element model to evaluate the structural performance of such sandwich pipes with two different options of core material. After calibrated in view of the experimental results, a three-dimensional finite element model incorporating nonlinear geometric and material behavior is employed to perform strength analyses of sandwich pipes under combined external pressure and longitudinal bending. Ultimate strength envelopes for sandwich pipes are compared with those generated for single-wall steel pipes with equivalent collapse pressures. The study shows that sandwich pipe systems with either cement or polypropylene cores are feasible options for ultra deepwater applications. [DOI: 10.1115/1.1940667]

Introduction

One of the challenges that the offshore oil industry faces as it moves to ultra deepwater is to design well-insulated pipelines and risers capable to withstand high internal and ambient external pressures. Pipe-in-pipe systems have been lately proposed (Fig. 1) as viable solutions to such problems [1–3]. They consist of two concentric metal pipes in which the annulus is either filled with a nonstructural insulating material or is used to carry water for well injection, umbilical cables, etc. Usually, internal and external pipes are designed independently against failure under internal and external pressures, respectively, combined with installation loads, mainly longitudinal bending. The factors governing the collapse and propagation of buckles in single pipes and pipe-in-pipe systems under external pressure have been extensively studied in the past so that nowadays, deepwater pipes can be safely designed [4–10]. Furthermore, a significant body of work for the design of submarine pipelines under combined external pressure and bending exists in the literature [11–16].

The concept presented in this paper aims at combining structural strength and thermal insulation in an optimized sandwich pipe with three layers, which are able to work together to resist combined high external pressure and bending loads, typical of installation processes of pipelines in ultra deepwaters.

Different types of ceramic or polymeric materials have been considered to fill the annulus space of the sandwich pipe. The selected material must provide good insulation and high compressive strength. Because of their wide availability and relatively low cost, cement and polypropylene were considered as core materials in this paper.

Laboratory tests of small-scale sandwich pipes under external pressure are initially described along with the nonlinear finite element model which is used to analyze the problem. After being carefully measured, small-scale sandwich pipes with two different geometries and core materials (cement and polypropylene) are tested under hydrostatic external pressure until collapse. The tests

are performed quasi-statically in an approximate volume-control scheme inside a pressure vessel. The experimental results are then compared with the calculated collapse pressures obtained by simulating numerically the physical experiments.

Finally, the structural performance of full scale sandwich pipes is studied through the numerical evaluation of the ultimate strength of different geometries under combined external pressure and longitudinal bending. The ultimate strength envelopes then generated are used to compare the structural performance of different geometries of sandwich pipes under combined loading with single-wall steel pipes with equivalent collapse pressures.

Experiments

The tubes used to manufacture the small-scale specimens were cut out of 6 m long aluminum tubes. The nomenclature used henceforth to identify the various specimens is given below:
PIP.XX.YY.ZZ where

XX =	core material (M1=cement or M2 =polypropylene);	
YY =	geometric parameter identifier (G1 or G2);	
ZZ =	model string.	

The total length of the specimens was kept equal to 1000 mm. The main measured geometric parameters of the individual specimens tested are listed in Table 1. The outer diameters (D) and wall thickness (t) represent the average values of sets of measurements made for outer and inner tubes. The variable Ξ_o is the wall thickness eccentricity measured at the ends of these tubes and Δ_o is the maximum value of initial ovality. They are defined as follows:

$$\Xi_o = \frac{t_{\max} - t_{\min}}{t_{\max} + t_{\min}} \tag{1}$$

and

$$\Delta_o = \frac{D_{\max} - D_{\min}}{D_{\max} + D_{\min}}. \tag{2}$$

During the manufacturing process, inner and outer tubes were held in place by two centralizing steel plugs at each end. They were assembled in a way that the maximum and minimum diameter directions of the transverse sections with the biggest ovality were approximately coincident. Two different manufacturing processes were used to fill the annular space depending on the material (cement or polypropylene). In both cases, the adopted procedure was such to minimize bubbles and cracks in the annulus
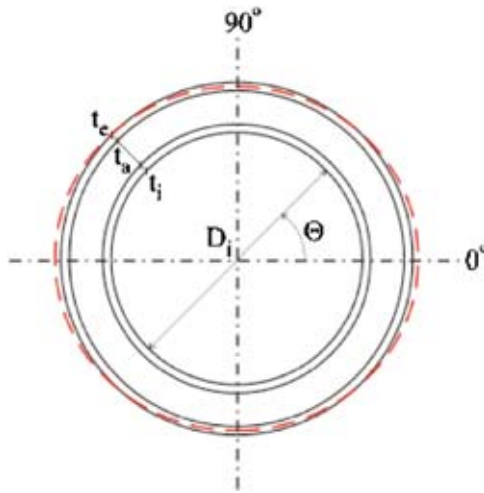
Fig. 1 Sandwich pipe geometry

Fig. 3 Polypropylene specimen

material and contact surfaces, and to provide a good adherence between layers.

The experimental setup used for cement injection is shown in Fig. 2. After mixing, the cement paste is poured in a reservoir sealed at the top by a piston that is driven by a hydraulic actuator. The apparatus is mounted on a compression frame, as indicated in the figure. The piston is then moved slowly, forcing the cement mix to flow from the reservoir into the annular space through the

lower centralizer. While the mixture is pumped, the assembly rests on a vibratory base and the air confined in the annulus is expelled through vents located at the top centralizer.

For the polypropylene specimens, polypropylene tubes were machined so as to fill in the annular space with a 0.5 mm radial tolerance. First the polypropylene tube is mounted on the inner tube and then slipped inside the outer tube (Fig. 3). The tubes are bounded together by an epoxy resin so to provide a good adherence between layers. Although it may not be feasible for large-scale production of full-scale specimens, this procedure proved to be very cost effective in the manufacturing of the specimens used in this experimental program.

Test coupons were cut in the longitudinal direction of each tube and tested in a conventional servo-hydraulic machine at a strain rate of approximately 10^{-4} s^{-1}. Figure 4 shows a typical engineering stress-strain curve obtained from such tests. The main measured material properties of the tubes are given in Table 2, where E is the Young's modulus, σ_o is the 0.2% offset yield stress, and σ_u is the ultimate stress. The procedure used to characterize the behavior of the cement under tension and compression followed the guidelines suggested by the Committee on Standardization of Laboratory and Field Tests (1975). Several cylindrical test specimens were made using the same paste injected in the sandwich pipe specimens. The average engineering stress-strain curve of all uniaxial compression tests is presented in Fig. 5. Typically, the ultimate tensile stress was approximately 10% of the correspon-

Table 1 Main geometric parameters of test specimens.

Specimen	Tube	D (mm)	D/t	Δ_o (%)	Ξ_o (%)
PIP.M1.G1.I01	Inner	49.15	30.34	0.225	4.172
	Outer	74.97	46.28	0.629	5.666
PIP.M1.G1.I02	Inner	50.36	30.90	0.120	1.534
	Outer	75.92	46.01	0.266	3.625
PIP.M1.G2.I01	Inner	50.76	30.21	0.205	3.650
	Outer	62.16	42.29	1.161	1.958
PIP.M1.G2.I02	Inner	50.73	30.38	0.260	5.090
	Outer	62.28	42.37	0.698	1.694
PIP.M2.G1.I02	Inner	49.64	29.37	0.456	4.451
	Outer	75.40	46.54	0.301	5.264
PIP.M2.G1.I03	Inner	49.76	30.72	0.186	3.887
	Outer	75.19	46.41	0.255	5.590
PIP.M2.G2.I01	Inner	49.94	29.55	0.364	4.451
	Outer	62.10	42.53	0.801	1.690
PIP.M2.G2.I02	Inner	50.03	29.60	0.547	4.451
	Outer	62.40	41.88	0.552	2.503

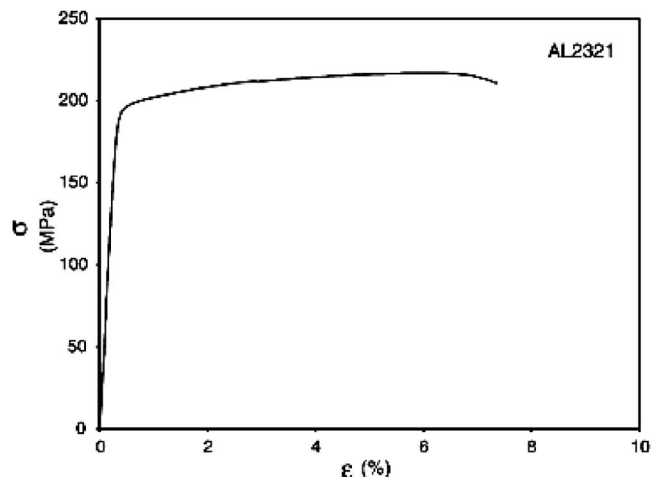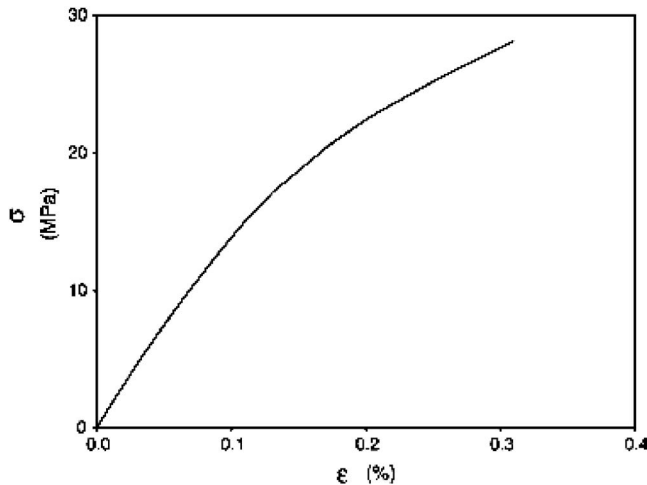**Fig. 2 Experimental setup for cement injection**

Fig. 4 Engineering stress-strain curve of aluminum tubes

Table 2 Aluminum pipe mechanical properties

Specimen	Tube	E (GPa)	σ_o (MPa)	σ_u (MPa)
PIP.M1.G1.I01	Inner	66.810	199.79	225.62
	Outer	61.967	200.93	223.40
PIP.M1.G1.I02	Inner	62.635	195.20	216.73
	Outer	61.507	192.10	213.36
PIP.M1.G2.I01	Inner	64.447	180.71	205.43
	Outer	63.130	141.57	171.27
PIP.M1.G2.I02	Inner	64.447	180.71	205.43
	Outer	63.130	141.57	171.27
PIP.M2.G1.I02	Inner	63.349	186.82	214.35
	Outer	64.909	206.52	230.97
PIP.M2.G1.I03	Inner	63.425	194.37	215.19
	Outer	64.909	206.52	230.97
PIP.M2.G2.I01	Inner	63.349	186.82	214.35
	Outer	63.130	141.57	171.27
PIP.M2.G2.I02	Inner	63.349	186.82	214.35
	Outer	64.975	160.37	190.21

Table 4 Polypropylene average parameters

Rod	Test	σ_u (MPa)	ε_u (%)
2	Tension	33.48	8.48
3	Tension	35.41	8.67
5	Tension	34.31	8.88
6	Tension	34.16	8.51
Average	**Tension**	**34.33**	**8.64**
2	Compression	43.40	16.64
3	Compression	50.11	21.41
5	Compression	46.04	20.19
6	Compression	44.38	17.72
Average	**Compression**	**45.98**	**18.99**

dent value in compression. Table 3 lists the main cement parameters obtained from both sets of tests. Polypropylene specimens and test procedures followed the specifications given by ASTM codes D638M-91 and D695-91. In general, three tensile and three compression test specimens were machined from the rods used to fabricate the polypropylene tubes. The material parameters of the individual rods are given in Table 4, along with the overall average parameters. The correspondent engineering tensile stress-strain curve is shown in Fig. 6. The specimens were tested under hydrostatic pressure in a 380 mm internal diameter, 5 m long pressure vessel with a pressure capacity of 50 MPa. The vessel is completely filled and pressurized with water using a positive displacement pump. The pressure is monitored by an electrical pressure transducer and by analog pressure gages connected to the pressurizing line. Four strain gages are mounted in the hoop direction at the mid-section of the specimens. The gages and connecting wires are insulated with a compliant coating.

Figure 7 shows a typical set of data from the four strain gages. The strains measured at the different points are very similar in the beginning of the test, when membrane strains are predominant. After an initial stiff response, bending strains caused by the in-

creasing ovality of the cross sections prevail. This is reflected by the divergence between strains measured circumferentially at points spanned by 90 deg. Eventually, the specimen locally collapses in the neighborhood of its weakest point. Local collapse is followed by a sudden drop in pressure caused by the spread of the buckle along the entire length of the specimen. The collapsed cross-section configuration of specimen PIP.M2.G1.I02 is shown in Fig. 8.
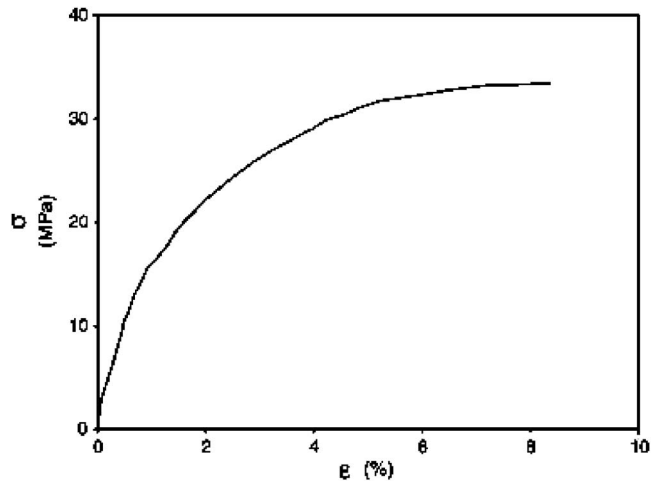
Fig. 6 Polypropylene tensile stress-strain curve
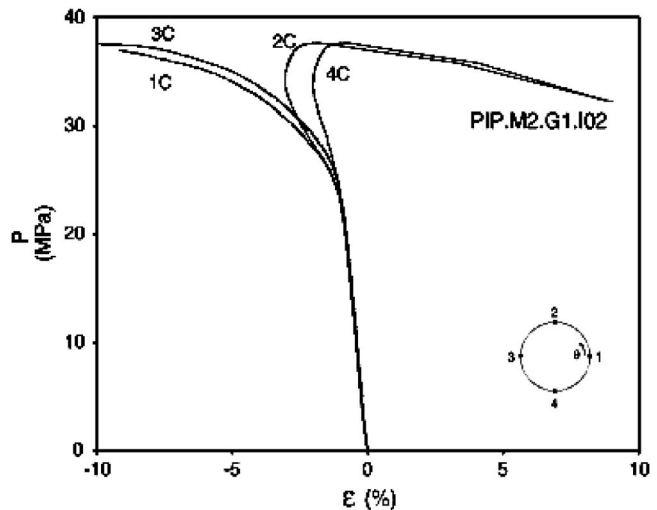
Fig. 7 Pressure-strain response recorded during the test of specimen PIP.M2.G1.I02

Fig. 5 Cement compressive stress-strain curve

Table 3 Cement mechanical properties

Test	E (GPa)	σ_u (MPa)	ε_u (%)
Compression	12.76	28.12	0.31
Tension	12.28	2.89	0.024

Fig. 8 Collapsed cross section of specimen PIP.M2.G1.I02

Fig. 9 Typical finite element mesh used in the correlation analyses

The collapse pressures (P_{co}) of the eight specimens tested are listed in Table 5, along with their main geometric parameters. Internal diameter D_i and thickness of the annulus (t_a) were calculated from the mean geometric parameters given in Table 1. The differences between the collapse pressures of specimens with the same nominal geometry and material can be partially attributed to slight variations from specimen to specimen in material properties, diameter, thickness, and initial imperfections (ovality and eccentricity). Other imperfections such as lack of adhesion between layers can also strongly affect the collapse pressure of sandwich pipes. This will be discussed later together with the results obtained from the numerical analyses.

Numerical Analyses of Sandwich Pipes Under External Pressure

Numerical Procedures. A numerical model based on the finite element method was developed within the framework of the software package ABAQUS [17] to simulate the behavior of sandwich pipes under hydrostatic pressure. Figure 9 illustrates a typical finite-element mesh used in the analyses. Eight-node, quadratic, nonlinear plane-strain elements (CPE8) with two degrees-of-freedom per node (displacements in directions 1 and 2) were used to model inner and outer pipes, and the annulus region. This simpler 2-D model with plane strain was assumed to simulate the uniform collapse of a long tube.

Symmetry conditions reduced the problem to a quarter of a ring, as shown in Fig. 9. The vast majority of experiments presented the well-known doubly symmetric mode of collapse, and that is the reason for this assumption in the numerical model. In fact, only two experiments ended up precipitating the U-mode of collapse, which is explained by the small eccentricity between

Table 5 Specimen geometric parameters and experimental collapse pressures

Specimen	D_i (mm)	t_i (mm)	t_a (mm)	t_e (mm)	P_{co} (MPa)
PIP.M1.G1.I01	45.91	1.62	11.29	1.62	43.35
PIP.M1.G1.I02	47.10	1.63	11.13	1.65	34.09
PIP.M1.G2.I01	47.40	1.68	4.23	1.47	10.98
PIP.M1.G2.I02	47.37	1.67	4.30	1.47	12.11
PIP.M2.G1.I02	46.28	1.68	11.26	1.62	37.64
PIP.M2.G1.I03	46.52	1.62	11.10	1.61	31.14
PIP.M2.G2.I01	46.54	1.70	4.62	1.46	20.31
PIP.M2.G2.I02	46.65	1.69	4.69	1.49	17.13

inner and outer pipes in these particular experiments.

Inner and outer tubes were discretized with ten elements in the hoop direction and one element through the thickness. Two to four elements were used to model the annulus region through the thickness depending on the geometry considered. The mesh density used was verified through a mesh sensitivity analysis in a similar work [18].

Finally, two layer interface conditions were simulated numerically—perfect adhesion and no adhesion between tubes and the annulus material. In the former case, annulus and tube meshes shared the same nodes along the interfaces. Lack of adhesion was simulated through a surface-based contact model between layers. The models allow for finite sliding between master and slave contacting surfaces, respectively tubes and annulus material.

Tubes were assumed to be J_2-flow, finite deforming, elastic-plastic solids with isotropic hardening. The cement was modeled in the plastic regime by a simplified associative flow rule with isotropic hardening. The yield surface, which is a function of the hydrostatic stress and the Mises equivalent stress at the integration points, is calibrated using the uniaxial stress-strain curves. The following failure ratios and parameters were adopted [17]:

- ratio between biaxial and uniaxial ultimate stresses=1.16;
- ratio between maximum plastic strains under biaxial and uniaxial compression=1.28;
- ratio between ultimate stresses under uniaxial tension and compression (absolute value)=0.10;
- Poisson ratio=0.15.

The polypropylene was modeled as a hyperelastic, incompressible material. The compressive and tensile uniaxial curves determined experimentally were used to calibrate the potential deformation energy model (Ogden) available in the ABAQUS library.

Correlation Between Numerical and Experimental Results. The numerical model was first employed to reproduce each of the physical experiments described in the previous section. The average geometric and material parameters of the individual specimens were used in the analyses considering both tied and untied material layers (i.e., perfect adhesion and no adhesion between tubes and annulus material). The measured maximum initial ovality (Δ_o) of the outer tube was reproduced via radial displacements (w_o) in relation to the perfect cylindrical surface as follows:

Table 6 Correlation between experimental and numerical results

Specimen	$\hat{P}_{CO}$ (MPa) perfect adhesion	$\hat{P}_{CO}$ (MPa) no adhesion	P_{CO} (MPa) experiment
PIP.M1.G1.I01	46.23	37.97	43.35
PIP.M1.G1.I02	44.78	38.05	34.09
PIP.M1.G2.I01	24.70	8.11	10.98
PIP.M1.G2.I02	25.74	10.03	12.11
PIP.M2.G1.I02	39.56	12.84	37.64
PIP.M2.G1.I03	38.27	12.52	31.14
PIP.M2.G2.I01	20.84	6.32	20.31
PIP.M2.G2.I02	22.42	6.89	17.13

$$w_o = -\frac{\Delta_o D}{2}\cos 2\theta \qquad (3)$$

where D is the outside diameter of the sandwich pipe and θ is the polar coordinate of a given node ($0 \leqslant \theta \leqslant 90$ deg). The calculated collapse pressures ($\hat{P}_{co}$) are given in Table 6 along with the correspondent experimental values (P_{co}).

The results show that, except for specimen PIP.M1.G1.I02, the experimental collapse pressures fell in between the calculated values considering perfect adhesion and no adhesion between layers. Actually, these two values can be expected to be the upper and lower estimates for a given material and geometric configuration. In general, the predictions assuming perfect adhesion between layers compared quite well with the experimental results for most of the studied configurations (e.g., nearly 5% average difference for specimens PIP.M1.G1.I01, PIP.M2.G1.I02, PIP.M2.G2.I01). Experimental collapse pressure of specimens PIP.M1.G1.I02, PIP.M1.G2.I01, and PIP.M1.G2.I02 were closer to the lower collapse pressure estimate (i.e., no adhesion assumption). A poor adhesion between the cement annulus and the pipes may have occurred after curing, which was not noticed before testing. This was not observed for the polypropylene specimens, in which the use of an epoxy resin promoted a good adhesion between layers.

It is also interesting to note the differences between lower and upper estimates, especially for both polypropylene specimens and for cement specimens with $t_a = 4.25$ mm. In these cases, the lack of adhesion between layers had a significant detrimental effect in the performance of the sandwich pipe, with a decrease of nearly 65% in the estimated collapse pressure. On the other hand, this effect was minor for specimens PIP.M1.G1.I01 and PIP.M1.G1.I02, probably because of the higher relative rigidity of the cement layer as compared to the overall sandwich pipe.

Parametric Study. A comprehensive parametric study on the collapse pressure of full-scale sandwich pipes with different geometries, materials, and initial imperfections was accomplished by Netto et al. [18]. It has been concluded that

- For both polypropylene and cement as annulus material, an approximately linear variation of $\hat{P}_{co}$ with increasing t_a was observed.
- Expectedly, because cement has a greater initial tangent modulus than polypropylene, the collapse pressure of a cement sandwich pipe is always greater than the one for a correspondent geometry with polypropylene. In addition, the collapse pressure for cement sandwich pipes presents a steeper growing rate when t_a is increased.
- As it occurs with single pipes, the collapse pressure of sandwich pipes is rather affected by geometric imperfections in the form of ovality of the cross sections.
- Lack of adhesion between layers can significantly decrease the maximum pressure capacity of sandwich pipes. Furthermore, the influence of the initial ovality on the collapse
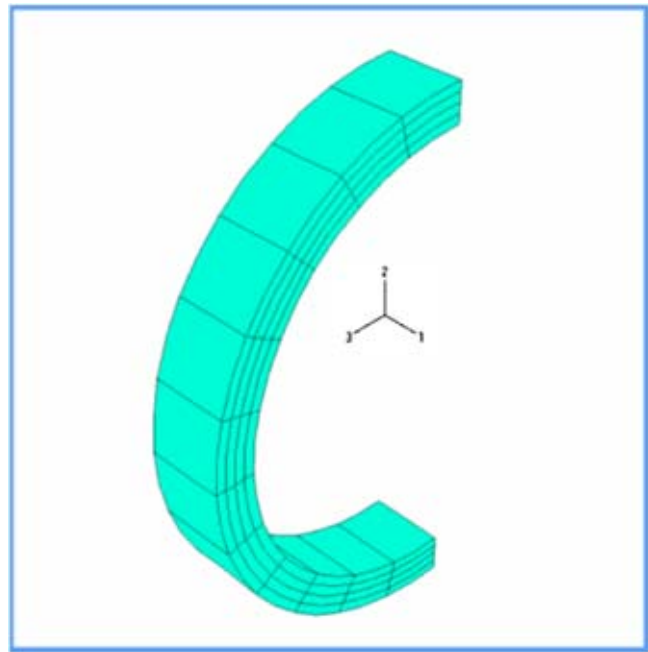
Fig. 10 Three-dimensional finite element mesh

pressure is amplified when there is lack of adhesion between layers, mainly for cement sandwich pipes.

As a continuation of the research work performed by Netto et al. [18], this paper will focus on the structural behavior of sandwich pipes under combined external pressure and bending, and their comparative performance with steel pipes. These results are shown next.

Numerical Analyses of Sandwich Pipes Under Combined External Pressure and Bending

Numerical Procedures. The problem consists of a long, circular, sandwich pipe under combined bending and external pressure loads. The pipe geometry (including initial imperfection in the form of ovality, eq. (3)), curvature K, and the deformation of its cross sections are assumed to be uniform along the length. It is also assumed that plane sections perpendicular to the pipe midsurface remain plane during loading. The coordinate system used to define the problem geometry is shown in Fig. 10. Symmetry about planes 1–2 and 2–3 are also assumed. These hypotheses allow the problem to be reduced to the analysis of a half-ring of unit length. In addition, for simplicity the layers were considered perfectly adhered to each other.

Again, the numerical model developed operates within the framework of the nonlinear FE code ABAQUS [17]. A typical mesh used in the analyses is shown in Fig. 10. The layers were discretized with three-dimensional, 27-node, quadratic brick elements (C3D27 for steel pipes and cement annulus and its hybrid version C3D27H for polypropylene annulus). In the circumferential direction, the 180 deg sector was discretized with 12 elements with equal angular spans of 15 deg each. Four elements were used through the thickness (one for inner and outer pipes and two for the annulus region). Although this mesh may seem coarser than that used for the correlation with experiments, the goal here is just to provide data to compare the performance of sandwich pipes and single wall pipes.

The loading history is simulated by applying external pressure followed by longitudinal bending until collapse ($P \rightarrow K$), i.e., past the maximum bending moment. External pressure is first prescribed through surface pressure load increments on the external
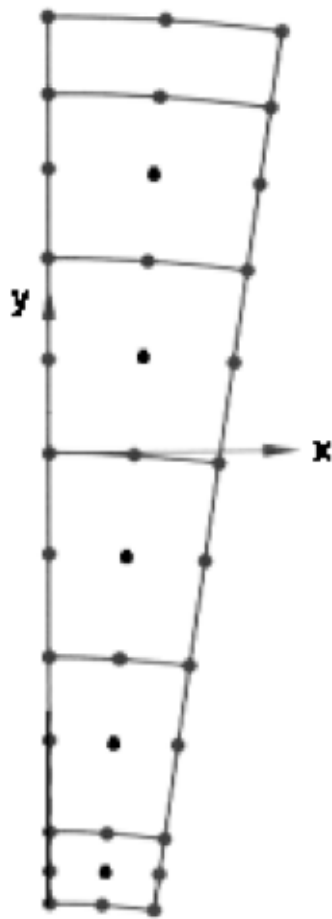
Fig. 11 Ring with unit length under longitudinal bending

Table 7 Geometric properties of the sandwich pipes

Case	D_i (mm)	t_i, t_e (mm)	t_a (mm)
1	152.4	3.175	6.35
2	152.4	4.7625	19.05
3	152.4	6.350	31.75

the hydrostatic pressure corresponding to water depths of 1000, 2000, and 3000 m, respectively. The same geometric parameters were then used for the cement-filled sandwich pipes.

The steel tubes were assumed to have Young modulus equal to 206.82 GPa and Poisson ratio equal to 0.3. Figure 12 shows the true stress-strain curve (nominal API X-60 grade steel) used in the numerical model.

Considering that the mechanical properties of the filled materials are dependent on both chemical composition and manufacturing process, simplified representative stress-strain curves were adopted for the numerical analyses. The uniaxial compressive stress-strain curve of the cement was fitted by three line segments where the first one represents the elastic regime with elastic modulus equal to 12.41 GPa. Therefore, the plastic regime is represented by the three points shown in Table 8 (ε^p is the engineering plastic strain).

The uniaxial tensile curve proposed by Sathe et al. [20] for the polypropylene (Fig. 13) was used to calibrate the potential deformation energy model (Ogden).

The collapses pressures of the three geometries with either polypropylene (PP) or cement (CMT) are presented in Table 9. The collapse pressures of the cement-filled sandwich pipes are, on average, 1.6 times the values obtained for respective geometries with polypropylene.

Pressure-curvature collapse envelopes for the analyzed sandwich pipes are shown in Figs. 14 and 15. Each envelope is described by six points, corresponding to constant pressures of 0%, 20%, 40%, 60%, 80%, and 100% of the corresponding collapse pressure (P_{co}).

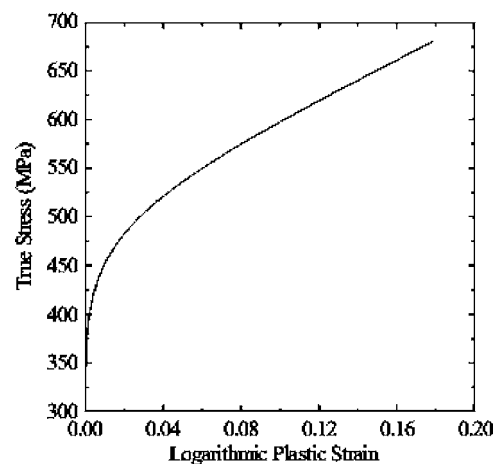For both materials, it can be verified that the differences among

face of the outer pipe. Bending is then applied with the aid of a reference node at $x_1=1$; $x_2=x_3=0$ which is coupled with the edge nodes on plane *2–3* at $x_1=1$. This kinematic coupling constraint makes the set of edge nodes to follow the rigid body motion of the reference node, as shown in Fig. 11. Therefore, bending load is applied by setting increments of rotation about the axis 3 at the reference node.

The failure load reported is associated with a limit-load type of instability in all cases considered. In fact, thin pipes under pure bending or combined bending and low pressure can experience bifurcation and wrinkle prior to attaining the limit load [5]. Pasqualino et al. [19] present results from a parametric study using similar geometries and materials and a 3-D finite element model which is able to capture both types of instabilities. The type of loading is exactly the same, and, in all cases, limit load was reached before bifurcation (wrinkling).

For the cement cases, the crack detection under tension is the main cause of failure in the pure bending and low pressure-bending analyses. This happens before failure due to crushing can occur, basically because of the cement different behaviors under tension and compression (maximum stresses in tension are about 10% of the compressive value).

Parametric Study. Sandwich pipe geometries considered for the parametric study are shown in Table 7. The internal diameter of the inner pipe (D_i) was fixed equal to 152.4 mm while t_i, t_e, and t_a were varied as indicated in the table. Thicknesses t_i and t_e were assumed identical for simplicity and an initial out-of-roundness of 1% was considered. As it will be seen later, these geometric parameters were selected so that the collapse pressure (P_{co}) of each sandwich pipe filled with polypropylene was at least 1.75 times

Fig. 12 True stress versus logarithmic plastic strain of the API steel grade X-60

Table 8 Points of the compressive stress-strain curve of the cement in the plastic regime

σ (MPa)	ε^p
20.00	0
34.30	0.0022
34.41	0.0100

Fig. 13 Polypropylene nominal stress-strain curve [11]

Fig. 15 Pressure-curvature ultimate strength for sandwich pipes filled with cement

collapse curvatures (K_{co}) for the three geometries considered are gradually reduced when the applied external pressure decreases. For instance, the collapse curvatures under pure bending for case 3 are smaller than those calculated for case 2, in spite of the lower rigidity of the cross section in case 2. This can be explained as follows: the increase of the external diameter from case 1 to case 3 (177.80, 209.55, and 241.30 mm, respectively) also increases the strains of the outer pipe for the same applied curvature. This leads to a premature collapse of the thickest sandwich pipe due to excessive yielding of the outer fibers.

When cracking is defined to occur at material points in tension, the model neglects any stress in the direction of the crack. For cement-filled sandwich pipes, the detection of cracks generates

Table 9 Collapse pressures of the sandwich pipes

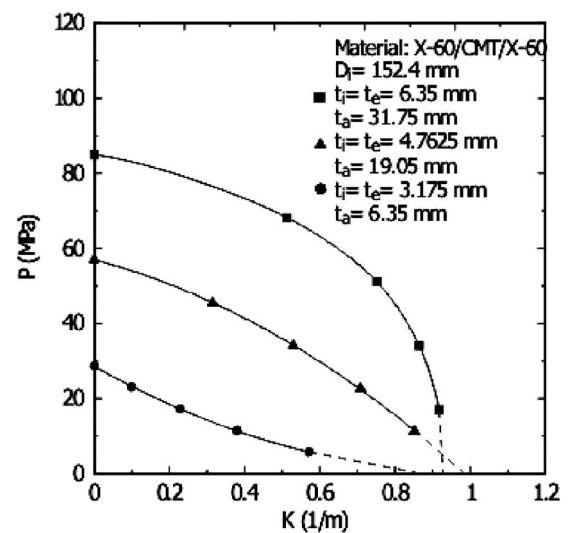| Case	Annulus	P_{co} (MPa)
1	PP	18.75
	CMT	28.77
2	PP	36.83
	CMT	57.01
3	PP	52.25
	CMT	84.99

Fig. 14 Pressure-curvature ultimate strength for sandwich pipes filled with polypropylene

numerical instabilities and is basically dependent on the mesh adopted. The dashed lines in Fig. 15 were obtained from extrapolation of the results of the analyses under combined loading.

Comparision Between Sandwich Pipes and Steel Single Wall Pipes

The structural performance of the sandwich pipe cases (Table 7) was compared with that of single-wall steel pipes (API grade X-60), with the same internal diameter, initial ovality (1%), and equivalent collapse pressures (±2% estimated using the FEM), in order to assess the feasibility of sandwich pipes for ultra deepwater applications. The geometric parameters of the single-wall pipes such as internal diameter D_i, external diameter D, and wall thickness (t) are presented in Table 10.

Ultimate strength curves for both sandwich and single-wall pipes, with similar collapse pressure, are showed in Figs. 16–21. The results are normalized by the collapse pressure (P_{co}) and the curvature parameter ($K_c = t/D^2$) of the single-wall pipe.

The results suggest that the structural performance of the sandwich pipe is superior to the corresponding single-wall pipe. Under the same external pressure level, the geometries considered for the polypropylene-filled sandwich pipes presented higher collapse curvatures than those for equivalent single wall pipes. The collapse curvature for cases 1 and 2 with polypropylene is, on average, twice the value calculated for the equivalent single wall pipe, as can be seen in Figs. 16 and 17. This behavior is not observed for case 3 (Fig. 18), where the thickness of the annulus is five times the metallic walls. The contribution of the filler to the strength of the sandwich pipe seems to decrease as the ratio between t_a and $t_i = t_e$ increases too much.

Table 10 Single-wall steel pipes with collapse pressure equivalent to respective sandwich pipes

Sandwich pipe		Single-wall pipe		
Case	Annulus	D_i (mm)	t (mm)	D/t
1	PP	152.4	6.985	23.8
	CMT	152.4	8.890	19.1
2	PP	152.4	10.287	16.8
	CMT	152.4	13.843	13.0
3	PP	152.4	13.208	13.5
	CMT	152.4	18.796	10.1

Fig. 16 Pressure-curvature ultimate strength for both sandwich (polypropylene) and single-wall pipes—case 1
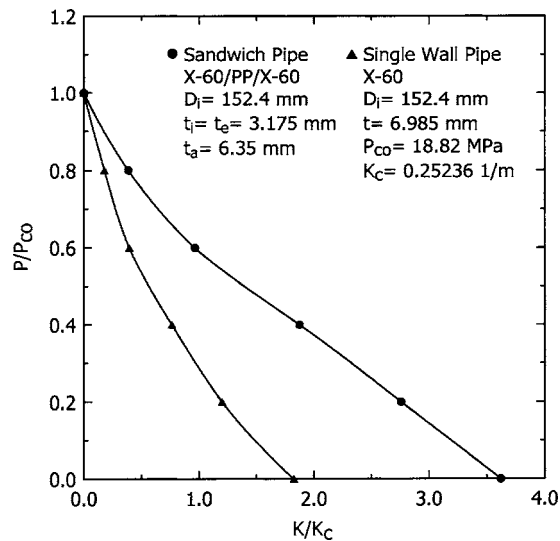

Fig. 18 Pressure-curvature ultimate strength for both sandwich (polypropylene) and single-wall pipes—case 3

Except for case 3, where the extrapolated result estimated for K_{co} in pure bending is lower than that presented for the equivalent single wall pipe (Fig. 21), an extra strength was presented for cement-filled sandwich pipes (Figs. 19 and 20). Nevertheless, ultimate curvatures under combined loading for case 3 were between 22% ($P=0.2P_{co}$) and 292% ($P=0.8P_{co}$) higher than those for equivalent single-wall pipes. Again, as observed for pure bending in Figs. 19–21, the structural benefits of the filler decreases with the relative increase of the thickness t_a.

Finally, the two concepts (sandwich and single-wall pipes) are compared through the steel weight (P_s), total weight (P_{tot}), and submerged weight (P_{sub}) per unit length in order to provide insight for the evaluation of material and manufacturing costs, as well as the feasibility of transportation and installation. Table 11 summarizes these results, where the following densities (ρ) were used to estimate the structural weight: $\rho_{steel}=7850$, $\rho_{PP}=910$, $\rho_{CMT}=1724$, and $\rho_{water}=1025$ Kg/m^3.

As can be observed from Table 11, there is a significant reduction in the steel weight for cement-filled sandwich pipes in relation to single-wall pipes, 25% on average. Polypropylene-filled
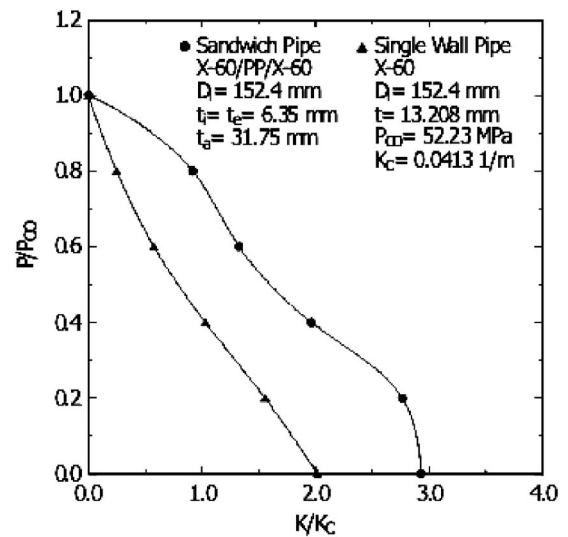
sandwich pipes presented almost the same steel weight as equivalent single-wall pipes (cases 1 and 2) and a slightly higher value, about 14%, for case 3. Moreover, in spite of having a superior overall weight in relation to the single-wall pipes, the submerged weight of sandwich pipes is usually smaller. It should be noted, however, that the core material of sandwich pipes works as a thermal insulation layer which is normally manufactured over the external surface of the single wall pipes, not being considered in Table 11. This layer depends on the scenario of application and cannot be estimated here.

Concluding Remarks

Sandwich pipes like the ones proposed here aim to fulfill concomitantly structural and thermal design requirements. Therefore, the core material must be selected in order to provide both thermal insulation and, combined with the internal and external pipes, sufficient strength against either burst or collapse of the system under installation and working loads. Certainly, many different combinations of materials (pipes and annulus) and geometries (t_i, t_e, and t_a) may have similar structural and insulation performance. Mate-

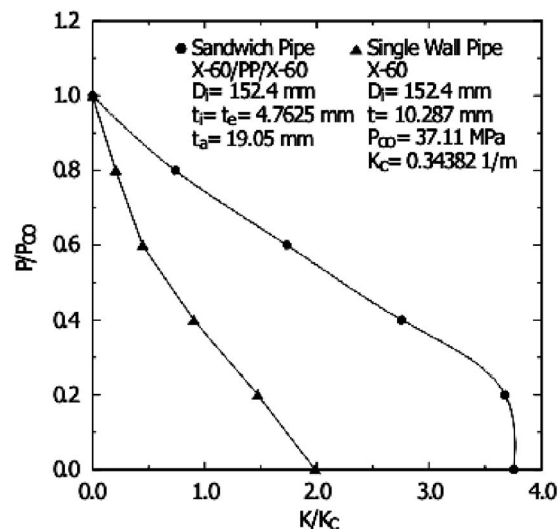**Fig. 17 Pressure-curvature ultimate strength for both sandwich (polypropylene) and single-wall pipes—case 2**
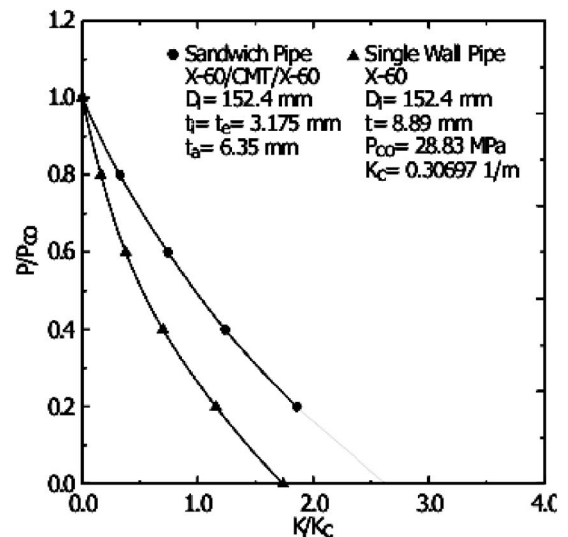

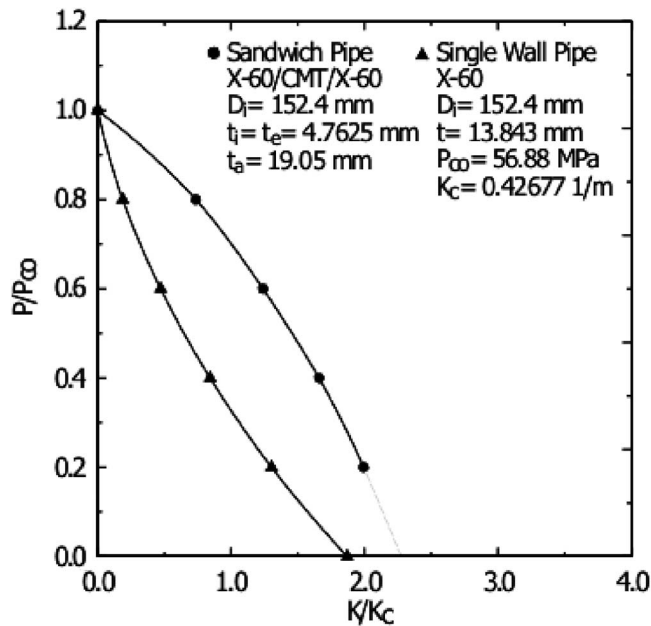**Fig. 19 Pressure-curvature ultimate strength for both sandwich (cement) and single-wall pipes—case 1**

Fig. 20 Pressure-curvature ultimate strength for both sandwich (cement) and single-wall pipes—case 2

rial availability, weight, thermal and mechanical properties, manufacturing process, installation method, and other cost-related variables are important parameters to be considered in the design
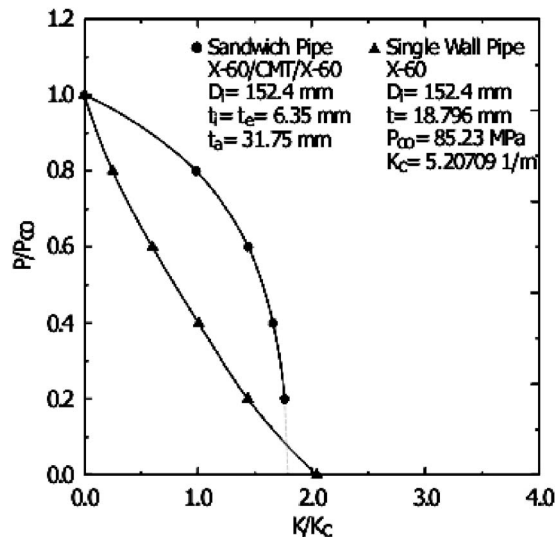

Fig. 21 Pressure-curvature ultimate strength for both sandwich (cement) and single-wall pipes—case 3

process. Because of the wide range of choices, the design of such pipes shall be tailored based on the specific requirements of each subsea production system.

Eight small-scale sandwich pipes, filled either with cement or polypropylene, have been tested under external pressure up to the collapse. The experimental results have been correlated with those from numerical analyses assuming both material and geometric nonlinearities. Except for one model, the experimental collapse pressure is within the numerical results obtained for full adhesion and lack of adhesion between metal pipes and annulus material. Special care should be exercise during the manufacturing process in order to ensure appropriate adhesion between layers, therefore increasing the sandwich pipe collapse pressure.

A numerical procedure for the structural analysis of sandwich pipes under combined loading of external pressure and longitudinal bending has been employed, based on the finite element method, with the aim of evaluating the feasibility of such structures for ultra deepwater applications. The proposed numerical model generated ultimate strength curves for three different geometries of sandwich pipes with two different filling materials, polypropylene and cement.

The structural performance under combined loading of six full-scale configurations of sandwich pipes was compared with that for steel single wall pipes, with the same internal diameter, initial ovality, and collapse pressure. The numerical results showed that the structural performance of sandwich pipes is superior to the corresponding single-wall pipes. Under the same constant external pressure, the polypropylene-filled sandwich pipes presented higher collapse curvatures than the equivalent single-wall pipes. Similar results were in general obtained for the cement-filled sandwich pipes. The submerged weight for sandwich pipes, inferior in most of the analyzed cases, and the reduction (cement) or equivalence (polypropylene) in weight of steel, when compared to the corresponding single-wall pipes, are parameters that contribute to evaluate positively the application of this alternative concept for ultra deepwater applications.

Although the results of pure pressure indicate that poor bonding can result in a significant reduction of the collapse pressure, shrinking the pressure-bending envelope, structural adhesive may be studied to guarantee perfect bonding. Additional full-scale prototype tests are recommended to confirm the efficiency of the proposed sandwich pipes for a particular offshore scenario.

Acknowledgment

Nomenclature

D_i = inner diameter
D = outer diameter
D_{max} = maximum diameter at the cross section
D_{min} = minimum diameter at the cross section

Table 11 Comparison between the structural weight of sandwich and single-wall pipes

		Sandwich pipes			Single-wall pipes		
Case	Annulus	P_s (kg/m)	P_{tot} (kg/m)	P_{sub} (kg/m)	t (mm)	P_s, P_{tot} (kg/m)	P_{sub} (kg/m)
1	PP	25.85	28.84	3.41	6.985	27.50	5.18
	CMT		31.53	6.09	8.890	35.36	12.04
2	PP	42.50	52.37	17.01	10.29	41.27	17.19
	CMT		61.18	25.83	13.84	56.75	30.64
3	PP	61.64	79.51	32.65	13.21	53.93	28.20
	CMT		95.50	48.63	18.80	79.35	50.29

E = material Young's modulus
K = applied curvature
K_c = curvature parameter (t/D^2)
K_{co} = collapse curvature
P = applied pressure
P_{co} = collapse pressure
P_s = steel weight
P_{sub} = submersed pipe weight
P_{tot} = total weight
w_o = radial geometric imperfection
t = wall thickness
t_a = annulus thickness
t_i = inner pipe thickness
t_e = outer pipe thickness
t_{max} = maximum thickness at the cross section
t_{min} = minimum thickness at the cross section

Greek

Δ_o = initial ovality
Ξ_o = thickness eccentricity
ε^p = plastic strain
ε_u = ultimate strain
ρ_{CMT} = cement density
ρ_{PP} = polypropylene density
ρ_{steel} = polypropylene density
ρ_{water} = water density
σ = compressive stress
σ_o = ultimate stress
σ_u = ultimate stress
θ = polar coordinate

References

[1] Nutall, R. H., and Rogers, M. G., 1998, "Insulated Pipe-in-Pipe Subsea Hydrocarbon Flowlines," *Proceedings 17th International Conference on Offshore Mechanics and Arctic Engineering*, Lisbon, Portugal.

[2] Franklin, J. F., and Wright, A., 1999, "The Development of Thermal Insulation Systems for Deep Water Pipelines," *Proceedings 13th International Conference on Pipeline Protection*, Edinburgh, UK, pp. 217–229.

[3] Trout, S., and Sahota, B., 1999, "Shell ETAP High Pressure and Temperature Pipe-In-Pipe Pipeline Design and Fabrication," *Proceedings of the 18th International Conference on Offshore Mechanics and Arctic Engineering*, Newfoundland, Canada.

[4] Murphey, C. E., and Langner, C. G., 1985, "Ultimate Pipe Strength under Bending, Collapse and Fatigue," *Proceedings of the 4th International Conference on Offshore Mechanics and Arctic Engineering*, Vol. 1, pp. 467–477.

[5] Yeh, M. K., and Kyriakides, S., 1986, "On the Collapse of Inelastic Thick-Walled Tubes Under External Pressure," ASME J. Energy Resour. Technol., **108**, pp. 35–47.

[6] Dyau, J. Y., and Kyriakides, S., 1993, "On the Localization of Collapse in Cylindrical Shells Under External Pressure," Int. J. Solids Struct., **30**(4), pp. 463–482.

[7] Dyau, J. Y., and Kyriakides, S., 1993, "On the Propagation Pressure of Long Cylindrical Shells Under External Pressure," Int. J. Mech. Sci., **35**(8), pp. 675–713.

[8] Pasqualino, I. P., and Estefen, S. F., 2001, "A Nonlinear Analysis of the Buckle Propagation Problem in Deepwater Pipelines," Int. J. Solids Struct., **38**(46–47), pp. 8481–8502.

[9] Kyriakides, S., 2002, "Buckle Propagation in Pipe-in-pipe Systems, Part I: Experiments," Int. J. Solids Struct., **39**, pp. 351–366.

[10] Kyriakides, S., and Vogler, T. J., 2002, "Buckle Propagation in Pipe-in-Pipe Systems, Part II:Analysis," Int. J. Solids Struct., **39**, pp. 367–392.

[11] Fabian, O., 1977, "Collapse of Cylindrical, Elastic Tubes Under Combined Bending, Pressure and Axial Loads," Int. J. Solids Struct., **13**(12), pp. 1257–1270.

[12] Fabian, O., 1981, "Elastic-Plastic Collapse of Long Tubes Under Combined Bending and Pressure Load," Ocean Eng., **8**(3), pp. 295–330.

[13] Kyriakides, S., and Shaw, P. K., 1982, "Response and Stability of Elastoplastic Circular Pipes Under Combined Bending and External Pressure," Int. J. Solids Struct., **18**(11), pp. 957–973.

[14] Kyriakides, S., and Corona, E., 1988, "On the Collapse of Inelastic Tubes Under Combined Bending and Pressure," Int. J. Solids Struct., **24**(5), pp. 505–535.

[15] Kyriakides, S., and Corona, E., 1991, "An Experimental Investigation of the Degradation and Buckling of Circular Tubes Under Cyclic Bending and External Pressure," Thin-Walled Struct., **12**(3), pp. 229–263.

[16] Netto, T. A., and Estefen, S. F., 1994, "Ultimate Behaviour of Submarine Pipelines Under External Pressure and Longitudinal Bending," J. Constr. Steel Res., **28**, pp. 137–151.

[17] *ABAQUS User's and Theory Manuals*, 2000, Version 6.1, Hibbit, Karlsson, Sorensen, Inc.

[18] Netto, T. A., Santos, J. M. C., and Estefen, S. F., 2002, "Sandwich Pipes for Ultra-Deep Waters," *Proceedings of the 4th International Pipeline Conference*, Calgary, Canada, IPC02-27426.

[19] Pasqualino, I. P., Pinheiro, B. C., and Estefen, S. F., 2002 "Comparative Structural Analyses Between Sandwich and Steel Pipelines for Ultra-Deep Water," 21st International Conference on Offshore Mechanics and Arctic Engineering, June 23–28, Oslo, Norway.

[20] Sathe, S. N., Rao, G. S. S., Rao, K. V., and Devi, S., 1996, "The Effect of Composition on Morphological, Thermal, and Mechanical Properties of Polypropylene/Nylon-6/Polypropylene-G-Butyl Acrylate Blends," Polym. Eng. Sci., **36**, pp. 2443–2450.

Improved Form of a Fracture Mechanics Based Failure Probability Model for Brittle Materials

Rentong Wang
Graduate Student

Noriko Katsube
e-mail: katsube.1@osu.edu

Department of Mechanical Engineering,
Ohio State University,
Columbus, OH 43210

Robert R. Seghi
College of Dentistry,
Ohio State University,
Columbus, OH 43210

1 Introduction

Batdorf and Crose [1] combined the statistical analysis of failure for brittle materials by Weibull [2] with an appropriate fracture criterion based on fracture mechanics theory and extended this notion to multiaxial stress states. If an appropriate form of crack distribution is chosen, the cumulative failure probability function proposed by Batdorf and Crose [1] reduces to the Weibull distribution for uniaxial tensile stress states. In this work, we will show that the approximation of an infinitesimally small volumetric element may have been prematurely employed by Batdorf and Crose [1] in obtaining failure probability for an arbitrary volumetric element ΔV. The widely used failure probability formula based on this approximation may present some errors under certain conditions. We will derive an alternative formula without the use of this unnecessary approximation.

2 Theoretical Derivation

Batdorf and Crose [1] introduced the solid angle Ω containing the normals to all orientations for which the component of the applied stress normal to the crack plane is larger than the critical stress, i.e., $\sigma_n > \sigma_{cr}$. The solid angle Ω varies from zero to 4π for cracks contained inside a three-dimensional body, and it varies from zero to 2π for surface cracks based on the definition.

In general, a problem of crack propagation can only be one of the following two cases: (i) $\Omega/4\pi < 1$, propagation of a crack

depends on its orientation and there exists a range of orientation angle where a crack does not propagate and (ii) $\Omega/4\pi = 1$, propagation of a crack is independent of its orientation.

2.1 Case (I): $\Omega/4\pi < 1$. If there is only one crack and its crack plane is randomly orientated, the probability of failure caused by this single crack [Eq. (1) in Batdorf and Crose [1]] is given by

$$P_f = \Omega(\Sigma, \sigma_{cr})/4\pi, \tag{1}$$

where Σ is the applied stress, and σ_{cr} is the critical stress of the crack.

If the crack density is N, $N\Delta V$ represents the number of cracks inside the volumetric element ΔV. Therefore, the overall survival probability P_s is the multiplication of the survival probability of each crack.

$$P_s = [1 - \Omega(\Sigma, \sigma_{cr})/4\pi]^{N\Delta V}. \tag{2}$$

The overall failure probability of this volumetric element is given by

$$P_f = 1 - [1 - \Omega(\Sigma, \sigma_{cr})/4\pi]^{N\Delta V}. \tag{3}$$

In Batdorf and Crose [1], the failure probability is given by

$$P_f = (N\Delta V)(\Omega/4\pi). \tag{4}$$

If we let an arbitrary volumetric element ΔV approach zero, then $N\Delta V$ becomes small. In this case, by neglecting the higher order terms in the Taylor expansion, Eq. (3) can be reduced to Eq. (4). However, this approximation is premature and unnecessary. The theoretical derivation of failure probability prediction formula by Batdorf and Crose [1] is based on Eq. (4). In this work, we will derive the failure probability for the total volume V based on Eq. (3) instead of Eq. (4).

As in Batdorf and Crose [1], we will introduce the crack density function $N(\sigma_{cr})$ representing the number of cracks per unit volume with their critical stress less than or equal to σ_{cr}. The survival probability of this volumetric element ΔV for any possible cracks under stress Σ, $P_s(\Delta V, \Sigma)$, is the product of survival probability for every specific size crack with its critical stress in the range between σ_{cr}^{Min} and σ_{cr}^{Max}, where the values of the minimum critical stress σ_{cr}^{Min} and the maximum critical stress σ_{cr}^{Max} are determined by the actual stress status Σ and the fracture criterion.

$$P_s(\Delta V, \Sigma) = \prod_{m=1}^{M} P_s(\Delta V, \sigma_{cr}^m) = \prod_{m=1}^{M} [1 - \Omega(\Sigma, \sigma_{cr}^m)/4\pi]^{\Delta V dN(\sigma_{cr}^m)/d\sigma_{cr}\Delta\sigma_{cr}}, \tag{5}$$

where the critical stress range is divided into M equal increments $\Delta\sigma_{cr}$, and the critical stress in the m-th increment is denoted by σ_{cr}^m. By applying logarithmic operation to Eq. (5), letting $\Delta\sigma_{cr}$ approach zero, and reorganizing the resulting equations, we obtain

$$P_s(\Delta V, \Sigma) = \exp\left\{ \Delta V \int_{\sigma_{cr}^{Min}}^{\sigma_{cr}^{Max}} \ln[1 - \Omega(\Sigma, \sigma_{cr})/4\pi] \frac{dN(\sigma_{cr})}{d\sigma_{cr}} d\sigma_{cr} \right\}. \tag{6}$$

Therefore, the failure probability for the total volume V is given by

$$P_f^I(V) = 1 - \exp\left\{ \int_V \int_{\sigma_{cr}^{Min}}^{\sigma_{cr}^{Max}} \ln[1 - \Omega(\Sigma, \sigma_{cr})/4\pi] \frac{dN(\sigma_{cr})}{d\sigma_{cr}} d\sigma_{cr} dV \right\}, \tag{7}$$

where a superscript I is introduced to indicate that this equation is suitable for calculating the failure probability caused by the cracks whose propagation is dependent on orientation, i.e., Case (I): $\Omega/4\pi < 1$. Only in the special cases where $\Omega/4\pi$ is very small, i.e., $\Omega/4\pi \ll 1$, can Eq. (7) be reduced to the equation obtained by Batdorf and Crose [1] (Eq. (11) in Batdorf and Crose [1]) by approximating $\ln[1 - \Omega/4\pi]$ by $-\Omega/4\pi$.

2.2 Case (II): $\Omega/4\pi = 1$. In this case, propagation of a crack is independent of its orientation and it is solely determined by the size of the crack. Therefore, the failure probability is the probability of finding at least one sufficiently large crack. If a sufficiently large single crack is contained inside a total volume V, the failure probability for an arbitrary volumetric element ΔV is given by $\Delta V/V$. The survival probability for ΔV is then $(1 - \Delta V/V)$. Therefore, for a total volume V with K number of sufficient large cracks, the survival probability for the volumetric element ΔV is

$$P_s(\Delta V) = [1 - \Delta V/V]^K. \tag{8}$$

As in Case (I), we again introduce the crack density function $N(\sigma_{cr})$. The survival probability of ΔV for all cracks with size equal to or larger than the minimum critical size (with the maximum critical stress σ_{cr}^M corresponding to a stress state Σ is the product of survival probability for each specific size crack as in Case (I).

$$P_s(\Delta V, \Sigma) = \prod_{m=1}^{M} P_s(\Delta V, \sigma_{cr}^m) = \prod_{m=1}^{M} [1 - \Delta V/V]^{VdN(\tilde{\sigma}_{cr}^m)/d\sigma_{cr}\Delta\sigma_{cr}}$$
$$= [1 - \Delta V/V]^{\sum_{m=1}^{M} VdN(\sigma_{cr}^m)/d\sigma_{cr}\Delta\sigma_{cr}}. \tag{9}$$

Total volume V is divided into n number of volumetric elements and σ_{cr}^{Mi} and Σ^i, respectively, denote the maximum critical stress and the stress level in the i-th volumetric element ΔV^i. If we let $\Delta\sigma_{cr}$ approach zero in the i-th volumetric element ΔV^i, we obtain

$$P_s(\Delta V^i, \Sigma^i) = [1 - \Delta V^i/V]^{V\int_0^{\sigma_{cr}^{Mi}} dN(\sigma_{cr})/d\sigma_{cr} d\sigma_{cr}}. \tag{10}$$

Furthermore, we assume that the total volume V is equally divided. This leads to the overall survival probability in the entire volume V given by

$$P_s(V) = \prod_{i=1}^{n} P_s(\Delta V^i) = [1 - \Delta V/V]^{V/\Delta V \sum_{i=1}^{n} \Delta V^i \int_0^{\sigma_{cr}^{Mi}} dN(\sigma_{cr})/d\sigma_{cr} d\sigma_{cr}}. \tag{11}$$

As the number of volumetric elements n approaches infinity, $\Delta V/V$ approaches zero. In this case [3],

$$\lim_{\Delta V/V \to 0} [1 - \Delta V/V]^{V/\Delta V} = e^{-1}. \tag{12}$$

Therefore, the failure probability for the total volume V, $P_f(V)$, is given by

$$P_f^{II}(V) = 1 - \exp\left(- \int_V \int_0^{\sigma_{cr}^M} \frac{dN(\sigma_{cr})}{d\sigma_{cr}} d\sigma_{cr} dV \right), \tag{13}$$

where a superscript II is introduced to indicate that this equation is suitable for calculating the failure probability caused by the cracks whose propagation is not dependent on orientation, i.e., Case (II): $\Omega/4\pi = 1$.

In Summary, the probability of failure caused by all types of cracks $P_f(V)$ is given by

$$P_f(V) = 1 - (1 - P_f^I)(1 - P_f^{II}). \tag{14}$$

3 Examples

3.1 Example 1: Failure Probability Predictions for Uniform Stress States. We will examine the effect of differences in the above formulations on the failure probability predictions caused only by surface cracks given the same crack density parameters by using three simple stress states. Since only surface cracks are considered, $\Omega/2\pi$ is used instead of $\Omega/4\pi$, and surface integral instead of volumetric integral will be used in the above formulations. In order to evaluate the above equations and visualize the differences, we assume the following form of the crack density function $N(\sigma_{cr})$ as in Chao and Shetty [4]

$$N(\sigma_{cr}) = k\sigma_{cr}^m \tag{15}$$

where k and m, respectively, are scale and shape parameters. Here, we choose them as $m = 4.917$ and $k = 4.95 \times 10^{-11}$ mm^{-2} MPa$^{-4.917}$. These numerical values are reasonable in representing glass surfaces sanded by 600 grit S_iC sand papers. In addition, we consider the case where the stress status is uniform, and the specimen surface area is $A = 1$ mm^2.

To determine Ω, a fracture criterion is required. Here, we assume that a crack propagates when the stress normal to the crack surface σ_n reaches its critical value σ_{cr}, i.e.,

$$\sigma_n = \sigma_{cr}. \tag{16}$$

In the following examples, σ_1 and σ_2 are the two principal stresses on the specimen surface.

Example 1A. $\sigma_1 = \sigma_2 \geq 0$

Since the normal stress in any direction is σ_1 (or σ_2), if the critical stress σ_{cr} for a crack is less than σ_1, i.e., $0 \leq \sigma_{cr} \leq \sigma_1$, the crack will propagate regardless of its orientation, i.e., $\Omega/2\pi = 1$. Equation (13) is used by replacing the volume integral with the area integral. If $\sigma_{cr} > \sigma_1$ for a crack, the crack will not propagate regardless of its orientation. The formulation by Batdorf and Crose [1] becomes identical to Eq. (13) obtained in this work.

Example 1B. $\sigma_1 > \sigma_2 \geq 0$

If the critical stress σ_{cr} for a crack is between zero and σ_2, i.e., $0 \leq \sigma_{cr} \leq \sigma_2$, the stress normal to the crack plane is always larger than the critical stress, i.e., $\sigma_n > \sigma_{cr}$, regardless of its direction. This again leads to $\Omega/2\pi = 1$, and Eq. (13) is used with the substituted area integral. If the critical stress σ_{cr} for a crack is between σ_2 and σ_1, i.e., $\sigma_2 < \sigma_{cr} \leq \sigma_1$, the crack propagation is influenced by the crack orientation, and the critical range of angle Ω needs to be calculated from the fracture criterion Eq. (16). Under this condition, we have

$$\Omega = 2\cos^{-1}\left(\frac{2\sigma_{cr} - \sigma_1 - \sigma_2}{\sigma_1 - \sigma_2} \right). \tag{17}$$

Since $\Omega/2\pi < 1$, Eq. (7) with the substituted area integral is used. The overall failure probability given by Eq. (14) leads to
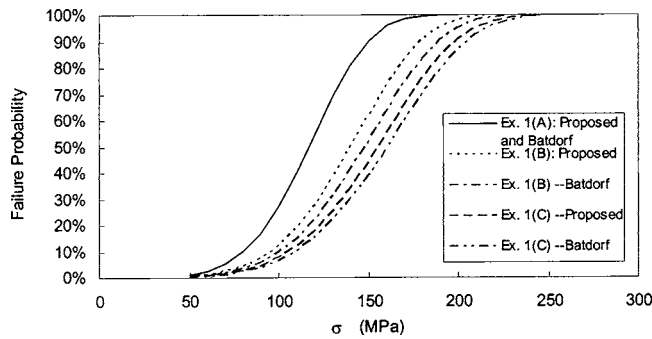
Fig. 1 Comparison of failure probability prediction

Fig. 2 Biaxial flexure test data

$$P_f = 1 - \exp\left(-A\int_0^{\sigma_2} \frac{dN(\sigma_{\mathrm{cr}})}{d\sigma_{\mathrm{cr}}}d\sigma_{\mathrm{cr}} + A\int_{\sigma_2}^{\sigma_1}\ln\left[1\right.\right.$$
$$\left.\left. - \Omega/2\pi\right]\frac{dN(\sigma_{\mathrm{cr}})}{d\sigma_{\mathrm{cr}}}d\sigma_{\mathrm{cr}}\right). \tag{18}$$

Correspondingly, the overall failure probability based on Batdorf and Crose [1] can be obtained as follows:

$$P_r = 1 - \exp\left(-A\int_0^{\sigma_1} \frac{\Omega}{2\pi}\frac{dN(\sigma_{\mathrm{cr}})}{d\sigma_{\mathrm{cr}}}d\sigma_{\mathrm{cr}}\right). \tag{19}$$

Example 1C. $\sigma_1 \geqslant 0 > \sigma_2$

In this case, the critical stress σ_{cr} is between 0 and σ_1, and the crack propagation depends on the crack orientation. The range of critical angle Ω can be calculated from Eq. (17). For this case, any crack propagation depends on its orientation, i.e., $\Omega/2\pi < 1$. Therefore, Eq. (7) is employed with the substituted area integral.

In order to graphically present some results, we choose σ_1 and σ_2 as follows:

For Example 1A, $\sigma_1 = \sigma_2 = \sigma$ ($\sigma > 0$)
For Example 1B, $\sigma_1 = \sigma$, $\sigma_2 = \sigma/3$
For Example 1C, $\sigma_1 = \sigma$, $\sigma_2 = -\sigma/3$

Based on Eq. (15) and the above assumptions, the failure probability as a function of σ is calculated for both formulations as shown in Fig. 1.

3.2 Example 2: Crack Density Parameter Determination Based on Biaxial Flexure Tests. Crack density parameters are often determined by curve fitting of data from failure strength experiments. We will examine the effect of differences in the above formulations on the resulting statistical parameters determined from a set of experimental data. For this purpose, biaxial flexure tests were conducted using borosilicate glass specimens. The 1 mm thick glass disks with 15.9 mm diameter were supported at the edge by a ring of bearings and loaded on the top center through a tungsten carbide (WC) spherical ball indenter with a diameter of 10 mm. The surfaces of the glass disks were sanded on a rotating wheel with 600 grit SiC sandpaper under water coolant. The experiments were carried out on the Universal Testing Machine (Instron Model 4020, Canton, Mass.) at a crosshead speed of 0.01 mm/min. A total of 34 specimens were used to obtain the experimental failure probability distribution. The fracture initiation load P was recorded for each specimen and the cumulative probability of crack initiation was obtained.

The crack density function $N(\sigma_{\mathrm{cr}})$ is assumed to be in the form of Eq. (15). By using this equation in the formulation by Batdorf and Crose [1] or Eq. (14) obtained in this work, we can derive the failure probability function P_f as follows:
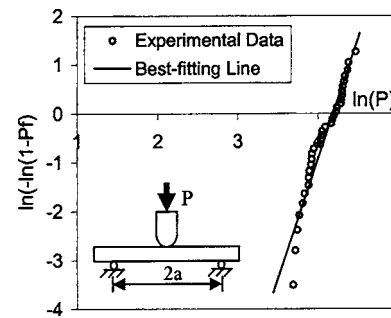
$$P_f = 1 - \exp(-e^B P^m), \tag{20}$$

where

$$B = \ln\left[\frac{a^2 m k \mathrm{I}_D}{(\pi t^2)^m}\right]. \tag{21}$$

For the formulation based on Batdorf and Crose [1], I_D in Eq. (21) is given by

$$\mathrm{I}_D = \int_{r/a=0}^{1}\int_{\sigma_{\mathrm{cr}}/P/\pi t^2=0}^{\sigma_1/P/\pi t^2} 2\pi\left(\frac{\Omega}{2\pi}\right)\left(\frac{\sigma_{\mathrm{cr}}}{P/\pi t^2}\right)^{m-1}\left(\frac{r}{a}\right)d\left(\frac{\sigma_{\mathrm{cr}}}{P/\pi t^2}\right)d\left(\frac{r}{a}\right). \tag{22}$$

For the formulation obtained in this work, I_D is given by

$$\mathrm{I}_D = \int_{r/a=0}^{1} 2\pi\left(\frac{r}{a}\right)d\left(\frac{r}{a}\right)\left[\int_{\sigma_{\mathrm{cr}}/P/\pi t^2=0}^{\sigma_2/P/\pi t^2}\left(\frac{\sigma_{\mathrm{cr}}}{P/\pi t^2}\right)^{m-1}d\left(\frac{\sigma_{\mathrm{cr}}}{P/\pi t^2}\right)\right.$$
$$\left. - \int_{\sigma_{\mathrm{cr}}/P/\pi t^2=\sigma_2/P/\pi t^2}^{\sigma_1/P/\pi t^2}\ln\left(1-\frac{\Omega}{2\pi}\right)\left(\frac{\sigma_{\mathrm{cr}}}{P/\pi t^2}\right)^{m-1}d\left(\frac{\sigma_{\mathrm{cr}}}{P/\pi t^2}\right)\right], \tag{23}$$

where a is the radius of the support ring, t is the thickness of the specimen, and Ω is the critical angle as defined before.

By taking logarithmic operation twice on Eq. (20) and best fitting the experimental data points with a straight line as in Fig. 2, we determine that the parameters m and B are 4.917 and -20.60, respectively. We can also obtain the crack density parameter k from Eq. (21). The determined crack density parameters m and k based on the two distinct formulations are summarized in Table 1.

4 Discussions

In Example 1A, there is no difference in terms of actual formulations between the work based on Batdorf and Crose [1] and the present work as shown in Fig. 1. In Examples 1B and 1C, the results shown in Fig. 1 reflect the differences caused by the two formulations. While these are not dramatic differences, the failure probability of the proposed formulation is higher than that of the formulation by Batdorf and Crose [1] given the same stress level and therefore it provides a more conservative estimate.

In Example 1B, if σ_{cr} falls between σ_2 and σ_1, the two formulations are significantly different as in Eqs. (18) and (19). When the value of σ_{cr} approaches that of σ_1, the value of $\Omega/2\pi$ approaches zero [see Eq. (17)]. In this case, $\ln(1-\Omega/2\pi)$ in Eq. (18) can be approximated by $-\Omega/2\pi$ in Eq. (19) by ignoring the higher order terms of the Taylor expansion. Because of this, the contribution to the overall failure probability by cracks with critical

Table 1 Crack density parameters determined by biaxial flexure tests

	m	$k \; [\mathrm{mm}^{-2}\,\mathrm{MPa}^{-4.917}]$
Batdorf and Crose Formulation	4.917	4.95×10^{-11}
Proposed Formulation	4.917	3.99×10^{-11}

stress close to σ_1 is very similar between the two formulations. When the value of σ_{cr} is larger than σ_2 but not close to σ_1, however, $\ln(1-\Omega/2\pi)$ cannot be approximated by $-\Omega/2\pi$ since $\Omega/2\pi$ is not small. As in Eq. (15), the number of cracks increases rapidly as critical stress σ_{cr} increases, indicating that the number of small cracks is much larger than the number of large cracks. Therefore, the crack density function, serving as a weight function in Eqs. (18) and (19), favors the small cracks which give a range of small $\Omega/2\pi$. In other words, the numerical difference between the two formulations in Example 1B can be attributed to the failure probability primarily due to cracks with critical stress larger than σ_2 but not close to σ_1.

Similarly in Example 1C, there are significant differences between $\ln(1-\Omega/2\pi)$ and $-\Omega/2\pi$ when the critical stress is larger than zero but not close to σ_1. However, the number of cracks corresponding to this range of critical stress is relatively small. Therefore, the contribution to the overall failure probability by such cracks is relatively small. On the contrary, when the critical stress is close to σ_1, $\ln(1-\Omega/2\pi)$ can be approximated by $-\Omega/2\pi$. The number of cracks corresponding to this range of critical stress is relatively large. Therefore, the contribution to the overall failure probability of such cracks is large, and the numerical differences between the two formulations remain small. By adding these small differences in two regions of critical stress, the numerical difference between the two formulations in Example 1C remains relatively small.

We have discussed the reason that there are not significant numerical differences between the two formulations in the results of Example 1 while the failure probability prediction formulas are significantly different. It is important to note that we chose a specific form of crack density function N expressed in Eq. (15). Given a more general form of crack density function, the numerical results based on these two formulations may be different.

In Example 2, the values of k are different for the two formulations while the values of m are the same. The difference is 24% using the result based on the proposed formulation as the baseline. This difference may lead to possible errors in failure probability prediction.

While the formulation by Batdorf and Crose [1] has been successfully used for many practical applications, the formulation is based on Eq. (4) where the premature assumption infinitesimally small volume element is implicitly employed. In order to improve accuracy of failure probability prediction, it is best to employ Eq. (14).

References

[1] Batdorf, S. B., and Crose, J. G., 1974, "A Statistical Theory for the Fracture of Brittle Structures Subjected to Nonuniform Polyaxial Stress," ASME J. Appl. Mech., **41**(2), pp. 459–464.
[2] Weibull, W. A., 1951, "Statistical Distribution Function of Wide Applicability," ASME J. Appl. Mech., **18**(3), pp. 293–297.
[3] Leithold, L., 1976, *The Caluculus with Analytica Geometry*, Harper & Row, New York, pp. 411–412.
[4] Chao, L. Y., and Shetty, D. K., 1991, "Reliability Analysis of Structural Ceramics Subjected to Biaxial Flexure," J. Am. Ceram. Soc., **74**(2), pp. 333–344.

The Tallest Two-Segment Column

C. Y. Wang

Member ASME

Department of Mathematics and Department of
Mechanical Engineering,
Michigan State University,
East Lansing, MI 48824
e-mail: cywang@mth.msu.edu

1 Introduction

The study of the tallest standing column is basic in the design of tall structures, especially antennae. Greenhill [1] first found the maximum height of the uniform column that does not buckle under its own weight. The height can be extended if the column tapers towards the top. The optimum shape of the tallest nonuniform column is more difficult to determine. Keller and Niordson [2], and Cox and McCarthy [3] discussed its asymptotic properties, and Hu and Kirmser [4] integrated the shape numerically. It was found that the optimum column shrinks very fast at the base, then more slowly, and becomes very sharp at the apex.

However, such an optimal column is very difficult to manufacture due to the vastly nonuniform radius variations. A more practical solution is to weld two uniform rods of different radii together. The purpose of the present paper is to find the optimum dimensions (lengths and diameters) of a compound column such that for a given total weight, the maximum height can be attained.

2 Formulation

We assume the column is composed of two segments, each with uniform density (weight per length) and uniform flexural rigidity (Fig. 1). Let the subscript 1 denote the top segment of length $(1-a)L$, and the subscript 2 denote the bottom segment of length aL. For small deflections, a moment balance on an elemental length for the top segment gives

$$dm + \rho_2(L - s')\theta_2 \, ds' = 0. \tag{1}$$

The moment is proportional to the curvature

$$m = EI_2 \frac{d\theta_2}{ds'} \tag{2}$$

where s' is the arc length from the base, θ is the local angle of inclination, and ρ, EI are the weight per length and the flexural rigidity, respectively. Equations (1) and (2) give

$$\frac{d^2\theta_2}{ds^2} + \beta_2(1 - s)\theta_2 = 0. \tag{3}$$

Here $s = s'/L$ is the normalized distance and $\beta_2 \equiv \rho_2 L^3/EI_2$ is a nondimensional weight parameter. The general solution to Eq. (3) is in terms of Bessel functions

$$\theta_2 = A s^t J_t(s) + B s^t J_{-t}(s) \tag{4}$$

where $t \equiv 1/3$, A and B are constants and

$$s = \frac{2}{3}(1 - s)^{3/2}\beta_2^{1/2}. \tag{5}$$

For the bottom segment a similar moment balance yields

$$\frac{d^2\theta_1}{ds^2} + \beta_1[\lambda(1 - a) + a - s]\theta_1 = 0 \tag{6}$$

where $\beta_1 \equiv \rho_1 L^3/EI_1$ and $\lambda \equiv \rho_2/\rho_1$ is the density ratio. The general solution to Eq. (6) is

$$\theta_1 = C\eta^t J_t(\eta) + D\eta^t J_{-t}(\eta) \tag{7}$$

where C and D are constants and

$$\eta = \frac{2}{3}[\lambda(1 - a) + a - s]^{3/2}\beta_1^{1/2}. \tag{8}$$

The boundary conditions are that the top is free

$$\frac{d\theta_2}{ds}(1) = 0 \tag{9}$$

the base is fixed

$$\theta_1(0) = 0 \tag{10}$$

and the angles and moments match at the joint

$$\theta_1(a) = \theta_2(a) \tag{11}$$

$$EI_1\frac{d\theta_1}{ds}(a) = EI_2\frac{d\theta_2}{ds}(a). \tag{12}$$

For nontrivial solutions, after some work, Eqs. (9)–(12) give the characteristic equation

$$\gamma\lambda[J_t(\eta_1)J_{-t}(\eta_a) - J_{-t}(\eta_1)J_t(\eta_a)]\{s_a[J_{1-t}(s_a) - J_{-1-t}(s_a)]$$
$$- 2tJ_{-t}(s_a)\} + J_{-t}(s_a)\{\eta_a[J_{-t}(\eta_1)\langle J_{1+t}(\eta_a) - J_{t-1}(\eta_a)\rangle + J_t(\eta_1)$$
$$\times\langle J_{-1-t}(\eta_a) - J_{1-t}(\eta_a)\rangle + 2t[J_{-t}(\eta_a)J_t(\eta_1) - J_{-t}(\eta_1)J_t(\eta_a)]\}$$
$$= 0 \tag{13}$$

where $\gamma \equiv EI_2/EI_1$ and

$$\eta_1 = \eta|_{s=1}, \quad \eta_a = \eta|_{s=a}, \quad s_a = s|_{s=a}. \tag{14}$$

For given a, γ, λ and since $\beta_1 = \gamma\beta_2/\lambda$, Eq. (13) can be solved numerically for the lowest value of β_2, which can represent the buckling weight.

3 Optimum Shape

Consider the important case of the two-segment solid column made of the same elastic material (E=Young's modulus, σ =density) and a similar circular cross section (radii r_1 and $r_2 < r_1$). Thus

$$\gamma = \lambda^2 = (r_2/r_1)^4. \tag{15}$$

Actually Eq. (15) is satisfied by a variety of other similar cross-sectional shapes. For each λ and a Eq. (13) is solved by a root finder algorithm yielding the buckling parameter β_2 in Table 1. When $a=0$ or $\lambda=1$ the compound column is equivalent to a single uniform column. The value for the buckling parameter is 7.837 35 (see, e.g., [5]). On the other hand, if $a=1$ the buckling parameter is 7.837 35/λ. The total volume of the column is

$$V' = aL\pi r_1^2 + (1 - a)L\pi r_2^2. \tag{16}$$

The normalized volume is

$$V = \frac{V'}{4\pi\sigma L^4/E} = \frac{1}{\beta_2}\left(1 - a + \frac{a}{\lambda}\right). \tag{17}$$

For given a, λ we obtain β_2 from Table 1 and the value of V from Eq. (17). The dimensions of the tallest two-segment column for a given amount of weight or volume is equivalent to that which minimizes the volume for given height. Thus we search for the minimum V in Table 2. After some refinement, we find minimum V is 0.038 41 occurring at $a=0.661$ and $\lambda=0.131$. For given height L that needs to be reached, the optimum dimensions are
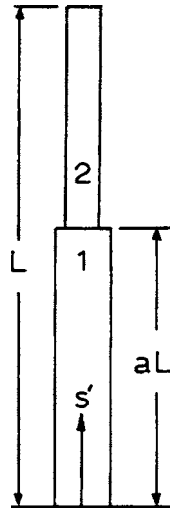
Fig. 1 The two-segment column

$$r_1 = \sqrt{\frac{4V\sigma L^3}{E[a-(1-a)\lambda]}} = 0.499\sqrt{\frac{\sigma L^3}{E}}, \quad r_2 = \sqrt{\lambda}\,r_1 = 0.362\,r_1$$

$$(18)$$

and the lengths of the two segments are $0.661L$ and $0.339L$, respectively. Notice that the radial dimensions depend on the material properties of the column.

The present analysis can be extended to two-segment columns with different material properties or to multisegment columns.

References

[1] Greenhill, A. G., 1881, "Determination of the greatest height consistent with stability that a vertical pole or mast can be made, and of the greatest height to which a tree of given proportions can grow," Proc. Cambridge Philos. Soc. **4**, pp. 65–73.

[2] Keller, J. B. and Niordson, F. I., 1966, "The tallest column," J. Math. Mech. **16**, pp. 433–446.

[3] Cox, S. J. and McCarthy, C. M., 1998, "The shape of the tallest column," SIAM J. Math. Anal. **29**, pp. 547–554.

[4] Hu, K. K. and Kirmser, P. G., 1980, "Numerical solution of a non-linear differential-integral equation for the optimal shape of the tallest column," Int. J. Eng. Sci. **18**, pp. 333–339.

[5] Wang, C. Y., 1986, "A critical review of the heavy elastica," Int. J. Mech. Sci. **28**, pp. 549–559.

Table 1 Buckling parameter $\beta_2 = \rho_2 L^3/EI_2$

λ/a	0.1	0.3	0.5	0.7	0.9
0.1	10.72	22.55	60.67	183.6	103.7
0.3	10.43	20.28	43.44	48.40	32.28
0.5	9.898	16.42	24.43	23.45	18.16
0.7	9.160	12.34	14.51	13.99	12.20
0.9	8.293	9.096	9.448	9.332	8.954

Table 2 The total volume V

λ/a	0.1	0.3	0.5	0.7	0.9
0.1	0.177	0.164	0.0907	0.0398	0.0877
0.3	0.118	0.0838	0.0499	0.0544	0.0960
0.5	0.111	0.0792	0.0614	0.0725	0.105
0.7	0.114	0.0915	0.0837	0.0929	0.114
0.9	0.122	0.114	0.112	0.116	0.123

The Fault in the Stress Analysis of Pseudo-Stress Function Method

B. X. Xu
e-mail: xubaixiang@sohu.com

M. Z. Wang
e-mail: wangminz@pku.edu.cn

Department of Mechanics and Engineering Science,
Peking University,
100871, Beijing, PRC

Although Peng Yafei and his co-workers discovered some faults with the pseudo-stress function method suggested by Y. S. Lee in 1987, the authors did not provide convincing arguments. We investigate the crucial assumption in Lee's method by rewriting it as the form of real part and imaginary part. Through a specific counterexample, we point out that the crucial assumption in Lee's theory is untenable. Namely, for given Airy's stress function, it cannot be guaranteed that the pseudo-stress function $\Lambda(x,y)$ exists. The root cause of the fault with Lee's method is found in this paper. [DOI: 10.1115/1.1935526]

1 Introduction

In 1987, Y. S. Lee and his co-workers [1–4] came up with an analytic method that is known as the pseudo-stress function method. The pseudo-stress function method is similar to the Airy stress function method in linear elasticity and has been used to analyze the stress field of power-hardening materials. However, this method was challenged by Peng Yafei and his co-workers in 1996 [5]. In Peng's paper, the authors declared that they had found some crucial problems in Lee's theory and presented some discus-sion. But the authors did not explain why Lee's method was wrong and failed to point out the major cause that is responsible for the fault with Lee's theory. In order to show the fallacy in Lee's theory, a more convincing reasoning process is needed.

In the present paper, efforts are paid to find the radical fault in Lee's theory. Through a specific counterexample, we prove that the basic assumption used in Lee's theory cannot always hold true. Namely, for given Airy's stress function U and constant n ($n \neq 0$), it cannot be guaranteed that the pseudo-stress function exists.

2 Discussion

We can see that Lee's theory is completely based on the following assumption [1],

$$
\begin{cases}
\dfrac{\partial^2 \Lambda}{\partial z^2} = \xi^n \dfrac{\partial^2 U}{\partial z^2} \\[2mm]
\dfrac{\partial^2 \Lambda}{\partial \bar{z}^2} = \xi^n \dfrac{\partial^2 U}{\partial \bar{z}^2}
\end{cases},
\tag{1}
$$

where

$$
\xi^2 = \frac{\partial^2 U}{\partial z^2}\frac{\partial^2 U}{\partial \bar{z}^2},
\tag{2}
$$

$$
n = \frac{1-m}{m}.
\tag{3}
$$

Then, Lee obtained the compatibility equation in terms of the pseudo-stress function Λ as follows:

$$
\frac{\partial^4 \Lambda}{\partial z^2 \partial \bar{z}^2} = 0.
\tag{4}
$$

However, Eq. (4) cannot hold true because there is a serious problem in Lee's assumption Eqs. (1). The following discussion in this section will show that the pseudo-stress function $\Lambda(x,y)$ does not always exist for given Airy's stress function U and constant n ($n \neq 0$). This problem is fatal to the whole theory.

For given Airy's stress function U and n, we suppose $\Lambda(x,y) = P(x,y) + iQ(x,y)$. Rewriting Eqs. (1), we have

$$
\begin{cases}
\dfrac{1}{4}\left\{\dfrac{\partial^2 P}{\partial x^2} - \dfrac{\partial^2 P}{\partial y^2} + 2\dfrac{\partial^2 Q}{\partial x\partial y} + i\left[-2\dfrac{\partial^2 P}{\partial x\partial y} + \left(\dfrac{\partial^2 Q}{\partial x^2} - \dfrac{\partial^2 Q}{\partial y^2}\right)\right]\right\} = \dfrac{1}{4}\xi^n\left(\dfrac{\partial^2 U}{\partial x^2} - \dfrac{\partial^2 U}{\partial y^2} - 2i\dfrac{\partial^2 U}{\partial x\partial y}\right) \\[4mm]
\dfrac{1}{4}\left\{\dfrac{\partial^2 P}{\partial x^2} - \dfrac{\partial^2 P}{\partial y^2} - 2\dfrac{\partial^2 Q}{\partial x\partial y} + i\left[2\dfrac{\partial^2 P}{\partial x\partial y} + \left(\dfrac{\partial^2 Q}{\partial x^2} - \dfrac{\partial^2 Q}{\partial y^2}\right)\right]\right\} = \dfrac{1}{4}\xi^n\left(\dfrac{\partial^2 U}{\partial x^2} - \dfrac{\partial^2 U}{\partial y^2} + 2i\dfrac{\partial^2 U}{\partial x\partial y}\right)
\end{cases}.
\tag{5}
$$

Considering Airy's stress function U is a real function, from Eq. (2) we know ξ^n is also a real function and has the following form:

$$
\xi^n = \left\{\frac{1}{16}\left[\left(\frac{\partial^2 U}{\partial x^2} - \frac{\partial^2 U}{\partial y^2}\right)^2 + 4\left(\frac{\partial^2 U}{\partial x\partial y}\right)^2\right]\right\}^{n/2}.
\tag{6}
$$

Consequently, from Eqs. (5), we obtain

$$
\begin{cases}
\dfrac{\partial^2 P}{\partial x^2} - \dfrac{\partial^2 P}{\partial y^2} \pm 2\dfrac{\partial^2 Q}{\partial x\partial y} = \xi^n\left(\dfrac{\partial^2 U}{\partial x^2} - \dfrac{\partial^2 U}{\partial y^2}\right) \\[4mm]
\mp 2\dfrac{\partial^2 P}{\partial x\partial y} + \left(\dfrac{\partial^2 Q}{\partial x^2} - \dfrac{\partial^2 Q}{\partial y^2}\right) = \mp 2\xi^n\dfrac{\partial^2 U}{\partial x\partial y}
\end{cases}.
\tag{7}
$$

Thus,

$$
\begin{cases}
\dfrac{\partial^2 P}{\partial x^2} - \dfrac{\partial^2 P}{\partial y^2} = \xi^n\left(\dfrac{\partial^2 U}{\partial x^2} - \dfrac{\partial^2 U}{\partial y^2}\right) \\[4mm]
\dfrac{\partial^2 P}{\partial x\partial y} = \xi^n\dfrac{\partial^2 U}{\partial x\partial y}
\end{cases}.
\tag{8}
$$

$$\begin{cases} \dfrac{\partial^2 Q}{\partial x \partial y} = 0 \\[2mm] \dfrac{\partial^2 Q}{\partial x^2} - \dfrac{\partial^2 Q}{\partial y^2} = 0 \end{cases}. \qquad (9)$$

Equations (8) and (9) are the differential equations that the real part and imaginary part of $\Lambda(x,y)$ should satisfy respectively. From Eqs. (9) one obtains

$$Q(x,y) = A(x^2 + y^2) + Bx + Cy + D, \qquad (10)$$

where A, B, C, D are constants.

However, when $n \neq 0$, Eqs. (8) usually have no solutions. We consider the case of $n=2$ and $U=xy^2$. Substituting $n=2$ and $U=xy^2$ into Eqs. (8), we obtain

$$\begin{cases} \dfrac{\partial^2 P}{\partial x^2} - \dfrac{\partial^2 P}{\partial y^2} = -\dfrac{1}{2}x^3 - 2xy^2 \\[2mm] \dfrac{\partial^2 P}{\partial x \partial y} = \dfrac{1}{2}x^2 y + 2y^3 \end{cases}. \qquad (11)$$

It is not difficult for us to know that Eqs. (11) have no solutions at all. Therefore, the existence of $\Lambda(x,y)$ is not ensured, at least in the case of $n=2$, $U=xy^2$. Generally speaking, Eqs. (8) do not have solutions since a single function $P(x,y)$ is required to satisfy two distinct differential equations.

3 Conclusion

From what have been discussed above, we find, for given Airy's stress function U and constant n ($n \neq 0$), it cannot be guaranteed that the function $\Lambda(x,y)$ exists, which is required to satisfy Eqs. (1). So, we can safely draw our conclusion that Lee's assumption is unwarranted and the pseudo-stress function method suggested by Lee collapses.

Acknowledgment

References

[1] Lee, Y. S., and Smith, L. C., 1988, "Analysis of a Power-law Material Containing a Single Hole Subjected to a Uniaxial Tensile Stress Using the Complex Pseudo-Stress Function," ASME J. Appl. Mech., **55**, pp. 267–274.
[2] Lee, Y. S., 1987, "Stress Analysis and Investigation of Void Formation Under External Stress of Power-law Creep Materials Using Complex Variables and the Pseudo-stress Function," Dissertation abstracts, University of Pittsburgh, Pittsburgh, PA.
[3] Lee, Y. S., and Gong, H., 1987, "Application of Complex Variables and Pseudo-stress Analysis of Single Rigid Inclusion in Power-law Materials Subjected to Simple Tension and Pure Shear," Int. J. Mech. Sci., **29**, pp. 669–694.
[4] Lee, Y. S., and Batt, T. J., 1989, "Investigation of Void Formation on a Bonded Interface of Power Law Creep Materials Containing a Cylindrical Particle," Acta Mech., **79**, pp. 183–205.
[5] Peng Yafei, Shen Yapeng, and Chen Yiheng, 1996, "Some Faults in the Stress Analysis of Power-Hardening Materials with Pseudo-Stress Function Method," ASME J. Appl. Mech., **63**, pp. 554–555.

Inaccuracy in the Detection of Molecules Using Two Microcantilever-Based Methods

Cheng Luo

Biomedical Engineering and Institute for Micromanufacturing, Louisiana Tech University, 911 Hergot Avenue, Ruston, LA 71272
e-mail: chengluo@latech.edu

Based on mass-loading effect on a microcantilever, there are two approaches for sensing the presence of molecules: dynamic and static methods. In this paper, we demonstrate that the two methods actually use the same form of relationships for their sensing purposes and, if the designed adhesion region of a cantilever is only partially occupied by molecules, then neither method can be applied to accurately determine the number of molecules adsorbed. [DOI: 10.1115/1.1938201]

1 Introduction

The microcantilever-based dynamic method uses resonant frequency changes due to mass loading. This method has been applied to detect vapor adsorption [1], the amount of uniformly coated material [2], the number of Escherichia coli cells bound [3], the mass of an Escherichia coli cell [4], etc. In this work, using the general relationship derived for the method, we demonstrate that, when the designed adhesion region of a cantilever is not fully occupied by molecules, this method cannot be applied to accurately determine the number of molecules adsorbed, which is different from what was claimed in previous studies [3,4]. We also discuss the possibility of determining the number of molecules adsorbed according to the relationship between the static bending of a cantilever and the mass loading.

2 Dynamic Method

Suppose that n molecules of the same type are adsorbed on a rectangular cantilever. Our assumptions for the dynamic method are as follows: (i) the length, width, and thickness of the cantilever are much larger than the size of a molecule; (ii) the molecules have identical sizes; (iii) the damping effect is negligible; and (iv) after adsorption of molecules, the higher natural frequencies of the mass-loaded cantilever are much larger than the fundamental natural frequency. Set up a coordinate system at the top surface of the cantilever (Fig. 1), and the x-axis coincides with the middle line of the top surface. Only translational vibration in the x-z plane is considered. Then, according to Dunkerley's equation [5], we have

$$\frac{1}{(2\pi f)^2} - \frac{1}{(2\pi f_0)^2} = a_1 m_1 + a_2 m_2 + \cdots + a_n m_n \qquad (1)$$

where $m_i \, (i=1,2,\cdots,n)$ is the mass of the ith molecule and, f_0 and f are the fundamental natural frequencies of the cantilever before and after mass (molecule) loading, respectively. Both f and f_0 can be measured by experiments [1–4]. The expression of f_0 is [6]

$$f_0 = 0.559 \sqrt{\frac{EI}{m_0 l^3}} \qquad (2)$$

where m_0 is the mass of the cantilever, E Young's modulus, and I moment of inertia. It can be observed from Eq. (1) that f is always less than f_0. Since $f_0 - f$ (the frequency change) is monotonously increasing with $1/(2\pi f)^2 - 1/(2\pi f_0)^2$ and the value of $f_0 - f$ has one-to-one correspondence with that of $1/(2\pi f)^2 - 1/(2\pi f_0)^2$, then $1/(2\pi f)^2 - 1/(2\pi f_0)^2$ can be considered a generalized frequency change. According to Eq. (1), it is more straightforward to use $1/(2\pi f)^2 - 1/(2\pi f_0)^2$, instead of $f_0 - f$, to consider the frequency shift caused by the adsorption of molecules. Therefore, in this paper, $1/(2\pi f)^2 - 1/(2\pi f_0)^2$ is used to represent the frequency change. In Eq. (1), a_i is the so-called influence coefficient of the ith molecule and is defined as the vertical deflection of the cantilever at the position of the ith molecule due to a unit force at the same point (Fig. 1). The expression of a_i is given by [7]

$$a_i = \frac{x_i}{3EI} \qquad (3)$$

where x_i is the x coordinate of the position occupied by the ith molecule. The relationship (1) can also be rewritten in the integration form as

$$\frac{1}{(2\pi f)^2} - \frac{1}{(2\pi f_0)^2} = \int_A \frac{\rho x^3}{3EI} dx dy \qquad (4)$$

where A is the surface area occupied by the monolayer of adsorbed molecules. In Eq. (4), ρ is a constant representing the areal mass density of molecules in A and depends on the mass and size of the molecule. Using Eq. (1) or (4), the mass of a single molecule and the number of molecule layers adsorbed can be determined as shown below.

i. Determination of the mass of a single molecule. When only one molecule is adsorbed on the cantilever, the relationship (1) becomes

$$m = \frac{3EI}{x_0^3} \left[\frac{1}{(2\pi f)^2} - \frac{1}{(2\pi f_0)^2} \right] \qquad (5)$$

where x_0 is the x coordinate of the position occupied by the molecule. After f is measured experimentally and x_0 is found, m can be determined using Eq. (5). This idea was used in [4] to determine the mass of an Escherichia coli cell of 6.65×10^{-13} g with the assistance of a silicon-nitride microcantilever. In order to detect the mass of a much lighter particle, it is expected that further miniaturization of a cantilever from micro- to nanoscale is needed because nanocantilevers have much higher sensitivity in detecting molecules [8,9].

According to the size of the particle, its position x_0 can be found using a microscope of the comparative resolution. For example, for a microparticle the commonly used optical microscope of a resolution of 1 μm can be employed to find its location on the cantilever, whereas for a nanoparticle the LEO Transmission Electron Microscope of a resolution of 0.2 nm [10] is a good tool for the purpose of observation.

ii. Determination of the number of molecule layers adsorbed. In the case the top surface is fully occupied by k $(k=1,2,\dots)$ layer(s) of molecules, according to the relationship (4), we have

$$k = \frac{12EI}{\rho t l^4} \left[\frac{1}{(2\pi f)^2} - \frac{1}{(2\pi f_0)^2} \right] \qquad (6)$$

where l is the cantilever length and t the cantilever width. The mass of the molecule can be obtained, for instance, by the approach of case i, and the size of the molecule is normally available in a physical or chemical manual. After these two are known, ρ can be calculated accordingly. After f is further measured experimentally, k can be determined using Eq. (6). This equation can
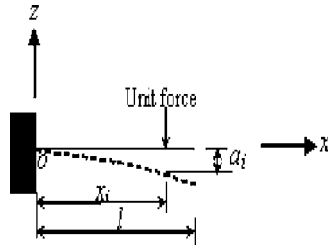
Fig. 1 Cross-sectional view of a cantilever and the geometric meaning of the influence coefficient a_i in the dynamic method

Fig. 2 Relationship between the generalized frequency change and the normalized x coordinate of a molecule

also be used to determine the amount of material uniformly adsorbed or coated on the whole top surface [1,2].

In the above two cases, when the position of a single particle is known or the top surface of the cantilever is fully occupied by molecules, the mass of a single particle or the number of molecule layers can be found using Eq. (1) or (4). On the other hand, in case ρ is known beforehand and the designed adhesion region on top surface of the cantilever is only partially occupied by a monolayer of the molecules, the number of molecules adsorbed cannot be accurately determined using the relationship (1) or (4). The determination of the number of molecules adsorbed in this case is equivalent to finding the surface area occupied by the molecules since the number of molecules adsorbed equals the division of the occupied surface area by the cross-sectional area of a single molecule. According to the relationship (4), this occupied surface area cannot be accurately found. Consequently, the amount of molecules adsorbed cannot be accurately determined. The inaccuracy in determination of the occupied surface area is because $1/(2\pi f)^2 - 1/(2\pi f_0)^2$ does not have a one-to-one correspondence with this occupied area since $x^3/3EI$ is involved in Eq. (4). For example, according to Eq. (4), eight molecules adsorbed at the positions, which have the same x coordinate x_1 ($x_1 \leqslant l/2$) but different y coordinates, generate the same frequency change as a single molecule adsorbed at the location whose x coordinate is $2x_1$. However, the area occupied by the eight molecules is eight times that covered by the single molecule. Another example of addressing this inaccuracy in detection is as follows. We claim that given a fully occupied region Ω, in which $0 < c \leqslant x \leqslant d < l$, $-(t/2) \leqslant y \leqslant t/2$, and $c < d$, we can always find a fully occupied region Ψ, in which $0 \leqslant e \leqslant x \leqslant h \leqslant l$ and $-(t/2) \leqslant y \leqslant t/2$, such that Ψ and Ω have different areas but are corresponding to the same frequency change. For this claim to hold, e and h need to meet the following criteria:

$$d^4 - c^4 = h^4 - e^4 \tag{7a}$$

$$d - c \neq h - e \tag{7b}$$

where use is made of Eq. (4). According to Eq. (7a), we have

$$h = \sqrt[4]{d^4 - c^4 + e^4} \tag{8}$$

For every chosen e, which satisfies $0 \leqslant e < c$, we obtain a unique h from Eq. (8). After we substitute $d - c = h - e$ into Eq. (7a), it can be readily shown that Eq. (7a) does not hold. Consequently, Eq. (7b) must hold. Therefore, it is concluded that our claim is true. Since e can vary in $[0, c)$, there actually exist numerous Ψ.

The above two examples indicate that due to the integrand $x^3/3EI$ in Eq. (4) different occupied areas may lead to the same frequency change and, subsequently, that the number of molecules adsorbed cannot be accurately found using Eq. (4). Our results are different from what was claimed in [3,4], which stated that the frequency change is proportional to the number of the molecules adsorbed.
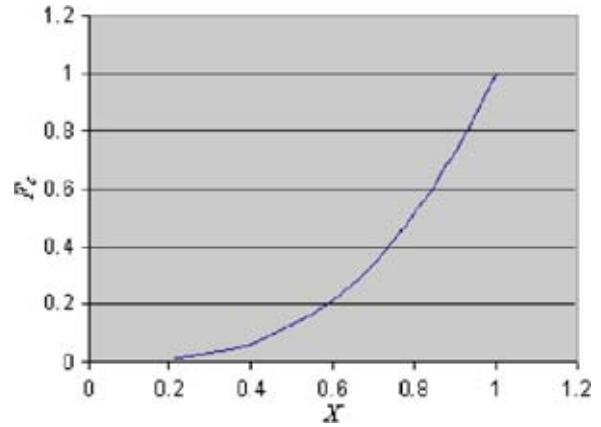
According to Eq. (3), $x^3/3EI$ represents how much a molecule affects the frequency change when the molecule is located at different positions on the cantilever. In order to see this effect clearly, let us rewrite Eq. (5) as

$$F_c = X^3 \tag{9}$$

where $F_c = [1/(2\pi f)^2 - 1/(2\pi f_0)^2]3EI/ml^3$ and $X = x/l$. F_c can be considered as another generalized frequency change. X is the normalized x coordinate of a molecule and ranges from 0 to 1. Figure 2 shows the relationship between F_c and X and illustrates how the molecule position affects the frequency change. It can be observed from this figure that the closer the molecule is located to the open end of the cantilever, the bigger the frequency change is. This implies that in order to have a single-molecule sensitivity in detection of the number of molecules adsorbed, the designed adhesion places should be away from the fixed end such that frequency change induced by the absorption of an additional molecule is high enough to be detected. The area under the curve in Fig. 2 represents this generalized frequency change due to the loading of a monolayer of molecules that occupy the whole top surface of the cantilever.

3 Static Method

In Sec. 2, we demonstrated that the dynamic method cannot be used to find the number of molecules adsorbed according to the mass-loading-induced frequency change. Naturally a question arises: can we detect the number of molecules adsorbed in view of the mass-loading-induced static bending? In this section, we discuss the possibility. The coordinate system of Sec. 2 is adopted in this section, and unless otherwise noted, the symbols used in both sections have the same meanings. Assume that the mass loading predominates the bending of the cantilever. Then, because of the absorption of a particle, we have [7]

$$w = \frac{mg\hat{x}_0^2}{6EI}(3l - \hat{x}_0) \tag{10}$$

where w denotes the vertical deflection of the open end of the cantilever and can be measured using an optical laser [11–14]. In Eq. (10), $\hat{x}_0$ is the x coordinate of the position occupied by the molecule and g represents the gravitational acceleration (Fig. 3). Sophisticated optical lasers can be used to measure the bending of a cantilever with high accuracy. For instance, the vertical cavity surface-emitting laser used in [14] was stated in the same paper to have an accuracy of 0.1 nm. The silicon cantilever used in [14] is 1 μm thick, 100 μm wide, and 500 μm long. If the same laser and silicon cantilever as those of [14] are adopted in measurement, Young's modulus of silicon is chosen to be 150 GPa [15], and the molecule is located at the open end of the cantilever, then
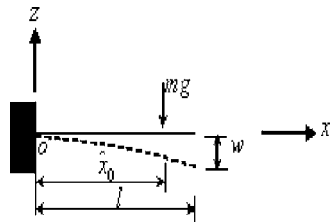
Fig. 3   Cross-sectional view of a cantilever and the bending configuration under the mass loading of a molecule in the static method

according to Eq. (10) the minimum detectable mass is 3.06×10^{-13} g. In order to measure the mass of a lighter molecule using the same laser, it can be observed from Eq. (10) that another cantilever with a smaller Young's modulus and/or smaller geometric dimensions (i.e., width and height) should be chosen for detection.

When a monolayer of molecules are adsorbed onto the cantilever, in view of Eq. (10) and the principle of superposition principle [16], we obtain

$$w = \int_A \frac{\rho g x^2 (3l - x)}{6EI} dx dy \qquad (11)$$

If w and $[g x^2 (3l - x)]/6EI$ in Eq. (11) are interpreted as $1/(2\pi f)^2 - 1/(2\pi f_0)^2$ and $x^3/3EI$, respectively, then Eq. (11) actually has the same form as Eq. (4). As a result, we expect that the static method can perform the same functions as the dynamic method, such as the detection of the mass of a single particle, while the same limitations faced by the dynamic methods in the detection of the number of molecules adsorbed also apply to the static method. Using Eq. (11) and following the same procedures in Sec. 2, which are based on Eq. (4), it can be readily shown that this expectation is true.

Finally, it should be noted that, in order to make the static method valid for detecting the presence of molecules based on the relationship between the static bending of the free end of a cantilever and the mass loading, the mass loading must predominate the bending. As a result, the Eq. (11)-based static method is not applicable to cases where the surface stress generated because of the absorption of molecules affects the bending more than the mass loading [10–13]. In those cases, the dynamic method should be chosen to detect the presence of molecules since this method is not affected by the surface stress, in principle.

4   Summary

In this work, we demonstrate that both dynamic and static methods use the same form of relationships for their sensing pur-

poses. We also show that when the designed adhesion region is only partially occupied by molecules of interest neither method can be applied to accurately determine the amount of molecules adsorbed on a cantilever, although it is possible to use both methods to find the mass of a single particle and the amount of uniformly coated materials.

References

[1] Thundat, T., Warmack, R. J., Chen, G. Y., and Allison, D. P., 1994, "Thermal and Ambient-Induced Deflections of Scanning Force Microscope Cantilevers," Appl. Phys. Lett., **64**, pp. 2894–2896.

[2] Thundat, T., Sharp, S. L., Fisher, W. G., Warmack, R. J., and Wachter, E. A., 1995, "Micromechanial Radiation Dosimeter," Appl. Phys. Lett., **66**, pp. 1563–1565.

[3] Ilic, B., Czaplewski, D., Craighead, H. G., Neuzil, P., Campagnolo, C., and Batt, C., 2000, "Mechanical Resonant Immunospecific Biological Detector," Appl. Phys. Lett., **77**, pp. 450–452.

[4] Ilic, B., Czaplewski, D., Craighead, H. G., Neuzil, P., Campagnolo, C., and Batt, C., 2001, "Single Cell Detection With Micromechanical Oscillators," J. Vac. Sci. Technol. B, **19**, pp. 2825–2828.

[5] Thomson, W. T., 1988, *Theory of Vibration with Applications*, Third Edition, Prentice-Hall, Englewood Cliffs, NJ, pp. 156–158 and 299–304.

[6] Weaver, W., Timoshenko, S. P., and Young, D. H., 1990, *Vibration Problems in Engineering*, Fifth Edition, Wiley, New York, p. 428.

[7] Timoshenko, S. P., and Gere, J. M., 1972, *Mechanics of Materials*, Van Nostrand Reinhold, New York.

[8] Davis, Z. J., Abadal, G., Kuhn, O., Hansen, O., Grey, F., and Boisen, A., 2000, "Fabriation and Characterization of Nanoresonating Devices for Mass Detection," J. Vac. Sci. Technol. B, **18**, pp. 612–616.

[9] Abadal, G., Davis, Z. J., Borrisé, X., Hansen, O., Boisen, A., Barniol, N., Pérez-Murano, F., and Serra, F., 2003, "Atomic Force Microscope Characterization of a Resonating Nanocantilever," Ultramicroscopy, **97**, pp. 127–133.

[10] *Manual of the LEO Transmission Electron Microscopy*, LEO Electron Microscopy Inc., Thornwood, NY.

[11] Berger, R., Delamarche, E., Lang, H. P., Gerber, C., Gimzewski, J. K., Meyer, E., and Güntherdot, H.-J., 1997, "Surface Stress in the Self-Assembly of Alkanethiols on Gold," Science, **276**, pp. 2021–2024.

[12] Lang, H. P., Berger, R., Battiston, F., Ramseyer, J.-P., Meyer, E., Andreoli, C., Brugger, J., Vettiger, P., Despont, M., Mezzacasa, T., Scandella, L., Güntherdot, H.-J., Gerber, Ch., and Gimzewski, J. K., 1998, "A Chemical Sensor Based on a Micromechanical Cantilever Array for the Identification of Gases and Vapors," Appl. Phys. A: Mater. Sci. Process., **66**, S61–S64.

[13] Moulin, A. M., O'Shea, S. J., Badley, R. A., Doyle, P., and Welland, M. E., 1999, "Measuring Surface-Induced Conformational Changes in Protein," Langmuir, **15**, pp. 8776–8779.

[14] Fritz, J., Baller, M. K., Lang, H. P., Rothuizen, H., Vettiger, P., Meyer, E., Güntherdot, H.-J., Gerber, Ch., and Gimzewski, J. K., 2000, "Translating Biomolecular Recognition Into Nanomechanics," Science, **288**, pp. 316–318.

[15] Spiering, V. L., Bouwstra, S., and Spiering, R., 1993, "On Chip Decoupling Zone for Package-Stress Reduction," Sens. Actuators, A, **39**, pp. 149–156.

[16] Hibbeler, R. C., 1994, *Mechanics of Materials*, Prentice-Hall, Englewood Cliffs, NJ, p. 134.

Buckling Mode Localization in Restrained Columns Under Cyclic Loading

Edmundo Corona

Department of Aerospace and Mechanical Engineering,
University of Notre Dame, 365 Fitzpatrick Hall
of Engineering, Notre Dame, IN 46556
e-mail: ecorona@nd.edu

The buckling mode localization of a multiply supported column subjected to fully reversed, displacement-controlled cyclic loading is studied numerically. The results show that localization can take place and lead to limit loads similar to those reported in other structural members in the literature. Analysis of the results reveals the mechanisms by which localization takes place.
[DOI: 10.1115/1.1938203]

Introduction

It is well known that metal structural members, such as open-section beams and tubes, can collapse when subjected to persistent cyclic loading into the plastic range of the material [1–6]. These references illustrate that structural integrity degrades with cycling due to the accumulation of deflections. For some relatively long structural members, deflections are consistent with buckling modes that display periodic variation along the length [4,5]. Under persistent loading, localization of such buckling modes can take place and induce collapse. Under load control the deflections are unconstrained, and localization can take place in a manner that is reminiscent of monotonic loading. The events that lead to localization under displacement-controlled cycles of constant amplitude are less obvious because the global prescribed deflections of the structure are constrained, and one may think that this would impair localization and result in shakedown. Yet, several examples in the literature indicate that localization can set in and result in collapse.

In this note an attempt is made to illustrate how localization can develop under displacement-controlled cyclic loading as in [4,5] by using a simple model of a multiply supported column similar to that considered by Goto et al. [7]. They used this model to demonstrate that localization can take place under displacement-controlled cyclic loading. Here, the model will be used to illustrate how localization develops.

Formulation

The problem consists of a pinned-pinned elastic-plastic column with several intermediate pin supports as shown in Fig. 1(*a*). The spacing between the supports is λ. All but the left-most support allow axial deflections. The total length of the column is L, and it has a rectangular cross section of width b and height h. The column carries an axial load P (positive in tension) at the right end. The intermediate supports of the column are modeled as elastic springs, as shown in Fig. 1(*b*). The stiffness of the springs k is set to be relatively high, as in the penalty approach. The coordinate system is also shown in the figure, with x and z being the axial and

transverse coordinates, respectively. The respective displacement components are u and w. The column is loaded by prescribing the axial deflection of the right end u^*.

Kinematics. The kinematics are based on the common assumption that plane sections, originally perpendicular to the axis of the column, remain plane and perpendicular to the deformed axis. An intermediate class of kinematics [8], which allows moderate rotations but requires small strains, is adopted. The axial strain is given by

$$\varepsilon = \varepsilon^o + \kappa z \tag{1}$$

where

$$\varepsilon^o = u' + \frac{1}{2}w'^2 \tag{2}$$

is the membrane strain and

$$\kappa = -w'' \tag{3}$$

is the curvature.

Constitutive Model. Since we expect progressive deformation with cyclic loading, the constitutive model adopted needs to model ratcheting in a reasonable manner. The two-surface, Dafalias-Popov cyclic plasticity model [9,10] has proven successful in this respect (see [11]), and is adopted here. This model is based on classical incremental plasticity with kinematic hardening.

Principle of Virtual Work. Equilibrium is satisfied using the Principle of Virtual Work (PVW). It can be written as

$$b \int_0^L \int_{-h/2}^{h/2} \sigma \delta\varepsilon \, dz dx - \sum_{i=1}^{S} kw_i \delta w_i - \ell \, \delta g = 0 \tag{4}$$

where w_i is the transverse deflection at the ith support, S is the number of intermediate supports and

$$g = [u(L) - u^*]\text{EA} = 0 \tag{5}$$

is a constraint used to prescribe the axial deflection u^*. EA is the elastic axial rigidity of the column. The Lagrange multiplier $\ell = P/\text{EA}$ yields the axial load in the column.

Solution. An approximate solution is attempted by discretizing the displacement components u and w using trigonometric series expansions that satisfy the boundary conditions

$$w = \sum_{n=1}^{N_s} a_n \sin\frac{n\pi x}{L}$$

$$\tag{6}$$

$$u = b_o x + \sum_{n=1}^{N_s} b_n \sin\frac{n\pi x}{L}$$

Substituting (6) into the kinematics (1)–(3) and into the PVW (4) yields a system of $2N_s+2$ nonlinear algebraic equations with unknowns $\{a_1,\ldots,a_{N_s},b_o,b_1,\ldots,b_{N_s},\ell\}$. The equations are solved iteratively for each prescribed increment in u^* using the Newton-Raphson method. The integrations in (4) are carried out using Gaussian quadrature.

Results

In order to illustrate localization, columns with several segments must be considered. A column with no internal supports ($\lambda=L$) does not display localization [12], and its response under cyclic loading is very similar to that predicted by a simple Shanley-type model [13]. Localization events become easier to see as the number of segments increases. A column with eight segments ($\lambda=L/8$) has been chosen to illustrate the development
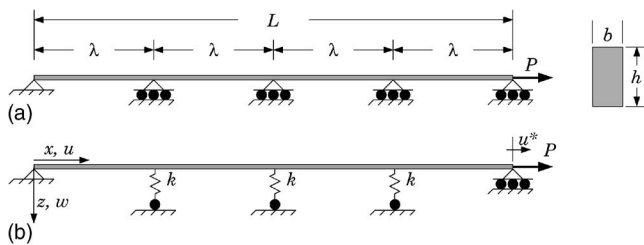
Fig. 1 (*a*) **Schematic of the parameters of the problem and (*b*) model of the problem**

of localization. The slenderness ratio of each segment is $\lambda/r=30$, where r is the radius of gyration of the cross section. The plastic buckling load P_{cr} and corresponding axial displacement at $x=L$, u_{cr}, of this column can be found using a simple plastic bifurcation analysis. In the example to be shown, the column starts with a periodic geometric imperfection in the transverse direction with wavelength λ and amplitude $w_o/\lambda=0.033\%$. In order to excite localization, the left-most segment has a slightly larger imperfection amplitude w_b+w_o, where $w_b/w_o=0.1$. The material parameters used were obtained from experimental data in [11] and are given in Table 1. The results shown below were generated with $N_s=24$. Five and sixty integration points were used through the height and length of the column, respectively.

The column is loaded by prescribing the displacement u^*. A symmetric cycle with amplitude $\tilde{u}/|u_{cr}|=0.975$ has been used. Loading has been carried out for 90 cycles. The load-deflection response of the column is shown in Fig. 2(*a*) with P plotted positive when in tension. P has been normalized by $|P_{cr}|$. It is clear that the peak compressive loads decrease rapidly over the first five cycles, as the amplitude of the transverse deflections grows, as shown in Fig. 2(*b*). In this figure, w_p is the lateral deflection of the column at the peak compressive load, plotted in increments of 10 cycles. The initial imperfection is also shown, but it is almost indistinguishable from the horizontal axis. It is clear that the transverse deflections of the column have increased significantly by the tenth cycle, but are uniform along the length. After that point, the transverse deflections continue to grow and the compressive load peaks continue to decrease, but at a much slower rate. The deflection amplitude in the biased segment grows at a slightly faster rate. This continues until approximately cycle 30 when the segments near the right end begin to experience a decrease in transverse amplitude. By cycle 75, the localization of transverse deflections becomes significant enough to cause the decrease in peak compressive loads to accelerate. In fact, limit loads appear within each cycle starting with the 78th cycle. By the 90th cycle, it is clear that the deformation is heavily localized at the left end of the column.

Now that the localization behavior of the column has been established, it is of interest to investigate how it takes place. It has already been shown that the amplitude of the transverse deflection in segments on the right part of the column decreases with localization and provides at least part of the length necessary to feed the localization process. This is illustrated in Fig. 3, where the axial deflections at the supports at the peak compressive displacement u_p are plotted versus the number of cycles N. The displace-
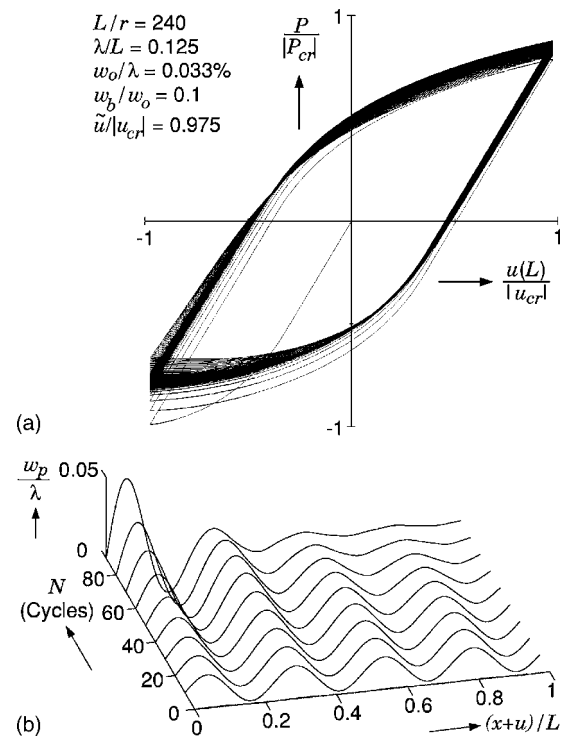
Fig. 2 (*a*) **Axial load-deflection response and (*b*) column deflection as function of number of cycles**

ment at point h is the prescribed quantity and is shown in dashed line. The increase in the deflection at point a is a direct result of the localization process and is expected, but note that the magnitude of the deflection of points b–g eventually becomes larger than that of point h. Considering point g, for example, the fact that it has an axial deflection larger than point h can be accounted for by an increase in length of the segment between these two points. The change in length of a segment between the nth and $(n+1)$th supports can be calculated from

$$\Delta\lambda = \int_{n\lambda}^{(n+1)\lambda}\left(\sqrt{(1+u')^2+w'^2}-1\right)dx \qquad (7)$$

Figure 4 shows the length change of several segments in solid lines and the change in length of the column in dashed line. It is clear that segments close to the right end of the column increase in length, whereas the segment where localization occurs decreases in length, and that the overall length of the column increases somewhat as well.

In summary, the localization process is a global event in the column, where the amplitude of the lateral deflections away from the localization zone decreases as the peak compressive load decreases and the length of the same segments increases. These events provide the necessary kinematic events to make localization possible. In this model localization occurs in a very smooth

Table 1 **Material properties (see [11] for parameter definitions)**

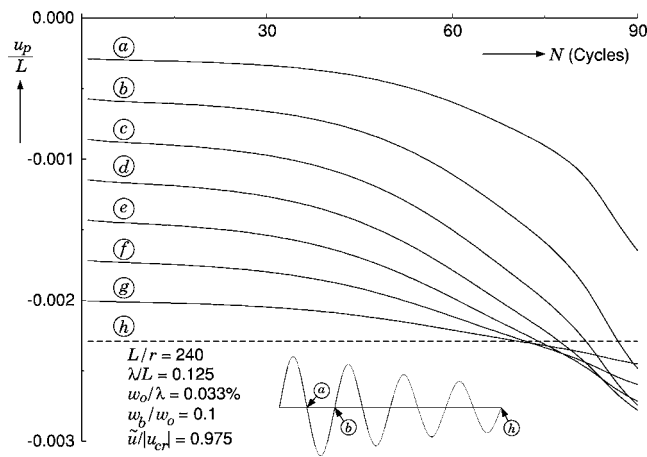| E, GPa (ksi)	E_o^P, GPa (ksi)	σ_y, MPa (ksi)	σ_b, MPa (ksi)	α, GPa (ksi)	β	m
181 (26,320)	1.38 (200)	131 (19)	269 (39)	490 (71,100)	27	2

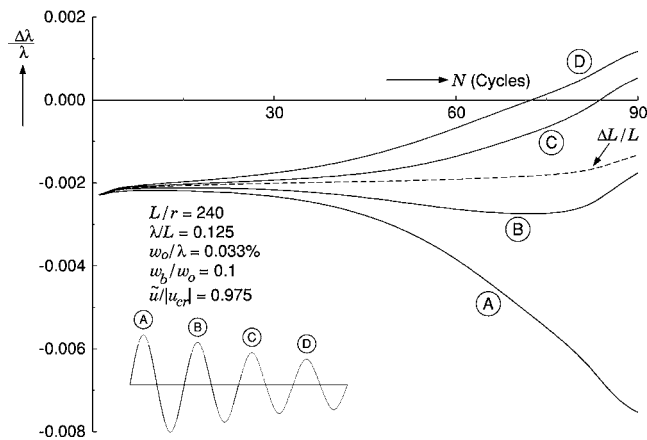**Fig. 3 Axial displacement of column supports**

Fig. 4 Length change of column segments

manner. Other column models demonstrate much more severe changes in the load-deflection response at localization. The model in [12] of a column on an elastic-plastic foundation is one such example.

Acknowledgment

References

[1] Bertero, V. V., and Popov, E. P., 1965, "Effect of Large Alternating Strains of Steel Beams," ASCE J. Struct. Div. **91**, pp. 1–12.

[2] Kyriakides, S., and Shaw, P. K., 1987, "Inelastic Buckling of Tubes Under Cyclic Bending," ASME J. Pressure Vessel Technol. **109**, pp. 169–178.

[3] Corona, E., and Kyriakides, S., 1991, "An Experimental Investigation of the Degradation and Buckling of Circular Tubes Under Cyclic Bending and External Pressure," Thin-Walled Struct. **12**, pp. 229–263.

[4] Vaze, S. P., and Corona, E., 1998, "Degradation and Collapse of Square Tubes Under Cyclic Bending," Thin-Walled Struct. **31**, pp. 325–341.

[5] Ellison, M. S., and Corona, E., 1998, "Buckling of T-Beams Under Cyclic Bending," Int. J. Mech. Sci. **40**(9), pp. 835–855.

[6] Yin, S., Corona, E., and Ellison, M. S., 2004, "Degradation and Buckling of I-Beams Under Cyclic Pure Bending," J. Eng. Mech. **130**, 809–817.

[7] Goto, Y., Toba, Y., and Matsuoka, H., 1995, "Localization of Plastic Buckling Patterns Under Cyclic Loading," J. Eng. Mech. **121**(4), pp. 493–501.

[8] Brush, D. O., and Almroth, B. O., 1975, *Buckling of Bars, Plates and Shells*, McGraw-Hill, New York, Chap. 1.

[9] Dafalias, Y. F., and Popov, E. P., 1975, "A Model of Nonlinearly Hardening Materials for Complex Loading," Acta Mech. **21**, pp. 173–192.

[10] Dafalias, Y. F., and Popov, E. P., 1976, "Plastic Internal Variable Formalism in Cyclic Plasticity," ASME J. Appl. Mech. **43**, pp. 645–651.

[11] Hassan, T., and Kyriakides, S., 1992, "Ratcheting in Cyclic Plasticity, Part I: Uniaxial Behavior," Int. J. Plast. **8**, pp. 91–116.

[12] Yin, S., 2002, "Buckling of Structural Members Under Cyclic Loading," PhD dissertation, University of Notre Dame.

[13] Corona, E., 2001, "Stability of the Shanley Column Under Cyclic Loading," ASME J. Appl. Mech. **68**, pp. 324–331.

Synthesis and Experimental Validation of a Delayed Reference Controller for Active Vibration Suppression in Mechanical Systems

P. Gallina

Dip. Energetica, Universita di Trieste, Via A. Valerio 10, 34127 Trieste, Italy
e-mail: pgallina@units.it

A. Trevisani[1]

DIMEG, Universita di Padova, Via Venezia 1, 35131 Padova, Italy
e-mail: alberto.trevisani@unipd.it

This paper introduces a non-time-based control scheme for active position and vibration control of two-degree-of-freedom systems by applying it to the path-tracking and swing control of a system composed of a trolley and a simple pendulum. The basic idea behind such a scheme is to make the path reference of the trolley a function of the time and of a time delay. This latter, which is affected by the measured oscillation, is calculated on-the-fly in order to reduce the swing phenomenon. The effectiveness of the proposed control scheme, which may have application to the control of overhead cranes, is proved experimentally.
[DOI: 10.1115/1.1940663]

1 Introduction

There exists a considerable number of industrial applications demanding for innovative controllers guaranteeing, at the same time, accurate tracking of planned positions and effective damping of undesired vibrations. This paper proposes an innovative "non-time-based" control strategy fitting this new and challenging task, and which is suitable for path tracking and active vibration suppression on the mechanical systems that can be modeled as two-mass systems. In order to validate experimentally the proposed strategy, it is applied to the control of a "crane-like" system with a single resonant mode, modeled through a concentrated mass and a simple pendulum.

The simultaneous tracking control and swing suppression in overhead cranes is an interesting control task since point-to-point operations lead to significant control problems due to their pendulum-like dynamic behavior. In the last decades many solutions have been proposed for the control of crane-like systems (the contributions from [1–5], and the references therein, are representative of the research conducted to date). Some of them have led to industrial applications and commercial diffusion (e.g., [3]). Nevertheless, researchers keep working in this field so as to reduce maneuver time and residual vibrations in rest-to-rest operations (e.g., [4,5]).

This paper proposes a novel non-time-based strategy: the reference path is made a function of a parameter depending on the time and on another variable which is related to the measured vibration (i.e., the swing angle), and can be thought of as a time delay. For this reason the proposed control is named "delayed reference control" (DRC).

Non-time-based control strategies are being employed more and more widely in a variety of fields, in particular, when following a desired path defined only through space ("path tracking") becomes the primary goal. In particular, an interesting "event-based" (non-time-based) control strategy has been recently introduced and applied to several control problems in the robotic field, such as robot motion control [6], robotic teleoperation [7], force and impact control [8], and manufacturing automation [9]. According to the event-based method, the reference input is a function of an action reference parameter instead of a function of time. The reference input is modified on-the-fly since the action reference parameter is sensitive to the sensory measurement and the task. Clearly, the role of the planner in the event-based approach changes with respect to the classical control theory.

The DRC scheme proposed in this paper openly belongs to the category of these new non-time-based control approaches. In a previous work [10] the effectiveness of a DRC scheme was proven by means of simulation results in the problem of path-tracking and vibration suppression in a generic two-mass system. With this work, the DRC scheme is extended to a real application, and a first experimental validation of the scheme is provided.

2 The System Model

The free-body diagrams of the studied system is shown in Fig. 1. The trolley can move on a horizontal plane driven by the external force F and is connected to a simple pendulum oscillating on a vertical plane. The nonlinear equations of motion of the trolley and of the pendulum can be easily inferred by applying Newton's law. If friction is neglected, the following equations hold:

$$\text{Trolley: } (M+m)\ddot{x}(t) + mL\ddot{\theta}(t)\cos\theta(t) - mL\dot{\theta}(t)^2\sin\theta(t) = F(t) \tag{1}$$

$$\text{Pendulum: } (J+mL^2)\ddot{\theta}(t) + mgL\sin\theta(t) = -mL\ddot{x}(t)\cos\theta(t) \tag{2}$$

In order to employ the generic relations proposed in [10] for the synthesis of the DRC, the differential equations above are linearized, under the hypothesis that the pendulum has small motion about the vertical axis. The following approximations are therefore introduced: $\sin\theta \cong \theta, \cos\theta \cong 1, \dot{\theta}^2\theta \cong 0$. The linearized equations take the form

$$\begin{bmatrix} M+m & mL \\ mL & J+mL^2 \end{bmatrix}\begin{Bmatrix} \ddot{x}(t) \\ \ddot{\theta}(t) \end{Bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & mgL \end{bmatrix}\begin{Bmatrix} x(t) \\ \theta(t) \end{Bmatrix} = \begin{Bmatrix} F(t) \\ 0 \end{Bmatrix} \tag{3}$$

3 Synthesis of the Delayed Reference Controller

Let us rewrite the second row of Eq. (3) in the form

$$(J+mL^2)\ddot{\theta}(t) + mgL\theta(t) = -mL\ddot{x}(t). \tag{4}$$

The term on the right-hand side may be thought of as an external force $q(t)$ acting on the pendulum, and whose value can be appropriately set by regulating the trolley acceleration. The values taken by $q(t)$ depend on the control problem considered. A very basic control problem which fits the studied system is keeping the pendulum swing to a minimum while the trolley moves. This is a regulation problem, in which the reference is set equal to zero. The following corrective action can therefore be established:

$$q(t) = -mL\ddot{x}(t) = k_{cD}[\dot{\theta}_r(t) - \dot{\theta}(t)] = -k_{cD}\dot{\theta}(t) \tag{5}$$
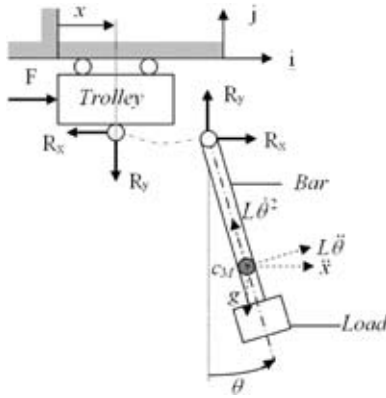
Fig. 1 Free-body diagrams of the trolley and the pendulum

The corrective action can be straightforwardly related to the acceleration of the trolley position reference ($\ddot{x}_r$) if the servo system ensuring the motion of the trolley has a very low time constant with respect to the swing period. In other words, the trolley position controller is required to keep the position error $x_r - x$ small, so that $x \, x_r$ can be assumed. Under such a hypothesis Eq. (5) can be rewritten as follows:

$$k_{cD}\dot{\theta}(t) = mL\ddot{x}(t) \approx mL\ddot{x}_r(t) \qquad (6)$$

Now, assume that x_r is not only a function of time, but also of an additional parameter acting as a time delay, and which can be modified continuously. It is basically being assumed that the control problem consists in ensuring that the trolley moves along a desired path through space (no matter how long it takes), rather than in accordance with a trajectory in time (as in classical control theory). Without loss of generality, the path through space the

trolley is desired to follow may be defined by the parametric equation $x_r = x_r(l)$, where $l \in \nabla$ is called the action reference parameter. In the DRC scheme l is of the form $l = t - \tau$, with $\tau \in \nabla$. The time varying quantity $\tau = \tau(t)$ will henceforth be called a time delay, although it does not necessarily take positive values. τ has been introduced to allow modifying the action reference parameter, and, consequently, the instantaneous position reference x_r and its time derivatives. This ensures the possibility of generating the control action specified by Eq. (5) without drifting away from the desired path.

By computing the second derivative of $x_r(l)$ with respect to time, and by substituting it for $\ddot{x}_r(t)$ into Eq. (6), one gets the following equation:

$$k_{cD}\dot{\theta}(t) = mL\left[\frac{d^2 x_r}{dl^2}(1-\dot{\tau}(t))^2 - \frac{dx_r}{dl}\ddot{\tau}(t)\right] \qquad (7)$$

which can be integrated over time twice yielding:

$$k_{cD}\int_0^t \theta(\zeta)d\zeta = mL\int\int_0^t \frac{d^2 x_r}{d(\zeta - \tau(\zeta))^2}(1-\dot{\tau}(\zeta))^2 d\zeta$$
$$- mL\int\int_0^t \frac{dx_r}{d(\zeta - \tau(\zeta))}\ddot{\tau}(\zeta)d\zeta \qquad (8)$$

If we write $\beta = \beta(l) = (dx_r/dl)(l)$, $\beta' = \beta'(l) = (d^2 x_r/dl^2)(l)$, Eq. (8) clearly becomes

$$k_{cD}\int_0^t \theta(\zeta)d\zeta = mL\int\int_0^t \beta'(1-\dot{\tau}(\zeta))^2 d\zeta - mL\int\int_0^t \beta\ddot{\tau}(\zeta)d\zeta \qquad (9)$$

By integrating by parts the last term of the right-hand side, there follows

$$k_{cD}\int_0^t \theta(\zeta)d\zeta = mL\int\int_0^t \beta'(1-\dot{\tau}(\zeta))^2 d\zeta - mL\int_0^t\left(\beta\dot{\tau}(\zeta) - \int_0^t \beta'(1-\dot{\tau}(\zeta))\dot{\tau}(\zeta)d\zeta\right)d\zeta = mL\int\int_0^t \beta'(1-\dot{\tau}(\zeta))^2 d\zeta - mL\left(\beta\tau(t)\right.$$

$$-\int_0^t \beta'(1-\dot{\tau}(\zeta))\dot{\tau}(\zeta)d\zeta - \int\int_0^t \beta'(1-\dot{\tau}(\zeta))\dot{\tau}(\zeta)d\zeta\right) = mL\left(\int\int_0^t \beta'(1-\dot{\tau}(\zeta))d\zeta - \int_0^t \beta'(1-\dot{\tau}(\zeta))\dot{\tau}(\zeta)d\zeta\right)$$

$$- mL\beta\tau(t) \qquad (10)$$

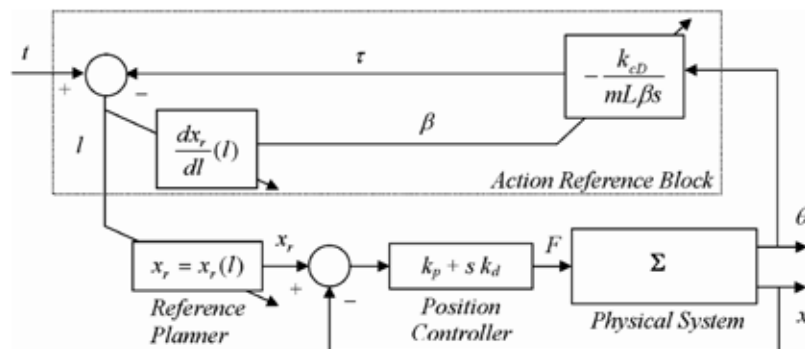**Fig. 2 Block scheme of the DRC**

Fig. 3 The experimental apparatus

Table 1 Geometric and inertial characteristics of the prototype

M (kg)	32.318
m (kg)	0.7589
J (kg m^2)	584×10^4
L (m)	0.459

If the parametric equation $x_r = x_r(l)$ is smooth enough to cause all the integrals containing its second derivative β' to be negligible, Eq. (10) becomes

$$k_{cD} \int_0^t \theta(\zeta)d\zeta = -mL\beta\tau(t) \qquad (11)$$

Although the assumption under which Eq. (11) holds might appear too restrictive, in many practical applications (e.g., material handling by cranes or robots) the second derivative of the path $x_r(l)$ is small. In these applications neglecting the terms containing β' does not affect the expected dynamic behavior of the controlled system, as it is shown by the experimental results in Sec. 4.
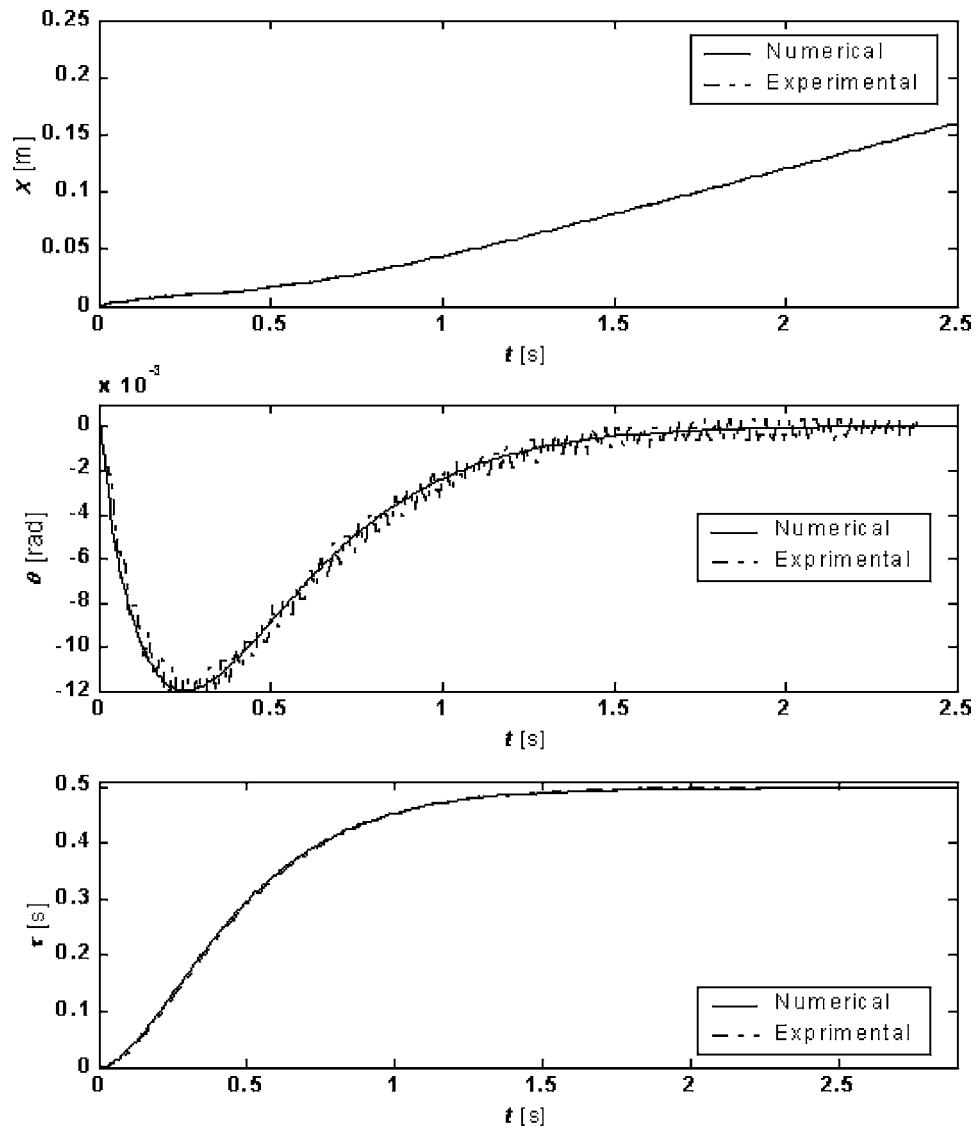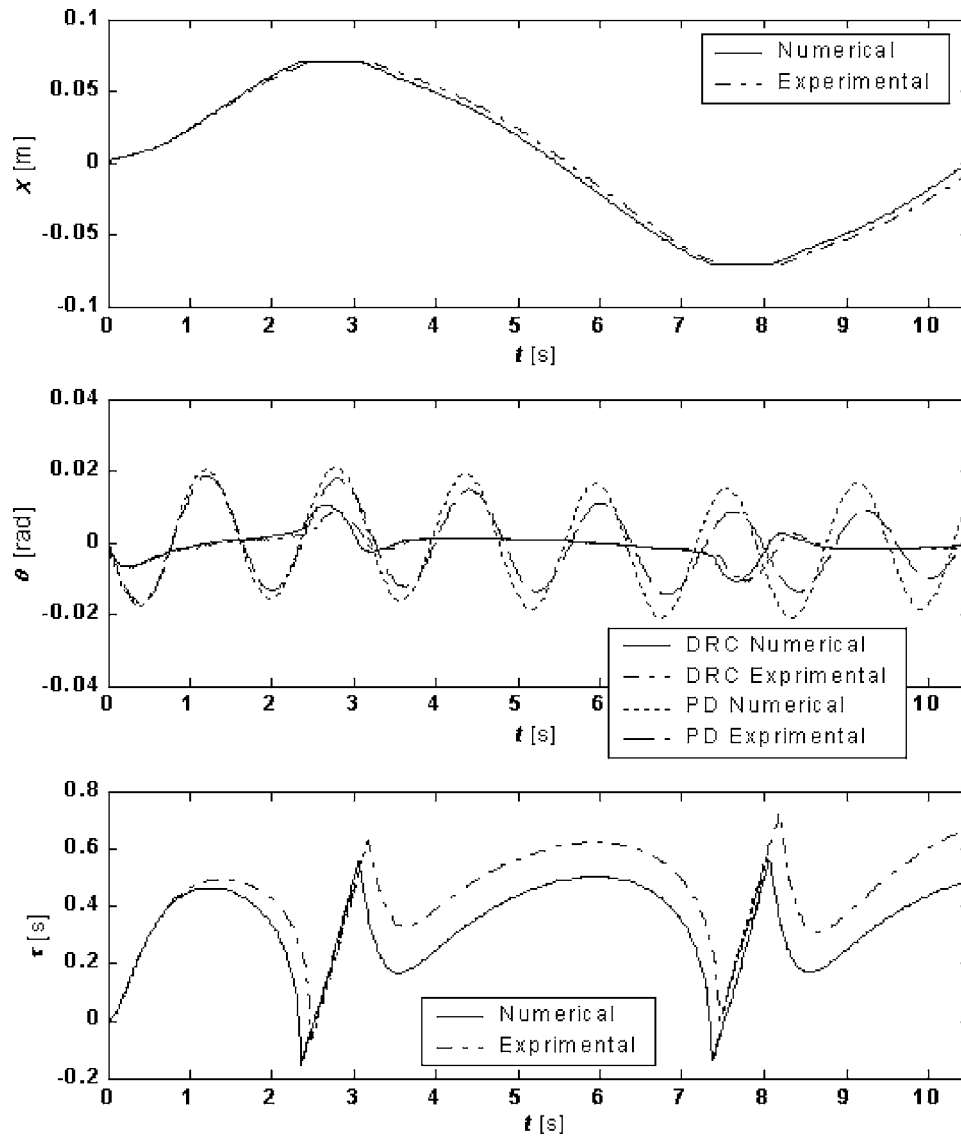

Fig. 4 Comparison of the experimental and numerical results. Linear reference path. x, θ, and τ vs. t.

Fig. 5 Comparison of the experimental and numerical results. Nonlinear reference path. x, θ (DRC and PD), and τ vs. t.

As a result, the following relation between the delay τ and the pendulum swing is established:

$$\tau(t) = -\frac{k_{cD}}{mL\beta}\int_0^t \theta(\zeta) \cdot d\zeta. \tag{12}$$

Hence, as long as the actual position of the trolley (x) can be approximated by the path reference (x_r), Eq. (12), or its Laplace transform,

$$T(s) = -\frac{k_{cD}}{mL\beta s}\Theta(s), \tag{13}$$

allows computing the time τ the reference has to be delayed to perform the desired control action given by Eq. (5).

As far as the parameter k_{cD} is concerned, its value should be set so that the pendulum exhibits a critically damped response. The combination of Eqs. (4) and (5) immediately yields

$$k_{cD} = 2\sqrt{mgL(J + mL^2)}. \tag{14}$$

Figure 2 shows the block scheme of the DRC designed for the studied system. It comprises the following blocks:

- *Reference planner*: it computes the instantaneous reference x_r on the basis of the action reference parameter l;
- *Action reference block*: it computes l on the basis of t, the instantaneous values of $\beta(l)$ and θ;
- *Position controller*: a standard proportional-plus-derivative (PD) regulator is employed to compute the control action $F(l)$.

The DRC scheme therefore has a cascade structure, which consists of an inner position (x) loop and an outer loop closed to suppress the pendulum oscillation (θ) actively. A major advantage of the DRC scheme is the possibility of employing any kind of regulator in the inner feedback loop, which provides a great flexibility in the design of DRC-based regulators. The inner loop regulator, in fact, does not play any role in the DRC synthesis: its sole role is to ensure an accurate tracking of x_r.

4 DRC Experimental Validation

The theory developed in the foregoing sections has been applied to the prototype shown in Fig. 3. A pendulum, composed of a steel load at the end of a slender steel bar with square cross

section, is connected to the inner block of a ballscrew linear actuator THK LM KR3306A by means of a horizontal shaft supported by two ball bearings. The linear actuator is driven by the brush type dc motor MAE M540-0741. The angular position of the shaft supporting the pendulum, coinciding with θ, is measured by the incremental encoder Eltra EH 63G. An identical encoder is employed to measure the angular position of the motor shaft, and, consequently, to compute x. Table 1 reports the basic geometric and inertial characteristics of the physical system.

As far as the tuning of the trolley position controller parameters is concerned, the following values have been employed: $k_p = 2 \times 10^5$, $k_d = 3 \times 10^3$. The value employed for the DRC gain is $k_{cD} = 1.7$, which has been computed through Eq. (14). The DRC scheme has been implemented on a PC with a real-time operating system, running at a sample rate of 1 kHz.

The experimental analysis has been carried out performing two significant tests aimed at assessing the DRC performances in tracking two different paths $x_r(l)$, characterized by either a constant or a continuously varying β. In these tests the pendulum swing is caused by the trolley motion, and in particular by changes in the direction of motion, while no external disturbances are added.

Figure 4 shows the results obtained in the first test, in which β is equal to 0.08 (m s^{-1}). In this case, the desired position x_r against the action reference parameter l is a straight line, and Eq. (13) holds true without any approximation. The time histories of the controlled variables x and θ recorded experimentally and computed in simulation prove the effectiveness of the scheme and show that the real system performs as predicted by the numerical model: only negligible differences can be discriminated, in particular, the residual high-frequency oscillation which can be noticed in the θ versus t subplot is the first flexural mode of the pendulum, which is not accounted for in the model. Further evidences of the effectiveness of the control scheme have been gathered carrying out a second test that allows assessing the performance of the DRC when a nonlinear path is to be followed. The reference path $x_r(l)$ employed in the test is $x_r(l) = x_0 \sin(\omega l)$, where $x_0 = 0.07$ (m) and $\omega = 0.2\pi$ (rad s^{-1}). Hence, the explicit expression for $\beta(l)$ is $\beta(l) = x_0\omega \cos(\omega l)$, and contrary to the previous test, β changes with l and also becomes equal to zero when $l = \lambda\pi/2\omega$, for any $\lambda \in \wedge$. However, if β tends to zero, Eq. (12) goes to infinity and no delay τ can be computed. This drawback of the scheme has been overcome by introducing a discontinuity offset at zero in β: a coefficient $\delta \in \nabla^+$ has been defined, so that, if $|\beta(\bar{l})| < \delta \Rightarrow \beta(\bar{l}) := \delta \cdot \{\beta(\bar{l})/|\beta(\bar{l})|\}$. The value of the coefficient δ employed in this test is $\delta = 0.001$ (m s^{-1}).

Figure 5 compares the experimental and numerical time histories of x, θ, and τ. In the subplot showing θ vs. t, the results obtained without controlling the pendulum swing (i.e., just employing the PD regulator of the inner loop) have also been plotted for comparison. Figure 5 proves that also in this case the DRC ensures an accurate tracking of the delayed reference x_r, even when β approaches and crosses zero, i.e., when the reversal of the direction of motion takes place. On the other hand, swing control is effective along all the path except when β approaches zero. Such an expected "path-dependent" decline in the DRC performances is caused by the fact that when β approaches zero no compensation of the oscillatory motion can be realized without theoretically drifting away from the desired path, and the value of β employed in the computation of τ is either kept constant or changed discontinuously to avoid the aforementioned singularity. The limited discrepancies between the experimental recordings and the expected behavior are mostly due to friction in the ball

bearings supporting the shaft about which the pendulum oscillates. The fact that when the sole PD is employed, the swing of the real system shows a faster decrease in amplitude supports this evidence.

5 Conclusions

In this paper a new non-time-based control strategy has been introduced which is suitable for path tracking and vibration suppression on the mechanical systems that can be modeled as two-mass systems.

The main novelty of the DRC is the introduction of an action reference parameter, which is a function of a time delay. A chief advantage of the scheme is ensuring a correct path tracking overcoming detrimental time constraints. A further advantage is that it can be obtained by simply adding an outer loop to a standard tracking controller, which enlarges the range of applications to which it can be applied.

The experimental results from the test case of a pendulum-like mechanism have proved that the DRC performs effectively, both in terms of vibration suppression and path following accuracy.

The extension of the method to fit systems with more than two degrees of freedom requires further theoretical insight and will be the subject of future numerical and experimental developments.

Nomenclature

c_M = mass center of the pendulum
m = overall mass of the pendulum
x = trolley absolute displacement
J = moment of inertia of the pendulum about the mass center
L = distance of the mass center from the pendulum axis of rotation
R_x, R_y = horizontal and vertical component of the reaction force between the trolley and the pendulum
F = trolley driving force
M = overall mass of the trolley
θ = angle of the pendulum from the vertical
g = gravity acceleration.

References

[1] Suzuki, H., and Terashima, E., 2000, "Semi-Automatic Control of Overhead Crane Using Potential Method," Proc. 2000 IEEE Int. Conf. on Systems, Man, and Cybernetics, Nashville, Vol. 5, pp. 3224–3229.
[2] Yoshida, K., 1997, "Nonlinear Control for a Crane System Having Constrained Trolley Stroke and Pendulum," Proc. 36th IEEE Conf. on Length Decision and Control, San Diego, Vol. 3, pp. 2895–2896.
[3] Parker, G. G., Petterson, B., Dohrmann, C., and Robinett, R. D., 1995, "Command Shaping for Residual Vibration Free Crane Maneuvers," Proc. American Control Conference, Seattle, Vol. 1, pp. 934–938.
[4] Bartolini, G., Pisano, A., and Usai, E., 2002, "Second-Order Sliding-Mode Control of Container Cranes," Automatica, 38, pp. 1783–1790.
[5] Al-Garni, A. Z., Moustafa, K. A. F., and Nizami, S. S. A. K., 1995, "Optimal Control of Overhead Cranes," Control Eng. Pract., 3, pp. 1277–1284.
[6] Kang, W., Xi, N., and Tan, J., 1999, "Analysis and Design of Non-Time Based Motion Controller for Mobile Robots," Proc. 1999 IEEE Int. Conf. on Robotics and Automation, Detroit, Vol. 2, pp. 919–924.
[7] Tarn, T. J., Xi, N., Guo, C., and Bejczy, A. K., 1995, "Fusion of Human and Machine Intelligence for Telerobotic Systems," Proc. 1995 IEEE/ICRA, Nagoya, Vol. 3, pp. 3110–3115.
[8] Wu, Y., Tarn, T. J., and Xi, N., 1995, "Force and Transition Control with Environmental Uncertainties," Proc. 1995 IEEE/ICRA, Nagoya, Vol. 1, pp. 899–904.
[9] Xi, N., and Tarn, T. J., 1997, "Integrated Task Scheduling and Action Planning/Control for Robotic Systems Based on a Max-Plus Algebra Model," Proc. 1997 IEEE/RSJ Int. Conf. on Intelligent Robotics and Systems, Grenoble, Vol. 2, pp. 926–931.
[10] Gallina, P., and Trevisani, A., 2004, "Delayed Reference Control of a Two Mass Elastic System," J. Vib. Control, 10, pp. 135–159.

Errata

Erratum: "Size-Dependent Eshelby's Tensor for Embedded Nano-Inclusions Incorporating Surface/Interface Energies" [Journal of Applied Mechanics, 2004, 71(5), pp. 663–671]

P. Sharma and S. Ganti

Equations (21*b*) and (21*c*) and (22*b*) and (22*c*) in the published paper contain typographical errors. The error in Eqs. (21*b*) and (21*c*) involves a switching of the sign and a factor of 2 while the error in Eqs. (22*b*) and (22*c*) involves a switching of the signs. The corrected equations are written below.

Spherical Inclusion:

$$\varepsilon_{rr}(r) = -\left[\frac{3K^M \varepsilon^* - 2\tau_o/R_o}{4\mu^M + 3K^M + 2K^s/R_o}\right] \frac{2R_o^3}{r^3}\bigg| r > R_o \tag{21b}$$

$$\varepsilon_{\theta\theta}(r) = \varepsilon_{\phi\phi}(r) = \left[\frac{3K^M \varepsilon^* - 2\tau_o/R_o}{4\mu^M + 3K^M + 2K^s/R_o}\right] \frac{R_o^3}{r^3}\bigg| r > R_o. \tag{21c}$$

Cylindrical Inclusion:

$$\varepsilon_{rr}(r) = -\left[\frac{3K'^M \varepsilon^* - \tau_o/R_o}{2\mu^M + 3K'^M + K'^s/R_o}\right] \frac{R_o^2}{r^2}\bigg| r > R_o \tag{22b}$$

$$\varepsilon_{\theta\theta}(r) = \left[\frac{3K'^M \varepsilon^* - \tau_o/R_o}{2\mu^M + 3K'^M + K'^s/R_o}\right] \frac{R_o^2}{r^2}\bigg| r > R_o. \tag{22c}$$

Further, in the case of the cylindrical problem, an eigenstrain of $\varepsilon_{rr}^* = \varepsilon_{\theta\theta}^* = \varepsilon^*$ and $\varepsilon_{zz}^* = 0$ was assumed.

We are grateful to Changwen Mi and Professor Demitris Kouris regarding discussion on the aforementioned issues.

Erratum: "Analytical Solution of a Dynamic System Containing Fractional Derivative of Order One-Half by Adomian Decomposition Method" [Journal of Applied Mechanics, 2005, 72(2), pp. 290–295]

S. Saha Ray and R. K. Bera

The correct authors and their affiliations should have been listed as follows:

S. Saha Ray
Lecturer,
Department of Mathematics,
B. P. Poddar Institute of Management and Technology,
137, V.I.P. Road,
Poddar Vihar,
Kolkata 700052, West Bengal, India

R. K. Bera
Heritage Institute of Technology,
Chowbaga Road,
Anandapur, Kolkata 700107, India