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

Journal of Applied Mechanics

Published Bimonthly by ASME

VOLUME 71 • NUMBER 4 • JULY 2004

TECHNICAL PAPERS

(**Contents continued on inside back cover**)

This journal is printed on acid-free paper, which exceeds the ANSI Z39.48-1992 specification for permanence of paper and library materials. ♾™
♻ 85% recycled content, including 10% post-consumer fibers.

Journal of Applied Mechanics Volume 71, Number 4 JULY 2004

BRIEF NOTES

DISCUSSION

ANNOUNCEMENTS AND SPECIAL NOTICES

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

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

On the Use of a Kolsky Torsion Bar to Study the Transient Large-Strain Response of Polymer Melts at High Shear Rates

Y. Hu
Graduate Research Assistant

R. Feng[1]
Assistant Professor
Mem. ASME
e-mail: rfeng1@unl.edu

Department of Engineering Mechanics,
University of Nebraska-Lincoln,
Lincoln, NE 68588-0526

A Kolsky torsion bar is utilized successfully in a novel rheometric experiment for measuring the transient large-strain response of polymer melts under high shear-rate loading. A molten low-density polyethylene is studied with the new technique. The results show that the high-rate shear response of the material has an instantaneous rate dependence that may not be discernible at low rates and a strain-dependent hardening that saturates at large strains instead of fading. The usefulness of the technique and the significance of the findings are discussed in comparison with a modified rubberlike liquid theory and high-rate capillary measurements for low-density polyethylene melts.
[DOI: 10.1115/1.1756142]

1 Introduction

During a rapid extrusion of molten polymer, a common process in injection molding, wire extrusion, and film blowing of plastics, the polymer melt is subjected to high strain-rate deformation as well as varying temperature and pressure. Local shear rate may reach $10^3\,s^{-1}$ near the extruder barrel surface and $10^5\,s^{-1}$ or higher in the die lips, [1]. The cumulative strain in the material may be far beyond 100%. How to characterize accurately the response of viscoelastic polymer melts under these dynamic loading conditions remains a scientific challenge even though the rheology of polymer melts has been studied extensively, [2–4].

The capillary and oscillatory rheometrics, [5], are the two most commonly used methods for measuring the rheological response of polymer melts under high shear-rate loading, [6–9]. A capillary rheometer is essentially a small-size plunge extruder. The overall response of a *steady* but generally *nonuniform* melt flow through the entire extruder is measured. With an a priori assumed flow field, an ''apparent shear viscosity,'' the ratio of shear stress to shear rate is determined from the measurement, [5]. Such a result is meaningful if the material behaves as a viscous fluid (rate-dependent only) under given temperature and pressure. If the material response is of a viscoelastic fluid, [5], however, the steady-state apparent shear viscosity determined by the capillary rheometry is, in general, an ambiguous measurement since it may be significantly affected by the cumulative strain, which is an unknown in the experiment. Though useful as a reference for regulating the processing conditions in manufacturing, [8], such data are not sufficient for developing an accurate material model for a viscoelastic polymer melt.

The oscillatory rheometrics typically involves the use of an oscillatory mechanical spectrometer to subject a molten polymer specimen to sinusoidal loadings of small strains. A uniform flow field may be achieved in the specimen by using a concentric Couette cell or a cone-and-plate rheometric configuration, [5]. The dynamic viscosity and storage modulus of the sample material can be determined for these conditions. The frequency dependence of the dynamic viscosity is often related to the rate dependence of the apparent shear viscosity of the material in steady flow, [6,9,10], though the two may bear entirely different origins as pointed out by Khanna [11]. The polymer melt elasticity may be inferred from the storage modulus measurements. However, significant discrepancy between the two has been observed for some polymer melts, [12]. Transient measurements are possible with an oscillatory rheometer but only for shear rates and strains significantly smaller than those typical in a rapid extrusion process. Large-amplitude oscillatory rheometer has been reported by Giacomin et al. [13]. However, the long rise time of the device (~200 ms) makes it only suitable for steady-state measurement. An experimental technique capable of measuring the transient large-strain response of polymer melts under high strain-rate loading is currently lacking.

The limitation of the available rheological data on polymer melts also affects the theoretical modeling of these materials. The rheological descriptions used in polymer melt extrusion analysis and simulations are mostly empirical viscous models (see, e.g., [14]). This is, to a large extent, because the capillary measurements of the apparent viscosity have been the only experimental data available for the high strain rates of interest. Unfortunately, as discussed earlier, such measurements are ambiguous and insufficient for modeling viscoelastic polymer melts. The more rigorous polymer melt models based on continuum mechanics principle (e.g., models based on the rubberlike liquid theory and its modification, [2,3,9]), or macromolecule theory (e.g., various reptation models of entangled polymers, [4,15,16]), or a combination of the two, [17], on the other hand, have been developed and calibrated mainly for low-rate loading, [9,16]. Transient large-strain measurements at high strain rates are necessary to verify the usefulness of such a model for predicting the high-rate response of viscoelastic polymer melts.

In order to measure accurately the transient large-strain response of polymer melts under high shear-rate loading, we need an experimental technique that accelerates a melt flow very rapidly to minimize viscous heating and relaxation during the spin-up process, that retains a uniform steady flow field in the sample for

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, September 6, 2001; final revision, November 20, 2003. Associate Editor: K. R. Rajagopal. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

a sufficiently long time to render a homogeneous large-strain deformation, and that tracks accurately the time histories of the loading and the sample response to enable a time-resolved measurement. In principle, the Kolsky torsion bar technique can be a good solution for these issues. The technique is based on an experimental configuration introduced by Kolsky [18] in 1949, i.e., to sandwich a specimen between two long cylindrical bars functioning as two elastic waveguides. With the bars, dynamic loading can be applied to the specimen as an incident stress wave and the specimen response, which governs the stress waves leaving the specimen, can be detected far away from the specimen. The method has also been referred to as the split-Hopkinson bar technique. The method of generating dynamic *torsional* loading in the Kolsky bar system by rapid release of a pre-stored torque was introduced by Baker and Yew [19] and further developed by a number of researchers, [20,21]. An overview of the state of the art of the Kolsky torsion bar technique can be found in [21].

However, the Kolsky torsion bar (KTB) technique has mostly been used to study the dynamic plasticity of metals, [18–25]. In extending the technique for polymer melt rheometrics, there are three technical issues: (1) integrating a proper rheometric cell with the system, (2) maintaining the specimen temperature, and (3) measuring very low stress sustained in a polymer melt. The first KTB rheometer was developed by Clyens et al. [26] for studying super-cooled lubricants. The bar ends sandwiching the specimen were machined to form a parallel-annular-plates rheometer as well as a cone-and-plate rheometer. No stress signal enhancement was used since the stress levels in their super-cooled lubricants were sufficiently high. To investigate the low-stress response of lubricants near the room temperature, Feng and Ramesh [27] designed a KTB rheometer with a tubular input bar and a thinner solid output bar. The rheometric cell was a concentric Couette cell with a cone-and-plate cup end so that the area of specimen-bar interface was significantly greater than that of the cross section of the output bar, resulting in a signal enhancement that was sufficient for measuring a sample stress as low as 0.015 MPa accurately.

In this paper, we report a KTB rheometer developed specifically for polymer melts. This new rheometer provides a novel technique for measuring the transient response of polymer melts under shear rate rates of $10^2 \sim 10^4\ s^{-1}$ and initial temperatures up to 300°C, and for shear strains up to 1000%. To demonstrate the usefulness of the technique, the results of a series of KTB rheometric experiments on a molten low-density polyethylene (LDPE) are also presented. These transient high-rate large-strain measurements are, to our knowledge, the first of this kind and reveal important new information on the rheological response of branched LDPE melts. The significance of our new findings will be discussed in comparison with a low-rate constitutive theory as well as high-rate capillary measurements reported in the literature for LDPE melts.

In what follows, the experimental technique is described first. The experimental results on the molten LDPE are presented in Section 3. The results are further discussed in Section 4. Finally, the main conclusions are summarized in Section 5.

2 Experimental Technique

2.1 Kolsky Torsion Bar Rheometer. The new KTB rheometer for polymer melts has three components: a KTB system, a thermal chamber, and a sealed cone-and-plate rheometric cell. The experimental setup is shown schematically in Fig. 1(*a*). The KTB system consists essentially of two long aluminum-alloy circular bars—the input and output bars, respectively, a friction clamp of Duffy's design, [21], (detailed in A–A view), and a torsion driver (pulley). The thermal chamber has a through hole to allow the bars to sandwich the test specimen inside the chamber. In operation, the friction clamp is activated and the segment of the input bar from the pulley to the clamp is twisted to store a desired torque. Forcing the pre-notched bolt that locks the clamp to rapture releases the stored torque very rapidly, giving rise to a torsional loading wave that propagates towards the specimen and a

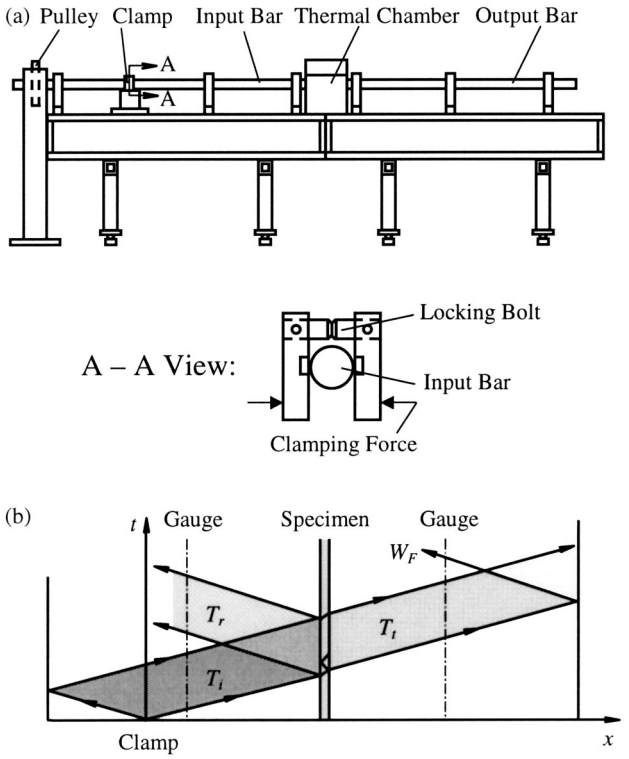


Fig. 1 Schematic of KTB rheometric experiment. (*a*) Experimental setup based on a Kolsky torsion bar system. (*b*) Torsional wave propagation during the experiment.

partial release wave of the same magnitude that travels towards the pulley and then reflects from it (which is nearly rigid in torsion) as a full unloading wave as shown in Fig. 1(*b*), which is a distance-time $(x-t)$ diagram illustrating the torsional wave propagation during the experiment. The loading and unloading waves form a torsional incident (loading) pulse T_i. By design, the torsional impedances of the bars and the shear stress associated with the incident pulse are significantly greater than the impedance and the material strength of the specimen, respectively. In consequence, most of the mechanical energy associated with T_i reflects from the specimen as a reflected pulse T_r while a small fraction of the energy (governed by the material strength of the specimen) transmits through the specimen as a transmitted pulse T_t propagating into the output bar. The primary mode of a linear elastic torsional wave in a circular waveguide is non-dispersive. Hence, the relative rotation between the two bars and the torque transmitted through the specimen can be determined straightforwardly from the shear strain gauge measurements of the three pulses away from the specimen (Fig. 1(*b*)). The related experimental analysis will be described later.

Several special designs and techniques have been employed to satisfy the particular needs arising from the rheometric application of interest. First, an asymmetric bar pairing is used. The input bar is a 7075-T6 aluminum alloy rod (25.4 mm diameter) whereas the output bar is a thin-walled 6061-T6 aluminum alloy tube (25.4 mm outer diameter and 0.89 mm wall thickness) as shown in Fig. 2(*a*). This new design retains a very large impedance difference between the input bar and specimen as necessary for steady high-rate loading while minimizing the torsional rigidity of the output bar to improve the accuracy of the measurements at low stresses. The smaller the output bar rigidity, the larger the shear strain for a given transmitted pulse, and thus the larger the signal-to-noise ratio.

Second, controlled convection heating is used to heat the speci-

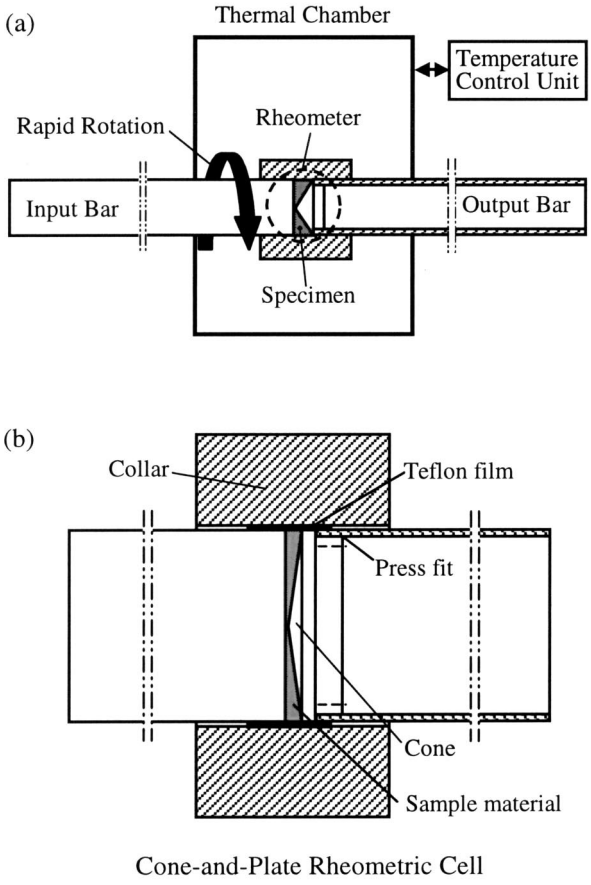


Fig. 2 Schematic of the rheometer assembly. (*a*) Thermal chamber and asymmetrically paired bars. (*b*) Cone-and-plate rheometric cell.

men and to maintain its initial temperature at a designed value. The heating is provided by resistance rope heaters and a forced air circulation system inside the thermal chamber (Fig. 2(*a*)). A thermocouple sensor and an OMEGA thermal control module form the temperature control unit that can regulate the temperature fluctuation to within $\pm 1°C$.

Finally, a sealed cone-and-plate rheometric cell is integrated into the system. As shown in Fig. 2(*b*), the cone-and-plate configuration is formed directly by a cone press-fitted into the tubular output bar and the flat end of the input bar. The motivation for this arrangement is that the dimension of a separately machined cone can be determined accurately with an optical profilometer. Although using the flat side to drive specimen shear flow causes small flow field nonuniformity in the vicinity of the outer edge, the effect is negligibly small compared to the other sources of experimental error. A thick-walled aluminum alloy collar that can slide smoothly on the bars is used for two purposes. One is to prevent the specimen from any large outward radial motion (driven by the centrifugal force) at high-speed rotations. The other is to improve the temperature uniformity across the specimen. The solid input bar conducts heat much quicker than the tubular output bar and a polymer melt is typically a poor conductor. Therefore, a significant temperature gradient may develop across the specimen. The collar provides a heat conduction bypass between the two bars and increases the thermal inertia of the rheometric cell. In addition, an experimental study was done to optimize the position of the rheometric cell inside the thermal chamber. As the result, the temperature variation in the specimen is within 1°C.

The use of the collar, however, gives rise to an edge effect. The specimen sticks to the inner surface of the collar so that the melt flow field near the surface is significantly different from that away from it. This problem is resolved by inserting a thin Teflon film between the collar and the specimen (Fig. 2(*b*)). For the temperatures of interest (below 300°C), the Teflon film remains an excellent *solid* lubricant. The torque transmitted by the soft Teflon film is negligibly small compared to that transmitted by the specimen. An experimental verification will be given later.

Foil resistance shear strain gauges forming a full Wheatstone bridge circuitry are installed on each bar for measuring the torsional pulses. The gauges used for the output bar have a nominal resistance of 3000 Ω to allow a high excitation voltage of 62 volts. The combination of high gauge excitation and low-impedance output bar provides a sufficient signal-to-noise ratio for determining shear stresses as low as 0.01 MPa. During the experiment, the strain gauge signals are fed to a NICOLET multichannel digitizer operated in common mode rejection and recorded at a 10-MHz sample rate and a 12-bit resolution. The data acquisition process is synchronized to obtain time-resolved measurements of the relevant torsional pulses. The analysis to determine the specimen shear rate and stress from the measurements is presented below.

2.2 Analysis of Experiment. Hartley et al. [21] described in detail the analysis of the conventional KTB experiment, in which the symmetric bar pairing allows the determination of the specimen shear rate $\dot{\gamma}_s$ solely based on the reflected pulse measurement. The asymmetric bar pairing used in this work requires, however, the use of all three pulses to determine $\dot{\gamma}_s$. An analysis specialized for this experimental configuration is therefore in order. In the derivation follows, subscripts 1 and 2 denote the quantities of the input and output bars, respectively, and *i*, *r*, and *t* the quantities associated with the incident, reflected and transmitted pulses, respectively.

Assume that the velocity field in the specimen is linear after the spin-up process. Then $\dot{\gamma}_s$ can be related to the angular velocities of the two bar-specimen interfaces $\omega_k\,(k=1,2)$ as

$$\dot{\gamma}_s=\frac{\omega_1 R_s-\omega_2 R_s \sin^2\alpha}{h\sin\alpha}, \tag{1}$$

where R_s, h, and α are the radius, vertex height and half vertex angle of the cone, respectively. In the equation, the denominator is the distance from the edge of the flat end to the cone surface (Fig. 2(*b*)) and the numerator the velocity difference over the distance. The characteristic equations governing the propagation of elastic torsional waves in the bars give

$$\omega_1=\frac{T_r-T_i}{K_1} \quad \text{and} \quad \omega_2=-\frac{T_t}{K_2}. \tag{2}$$

Here $K_k\equiv\rho_k c_{Sk} I_k\,(k=1,2)$ are the torsional impedances of the bars with ρ_k, c_{Sk}, and I_k being the material densities, shear wave speeds and polar moments of inertia of the bars, respectively. The torque pulses $T_q\,(q=i,r,t)$ can be related, through the elementary torsion formula, to the corresponding shear strain pulses $\gamma_q\,(q=i,r,t)$ as

$$T_i=\frac{c_{S1}K_1}{R_1}\gamma_i, \quad T_r=\frac{c_{S1}K_1}{R_1}\gamma_r \quad \text{and} \quad T_t=\frac{c_{S2}K_2}{R_{2o}}\gamma_t, \tag{3}$$

where R_1 is the radius of the input bar and R_{2o} is the outer radius of the output bar. The relations $G_k=\rho_k c_{Sk}^2\,(k=1,2)$ with G_k being the shear moduli of the bars have been used. Combining Eqs. (1)–(3) and using the relation $\sin\alpha=R_s/\sqrt{R_s^2+h^2}$ lead to the following expression:

$$\dot{\gamma}_s=\frac{\sqrt{R_s^2+h^2}}{h}\left[\frac{c_{S1}}{R_1}(\gamma_r-\gamma_i)+\frac{h^2 c_{S2}}{R_{2o}(R_s^2+h^2)}\gamma_t\right]. \tag{4}$$

Note that the occurrence of $\dot{\gamma}_s$ and the recordings of the three shear strain pulses are four events usually at different times. The determination of the shear rate-time profile $\dot{\gamma}_s(t)$ requires time shifts for the measured shear strain pulses. For an unknown specimen, a rigorous determination of all the time differences is difficult. In practice, a subjective ''best match'' of the three pulses in time is used as an approximation. The error is within ± 2 μs. Consequently, the slope of a rapid variation in $\dot{\gamma}_s$ is less accurate.

The shear stress in the specimen τ_s can be related straightforwardly to T_t by applying the stress continuity across the specimen-output bar interface and making use of the assumption of uniform shear stress in the specimen, i.e.,

$$T_t = \frac{2\pi}{3} \tau_s R_s^2 \sqrt{R_s^2 + h^2}. \tag{5}$$

The right-hand side of Eq. (5) is the result of integrating the torque produced by τ_s over the cone surface (Fig. 2(*b*)). Using Eq. (3) and $K_2 = \rho_2 c_{S2} I_2$ in Eq. (5) gives

$$\tau_s = \frac{3\rho_2 c_{S2}^2 (R_{2o}^4 - R_{2i}^4)}{4 R_{2o} R_s^2 \sqrt{R_s^2 + h^2}} \gamma_t, \tag{6}$$

where R_{2i} is the inner radius of the output bar, and the relation $I_2 = \pi(R_{2o}^4 - R_{2i}^4)/2$ has been used. Note that the time correlation of the shear stress-time profile $\tau_s(t)$ with $\dot{\gamma}_s(t)$ is as good as the time-shifting approximation used to determine $\dot{\gamma}_s(t)$.

The maximum torque that can be stored in the KTB apparatus used is 560 N-m. On the other hand, a minimum torque of 40 N-m is required for producing a loading pulse of good quality. For $h = 2$ mm (which is used in this work), the possible range of plateau shear rate is from 700 s^{-1} to 10^4 s^{-1}. For shear rates beyond this range, h needs to be modified accordingly.

The deviation of the specimen velocity from the assumed linear field may be estimated with the analysis by Walters and Waters [28]. For $\alpha = 81$ deg used in this work, the maximum deviation is about 0.4%, an insignificant value. The primary sources of experimental error are: (1) the deviation of the *effective* gauge output-bar shear strain relation, which is not exactly known, from the *idealized* gauge calibration (assuming a perfect Wheatstone bridge), which is actually used, and (2) the uncertainties associated with our measurements of the geometric and material parameters used in Eqs. (4) and (6). The error bars estimated with our error analysis and elastic tests are 4% for $\dot{\gamma}_s$ and 5% for τ_s. Although the temperature of the heated end of each bar (inside the thermal chamber) may approach 50% of the melting temperature of aluminum alloy during the experiment, the temperature rise at each gauge location is small because the gauge resides far away from the thermal chamber and most of the heat is conducted away through the bearing blocks between the gauge and chamber (Fig. 1(*a*)). As long as a temperature correction for the gauge factor is not necessary and the bars remain elastic during the experiment, even if a torsional impedance change does occur in the heated portions of the bars, it may only affect the measurements for the *initial slopes* of sudden changes in $\dot{\gamma}_s(t)$ and $\tau_s(t)$, for which the KTB technique is not very accurate anyway. However, for the temperatures covered in this work, there has been no sign that may indicate a significant torsional impedance change in the bars due to the heating of the specimen.

2.3 Description of Sample Material. The new KTB rheometer for polymer melts has been evaluated in a series of experiments on a molten low-density polyethylene (LDPE). The raw material was DuPont 20 Series LDPE resin, which had a material density of 920 kg/m^3 at 23°C, a melt index of 1.9 g/10 min and a melting point of 109°C (manufacturer's data). The resin pallets were melted in a vacuum oven and resolidified into a brick. Specimens of 2.5 mm thick and 25 mm in diameter were then machined from the brick. Each specimen was reheated sandwiched between the two bars and shaped to fill the gap of the rheometric cell (Fig. 2(*b*)). Excessive material (if any) was removed, a layer of thin Teflon film was applied, and the collar was slid into the place to complete the assembly. Finally, the assembled specimen was heated up again slowly to a designed initial temperature. For the temperatures covered in this work (150°C~210°C), the material, referred to hereafter as the DuPont-20 LDPE, is colorless. The molecular structure of the material is known to contain irregularly spaced long-chain branches.

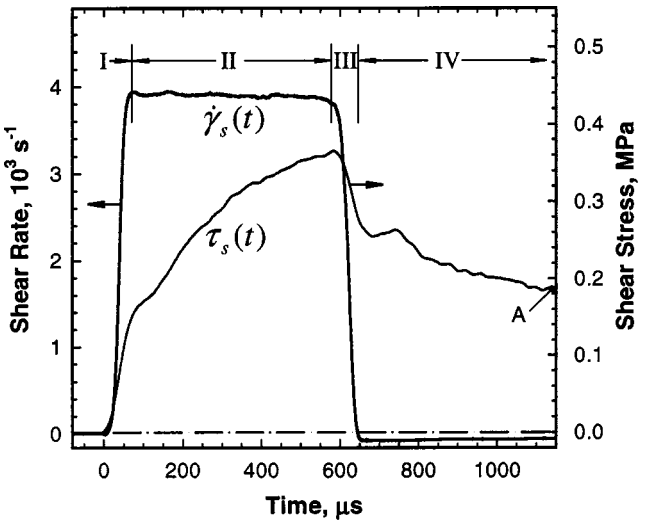


Fig. 3 Shear rate and shear stress profiles from a typical KTB rheometric experiment on the DuPont-20 LDPE melt at 190°C. The thicker line with respect to the left axis is the rate profile. The thinner line with respect to the right axis is the stress profile.

3 Experimental Results

The experiments presented below were designed to examine first the shear rate and shear strain dependences of the transient response of the DuPont-20 LDPE heated to 190°C and then the material's temperature dependence at a chosen shear rate. The shear rates examined range from 800 s^{-1} to 9900 s^{-1} (nearly the full span of the designed capacity). The material melts *completely* for temperatures above 130°C and color change was observed for the material heated in air to 230°C and beyond. Therefore, the variation of initial temperature was chosen to be from 150°C to 210°C.

The data from a typical experiment are presented in Fig. 3 in terms of the time-resolved profiles of $\dot{\gamma}_s$ and τ_s (the shear rate and shear stress in the specimen, respectively). The initial temperature of this experiment was 190°C. Under the applied rapid rotation, $\dot{\gamma}_s$ (the thicker line with respect to the left axis) increases to 3900 s^{-1} in a rise time of about 65 μs. This rapid spin-up process will hereafter be referred to as stage I. After stage I, $\dot{\gamma}_s$ remains nearly constant as designed until the arrival of unloading approximately 550 μs later. The unloading is as rapid as the loading. The period of the shear rate plateau and the subsequent rate releasing process will be referred to as stages II and III, respectively. Different from typical KTB experiments on metals, the present experiment did not end at the completion of the rate releasing that brings $\dot{\gamma}_s$ back to zero. Instead, $\dot{\gamma}_s$ runs across the baseline and evolves at small but measurable negative values. This period of negative $\dot{\gamma}_s$ will be referred to as stage IV. In response to the loading, τ_s (the thinner line with respect to the right axis) displays interesting behavior. It increases with $\dot{\gamma}_s$ in stage I as expected. However, in stage II, τ_s increases continuously with time while $\dot{\gamma}_s$ remains nearly constant. In other words, τ_s increases with the cumulative shear strain in the specimen γ_s at a constant $\dot{\gamma}_s$. Clearly, the material response is other than purely viscous (rate-dependent only). It appears to

have a significant *shear strain hardening*, a phenomenon that may occur in a viscoelastic fluid. For this experiment, the increase in τ_s during stage II is actually greater than that in stage I. Although the rising of τ_s in stage I also involves the contribution of strain hardening, it is due primarily to the increase in $\dot{\gamma}_s$ since the rising slope in stage I is much steeper than that in stage II.

The material response in stages III and IV is particularly interesting. In stage III, τ_s decreases with $\dot{\gamma}_s$ (Fig. 3). However, τ_s retains a large residual value when $\dot{\gamma}_s$ reaches zero. The evolution of τ_s thereafter (in stage IV) corresponds to a time-dependent elastic recovery of the specimen. Because the solid input bar is significantly more rigid than the thin-walled tubular output bar in torsion, this elastic recovery causes the output bar to rotate further in the same direction as in the previous stages but at a very small and time-dependent decaying angular velocity while the input bar-specimen interface remains essentially motionless. The result is the small-amplitude reverse shear rate profile seen in stage IV (Fig. 3). This elastic release appears to require much longer time to complete than the permissible time window of the experiment. Point A in Fig. 3 indicates the time when the reflection of the leading transmitted wave from the free end of output bar (W_F, Fig. 1(*b*)) arrives at the output bar gauge and the gauge signal can no longer be unambiguously related to τ_s. There seems a shear stress oscillation at the beginning of stage IV. It is not yet clear, however, whether this is real material response or a manifestation of the dynamic interactions between the specimen and the cone-tube assembly (Fig. 2(*b*)) as $\dot{\gamma}_s$ changes direction. A definite answer requires further study.

It should be pointed out that whereas τ_s appears to decrease *instantaneously* with $\dot{\gamma}_s$ in stage III, the decrease of τ_s in stage IV is entirely *time-dependent*. The data in this period reflect the relaxation behavior of the rapidly sheared material. In a sense, the KTB rheometric experiment is a high-rate transient test in loading and a step shear-relaxation test in unloading, considering the similarity between stage IV and the step shear-relaxation experiment, [3,5,9].

To verify that the use of a thin Teflon film as lubricant between the specimen and collar does not affect adversely the accuracy of shear stress measurement, an experiment without Teflon film was carried out. The collar used in the experiment had a recess to avoid direct contact with the specimen. A moderate input angular velocity was employed so that the centrifugal force during the experiment was insignificant. The initial temperature was also 190°C and the plateau shear rate was 3970 s^{-1}. In Fig. 4, the shear stress profile obtained from this experiment (the broken line) is compared with that from the previously described experiment (the solid line), which used a thin Teflon film and had a very close plateau shear rate of 3900 s^{-1}. The two profiles are seen to be nearly identical, indicating that the effect of thin Teflon film on the shear stress measurement is negligible compared to the expected experimental error.

Figure 5 presents the shear stress-time $\tau_s(t)$ profiles obtained from six experiments with an initial temperature of 190°C and shear rates ranging from 800 s^{-1} to 9900 s^{-1}. The plateau shear rate reached in each of the experiments is specified in the figure. For clarity, the data are shown for the rate loading and releasing cycle only without the time-dependent elastic release. The corresponding shear rate profiles are similar in pattern to that shown in Fig. 3 and will not be presented for the lack of space. The stress profiles for plateau shear rates up to 5510 s^{-1} bear the characteristic features similar to those described earlier. However, the rate of τ_s increase in stage II decreases with increasing plateau shear rate. The data from 7850 s^{-1} experiment seem to indicate the development of a shear stress plateau right before the arrival of rate releasing. Note that the loading pulse duration is the same for all the experiments shown. The higher the plateau shear rate, the larger the value of γ_s before the rate releasing. Hence, the later response in 7850 s^{-1} experiment may be a sign of the saturation of shear strain hardening. This becomes apparent in the data from 9900 s^{-1} experiment. If the initial rise in τ_s is viewed as predominantly rate-dependent response and the increase of τ_s afterwards as strain-dependent response, the former is more pronounced at higher shear rates while the latter is more pronounced at lower shear rates. In other words, the material response at higher shear rates is more viscous than viscoelastic and vice versa at lower shear rates.

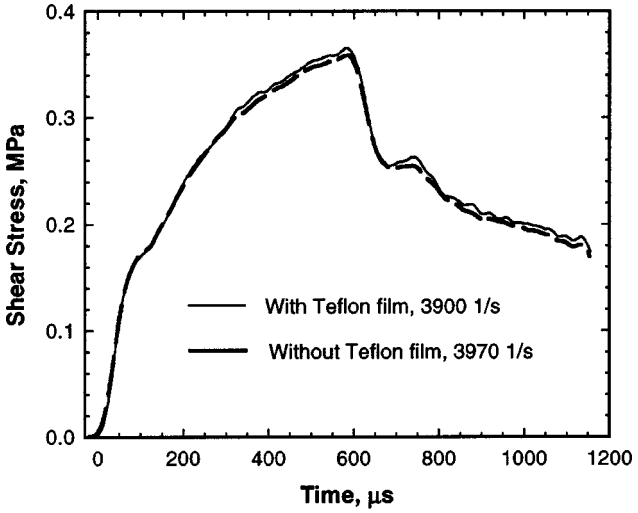


Fig. 4 Comparison of shear stress measurements with and without thin Teflon film. The initial temperatures for both experiments were 190°C. The solid line is the measurement with Teflon film and under a plateau shear rate of 3900 s^{-1}. The broken line is the measurement without Teflon film and under a plateau shear rate of 3970 s^{-1}.

The shear strain-time $\gamma_s(t)$ profile for each experiment was determined by integrating the $\dot{\gamma}_s(t)$ profile over the time so that the $\tau_s - \gamma_s$ relation is obtained. The results for the six experiments described in the previous paragraph are summarized in Fig. 6. Note that the initial rise is due primarily to the increase in $\dot{\gamma}_s$. The slope should not be viewed as a measure of elastic modulus. The strain hardening after the initial rise is apparently nonlinear. However, the hardening trends of all the experiments are surprisingly similar considering that the variation of the plateau shear rate

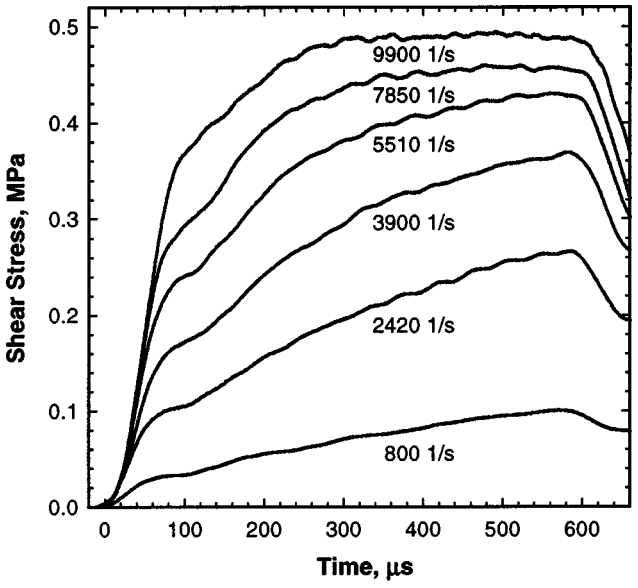


Fig. 5 Shear stress profiles for various shear rates and 190°C initial temperature

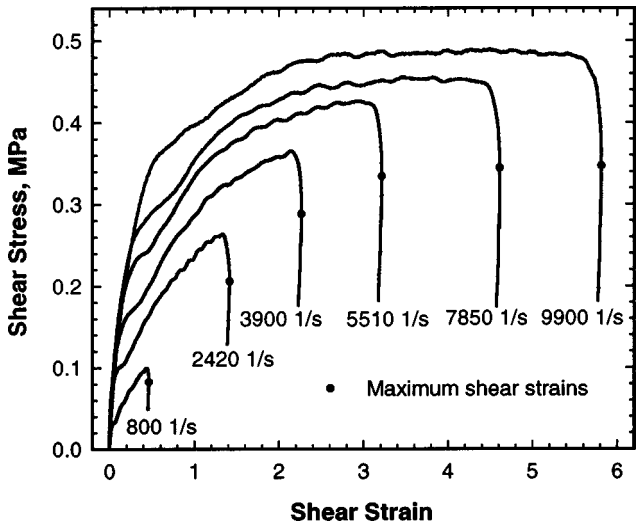


Fig. 6 Shear stress-shear strain relations for various shear rates and 190°C initial temperature. The circles indicate the maximum shear strains reached in the experiments.

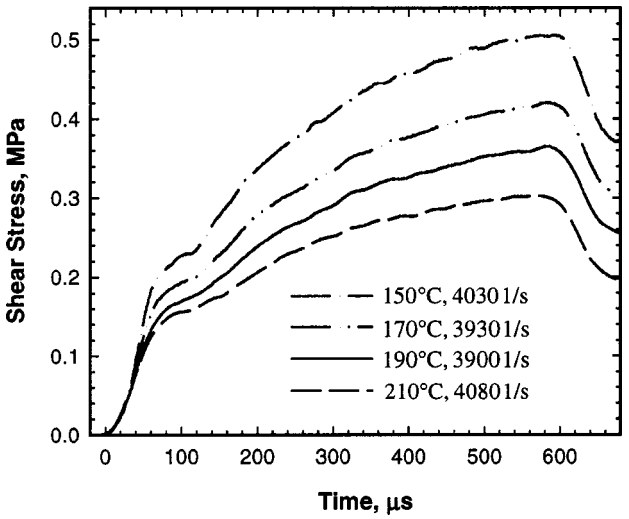


Fig. 7 Variation of transient response with initial temperature. The mean plateau shear rate is 3990 s^{-1} and the deviation is ±90 s^{-1}.

from one experiment to another spans more than one order of magnitude. This indicates that the observed shear strain hardening is, by and large, independent of $\dot{\gamma}_s$. A further implication is that a master curve of shear strain hardening can be extracted from the $\tau_s - \gamma_s$ curves. The maximum shear strain reached in 9900 s^{-1} experiment is approximately 580%. The data from the three highest-rate experiments together seem to indicate that the shear strain hardening saturates (or remains at a steady value) when γ_s reaches 300% and beyond.

If the $\tau_s - \gamma_s$ curve for a complete loading-unloading cycle is measured, the density of energy dissipation during the experiment can be determined by calculating the area enclosed by the $\tau_s - \gamma_s$ curve. Although each $\tau_s - \gamma_s$ curve in Fig. 6 has a significant portion of the unloading data, it does not contain the later part of release including the further elastic stress relaxation beyond the time window of experiment and the elastic recoil process under macroscopically ''zero-stress'' condition, [9]. In other words, the energy density calculated using the area underneath each $\tau_s - \gamma_s$ curve in Fig. 6 is somewhat greater than the true density of energy dissipation during the experiment. However, for an upper bound estimate of the temperature rise during the experiment, such an approximation is conservative and sufficiently accurate. The numerical result of such an estimate will be presented later.

The experiments to study the dependence of the transient large-strain response of the material on the initial temperature θ_0 were designed for a plateau shear rate of 4000 s^{-1}. Four experiments were carried out for $\theta_0 = 150 \sim 210$°C at a 20°C increment. The results are presented in Fig. 7 in terms of a comparison of $\tau_s(t)$ profiles. The plateau shear rates of the four experiments (3900~4090 s^{-1}) are very close to the targeted value so that the comparison is not biased by the shear rate. Qualitatively, the trends of the stress profiles are similar. Quantitatively, however, both the initial rise of τ_s (responding to the rapid increase of $\dot{\gamma}_s$) and the shear strain hardening after the initial rise increase with decreasing θ_0. Since, by design, the value of γ_s right before the unloading varies very slightly for these experiments, the significant variation in the strain hardening is due entirely to the change in θ_0. In comparison, the variation of the initial rise is less significant. In other words, the strain-dependent part of the material response is more sensitive to θ_0 than the rate-dependent (viscous) part of the material response. If the strain hardening continues to decrease with increasing θ_0, the material response may eventually become predominantly viscous rather than viscoelastic.

4 Discussion

It has been shown that the low-rate large-shear response of branched LDPE melts can be modeled quite well by incorporating a memory function into the rubberlike liquid theory, [29]. In particular, Laun [9] demonstrated that both the transient shear measurements (for $\dot{\gamma}_s$ up to 10 s^{-1}) and the steady-state apparent shear viscosity data (for $\dot{\gamma}_s$ up to 10^3 s^{-1}) for a molten branched LDPE can be well represented by an integral constitutive equation based on such a modified rubberlike liquid theory. For a step shear rate loading of intensity $\dot{\gamma}_s$ and starting at $t=0$, the analytical result of the integral equation is

$$\tau_s(t) = \dot{\gamma}_s t \sum_i \frac{n\dot{\gamma}_s \tau_i g_i}{1+n\dot{\gamma}_s\tau_i} \exp\left(-\frac{1+n\dot{\gamma}_s\tau_i}{\tau_i} t\right) + \dot{\gamma}_s \sum_i \frac{\tau_i g_i}{(1+n\dot{\gamma}_s\tau_i)^2}\left[1-\exp\left(-\frac{1+n\dot{\gamma}_s\tau_i}{\tau_i} t\right)\right], \quad (7)$$

where n, τ_i, and g_i are material parameters. Although $\dot{\gamma}_s$ appears in the both terms on the right-hand side of Eq. (7), the first one describes the effect of the shear strain ($\dot{\gamma}_s t$) and the second one accounts for the explicit response to $\dot{\gamma}_s$. Comparing Eq. (7) with our high-rate data on the molten DuPont-20 LDPE helps to identify the key difference between the response of a viscoelastic polymer melt under low shear-rate loading and that under high shear-rate loading. For short-time response, τ_s given by Eq. (7) does not jump with $\dot{\gamma}_s$ since for a very small t both the shear strain and shear rate terms are nearly zero. In contrast, τ_s shown in Fig. 3 changes *instantaneously* with a change in $\dot{\gamma}_s$. To some extent, the high-rate response is more of Kelvin-Voigt (i.e., $\tau_s = G\gamma_s + \eta\dot{\gamma}_s$, where G and η denote shear modulus and viscosity, respectively) than rubberlike liquid (Eq. (7)). The fact that the rate-dominant initial rise in τ_s decreases with the plateau shear rate (Fig. 5) suggests that the instantaneous viscous response of a viscoelastic fluid may not be discernible in low-rate rheometric experiments. The type of high-rate transient measurements presented here is important in developing better material models for viscoelastic polymer melts.

For long-time (large-strain) response, the exponential terms in Eq. (7) vanish giving rise to ''viscous like'' shear response, which depends only on the steady-state apparent shear viscosity, the first term in the second summation. The high-rate $\tau_s - \gamma_s$ data shown in Fig. 6, on the other hand, indicate that whereas the shear strain hardening saturates for $\gamma_s > 300\%$, the effect of γ_s does not fade

in such a steady state. Therefore, although a relationship between the steady-state $\tau_s/\dot{\gamma}_s$ ratio and $\dot{\gamma}_s$ may also be determined for high shear-rate loading, it is fundamentally different from the steady-state apparent shear viscosity as in Eq. (7). Another evidence to support this conclusion is the fact that the full rate releasing only releases a fraction of the shear stress sustained in the steady state (the three higher-rate experiments, Figs. 5 and 6).

The observed viscoelastic flow behavior may be explained at the molecular level as the following. Driven by active thermal motions in the molten polymer, vastly many molecular segments may be disentangled from surrounding molecules at any given instance so that they are momentarily ''free'' to move. Under a biased mechanical field, these disentangled molecular segments may either move in the direction dictated by the applied field to new locations thus forming a flow at the macroscopic level or swing to align with the flow direction. In each case, such a molecular segment needs to pass through the network of surrounding molecules. The viscous drag of the molecular network against this passing-through motion gives rise not only to viscous stress that depends on the velocity of relative motion but also to elastic stretching of the molecular segment as well as the molecular chain it associates with. At the same time, the moving and stretched molecular segment may be part of molecular network for other moving segments to pass through. Hence, as the flow evolves, the molecular chains are quickly stretched and consequently stiffened. The molecular network becomes more viscously resistive against the passing-through molecular segments, resulting in a higher viscous stress for the same applied strain rate. However, once the molecular chains are stretched to such an extent that the elastic stress increment due to further stretching will exceed the corresponding viscous stress increment, the hardening will stop as required by the equilibrium condition. From this perspective, the observed shear strain hardening can be viewed as the result of a material state variation from an undeformed, less viscous molecular network to a highly stretched, more viscous one. As such, the material response during loading may be described with a viscous model having a strain-dependent apparent viscosity. However, such a model will be invalid under rapid unloading or oscillatory loading. This is because the reverse transition through relaxation to a more compliant molecular network is *time-dependent* and can be significantly slower than the applied load variation. For a strained viscoelastic polymer melt, the effect of elastic deformation of molecular network will not fade quickly. A useful model for this type of material behavior needs to keep track of the contribution of elastic strain.

As an attempt to separate empirically the contributions of shear rate and shear strain to the viscoelastic flow response of the molten DuPont-20 LDPE at high shear rates, the *loading* portion of the data obtained for 190°C initial temperature is treated as the sum of a rate-dependent viscous jump $\Delta\tau_s(\dot{\gamma}_s)$ and a rate-independent master curve of shear strain hardening, for which the function $0.2592\tanh(0.72\gamma_s)$ MPa was found to fit the data well with deviations being within 4%. Figure 8 shows the results of subtracting the master curve from the loading portion of the $\tau_s - \gamma_s$ curves shown in Fig. 6. Step jump appears to be a good approximation to all the curves, demonstrating that the proposed separation is a reasonable treatment. According to the master curve, τ_s reached a steady state (the saturation of shear strain hardening) in the experiments with a plateau shear rate of 5510 s^{-1} or higher (Fig. 5). For the three experiments at lower rates, the steady-state τ_s can be estimated using the master curve. In Fig. 9, both the steady-state $\tau_s/\dot{\gamma}_s$ ratios measured (the filled squares) and estimated (the open squares) are presented and compared with the steady-state apparent shear viscosity results on a PEMEX LDPE melt (the circles on the solid line), [6], and a Union Carbide LDPE melt (the dashed line), [14], both from the capillary measurements at 190°C. In terms of absolute value, the three sets of results clearly disagree one with another, indicating significant material difference (likely in molecular weight). However, the decreasing trend of the steady-state $\tau_s/\dot{\gamma}_s$ ratio with increasing $\dot{\gamma}_s$ (or the shear shinning) is close to those of the steady-state apparent shear viscosity results. It should be pointed out that derived from the capillary data the latter rely on an a priori assumed flow field in the extruder, [5]. Therefore, the steady-state $\tau_s/\dot{\gamma}_s - \dot{\gamma}_s$ relation determined directly by the KTB rheometric experiment can be used to verify the high-rate capillary results, the traditional database for polymer processing design and control. Also shown in Fig. 9 are the ratios of $\Delta\tau_s(\dot{\gamma}_s)$ to the plateau shear rate (the triangles). The shear thinning in this set of data is seen to be much weaker than that of the steady-state $\tau_s/\dot{\gamma}_s$ ratios. In fact, there is almost no thinning for shear rates up to about 5500 s^{-1}. This

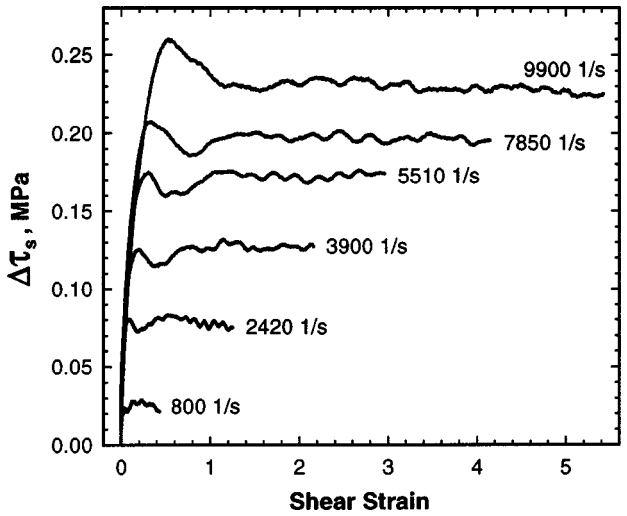


Fig. 8 Viscous stress jumps for various shear rates and 190°C initial temperature

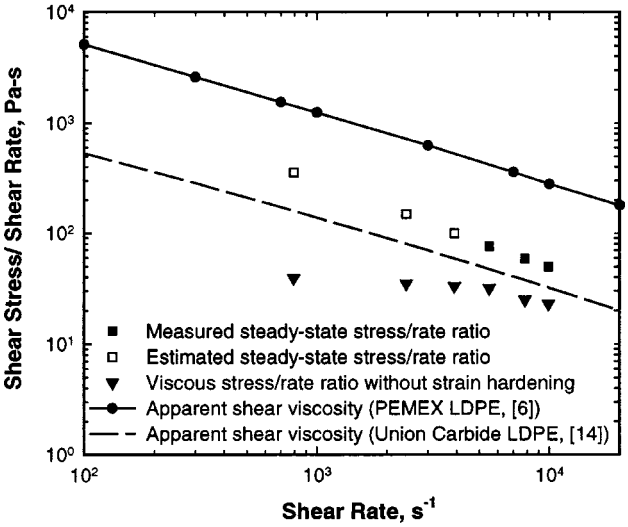


Fig. 9 Comparison of the ratios of shear stress to shear rate with the apparent shear viscosity results on similar LDPE melts. The squares are the steady-state data and estimates for the DuPont-20 LDPE melt at 190°C. The triangles are the results corrected for the shear strain hardening. The circles and the broken line are respectively the capillary results of the apparent shear viscosity of a PEMEX LDPE melt, [6], and that of a Union Carbide LEDP melt, [14], both at 190°C.

demonstrates, from a different perspective, the ambiguity of using the steady-state apparent shear viscosity in describing the high-rate response of a viscoelastic fluid.

Finally, it should be pointed out that the effects of temperature increase during the KTB rheometric experiment have not been included in the above analysis and discussion. The experiment is nearly adiabatic because the time scale is much shorter than that for significant heat conduction. Hence, the specimen temperature is expected to increase somewhat with the energy dissipation during the experiment. Although it is difficult to measure the temperature change during the experiment, its upper bond can be estimated as discussed in the previous section. For the experiments with 190°C initial temperature, the upper bond of temperature rise is less than 1.35°C, an insignificant value on the order of the temperature measurement uncertainty. However, for larger strains (under longer loading pulses) and higher stresses (at lower temperatures), the viscous heating may become significant and the calculation for the temperature change during the experiment may become necessary. A related issue is that the viscous heating together with the saturation of the strain hardening may initiate the development of an inhomogeneous flow field and further an adiabatic shear bending in the specimen, [30]. The evolution of such a shear bending should, however, be accompanied by a significant reduction in τ_s. There is no such a sign in the experimental data obtained in this work.

5 Closing Remarks

The Kolsky torsion bar technique has been successfully utilized in a novel rheometric experiment for characterizing the transient large-strain response of polymer melts under shear rates from $10^2\ s^{-1}$ to $10^4\ s^{-1}$ and temperatures up to 300°C. The new KTB rheometer measures the specimen response under a constant shear rate and at a microsecond resolution and is particularly useful for viscoelastic polymer melts in that the effect of shear strain can be examined separately from that of shear rate. The loading conditions realized with this technique enable revelations of important characteristic features of the high-rate response of viscoelastic polymer melts that can not be observed with the existing rheometric techniques for polymer melts.

A series of KTB rheometric experiments have been carried out on molten samples of the DuPont-20 (branched) LDPE. The experiments cover a range of plateau shear rates from $800\ s^{-1}$ to $9900\ s^{-1}$ and several initial temperatures from 150°C to 210°C in a 20°C increment. The maximum shear strain reached in the experiments is 580%. The results show that the transient response of the material under the high shear-rate loading not only is time-dependent (or strain-dependent) but also has an *instantaneous* dependence on the shear rate. The latter phenomenon is qualitatively different from the rate dependence described by the modified rubberlike liquid theory, which has been shown to represent the low-rate response of branched LDPE melt well, [9]. The experimental data further suggest that the explicit rate dependence decreases with the shear rate and may become indiscernible at low shear rates. Transient high-rate large-strain measurements are necessary for characterizing this part of the material response.

The strain dependence of the material under the high shear-rate loading appears as a significant shear strain hardening for shear strains up to 300% where the strain hardening saturates and a steady state is reached. However, the strain effect does not fade for the time duration of the experiment. The ratio of shear stress to shear rate is a function of shear strain. An important implication is that the shear thinning shown in the available capillary measurements of the high-rate apparent shear viscosity of LDPE melts may actually be a manifestation of the saturation of *shear strain* hardening at large shear strains instead of a measure of the *shear rate* dependence of the viscosity. If the rate-dependent viscous jump is separated out of the high-rate transient response of the DuPont-20 LDPE melt at 190°C, it shows little shear (rate) thinning for shear rates up to $5500\ s^{-1}$ (Fig. 9).

The dependence of the material response on the initial temperature has been studied at a plateau shear rate of about $3990\ s^{-1}$. There is no qualitative difference in the material response for the range of initial temperatures examined (from 150°C to 210°C). Quantitatively, however, the strain-dependent part of the response softens more rapidly with increasing temperature than the instantaneous rate-dependent part of the response. Predominantly viscous response can be expected for the material at significantly higher temperatures. At lower temperatures, the shear strain hardening of the material will stiffen. However, the material response at temperatures above the melting point is still expected to be different from that of a rubberlike liquid by a noticeable instantaneous dependence on the shear rate.

Acknowledgments

The authors would like to thank Dr. M. Negahban for providing the raw material and the guidance in specimen preparation, and for many useful discussions, Dr. R. J. Clifton of Brown University for pointing out the importance of the release data, M. D. Kollath for designing and constructing the thermal chamber, and H. Huang for assistance in conducting the experiments. Financial support from the Nebraska Research Initiative Program and the University of Nebraska Foundation through a Layman Award are also gratefully acknowledged.

References

[1] Levy, S., and Carley, J. F., 1989, *Plastics Extrusion Technology Handbook*, Industrial Press, New York.
[2] Lodge, A. S., 1964, *Elastic Liquids*, Academic Press, London.
[3] Tanner, R. I., 1985, *Engineering Rheology*, Clarendon Press, Oxford.
[4] Doi, M., and Edwards, S. F., 1986, *The Theory of Polymer Dynamics*, Clarendon Press, Oxford.
[5] Whorlow, R. W., 1980, *Rheological Techniques*, Ellis Horwood, Chichester, UK.
[6] Montes, S. A., 1984, "Rheological Properties of Blown Film Low-Density Polyethylene Resins," Polym. Eng. Sci., **24**, pp. 259–263.
[7] Samurkas, T., and Rogers, M. G., 1992, "The Melt Rheology of Poly-ε-Caprolactam/Polypropylene Alloy: A VAMAS Round-Robin Study," Polym. Eng. Sci., **32**, pp. 1727–1733.
[8] Laun, H. M., and Schuch, H., 1989, "Transient Elongational Viscosities and Drawability of Polymer Melts," J. Rheol., **33**, pp. 119–175.
[9] Laun, H. M., 1986, "Prediction of Elastic Strains of Polymer Melts in Shear and Elongation," J. Rheol., **30**, pp. 459–501.
[10] Kobayashi, M., Takahashi, T., Takimoto, J., and Koyama, K., 1996, "Influence of Glass Beads on the Elongational Viscosity of Polyethylene With Anomalous Strain Rate Dependence of the Strain-Hardening," Polymer, **37**, pp. 3745–3747.
[11] Khanna, Y. P., 1991, "Dynamic Melt Rheology. I: Re-Examining Dynamic Viscosity in Relationship to the Steady Shear Flow Viscosity," Polym. Eng. Sci., **31**, pp. 440–444.
[12] Khanna, Y. P., and Slusarz, K. R., 1993, "Dynamic Melt Rheology. II: Re-Examining the Relationship of G′ in Oscillatory Rheometry to the Melt Elasticity," Polym. Eng. Sci., **33**, pp. 122–124.
[13] Giacomin, A. J., Samurkas, T., and Dealy, J. M., 1989, "A Novel Sliding Plate Rheometer for Molten Plastics," Polym. Eng. Sci., **29**, pp. 499–504.
[14] Agur, E. E., and Vlachopoulos, J., 1982, "Numerical Simulation of a Single-Screw Plasticating Extruder," Polym. Eng. Sci., **22**, pp. 1084–1094.
[15] Öttinger, H. C., 1999, "A Thermodynamically Admissible Reptation Model for Fast Flows of Entangled Polymers," J. Rheol., **43**, pp. 1461–1494.
[16] McLeish, T. C. B., and Larson, R. G., 1998, "Molecular Constitutive Equations for a Class of Branched Polymers: The Pom-Pom Polmer," J. Rheol., **42**, pp. 81–110.
[17] Porter, D., 1997, "Combining Molecular and Continuum Mechanics Concepts for Constitutive Equations of Polymer Melt Flow," J. Non-Newtonian Fluid Mech., **68**, pp. 141–152.
[18] Kolsky, H., 1949, "An Investigation of the Mechanical Properties of Materials at Very High Rates of Loading," Proc. R. Soc. London, Ser. B, **62**, pp. 676–700.
[19] Baker, W. W., and Yew, C. H., 1966, "Strain Rate Effects in the Propagation of Torsional Plastic Waves," ASME J. Appl. Mech., **33**, pp. 917–923.
[20] Lewis, J. L., and Campbell, J. D., 1972, "The Development and Use of a Torsional Hopkinson Bar Apparatus," Exp. Mech., **12**, pp. 520–524.
[21] Hartley, K. A., Duffy, J., and Hawley, R. H., 1985, "The Torsional Kolsky (Split-Hopkinson) Bar," ASM Metals Hand., **43**, pp. 873–896.
[22] Duffy, J., Campbell, J. D., and Hawley, R. H., 1971, "On the Use of a Torsional Split Hopkinson Bar to Study Rate Effects in 1100-O Aluminum," ASME J. Appl. Mech., **38**, pp. 83–91.
[23] Frantz, R. A., and Duffy, J., 1972, "The Dynamic Stress-Strain Behavior in

Torsion of 1100-O Aluminum Subjected to a Sharp Increase in Strain Rate,'' ASME J. Appl. Mech., **39**, pp. 939–945.
[24] Pao, Y. H., and Gilot, A., 1992, ''High Strain Rate Deformation and Failure of A533B Steel at Various Temperatures,'' Acta Metall., **40**, pp. 1271–1280.
[25] Chichili, D. R., and Ramesh, K. T., 1999, ''Recovery Experiments for Adiabatic Shear Localization: A Novel Experimental Technique,'' ASME J. Appl. Mech., **66**, pp. 10–20.
[26] Clyens, S., Evans, C. R., and Johnson, K. L., 1982, ''Measurement of the Viscosity of Supercooled Liquids at High Shear Rates With a Hopkinson Torsion Bar,'' Proc. R. Soc. London, Ser. A, **381**, pp. 215–222.
[27] Feng, R., and Ramesh, K. T., 1993, ''The Rheology of Lubricants at High Shear Rate,'' ASME J. Tribol., **115**, pp. 640–649.
[28] Walters, K., and Waters, N. D., 1968, ''On the Use of a Rheogoniometer. Part I—Steady Shear,'' *Polymer Systems*, R. E. Wetton and R. W. Whorlow, eds., William Clowes and Sons, London, pp. 211–235.
[29] Wagner, M. H., and Stephenson, S. E., 1979, ''The Irreversibility Assumption of Network Disentanglement in Flowing Polymer Melts and Its Effects on Elastic Recoil Predictions,'' J. Rheol., **23**, pp. 489–504.
[30] Feng, R., Ramesh, K. T., and Douglas, A. S., 1996, ''An Analytical and Computational Investigation of High-Rate Rheometry,'' ASME J. Tribol., **118**, pp. 601–607.

Dynamic Stability of Disks With Periodically Varying Spin Rates Subjected to Stationary In-Plane Edge Loads

T. H. Young
Professor

M. Y. Wu
Graduate Student

Department of Mechanical Engineering,
National Taiwan University of Science and Technology,
Taipei, Taiwan

This paper presents an analysis of dynamic stability of an annular plate with a periodically varying spin rate subjected to a stationary in-plane edge load. The spin rate of the plate is characterized as the sum of a constant speed and a small, periodic perturbation. Due to this periodically varying spin rate, the plate may bring about parametric instability. In this work, the initial stress distributions caused by the periodically varying spin rate and the in-plane edge load are analyzed first. The finite element method is applied then to yield the discretized equations of motion. Finally, the method of multiple scales is adopted to determine the stability boundaries of the system. Numerical results show that combination resonances take place only between modes of the same nodal diameter if the stationary in-plane edge load is absent. However, there are additional combination resonances between modes of different nodal diameters if the stationary in-plane edge load is present. [DOI: 10.1115/1.1753267]

Introduction

Spinning disks find wide applications in mechanical engineering, from early circular saw blades, turbine rotors to recent floppy and compact disks, etc.. With the progress of industrial technology, the speeds of spinning disks have been increasing, and the ensuing vibration problem has been exacerbated. Thus, the topics pertaining to the vibration and stability of spinning disks have attracted much attention to researchers since the early 1920s. Early investigations dealt primarily with the determination of natural frequencies and critical speeds of spinning disks. Lamb and Southwell [1] first derived the frequency equation and plotted the mode shapes of a spinning circular plate. Southwell [2] continued to analyze the free vibration of a spinning annular plate with clamped inner edge and free outer edge.

Spinning disks are usually subjected to in-plane loading in application. Mote [3,4] studied the free vibration of circular disks subjected to initial in-plane stresses introduced purposely by rolling or to thermal membrane stresses resulting from the cutting process. The free vibration of a spinning disk under a concentrated radial edge load was first investigated by Carlin and his coworkers, [5]. Redcliffe and Mote [6] extended the work of Carlin et al. by considering a general concentrated edge load with both normal and tangential components. In these two papers, the concentrated edge load is spinning along with the spinning disk, that is, there is no relative motion between the disk and the edge load.

In most real situations, the loads of spinning disks are fixed in space. Iwan and Moeller's work [7] appears to be the first publication on this subject. They investigated the free vibration of a spinning disk transversely in contact with a stationary load system consisting of a mass, spring, and dashpot. Ono et al. [8] extended Iwan and Moeller's work to include the pitching parameters as well as the friction force between the spinning disk and the stationary load system. A further study of effects of load parameters, such as friction force, transverse mass, damping and stiffness, and the analogous pitching parameters, of a stationary load system on the free vibration of a spinning disk was conducted by Chen and Bogy [9].

Afterward Chen [10] continued his research efforts to study the stability of a spinning disk under a stationary concentrated edge load in the radial direction. The edge load may be a conservative or follower force, and the analysis is based on the inertial coordinates. Numerical results show that a stationary, conservative, and compressive edge load decreases the natural frequencies of the forward and backward traveling waves but increases the natural frequencies of the so-called reflected waves. The compressive edge load induces a divergence-type instability before the critical speed and a flutter-type instability beyond the critical speed when a reflected wave meets a forward or backward wave. In addition, the effect of the conservative edge load on the natural frequencies of the spinning disk is mostly through the transverse component of the edge load and is much more important than the effect of the follower edge load. The parametric instability of a spinning disk under the action of space-fixed pulsating edge loads was further investigated by Chen [11]. He found that sum-type resonances occur between both reflected modes or both nonreflected modes, but difference-type resonances take place when one mode is reflected and the other is nonreflected.

In the mean time, Shen and Song [12] also studied independently the stability of a spinning disk under a stationary concentrated in-plane edge load. The edge load is more general, having not only the normal component but also the tangential component. The analysis is based on the rotating coordinates attached to the spinning disk, and parametric instability is shown to exist. The results reveal that the radial component of the in-plane edge load dominates the rotational speed at which parametric instability occurs, and the tangential component of the in-plane edge load dominates the width of instability zones.

The spin rates of the disks considered in all abovementioned references are constant. However, in the real world, the spin rates of the disks usually fluctuate within a small interval. The work by Kammer and Schlack [13] appears to be the first one concerning this topic. In this paper, the angular speed is expressed as the sum of a constant and a small periodic perturbation, and the KBM

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, September 19, 2001; final revision, June 6, 2002. Associate Editor: N. C. Perkins. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

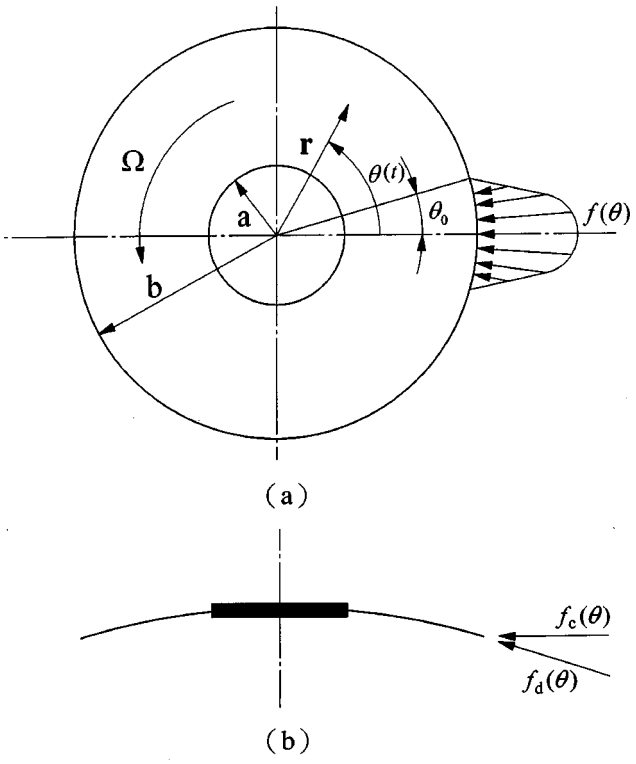


Fig. 1 Configuration of a spinning disk subjected to a stationary in-plane edge load

method is used to derive approximate solutions. Later Young and Liou [14,15] applied the method of multiple scales to investigate the parametric instability of a rotating cantilever plate and a rotor-bearing system with nonconstant rotating speeds. The nonlinear response of spinning disks with nonconstant spin rates was also studied by Young [16]. Sum-type resonances are shown to exist between modes of the same nodal diameter. This work extends the past research efforts of the authors to analyze the dynamic stability of annular plates with periodically varying spin rates subjected to stationary in-plane edge loads.

Equation of Motion

Figure 1 shows an annular plate which is clamped at the inner radius $r=a$ and subjected to a radial traction $f(\theta)$ at the outer radius $r=b$, where (r,θ) is an inertial coordinate system fixed in space. The disk is spinning with an angular speed Ω, while the radial edge traction is fixed in space. The radial edge traction may be a conservative force f_c or a follower one f_d and is assumed to be symmetrically distributed. The equations governing the in-plane stresses σ^*_{rr}, $\sigma^*_{\theta\theta}$, and $\sigma^*_{r\theta}$ of the plate due to rotation are ([17])

$$\frac{\partial\sigma^*_{rr}}{\partial r}+\frac{\sigma^*_{rr}-\sigma^*_{\theta\theta}}{r}+\frac{1}{r}\frac{\partial\sigma^*_{r\theta}}{\partial\theta}=-\rho r\Omega^2$$
$$\frac{\partial\sigma^*_{r\theta}}{\partial r}+\frac{1}{r}\frac{\partial\sigma^*_{\theta\theta}}{\partial\theta}+\frac{2\sigma^*_{r\theta}}{r}=\rho r\frac{\partial\Omega}{\partial t} \tag{1}$$

where ρ is the mass density of the plate. Equation (1) is linear; consequently, the resulting stress field will be the sum of two parts: one due to the centripetal acceleration and the other due to the angular acceleration. The former is axisymmetric and can be obtained as

$$\sigma^*_{rr1}=c^*_o+c^*_1\left(\frac{a}{r}\right)^2-\frac{3+\nu}{8}\rho r^2\Omega^2$$
$$\sigma^*_{\theta\theta1}=c^*_o-c^*_1\left(\frac{a}{r}\right)^2-\frac{1+3\nu}{8}\rho r^2\Omega^2 \tag{2}$$
$$\sigma^*_{r\theta1}=0$$

and the latter is found to be

$$\sigma^*_{rr2}=\sigma_{\theta\theta2}=0$$
$$\sigma^*_{r\theta2}=c^*_2\left(\frac{b}{r}\right)^2+\frac{1}{4}\rho r^2\frac{\partial\Omega}{\partial t} \tag{3}$$

where c^*_o, c^*_1, and c^*_2 are undetermined coefficients, and ν is Poisson's ratio of the plate. By using the boundary conditions $u^*(a)=\sigma^*_{rr}(b)=\sigma^*_{r\theta}(b)=0$, where u^* is the radial displacement of the plate, these undetermined coefficients can be solved as

$$c^*_o=\frac{1+\nu}{8}\frac{(1-\nu)a^4+(3+\nu)b^4}{(1-\nu)a^2+(1+\nu)b^2}\rho\Omega^2$$
$$c^*_1=\frac{1-\nu}{8}\frac{(3+\nu)b^2-(1+\nu)a^2}{(1-\nu)a^2+(1+\nu)b^2}\rho r^2\Omega^2 \tag{4}$$
$$c^*_2=-\frac{1}{4}\rho b^2\frac{\partial\Omega}{\partial t}.$$

The radial edge traction is assumed to be symmetrically distributed with respect to θ; hence, it can be expended into a Fourier cosine series,

$$f(\theta)=\sum_{k=0}^{\infty}f_k\cos k\theta \tag{5}$$

where f_k are Fourier coefficients. By the symmetry class of the stresses, the in-plane stresses $\tilde{\sigma}_{rr}$, $\tilde{\sigma}_{\theta\theta}$, and $\tilde{\sigma}_{r\theta}$ of the plate due to the radial edge traction may be assumed of the form

$$\tilde{\sigma}_{rr}(r,\theta)=\sum_{k=0}^{\infty}\tilde{\sigma}_{rrk}(r)\cos k\theta$$
$$\tilde{\sigma}_{\theta\theta}(r,\theta)=\sum_{k=0}^{\infty}\tilde{\sigma}_{\theta\theta k}(r)\cos k\theta \tag{6}$$
$$\tilde{\sigma}_{r\theta}(r,\theta)=\sum_{k=0}^{\infty}\tilde{\sigma}_{r\theta k}(r)\sin k\theta$$

where $\tilde{\sigma}_{rrk}$, $\tilde{\sigma}_{\theta\theta k}$, and $\tilde{\sigma}_{r\theta k}$ are Fourier coefficients of $\tilde{\sigma}_{rr}$, $\tilde{\sigma}_{\theta\theta}$, and $\tilde{\sigma}_{r\theta}$, respectively. Using the generalized two-dimensional elasticity solution in polar coordinates and imposing the periodicity conditions of the radial and circumferential displacements $\tilde{u}$ and $\tilde{v}$, and the disk boundary conditions, $\tilde{u}(a)=\tilde{v}(a)=\tilde{\sigma}_{r\theta k}(b)=0$, $\tilde{\sigma}_{rrk}(b)=-f_k/h$, where h is the thickness of the plate, the in-plane stress components $\tilde{\sigma}_{rrk}$, $\tilde{\sigma}_{\theta\theta k}$, and $\tilde{\sigma}_{r\theta k}$ can be solved. The solutions are given in Carlin et al. [5].

With respect to the inertial polar coordinates, the transverse displacement of the plate can be expressed as $w=w(r,\theta(t),t)$. Therefore, the equation of motion of a spinning disk with viscous damping subjected to stationary conservative edge loads can be written as

$$\frac{D}{h}\nabla^4 w+\frac{c}{h}\left(\frac{\partial w}{\partial t}+\Omega\frac{\partial w}{\partial \theta}\right)+\rho\left(\frac{\partial^2 w}{\partial t^2}+2\Omega\frac{\partial^2 w}{\partial t\partial\theta}+\Omega^2\frac{\partial^2 w}{\partial\theta^2}\right)$$
$$=\frac{1}{r}\frac{\partial}{\partial r}\left(\sigma^*_{rr}r\frac{\partial w}{\partial r}\right)+\frac{\sigma^*_{\theta\theta}}{r^2}\frac{\partial^2 w}{\partial\theta^2}+2\sigma^*_{r\theta}\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial w}{\partial\theta}\right)$$
$$+\frac{1}{r}\left[\frac{\partial}{\partial r}\left(\tilde{\sigma}_{rr}r\frac{\partial w}{\partial r}+\tilde{\sigma}_{r\theta}\frac{\partial w}{\partial\theta}\right)+\frac{\partial}{\partial\theta}\left(\tilde{\sigma}_{r\theta}\frac{\partial w}{\partial r}+\frac{\tilde{\sigma}_{\theta\theta}}{r}\frac{\partial w}{\partial\theta}\right)\right]$$
$$+\frac{f(\theta)}{h}\delta(r-b)\frac{\partial w}{\partial r} \tag{7}$$

where c is the viscous damping coefficient of the plate; D is the flexural rigidity of the plate, $D=Eh^3/12(1-\nu^2)$, in which E is Young's modulus of the plate; $\delta(\cdot)$ is the Dirac delta function; and ∇^4 is the biharmonic operator in the polar coordinates. If the stationary edge load is of follower type, the last term in Eq. (7) is dropped. Note that when the spin rate is not constant as the case studied in this work, there exists a term $\rho(\partial\Omega/\partial t)(\partial w/\partial\theta)$ on both sides of the above equation and is canceled out each other. The boundary conditions of the disk in terms of the transverse displacement w are

$$w=\frac{\partial w}{\partial r}=0, \quad \text{at } r=a$$

$$\frac{\partial}{\partial r}\left(\frac{\partial^2 w}{\partial r^2}+\frac{1}{r}\frac{\partial w}{\partial r}+\frac{1}{r^2}\frac{\partial^2 w}{\partial\theta^2}\right)+\frac{1-\nu}{r^2}\left(\frac{\partial^3 w}{\partial r\partial\theta^2}-\frac{1}{r}\frac{\partial^2 w}{\partial\theta^2}\right)$$
$$=0, \quad \text{at } r=b \tag{8}$$
$$\frac{\partial^2 w}{\partial r^2}+\nu\left(\frac{1}{r}\frac{\partial w}{\partial r}+\frac{1}{r^2}\frac{\partial^2 w}{\partial\theta^2}\right)=0.$$

Solution Methods

Equation (7) is a fourth-order partial differential equation with variable coefficients and is unable to be solved directly. A solution of this equation is assumed to be in the form of a Fourier series,

$$w(r,\theta,t)=\sum_{j=0}^{J}[P_j(r,t)\cos j\theta+Q_j(r,t)\sin j\theta]. \tag{9}$$

Substituting Eq. (9) into Eq. (7) and equating the coefficients of each $\cos j\theta$ and $\sin j\theta$ yields a set of partial differential equations for the functions $P_j(r,t)$ and $Q_j(r,t)$. This set of partial differential equations is still too complicated to have exact solutions. Thus, the finite element method is adopted to eliminate the dependence upon the spatial coordinate r. By using the two-noded beam element, the functions $P_j(r,t)$ and $Q_j(r,t)$ within an element can be assumed as ([18])

$$P_j^e=\boldsymbol{\psi}^T\mathbf{u}_j^e, \quad Q_j^e=\boldsymbol{\psi}^T\mathbf{v}_j^e \tag{10}$$

where $\boldsymbol{\psi}$ is the shape function vector, and $\mathbf{u}_j^e$ and $\mathbf{v}_j^e$ are nodal parameter vectors within an element. Going through the finite element formulation yields the following two sets of discretized equations,

$$[M_1]\ddot{\mathbf{u}}+\frac{c}{\rho h}[C_1]\dot{\mathbf{u}}+2\Omega[G_1]\dot{\mathbf{v}}+\left(\frac{D}{\rho hb^4}[K_{e1}]+\Omega^2[K_{r1}]\right)\mathbf{u}$$
$$+\left(\dot{\Omega}[H_1]+\frac{c\Omega}{\rho h}[G_1]\right)\mathbf{v}=\mathbf{0} \tag{11}$$

$$[M_2]\ddot{\mathbf{v}}+\frac{c}{\rho h}[C_2]\dot{\mathbf{v}}+2\Omega[G_2]\dot{\mathbf{u}}+\left(\frac{D}{\rho hb^4}[K_{e2}]+\Omega^2[K_{r2}]\right)\mathbf{v}$$
$$+\left(\dot{\Omega}[H_2]+\frac{c\Omega}{\rho h}[G_2]\right)\mathbf{u}=\mathbf{0}$$

where $[M]$ and $[K_e]$ are the mass and elastic stiffness matrices, respectively; $[C]$ and $[G]$ are the damping and gyroscopic matrices, respectively; $[K_r]$ and $[H]$ are geometric stiffness matrices due to the centripetal acceleration and the angular acceleration, respectively; $\mathbf{u}$ and $\mathbf{v}$ are vectors formed by all nodal parameter vectors $\mathbf{u}_j^e$ and $\mathbf{v}_j^e$, respectively, and a overdot denotes a differentiation with respect to time t. Note that all the matrices except $[K_e]$ are of block-diagonal forms.

Equation (11) is a set of coupled ordinary differential equations with variable coefficients. If the spin rate of the disk can be expressed as the sum of a constant speed Ω_o and a small periodic perturbation $\Omega_1(t)$, i.e., $\Omega=\Omega_o+\Omega_1(t)$, Eq. (11) can be combined together and rewritten in the nondimensional form as

$$[M]\mathbf{p}''+2\mu\Omega_o(\alpha[C]+[G])\mathbf{p}'+\{[K_e]+\mu^2\Omega_o^2([K_r]$$
$$+2\alpha[G])\}\mathbf{p}=-2\mu\Omega_1[G]\mathbf{p}'-\{\mu^2(2\Omega_o\Omega_1+\Omega_1^2)[K_r]$$
$$+\mu\Omega_1'[H]+2\alpha\mu^2\Omega_o\Omega_1[G]\}\mathbf{p} \tag{12}$$

where $\alpha=c/2\rho h\Omega_o$, $\mu=\sqrt{\rho hb^4/D}$, $\mathbf{p}=\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\end{bmatrix}$,

$$[M]=\begin{bmatrix}[M_1] & [0]\\ [0] & [M_2]\end{bmatrix} \quad [C]=\begin{bmatrix}[C_1] & [0]\\ [0] & [C_2]\end{bmatrix}$$
$$[G]=\begin{bmatrix}[0] & [G_1]\\ [G_2] & [0]\end{bmatrix},$$
$$[K_e]=\begin{bmatrix}[K_{e1}] & [0]\\ [0] & [K_{e2}]\end{bmatrix} \quad [K_r]=\begin{bmatrix}[K_{r1}] & [0]\\ [0] & [K_{r2}]\end{bmatrix}$$
$$[H]=\begin{bmatrix}[0] & [H_1]\\ [H_2] & [0]\end{bmatrix},$$

and a prime denotes a differentiation with respect to the dimensionless temporal variable $\tau=\mathrm{t}\sqrt{\mathrm{D}/\rho\mathrm{hb}^4}$. Note that the matrices $[M]$, $[C]$, and $[K_r]$ are symmetric, while $[G]$, $[H]$, and $[K_e]$ are asymmetric. To improve the solvability of Eq. (12), a modal analysis is then applied to uncouple the undamped, autonomous terms in the system equation. Since the matrix $[K_e]$ is asymmetric in this case, rewrite Eq. (12) into the following set of the first-order differential equations,

$$\begin{bmatrix}[M] & [0]\\ [0] & [I]\end{bmatrix}\mathbf{q}'+\begin{bmatrix}2\mu\Omega_o[G] & [K_t]\\ -[I] & [0]\end{bmatrix}\mathbf{q}$$
$$=-2\alpha\begin{bmatrix}\mu\Omega_o[C] & [0]\\ [0] & [0]\end{bmatrix}\mathbf{q}-2\alpha\left(1+\frac{\Omega_1}{\Omega_o}\right)$$
$$\times\begin{bmatrix}[0] & \mu^2\Omega_o^2[G]\\ [0] & [0]\end{bmatrix}\mathbf{q}$$
$$-2\frac{\Omega_1}{\Omega_o}\begin{bmatrix}\mu\Omega_o[G] & [0]\\ [0] & [0]\end{bmatrix}\mathbf{q}-\frac{\Omega_1'}{\Omega_o}\begin{bmatrix}[0] & \mu\Omega_o[H]\\ [0] & [0]\end{bmatrix}\mathbf{q}$$
$$-\left(2\frac{\Omega_1}{\Omega_o}+\frac{\Omega_1^2}{\Omega_o^2}\right)\begin{bmatrix}[0] & \mu^2\Omega_o^2[K_r]\\ [0] & [0]\end{bmatrix}\mathbf{q} \tag{13}$$

where $[K_t]=[K_e]+\mu^2\Omega_o^2[K_r]$; $[I]$ is an identity matrix, and $\mathbf{q}=\begin{bmatrix}\mathbf{p}'\\ \mathbf{p}\end{bmatrix}$.

The eigenvalues of the corresponding undamped, autonomous system in Eq. (13) appear in complex conjugate pairs, i.e., $\lambda_n=\gamma_n\pm i\omega_n$, $n=1,2,\ldots,N$, where N is the total degrees-of-

freedom of the discretized system defined by Eq. (12), and ω_n are the nondimensional natural frequencies of the disk with a constant spin rate Ω_o. When the natural mode is stable, the real part of the corresponding eigenvalue γ_n is zero. If γ_n becomes positive, the natural mode corresponding to this eigenvalue is unstable. If, in addition, the imaginary part of this eigenvalue ω_n is equal to 0, the mode experiences the so-called divergence-type instability. If the imaginary parts of two eigenvalues λ_m and λ_n which have positive real parts are equal, i.e., $\omega_m=\omega_n$, these two modes experience the so-called flutter-type instability. The normalized right and left eigenvectors of the corresponding undamped, autonomous system in Eq. (13) also appear in complex conjugate pairs, i.e.,

$$\mathbf{x}_{2n-1}=\mathbf{a}_n+i\mathbf{b}_n, \quad \mathbf{x}_{2n}=\mathbf{a}_n-i\mathbf{b}_n$$

$$\mathbf{y}_{2n-1}=\mathbf{d}_n+i\mathbf{e}_n, \quad \mathbf{y}_{2n}=\mathbf{d}_n-i\mathbf{e}_n, \quad n=1,2,\ldots,N$$

where $\mathbf{x}_j$ and $\mathbf{y}_j$ are normalized right and left eigenvectors, respectively; $\mathbf{a}_n$ and $\mathbf{d}_n$ are real parts of $\mathbf{x}_{2n-1}$ and $\mathbf{y}_{2n-1}$, respectively; $\mathbf{b}_n$ and $\mathbf{e}_n$ are imaginary parts of $\mathbf{x}_{2n-1}$ and $\mathbf{y}_{2n-1}$, respectively. Due to the biorthogonality of the right and left eigenvectors, $\mathbf{a}_n$, $\mathbf{b}_n$, $\mathbf{d}_n$ and $\mathbf{e}_n$ have the following properties ([19]),

$$\begin{gathered}
\mathbf{d}_j^T\begin{bmatrix}[M] & [0]\\ [0] & [\mathrm{I}]\end{bmatrix}\mathbf{a}_k=-\mathbf{e}_j^T\begin{bmatrix}[M] & [0]\\ [0] & [I]\end{bmatrix}\mathbf{b}_k=\frac{1}{2}\delta_{jk}\\
\mathbf{e}_j^T\begin{bmatrix}[M] & [0]\\ [0] & [\mathrm{I}]\end{bmatrix}\mathbf{a}_k=-\mathbf{d}_j^T\begin{bmatrix}[M] & [0]\\ [0] & [I]\end{bmatrix}\mathbf{b}_k=0\\
\mathbf{d}_j^T\begin{bmatrix}2\mu\Omega_o[G] & [K_t]\\ -[I] & [0]\end{bmatrix}\mathbf{a}_k=-\mathbf{e}_j^T\begin{bmatrix}2\mu\Omega_o[G] & [K_t]\\ -[I] & [0]\end{bmatrix}\mathbf{b}_k=-\frac{1}{2}\gamma_j\delta_{jk}\\
\mathbf{e}_j^T\begin{bmatrix}2\mu\Omega_o[G] & [K_t]\\ -[I] & [0]\end{bmatrix}\mathbf{a}_k=-\mathbf{d}_j^T\begin{bmatrix}2\mu\Omega_o[G] & [K_t]\\ -[I] & [0]\end{bmatrix}\mathbf{b}_k=-\frac{1}{2}\omega_j\delta_{jk}
\end{gathered} \tag{14}$$

where δ_{jk} is the Kronecker delta function.

Consider the case that the disk is stable when the spin rate is constant, i.e., all the real parts of the eigenvalues γ_n are zero. Introduce now a linear transformation $\mathbf{q}=[\mathbf{a}_1,\mathbf{b}_1,\ldots,\mathbf{a}_N,\mathbf{b}_N]\boldsymbol{\zeta}$. Substituting this transformation into Eq. (13), premultiplying the matrix $[\mathbf{d}_1,-\mathbf{e}_1,\ldots,\mathbf{d}_N,-\mathbf{e}_N]^T$ and using the orthogonality in Eq. (14) yields the following partially uncoupled equation:

$$\begin{aligned}
\boldsymbol{\zeta}'+[\Lambda]\boldsymbol{\zeta}=-2\Bigg\{&2\alpha[C^*]+2\alpha\left(1+\frac{\Omega_1}{\Omega_o}\right)[S^*]+2\frac{\Omega_1}{\Omega_o}[G^*]\\
&+\frac{\Omega_1'}{\Omega_o}[H^*]+\left(2\frac{\Omega_1}{\Omega_o}+\frac{\Omega_1^2}{\Omega_o^2}\right)[K^*]\Bigg\}\boldsymbol{\zeta}
\end{aligned} \tag{15}$$

where $[\Lambda]=\text{block}-\text{diag.}\left[\begin{smallmatrix}0 & -\omega_n\\ \omega_n & 0\end{smallmatrix}\right]$

$$\begin{aligned}
[C^*]&=[\mathbf{d}_1,-\mathbf{e}_1,\ldots,\mathbf{d}_N,-\mathbf{e}_N]^T\begin{bmatrix}\mu\Omega_o[C] & [0]\\ [0] & [0]\end{bmatrix}\\
&\quad\times[\mathbf{a}_1,\mathbf{b}_1,\ldots,\mathbf{a}_N,\mathbf{b}_N]\\
[S^*]&=[\mathbf{d}_1,-\mathbf{e}_1,\ldots,\mathbf{d}_N,-\mathbf{e}_N]^T\begin{bmatrix}[0] & \mu^2\Omega_o^2[G]\\ [0] & [0]\end{bmatrix}\\
&\quad\times[\mathbf{a}_1,\mathbf{b}_1,\ldots,\mathbf{a}_N,\mathbf{b}_N]\\
[G^*]&=[\mathbf{d}_1,-\mathbf{e}_1,\ldots,\mathbf{d}_N,-\mathbf{e}_N]^T\begin{bmatrix}\mu\Omega_o[G] & [0]\\ [0] & [0]\end{bmatrix}\\
&\quad\times[\mathbf{a}_1,\mathbf{b}_1,\ldots,\mathbf{a}_N,\mathbf{b}_N]\\
[H^*]&=[\mathbf{d}_1,-\mathbf{e}_1,\ldots,\mathbf{d}_N,-\mathbf{e}_N]^T\begin{bmatrix}[0] & \mu\Omega_o[H]\\ [0] & [0]\end{bmatrix}\\
&\quad\times[\mathbf{a}_1,\mathbf{b}_1,\ldots,\mathbf{a}_N,\mathbf{b}_N]\\
[K^*]&=[\mathbf{d}_1,-\mathbf{e}_1,\ldots,\mathbf{d}_N,-\mathbf{e}_N]^T\begin{bmatrix}[0] & \mu^2\Omega_o^2[K_r]\\ [0] & [0]\end{bmatrix}\\
&\quad\times[\mathbf{a}_1,\mathbf{b}_1,\ldots,\mathbf{a}_N,\mathbf{b}_N].
\end{aligned}$$

The terms on the left-hand side of Eq. (15) are uncoupled in a blockwise sense; however, those on the right-hand side of the equation are still coupled together. To match the form of the matrix $[\Lambda]$, the matrices on the right-hand side are partitioned into N^2 blocks of 2×2 matrices. Consequently, Eq. (15) can be rewritten into the following form:

$$\begin{aligned}
\xi_n'-\omega_n\eta_n=&-4\alpha\sum_{r=1}^{N}(c_{nr}^{11}\xi_r+c_{nr}^{12}\eta_r)-4\alpha\left(1+\frac{\Omega_1}{\Omega_o}\right)\\
&\times\sum_{r=1}^{N}(s_{nr}^{11}\xi_r+s_{nr}^{12}\eta_r)-4\frac{\Omega_1}{\Omega_o}\sum_{r=1}^{N}(g_{nr}^{11}\xi_r+g_{nr}^{12}\eta_r)\\
&-2\frac{\Omega_1'}{\Omega_o}\sum_{r=1}^{N}(h_{nr}^{11}\xi_r+h_{nr}^{12}\eta_r)\\
&-2\left(2\frac{\Omega_1}{\Omega_o}+\frac{\Omega_1^2}{\Omega_o^2}\right)\sum_{r=1}^{N}(k_{nr}^{11}\xi_r+k_{nr}^{12}\eta_r)\\
\eta_n'+\omega_n\xi_n=&-4\alpha\sum_{r=1}^{N}(c_{nr}^{21}\xi_r+c_{nr}^{22}\eta_r)-4\alpha\left(1+\frac{\Omega_1}{\Omega_o}\right)\\
&\times\sum_{r=1}^{N}(s_{nr}^{21}\xi_r+s_{nr}^{22}\eta_r)-4\frac{\Omega_1}{\Omega_o}\sum_{r=1}^{N}(g_{nr}^{21}\xi_r+g_{nr}^{22}\eta_r)\\
&-2\frac{\Omega_1'}{\Omega_o}\sum_{r=1}^{N}(h_{nr}^{21}\xi_r+h_{nr}^{22}\eta_r)\\
&-2\left(2\frac{\Omega_1}{\Omega_o}+\frac{\Omega_1^2}{\Omega_o^2}\right)\sum_{r=1}^{N}(k_{nr}^{21}\xi_r+k_{nr}^{22}\eta_r)\\
&n=1,2,\ldots,N
\end{aligned} \tag{16}$$

where ξ_n and η_n are the $(2n-1)$th and $2n$th entries of $\boldsymbol{\zeta}$; c_{nr}^{ij}, s_{nr}^{ij}, g_{nr}^{ij}, h_{nr}^{ij}, and k_{nr}^{ij} are the $i-j$th entries of the $n-r$th blocks of $[C^*]$, $[S^*]$, $[G^*]$, $[H^*]$, and $[K^*]$, respectively.

Perturbation Analysis

Equation (16) represents a set of simultaneous ordinary differential equations and has no exact solutions. However, if the right-hand side of the equation is small in some senses, perturbation methods can be utilized to obtain analytical solutions. In this work, the angular speed perturbation Ω_1 is assumed to be periodic. Hence, it can be expanded into a Fourier series of the form

$$\Omega_1(\tau)=\sum_{m=-M}^{M}\kappa_m e^{im\beta\tau} \tag{17}$$

where β is called the perturbation frequency. Since the speed perturbation Ω_1 is small as compared with the steady speed Ω_o, the magnitude of each Fourier coefficient κ_m is also small as compared with Ω_o. Therefore, the small parameter ϵ appearing in the perturbation technique is defined as $\epsilon=|\kappa_M|/\Omega_o$ here.

In this work, the method of multiple scales is used to find an analytical solution of Eq. (16). By introducing new independent variables $T_j=\epsilon^j\tau$, $j=0,1,2,\ldots$, it follows that the derivatives with respect to τ become expansions in terms of the partial derivatives with respect to T_j, [20], i.e., $d/d\tau=D_0+\epsilon D_1+\ldots$. It is assumed that the solution of Eq. (16) can be represented by a uniformly valid expansion having the form

$$\xi_n(\tau,\epsilon)=\xi_{n0}(T_0,T_1,\ldots)+\epsilon\xi_{n1}(T_0,T_1,\ldots)+\ldots \tag{18}$$

$$\eta_n(\tau,\epsilon)=\eta_{n0}(T_0,T_1,\ldots)+\epsilon\eta_{n1}(T_0,T_1,\ldots)+\ldots.$$

Due to the complexity of the problem, the expansion is carried out to the order of ϵ. Substituting Eq. (18) into Eq. (16) and equating the coefficients of like powers of ϵ yields the following equations:

Order 1

$$\begin{aligned} D_0\xi_{n0}-\omega_n\eta_{n0}&=0\\ D_0\eta_{n0}+\omega_n\xi_{n0}&=0 \end{aligned} \tag{19}$$

Order ϵ

$$\begin{aligned} &D_0\xi_{n1}-\omega_n\eta_{n1}\\ &\quad=-D_1\xi_{n0}-4\hat{\alpha}\sum_{r=1}^{N}(c_{nr}^{11}\xi_r+c_{nr}^{12}\eta_r)\\ &\qquad-4\hat{\alpha}\sum_{r=1}^{N}(s_{nr}^{11}\xi_r+s_{nr}^{12}\eta_r)\\ &\qquad-4\sum_{m=1}^{M}\hat{\kappa}_m e^{i\beta_m\tau}\cdot\sum_{r=1}^{N}(g_{nr}^{11}\xi_r+g_{nr}^{12}\eta_r)\\ &\qquad-2\sum_{m=1}^{M}i\beta_m\hat{\kappa}_m e^{i\beta_m\tau}\cdot\sum_{r=1}^{N}(h_{nr}^{11}\xi_r+h_{nr}^{12}\eta_r)\\ &\qquad-4\sum_{m=1}^{M}\hat{\kappa}_m e^{i\beta_m\tau}\cdot\sum_{r=1}^{N}(k_{nr}^{11}\xi_r+k_{nr}^{12}\eta_r)+\text{c.c.}\\ &D_0\eta_{n1}+\omega_n\xi_{n1}\\ &\quad=-D_1\eta_{n0}-4\hat{\alpha}\sum_{r=1}^{N}(c_{nr}^{21}\xi_r+c_{nr}^{22}\eta_r)\\ &\qquad-4\hat{\alpha}\sum_{r=1}^{N}(s_{nr}^{21}\xi_r+s_{nr}^{22}\eta_r)\\ &\qquad-4\sum_{m=1}^{M}\hat{\kappa}_m e^{i\beta_m\tau}\cdot\sum_{r=1}^{N}(g_{nr}^{21}\xi_r+g_{nr}^{22}\eta_r)\\ &\qquad-2\sum_{m=1}^{M}i\beta_m\hat{\kappa}_m e^{i\beta_m\tau}\cdot\sum_{r=1}^{N}(h_{nr}^{21}\xi_r+h_{nr}^{22}\eta_r)\\ &\qquad-4\sum_{m=1}^{M}\hat{\kappa}_m e^{i\beta_m\tau}\cdot\sum_{r=1}^{N}(k_{nr}^{21}\xi_r+k_{nr}^{22}\eta_r)+\text{c.c.} \end{aligned} \tag{20}$$

where $\hat{\kappa}_m=\kappa_m/|\kappa_M|$, $\beta_m=m\beta$, and c.c. denotes the complex conjugate of the preceding terms. In the above equation, $\kappa_0=0$ is assumed without loss of generality, and $\alpha=\epsilon\hat{\alpha}$ is assumed to have the damping term appearing in the same order as the first excitation term. To this order of approximation, four cases of resonant combinations of frequencies have to be considered.

I The Case of β_m Away From $\omega_p\pm\omega_q$. When the requency of the excitation β_m is away from $\omega_p\pm\omega_q$, one speaks of a nonresonant case. In this case, the secular terms are eliminated from Eq. (20) if the amplitude of ξ_{n0} has the form,

$$A_n(T_1)=a_n e^{-2\hat{\alpha}c_n T_1} \tag{21}$$

where a_n is an arbitrary function of T_2, and $c_n=c_{nn}^{11}+c_{nn}^{22}+s_{nn}^{11}+s_{nn}^{22}$. Note that the above equation is obtained as a result of $c_{nn}^{12}=c_{nn}^{21}$ and $s_{nn}^{12}=s_{nn}^{21}$. The former is due to the symmetry of the matrix $[C^*]$, and the latter is observed during the course of numerical computation. The amplitude A_n will decay with time, and hence the system is always stable if $c_n>0$. In the course of numerical computation, it is interesting to find that c_n is always equal to 0.5 unless the system configuration falls within divergence–or flutter-type instability zones, where c_n is negative. Therefore, unless the system is experiencing divergence–or flutter-type instability, the system is stable when the excitation frequency is away from the sum or difference of any two natural frequencies.

II The Case of β_m Near $\omega_p+\omega_q$. When the frequency of the excitation β_m is near $\omega_p+\omega_q$, a combination resonance of the summed-type exists between the pth and qth modes. By eliminating the secular terms from Eq. (20) yields the expressions for the transition curves which separate the stable regions from the unstable regions,

$$\beta_m=\omega_p+\omega_q\pm\epsilon\left[\left(\frac{\Lambda_{pq}\overline{\Lambda}_{qp}}{c_pc_q}-\hat{\alpha}^2\right)(c_p+c_q)^2\right]^{1/2}+O(\epsilon^2) \tag{22}$$

where

$$\begin{aligned} c_p&=c_{pp}^{11}+c_{pp}^{22}+s_{pp}^{11}+s_{pp}^{22},\quad c_q=c_{qq}^{11}+c_{qq}^{22}+s_{qq}^{11}+s_{qq}^{22},\\ \Lambda_{pq}&=[(-2ik_{pq}^{21}-2k_{pq}^{22}-2ig_{pq}^{21}-2g_{pq}^{22}+\beta_m h_{pq}^{21}-i\beta_m h_{pq}^{22})\\ &\quad-(-2k_{pq}^{11}+2ik_{pq}^{12}-2g_{pq}^{11}+2ig_{pq}^{12}-\mathrm{i}\beta_m h_{pq}^{11}-\beta_m h_{pq}^{12})]\hat{\kappa}_m\\ \overline{\Lambda}_{qp}&=[(2ik_{qp}^{21}-2k_{qp}^{22}+2ig_{qp}^{21}-2g_{qp}^{22}+\beta_m h_{qp}^{21}+i\beta_m h_{qp}^{22})\\ &\quad-(-2k_{qp}^{11}-2ik_{qp}^{12}-2g_{qp}^{11}-2ig_{qp}^{12}+i\beta_m h_{qp}^{11}-\beta_m h_{qp}^{12})]\hat{\kappa}_m. \end{aligned}$$

If there is no damping in the system, the transition curves become

$$\beta_m=\omega_p+\omega_q\pm2\epsilon\sqrt{\Lambda_{pq}\overline{\Lambda}_{qp}}+O(\epsilon^2). \tag{23}$$

In the course of numerical computation, one finds that Λ_{pq} and $\overline{\Lambda}_{qp}$ are either real or pure imaginary simultaneously. Therefore, transition curves exist if Λ_{pq} and $\overline{\Lambda}_{qp}$ have the same sign when they are real, and have opposite signs when they are pure imaginary.

III The Case of β_m Near $\omega_p-\omega_q$. When the frequency of the excitation β_m is close to $\omega_p-\omega_q$, a combination resonance of the difference-type exists between the pth and qth modes. Similarly, by eliminating the secular terms from Eq. (20) yields the expressions for the transition curves

$$\beta_m=\omega_p-\omega_q\pm\epsilon\left[\left(\frac{\Lambda_{pq}\Lambda_{qp}}{c_pc_q}-\hat{\alpha}^2\right)(c_p+c_q)^2\right]^{1/2}+O(\epsilon^2) \tag{24}$$

where c_p and c_q are defined in the same way as the previous case, while

$$\begin{aligned} \Lambda_{pq}&=[(-2ik_{pq}^{21}+2k_{pq}^{22}-2ig_{pq}^{21}+2g_{pq}^{22}+\beta_m h_{pq}^{21}+i\beta_m h_{pq}^{22})\\ &\quad-(-2k_{pq}^{11}+2ik_{pq}^{12}-2g_{pq}^{11}+2ig_{pq}^{12}-i\beta_m h_{pq}^{11}+\beta_m h_{pq}^{12})]\hat{\kappa}_m\\ \Lambda_{qp}&=[(-2ik_{qp}^{21}+2k_{qp}^{22}-2ig_{qp}^{21}+2g_{qp}^{22}-\beta_m h_{qp}^{21}-i\beta_m h_{qp}^{22})\\ &\quad-(-2k_{qp}^{11}-2ik_{qp}^{12}-2g_{qp}^{11}-2ig_{qp}^{12}+i\beta_m h_{qp}^{11}-\beta_m h_{qp}^{12})]\hat{\kappa}_m. \end{aligned}$$

Again that Λ_{pq} and Λ_{qp} are either real or pure imaginary simultaneously is observed during the course of numerical computation. Therefore, transition curves exist if Λ_{pq} and Λ_{qp} have the same sign when they are real, and have opposite signs when they are pure imaginary.

IV The Case of β_m Near $\omega_p+\omega_q$ and $\omega_q+\omega_r$. In this case, β_m is simultaneously near $\omega_p+\omega_q$ and $\omega_q+\omega_r$, and there are no other resonances to this order. To express the nearness of β_m to $\omega_p+\omega_q$ and $\omega_q+\omega_r$, one introduces two detuning parameters σ_1 and σ_2 defined by

$$\beta_m=\omega_p+\omega_q+\epsilon\sigma_1,\quad \beta_m=\omega_q+\omega_r+\epsilon\sigma_2. \tag{25}$$

It follows that the secular terms in Eq. (20) are eliminated if

$$\begin{aligned} D_1A_p+\hat{\alpha}c_pA_p+\Lambda_{pq}\overline{A}_q e^{i\sigma_1T_1}&=0\\ D_1A_q+\hat{\alpha}c_qA_q+\Lambda_{qp}\overline{A}_p e^{i\sigma_1T_1}+\Lambda_{qr}\overline{A}_r e^{i\sigma_2T_1}&=0\\ D_1A_r+\hat{\alpha}c_rA_r+\Lambda_{rq}\overline{A}_q e^{i\sigma_2T_1}&=0 \end{aligned} \tag{26}$$

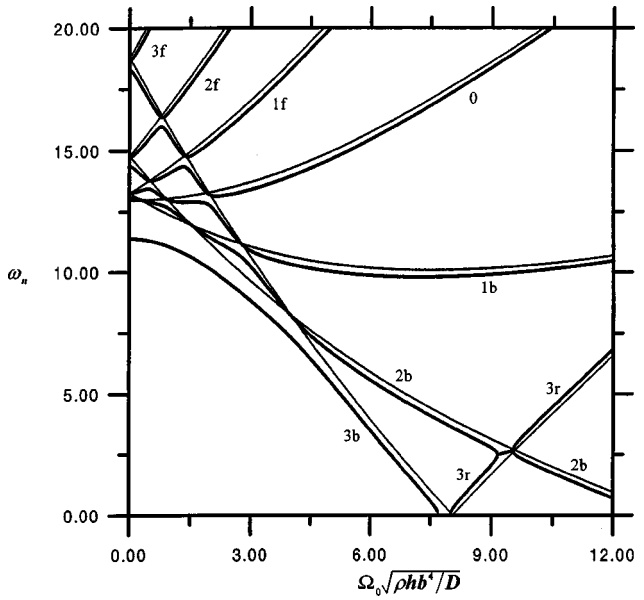


Fig. 2 The natural frequencies of the zero-nodal-circle modes of a spinning disk with a constant spin rate subjected to a conservative, concentrated and compressive edge load. $a/b=0.5$, $\nu=0.27$. Heavy line: $F_c b^2/D=3.5$, light line: $F_c b^2/D=0$.

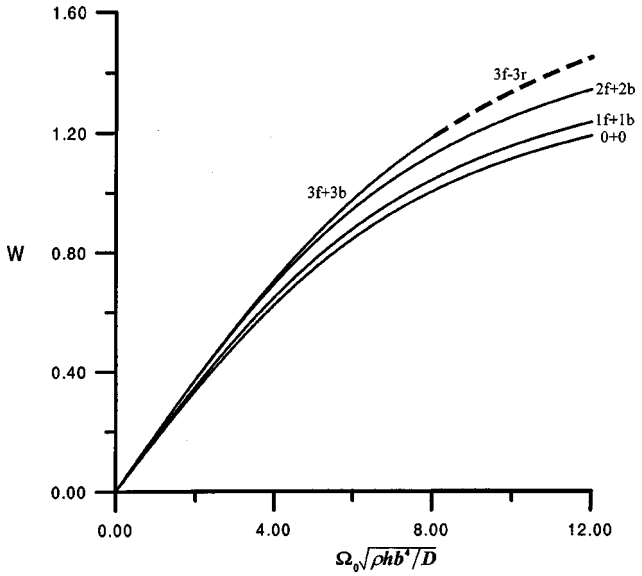


Fig. 3 The width parameters of unstable regions of a freely spinning disk with a harmonically varying spin rate. $a/b=0.5$, $\nu=0.27$, $\alpha=0$, $F_c b^2/D=0$.

where A_p, A_q, and A_r are the amplitudes of ξ_{p0}, ξ_{q0}, and ξ_{r0}, respectively. Λ_{pq} and Λ_{qp} are the same as those defined in Eq. (22), and Λ_{qr} and Λ_{rq} are defined in a similar way. The above equation admits solutions having the form

$$A_p=a_p e^{i(\bar{\lambda}+\sigma_1)T_1},\quad A_q=a_q e^{-i\lambda T_1},\quad A_r=a_r e^{i(\bar{\lambda}+\sigma_2)T_1} \tag{27}$$

where a_p, a_q, and a_r are complex functions of T_2, and λ is a complex number. Substituting Eq. (27) into Eq. (26) yields a set of algebraic equations in a_p, a_q, and a_r. To have nontrivial solutions of a_p, $\bar{a}_q$, and a_r, the determinant of the coefficient matrix must vanish, and λ must be real also to have stable response of the system. This results in a cubic equation in λ,

$$\lambda^3+\hat{p}(\sigma_1,\sigma_2,\epsilon)\lambda^2+\hat{q}(\sigma_1,\sigma_2,\epsilon)\lambda+\hat{r}(\sigma_1,\sigma_2,\epsilon)=0, \tag{28}$$

where the coefficients $\hat{p}$, $\hat{q}$, and $\hat{r}$ are functions of σ_1, σ_2, and ϵ. To guarantee all three roots being real, the coefficients of the cubic equation, so are σ_1, σ_2, and ϵ, must satisfy the following relationship:

$$\frac{1}{4}(2\hat{p}^3-9\hat{p}\hat{q}+27\hat{r})^2+(3\hat{q}-\hat{p}^2)^3\leqslant 0. \tag{29}$$

From the above relationship, the transition curves can be obtained. The transition curves correspond to the values of β_m for which λ has at least two real, equal roots.

Numerical Results and Discussions

Before presenting the numerical results for the stability analysis, verification of the finite element formulation is taken by considering the free vibration problem of a spinning disk with a constant spin rate subjected to a stationary concentrated edge load in the radial direction. According to Chen's study [10], the maximum number of nodal diameters J in Eq. (9) and the number of elements in the radial direction are chosen to be 3 and 6, respectively. Figure 2 shows the natural frequencies of the zero-nodal-circle modes of a spinning disk with a constant spin rate subjected to a conservative, concentrated and compressive edge load F_c. There are two sets of curves in this figure. The light curves are the results for the freely spinning disk, while the heavy curves are the results for the disk subjected to an edge load $F_c b^2/D=3.5$. In this figure, nf, nb, and nr denote the forward, backward and reflected modes with n nodal diameters, respectively. It is observed from the figure that the compressive edge load tends to decrease the natural frequencies of the forward and backward modes but to increase the natural frequencies of the reflected modes. Furthermore, the conservative, compressive edge load induces a divergence-type instability (an unstable mode with a zero natural frequency) before the critical speed, the speed at which a natural frequency of a freely spinning disk becomes zero, and a flutter-type instability (two unstable modes having the same natural frequency) beyond the critical speed when the $3r$ reflected mode meets the $2b$ backward mode. At the intersection point between two light curves, i.e., the degenerate modes of a freely spinning disk, the natural frequencies will be separated upon the application of the edge load. One of the natural frequencies remains unchanged, while the other lowers. Note that the results in Fig. 2 are identical to those obtained previously by Chen [10], which proves that the finite element formulation in this work is correct. The natural frequencies of the one-nodal-circle modes are several times higher than those of the zero-nodal-circle modes. Since the lowest few modes are relatively important in vibration, we will focus on the zero-nodal-circle modes hereafter.

As an example of application of the general solutions obtained in this work, the speed perturbation $\Omega_1(\tau)$ is taken as $\Omega_1(\tau)=\kappa\cos\beta\tau$, where κ is assumed to be small as compared with the average spin rate Ω_o, and β is the nondimensionalized perturbation frequency. Hence, the small parameter ϵ appearing in the perturbation technique is defined as $\epsilon=\kappa/\Omega_o$. Figure 3 presents the width parameters W of unstable regions of an undamped, freely spinning disk with a harmonically varying spin rate. The width parameter W is defined as $\sqrt{\Lambda_{pq}\Lambda_{qp}}$ for sum-type resonances or $\sqrt{\Lambda_{pq}\Lambda_{qp}}$ for difference-type resonances. In this figure, solid curves represent sum-type resonances, while a dash curve represents a difference-type resonance. It is found that combination resonances will not take place between modes having different nodal diameters. This can be explained as follows: Recall that all the matrices except $[K_e]$ in Eq. (11) have block-diagonal forms, and coupling is caused by the initial stresses induced by the stationary edge load. For a freely spinning disk, the equations for modes having different nodal diameters are mutually independent. Consequently, combination resonances will not occur between

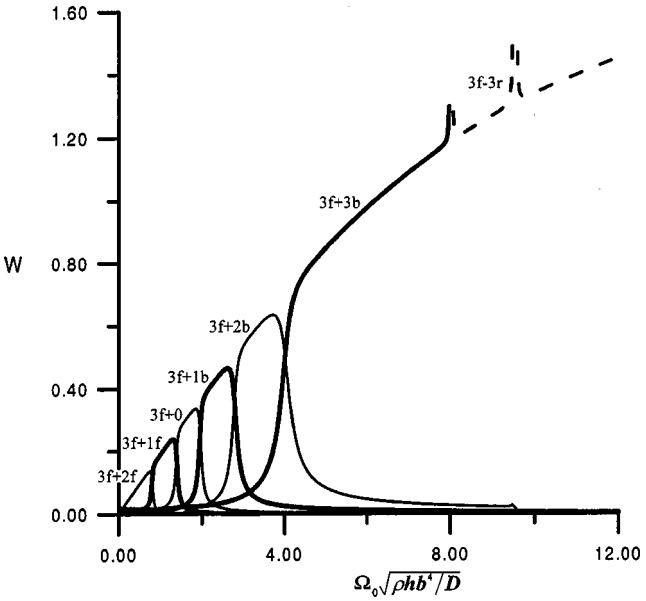


Fig. 4 The width parameters of specific unstable regions of a spinning disk with a harmonically varying spin rate subjected to a conservative, concentrated and compressive edge load. $a/b=0.5$, $\nu=0.27$, $\alpha=0$, $F_c b^2/D=1.0$.

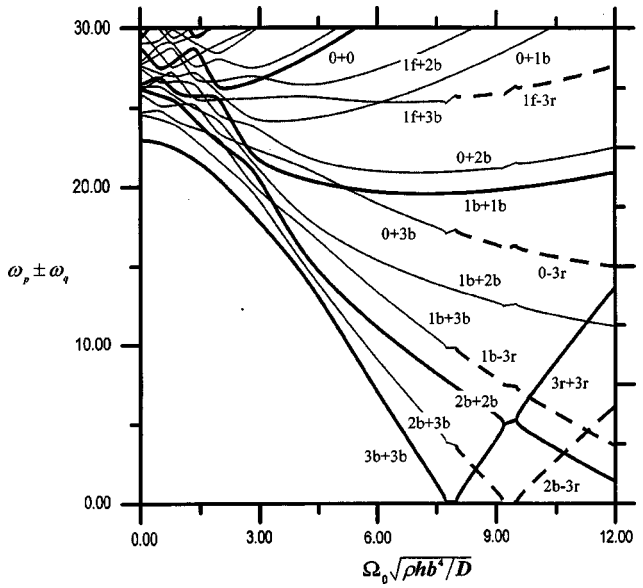


Fig. 5 The central frequencies of unstable regions of a spinning disk with a harmonically varying spin rate subjected to a conservative, concentrated and compressive edge load. $a/b=0.5$, $\nu=0.27$, $\alpha=0$, $F_c b^2/D=3.5$.

modes with different nodal diameters in the absence of the stationary edge load. In addition, sum-type resonances occur between forward and backward modes, while difference-type resonances occur between forward and reflected modes. The width parameters of sum-type and difference-type resonances rise with increasing spin rates, but the width parameters of main resonances ($\beta=2\omega_n$) are always 0.

Figure 4 depicts the width parameters W of some specific unstable regions of an undamped, spinning disk with a harmonically varying spin rate subjected to a conservative, concentrated and compressive edge load. Again six solid curves represent the sum-type resonances between the $3f$ mode and the $2f$, $1f$, 0, $1b$, $2b$, and $3b$ modes, respectively, while the dash curve represents the difference-type resonance between the $3f$ mode and $3r$ mode. In this figure, there are two protrusions, where the width parameter tends to infinity. The left one corresponds to a divergence-type instability, and the right one corresponds to a flutter-type instability. Compared with Fig. 3, one discovers that the envelope of these seven curves is closely related to but somewhat higher than the $(3f+3b)$ and $(3f-3r)$ curve in Fig. 3. This is attributed to the fact that the $3b$ mode of the freely spinning disk resolves into parts of the $2f$, $1f$, 0, $1b$, $2b$, and $3b$ modes of the disk subjected to the edge load, as shown in Fig. 2, and the disk is destabilizing when subjected to compressive edge loads. Similar phenomena can also be observed between the $2f$ mode and the $1b$, $2b$, and $3b$ modes, the $1f$ mode and the 0, $1b$, and $2b$ modes, and the 0 mode and the 0, $1b$, and $2b$ modes. The maximum width parameters of all other resonances are relatively small.

The central frequencies of unstable regions of a spinning disk with a harmonically varying spin rate subjected to a conservative, concentrated and compressive edge load are illustrated in Fig. 5. In this figure, heavy solid curves, light solid curves and dash curves denote main resonances, sum-type resonances and difference-type resonances, respectively. The figure shows that sum-type resonances occur between two modes being both reflected or both nonreflected modes, while difference-type resonances take place between one reflected and one nonreflected modes. This phenomenon was also observed previously by Chen [11] when he studied the problem of a spinning disk under space-fixed pulsating edge loads. It is evident that whether sum-type or difference-type resonances will come about is decided by the characteristics of each mode of the disk but not by the characteristics of the excitation.

Figure 6 presents the lowest few unstable regions of a spinning disk with a harmonically varying spin rate subjected to a conservative, uniformly distributed and compressive edge load of intensity $f_c b^3/D=3.5$. In this figure, the average spin rate Ω_o is equal to half of the lowest natural frequency of the corresponding free, nonrotating disk. According to Fig. 4, the maximum width parameters of the unstable regions shown in this figure are relatively small as compared with those of the sum-type resonances between the $3f$ mode and the $3b$, $2b$, $1b$ modes, etc. Fortunately, those unstable regions with larger maximum width parameters belong to the higher frequency domain. Looking into (*a*) through (*c*), one finds that the unstable regions, which lie within a pair of stability boundaries, enlarge initially as the load distribution angle increases and reach maximum as the edge load distributed over half a circle. A further increase in the load distribution angle will generally reduce their size. However, the unstable regions corresponding to the sum-type resonances between the $3f$ mode and the $3b$, $2b$, $1b$ modes, etc., will still enlarge for a further increase in the load distribution angle. The reason is that the envelope of the width parameter curves of these resonances will rise once the load distribution widens because the total edge load applied increases also. As the edge load distributed over the whole circle, the natural frequency curves look like the light curves shown in Fig. 2 for the freely spinning disk, and the unstable regions shown in this figure will disappear, leaving only unstable regions of combination resonances between modes having the same nodal diameter, like the case of the freely spinning disk. In addition, all unstable regions shift toward the lower frequency domain as the load distribution widens.

Conclusions

Dynamic stability of a disk with a periodically varying spin rate subjected to a stationary in-plane edge load has been studied both analytically and numerically. Due to the complexity of the problem, only the first order approximation was presented, and a simple perturbation function was considered to provide numerical

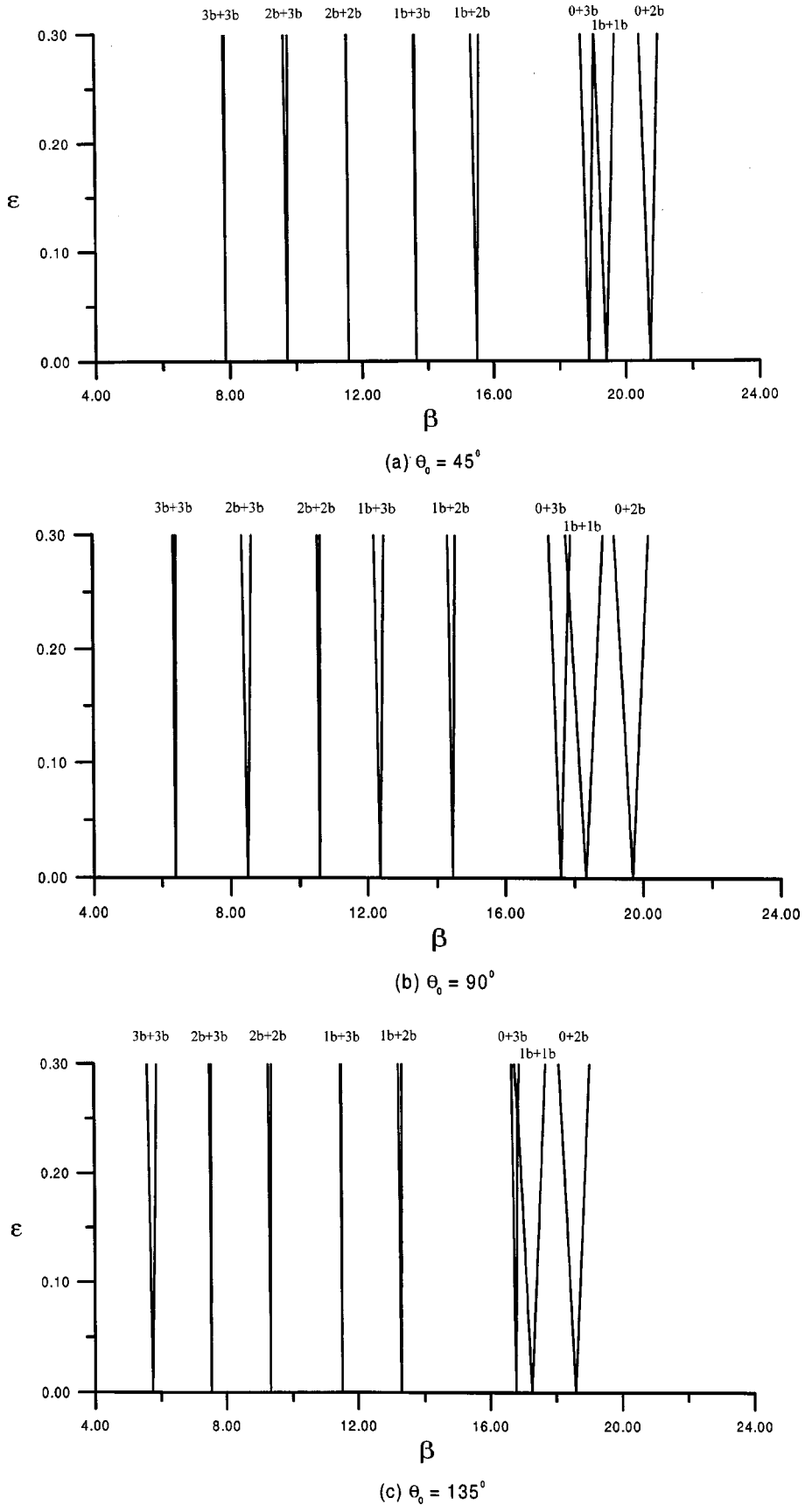


Fig. 6 Effects of the distribution angle of a conservative, uniformly distributed and compressive edge load on the stability boundaries of a spinning disk with a harmonically varying spin rate. $a/b=0.5$, $f_c b^3/D=3.5$, $\Omega_o\sqrt{\rho h b^4/D}=5.6372$.

illustrations of the general solutions. However, solutions for more general periodic functions can be easily generated for the first few approximations.

From the above numerical results, the following conclusions can be drawn:

1. For a freely spinning disk with a harmonically varying spin rate, combination resonances will not take place between modes having different nodal diameters; moreover, sum-type resonances occur between forward and backward modes, while difference-type resonances occur between forward and reflected modes.
2. When the stationary in-plane edge load is applied, combination resonances may occur between modes having different nodal diameters; sum-type resonances come about between two modes being both reflected or both nonreflected, while difference-type resonances come about between one reflected and one nonreflected modes.
3. The envelope of the width parameter curves of combination resonances between the forward mode and the modes whose portions constitute the corresponding backward mode of the freely spinning disk is closely related to the width parameter curve of the combination resonance between the modes of the freely spinning disk having the same nodal diameter. The maximum width parameters of all other resonances are relatively small.
4. Generally speaking, when the stationary in-plane edge load is uniformly distributed, the lowest few unstable regions, whose maximum width parameters are relatively small, tend to enlarge at first as the load distribution widens and to reach maximum as the load distributed over half a circle. A further increase in the load distribution angle tends to reduce the lowest few unstable regions.

Acknowledgment

This work was sponsored by the National Science Council of the Republic of China under Grant No. NSC 90-2212-E-011-029.

References

[1] Lamb, H., and Southwell, R. V., 1921, ''The Vibration of a Spinning Disk,'' Proc. Phys. Soc. London, **99**, pp. 272–280.
[2] Southwell, R. V., 1922, ''On the Free Transverse Vibrations of a Uniform Circular Disk Clamped at Its Center and on the Effects of Rotation,'' Proc. Phys. Soc. London, **101**, pp. 133–153.
[3] Mote, Jr., C. D., 1965, ''Free Vibrations of Initially Stressed Circular Disks,'' J. Eng. Ind., **87**, pp. 258–264.
[4] Mote, Jr., C. D., 1967, ''Natural Frequencies in Annuli With Induced Thermal Membrane Stress,'' J. Eng. Ind., **89**, pp. 611–618.
[5] Carlin, J. F., Appl, F. C., Bridwell, H. C., and Dubois, R. P., 1975, ''Effects of Tensioning on Buckling and Vibration of Circular Saw Blades,'' J. Eng. Ind., **99**, pp. 37–49.
[6] Redcliffe, C. J., and Mote, Jr., C. D., 1977, ''Stability of Stationary and Rotating Discs Under Edge Load,'' Int. J. Mech. Sci., **19**, pp. 567–574.
[7] Iwan, W. D., and Moeller, T. L., 1976, ''The Stability of a Spinning Elastic Disk With a Transverse Load System,'' ASME J. Appl. Mech., **43**, pp. 485–490.
[8] Ono, K., Chen, J.-S., and Bogy, D. B., 1991, ''Stability Analysis for the Head-Disk Interface in a Flexible Disk Drive,'' ASME J. Appl. Mech., **58**, pp. 1005–1014.
[9] Chen, J.-S., and Bogy, D. B., 1992, ''Effects of Load Parameters on the Natural Frequencies and Stability of a Flexible Spinning Disk With a Stationary Load System,'' ASME J. Appl. Mech., **59**, pp. 230–235.
[10] Chen, J.-S., 1994, ''Stability Analysis of a Spinning Elastic Disk Under a Stationary Concentrated Edge Load,'' ASME J. Appl. Mech., **61**, pp. 788–792.
[11] Chen, J.-S., 1997, ''Parametric Resonance of a Spinning Disk Under Space Fixed Pulsating Edge Loads,'' ASME J. Appl. Mech., **64**, pp. 139–143.
[12] Shen, I. Y., and Song, Y., 1996, ''Stability and Vibration of a Rotating Circular Plate Subjected to Stationary In-Plane Edge Loads,'' ASME J. Appl. Mech., **63**, pp. 121–127.
[13] Kammer, D. C., and Schlack, Jr., A. L., 1987, ''Effects of Nonconstant Spin Rate on the Vibration of a Rotating Beam,'' ASME J. Appl. Mech., **54**, pp. 305–310.
[14] Young, T. H., and Liou, G. T., 1992, ''Coriolis Effect on the Vibration of a Cantilever Plate With Time-Varying Rotating Speed,'' ASME J. Vibr. Acoust., **114**, pp. 232–241.
[15] Young, T. H., and Liou, G. T., 1993, ''Dynamic Response of Rotor-Bearing Systems With Time-Dependent Spin Rate,'' ASME J. Eng. Gas Turbines Power, **115**, pp. 239–245.
[16] Young, T. H., 1992, ''Nonlinear Transverse Vibration and Stability of Spinning Disks With Nonconstant Spin Rate,'' ASME J. Vibr. Acoust., **114**, pp. 506–513.
[17] Barber, J. R., 1992, *Elasticity*, Kluwer Academic Publishers, Boston.
[18] Reddy, J. N., 1984, *An Introduction to the Finite Element Method*, McGraw-Hill, New York.
[19] Young, T. H., Tseng, T. C., and Song, L. S., 2001, ''Dynamic Stability of Fluttered Systems Subjected to Parametric Random Excitations,'' J. Vib. Control, to appear.
[20] Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillations*, John Wiley and Sons, New York.

S. Bhattacharyya
Associate Professor

S. Mahapatra
Ph.D. Student

Department of Mathematics,
Indian Institute of Technology,
Kharagpur 721 302, India

F. T. Smith
Goldsmid Professor of Applied Mathematics,
Department of Mathematics,
University College London,
Gower Street,
London WC1E 6BT, UK

Influence of Surface Roughness on Shear Flow

The local planar flow of incompressible fluid past an obstacle of semi-circular cross section is considered, the obstacle being mounted on a long flat surface. The far-field motion is one of uniform shear. Direct numerical solutions of the Navier-Stokes equations are described over a range of Reynolds numbers. The downstream eddy length and upstream position of maximum pressure gradient are found to agree with increased Reynolds number theory; in particular the agreement for the former quantity is close for Reynolds numbers above about 50. [DOI: 10.1115/1.1767842]

Introduction

A distribution of roughness elements along an otherwise smooth wall boundary surface can significantly alter the dynamic process of the flow. It is also connected with turbulence promoters located on a surface for the sake of mass transfer enhancement. The fluid flow past a roughness or roughnesses on an otherwise smooth solid surface is of concern in several physical settings such as atmospheric boundary layer over hills, wind over water, uneven terrain and buildings in the atmospheric boundary layer. It is also of concern in external aerodynamics, turbomachinary flows and physiological flows. Flow past a surface mounted obstacle is further relevant to a variety of engineering problems such as flow past excrescences in pipes, the Gurney-flap device placed on an airfoil or blade, chips mounted on a PC-board in electronic equipment and so on.

The surface mounted blunt obstacle provokes both a separation upstream and a large eddy downstream of the obstacle. Flow separation has a considerable impact on the flow structure and has been the subject of intensive study for many years. A number of experimental and computational studies on flow past large-scale obstacle have been made experimentally as well as computationally. One striking phenomenon is the upstream appearance of significant three-dimensional separation, due to inviscid feedback of pressure, along with so-called necklace vortices which then skirt around the extreme spanwise edges or wing-tips of the obstacle before trailing downstream; near the outer boundary of the upstream separation bubble in two dimension, an adverse pressure gradient exists in the flow in the sense that there is increasing pressure in the direction of the main flow. The form of the upstream recirculating eddy and the parameters which influence the upstream separation were studied extensively by Smith and Walton [1] and later by Bhattacharyya et al. [2]. There it was shown through theoretical analysis using nonlinear reasoning and direct computation on the Navier-Stokes equations that the upstream separation distance elongates with increasing Reynolds number.

A number of experimental and computational studies of boundary layer flows (either external or internal) over surface-mounted obstacles have focused on the overall features of the flow, such as velocity, downstream reattachment length and surface pressure distribution, e.g., Durst and Loy [3], Williams and Baker [4], Chang and Sheu [5], and Giguere [6] has shown through experiment that, by placing a tiny normal flap (a so-called Gurney flap) near the trailing edge of an airfoil or blade, an increase in lift and reduction in drag could be produced. Recently Smith [7] and others have studied the advantages of a flap being buried within the trailing-edge boundary layer.

Martinuzzi and Tropea [8] studied experimentally the flow field around surface-mounted prismatic obstacle. There they observed that the flow in the recirculation region upstream of the obstacle develops a cellular structure and passes over the obstacle along preferred paths. The turbulent flow structure and distribution of the heat transfer coefficient of a wall-mounted cube in a spatially periodic matrix flow were studied experimentally by Meinders and Hanjalic [9].

In this study, we consider the steady planar flow past an obstacle so small and close to the flat surface that the surrounding fluid motion, that is the far-field flow as far as the obstacle is concerned may be viewed as uniform shear flow. The shape of the obstacle considered is semi-circular. The obstacle is mounted on a fixed surface and surface is infinitely long. The fluid occupies the semi-infinite region above the surface. A prime reason for addressing this problem is that, although Bhattacharyya et al. [2] showed good agreement with theory in the case of particular straight-sided obstacles in the shape of a normal flap or rectangular block, the theory is supposedly valid for any shape; hence the present focus on a fundamental curved shape, the semi-circle. A direct numerical simulation of the Navier-Stokes equations is undertaken below to compute the flow fields at various Reynolds numbers, following which comparisons are made with the theoretical analysis.

Formulation and Mathematical Equations

We consider a surface lying along the x-axis and supporting an obstacle in the form of a semi-circle of radius R with center at the origin (Fig. 1). The viscous incompressible fluid occupies the semi-infinite region $y>0$. A polar coordinate system is used such that $(x,y)=(r\cos\theta, r\sin\theta)$ with $y=0$ as the initial line. Introducing the stream function Ψ such that

$$v_r=\frac{1}{r}\frac{\partial\Psi}{\partial\theta}, \quad v_\theta=-\frac{\partial\Psi}{\partial r}. \tag{1}$$

The two momentum equations can be combined to form the vorticity transport equation as

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, October 7, 2001; final revision, March 8, 2004. Associate Editor: W. S. Saric. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

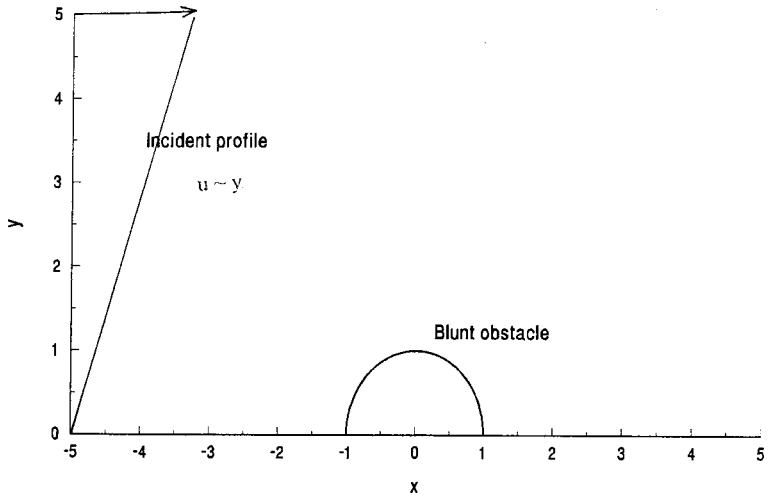


Fig. 1 Sketch of the flow configuration in nondimensional terms

$$\frac{\partial \omega}{\partial \tau}+\frac{1}{r}\left[\frac{\partial \Psi}{\partial \theta}\frac{\partial \omega}{\partial r}-\frac{\partial \Psi}{\partial r}\frac{\partial \omega}{\partial \theta}\right]=\nu\nabla^2\omega \tag{2}$$

where the vorticity ω is given by

$$\omega=-\nabla^2\Psi \tag{3}$$

and

$$\nabla^2=\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}. \tag{4}$$

These are subject to the boundary conditions

$$u=v=0$$

on the obstacle and flat plate,

$$u\sim y, \quad v=0 \quad \text{as} \quad x^2+y^2\to\infty.$$

The form of the obstacle is taken to be $x^2+y^2=R^2$. In order to achieve a more accurate numerical solution, it is essential to have a finer grid near the obstacle. This can be accomplished by use of the log-polar coordinate transformation given by

$$\frac{r}{R}=\exp(a\xi), \quad \theta=a\eta,$$

where a is the transformation parameter which is equal to π in this study. This log-polar coordinate transformation allows us to have a uniform grid in a transformed rectangular domain.

Using the shear to define the Reynolds number as here is equivalent to defining the Reynolds number based on a height typical of the roughness element height, and the shear-layer velocity at the former height, as the normalizing length and velocity, respectively. The nondimensional variables are defined as

$$\psi=\frac{\Psi}{RU}, \quad \zeta=\frac{\omega R}{U}, \quad t=\frac{\tau U}{R}, \quad \mathrm{Re}=\frac{RU}{\nu}.$$

The vorticity transport equation in its dimensionless form then reads

$$g(\xi)\frac{\partial \zeta}{\partial t}+\frac{\partial \psi}{\partial \eta}\frac{\partial \zeta}{\partial \xi}-\frac{\partial \psi}{\partial \xi}\frac{\partial \zeta}{\partial \eta}=\frac{1}{\mathrm{Re}}\left(\frac{\partial^2\zeta}{\partial \xi^2}+\frac{\partial^2\zeta}{\partial \eta^2}\right) \tag{5}$$

$$g(\xi)\zeta=-\left(\frac{\partial^2\psi}{\partial \xi^2}+\frac{\partial^2\psi}{\partial \eta^2}\right) \tag{6}$$

where $g(\xi)=a^2\exp(2a\xi)$.

The boundary conditions are given by

$$u=v=0 \quad \text{at} \quad \xi=0, \quad 0\leqslant\eta\leqslant 1,$$

on the obstacle

$$u=v=0 \quad \text{at} \quad \eta=0 \quad \text{and} \quad \eta=1, \quad 0<\xi<\infty$$

$$u\sim\exp(a\xi)\sin(a\eta), \quad v=0 \quad \text{as} \quad \xi\to\infty.$$

We define the surface pressure coefficient as $C_p=(p-p_\infty)/\rho U^2$, where p_∞ is the pressure at the far-field and ρ is the density of the fluid.

Computational Method

The vorticity transport Eq. (5) and the Poisson equation for the stream function (6) are solved using an upwind spatial difference scheme and the ADI scheme for the time derivative discretization.

Flows that are dominated by convection as here suffer form numerical instability at high values of the Reynolds number. To suppress such instability upwind differencing in the convective terms can be employed. The artificial dissipation introduced by the upwind difference scheme stabilizes the numerical solution at high values of the Reynolds numbers. The following quasilinearization is used for the convective terms:

$$[u\zeta_x]^{n+1}=u^n[\zeta_x]^{n+1}, \tag{7}$$

with the superscript n represents the time step. The spatial derivatives in (7) are approximated as

$$u^n_{i,j}[\zeta_x]^{n+1}=u^n_{i,j}\{\zeta^{n+1}_{i+2,j}-2\zeta^{n+1}_{i+1,j}+9\zeta^{n+1}_{i,j}-10\zeta^{n+1}_{i-1,j}+2\zeta^{n+1}_{i-2,j}\}/(6\delta x) \quad \text{for } u^n_{i,j} \text{ positive}, \tag{8}$$

and

$$u^n_{i,j}\{-2\zeta^{n+1}_{i+2,j}+10\zeta^{n+1}_{i+1,j}-9\zeta^{n+1}_{i,j}+2\zeta^{n+1}_{i-1,j}-\zeta^{n+1}_{i-2,j}\}/(6\delta x) \quad \text{for } u^n_{i,j} \text{ negative}. \tag{9}$$

The truncation error of this scheme is given by

$$\frac{1}{4}(\delta x)^3\left[u^n_{i,j}\frac{\partial^4\zeta}{\partial x_4}\right].$$

The diffusion terms (second-order derivatives) are discretized through the central difference scheme. At every fractional time step the Poisson Eq. (6) for the stream function is solved iteratively using SOR technique.

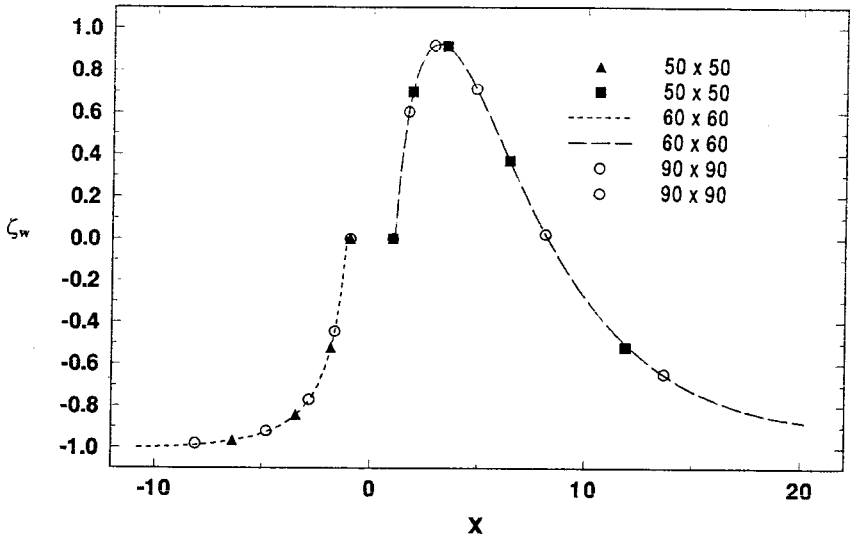


Fig. 2 Influences of the grid sizes on the wall shear ζ_w along the flat surface for Re=50

To check the grid dependency of the numerical solutions, number of grids were made to range between 50×50 and 90×90 at Reynolds number 50. Figure 2 shows the effects of grid size on the wall vorticity ζ_w. The effects of gird size on the solution are found to be minimal. We found that a nonuniform grid size .0159 ×.0167 near the obstacle produces the optimal solution at Re=50. For higher Re we chose a much finer grid so as to reduce the grid size near the far-field in the physical domain. The time step was taken as 0.001 originally but was increased in subsequent time *t*.

Results and Discussions

Solutions were obtained for various values of Reynolds number (Re≤400) withs the obstacle taken as a semi-circle mounted on a flat surface. In the upstream part of the flow a viscous wall layer close to the flat surface must be produced, for increased Re values. This boundary layer separates at an upstream position and joins to the thinner boundary layer on the obstacle itself. This thin boundary layer proceeds downstream through a second separation from the obstacle to form a thin free-shear layer. The free-shear layer then emerges almost horizontally atop the long eddy far downstream. Closure of that downstream eddy and a flow-reattachment process to the surface then take place on a much longer length scale.

Figure 3 shows the streamlines for Reynolds number 400. We found the presence of a downstream eddy even at smaller Reynolds numbers of 5 and less. According to Dean-Moffatt local analysis both downstream and upstream eddy exist in the Stokes limit of zero Re for a blunt two-dimensional roughness element in a shear flow. The length of the downstream large-scale separation eddy increases in size with an increase of Reynolds number. An increase in Re means that (for example) the velocity increases and it makes the separated flow regions expand and the separated flow

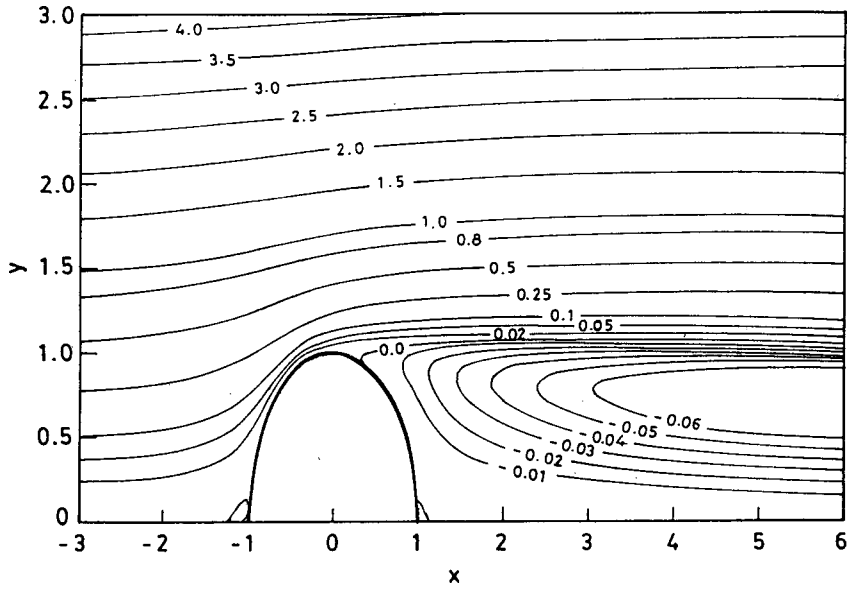


Fig. 3 Streamlines at Re=400

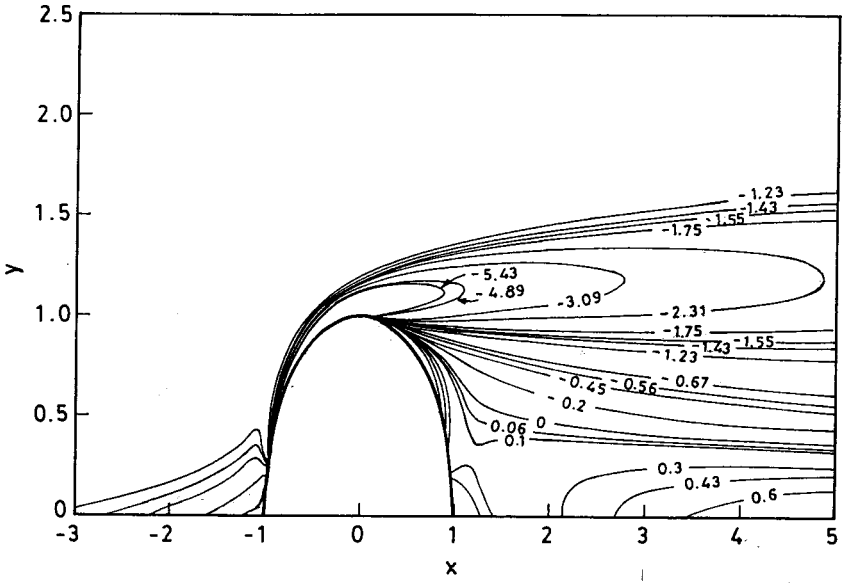


Fig. 4 Vorticity contours at Re=300

strength increased. It is clear from the figures that the center of the eddy shifts away from the obstacle as Re increases. The large-scale separation itself is a sudden localized process involving a relatively abrupt adverse pressure gradient.

The streamline patterns show the appearance of an upstream recirculation region very close to the obstacle. The upstream separation point x_1 lies between $-1.17<x_1<-1.13$ at Re=100 which shifts towards upstream to $-1.28<x_1<-1.25$ at Re=400. Part of the high-inertia fluid flow is deflected on the front face of the obstacle back upstream in a jet adjacent to the wall. This fluid moves with the pressure gradient. As the flow there loses energy, it rolls up to from recirculating zones. We found that the upstream separation length is small and remains almost invariant under variations in the Reynolds number.

Figure 4 gives the vorticity contours for Re=300. In the above figure we have presented the equi-vorticity lines. The decay of the vortex strength in the transverse direction is rather fast compared with that in the stream wise direction. Within the recirculation regions the value of the vorticity changes sign when one follows a closed streamline. Other numerical results not presented here show that an increase in Reynolds number leads to an increase in vortex strength. This is basically because an increase in Reynolds number reduces the boundary layer thickness on the obstacle which results in a higher vorticity distribution in the separated region.

The surface vorticity distributions are presented in Fig. 5 for Re of 100, 200, 300, 400. The wall vorticity on the upstream flat surface near the obstacle rapidly passes through zero and changes sign, compared with the more gradual development over the rest of the surface.

The effect on the downstream separation length (x_2) due to variations of Reynolds number is presented in Fig. 6. The downstream separation length increases as the Reynolds number increases. A near-linear variation of the downstream separation

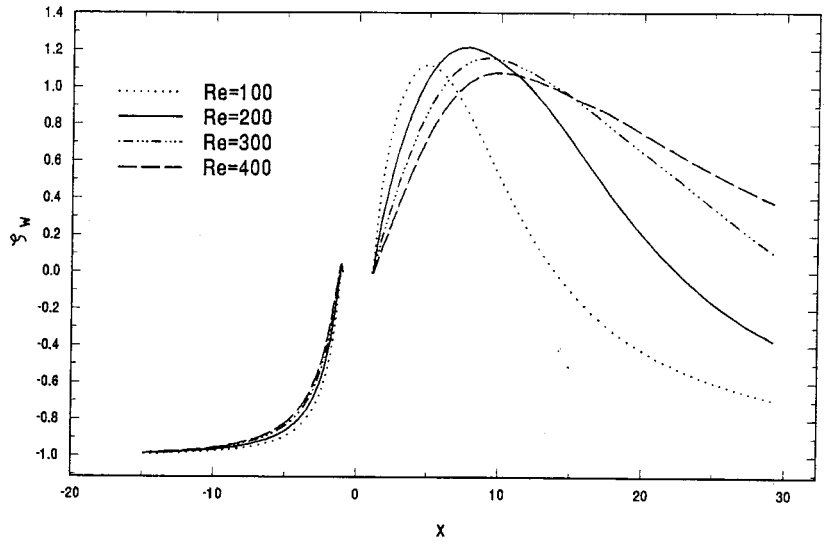


Fig. 5 Effects of Reynolds number on surface vorticity ζ_w along the flat surface

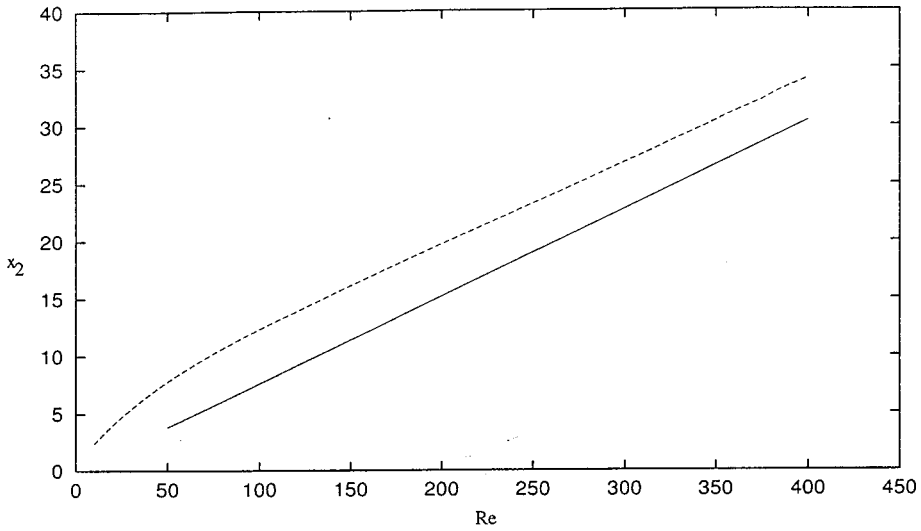


Fig. 6 Effects of Reynolds number on downstream reattachment point x_2;—theoretically predicted results

length with Reynolds number is evident from the figure, over almost all Reynolds numbers. This observation is in agreement with the theoretical prediction of Smith and Daniels [10]. Further, the eddy length tends to a finite nonzero value as Re tends to zero, as in the flow studied by Dennis and Smith [11].

The distribution of surface pressure along the obstacle is presented in Fig. 7. The maximum pressure occurs at or near the point where the flow separates from the obstacle.

The pressure distribution on the flat surface both upstream and downstream is presented on Fig. 8 at Re of 100, 200, 300, 400. There is not much change in pressure in the upstream portion of the flow. In the downstream position the surface pressure is quite high. The downstream surface pressure alters with a variation of Reynolds number. This difference in pressure distribution is associated with the variation of the distance required for vortex formation behind the obstacle and the strength of the formed vortices, as described earlier in this section.

Finally, it is worth pointing out that the theory above also predicts a gradual increase of upstream eddy length with Re. Indeed, at all Reynolds numbers there must be eddies, if only very small ones, sufficiently close to the geometrical corners, according to the Dean-Moffatt local analysis. It appears that significant growth of the upstream eddy with Re is delayed for the present semi-circular obstacle, which makes physical sense when compared with the blunt noncurved cases studied previously. The close agreement, however, on downstream eddy length for Reynolds numbers above approximately 50 and on the position of maximum pressure gradient is an encouraging feature in the current setting.

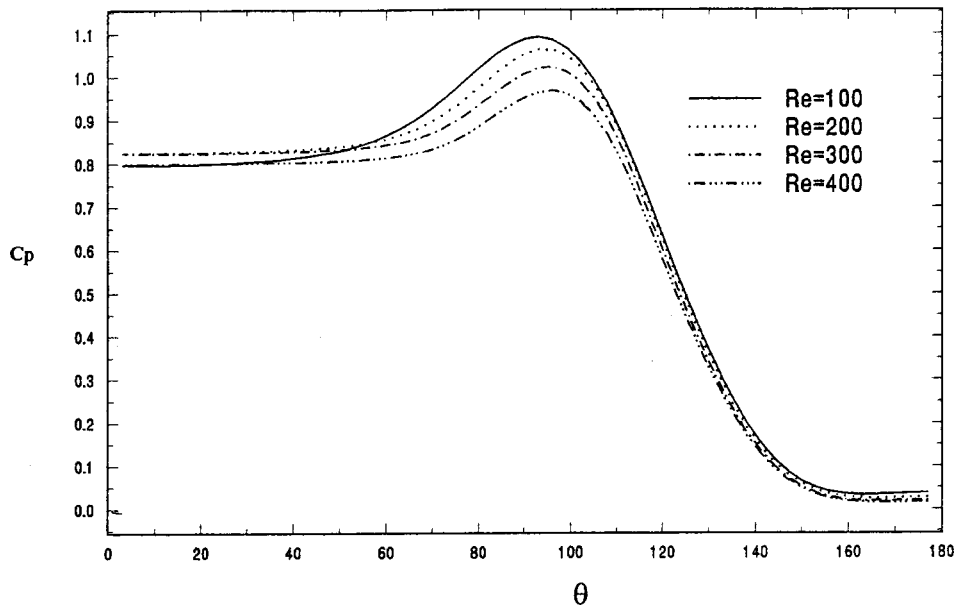


Fig. 7 Surface pressure distribution C_p along the semi-circular obstacle at different Re

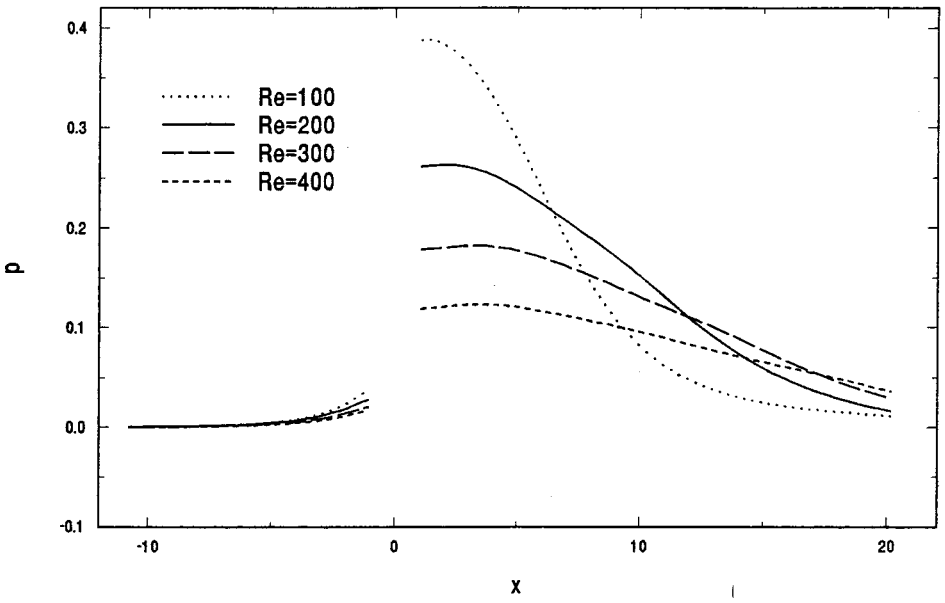


Fig. 8 Pressure distribution along the flat surface at different Re

References

[1] Smith, F. T., and Walton, A. G., 1998, ''Flow Past a Two- or Three-Dimensional Steep-Edged Roughness,'' Proc. R. Soc. London, Ser. A, **454**, pp. 31–69.

[2] Bhattacharyya, S., Dennis, S. C. R., and Smith, F. T., 2001, ''Separating Shear Flow Past a Surface-Mounted Blunt Obstacle,'' J. Eng. Math., **39**, pp. 47–62.

[3] Durst, F., and Loy, T., 1985, ''Investigation of Laminar Flow in a Pipe With Sudden Contraction of Cross Section Area,'' Comput. Fluids, **13**, pp. 15–36.

[4] Williams, P. T., and Baker, A. J., 1997, ''Numerical Simulations of Laminar Flows Over a 3D Backward-Facing Step,'' Int. J. Numer. Methods Fluids, **24**, pp. 1159–1183.

[5] Chang, T. P., and Sheu, Tony W. H., 1999, ''Time Evaluation of Laminar Flow Over a Three-Dimensional Backward-Facing Step,'' Int. J. Numer. Methods Fluids, **31**, pp. 721–745.

[6] Giguere, P., Dumes, G., and Lemay, J., 1997, ''Gurney Flap Scaling for Optimum Lift-to- Drag Ratio,'' AIAA J., **35**, pp. 1888–1890.

[7] Smith, F. T., 2000, ''On Physical Mechanisms in Two- and Three-Dimensional Separations,'' Philos. Trans. R. Soc. London, Ser. A, **358**, pp. 3091–3111.

[8] Martinuzzi, E. R., and Tropea, C., 1993, ''The Flow Around Surface Mounted Prismatic Obstacles Placed in a Fully Developed Channel Flow,'' ASME J. Fluids Eng., **115**, pp. 85–92.

[9] Meinders, E. R., and Hanjalic, K., 1999, ''Vortex Structure and Heat Transfer in Turbulent Flow Over a Wall-Mounted Matrix of Cubes,'' Int. J. Heat Fluid Flow, **20**, pp. 255–267.

[10] Smith, F. T., and Daniels, P. G., 1981, ''Removal of Goldstein's Singularity at Separation in Flow Past Obstacles in Wall Layers,'' J. Fluid Mech., **110**, pp. 1–37.

[11] Dennis, S. C. R., and Smith, F. T., 1980, ''Steady Flow Through a Channel With a Symmetrical Constriction in the Form of a Step,'' Proc. R. Soc. London, Ser. A, **372**, pp. 393–414.

The Near-Tip Stress Intensity Factor for a Crack Partially Penetrating an Inclusion

Zhonghua Li
e-mail: zhli@sjtu.edu.cn

Lihong Yang

School of Civil Engineering and Mechanics,
Shanghai Jiaotong University,
200240 Shanghai Minhang,
P.R. China

When a crack is lodged in an inclusion, the difference between the elastic modulus of the inclusion and matrix material will cause the near-tip stress intensity factor to be greater or less than that prevailing in a homogeneous material. A method is derived for calculation of the near-tip stress intensity factor for the inclusion with arbitrary shape. The derivation of the fundamental formula is based on the transformation toughening theory. The equivalent transformation strain contributed from modulus difference between inclusion and matrix is calculated from Eshelby equivalent inclusion approach. As validated by numerical examples, the developed formula has excellent accuracy.
[DOI: 10.1115/1.1651539]

1 Introduction

In two-phase materials, the crack-tip may be lodged within a second-phase particle. Depending on the geometry and modulus difference between the inclusion and matrix material, the crack-tip stress intensity factor (SIF) may be increased (stress amplification) or decreased (stress shielding or toughening). Hence, the analysis of stress state for the crack-tip within an inclusion is of importance in understanding the fracture behavior of composite materials and has received considerable attention, [1–4]. Another motivation for studying this problem is microcrack toughening in brittle materials, [1,5,6], where the microcracking process zone was treated as a particle with reduction in moduli resulting from the microcracking.

By using the techniques from complex variable elasticity together with numerical computation, the near-tip SIF for a circular inclusion centered at the tip of a semi-infinite crack can be calculated for arbitrary differences in the sets of moduli, [2]. For an inclusion with arbitrary shape, a closed-form solution to the lowest order effect of the modulus difference is given by Hutchinson [1]. In this study, we will formulate a method for calculation of the SIF for the crack-tip within an inclusion of arbitrary shape. This method is based on transformation toughening theory and Esbelby equivalent inclusion approach. The accuracy of the solution is validated by numerical examples.

2 Mode and Formulation

As shown in Fig. 1(*a*), the crack-tip is partially penetrating an inclusion, which may have an arbitrary shape, but is assumed to be symmetrical with respect to the crack plane. The Young's modulus and Poisson's ratio of the inclusion, E_I and ν_I, differ from those of the matrix material, E_M and ν_M. It is assumed that the size of the inclusion is small compared with crack length and other dimensions of the crack body. Therefore, the inclusion is within applied remote *K*-field and the near-tip fields are assumed to have the same classical form, denoted by K_{tip}:

$$\sigma_{ij}=\frac{K}{\sqrt{2\pi r}}\tilde{\sigma}_{ij}(\theta) \quad r\to\infty \tag{2.1}$$

$$\sigma_{ij}=\frac{K_{\text{tip}}}{\sqrt{2\pi r}}\tilde{\sigma}_{ij}(\theta) \quad r\to 0. \tag{2.2}$$

Now, we consider a differential element dA within the inclusion, which sustains an applied strain field $\mathbf{e}^A$, exerted by the remote stress intensity *K*. The equivalent transformation strain in dA, $\mathbf{e}^T$, is given by

$$\mathbf{e}^T=[(\mathbf{C}_I-\mathbf{C}_M)\mathbf{S}+\mathbf{C}_M]^{-1}(\mathbf{C}_M-\mathbf{C}_I)\mathbf{e}^A \tag{2.3}$$

according to Eshelby equivalent inclusion approach, [7,8], where $\mathbf{S}$ is the Eshelby tensor, dependent solely upon the inclusion shape and the Poisson's ratio of the matrix. $\mathbf{C}_I$ and $\mathbf{C}_M$ are the elastic tensors of the inclusion and matrix, respectively. As shown in (2.3), the equivalent transformation strain $\mathbf{e}^T$ in dA varies with the applied strain $\mathbf{e}^A$, and is not zero for an inhomogeneous inclusion ($C_I\neq C_M$).

The Eshelby approach is mathematically rigorous for an infinite matrix containing a single ellipsoidal inclusion. When the inclusion undergoes a uniform stress-free transformation strain, the stress and strain within the inclusion are uniform. However, in order to utilize the approach in more realistic situations, there has been considerable activity in extending Eshelby approach to various problems, such as the interaction of two ellipsoidal inclusions, [9], the behavior of hybrid composite, [10], and short fiber-reinforced composites, [8], the calculation of the stress fields inside a nonellipsoidal inclusion which are not uniform, [11], to cite only a few examples. In the present study, we extend the Eshelby approach to the case of an inclusion with arbitrary shape embedded in a crack-tip field. Either the nonellipsoidal shape of the inclusion considered or the singular crack-tip field will result in a nonuniform stress-strain field within the inclusion. However, we assume that the Eshelby theory can be used to each differential element within the inclusion, which undergoes uniform transformation strain determined by (2.3) and the resultant stresses in which are uniform. Then a nonuniform transformation strains, therefore also the stresses, inside the inclusion, can be obtained by integrating (2.3) in the domain of the inclusion.

For simplicity, it is assumed in the present paper that the inclusion and matrix material are elastic and isotropic and their Poisson's ratios are the same, denoted by ν. Then we have

$$\mathbf{C}_I=\alpha\mathbf{C}_M \tag{2.4}$$

where

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Oct. 10, 2001; final revision, Sept. 19, 2003. Associate Editor: B. M. Moran. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

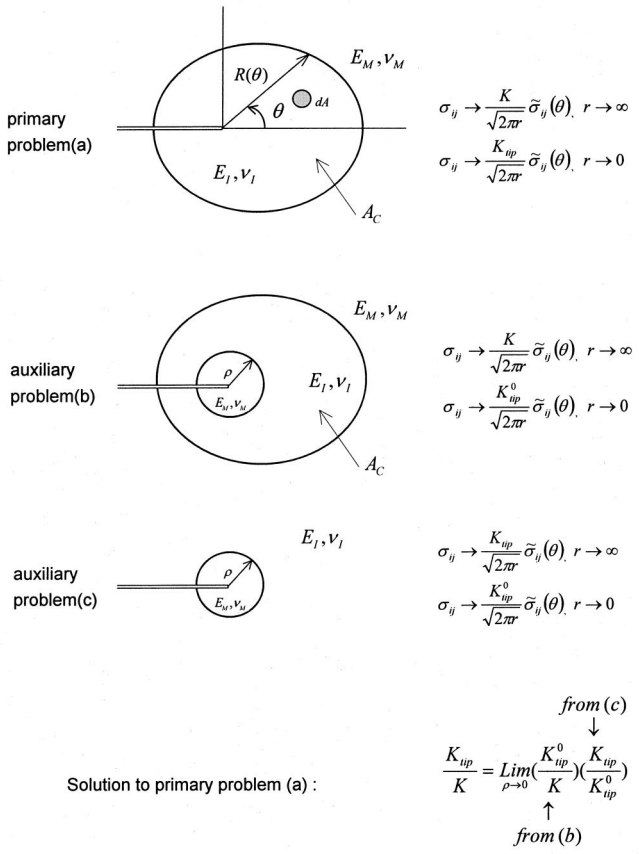


Fig. 1 Definitions of auxiliary problems (b) and (c) and their use in construction of the solution to the primary problem (a)

$$\alpha = E_I / E_M . \tag{2.5}$$

Combining (2.3) and (2.4), it gives

$$\mathbf{e}^T = \mathbf{L}\mathbf{e}^A \tag{2.6}$$

where

$$\mathbf{L} = [(\alpha - 1)\mathbf{S} + \mathbf{I}]^{-1}(1 - \alpha). \tag{2.7}$$

Here, $\mathbf{I}$ is the identity tensor. Thus, the tensor $\mathbf{L}$ relates the equivalent transformation strain $\mathbf{e}^T$ in the inclusion to the applied strain $\mathbf{e}^A$ without going into the details of the form of the $\mathbf{C}_I$ and $\mathbf{C}_M$ tensors. According to transformation toughening theory, [12,13], the increment in SIF due to the differential element with transformation strain $\mathbf{e}^T$ defined in (2.6) is given by

$$dK^0_{tip} = \frac{1}{2\sqrt{2\pi}} \frac{E_M}{1 - \upsilon^2} r^{-3/2} \Omega(e^T_{\alpha\beta}, \theta) dA \tag{2.8}$$

for a plane-strain mode I crack, where

$$\Omega(e^T_{\alpha\beta}, \theta) = (e^T_{11} + e^T_{22}) \cos\frac{3\theta}{2} + 3e^T_{12} \cos\frac{5\theta}{2} \sin\theta + \frac{3}{2}(e^T_{22} - e^T_{11}) \sin\theta \sin\frac{5\theta}{2}. \tag{2.9}$$

It is essential to note that Eq. (2.8) is derived for the case that the transformation area have the same moduli with its surrounding. Therefore, the elastic modulus E_M used in (2.8), and the K^0_{tip} is the stress intensify factor for the crack-tip within a medium of same moduli as matrix materials, not the desired factor K_{tip} for the case shown in Fig. 1(*a*).

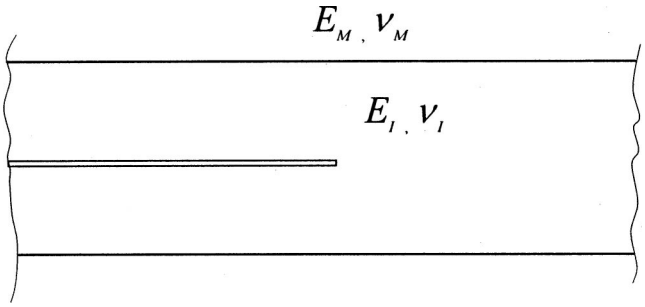


Fig. 2 Spherical geometry used to infer solution to auxiliary problem (c) shown in Fig. 1

To obtain the K_{tip}/K solution for the primary problem in Fig. 1(*a*), we adopt the method developed by Hutchinson [1]. The solution is constructed using solutions to two auxiliary problems denoted by (b) and (c) in Fig. 1. Once the solutions to the two auxiliary problems are in hand, the ratio of the stress intensify factors sought is given by

$$\frac{K_{\text{tip}}}{K} = \left(\frac{K^0_{\text{tip}}}{K}\right) \Big/ \left(\frac{K^0_{\text{tip}}}{K_{\text{tip}}}\right) \tag{2.10}$$

where K^0_{tip}/K is the ratio of the near-tip to remote intensity factor in the auxiliary problem (b), and $K^0_{\text{tip}}/K_{\text{tip}}$ is the corresponding ratio in the auxiliary problem (c).

The solution to the auxiliary problem (b) can be written down immediately using (2.8)

$$\frac{K^0_{\text{tip}}}{K} = 1 + \frac{E_M}{2\sqrt{2\pi}(1 - \nu^2)K} \int_A r^{-3/2} \Omega(e^T_{\alpha\beta}, \theta) dA \tag{2.11}$$

where the area integral extends over the upper half of A_C, excluding the inner circular region.

To generate the solution to the auxiliary problem (c), Hutchinson [1] exploited a special region A_C for which K_{tip}/K is known. The special region A_C is the infinite strip with a centered semi-infinite crack shown in Fig. 2. From a simple energy argument or application of the J-integral (seen (3.15) in the next section for the case of $\nu_I = \nu_M$) the following relation holds:

$$\frac{K_{\text{tip}}}{K} = \sqrt{\alpha} \tag{2.12}$$

exactly for this problem. Then by evaluating K^0_{tip}/K for this special mode, one can "back out" the desired universal result $K^0_{\text{tip}}/K_{\text{tip}}$ for the auxiliary problem (c) used in (2.10). The details for evaluating $K^0_{\text{tip}}/K_{\text{tip}}$ for the auxiliary problem (c), and K^0_{tip}/K for the auxiliary problem (b) will be given in the following section.

3 The Crack-Tip Stress Intensity Factor

Throughout this paper, it is assumed that the size of the inclusion is small compared with the length of the crack. Under this condition the applied strain field to the inclusion shown in Figs. 1(*a*) and (*b*) is that of the mode I crack controlled by the remote stress intensity factor K

$$\left.\begin{aligned} e_{11}^{A} &= \frac{K}{E_M\sqrt{2\pi r}}(1+v)\cos\frac{\theta}{2}\left[(1-2v)-\sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right] \\ e_{22}^{A} &= \frac{K}{E_M\sqrt{2\pi r}}(1+v)\cos\frac{\theta}{2}\left[(1-2v)+\sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right] \\ e_{12}^{A} &= \frac{K}{E_M\sqrt{2\pi r}}(1+v)\cos\frac{\theta}{2}\sin\frac{\theta}{2}\cos\frac{3\theta}{2} \\ e_{33}^{A} &= e_{13}^{A}=e_{23}^{A}=0 \quad \text{for plane strain} \end{aligned}\right\}. \tag{3.1}$$

For a differential element with circular section inside A_C, the components of the Esbelby tensor are given by, [14],

$$\left.\begin{aligned} S_{1111} &= S_{1122}=\frac{5-4\nu}{8(1-\nu)}, \quad S_{1122}=S_{2211}=\frac{4\nu-1}{8(1-\nu)} \\ S_{1133} &= S_{2233}=\frac{\nu}{2(1-\nu)}, \quad S_{1212}=\frac{3-4\nu}{4(1-\nu)} \\ S_{1313} &= S_{2323}=\frac{1}{2}, \end{aligned}\right\} \tag{3.2}$$

and other components are zero. Substituting (3.2) into (2.7), it gives

$$\left.\begin{aligned} L_{1111} &= L_{2222}=\frac{(1-\alpha)(1-\nu)(3-4\nu+5\alpha-4\nu\alpha)}{(1+\alpha-2\nu)(1+3\alpha-4\nu\alpha)} \\ L_{1122} &= L_{2211}=-\frac{(1-\alpha)^2(1-\nu)(1-4\nu)}{(1+\alpha-2\nu)(1+3\alpha-4\nu\alpha)} \\ L_{1133} &= L_{2233}=\frac{(1-\alpha)^2\nu}{(1+\alpha-2\nu)}, \quad L_{3333}=(1-\alpha) \\ L_{1212} &= \frac{4(1-\alpha)(1-\nu)}{(1+3\alpha-4\nu\alpha)}, \quad L_{1313}=L_{2323}=\frac{2(1-\alpha)}{1+\alpha} \end{aligned}\right\}, \tag{3.3}$$

and other components of the **L** tensor are zero.

The transformation strains used in (2.9) are determined by (2.6)

$$\left.\begin{aligned} e_{\alpha\alpha}^{T} &= (L_{1111}+L_{1122})(e_{11}^{A}+e_{22}^{A})=\frac{4(1-\alpha)(1-2v)(1-v^2)K}{(1+\alpha-2v)E_M\sqrt{2\pi r}}\cos\frac{\theta}{2} \\ e_{22}^{T}-e_{11}^{T} &= (L_{1111}-L_{1122})(e_{22}^{A}-e_{11}^{A})=\frac{4(1-\alpha)(1-v^2)K}{(1+3\alpha-4v\alpha)E_M\sqrt{2\pi r}}\sin\theta\sin\frac{3\theta}{2} \\ e_{12}^{T} &= L_{1212}e_{12}^{A}=\frac{2(1-\alpha)(1-v^2)K}{(1+3\alpha-4v\alpha)E_M\sqrt{2\pi r}}\sin\theta\cos\frac{3\theta}{2} \end{aligned}\right\}. \tag{3.4}$$

Substituting (3.4) and (2.9) into (2.11), we have

$$\begin{aligned} \frac{K_{\text{tip}}^{0}}{K} &= 1+\frac{1}{\pi}\int_0^{\pi}\{\ln[R(\theta)]-\ln\rho\} \\ &\quad\times\left(C_1\cos\frac{\theta}{2}\cos\frac{3\theta}{2}+C_2\sin^2\theta\cos\theta\right)d\theta \\ &= 1+\beta_1C_1+\beta_2C_2 \end{aligned} \tag{3.5}$$

where

$$C_1=\frac{(1-\alpha)(1-2\nu)}{(1+\alpha-2\nu)} \tag{3.6}$$

$$C_2=\frac{3(1-\alpha)}{2(1+3\alpha-4\nu\alpha)} \tag{3.7}$$

$$\beta_1=\frac{1}{\pi}\int_0^{\pi}\ln[R(\theta)]\cos\frac{\theta}{2}\cos\frac{3\theta}{2}d\theta \tag{3.8}$$

$$\beta_2=\frac{1}{\pi}\int_0^{\pi}\ln[R(\theta)]\sin^2\theta\cos\theta d\theta. \tag{3.9}$$

Since the integral $\int_0^{\pi}(C_1\cos\theta/2\cos 3\theta/2+C_2\sin^2\theta\cos\theta)d\theta=0$ in (3.5), the ratio of K_{tip}^0/K is independent of the radius of the inner circular core ρ. It can be also seen that β_1 and β_2 are unchanged when $R(\theta)$ is replaced by $\lambda R(\theta)$ and are thus dependent on the shape, but not on the size of the inclusion. When A_C is a circular region centered at the crack tip, $\beta_1=\beta_2=0$. These characters of the solution to the auxiliary problem (b) are consistent with those obtained by Hutchison [1].

We return to the problem shown in Fig. 2. The integral in the definitions of β_1 and β_2 are easily evaluated for this geometry ($\beta_1=0.5$, $\beta_2=-0.125$). Then, the solution to the auxiliary problem (c) for this geometry is given by

$$\frac{K_{\text{tip}}^{0}}{K}=1+0.5C_1-0.125C_2 \tag{3.10}$$

from (3.10) and (2.12), one immediately obtains the solution for the auxiliary problem (c) shown in Fig. 1(*c*).

$$\frac{K_{\text{tip}}^{0}}{K_{\text{tip}}}=\frac{1+0.5C_1-0.125C_2}{\sqrt{\alpha}} \tag{3.11}$$

By combining (3.11) and (3.5) according to (2.10), the general solution for the primary problem shown in Fig. 1(*a*) can be finally given by

$$\frac{K_{\text{tip}}}{K}=\frac{\sqrt{\alpha}(1+\beta_1C_1+\beta_2C_2)}{1+0.5C_1-0.125C_2}. \tag{3.12}$$

To obtain an explicit solution, it was assumed in the previous derivation that the Poisson's ratios of the inclusion and matrix are the same. This assumption sets a limit to apply the Eq. (3.12) for more general case. However, this limitation may be relaxed by introduction a modified factor. The J-integral from integration contours that circle the crack tip lying outside inclusion is given by

$$J=(1-\nu_M^2)K^2/E_M \tag{3.13}$$

while the J-integral for all such contours lying inside inclusion is

$$J_{tip}=(1-\nu_I^2)K_{tip}^2/E_I. \tag{3.14}$$

Then one obtains

$$\frac{K_{\text{tip}}}{K}=\sqrt{\frac{J_{\text{tip}}\alpha}{J}\frac{(1-\nu_M^2)}{(1-\nu_I^2)}}. \tag{3.15}$$

for the plane-strain condition.

For the case of $\nu_I=\nu_M$, (3.12) and (3.15) are identical if the effect of the inclusion on J can be neglected. This is true because we have assumed that the remote stress-strain field is controlled by K, i.e., the perturbation of the small inclusion on the remote K-field is neglected. Consequently, (3.12) and (3.15) leads to a modified factor $\sqrt{(1-\nu_M^2)/(1-\nu_I^2)}$ in (3.12),

$$\frac{K_{\text{tip}}}{K}=C_0(1+\beta_1 C_1+\beta_2 C_2) \tag{3.16}$$

where

$$C_0=\frac{1}{1+0.5C_1-0.125C_2}\sqrt{\frac{\alpha(1-\nu_M^2)}{(1-\nu_I^2)}} \tag{3.17}$$

for the case of $\nu_I\neq\nu_M$. It should be noted that the Poisson's ratio used in C_1 and C_2 is ν_M because C_1 and C_2 are derived from the remote stain field. As well be seen in following section, the modified formula gives good approximation for the case of $\nu_I\neq\nu_M$.

4 Numerical Examples

The accuracy of the fundamental formula (3.12) and the modified formula (3.16) were examined by three numerical examples.

An exactly numerical solution to circular inclusion centered at the tip of a semi-infinite crack has been given in [2], which is firstly used to measure the accuracy of (3.12). As shown in Fig. 3, good agreement between our solution and Steif's results is found. For comparison, the lowest order solution and the modified lowest order results, [1], are also plotted in Fig. 3. Though the agreement is good for the modified lowest order results when $E_I/E_M<2$, it fails as E_I/E_M becomes larger than 2.

Detailed finite element analyses were performed for a center-cracked plate and a three-point bending specimen under plain-strain conditions. Identical circular inclusions are centered on each crack tip for the two calculation models. To calculate the K_{tip} in the finite element analyses, three J-integral contours within inclusion were set around the crack-tip. The K_{tip} was calculated by (3.14) from the mean value of the three contours (note, in fact, that they nearly have no difference), J_{tip}. The K used to normalize K_{tip} is calculated for models of the homogeneous matrix material at the same applied load. The configuration and size effects on the K_{tip}/K were analyzed for the two calculation models. It was found that there is nearly no specimen configuration and scale effects on K_{tip}/K so long as the specimen sizes are two or three times of the crack length and the ratio R/a of the inclusion radius R to the crack length a is less than 0.15. The same conclusion was obtained in [4]. In the present finite element analyses, the calculation models are constructed to be no scale effect on K_{tip}/K, and therefore their geometries are unnecessary to specify. The finite element results, for a circular inclusion centered at the crack-tip are also plotted in Fig. 3. They are excellent agreement with the exact solution obtained by Steif [2]. Therefore, the finite element analyses can be used to measure the accuracy of (3.12) and (3.16) for the inclusion with arbitrary shape.

Figure 4 compares the results calculated from (3.12) and the modified lowest order solution, [1], with those obtained from finite element analyses for a square inclusion. The crack-tip partially penetrates the inclusion (1/4 edge length). As shown in Fig. 4, our results are in good agreement with those of the finite element analyses.

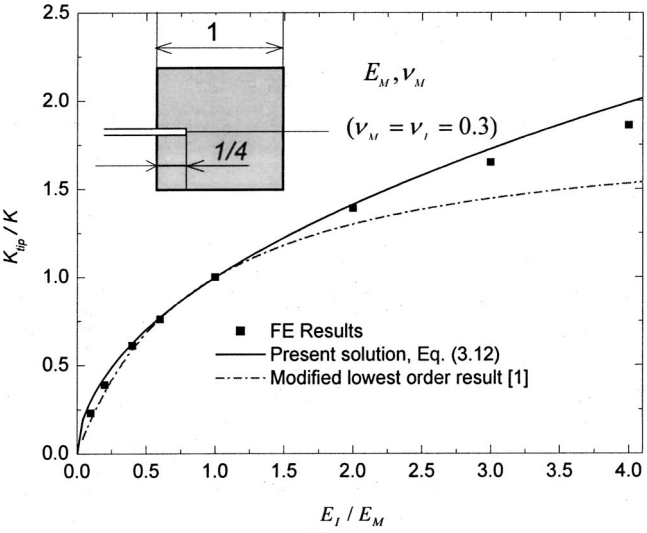


Fig. 4 A comparison of the results calculated from Eq. (3.12), the modified lowest-order solution and finite element analysis for a square inclusion

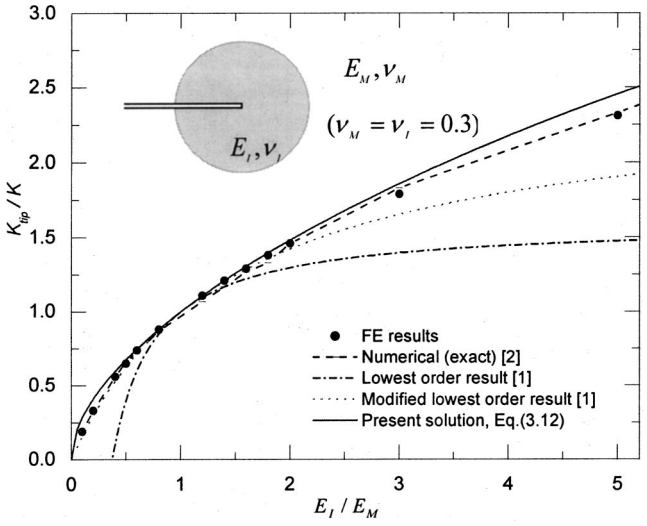


Fig. 3 A comparison of the selected results for a circular inclusion centered at the crack tip

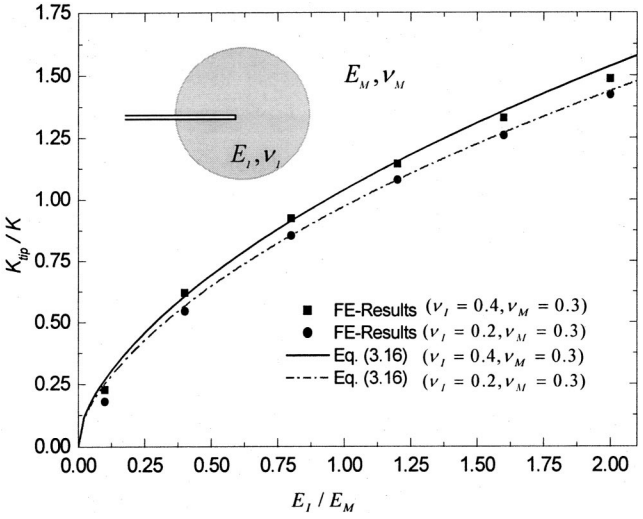


Fig. 5 A comparison of the results calculated from the modified Eq. (3.16) and finite element analyses for a circular inclusion in which both elastic modulus and Poisson's ratio differ from those of matrix material

In Fig. 5, the values of K_{tip}/K for a circular inclusion centered at the crack-tip predicted from the modified formula (3.16) are compared with the results of finite element analyses for the case where ν_I and ν_M are different, and good agreement can be found.

5 Conclusions and Discussions

When a crack is lodged in an inhomogeneous inclusion the near-tip field will be changed by the modulus difference between inclusion and matrix material. A closed-form solution for predicting the K_{tip}-field is derived based on transformation toughening theory and Eshelby equivalent inclusion approach. The numerical examples for different inclusion shape and modulus ratio between inclusion and matrix provide compelling evidence that the fundamental formula is satisfactory in accuracy. The fundamental formula (3.12) and its modified form (3.16) provide a quick estimate for the effects of shape, location, and stiffness of an inclusion surrounding crack-tip on the crack-tip field.

The limitation of the fundamental formula is that the remote applied K field was used to calculate the equivalent transformation strain of the inclusion. Hence, the size of the inclusion must be small compared with the length of crack and other dimensions of the cracked body. Consequently, the derived equations are independent of the geometry configurations of the mode I crack.

References

[1] Hutchinson, J. C., 1987, ''Crack Tip Shielding by Micro-Cracking in Brittle Solids,'' Acta Metall., **35**, pp. 1605–1619.

[2] Steif, P. S., 1987, ''A Semi-Infinite Crack Partially Penetrating a Circular Inclusion,'' ASME J. Appl. Mech., **54**, pp. 87–92.

[3] Erdogan, F., and Gaupta, G. D., 1975, ''The Inclusion Problem With a Crack Crossing the Boundary,'' Int. J. Fract., **11**, pp. 13–27.

[4] Li, R., and Chudnovsky, A., 1993, ''Energy Analysis of Crack Interaction With an Elastic Inclusion,'' Int. J. Fract., **63**, pp. 247–261.

[5] Evans, A. G., and Faber, K. T., 1981, ''Toughening of Ceramics by Circumferential Microcracking,'' J. Am. Ceram. Soc., **64**, pp. 394–398.

[6] McMeeking, R. M., and Evans, A. G., 1982, ''Mechanics of Transformation Toughening in Bittle Materials,'' J. Am. Ceram. Soc., **65**, pp. 242–246.

[7] Eshelby, J. D., 1957, ''The Determination of the Elastic Fields of an Ellipsoidal Inclusion, and Related Problems,'' Proc. R. Soc. London, Ser. A, **241**, pp. 376–396.

[8] Withers, D. J., Stobbs, W. M., and Pedersen, O. B., 1989, ''The Application of the Eshelby Method of Internal Stress Determination to Short Fibre Metal Matrix Composites,'' Acta Metall., **37**, pp. 3061–3084.

[9] Moschobidis, Z. A., and Mura, T., 1975, ''Two Ellipsoidal Inhomegeneities by the Equivalent Inclusion Method,'' ASME J. Appl. Mech., **42**, pp. 847–852.

[10] Taya, M., and Chou, T. W., 1981, ''On Two Kinds Ellipsoidal Inhomegeneities in an Infinite Elastic Body: An Application to a Hybrid Composites,'' Int. J. Solids Struct., **17**, pp. 553–563.

[11] Johnson, W. C., Earmme, Y. Y., and Lee, J. K., 1980, ''Approximation of the Strain Field Associated With an Inhomegeneous Pricipitate,'' ASME J. Appl. Mech., **47**, pp. 775–780.

[12] Lambropoulos, J. C., 1986, ''Shear, Shape and Orientation Effects in Transformation Toughening in Ceramics,'' Int. J. Solids Struct., **22**, pp. 1083–1106.

[13] Budiansky, B., Hutchinson, J. W., and Lambropoulos, J. C., 1983, ''Continuum Theory of Dilatant Transformation Toughening in Ceramics,'' Int. J. Solids Struct., **19**, pp. 337–355.

[14] Mura, T., 1987, *Micromechanics of Defects in Solids*, Second Rev. Ed., Kluwer, Dordrecht, The Netherlands.

Brittle to Plastic Transition in the Dynamic Mechanical Behavior of Partially Saturated Granular Materials

Simon M. Iveson[1]

Neil W. Page

Faculty of Engineering and Built Environment,
University of Newcastle,
Callaghan NSW 2308, Australia

The effect of liquid viscosity, surface tension and strain rate on the deformation behavior of partially saturated granular material was studied over a ten order of magnitude range of capillary number (the ratio of viscous to capillary forces). Glass spheres of average size 35 microns were used to make pellets of 35% porosity and 70% liquid saturation. As the capillary number increased, the failure mode changed from brittle cracking to ductile plastic flow. This change coincided with the transition from strain-rate independent flow stress to strain-rate dependent flow stress noted previously [Iveson, S. M., Beathe, J. A., and Page, N. W., 2002, "The Dynamic Strength of Partially Saturated Powder Compacts: The Effect of Liquid Properties," Powder Technol., **127***, pp. 149–161]. This change in failure mode is somewhat counter-intuitive, because it is the opposite of that observed for fully saturated slurries and pastes, which usually change from plastic to brittle with increasing strain rate. A model is proposed which predicts the functional dependence of flow stress on capillary number and also explains why the flow behavior changes. When capillary forces dominate, the material behaves like a dry powder: Strain occurs in localized shear planes resulting in brittle failure. However, when viscous forces dominate, the material behaves like a liquid: Shear strain becomes distributed over a finite shear zone, the size of which increases with strain rate. This results in less strain in each individual layer of material, which promotes plastic deformation without the formation of cracks. This model also explains why the power-law dependency of stress on strain rate was significantly less than the value of 1.0 that might have been expected given that the interstitial liquids used were Newtonian.* [DOI: 10.1115/1.1753269]

1 Introduction

The mechanical behavior of wet granular material is of interest in a number of fields, including agitated wet agglomeration, handling of filter cakes, extrusion and pumping of pastes, soil mechanics, and shape forming of ceramic parts, [1–6]. High strain rate deformation is involved in many cases. For example, high impact speeds and viscous binders are frequently present in agitated granulation. However, traditionally most research on wet granule strength has been performed at low and invariant strain rates with relatively nonviscous binders, [7–9]. Saturated pastes have been found to have some strain-rate dependency, but only a limited range of strain rates has been studied, [10,11]. Dynamic effects have also received some attention in the soil mechanics and slurry pumping literature due to the phenomena of vibrational liquefaction, but this work is restricted to saturated systems with water as the interstitial fluid, a relatively nonviscous liquid, [12,13].

Recently, more of an interest has been taken in the dynamic behavior of wet powders with viscous binders. Iveson and Litster [14] found that using glycerol (a viscous liquid) instead of water greatly reduced the amount of impact deformation experienced by granules. Discrete element simulations of granule impacts have also predicted that increasing binder viscosity will reduce granule deformation, [15]. In contrast, at low strain rates (5 mm/min), Mills et al. [16] found that the crushing strength of granules made with 90–180 μm sized sand was unaffected by binder viscosity for four different silicone oils of viscosity ranging from 0.02 to 0.5 Pa·s.

Iveson and co-workers [17–19] have directly measured the compressive strength of pellets in a high-speed load frame at crosshead speeds ranging from 0.01 to 150 mm/s (equivalent to bulk strain rates varying from $4\times10^{-4}\,s^{-1}$ to $6\,s^{-1}$). In initial work they used spherical glass particles with an average size d_p of 35 microns, [17]. The pellet porosity and liquid saturation were held constant at 35% and 70%, respectively. Six different liquid binders were used to cover a range of surface tension γ and viscosity μ from 0.025 to 0.072 N/m and 0.001 to 60 Pa·s, respectively.

The measured peak flow stress σ_{pk} results all collapsed onto a single curve when plotted as the dimensionless flow stress Str* versus the capillary number Ca. This curve can be empirically described by the equation

$$\mathrm{Str}^* = k_1 + k_2\overline{\mathrm{Ca}}^n, \tag{1}$$

where $\mathrm{Str}^* = \sigma_{pk} d_p/\gamma\cos\theta$, d_p is the average particle size, θ is the contact angle, $\overline{\mathrm{Ca}} = d_p\mu\dot{\varepsilon}/(\gamma\cos\theta)$ and $\dot{\varepsilon}$ is the bulk strain rate (deformation velocity divided by instantaneous pellet length). Str* is the ratio of peak flow stress to capillary (surface tension) forces. Ca is the ratio of viscous to capillary forces. The best-fit values of the three parameters were $k_1 = 5.3\pm0.4$, $k_2 = 280\pm40$, and $n = 0.58\pm0.04$. Subsequent work has also varied particle size and morphology, [18,19].

From Eq. (1) it can be seen that at low strain rates the flow stress was strain-rate independent while at high strain rates the flow stress was proportional to strain rate to the power 0.58. Ive-

[1]Currently at the Department of Chemical Engineering, Universitas Pembangunan Nasional "Veteran" Yogyakarta, Jl. SWK 104 Condongcatur Yogyakarta 55283, Indonesia. e-mail: ivesonsj@omf.net

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 1, 2002; final revision, Nov. 20, 2003. Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering University of California—Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

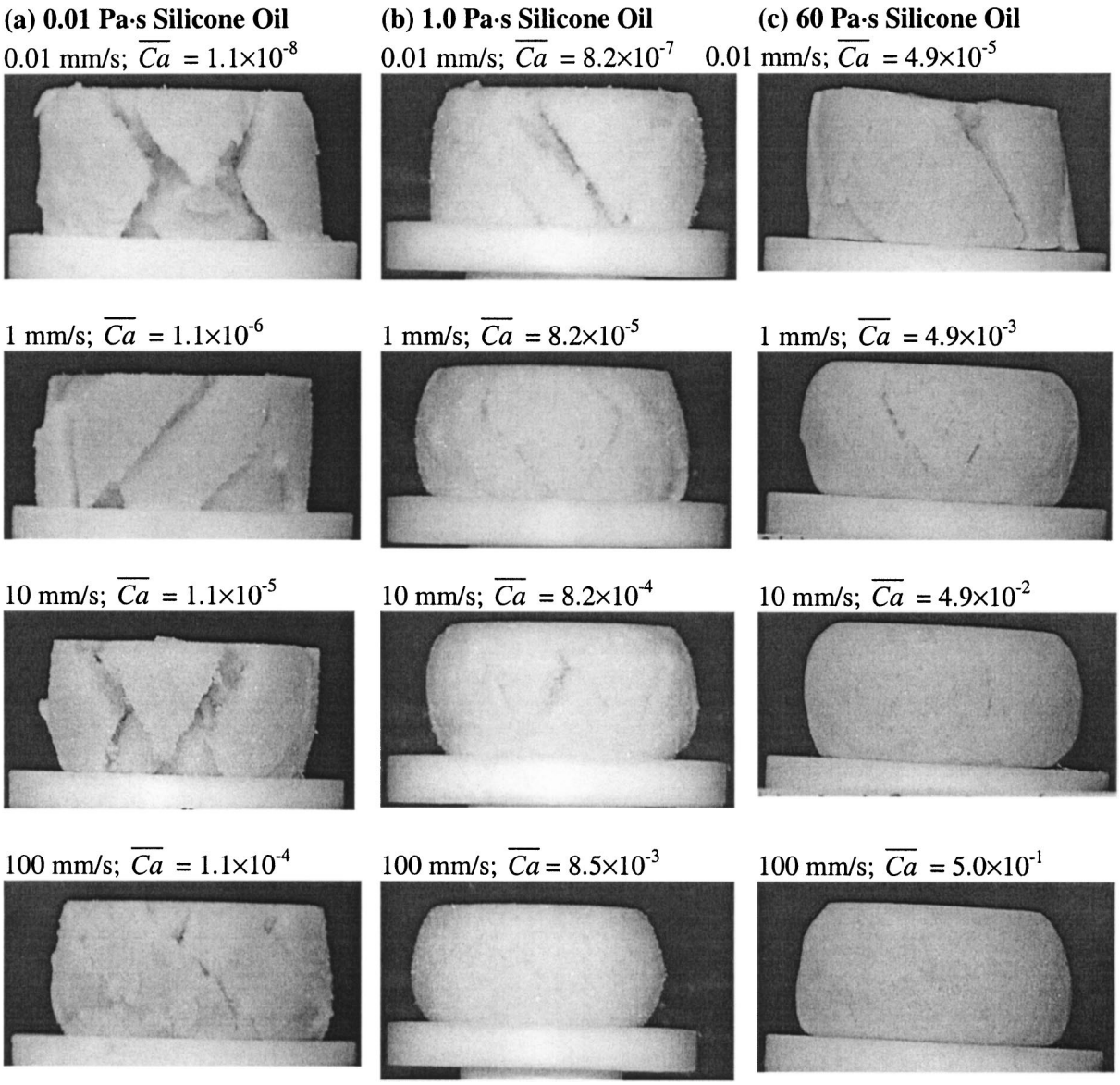


Fig. 1 Photos of pellets after 10 mm deformation at various speeds for (*a*) 0.01 Pa·s, (*b*) 1.0 Pa·s, and (*c*) 60 Pa·s viscosity silicone oils. Initial pellet aspect ratio 1.25 (25 mm×ϕ20 mm). Scale: diameter of bottom platen is 38 mm.

son et al. [17] commented that this power-law dependence was less than the value of 1.0 expected when viscous effects dominate. They speculated that this might be attributed to the width of the shear zone increasing with strain rate. In this paper, we develop a model based on this hypothesis and also present photographs of the deformed pellets that show a transition from brittle to plastic failure as viscous effects become significant.

2 Experimental

Figure 1 shows photographs of the pellets bound with 0.01, 1.0 and 60 Pa·s viscosity silicone oils after a total compressive strain of 0.4 at strain rates ranging from $4\times10^{-4}\,s^{-1}$ to $6\ s^{-1}$ (full details of the sample preparation and testing procedures are reported in Iveson et al. [17]). There were two different modes of failure. At low velocities, all pellets displayed brittle behavior and failed by large-scale cracking. Several of the recovered specimens displayed the commonly reported ''apple core'' shape, where pellet failure occurs by shearing along planes at approximately 45 deg to the direction of the applied load. These specimens fell apart and could not be recovered in one piece. At high velocities, however, the pellets bound with 1.0 and 60 Pa·s silicone oil showed no visible cracking (Figs. 1(*b*) and 1(*c*)). These pellets appear to have failed by uniform plastic flow. They retained their integrity and could be easily removed in one piece after the test.

For the 1.0 and 60 Pa·s silicone oil bound pellets, the failure mode changed from brittle to plastic as the deformation velocity increased. Even for the 0.01 Pa·s silicone oil bound pellets (Fig. 1(*a*)), the extent of brittle failure appears to become less extreme as the deformation velocity increased. Hence the velocity at which the transition from brittle to plastic flow occurred decreased with increasing liquid viscosity.

These observations suggest that the brittle-to-plastic transition depends on the relative importance of viscous forces. This is highlighted in Fig. 2 where schematic drawings of a selection of the deformed pellets are shown above the previously measured flow stress data, [17]. The transition from brittle to plastic flow behavior occurred at Ca in the range 10^{-4} to 10^{-3}. This coincides with the transition from Region I (strain-rate independent flow stress) to Region II (strain-rate dependent flow stress).

This same pattern of brittle failure at low speeds and plastic

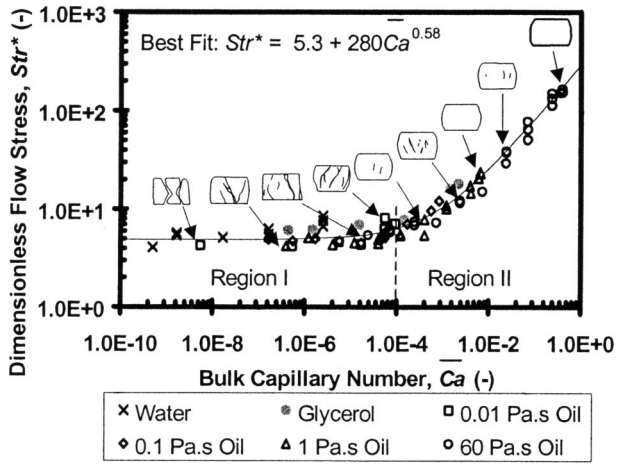


Fig. 2 The dimensionless flow stress results of Iveson et al. [17] with added schematics of deformed pellets shown above selected points. Pellets: 35% porosity, 70% saturation, 35 micron glass ballotini. Line shows best fit of Eq. (1).

failure at high speeds and with viscous binders was also observed when the platens were lubricated, [17], and when particle size was varied, [18]. This transition was also present, although less pronounced, with particles of nonspherical shape, [19]. Hence we believe that this is a general phenomenon. Note that the liquid content was held constant, so the transition to more liquid-like flow at high strain rates is nothing to do with changes in liquid saturation.

The response of materials to increasing strain rate is material dependent. For some metals, high strain rates causes adiabatic heating which softens the metal and hence promotes ductile flow. In contrast, for most polymers, increasing the strain rate reduces the time available for the long-chain polymer molecules to relax and rearrange. This results in a shift from ductile to brittle behavior, [20,21]. Concentrated pastes and suspensions usually exhibit shear dilatancy. This causes a reduction in pore pressure. At high strain rates, liquid is unable to flow in fast enough to fill the dilated regions and cavitation occurs. This results in a transition from ductile to brittle behavior at high strain rates (e.g., materials such as Silly Putty™), [22–24].

Saturated suspensions closely resemble the partially-saturated granular materials we are studying. Hence, one might expect that the effect of strain rate on both systems would be similar. However, the brittle-to-plastic transition seen in Fig. 1 and Fig. 2 is the exact opposite of what is generally observed with saturated suspensions and pastes. We now propose a theory to explain this transition from brittle to plastic flow behavior. This theory also predicts the lower-than-one power-law dependence of flow stress on strain rate observed in the strain-rate dependent region.

3 Theory

A dry granular material whose strength is determined by inter-particle friction will begin to fail along its weakest plane at an orientation in which the yield criteria is met. Once initiated, failure is generally confined to a narrow shear zone with the rest of the bulk material remaining relatively undeformed (Fig. 3(*a*)). This is because inter-particle friction forces are relatively strain-rate independent. Also, the kinematic coefficient of friction is generally lower than the static coefficient and dilation of material in the shearing layer further lowers its strength, [25].

In contrast, when a Newtonian fluid is sheared under laminar conditions the shear strain is uniformly distributed across the entire shear zone (Fig. 3(*b*)). This is because liquids have a strain-rate dependent shear stress. Hence it is not energetically favorable for all the shear to occur in one thin layer.

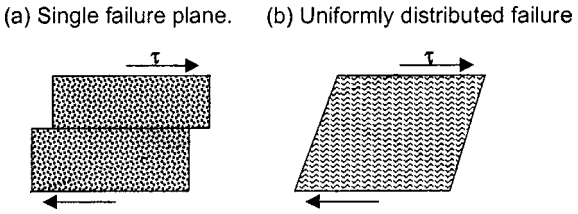


Fig. 3 Schematic of shear failure modes for (*a*) dry granular material and (*b*) Newtonian liquid in laminar flow

This change from single-layer to distributed multilayer deformation is similar to the transition observed in Figs. 1 and 2. Failure in single isolated planes will result in brittle cracking because a large amount of strain occurs in the planes of slip (Region I). However, when failure occurs over an extended zone, each layer only experiences a small portion of the bulk strain. This inhibits the formation of cracks and results in plastic flow (Region II). Hence, we propose that a bulk granular material with interstitial liquid can exhibit yield behavior similar to either a dry bulk solid or a liquid. Which behavior occurs depends on whether it is strain-rate dependent viscous forces or strain-rate insensitive capillary and friction forces that dominate.

Granular materials in unconfined compression fail by shear along some plane. In a continuum model there are potentially an infinite number of planes along which the material could fail. The normals to these planes are inclined at an angle ω to the direction of the imposed compressive force given by

$$\omega = \frac{\pi}{4} + \frac{\phi}{2}, \tag{2}$$

where ϕ is the material's internal angle of friction, [25].

The simplest yield condition is defined by the Mohr-Coulomb criteria

$$\tau_n = \sigma_n \tan \phi + c, \tag{3}$$

where τ_n is the limiting shear stress on that plane once friction is fully activated, σ_n is the normal stress on the failure plane arising from the applied load and c is the internal friction and cohesivity of the material, [25]. For a cylindrical sample subjected to uniaxial compression and laterally unrestrained, as shown in Fig. 4, a Mohr-Coulomb analysis of the stress conditions on the failure plane gives

$$\sigma_n = \frac{\sigma_{az}}{2}(1 - \sin \phi) \tag{4a}$$

and

$$\tau_n = \frac{\sigma_{az}}{2} \cos \phi, \tag{4b}$$

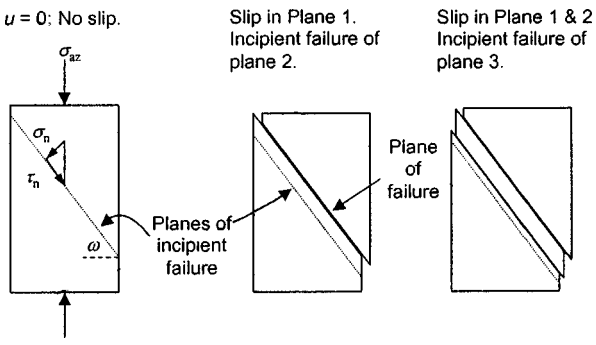


Fig. 4 Schematic of model showing how successive planes simultaneously fail as applied bulk strain rate increases

where σ_{az} is the compressive stress applied along the axis of a cylinder, [25].

For a wet granular body, the cohesivity in Eq. (3) arises predominantly from liquid adhesion forces. Surface tension effects impose an isostatic normal stress on the granular material (σ_c). To a first approximation then

$$c = \sigma_c \tan \phi. \tag{5}$$

The static cohesive strength σ_c of liquid-bound granular materials arising from the capillary force has been found to be proportional to the liquid-solid adhesion tension $\gamma_{LV} \cos \theta$ and inversely proportional to the average particle size d_p, [26]. It is also some function Ψ_1 of the liquid saturation s and particle packing fraction Φ, thus

$$\sigma_c = \frac{\gamma \cos \theta}{d_p} \psi_1(s,\Phi). \tag{6}$$

Substituting Eqs. (4*a*), (5), and (6) into (3), the static limiting shear stress on the slip plane is

$$\tau_n = \left(\frac{\sigma_{az}}{2}(1-\sin\phi) + \frac{\gamma \cos\theta}{d_p}\psi_1(s,\Phi) \right) \tan\phi. \tag{7}$$

Once failure occurs, there will also be a dynamic component of the shear stress due to viscous forces in the liquid phase. For a continuous liquid film of viscosity μ_v between two plates a distance δ apart and sheared at a speed of u, this viscous shear force would simply be

$$\tau_v = \mu \left(\frac{u}{\delta} \right). \tag{8}$$

If we assume that the average gap distance between particles is proportional to the particle size, then we may write a general expression for the viscous shear stress for shearing a plane of wet granular material as

$$\tau_v = \mu \frac{u}{d_p} \Psi_2(s,\Phi), \tag{9}$$

where $\psi_2(s,\Phi)$ is some unknown function of liquid saturation and packing fraction. The exact form of the functional relationships $\psi_1(s,\Phi)$ and $\psi_2(s,\Phi)$ is not important in this work because s and Φ were both held constant in the experiments shown in Figs. 1 and 2.

In this analysis we have assumed the pore pressure is constant. In a saturated granular system, if dilation occurs in response to shear, then the pore pressure is lowered. This low pore pressure resists particle motion and causes liquid to be sucked into the sheared region (this effect can be seen when walking on wet beach sand). At high enough strain rates, cavitation will eventually occur. However, in partially saturated granular materials, the pore pressure remains essentially constant because air can flow freely into and out of the assembly. Hence this model does not apply to saturated systems.

If the internal friction coefficient used in Eq. (7) was that valid for sustained shear (critical consolidation) of the granular material, then the static component of the shear stress (Eq. (7)) would also be valid after yield has been initiated. Then Eqs. (7) and (9) can be added to give the total shear stress on the slip plane once yield has been initiated, namely

$$\tau_t = \tau_n + \tau_v = \left(\frac{\sigma_{az}}{2}(1-\sin\phi) + \frac{\gamma_{LV}\cos\theta}{d_p}\psi_1(s,\Phi) \right) \tan\phi + \frac{\mu u}{d_p}\psi_2(s,\Phi). \tag{10}$$

At the onset of slip, the shear stress on the slip plane is given by Eq. (4*b*). Once slip is established, the normal applied stress σ_{az} becomes the flow stress σ_f such that it must overcome both frictional and viscous components to maintain flow. Hence from Eqs. (4*a*) and (10)

$$\frac{\sigma_f d_p}{\gamma \cos\theta} = [\cos\phi - \tan\phi(1-\sin\phi)]^{-1} \left(2\tan(\phi)\cdot\psi_1(s,\Phi) + 2\psi_2(s,\Phi)\frac{\mu u}{\gamma\cos\theta} \right), \tag{11}$$

which is of the form

$$\mathrm{Str}^* = a + b\mathrm{Ca}^{1.0}, \tag{12}$$

where

$$a = [\cos\phi - \tan\phi(1-\sin\phi)]^{-1}[2\tan\phi\cdot\psi_1(s,\Phi)], \tag{12a}$$

$$b = [\cos\phi - \tan\phi(1-\sin\phi)]^{-1}[2\psi_2(s,\Phi)], \tag{12b}$$

and

$$\mathrm{Ca} = (\mu u)(\gamma\cos\theta)^{-1}. \tag{12c}$$

Ca is the capillary number defined in terms of the slip velocity *in a single plane*, whereas the capillary number $\overline{\mathrm{Ca}}$ in Eq. (1) is based on the average bulk strain rate of the pellet. Equation (12) describes the strength of every potential failure plane. At steady state (whether static or in failure) the transmitted stress between all these planes of material will be equal.

Equation (12) has a similar form to Eq. (1), except that the exponent of the capillary number is one, whereas the experimentally fitted exponent was $n = 0.58 \pm 0.04$. We now introduce an assumption which allows us to resolve this apparent discrepancy between the expected and observed power-law dependency of stress on strain rate.

In any granular material there will be a distribution in the strength of planes due to irregularities in the packing and liquid distribution. Therefore, we shall assume that the magnitude of the static friction coefficient a varies between planes. For simplicity we shall take the viscous coefficient b of all planes as a constant. In reality b would also vary for the same reasons that a does, however, this assumption greatly simplifies the analysis that follows and does not alter the general conclusions of the model. Let the plane with the lowest static friction coefficient be known as plane 1. This is the plane in which failure will initiate. Let the next weakest plane be plane 2, and so on, such that $a_{i+1} > a_i$ for all i. The shear strength of plane i is then given by

$$\mathrm{Str}_i^* = a_i + b\mathrm{Ca}_i^{1.0}, \tag{13}$$

where Ca_i is the layer capillary number in plane i, based on the slip velocity u_i in that plane of material.

When a compressive load is first applied to the material, failure does not occur until the applied stress exceeds the static strength of the weakest layer. For $\mathrm{Str}^* > a_1$, the material will slip on plane 1, with the applied stress varying with velocity according to Eq. (13) with i=1.

There is a velocity below which failure will remain confined to plane 1. Above this velocity, however, the applied stress will exceed the static strength of the second weakest plane, plane 2. Failure will then start to occur simultaneously in two planes. This occurs when

$$\mathrm{Str}^* = a_1 + b\mathrm{Ca}_1 > a_2, \tag{14a}$$

or equivalently

$$\mathrm{Ca}_1 > (a_2 - a_1)/b. \tag{14b}$$

For even higher velocities, a third plane will start failing, and so on. Once more than one plane is slipping, the applied bulk shear velocity u will be distributed between all failing planes such that

$$u = \sum_i u_i. \tag{15}$$

Hence, the capillary number based on the overall bulk strain rate can be expressed as the sum of the layer capillary numbers in each plane, i.e.,

$$\overline{\mathrm{Ca}} = \sum_i \mathrm{Ca}_i . \tag{16}$$

Note that once multiple planes are in simultaneous failure, further increases in u cause the slip velocity in *all* of the failing planes to increase so that a constant transmitted stress is maintained between each layer i.e., u_1 does not remain at the value it was when plane 2 began slipping.

This model assumes a steady, non-accelerating state of failure. This requires that inertial effects be negligible. This assumption is partially supported by recent work comparing copper and glass powders which found negligible difference in flow stress for a fourfold increase in solid-phase density, [19]. Liquid density ρ_L has not been altered in any experiments to date, so the significance of inertial effects in the liquid phase has not been established. However, we may estimate the film Reynolds number in order to check whether the flow conditions are laminar or turbulent. If we assume the film thickness is of the same magnitude as particle diameter, then for the highest deformation rate studied (water-bound pellets at 150 mm/s)

$$\mathrm{Re} \approx \frac{\rho_L u d_p}{\mu} = \frac{(1000\ \ \mathrm{kg \cdot m^{-3}})(0.15\ \ \mathrm{m \cdot s^{-1}})(3\times 10^{-5}\ \ \mathrm{m})}{(0.001\ \ \mathrm{Pa \cdot s})}$$

$$= 4.5.$$

This is well within the laminar regime, so inertial effects in the fluid films between particles should be negligible.

In order to proceed further with this analysis, we need to know how the static strength varies between planes. In reality this is likely to vary in a stochastic fashion, dependent on slight variations in packing, particle orientation and shape between planes within the granular material. For simplicity we will here assume that each plane is a constant amount, Δa, stronger then the previous weakest plane i.e., $a_{i+1} = a_i + \Delta a$. Therefore

$$a_n = a_1 + (n-1)\Delta a. \tag{17}$$

Let us consider what happens when plane n is on the verge of starting to slip. This occurs when

$$\mathrm{Str}^* = a_n = a_1 + (n-1)\Delta a. \tag{18}$$

At this point $(n-1)$ planes are already simultaneously slipping. In each of these planes the relationship between Str* and Ca is given by Eq. (12). Summing the two sides of Eq. (12) for all $(n-1)$ slipping planes gives

$$(n-1)\mathrm{Str}^* = \sum_{i=1}^{n-1} a_i + b \sum_{i=1}^{n-1} \mathrm{Ca}_i . \tag{19}$$

Evaluating the sum of the arithmetic series a_1 to a_{n-1}, dividing both sides by $(n\text{-}1)$ and substituting the global capillary number for the sum of the layer capillary numbers (16) gives

$$\mathrm{Str}^* = a_1 + \frac{(n-2)\Delta a}{2} + b\frac{\overline{\mathrm{Ca}}}{(n-1)}. \tag{20}$$

Equating Eqs. (18) and (20) and rearranging gives

$$\overline{\mathrm{Ca}} = \frac{n(n-1)\Delta a}{2b}. \tag{21}$$

Equations (18) and (21) give the applied stress and the bulk capillary number, respectively, when plane n is about to fail.

These results can now be used to interpret the discrepancy between the expected and observed power-law dependency of Str* on Ca in the viscous controlled regime. When the bulk strain rate is increased, the number of layers in simultaneous failure increases. Hence the percentage increase of strain rate in individual layers is less than the overall percentage increase in bulk strain rate because the bulk strain is distributed over a greater number of layers than previously. In the limit as $n \to \infty$, $\overline{\mathrm{Ca}} \propto n^2$ and $\mathrm{Str}^* \propto n$. Therefore $\mathrm{Str}^* \propto \overline{\mathrm{Ca}}^{0.5}$ at high strain rates. This is illustrated in Fig. 5 using arbitrary values for the model parameters a_1, Δa and b. The single failure plane model has a power law exponent of one in the viscous region, whereas the multi-layer model has an exponent of 0.5. This is very close to the experimentally observed value of 0.58 ± 0.04, which is encouraging given the very simplistic assumptions of constant Δa and constant b that were made. When all the strain occurs in single isolated failure planes, cracks develop. However, when the strain is distributed over a large number of planes, the actual amount of strain in each plane is lessened, which inhibits crack growth and results in plastic behavior.

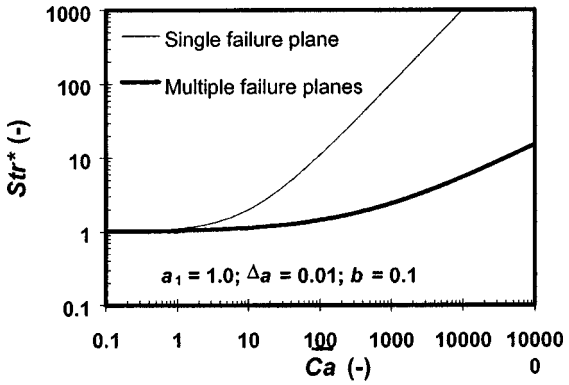


Fig. 5 Flow stress as a function of capillary number for the single-failure-plane model, Eq. (12), and the multiple-failure-plane model, Eqs. (18) and (21), using $a_1 = 1.0$, $\Delta a = 0.01$, and $b = 0.1$

4 Discussion and Conclusions

It has been shown that partially saturated liquid-bound granular material can exhibit the dynamic flow behavior of either a dry bulk solid or a liquid. Which behavior dominates depends on the ratio of liquid viscous forces to capillary forces, which can be represented by the capillary number. Inter-particle friction, controlled by the capillary forces, dominates yield behavior at low strain rates and results in dry-powder-like brittle behavior. Viscous forces dominate at high strain rates and result in liquid-like plastic flow. In effect, we have identified a strain-rate-dependent phase transition in the behavior of partially saturated granular systems. It is noteworthy that this is the opposite of the transition from plastic to brittle behavior with increasing strain rate that is usually seen in fully saturated slurries and pastes.

A model has been proposed to explain the functional role of binder viscosity in influencing the dynamic flow stress and deformation characteristics of wet granular materials. It is hypothesized that when viscous forces become significant there are multiple planes of material in simultaneous failure. This model successfully explains the observed brittle-to-plastic transition in pellet flow behavior and also shows good qualitative agreement with the experimentally observed dependence of compressive flow stress on the capillary number.

One important practical implication of these findings is that measurements performed on wet granular material at low strain rates cannot be used to draw conclusions about their strength or mode of failure at high strain rates. A strain sufficient to cause a liquid-bound granule to crack and fall apart at low strain rates may not cause any breakage at high strain rates. The applied force required to cause a given amount of strain may also be much higher. Hence, viscous effects must be considered when modelling the deformation and breakage of such material during impacts or other high strain rate events.

Acknowledgments

This work has been funded by an Australian Research Council Large Grant and S. Iveson is partially supported by an Australian Postdoctoral Fellowship. The authors would also like to thank Dr. Mark Hentschell and Dr. George Franks for their helpful comments during proofreading.

References

[1] Ouchiyama, N., and Tanaka, T., 1975, ''The Probability of Coalescence in Granulation Kinetics,'' Ind. Eng. Chem. Process Des. Dev., **14**, pp. 286–289.
[2] Thornton, C., and Ning, Z., 1998, ''A Theoretical Model for the Stick/Bounce Behavior of Adhesive, Elastic-Plastic Spheres,'' Powder Technol., **99**, pp. 154–162.
[3] Liu, L. X., Litster, J. D., Iveson, S. M., and Ennis, B. J., 2000, ''Coalescence of Deformable Granules in Wet Granulation Processes,'' AIChE J., **46**, pp. 529–539.
[4] Iveson, S. M., and Litster, J. D., 1998, ''Growth Regime Map for Liquid-Bound Granules,'' AIChE J., **44**, pp. 1510–1518.
[5] Irfan Khan, M., and Tardos, G. I., 1997, ''Stability of Wet Agglomerates in Granular Shear Flows,'' J. Fluid Mech., **347**, pp. 347–368.
[6] Iveson, S. M., Litster, J. D., Hapgood, K., and Ennis, B. J., 2001, ''Nucleation, Growth and Breakage Phenomena in Agitated Wet Granulation Processes: A Review,'' Powder Technol., **117**, pp. 3–39.
[7] Rumpf, H., 1962, ''The Strength of Granules and Agglomerates,'' *AIME, Agglomeration*, W. A. Knepper, ed., Interscience, New York, pp. 379–418.
[8] Schubert, H., Herrmann, W., and Rumpf, H., 1975, ''Deformation Behavior of Agglomerates Under Tensile Stress,'' Powder Technol., **11**, pp. 121–131.
[9] Kristensen, H. G., Holm, P., and Schæfer, T., 1985, ''Mechanical Properties of Moist Agglomerates in Relation to Granulation Mechanisms, Part 1: Deformability of Moist, Densified Agglomerates,'' Powder Technol., **44**, pp. 227–238.
[10] Benbow, J. J., Jazayeri, S. H., and Bridgewater, J., 1991, ''The Flow of Pastes Through Dies of Complicated Geometry,'' Powder Technol., **65**, pp. 393–401.
[11] Franks, G. V., and Lange, F. F., 1999, ''Plastic Flow of Saturated Alumina Powder Compacts: Pair Potential and Strain Rate,'' AIChE J., **45**, pp. 1830–1835.
[12] Cakmak, A. S., and Herrera, I., eds., 1989, ''Soil Dynamics and Liquefaction,'' *Proceedings of the 4th International Conference on Soil Dynamics and Earthquake Engineering*, Mexico City, Mexico, Oct., Computational Mechanics, Southhampton, UK.
[13] Deysarkar, A. K., and Turner, G. A., 1980, ''The Effect of Vibrations on the Flow Properties of a Saturated Paste of Iron Ore and Water,'' Int. J. Min. Process., **6**, pp. 257–276.
[14] Iveson, S. M., and Litster, J. D., 1998, ''Liquid-Bound Granule Impact Deformation and Coefficient of Restitution,'' Powder Technol., **99**, pp. 234–242.
[15] Lian, G., Thornton, C., and Adams, M. J., 1998, ''Discrete Element Simulation of Agglomerate Impact Coalescence,'' Chem. Eng. Sci., **53**, pp. 3381–3391.
[16] Mills, P. J. T., Seville, J. P. K., Knight, P. C., and Adams, M. J., 2000, ''The Effect of Binder Viscosity on Particle Agglomeration in a Low Shear Mixer/Agglomerator,'' Powder Technol., **113**, pp. 140–147.
[17] Iveson, S. M., Beathe, J. A., and Page, N. W., 2002, ''The Dynamic Strength of Partially Saturated Powder Compacts: The Effect of Liquid Properties,'' Powder Technol., **127**, pp. 149–161.
[18] Iveson, S. M., and Page, N. W., 2002, ''Dynamic Mechanical Properties of Liquid-Bound Powder Compacts,'' 3rd Australasian Congress on Applied Mechanics (ACAM 2002), Sydney, Australia, Feb. 20–22.
[19] Iveson, S. M., and Page, N. W., 2002, ''Dynamic Strength of Partially-Saturated Powder Compacts: Effects of Particle Shape and Density,'' World Congress of Particle Technology 4 (WCPT4), Sydney, Australia, July 21–25.
[20] Ravi-Chandar, K., Lu, J., Yang, B., and Zhu, Z., 2000, ''Failure Mode Transition in Polymers Under High Strain Rate Loading,'' Int. J. Fract., **101**, pp. 33–72.
[21] Gensler, R., Plummer, C. J. G., Grein, C., and Kausch, H. H., 2000, ''Influence of the Loading Rate on the Fracture Resistance of Isotactic Polypropylene and Impact Modified Isotactic Polypropylene,'' Polymer, **41**, pp. 3809–3819.
[22] Levenspiel, O., 1984, *Engineering Flow and Heat Exchange*, Plenum, New York.
[23] Smith, J. V., 1997, ''Shear Thickening Dilatancy in Crystal-Rich Flows,'' J. Volcanol. Geotherm. Res., **79**, pp. 1–8.
[24] Franks, G. V., Zhou, Z. W., Duin, N. J., and Boger, D. V., 2000, ''Effect of Interparticle Forces on Shear Thickening of Oxide Suspensions,'' J. Rheol., **44**, pp. 759–779.
[25] Nedderman, R. M., 1992, *Statics and Kinematics of Granular Materials*, Cambridge University Press, Cambridge, UK.
[26] Newitt, D. M., and Conway-Jones, J. M., 1958, ''A Contribution to the Theory and Practice of Granulation,'' Trans. Inst. Chem. Eng., **36**, pp. 422–441.

Brad Buckham
Department of Mechanical Engineering,
University of Victoria,
P.O. Box 3055,
Victoria, BC V8W 3P6, Canada
e-mail: bbuckham@uvic.ca

Frederick R. Driscoll
Department of Ocean Engineering,
Florida Atlantic University,
101 North Beach Road,
Dania Beach, FL 33004
e-mail: rdriscol@oe.fau.edu

Meyer Nahon
Department of Mechanical Engineering,
McGill University,
817 Sherbrooke Street West,
Quebec H3A 2K6, Canada
e-mail: Meyer.Nahon@mcgill.ca

Development of a Finite Element Cable Model for Use in Low-Tension Dynamics Simulation

To accurately simulate the motion of slack marine cables, it is necessary to capture the effects of the cable's bending and torsional stiffness. In this paper, a computationally efficient and novel third-order finite element is presented that provides a representation of both the bending and torsional effects and accelerates the convergence of the model at relatively large element sizes. Using a weighted residual approach, the discretized motion equations for the new cubic element are developed. Applying inter-element constraint equations, we demonstrate how an assembly of these novel elemental equations can be significantly reduced to prevent the growth of the system equations normallly associated with such higher order elements and allow for faster evaluation of the cable dynamics in either taut or low-tension situations.

1 Introduction

The use of numerical models to simulate the dynamics of underwater vehicles and moorings has been presented extensively in existing literature. Numerical models provide a means for the analysis of towed cable dynamics, [1–5], deep water moorings, [6–8], and marine risers, [9–14], prior to the actual development and/or deployment of the system. In addition, accurate dynamics models have proven to be effective mission planning tools, having been used to study both towcable management and turning strategies for towed underwater vehicles, [15–20] and in the design of systems to provide heave compensation for tethered deep-sea ROV platforms, [21,22]. An application that has not yet seen significant attention is the time domain simulation of underwater remotely operated vehicles (ROVs) with slack tethers.

Despite advances in autonomous technology, ROVs are still the predominant tool for complex intervention tasks. These vehicles typically follow omnidirectional paths during operation and deploy a twisted lay of low-tension (or slack) neutrally buoyant tether, or umbilical cable, along the path. Travel to the limits of the tether, sudden movements of the ROV and/or environmental loads can cause the tether to become taut. During such tensioning, the ROV response is dominated by the rate of and direction of tensioning, [23]. Given the complexity of the tasks that are attempted using ROVs, the ability to simulate an ROV deployment *a priori* is a valuable asset. In order to extend the capabilities of numerical dynamics models to include the simulation of ROV platforms, including the motion of the slack umbilical tether, it is necessary to develop an accurate representation of the higher order internal effects that contribute significantly to the motion of the slackened tether.

To date, efforts to model underwater vehicles with slack tethers have included a two-dimensional implementation by Grosenbaugh et al. [24], and a three-dimensional model formulated by Banerjee and Do [25]. Both works provide a representation of the tether's bending stiffness. In a previous work, the authors presented a first-order lumped parameter model to simulate a three-dimensional ROV tether, [26]. By estimating the curvature of the tether from the nodal positions, bending effects were applied to the linear elements of this model as equivalent transverse nodal forces. The experimental results of this work showed the importance of including a bending stiffness model in the slack tether dynamics. However, the linear elements of this approach, like the methods of [24] and [25], require a fine spatial discretization of the tether to capture any rapid variations of curvature along the tether, resulting in a potentially large computational overhead. To alleviate this problem, higher order curvilinear elements could be used to better approximate the twisted lay of a tether with a larger spatial discretization. To date, third order finite element techniques have been applied to the analysis of mooring lines, marine risers, and flexible pipelines, [9–12]. Depending on the particular element formulation, the state of each finite element presented in these works is defined by a vector of 12 or 13 state variables, as compared to the six state variables required for the first-order lumped parameter elements. This increased number of state variables detracts from the accuracy of the higher order element, and it is therefore desirable to develop a higher order element with a reduced number of state variables.

This work develops a novel finite element cable model based on a third-order element with a reduced-order state vector. Following the developments of Garret [12] and Nordgren [13,14], we develop the non-linear motion equations for the continuous tether in terms of an inertial frame of reference using the Frenet equations. We show how these motion equations can be discretized using Galerkin's method of weighted residuals to produce the system of ordinary differential equations, element equations, that govern the motion of the chosen cubic element. We will show how these elemental equations are assembled to form a global set of equations which define the motion of the entire tether. Finally, we will demonstrate that our particular choice of finite element allows the assembled system of equations to be reduced into a system that has the same dimension as the simpler linear lumped parameter model, but yet maintains an accurate and complete representation of the higher order bending and torsional effects.

2 Equations of Motion

To create a numerical model that includes the desired bending and torsional effects it is necessary to derive the dynamics equa-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 11, 2002; final revision, October 8, 2003. Associate Editor: N. Triantafyllidis. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

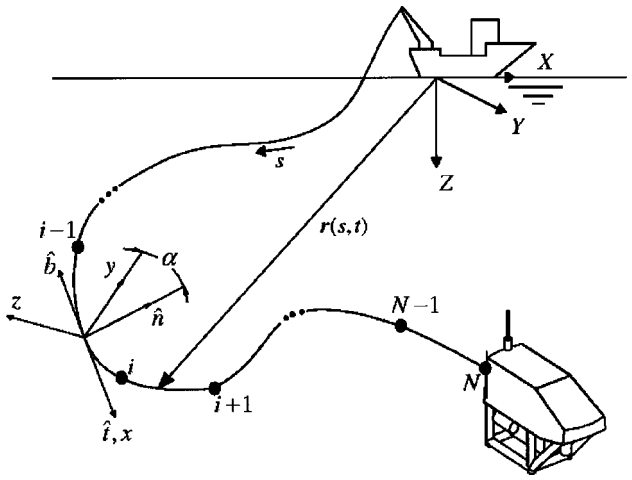


Fig. 1 A diagrammatic presentation of the coordinate systems, the Frenet and body-fixed frames, used to describe the tether element. The discretized tether is formed from an assembly of cubic elements, with the *i*th element extending between the *i*th and *i*+first nodes.

tions for a continuous tether, considering the tortuous profile that a tether forms in three-dimensional space. In existing literature, the cable is modeled as a slender flexible rod that sustains environmental, gravitational and buoyancy forces. Of particular relevance to this work was Nordgren's presentation of the vector equations of motion for a continuous cable, [13], (Figs. 1 and 2) which are

$$\boldsymbol{F}' + \boldsymbol{q} = \left(\frac{1}{4}\pi d^2 \rho_C\right)\ddot{\boldsymbol{r}} \tag{1}$$

$$\boldsymbol{M}' + \boldsymbol{r}' \times \boldsymbol{F} + \boldsymbol{m} = \boldsymbol{0}. \tag{2}$$

where ρ_C is the density of the cable, including the water entrained in the voids between the strands of the cable armor where appropriate, d is the diameter of the cable, $\boldsymbol{r}(s,t)$ is a position vector describing the space curve formed by the center line of the cable, $\boldsymbol{q}$ is the vector of applied forces per unit length, $\boldsymbol{m}$ represents the applied moment per unit length, $\boldsymbol{F}$ is the vector of internal forces, $\boldsymbol{M}$ is the vector of internal moments, $\mathbf{I}$ is a 3×3 identity matrix, $(')$ denotes differentiation with respect to s, the unstretched curvilinear coordinate along the cable, and $(\dot{})$ denotes differentiation with respect to time t. We express $\boldsymbol{r}(s,t)$ in terms of an inertial frame of reference $(X\ Y\ Z)$, where X and Y point in perpendicular horizontal directions, and Z is aligned with gravity. The rotational inertia of the cable is considered to be much smaller than the other terms of Eq. (2), and is thus neglected, [1]. The dynamics equations are formulated using two local frames of reference, the Frenet frame and a frame of reference oriented with the principal axes of the cable, Fig. 1.

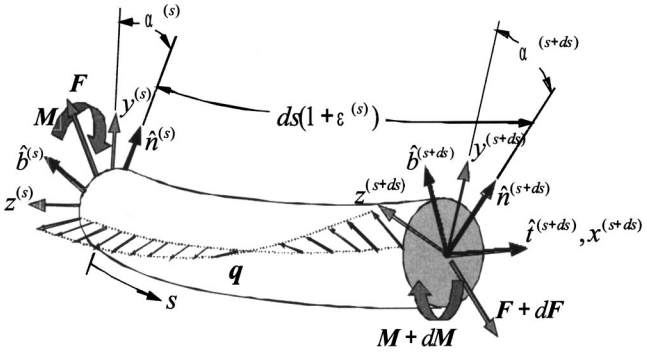


Fig. 2 A close up view of a differential segment of the ROV tether. The distributed load *q* contains weight, buoyancy and hydrodynamic loads. An additional degree-of-freedom, α, defines the orientation of the body-fixed frame relative to the Frenet frame.

The Frenet frame $(\hat{t}\ \hat{n}\ \hat{b})$ is oriented with the space curve formed by $\boldsymbol{r}(s,t)$. As the curve $\boldsymbol{r}(s,t)$ is traced, the changing orientation of the Frenet frame is quantified by two parameters: the curvature, κ, and the torsion, γ. Assuming $\varepsilon \ll 1$, the base unit vectors are defined as: $\hat{t} = \boldsymbol{r}'$, $\hat{n} = \boldsymbol{r}''/\kappa$, and $\hat{b} = \boldsymbol{r}' \times \boldsymbol{r}''/\kappa$ and their gradients are defined as $\hat{t}' = \kappa \cdot \hat{n}$, $\hat{n}' = \gamma \cdot \hat{b} - \kappa \cdot \hat{t}$, and $\hat{b}' = -\gamma \cdot \hat{n}$. The curvature, κ, defines the bend of the cable within an osculating plane that is formed by $\hat{t}$ and $\hat{n}$ at the point considered. Both κ and γ are defined in terms of the spatial derivatives of the space curve [27]:

$$\kappa = (\boldsymbol{r}'' \cdot \boldsymbol{r}'')^{1/2}, \quad \gamma = \frac{\boldsymbol{r}' \cdot (\boldsymbol{r}'' \times \boldsymbol{r}''')}{\boldsymbol{r}'' \cdot \boldsymbol{r}''}.$$

The torsion of the curve, γ, is the spatial rate of change of the osculating plane's orientation about the tangent vector, and consequently the orientation of the bend, about the tangent vector, $\hat{t}$. As such, the torsion represents the twist experienced within the tether due to the shape of the tether profile. However, when considering twist of the tether, one recognizes that torsional couples applied at the boundaries of a tether section create additional angular displacement, α (the torsion deformation), throughout the tether scope. Thus, we introduce the body fixed reference frame $(x\ y\ z)$ which remains aligned with the principal axes of the tether cross section, as shown in Figs. 1 and 2, and is separated from the Frenet frame by an angle α about the tangent direction.

As discussed by Love, the internal moment, $\boldsymbol{M}$, generated at any point within the tether is proportional to the curvature and the twist observed at the point [28]:

$$\boldsymbol{M} = EI\kappa\hat{b} + GJ\tau\hat{t} \tag{3}$$

where the overall twist, τ, of the cable at the point is given by

$$\tau = \gamma + \alpha'.$$

Following [12] and [13], and assuming that there are no external moments applied between the tether boundaries ($\boldsymbol{m} = \boldsymbol{0}$), Eq. (3) is substituted into Eq. (2) to produce a definition for the vector of internal forces at any location along the tether

$$\boldsymbol{F} = T \cdot \hat{t} - EI\kappa' \cdot \hat{n} - EI\kappa \cdot \gamma \cdot \hat{b} + GJ\tau \cdot \kappa \cdot \hat{b}. \tag{4}$$

The spatial derivative κ' is a nonlinear function of the spatial coordinate s and complicates the finite element procedure to follow. The second term on the right-hand side of (4) can be expressed as

$$EI\kappa' \cdot \hat{n} = (EI\kappa \cdot \hat{n})' - EI\kappa \cdot \gamma \cdot \hat{b} + EI\kappa^2 \cdot \hat{t} \tag{5}$$

which, eliminates κ' in (4). Thus, substituting Eq. (5) into Eq. (4) yields

$$\boldsymbol{F} = -(EI\kappa\hat{n})' + (T - EI\kappa^2)\hat{t} + GJ\tau\kappa\hat{b} \tag{6}$$

where the term $-EI\kappa^2$ is included to cancel the spurious component of the internal force induced when considering the tangential component of changing $\hat{n}$. The internal axial force, T, is defined using the constitutive relationship

$$T = EA\varepsilon + C_{ID}\dot{\varepsilon}, \quad \varepsilon = [(\boldsymbol{r}' \cdot \boldsymbol{r}')^{1/2} - 1] \tag{7}$$

where ε is the strain and C_{ID} is an internal viscous damping coefficient. This simple viscous damping model is used to represent the dissipation of energy within the tether via friction between the layers of conductors, armor, and polymer coatings of a typical tether. Equation (6) defines the internal force as an explicit function of the tether's elastic deformation as defined by the axial

strain, curvature, and twist. Substituting Eq. (6) into Eq. (1) reduces the two original vector equations of motion to a single vector equation that defines the tether's translational motion in terms of these three elastic deformations, and a scalar constraint equation that defines the variation of the twist along the cable:

$$-(EI\mathbf{r}'')''+[(T-EI\kappa^2)\mathbf{r}']'+[GJ\tau(\mathbf{r}'\times\mathbf{r}'')]'+\boldsymbol{q}=\left(\frac{1}{4}\pi d^2\rho_C I\right)\ddot{\mathbf{r}}$$

$$(GJ\tau)'=0. \tag{8}$$

In this work, the applied force vector, $\boldsymbol{q}$, consists of the distributed weight, buoyancy, and hydrodynamic force, which is defined using Morison's well-known approximation. Since the Keulegan-Carpenter number of ocean cables is always very large, Morison's approximation is valid.

$$\boldsymbol{q}=(\rho_C-\rho_W)\frac{\pi(d)^2}{4}\boldsymbol{g}-1/2\rho_W d\mathbf{L}_{IH}\begin{bmatrix} f_p C_D v_{\hat{p}_1}\sqrt{v_{\hat{p}_1}^2+v_{\hat{p}_2}^2} \\ f_p C_D v_{\hat{p}_2}\sqrt{v_{\hat{p}_1}^2+v_{\hat{p}_2}^2} \\ f_q C_D v_{\hat{q}}|v_{\hat{q}}| \end{bmatrix} -\mathbf{L}_{IH}\mathbf{M}_A\mathbf{L}_{IH}^{\mathrm{T}}\ddot{\mathbf{r}} \tag{9}$$

where $\boldsymbol{g}=[0\ 0\ g]^{\mathrm{T}}$ is the gravitational acceleration, ρ_W is the density of sea water, $\mathbf{L}_{IH}$ is a transformation matrix relating the axes of a hydrodynamic frame, to be described shortly, to the inertial frame of reference, C_D is the normal drag coefficient, $\mathbf{M}_A$ is the diagonal added mass matrix, expressed in terms of the hydrodynamic frame. The relative velocity of the cable through a water column with a current $\boldsymbol{j}$ is $\mathbf{v}=\mathbf{L}_{IH}^{\mathrm{T}}(\dot{\boldsymbol{r}}-\boldsymbol{j})$ where $v_{\hat{p}_1}$, $v_{\hat{p}_2}$, and $v_{\hat{q}}$ are the components of this velocity in the hydrodynamic frame of reference.

The hydrodynamic frame ($\hat{p}_1\ \hat{p}_2\ \hat{q}$) attached to the cable segment, composed of a $\hat{q}$ axis that is tangent to the cable, an $\hat{p}_1$ axis that is normal to the segment and a $\hat{p}_2$ axis that completes the right-handed coordinate system. The hydrodynamic frame is used to resolve the relative fluid velocity into components normal and tangential to the tether such that the hydrodynamic loads can be calculated based on the loading coefficients f_p and f_q. These coefficients are functions of the incidence angle of the relative fluid flow on the tether element and account for the nonlinear breakup of drag between normal and tangential directions, respectively, as discussed by Folb and Nelligan [29]. Since the calculated hydrodynamic load is transformed back into inertial representation for application in the motion equation of Eq. (8), the particular orientation of the $\hat{p}_1$ and $\hat{p}_2$ axes about the tangent direction is inconsequential. The transform between the axes of the hydrodynamic and inertial frames, $\mathbf{L}_{IH}$, is formed from a Z-Y'-X'' ($\psi\ \theta\ \xi$) set of Euler angles. These successive rotations bring the inertial Z-axis into alignment with the tangent direction of the cable segment, and the orientation of the $\hat{p}_1$ and $\hat{p}_2$ axes is constrained by setting one of the Euler angles, the initial ψ rotation about the inertial Z-direction, to zero. Thus,

$$\mathbf{L}_{IH}=\begin{bmatrix} \cos\theta & \sin\theta\sin\xi & \sin\theta\cos\xi \\ 0 & \cos\xi & -\sin\xi \\ -\sin\theta & \cos\theta\sin\xi & \cos\theta\cos\xi \end{bmatrix}$$

where θ and ξ can be considered as traditional pitch and roll of the cable.

The added mass values in $\mathbf{M}_A$ are zero for flow acceleration parallel to $\hat{q}$ and equal to the displaced mass of water for flow acceleration in the $\hat{p}_1$ and $\hat{p}_2$ directions, [30].

$$\mathbf{M}_A=\frac{1}{4}\pi d_C^2\rho_W\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

3 Finite Element Formalization

Equations (8) define the motion of a continuous tether in terms of space, s, and time, t. The true solution to this system of equations is a vector function $\boldsymbol{r}(s,t)$ and a scalar function $\alpha(s,t)$ that are defined over the entire tether scope and the time domain of the modeled ROV maneuver. Because of the tether's geometrically nonlinear profile, the strong nonlinearities contained in the hydrodynamic term, and the wide range of tether states that can occur during ROV operation, we look to find an approximation to the true solution of the differential equations of Eq. (8) using numerical techniques. The calculation of this approximate solution, or trial solution, using a finite element approach is accomplished in two stages:

1. We consider a small section, or element, of the tether and construct two trial solutions, $\boldsymbol{r}_i(s,t)$ and $\alpha_i(s,t)$, that are formed from interpolating polynomials in terms of the spatial coordinate s. These trial solutions apply over the length of this element, and the coefficients of these polynomials are chosen as the values and spatial derivatives of $\boldsymbol{r}(s,t)$ and $\alpha(s,t)$ at the end nodes of the element.
2. We substitute these trial solutions into Eq. (8) and apply an optimization criterion that transforms the continuous differential equations into a series of algebraic equations that defines the optimum coefficients of these two polynomial approximations.

3.1 Selecting the Trial Solution. As shown in Fig. 1, we consider the tether to be a contiguous set of N cubic segments, or elements, which have the same physical properties as the continuous cable. The trial solution, $\boldsymbol{r}_i(s,t)$, defines the three-dimensional profile of the ith, tether element which extends between the nodes $i-1$ and i of the discrete tether:

$$\boldsymbol{r}(s,t):s\in[s^{(i-1)},s^{(i)}]\approx\boldsymbol{r}_i=\sum_{j=1}^{p+1}\boldsymbol{\sigma}_{i,j}(t)\phi_{i,j}(s).$$

The coefficients $\boldsymbol{\sigma}_{i,j}$ are time-dependent generalized displacements, $\phi_{i,j}$ is a shape function that defines how the jth generalized displacement contributes to the shape of the ith tether element, and p is the order of the trial solution. In order to ensure that the concatenation of the elements maintains $C^1(s)$, continuity or smoothness across the node points, $\boldsymbol{r}_i$ is often chosen as a cubic Hermite interpolating polynomial, [12]. The generalized displacements then become the position, $\boldsymbol{r}^{(i-1)}$ and $\boldsymbol{r}^{(i)}$, and tangent vectors, $\boldsymbol{r}'^{(i-1)}$ and $\boldsymbol{r}'^{(i)}$, observed at the $i-1$ and i node points. In this case, twelve state variables are required to fully define the ith element: the components of $\boldsymbol{r}^{(i-1)}$, $\boldsymbol{r}'^{(i-1)}$, $\boldsymbol{r}^{(i)}$, and $\boldsymbol{r}'^{(i)}$.

Recalling that we assume no moments are applied along the tether, rather the tether is suspended in the water column and is subject to only distributed weight, buoyancy and hydrodynamic forces, we expect the curvature to vary continuously over the entire scope of the tether. Referring to Eq. (3) we recognize that this loading scenario produces a tether profile that exhibits $C^2(s)$ continuity. To replicate this behavior within the finite element model we model the tether with an assembly of finite elements that possess an additional degree of continuity relative to the cubic Hermite element. The twisted cubic spline element provides this level of continuity, [31]. We define the element profile, $\boldsymbol{r}_i$, as a twisted cubic spline segment:

$$\boldsymbol{\sigma}_{i,1}=\boldsymbol{r}^{(i-1)}\quad\boldsymbol{\sigma}_{i,2}=\boldsymbol{r}''^{(i-1)}\quad\boldsymbol{\sigma}_{i,3}=\boldsymbol{r}^{(i)}\quad\boldsymbol{\sigma}_{i,4}=\boldsymbol{r}''^{(i)}$$

$$\phi_{i,1}=\frac{s^{(i)}-s}{s^{(i)}-s^{(i-1)}},\phi_{i,2}=\frac{1}{6}(\phi_{i,1}^3-\phi_{i,1})L_i^2,$$

$$\phi_{i,3}=\frac{s-s^{(i-1)}}{s^{(i)}-s^{(i-1)}},\phi_{i,4}=\frac{1}{6}(\phi_{i,3}^3-\phi_{i,3})L_i^2 \tag{10}$$

where $(\,)^{(i-1)}$ and $(\,)^{(i)}$ represent evaluation at the left and right end nodes of element i, respectively, $L_i=s^{(i)}-s^{(i-1)}$ is the unstretched length of the ith element. Referring to Eq. (10), it is apparent that our trial solution, $\boldsymbol{r}_i$, is formed from a superposition of a linear component that interpolates the node positions and a cubic refinement that interpolates the curvatures at the node points and thus ensures second order continuity between the elements. Of course, it remains to ensure that a concatenation of the elements exhibits smoothness, or first-order continuity. Later we will show how this additional constraint on the element geometry is enforced.

To complete the trial solution, we model the torsional deformation within the element as a linearly varying quantity. This approach is a standard, and widely accepted, practice in structural finite element analysis that relies on the premise that the torsional deformations are limited and not fluctuating rapidly over the body, [32,33]. Thus, we use the $\phi_{i,1}(s)$ and $\phi_{i,3}(s)$ shape functions to define an approximate variation of the torsional deformation throughout the ith element, α_i:

$$\alpha_i=\alpha^{(i-1)}\phi_{i,1}(s)+\alpha^{(i)}\phi_{i,3}(s).$$

3.2 Application of the Galerkin Criterion. Given the trial solution, composed of both $\boldsymbol{r}_i$ and α_i, the solution of the dynamics equations given in Eq. (8) over the domain of the ith element is reduced to the problem of defining the coefficients of these two polynomials at any given instant in time. To transform the partial differential equation of Eq. (8) into a series of algebraic equations that define these coefficients, we apply Galerkin's method of weighted residuals, [33,34]. At this point it is important to outline how the choice of the twisted spline element affects this discretization process.

Referring to the vector equation of motion given in Eq. (8), one can see that the highest order differential of $\boldsymbol{r}(s,t)$ is of order four. The most common approach to the discretization of such a system is to integrate this differential equation by parts twice, [34]. This procedure ensures that the natural boundary conditions (the conditions on the highest derivatives of the element profile $\boldsymbol{r}''$ and $\boldsymbol{r}'''$) remain unconstrained, and that the essential boundary conditions (the conditions on $\boldsymbol{r}$ and $\boldsymbol{r}'$) are constrained by the element's mathematical definition. The benefit of this approach is that only $C^1(s)$ continuity is required of the trial solution $\boldsymbol{r}_i$, leading to the use of the cubic Hermite element in prior developments. In these more conventional developments, the second integration by parts step introduces moment terms in the boundary conditions developed for the $\mathbf{r}''$ derivatives. It is the specification of these moments at the element boundaries that provides the unconstrained conditions on the higher order derivatives.

The premise of this work is to take advantage of the fact that the elastic body being modeled should demonstrate a continuous second derivative, $\boldsymbol{r}''$, due to the type of loading it is subjected to (distributed environmental loads with no externally applied moments along the tether length). For this reason we have introduced an element with $C^2(s)$ continuity which, through its definition given in Eq. (24), provides a constraint on the curvature, or second derivative $\boldsymbol{r}''$, observed at the element nodes. Therefore, in applying Galerkin's criterion with this new element type, only the value of the third derivative at the element boundaries, $\boldsymbol{r}'''$, must remain unconstrained. Thus, we formulate our residual equations by integrating by parts one time. This produces the following 14 scalar residual equations:

$$\begin{aligned}\int_{s^{(i-1)}}^{s^{(i)}}&\Big((EI\boldsymbol{r}_i'''-[(T-EI\kappa^2)_i\boldsymbol{r}_i']-[GJ\tau_i(\boldsymbol{r}_i'\times\boldsymbol{r}_i'')])\phi_{i,j}'\\&+\Big(\boldsymbol{q}_i-\Big(\frac{1}{4}\pi d^2\rho_C\mathbf{I}\Big)\ddot{\boldsymbol{r}}_i\Big)\phi_{i,j}\Big)ds\\&=[((EI\boldsymbol{r}_i''')-(T-EI\kappa^2)_i\boldsymbol{r}_i'-(GJ\tau_i(\boldsymbol{r}_i'\times\boldsymbol{r}_i'')))\phi_{i,j}]\big|_{s^{(i-1)}}^{s^{(i)}},\\&\qquad j=1,2,3,4\end{aligned} \tag{11}$$

$$GJ\int_{s^{(i-1)}}^{s^{(i)}}(\gamma_i+\alpha_i')\phi_{i,j}'ds=[GJ\tau_i]\phi_{i,j}\big|_{s^{(i-1)}}^{s^{(i)}},\quad j=1,3. \tag{12}$$

We have assumed that the material properties are constant over the element length. Referring back to Eq. (6), it can be shown that the boundary term in Eq. (11) evaluates to

$$\begin{aligned}&[((EI\boldsymbol{r}_i''')-(T-EI\kappa^2)_i\boldsymbol{r}_i'-(GJ\tau_i(\boldsymbol{r}_i'\times\boldsymbol{r}_i'')))\phi_{i,j}]\big|_{s^{(i-1)}}^{s^{(i)}}\\&\quad=[(-\boldsymbol{F}_i)\phi_{i,j}]\big|_{s^{(i-1)}}^{s^{(i)}}.\end{aligned}$$

From Eq. (11) it is apparent that enforcing $C^2(s)$ continuity of the element, and thus eliminating the second integration by parts step, has eliminated any boundary terms involving moments from the analysis. Rather, we are left with boundary loads that are defined strictly in terms of the internal forces at the node points.

3.3 Evaluation of the Residual Equations. In order to evaluate Eqs. (11) and (12) numerically, it remains to evaluate the integrals of Eqs. (11) and (12) over the spatial domain of the ith tether element to produce the algebraic form of the residual equations. These integrals contain nonlinear functions of the spatial coordinate s and to complete the integration of Eq. (11) and Eq. (12) it is necessary to approximate these nonlinear terms with integrable functions. Since most cables are locally stiff in tension, the rate of change of these spatial functions is small and can be approximated by low-order polynomial functions. We approximate the axial force within the ith element as a constant quantity, $\lambda_i(s)$:

$$(T-EI\kappa^2)_i\approx\lambda_i(s)=(T_i-EI(\mathbf{r}''\cdot\mathbf{r}'')^{(i-1/2)}) \tag{13}$$

where $(\cdot)^{(i-1/2)}$ indicates evaluation at the midpoint of the ith element, and the internal elastic tension and damping forces within the ith element, T_i, are calculated following Eq. (7) using average values of strain and strain rate over the element, ε_i and $\dot{\varepsilon}_i$, respectively. To evaluate the average strain of the ith curvilinear element, a Romberg integration algorithm presented in [35] is adapted to approximate the integral expression:

$$\varepsilon_i=\frac{\Big(\int_{s^{(i-1)}}^{s^{(i)}}\mathbf{r}_i'\cdot ds-L_i\Big)}{L_i}.$$

To complete the calculation of the internal damping effects, the time rate of change of the element strain is calculated using the two end node velocities. The average element velocity, $\boldsymbol{v}_{RB}^i$, is subtracted from the velocities of the $i-1$th and ith node points leaving only the velocity components which create change in the element shape. The quantity $\dot{\varepsilon}_i$ is calculated by finding the portion of the remaining velocity vectors that lie in the tangential direction at each node and determining the net axial stretch that is induced. That is

$$\dot{\varepsilon}_i = \frac{(\boldsymbol{v}^{(i)} - \boldsymbol{v}^i_{RB}) \cdot \hat{t}^{(i)} - (\boldsymbol{v}^{(i-1)} - \boldsymbol{v}^i_{RB}) \cdot \hat{t}^{(i-1)}}{L_i}.$$

The hydrodynamic component of the distributed load vector $\boldsymbol{q}_i$, $\boldsymbol{h}_i$, is approximated as

$$\boldsymbol{h}_i \approx \boldsymbol{h}_i(s) = \boldsymbol{h}^{(i-1)}\phi_{i,1} + \boldsymbol{h}^{(i)}\phi_{i,3},$$

$$\boldsymbol{h}^{(j)} = -1/2\rho_W d \left(\mathbf{L}_{IH} \begin{bmatrix} f_p C_D v_{\hat{p}_1} \sqrt{v^2_{\hat{p}_1} + v^2_{\hat{p}_2}} \\ f_p C_D v_{\hat{p}_2} \sqrt{v^2_{\hat{p}_1} + v^2_{\hat{p}_2}} \\ f_q C_D v_{\hat{q}} |v_{\hat{q}}| \end{bmatrix} \right)^{(j)}. \quad (14)$$

Following the approach of Huang [3] and Buckham et al. [19], we represent the added mass effects as an additional inertia where the added mass component of the distributed load **q** is removed, and the tether element's mass is increased to reproduce the added mass forces. The added mass attributed to the ith element is approximated as the average of the added masses calculated at the bounding $i-1$th and ith nodes.

$$\mathbf{M}_i = \left(\frac{1}{4}\pi d_C^2 \rho_C \mathbf{I} \right) + \frac{1}{2}(\mathbf{L}_{IH}\mathbf{M}_A\mathbf{L}^T_{IH})^{(i-1)} + \frac{1}{2}(\mathbf{L}_{IH}\mathbf{M}_A\mathbf{L}^T_{IH})^{(i)}. \quad (15)$$

To facilitate evaluation of Eq. (12), we must approximate the torsion over the ith element, γ_i. Having selected a linear form for the trial solution α_i, we recognize that the α' term in Eq. (12) will evaluate to a constant quantity over the element. To maintain this level of approximation within Eq. (12) we model the torsion of the element with the constant quantity γ_i:

$$\gamma_i \approx \gamma^{(i-1/2)} = \left(\frac{\boldsymbol{r}'_i \cdot (\boldsymbol{r}''_i \times \boldsymbol{r}'''_i)}{\boldsymbol{r}''_i \cdot \boldsymbol{r}''_i} \right)^{(i-1/2)}. \quad (16)$$

Equation (16) thus allows the twist within the ith element to be modeled by

$$\tau_i = \gamma_i + \alpha'_i. \quad (17)$$

Substituting Eqs. (13)–(17) into Eq. (11), we evaluate the residuals to yield

$$([\mathbf{K}_B]_i + [\mathbf{K}_A]_i + [\mathbf{K}_\tau]_i)\boldsymbol{X}_i + \boldsymbol{W}_i + \boldsymbol{H}_i + \boldsymbol{B}_i = [\mathbf{M}_e]_i \ddot{\boldsymbol{X}}_i. \quad (18)$$

The 12×12 system matrices $[\mathbf{K}_B]_i$, $[\mathbf{K}_A]_i$, and $[\mathbf{K}_\tau]_i$ embody generalized bending, axial, and torsional forces, respectively, that are applied at the element nodes and result from the curvature, axial strain, and twist experienced throughout the cable element. The generalized load vectors $\boldsymbol{W}_i$, $\boldsymbol{H}_i$, and $\boldsymbol{B}_i$ define the weight and buoyancy, hydrodynamic, and boundary forces respectively, and the 12×1 vector, $\boldsymbol{X}_i = \{\boldsymbol{r}_i^{(i-1)^T}\ \boldsymbol{r}''^{i(i-1)^T}\ \boldsymbol{r}_i^{(i)T}\ \boldsymbol{r}''^{i(i)^T}\}^T$, defines the element trial solution, $\boldsymbol{r}_i$. The element mass matrix, $[\mathbf{M}_e]_i$, is a full 12×12 matrix constructed from multiples of the 3×3 mass matrix, $\mathbf{M}_i$:

$$[\mathbf{M}_e]_i = L_i \begin{bmatrix} \frac{1}{3}\mathbf{M}_i & \frac{-1}{45}L_i^2\mathbf{M}_i & \frac{1}{6}\mathbf{M}_i & \frac{-7}{360}L_i^2\mathbf{M}_i \\ & \frac{2}{945}L_i^4\mathbf{M}_i & \frac{-7}{360}L_i^2\mathbf{M}_i & \frac{31}{15120}L_i^4\mathbf{M}_i \\ & & \frac{1}{3}\mathbf{M}_i & \frac{-1}{45}L_i^2\mathbf{M}_i \\ \text{SYM} & & & \frac{2}{945}L_i^4\mathbf{M}_i \end{bmatrix}.$$

The stiffness matrices $[\mathbf{K}_B]_i$ and $[\mathbf{K}_A]_i$ are the following 12×12 element matrices:

$$[\mathbf{K}_B]_i = \frac{EI}{L_i} \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & -\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (19)$$

$$[\mathbf{K}_A]_i = \begin{bmatrix} -\frac{\lambda_i}{L_i}\mathbf{I} & \mathbf{0} & \frac{\lambda_i}{L_i}\mathbf{I} & \mathbf{0} \\ & -\frac{\lambda_i}{45}L_i^3\mathbf{I} & \mathbf{0} & -\frac{7\lambda_i}{360}L_i^3\mathbf{I} \\ & & -\frac{\lambda_i}{L_i}\mathbf{I} & 0 \\ \text{SYM} & & & -\frac{\lambda_i}{45}L_i^3\mathbf{I} \end{bmatrix}. \quad (20)$$

Due to the cross product within the torsional term of the residual equation, the generalized forces produced by the twist of the element are non-linear expressions. To express the torsional effects in first-order form, we choose the following factorization for $[\mathbf{K}_\tau]_i$:

$$[\mathbf{K}_\tau]_i = \frac{GJ\tau_i}{L_i} \begin{bmatrix} \frac{1}{2}(\mathbf{Q}^{(i-1)} + \mathbf{Q}^{(i)}) & -\frac{L_i^2}{24}\mathbf{Q}^{(i)} & -\frac{1}{2}(\mathbf{Q}^{(i-1)} + \mathbf{Q}^{(i)}) & \frac{L_i^2}{24}\mathbf{Q}^{(i-1)} \\ \frac{L_i^2}{24}(\mathbf{Q}^{(i-1)} - \mathbf{Q}^{(i)}) & \frac{L_i^4}{120}\mathbf{Q}^{(i)} & -\frac{L_i^2}{24}(\mathbf{Q}^{(i-1)} - \mathbf{Q}^{(i)}) & \frac{7L_i^4}{720}\mathbf{Q}^{(i-1)} \\ -\frac{1}{2}(\mathbf{Q}^{(i-1)} + \mathbf{Q}^{(i)}) & \frac{L_i^2}{24}\mathbf{Q}^{(i)} & \frac{1}{2}(\mathbf{Q}^{(i-1)} + \mathbf{Q}^{(i)}) & -\frac{L_i^2}{24}\mathbf{Q}^{(i-1)} \\ \frac{L_i^2}{24}(\mathbf{Q}^{(i-1)} - \mathbf{Q}^{(i)}) & \frac{7L_i^4}{720}\mathbf{Q}^{(i)} & -\frac{L_i^2}{24}(\mathbf{Q}^{(i-1)} - \mathbf{Q}^{(i)}) & \frac{L_i^4}{120}\mathbf{Q}^{(i-1)} \end{bmatrix} \quad (21)$$

where $\mathbf{Q}^{(i)}$ represents the 3×3 skew symmetric cross product matrix constructed from the components of the vector $\boldsymbol{r}''^{(i)}$. The load vectors $\boldsymbol{W}_i$, $\boldsymbol{H}_i$, and $\boldsymbol{B}_i$ take the following form:

$$\boldsymbol{H}_i = L_i \begin{Bmatrix} \left(\frac{1}{3}\boldsymbol{h}^{(i-1)} + \frac{1}{6}\boldsymbol{h}^{(i)}\right) \\ \left(-\frac{1}{45}\boldsymbol{h}^{(i-1)} - \frac{7}{360}\boldsymbol{h}^{(i)}\right)L_i^2 \\ \left(\frac{1}{6}\boldsymbol{h}^{(i-1)} + \frac{1}{3}\boldsymbol{h}^{(i)}\right) \\ \left(-\frac{7}{360}\boldsymbol{h}^{(i-1)} - \frac{1}{45}\boldsymbol{h}^{(i)}\right)L_i^2 \end{Bmatrix},$$

$$\boldsymbol{W}_i = L_i \frac{\pi d^2}{4} \begin{Bmatrix} \frac{1}{2}(\rho_c \boldsymbol{g} - \rho_w \boldsymbol{g}) \\ \frac{-L_i^2}{24}(\rho_c \boldsymbol{g} - \rho_w \boldsymbol{g}) \\ \frac{1}{2}(\rho_c \boldsymbol{g} - \rho_w \boldsymbol{g}) \\ \frac{-L_i^2}{24}(\rho_c \boldsymbol{g} - \rho_w \boldsymbol{g}) \end{Bmatrix}, \tag{22}$$

$$\boldsymbol{B}_i = \begin{Bmatrix} -\boldsymbol{F}_i^{(i-1)} \\ \mathbf{0} \\ \boldsymbol{F}_i^{(i)} \\ \mathbf{0} \end{Bmatrix}.$$

Substituting Eq. (16) into Eq. (12), yields additional 2×2 system of equations that defines the torsional deformation, α_i, over the element

$$\frac{GJ}{L_i}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{Bmatrix} \alpha^{(i-1)} \\ \alpha^{(i)} \end{Bmatrix} = GJ\begin{Bmatrix} \gamma_i \\ -\gamma_i \end{Bmatrix} + \begin{Bmatrix} -GJ\tau_i^{(i-1)} \\ GJ\tau_i^{(i)} \end{Bmatrix} \tag{23}$$

where the boundary terms $GJ\tau_i^{(i-1)}$ and $GJ\tau_i^{(i)}$ represent the internal restoring torque at the boundaries of the element.

4 Forming the Reduced Element Equations

The element equations given by Eq. (18) and Eq. (23), provide relationships between the observed state of a tether element, defined by $\boldsymbol{X}_i$ and α_i, and the resulting motion of that element, as defined by, $\ddot{\boldsymbol{X}}_i$ a 6×1 vector. The components of $\ddot{\boldsymbol{X}}_i$ are the acceleration of the $i-1$th and ith node points, which quantify the temporal change of the linear portion of the element shape, and the temporal change of the node curvatures, which quantify the changing curvilinear profile of the element between the $i-1$ and i node points. Applying the explicit solution scheme of [19,21,26] to this higher-order formulation, the model will be advanced in time by solving for the node accelerations, $\ddot{\boldsymbol{r}}^{(i)}$, and the second-order time differentials of the curvature vectors, $\ddot{\boldsymbol{r}}''^{(i)}$, and integrating these quantities to produce the updated generalized displacement values at the next time step. Thus, the discrete element motion equations are a 12×12 linear system: double the size of the conventional lumped mass motion equations.

Earlier, we demonstrated the second order continuity of the twisted spline element. Now, we will show that by enforcing first order continuity of the elements, or smoothness, across node points, we can reduce the number of variables required to define the element state, and correspondingly the element motion, to six matching that of conventional lumped mass approaches.

4.1 Enforcing First-Order Continuity. In order to enforce smoothness, or $C^1(s)$ continuity, of the twisted spline elements across the node points, the curvature vectors observed at the node points, $\boldsymbol{r}''^{(i)}|_{i=0}^{N}$, must satisfy a series of constraint equations defined over the scope of the concatenated twisted spline elements, [35]. For a cable with N elements that are concatenated to extend between nodes 0 and N, these constraint equations take the form:

$$\begin{bmatrix} \frac{L_1}{3}\mathbf{I} & \frac{L_1}{6}\mathbf{I} & 0 & 0 & 0 \\ \frac{L_1}{6}\mathbf{I} & \frac{L_1+L_2}{3}I & \frac{L_2}{6}\mathbf{I} & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \frac{L_{N-1}}{6}\mathbf{I} & \frac{L_{N-1}+L_N}{3}\mathbf{I} & \frac{L_N}{6}\mathbf{I} \\ 0 & 0 & 0 & \frac{L}{6}\mathbf{I} & \frac{L_N}{3}\mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{r}''^{(0)} \\ \boldsymbol{r}''^{(1)} \\ \vdots \\ \boldsymbol{r}''^{(N-1)} \\ \boldsymbol{r}''^{(N)} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\boldsymbol{r}^{(1)} - \boldsymbol{r}^{(0)}}{L_1} \\ \frac{\boldsymbol{r}^{(2)} - \boldsymbol{r}^{(1)}}{L_2} - \frac{\boldsymbol{r}^{(1)} - \boldsymbol{r}^{(0)}}{L_1} \\ \vdots \\ \frac{\boldsymbol{r}^{(N)} - \boldsymbol{r}^{(N-1)}}{L_N} - \frac{\boldsymbol{r}^{(N-1)} - \boldsymbol{r}^{(N-2)}}{L_{N-1}} \\ -\frac{\boldsymbol{r}^{(N)} - \boldsymbol{r}^{(N-1)}}{L_N} \end{bmatrix} + \begin{bmatrix} -\boldsymbol{r}'^{(0)} \\ 0 \\ \vdots \\ 0 \\ \boldsymbol{r}'^{(N)} \end{bmatrix}. \tag{24}$$

The solution of Eq. (24) requires the specification of boundary conditions. If the orientation of the tether at nodes 0 or N is constrained by a clamped termination to a vehicle or a stationary platform, this orientation must be specified through $\boldsymbol{r}'^{(0)}$ or, respectively. Conversely, in the case of a boundary idealized as a spherical joint, or a free boundary, at node 0 or N, $\boldsymbol{r}''^{(0)} = \mathbf{0}$ or $\boldsymbol{r}''^{(N)} = \mathbf{0}$. For terminations that only partially constrain the tether orientation, a combination of the free and clamped conditions can be applied across the X, Y, and Z dimensions at the boundary in question.

Equation (24) defines the $\boldsymbol{r}''^{(i)}$ values directly in terms of the node positions, $\boldsymbol{r}_i|_{i=0}^{N}$, and the element unstretched lengths, $L_i|_{i=1}^{N}$. Thus, for any given set of node positions and prescribed boundary conditions, there is a unique set of curvature vectors that produce a smooth $C^2(s)$ continuous tether between the tether ends. Thus, an assembly of the twisted cubic spline elements only requires specification of the node positions to fully define the profile of the cubic elements.

We choose to use the unstretched element lengths in the coefficient matrix of Eq. (24) to calculate the curvature vectors so that the coefficient matrix can be calculated prior to the simulation. This, however, introduces an error factor of $(1+\varepsilon)^2$ in the calculated curvatures, a factor which is significant only in high tension situations. However, in such circumstances, the magnitude of the tensile forces dominate the tether dynamics and dwarf any bending and torsional effects based on the third order component of the tether profile. Thus, the use of the unstretched lengths in Eq. (24) is expected to be acceptable in instances of both slack and taut tether. Another consequence of the low-tension assumption in Eq. (24) is that the clamped boundary conditions are completely specified by the orientations, $\hat{t}^{(0)}$ and $\hat{t}^{(N)}$ since $\boldsymbol{r}'^{(i)} \approx \hat{t}^{(i)}$ when the axial strain is very small.

4.2 Reduction of the Elemental Motion Equations. The system of smoothness constraint equations Eq. (24) transforms disturbances in the tether profile, given as changes in the series of node positions, into changes in the curvatures observed at the nodes of the assembled system. The resulting curvatures ensure

the twisted cubic spline fit through the resulting node configuration retains the desired smoothness and second-order continuity. The symbolic form of Eq. (24) is

$$\boldsymbol{R}''=\mathbf{K}_1\boldsymbol{R}+\mathbf{K}_2\boldsymbol{r}'^{(0)}+\mathbf{K}_3\boldsymbol{r}'^{(N)} \tag{25}$$

where $\boldsymbol{R}=\{r^{(0)^{\mathrm{T}}}\ r^{(1)^{\mathrm{T}}}\cdots\ r^{(N)^{\mathrm{T}}}\}^{\mathrm{T}}$ and the entries of the $3(N+1)\times 3(N+1)$ linear transformation $\mathbf{K}_1$ and the $3(N+1)\times 3$ transforms $\mathbf{K}_2$ and $\mathbf{K}_3$ are functions of the element lengths, $L_i|_{i=1}^{N}$. Considering a case in which the element lengths are constant, the $\mathbf{K}_1$, $\mathbf{K}_2$, and $\mathbf{K}_3$ transformations map, respectively, time rates of change of the node positions, the top boundary conditions, and the bottom boundary conditions over the evolving nodal curvatures. Differentiating (25) twice with respects to time in the case considered:

$$\ddot{\boldsymbol{R}}''=\mathbf{K}_1\ddot{\boldsymbol{R}}+\mathbf{K}_2\ddot{\boldsymbol{r}}'^{(0)}+\mathbf{K}_3\ddot{\boldsymbol{r}}'^{(N)} \tag{26}$$

Examining the structure of Eq. (24), it is apparent that the $\mathbf{K}_1$ matrix has no zero entries, reflecting the dependence of the acceleration of the ith curvature vector, and the corresponding changes in the ith element's curvilinear profile, on the acceleration of any node in the assembled system. Thus for the ith element, it is necessary to use Eq. (26) to replace $\ddot{\boldsymbol{r}}''^{(i-1)}$ and $\ddot{\boldsymbol{r}}''^{(i)}$ in Eq. (18) with linear combinations of the node accelerations, $\ddot{\boldsymbol{r}}^{(i)}|_{i=0}^{N}$, and boundary terms, $\ddot{\boldsymbol{r}}'^{(0)}$ and $\ddot{\boldsymbol{r}}'^{(N)}$, to ensure the evolving curvatures maintain the smoothness of the discrete cable.

This coupling of the element states goes against the premise of a finite element method. Conventionally, the state of the ith element would be defined strictly in terms of variables possessing a physical association with only the ith element. Element definitions following this rule ensure sparsity in the resulting element equations. However, due to a synergy between the twisted spline element and a lumped mass approximation to the cable continuum, it is possible to express Eq. (18) strictly in terms of node accelerations while maintaining the desired independence between the elements of the assembled. We first compress the element motion equations down to two vector equations by pre-multiplying both sides of Eq. (18) by $\mathbf{P}_i$ where

$$\mathbf{P}_i=\begin{bmatrix}\mathbf{1} & \dfrac{1}{L_i^2}\mathbf{I} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \dfrac{1}{L_i^2}\mathbf{I}\end{bmatrix}$$

The result is a reduced form of the motion equation for the ith element:

$$\begin{aligned}\mathbf{P}_i[([\mathbf{K}_B]_i+[\mathbf{K}_A]_i+[\mathbf{K}_\tau]_i)\boldsymbol{X}_i+\boldsymbol{W}_i+\boldsymbol{H}_i+\boldsymbol{B}_i]&\\ =L_i\begin{bmatrix}\mathbf{M}_i & \mathbf{0}\\ \mathbf{0} & \mathbf{M}_i\end{bmatrix}\Bigg(\begin{bmatrix}\dfrac{14}{45}\mathbf{I} & \dfrac{53}{360}\mathbf{I}\\ \dfrac{53}{360}\mathbf{I} & \dfrac{14}{45}\mathbf{I}\end{bmatrix}\begin{Bmatrix}\ddot{\boldsymbol{r}}^{(i-1)}\\ \ddot{\boldsymbol{r}}^{(i)}\end{Bmatrix}&\\ -L_i^2\begin{bmatrix}\dfrac{19}{945}\mathbf{I} & \dfrac{263}{15120}\mathbf{I}\\ \dfrac{263}{15120}\mathbf{I} & \dfrac{19}{945}\mathbf{I}\end{bmatrix}\begin{Bmatrix}\ddot{\boldsymbol{r}}''^{(i-1)}\\ \ddot{\boldsymbol{r}}''^{(i)}\end{Bmatrix}\Bigg).&\end{aligned} \tag{27}$$

The term in parentheses on the right-hand side of Eq. (27) describes how the mass of the tether element has been associated with the possible modes of element motion. The first matrix in the parentheses of Eq. (27) quantifies the mass of the element, consisting of both tether mass and added mass, that must be accelerated as the element's nodes, nodes $i-1$ and i, accelerate to produce translation or rotation of the element in the water column. The second matrix in the parentheses quantifies mass within the ith element that is accelerated as curvature changes at nodes $i-1$ and i produce a temporal variation of the higher order component of the element profile.

4.3 The Lumped Mass Approximation. Drawing on the results of previous developments in the area of cable dynamics modeling, one can argue that the added degree of complexity induced by the inclusion of the $\ddot{\boldsymbol{r}}''^{(i-1)}$ and $\ddot{\boldsymbol{r}}''^{(i)}$ terms in Eq. (27) is not warranted. Since its introduction by Walton and Polacheck [8] the success of the lumped mass approximation in the study of various underwater cable dynamics problems has been well documented, [2,3,5,7,16,19–21,25,26]. A detailed analysis of the lumped mass approach was presented by Huang in [3] which showed that the lumped mass approximation is consistent with the dynamics of an actual highly flexible cable in the limit of differential element size. Kamman and Huston rationalized the use of the lumped mass approach through its superior computational performance, and equivalent accuracy, compared to more complicated methods that used a more distributed representation of the cable mass, [16].

The premise of the lumped mass approach is a redistribution of the cable element's mass such that it is concentrated at the node points. The consequence of this approximation is that changes in the higher order shape of the element profile, caused by changing curvature at the element end nodes, do not accelerate any cable mass. Rather, only inertia associated with the acceleration of the node points is considered. The justification for this redistribution of the cable inertia is based on physical and heuristic reasoning. Since the ROV tether is a slender body, there is a limit to the amount of mass that is being shifted, and hence a limit to the error this shift incurs in the evaluation of the motion equations. Secondly, the lumped mass approximation serves to decouple the motion of the node points and this ensures computational simplicity in the solution of the resulting motion equations. The lumped mass approximation to the cable dynamics is realized by a direct manipulation of the inertial terms of Eq. (27). Using the lumped mass approximation the complete element motion equations can be given by

$$\begin{aligned}\mathbf{P}_i[([\mathbf{K}_B]_i+[\mathbf{K}_A]_i+[\mathbf{K}_\tau]_i)\boldsymbol{X}_i+\boldsymbol{W}_i+\boldsymbol{H}_i+\boldsymbol{B}_i]&\\ =L_i\begin{bmatrix}\dfrac{1}{2}\mathbf{M}_i & \mathbf{0}\\ \mathbf{0} & \dfrac{1}{2}\mathbf{M}_i\end{bmatrix}\begin{Bmatrix}\ddot{\boldsymbol{r}}^{(i-1)}\\ \ddot{\boldsymbol{r}}^{(i)}\end{Bmatrix}.&\end{aligned} \tag{28}$$

The aforementioned synergy between the lumped mass approximation and the twisted spline elements is evident when considering Eq. (24) and Eq. (28). Given that the curvature vectors, which completely define the curvilinear element profile and thus afford calculation of the bending and torsional effects, are an explicit function of the collection of node points, it is necessary to integrate only the time rates of change of the node positions to capture the complete tether state, including internal twist, at the next instant in time. Thus, the lumped mass motion equations, Eq. (28), are in terms of only six state variables—the node positions. Since the curvatures are calculated at each time step, the only consequence of eliminating the curvature time differentials from the element motion equations is the error associated with the heuristic mass redistribution.

5 Solution Procedure

The reduced motion equations, given by Eq. (28), and the torsional constraint equations for the element, Eq. (23), form the final element equations that are to be used in calculating the dynamic response of the ROV tether to changing boundary conditions. The tether is approximated by a series of the cubic cable elements and a global system of dynamics equations, that represent the behavior of the entire tether, is assembled from the final element equations. We will now illustrate the solution procedure.

5.1 Assembly of the Discrete Tether Model. To create a discrete representation of an ROV tether, we concatenate a series of the twisted cubic spline elements. This process, referred to as assembly, is represented mathematically by the application of Eq. (28) and Eq. (23) for $i=1,2,\ldots,N$. This assembly process forms the global system of motion equations:

$$[[\mathbf{K}_B]_G+[\mathbf{K}_A]_G+[\mathbf{K}_\tau]_G]\{\boldsymbol{r}^{(0)^{\mathrm{T}}}\boldsymbol{r}''^{(0)^{\mathrm{T}}}\boldsymbol{r}^{(1)^{\mathrm{T}}}\cdots\boldsymbol{r}''^{(N-1)^{\mathrm{T}}}\boldsymbol{r}^{(N)^{\mathrm{T}}}\boldsymbol{r}''^{(N)^{\mathrm{T}}}\}^{\mathrm{T}} + \boldsymbol{H}_G+\boldsymbol{W}_G+\boldsymbol{B}_G=[\mathbf{M}]_G\ddot{\boldsymbol{R}} \tag{29}$$

where $[\mathbf{K}_B]_G$, $[\mathbf{K}_A]_G$, and $[\mathbf{K}_\tau]_G$ are the global bending, axial, and torsional stiffness matrices, respectively; $\boldsymbol{H}_G$ is the global vector of hydrodynamic forces acting at the node points of the assembled model; $\boldsymbol{W}_G$ is the assembly of weight and buoyancy forces; and $\boldsymbol{B}_G$ defines the boundary loads applied over the assembled tether. For example, accounting for the pre-multiplication of each elemental hydrodynamic force vector by $\mathbf{P}_i$, the global vector of hydrodynamic forces is given by

$$\boldsymbol{H}_G=\begin{Bmatrix} \frac{14}{45}L_1\boldsymbol{h}^{(0)}+\frac{53}{360}L_1\boldsymbol{h}^{(1)} \\ \frac{53}{360}L_1\boldsymbol{h}^{(0)}+\left(\frac{14}{45}L_1+\frac{14}{45}L_2\right)\boldsymbol{h}^{(1)}+\frac{53}{360}L_2\boldsymbol{h}^{(2)} \\ \vdots \\ \frac{53}{360}L_{N-1}\boldsymbol{h}^{(N-2)}+\left(\frac{14}{45}L_{N-1}+\frac{14}{45}L_N\right)\boldsymbol{h}^{(N-1)}+\frac{53}{360}L_N\boldsymbol{h}^{(N)} \\ \frac{53}{360}L_N\boldsymbol{h}^{(N-1)}+\frac{14}{45}L_N\boldsymbol{h}^{(N)} \end{Bmatrix}$$

Previously, in Eq. (22), we expressed the boundary loads for the ith element in terms of the internal forces experienced at the boundaries of the element.

To define the global vector of boundary loads, we recognize that instantaneous changes of the internal force across the node points are created by external forces applied at the node points. At node i

$$\boldsymbol{F}_i^{(i)}-\boldsymbol{F}_{i+1}^{(i)}=\boldsymbol{f}_a^{(i)}$$

where $\boldsymbol{f}_a^{(i)}$ is an external force applied at node i of the assembled tether, $F_0^{(0)}=0$, and $F_{N+1}^{(N)}=0$. Using these substitutions, the global vector of boundary loads is given by

$$\boldsymbol{B}_G=\{\boldsymbol{f}_a^{(0)T}\boldsymbol{f}_a^{(1)T}\cdots\boldsymbol{f}_a^{(N)T}\}^T.$$

The global boundary load vector facilitates the application of external forces at any node point of the assembled system. These applied loads provide unconstrained boundary conditions on the third derivative of the tether profile, $\boldsymbol{r}'''$.

The inertial terms on the right-hand side of the equation consist of the global mass matrix that is formed by tiling the N instances of the element mass matrices and the vector of node accelerations. For instance, using the lumped mass assumption the global mass matrix takes the form

$$[\mathbf{M}]_G=\begin{bmatrix} \frac{1}{2}L_1\mathbf{M}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}L_1\mathbf{M}_1+\frac{1}{2}L_2\mathbf{M}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{2}L_{N-1}\mathbf{M}_{N-1}+\frac{1}{2}L_N\mathbf{M}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{2}L_N\mathbf{M}_N \end{bmatrix}$$

In addition to the global system of motion equations, the torsional constraint equations given by Eq. (23) are assembled to produce a global system of constraint equations that define the torsional deformation throughout the tether:

$$GJ\begin{bmatrix} \frac{1}{L_1} & -\frac{1}{L_1} & 0 & 0 & 0 \\ -\frac{1}{L_1} & \frac{1}{L_1}+\frac{1}{L_2} & -\frac{1}{L_2} & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & -\frac{1}{L_{N-1}} & \frac{1}{L_{N-1}}+\frac{1}{L_N} & -\frac{1}{L_N} \\ 0 & 0 & 0 & -\frac{1}{L_N} & \frac{1}{L_N} \end{bmatrix} \begin{Bmatrix} \alpha^{(0)} \\ \alpha^{(1)} \\ \vdots \\ \alpha^{(N-1)} \\ \alpha^{(N)} \end{Bmatrix} = GJ \begin{Bmatrix} \gamma_1 \\ \gamma_2-\gamma_1 \\ \vdots \\ \gamma_N-\gamma_{N-1} \\ -\gamma_N \end{Bmatrix} + \begin{Bmatrix} \mathrm{T}^{(0)} \\ 0 \\ 0 \\ 0 \\ \mathrm{T}^{(N)} \end{Bmatrix} \tag{30}$$

where $\mathrm{T}^{(0)}=-GJ\tau_1^{(0)}$ and $\mathrm{T}^{(N)}=GJ\tau_N^{(N)}$ are external torsional couples applied at the boundaries of the assembled tether. In writing Eq. (30) we have assumed that no torsional couples are applied at the interior nodes of the assembled system of elements.

5.2 Solution of the Initial Value Problem. Recalling that the node curvatures are an explicit function of the node positions, the left-hand side of Eq. (29) is seen to be an explicit function of the node positions and, through the hydrodynamic terms, node velocities. The series of global motion equations in Eq. (29) thus forms a series of $3(N+1)$ second-order differential equations of the form

$$\boldsymbol{F}(\boldsymbol{R},\dot{\boldsymbol{R}})=[\mathbf{M}]_G\ddot{\boldsymbol{R}}. \tag{31}$$

Equation (31) is a classic initial value problem that defines the motion of the assembled tether nodes, $\boldsymbol{r}^{(i)}|_{i=0}^{N}$, during a ROV maneuver. To advance the model in time using an integration routine, the initial value problem is recast as a series of $6N$ first-order differential equations by introducing the node velocities, $\dot{\boldsymbol{r}}^{(i)}|_{i=0}^{N}$, as additional state variables to be included in the integration process. Thus, given an initial state of the tether, $\boldsymbol{Y}(t_0)$

$$\boldsymbol{Y}(t_0)=\{\boldsymbol{r}^{(0)^{\mathrm{T}}}(t_0)\ \ \dot{\boldsymbol{r}}^{(0)^{\mathrm{T}}}(t_0)\ \ \boldsymbol{r}^{(1)^{\mathrm{T}}}(t_0)\ \ \cdots\ \ \dot{\boldsymbol{r}}^{(N-1)^{\mathrm{T}}}(t_0)\ \ \boldsymbol{r}^{(N)^{\mathrm{T}}}(t_0)\ \ \dot{\boldsymbol{r}}^{(N)^{\mathrm{T}}}(t_0)\}^{\mathrm{T}},$$

the integration scheme produces $\boldsymbol{Y}(t_1)$, $\boldsymbol{Y}(t_2),\ldots,\boldsymbol{Y}(t_k)$, $\boldsymbol{Y}(t_{k+1}),\ldots,\boldsymbol{Y}(t_f)$, where t_f is the duration of the simulation. To advance the model each time step we propose an explicit approach. For the kth step advancing the model from the state $\boldsymbol{Y}(t_k)$ at time t_k to t_{k+1}, the following sequence of calculations produces the next state, $\boldsymbol{Y}(t_{k+1})$:

1. Using the node positions of the current tether state, $\boldsymbol{r}^{(i)}|_{i=0}^{N}$, and given conditions at the tether boundaries, one of $\boldsymbol{r}'^{(0)}$ or $\boldsymbol{r}''^{(0)}$ and one of $\boldsymbol{r}'^{(N)}$ or $\boldsymbol{r}''^{(N)}$, Eq. (24) is evaluated to produce the set of curvature vectors, $\boldsymbol{r}''^{(i)}|_{i=0}^{N}$. This defines the profile of each element in the assembly.
2. Using the element profiles, the torsion within each of the elements, $\gamma_i|_{i=0}^{N}$, is calculated according to Eq. (16). Applying the torsion values in Eq. (30) and the given boundary conditions, one of $\alpha^{(0)}$ or $GJ\tau^{(0)}$ and one of $\alpha^{(N)}$ or $GJ\tau^{(N)}$, the torsional deformation throughout the discrete tether, $\alpha^{(i)}|_{i=0}^{N}$, is calculated.
3. The axial internal forces at the node points, $\lambda^{(i)}|_{i=0}^{N}$ are calculated using Eq. (13), and the element stiffness matrices $[\mathbf{K}_B]_i$, $[\mathbf{K}_A]_i$, and $[\mathbf{K}_\tau]_i$ are calculated according to Eqs. (19)–(21) respectively, and then premultiplied by $\mathbf{P}$.
4. The element mass matrices, $\mathbf{M}_i|_{i=0}^{N}$, are calculated according to Eq. (15).
5. The hydrodynamic loads at the node points $\boldsymbol{h}^{(i)}|_{i=0}^{N}$ are calculated using Eq. (14), and the generalized load vectors $\boldsymbol{W}_i|_{i=1}^{N}$ and $\boldsymbol{H}_i|_{i=1}^{N}$ are calculated according to Eq. (22), and premultiplied by $\mathbf{P}$.
6. The global system of equations is assembled. Specifying the boundary conditions, one of $\boldsymbol{F}^{(0)}$ or $\boldsymbol{r}^{(0)}$ and one of $\boldsymbol{F}^{(N)}$ or $\boldsymbol{r}^{(N)}$, Eq. (31) is solved for the node accelerations, $\ddot{\boldsymbol{r}}^{(i)}|_{i=0}^{N}$.
7. The node accelerations are augmented with the velocities of the current tether state to form the time derivative of the state vector $\dot{\boldsymbol{Y}}(t_k)$, which is integrated to generate the next state of the cable, $\boldsymbol{Y}(t_{k+1})$.

6 Conclusions

We have presented a cubic cable element that provides a more compact representation of low-tension cable motion than existing finite element models. Rather than the 12 or 13 state variables defining the cubic elements of these more conventional finite element approaches, the state of this new cubic element is defined at any instant by six variables: the positions of the end nodes. This reduction in the number of required state variables is achieved by choosing an element form that enforces second-order continuity of the cable profile. This choice of element reflects the fact that in ROV maneuvers, and in many other tethered systems where low-tension dynamics are likely, the tether is not subject to externally applied moments. The chosen element form allows the element motion equations to be formed strictly in terms of the forces acting on the node points and the resulting translation of the nodes. The ability of the model to capture the dynamics of slack cables in three dimensions while maintaining the advantages of simpler linear finite element approaches, makes it suitable for solving a much wider variety of undersea cable dynamics problems than its predecessors. The model's ability to capture long periods of low-tension cable motion makes it a desirable tool, not only in tethered ROV applications, but in any application that involves alternating taut and low-tension regimes.

References

[1] Burgess, J. J., 1993, "Bending Stiffness in the Simulation of Undersea Cable Deployment," Int. J. Offshore Polar Eng., **3**(3), pp. 197–204.
[2] Delmer, T. M., Stephens, T. C., and Coe, J. M., 1983, "Numerical Simulation of Towed Cables," Ocean Eng., **10**(2), pp. 119–132.
[3] Huang, S., 1994, "Dynamic Analysis of Three Dimensional Marine Cables," Ocean Eng., **21**(6), pp. 587–605.
[4] Sanders, J. V., 1982, "A Three Dimensional Dynamic Analysis of a Towed System," Ocean Eng., **9**(5), pp. 483–499.
[5] Vaz, M. A., Witz, J. A., and Patel, J. A., 1997, "Three Dimensional Transient Analysis of the Installation of Marine Cables," Acta Mech., **124**, pp. 1–26.
[6] Driscoll, F., and Nahon, M., 1996, "Mathematical Modeling and Simulation of a Moored Buoy System," *Proceedings of OCEANS 96 MTS/IEEE*, Ft. Lauderdale, FL, **1**, pp. 517–523.
[7] Thomas, D. O., and Hearn, G. E., 1994, "Deepwater Mooring Line Dynamics With Emphasis on Seabed Interference Effects," *Proceedings of the 26th Offshore Technology Conference*, Houston, TX, pp. 203–214.
[8] Walton, T. S., and Polacheck, H., 1960, "Calculation of Transient Motion of Submerged Cables," Math. Comput., **14**(69), pp. 27–46.
[9] Malahy, R. C., 1986, "A Non-linear Finite Element Method for the Analysis of Offshore Pipelines," *Proceedings of the 5th International Offshore Mechanics and Arctic Engineering Symposium*, **3**, pp. 471–478.
[10] McNamara, J. F., O'Brien, P. J., and Gilroy, S. G., 1988, "Nonlinear Analysis of Flexible Risers Using Hybrid Finite Elements," ASME J. Offshore Mech. Arct. Eng., **110**, pp. 197–204.
[11] O'Brien, P. J., and McNamara, J. F., 1988, "Analysis of Flexible Riser Systems Subject to Three-Dimensional Seastate Loading," *Proceedings of the*

International Conference on Behavior of Offshore Structures, **3**, pp. 1373–1388.
[12] Garret, D. L., 1982, ''Dynamic Analysis of Slender Rods,'' ASME J. Energy Resour. Technol., **104**, pp. 302–306.
[13] Nordgren, R. P., 1974, ''On the Computation of the Motion of Elastic Rods,'' ASME J. Appl. Mech., **41**, pp. 777–780.
[14] Nordgren, R. P., 1982, ''Dynamics Analysis of Marine Risers With Vortex Excitation,'' ASME J. Energy Resour. Technol., **104**, pp. 14–19.
[15] Chapman, D. A., 1982, ''Effects of Ship Motion on a Neutrally-Stable Towed Fish,'' Ocean Eng., **9**(3), pp. 189–220.
[16] Kamman, J. W., and Huston, R. L., 1999, ''Modeling of Variable Length Towed and Tethered Cable Systems,'' J. Guid. Control Dyn., **22**(4), pp. 602–608.
[17] Makarenko, A. I., Poddubnyi, V. I., and Kholopova, V. V., 1997, '' Study of the Non-Linear Dynamics of Self-Propelled Submersibles Controlled by A Cable,'' International Applied Mechanics, **22**(3), pp. 251–257.
[18] Paul, B., and Soler, A. I., 1972, ''Cable Dynamics and Optimum Towing Strategies for Submersibles,'' MTS Journal, **6**(2), pp. 34–42.
[19] Buckham, B., Nahon, M., Seto, M., Zhao, X., and Lambert, C., 2003, ''Dynamics and Control of a Towed Underwater Vehicle System, Part I: Model Development,'' Ocean Eng., **30**(4), pp. 453–470.
[20] Lambert, C., Nahon, M., Buckham, B., and Seto, M., 2003, ''Dynamics and Control of a Towed Underwater Vehicle System, Part II: Model Validation and Turn Maneuver Optimization,'' Ocean Eng., **30**(4), pp. 471–485.
[21] Driscoll, F., Lueck, R. G., and Nahon, M., 2000, ''Development and Validation of a Lumped-Mass Dynamics Model of a Deep-Sea ROV System,'' Appl. Ocean. Res., **22**(3), pp. 169–182.
[22] Driscoll, F., Buckham, B., and Nahon, M., 2000, ''Numerical Optimization of a Cage-Mounted Passive Heave Compensation System,'' *Proceedings of OCEANS 2000 MTS/IEEE*, Providence, RI, **2**, pp. 1121–1127.
[23] McLain, T. W., and Rock, S. M., 1992, ''Experimental Measurement of ROV Tether Tension,'' *Proceedings of the 10th Annual Subsea Intervention Conference and Exposition*, San Diego, CA, pp. 291–296.
[24] Grosenbaugh, M. A., Howell, C. T., and Moxnes, S., 1993, ''Simulating the Dynamics of Underwater Vehicles With Low-Tension Tethers,'' Int. J. Offshore Polar Eng., **3**(3), pp. 213–218.
[25] Banerjee, A. K., and Do, V. N., 1994, ''Deployment Control of a Cable Connecting a Ship to an Underwater Vehicle,'' J. Guid. Control Dyn., **17**(4), pp. 1327–1332.
[26] Buckham, B., and Nahon, M., 2001, ''Formulation and Validation of a Lumped Mass Model for Low-Tension ROV Tethers,'' Int. J. Offshore Polar Eng., **11**(4), pp. 282–289.
[27] O'Neill, B., 1966, *Elementary Differential Geometry*, Academic Press, San Diego, CA, pp. 51–77.
[28] Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, Cambridge University Press, New York, pp. 381–398.
[29] Folb, R., and Nelligan, J. J., 1983, ''Hydrodynamic Loading on Armoured Towcables,'' DTNSRDC Report 82/116.
[30] Newman, J. N., 1989, *Marine Hydrodynamics*, The M.I.T. Press, Cambridge, MA.
[31] Howard, B. E., and Syck, J. M., 1992, ''Calculation of the Shape of a Towed Underwater Acoustic Array,'' Journal of Oceanic Engineering, **17**(2), pp. 193–201.
[32] Logan, D. L., 1993, *A First Course in the Finite Element Method*, 2nd Ed., PWS, Boston, MA, pp. 195–239.
[33] Rao, S. S., 1989, *The Finite Element Method in Engineering*, 2nd Ed., Pergamon Press, Toronto, ON, pp. 101–198; 291–300.
[34] Burnett, D. S., 1987, *Finite Element Analysis. From Concepts to Applications*, Addison-Wesley, Don Mills, ON, pp. 53–197.
[35] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannetry, B. P., 1992, *Numerical Recipes in C*, Cambridge University Press, New York, pp. 113–117.

Collinear and Periodic Electrode-Ceramic Interfacial Cracks in Piezoelectric Bimaterials

Christoph Häusler[1]
e-mail: haeusler@mfk.mw.tu-dresden.de

Cun-Fa Gao

Herbert Balke
Institute of Solid Mechanics,
Technische Universität Dresden,
01062 Dresden, Germany

Field singularities of collinear and collinear periodic interface cracks between an electrode and a piezoelectric matrix are studied in terms of the Stroh formalism for mixed boundary conditions. In contrast to the relevant work done previously on this subject, the problem is solved based on the assumption that the upper and lower planes embedding the electrode consist of two arbitrary piezoelectric materials, and the cracks are assumed to be permeable. The problem is reduced to an interfacial crack problem equivalent to that in purely elastic media. Explicit expressions are presented for the complex potentials and field intensity factors. All the field variables exhibit oscillatory singularities, and their intensities are dependent on the material properties and the applied mechanical loads, but not on the applied electric loads. [DOI: 10.1115/1.1767168]

1 Introduction

Piezoelectric ceramics have been widely used in actuators, sensors and other components. In these devices, the most cost-efficient geometry is that of the cofired multilayer actuators with metal electrodes, [1]. Due to the different mechanical and electrical behavior of electrodes and ceramics, the field variables concentrate around the electrode edges. The nonuniform local field induces crack initiation and crack growth and finally leads to the failure of the device. Thus, with increasingly wide application of ceramic multilayer actuators made of piezoelectric materials, the problems of local fields near electrodes have received much attention in the recent decade.

Some important results have been presented by Yang and Suo [2], Hao et al. [3], Shindo et al. [4], Ye and He [5], and Dos Santos e Lucato et al. [6] for the cases of electrodes in a homogeneous material, and also Deng and Meguid [7] and Ru [8] for the case of interface electrodes between two dissimilar piezoelectric materials. These cited works are focused on the cases without cracks. In fact, experiments have demonstrated, [9,10], that cracks are formed preferentially at the interface between the electrode and the matrix. Thus, it is of both theoretical and practical importance to study the fracture problems of electrode-matrix interface cracks. On this subject, several analytical results have been obtained by Ru [11]. Recently, Wang and Shen [12] have developed a general treatment for generalized two-dimensional problems in anisotropic piezoelectric bimaterial with interface defects.

In the present work we study the generalized two-dimensional problem of collinear electrode-ceramic interface cracks in piezoelectric bimaterials. The solution of a single crack is included as a special case as well. The piezoelectric media located at the two sides of the electrode layer are assumed to be generally anisotropic, and the cracks to be permeable.

The assumption of collinear cracks based on the consideration that they may exist before a main crack emerges. The assumption of permeable cracks will be explained later.

Section 2 outlines the Stroh formalism. In Section 3 we derive the general expression for the complex potentials, in Section 4 several solutions of field intensity factors at the crack tips are given, and in Section 5 we specify numerical examples. Section 6 concludes the paper.

2 Stroh Formalism for Mixed Boundary Value Problem

In a fixed rectangular coordinate system x_i $(i=1,2,3)$, we identify the displacement by u_i, the electric potential by φ, the stress by σ_{ij}, the strain by γ_{ij}, the electric displacement by D_j and the electric field by E_i. The basic equations for a linear piezoelectric solid can be expressed as, [11],

$$\sigma_{ij,j}=0, \quad D_{i,i}=0 \tag{1}$$

$$\gamma_{ij}=\frac{1}{2}(u_{i,j}+u_{j,i}), \quad E_i=-\varphi_{,i} \tag{2}$$

$$\sigma_{ij}=C_{ijkl}\gamma_{kl}-e_{kij}E_k, \quad D_k=e_{kij}\gamma_{ij}+\varepsilon_{kl}E_l \tag{3}$$

where C_{ijkl}, e_{ijk}, and ε_{ij} are the elastic constants, the piezoelectric constants, and the dielectric constants, respectively.

For a generalized two-dimensional problem (all the field variables are independent of x_3) a displacement vector $\mathbf{u}$ can be introduced as, [13],

$$\mathbf{u}=[u_1 \;\; u_2 \;\; u_3 \;\; \varphi]^T=\mathbf{a}f(x_1+px_2) \tag{4}$$

where the superscript T represents the transpose, $f(x_1+px_2)$ is an analytic function, p is a constant, and $\mathbf{a}$ a constant four-element column. Equations (1), (2), and (3) are satisfied by (4) for arbitrary $f(x_1+px_2)$ if

$$[\mathbf{T}+p(\mathbf{R}+\mathbf{R}^{\mathrm{T}})+p^2\mathbf{W}]\mathbf{a}=\mathbf{0} \tag{5}$$

where the (4,4) matrices $\mathbf{T}$, $\mathbf{R}$, and $\mathbf{W}$ are given by

$$\mathbf{T}=\begin{bmatrix} C_{i1k1} & e_{11i} \\ e_{11i}^T & -\varepsilon_{11} \end{bmatrix}, \quad \mathbf{R}=\begin{bmatrix} C_{i1k2} & e_{21i} \\ e_{12i}^T & -\varepsilon_{12} \end{bmatrix},$$

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, October 17, 2002; final revision, January 30, 2004. Associate Editor: K. Ravi-Chandar. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

$$\mathbf{W}=\begin{bmatrix} C_{i2k2} & e_{22i} \\ e_{22i}^T & -\varepsilon_{22} \end{bmatrix}, \quad i,k=1,2,3.$$

The condition for nontrivial solutions of (5) requires $|\mathbf{T}+p(\mathbf{R}+\mathbf{R}^{\mathrm{T}})+p^2\mathbf{W}|=0$. The eight eigenvalues consist of four distinct conjugate pairs with positive imaginary parts, [13], such that p_α, $p_{\alpha+4}=\overline{p_\alpha}$ with $\mathrm{Im}\, p_\alpha>0$, $\alpha=1\ldots4$. Thus, the general solution of $\mathbf{u}$ is

$$\mathbf{u}=\sum_{\alpha=1}^{4}[\mathbf{a}_\alpha f_\alpha(z_\alpha)+\overline{\mathbf{a}_\alpha f_\alpha(z_\alpha)}] \text{ with } z_\alpha=x_1+p_\alpha x_2. \quad (6)$$

The associate stresses and electric displacement can be obtained by inserting (6) and (2) into (3), such that

$$[\sigma_{2j},D_2]^T=\sum_{\alpha=1}^{4}[\mathbf{b}_\alpha f'_\alpha(z_\alpha)+\overline{\mathbf{b}}_\alpha\overline{f'_\alpha(z_\alpha)}], \quad j=1,2,3 \quad (7)$$

$$[\sigma_{1j},D_1]^T=-\sum_{\alpha=1}^{4}[\mathbf{b}_\alpha p_\alpha f'_\alpha(z_\alpha)+\overline{\mathbf{b}}_\alpha\overline{p}_\alpha\overline{f'_\alpha(z_\alpha)}], \quad j=1,2,3 \quad (8)$$

where $\mathbf{b}_\alpha$ are determined by

$$\mathbf{b}_\alpha=(\mathbf{R}^{\mathrm{T}}+p_\alpha\mathbf{W})\mathbf{a}_\alpha=-\frac{1}{p_\alpha}(\mathbf{T}+p_\alpha\mathbf{R})\mathbf{a}_\alpha, \quad \alpha=1\ldots4. \quad (9)$$

From (6), (7), and (8) we have

$$\mathbf{u}=\mathbf{A}\mathbf{f}(z_\alpha)+\overline{\mathbf{A}\mathbf{f}(z_\alpha)} \quad (10)$$

$$\boldsymbol{\phi}=\mathbf{B}\mathbf{f}(z_\alpha)+\overline{\mathbf{B}\mathbf{f}(z_\alpha)} \quad (11)$$

where

$$\mathbf{A}=[\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3\ \mathbf{a}_4], \quad \mathbf{B}=[\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3\ \mathbf{b}_4],$$

$$\mathbf{f}(z_\alpha)=[f_1(z_1)\ f_2(z_2)\ f_3(z_3)\ f_4(z_4)]^T,$$

$$\boldsymbol{\phi}_{,1}=[\sigma_{2j}\ D_2]^T, \quad \text{and} \quad -\boldsymbol{\phi}_{,2}=[\sigma_{1j}\ D_1]^T.$$

To simplify the solution of problems with permeable cracks (this corresponds to a mixed boundary value problem), we define two appropriate vectors as

$$\hat{\mathbf{u}}=[u_1\ u_2\ u_3\ \phi_4] \text{ and } \hat{\boldsymbol{\phi}}=[\phi_1\ \phi_2\ \phi_3\ \varphi]$$

where $\hat{\mathbf{u}}$ and $\hat{\boldsymbol{\phi}}$ can be expressed as

$$\hat{\mathbf{u}}=\mathbf{I}_u\mathbf{u}+\mathbf{I}_t\boldsymbol{\phi} \text{ and } \hat{\boldsymbol{\phi}}=\mathbf{I}_t\mathbf{u}+\mathbf{I}_u\boldsymbol{\phi} \quad (12)$$

with the diagonal matrices $\mathbf{I}_u=\mathrm{diag}(1\,1\,1\,0)$ and $\mathbf{I}_t=\mathrm{diag}(0\,0\,0\,1)$. Note, for general mixed boundary value problems $\mathbf{I}_u$ and $\mathbf{I}_t$ would be any diagonal matrices, whose elements are either 1 or 0 and satisfy $\mathbf{I}_u+\mathbf{I}_t=\mathbf{I}$, $\mathbf{I}_u\cdot\mathbf{I}_t=\mathbf{0}$.

Substituting (10) and (11) into (12) leads to

$$\hat{\mathbf{u}}=\hat{\mathbf{A}}\mathbf{f}(z_\alpha)+\overline{\hat{\mathbf{A}}\mathbf{f}(z_\alpha)} \text{ and } \hat{\boldsymbol{\phi}}=\hat{\mathbf{A}}\mathbf{f}(z_\alpha)+\overline{\hat{\mathbf{B}}\mathbf{f}(z_\alpha)} \quad (13)$$

where

$$\hat{\mathbf{A}}=\mathbf{I}_u\mathbf{A}+\mathbf{I}_i\mathbf{B} \text{ and } \hat{\mathbf{B}}=\mathbf{I}_t\mathbf{A}+\mathbf{I}_u\mathbf{B}. \quad (14)$$

The nature of $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ was studied by Ting and Wang [14], Wang et al. [15], Homulka and Keer [16] for the case of purely anisotropic media, and also by Häusler and Balke [17], and Wang and Shen [12] for the case of piezoelectric materials. It is found that these matrices have a similar nature to that of $\mathbf{A}$ and $\mathbf{B}$, e.g., they are nonsingular and moreover satisfy for a proper normalization of $\mathbf{a}_\alpha$ and $\mathbf{b}_\alpha$ the following relation:

$$\begin{bmatrix} \hat{\mathbf{B}}^T & \hat{\mathbf{A}}^T \\ \overline{\hat{\mathbf{B}}}^T & \overline{\hat{\mathbf{A}}}^T \end{bmatrix}\begin{bmatrix} \hat{\mathbf{A}} & \overline{\hat{\mathbf{A}}} \\ \hat{\mathbf{B}} & \overline{\hat{\mathbf{B}}} \end{bmatrix}=\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad (15)$$

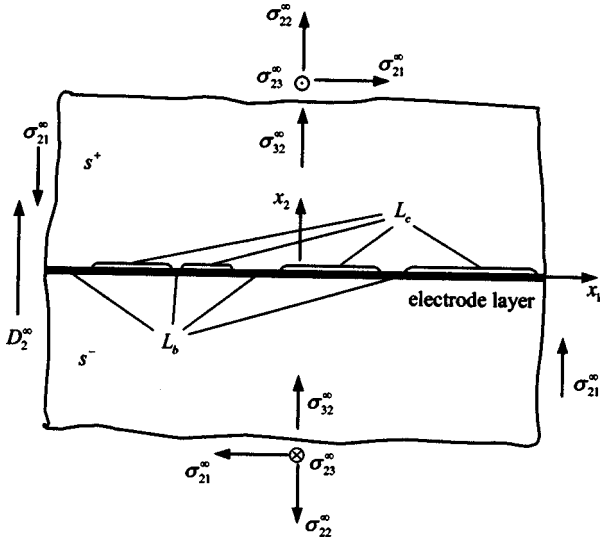


Fig. 1 Arbitrary collinear electrode-ceramic interfacial cracks

where $\mathbf{I}$ is a (4,4) unit matrix.

3 Complex Potentials

As shown in Fig. 1, we consider an electrode layer between two dissimilar piezoelectric half-infinite bodies distinguished by s^+ and s^-. Assume that these two bodies coexist in the state of generalized two-dimensional deformation under the mechanical-electric loads $(\sigma_{2j}^{\infty},D_2^{\infty})$ at infinity, containing collinear interfacial cracks between the electrode and s^+ denoted by L_c. The uncracked part along the x_1-axis is represented by L_b.

Since the electrode layer is commonly much thinner than the matrix, its mechanical properties will be neglected.

The cracks are assumed to be permeable. That means that the electric potential φ is continuous across the crack faces ($\varphi^+=\varphi^-$, [13,18]). We should like to mention besides that because of the electrode layer, the electric potential is not only continuous across the crack faces but also constant inside the cracks. That is just the condition for conducting cracks. Therefore, the reader may refer the considered cracks to conducting cracks. This makes no difference in solving the problem.

In [19], Balke et al. modeled fracture tests of compact tension and three point bending specimens (performed by Park and Sun [20]) by means of finite element method. In their numerical calculation they compare the total energy release rates for three different types of cracks that are impermeable, permeable, and limited permeable cracks, [21]. From their work one can realize that "the cracks behave more like permeable cracks, than like impermeable cracks." Additionally in [22], Schneider et al. study an indentation crack in a homogeneous PZT ceramic using Kelvin probe microscopy and atomic force microscopy. They determine the dielectric constant of the crack interior to be about 40, which is significantly higher than 1, which is expected for a crack filled with air. Schneider et al. state as one possible explanation for this high dielectric constant that " . . . the potential difference across the crack is reduced due to charge compensation at the crack surfaces. This would reduce the electric field interior to the crack, leading to an elevated apparent permittivity." In the case of electrode-ceramic interfacial cracks the presence of the electrode layer will support this effect. Furthermore, Schneider et al. illustrated the consequence of this elevated permittivity for a Griffith crack. They point out that "The theoretically predicted effect of an applied electric field in retarding crack growth decreases significantly with increasing permittivity. In practical situations in terms of crack length, applied electric load, and electric field

level, the retardation of crack growth is negligible when the dielectric constant of the crack interior is higher than 20.'' So we believe the studied permeable electrode-ceramic interfacial cracks are physically justified.

In consideration of the aforesaid assumptions the following boundary conditions hold:

$$E_1^+ = E_1^- = 0, \quad \sigma_{2j}^+ = \sigma_{2j}^- = 0, \quad x_1 \in L_c \tag{16}$$

$$E_1^+ = E_1^- = 0, \quad u_j^+ = u_j^-, \quad \sigma_{2j}^+ = \sigma_{2j}^-, \quad x_1 \in L_b. \tag{17}$$

Because of these boundary conditions and owing to the configuration of the problem, the one-complex-variable method introduced in [23] is appropriate. Each z_α from the function vector $\mathbf{f}(z_\alpha)$ has to be substituted by a common complex variable $z = x_1 + \zeta x_2$ where the imaginary part of ζ is positive. Instead of $\mathbf{f}(z_\alpha)$ we can write $\mathbf{f}(z)$ henceforth. After the solution of $\mathbf{f}(z)$ has been determined the z_α have to be resubstituted to compute field quantities.

From (13) we have

$$\hat{\mathbf{u}}_{,1} = \hat{\mathbf{A}}\mathbf{F}(z) + \overline{\hat{\mathbf{A}}\mathbf{F}(z)} \tag{18}$$

$$\hat{\boldsymbol{\phi}}_{,1} = \hat{\mathbf{B}}\mathbf{F}(z) + \hat{\mathbf{B}}\overline{\mathbf{F}(z)} \tag{19}$$

where $\mathbf{F}(z) = d\mathbf{f}(z)/dz$.

For the present problem, $\mathbf{F}(z)$ has the form

$$\mathbf{F}_k(z) = \mathbf{c}_k^\infty + \mathbf{F}_{k0}(z), \quad k = 1,2 \tag{20}$$

where $\mathbf{c}_k^\infty$ is a constant vector; $\mathbf{F}_{k0}(z)$ is an unknown function vector in s^+ $(k=1)$ or in s^- $(k=2)$, and $\mathbf{F}_{k0}(\infty) = \mathbf{0}$.

First let us determine $\mathbf{c}_k^\infty$. It is obvious that $\mathbf{c}_k^\infty$ is the complex potential corresponding to two completely bonded half-planes subjected to the applied uniform loads at infinity. For the subproblem, the continuity of deformation and stress on the entire x_1-axis requires (cf. (18) and (19))

$$\hat{\mathbf{A}}_1\mathbf{c}_1^\infty + \overline{\hat{\mathbf{A}}_1}\overline{\mathbf{c}_1^\infty} = \hat{\mathbf{A}}_2\mathbf{c}_2^\infty + \overline{\hat{\mathbf{A}}_2}\overline{\mathbf{c}_2^\infty} = \hat{\boldsymbol{\varepsilon}}_1^\infty \tag{21}$$

$$\hat{\mathbf{B}}_1\mathbf{c}_1^\infty + \overline{\hat{\mathbf{B}}_1}\overline{\mathbf{c}_1^\infty} = \hat{\mathbf{B}}_2\mathbf{c}_2^\infty + \overline{\hat{\mathbf{B}}_2}\overline{\mathbf{c}_2^\infty} = \hat{\boldsymbol{\sigma}}_2^\infty \tag{22}$$

where

$$\hat{\boldsymbol{\varepsilon}}_1^\infty = [0 \;\; 0 \;\; 0 \;\; D_2^\infty]^T, \quad \hat{\boldsymbol{\sigma}}_2^\infty = [\sigma_{21}^\infty \;\; \sigma_{22}^\infty \;\; \sigma_{23}^\infty \;\; 0]^T.$$

With respect to (15), from (21) and (22) we achieve

$$\mathbf{c}_k^\infty = \hat{\mathbf{B}}_k^T\hat{\boldsymbol{\varepsilon}}_1^\infty + \hat{\mathbf{A}}_k^T\hat{\boldsymbol{\sigma}}_2^\infty = \mathbf{A}_k^T[\sigma_{21}^\infty \;\; \sigma_{22}^\infty \;\; \sigma_{23}^\infty \;\; D_2^\infty]^T, \quad (k=1,2). \tag{23}$$

The remaining task is to find $\mathbf{F}_{k0}(z)$. Using (16) and (17) it follows

$$\hat{\boldsymbol{\phi}}_{,1}(x_1^+) = \hat{\boldsymbol{\phi}}_{,1}(x_1^-), \quad -\infty < x_1 < \infty. \tag{24}$$

Substituting (19) and (20) into (24), together with (21) and (22), leads to

$$\hat{\mathbf{B}}_1\mathbf{F}_{10}(x_1^+) + \overline{\hat{\mathbf{B}}_1}\overline{\mathbf{F}}_{10}(x_1^-) = \hat{\mathbf{B}}_2\mathbf{F}_{20}(x_1^-) + \overline{\hat{\mathbf{B}}_2}\overline{\mathbf{F}}_{20}(x_1^+), \quad -\infty < x_1 < \infty. \tag{25}$$

Equation (25) expresses that, [24],

$$\hat{\mathbf{B}}_1\mathbf{F}_{10}(z) - \overline{\hat{\mathbf{B}}_2}\overline{\mathbf{F}}_{20}(z) = \mathbf{0}, \quad z \in s_1 \tag{26}$$

$$\hat{\mathbf{B}}_2\mathbf{F}_{20}(z) - \overline{\hat{\mathbf{B}}_1}\overline{\mathbf{F}}_{10}(z) = \mathbf{0}, \quad z \in s_2. \tag{27}$$

Based on (18), we introduce a jump function:

$$i\Delta\hat{\mathbf{u}}_{,1} = i[\hat{\mathbf{u}}_{,1}(x_1^+) - \hat{\mathbf{u}}_{,1}(x_1^-)] = i[\hat{\mathbf{A}}_1\mathbf{F}_1(x_1^+) + \overline{\hat{\mathbf{A}}\mathbf{F}_1(x_1^+)}] - i[\hat{\mathbf{A}}_2\mathbf{F}_2(x_1^-) + \overline{\hat{\mathbf{A}}_2}\overline{\mathbf{F}_2(x_1^-)}]. \tag{28}$$

To rewrite Eqs. (19) and (28), we use (20), (21), (22), (26), and (27):

$$i\Delta\hat{\mathbf{u}}_{,1} = \mathbf{g}(x_1^+) - \mathbf{g}(x_1^-) \tag{29}$$

$$\hat{\boldsymbol{\phi}}_{,1}(x_1) = \hat{\boldsymbol{\sigma}}_2^\infty + \hat{\mathbf{H}}^{-1}\mathbf{g}(x_1^+) + \overline{\hat{\mathbf{H}}}^{-1}\mathbf{g}(x_1^-) \tag{30}$$

where

$$\mathbf{g}(z) = \begin{cases} \hat{\mathbf{H}}\hat{\mathbf{B}}_1\mathbf{F}_{10}(z), & z \in s_1 \\ \hat{\mathbf{H}}\overline{\hat{\mathbf{B}}}\overline{\hat{\mathbf{B}}}_2\mathbf{F}_{20}(z), & z \in s_2 \end{cases} \tag{31}$$

$$\hat{\mathbf{H}} = \hat{\mathbf{Y}}_1 + \overline{\hat{\mathbf{Y}}}_2, \quad \hat{\mathbf{Y}}_k = i\hat{\mathbf{A}}_k\hat{\mathbf{B}}_k^{-1}. \tag{32}$$

It can be shown that $\hat{\mathbf{Y}}_k$ and $\hat{\mathbf{H}}$ are positive definite, [25].

The continuity of the displacement on L_b and the jump of the displacement across the cracks give

$$[i\Delta\hat{\mathbf{u}}_{,1}(x_1)]_j \begin{cases} = 0, x_1 \in L_b \\ \neq 0, x_1 \in L_c \end{cases}, \quad j = 1,2,3. \tag{33}$$

Substituting (29) into (33) leads to

$$g_j(x_1^+) = g_j(x_1^-), \quad j = 1,2,3, \quad x_1 \in L_b \tag{34}$$

which means that the $g_j(z)$ are analytical in the entire z-plane except on the cracks.

Writing out the components of (30), we have

$$\hat{\phi}_{i,1}(x_1) = \hat{\sigma}_{2i}^\infty + \sum_{j=1}^4 h_{ij}g_j(x_1^+) + \sum_{j=1}^4 \overline{h_{ij}}g_j(x_1^-), \quad i = 1,2,3 \tag{35}$$

$$\hat{\phi}_{4,1}(x_1) = \sum_{j=1}^4 h_{4j}g_j(x_1^+) + \sum_{j=1}^4 \overline{h_{4j}}g_j(x_1^-) = 0, \quad -\infty < x_1 < +\infty \tag{36}$$

where $\mathbf{h} = \hat{\mathbf{H}}^{-1}$.

Multiplying both sides of (36) with $1/2\pi i\int_{-\infty}^{+\infty} dx_1/(x_1 - z)$, and using Cauchy integration theory, [24], we obtain

$$\sum_{j=1}^4 h_{4j}g_j(z) = 0, \quad z \in s^+ \tag{37}$$

$$\sum_{j=1}^4 \overline{h_{4j}}g_j(z) = 0, \quad z \in s^-. \tag{38}$$

Hence,

$$g_4(z) = \begin{cases} -\dfrac{1}{h_{44}}\displaystyle\sum_{j=1}^3 h_{4j}g_j(z), & z \in s^+ \\ -\dfrac{1}{h_{44}}\displaystyle\sum_{j=1}^3 \overline{h_{4j}}g_j(z), & z \in s^- \end{cases} \tag{39}$$

where h_{44} is real and greater than zero.

Inserting (39) into (35) gives

$$\hat{\phi}_{i,1}(x_1) = \hat{\sigma}_{2i}^\infty + \sum_{j=1}^3 \Lambda_{ij}g_j(x_1^+) + \sum_{j=1}^3 \overline{\Lambda_{ij}}g_j(x_1^-) \tag{40}$$

where

$$\Lambda_{ij} = h_{ij} - \frac{h_{i4}}{h_{44}}h_{4j}, \quad i,j = 1,2,3. \tag{41}$$

It can be shown that the (3,3) matrix $\boldsymbol{\Lambda} = [\Lambda_{ij}]$ is positive definite and is equivalent to $(\hat{\mathbf{H}}_{33})^{-1}$, in which $\hat{\mathbf{H}}_{33}$ is the upper left (3,3) hand block of $\hat{\mathbf{H}}$.

With

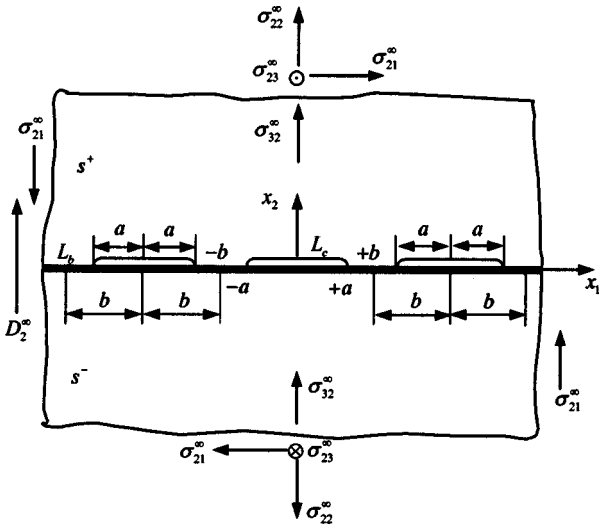


Fig. 2 Collinear periodic electrode-ceramic interfacial cracks

$$\boldsymbol{\phi}_{\sigma,1}=[\hat{\phi}_{1,1}\quad \hat{\phi}_{2,1}\quad \hat{\phi}_{3,1}]^T,\ \boldsymbol{\sigma}_2^{\infty}=[\sigma_{21}^{\infty}\quad \sigma_{22}^{\infty}\quad \sigma_{23}^{\infty}]^T,$$
$$\text{and}\ \ \mathbf{g}_{\sigma}=[g_1\quad g_2\quad g_3]^T$$

(40) reads in matrix notation as

$$\boldsymbol{\phi}_{\sigma,1}(x_1)=\boldsymbol{\sigma}_2^{\infty}+\boldsymbol{\Lambda}\mathbf{g}_{\sigma}^{+}+\overline{\boldsymbol{\Lambda}}\mathbf{g}_{\sigma}^{-}. \tag{42}$$

On the crack faces the stress boundary condition is

$$\boldsymbol{\phi}_{\sigma,1}(x_1)=\mathbf{0},\quad x_1\in L_c. \tag{43}$$

Substituting (43) into (42) results in

$$\boldsymbol{\Lambda}\mathbf{g}_{\sigma}^{+}+\overline{\boldsymbol{\Lambda}}\mathbf{g}_{\sigma}^{-}=-\boldsymbol{\sigma}_2^{\infty},\quad x_1\in L_c. \tag{44}$$

To determine the complete solution of (44) we have to consider the condition of unambiguous displacement. This requires

$$\int_{L_c}[\Delta\hat{\mathbf{u}}_{,1}(x_1)]_j dx_1=0,\ \ j=1,2,3. \tag{45}$$

Because the cracks are permeable, the electric potential is automatically unambiguous, unlike the displacement. Inserting (29) into (45) leads to

$$\int_{L_c}[\mathbf{g}^{+}(x_1)-\mathbf{g}^{-}(x_1)]_j dx_1=0,\ \ j=1,2,3. \tag{46}$$

Up to here, the problem has been reduced to an interfacial crack problem equivalent to that in purely elastic media. It can be solved according to traditional methods (e.g., Muskhelishvili [24], Suo [23], and Boniface and Banks-Sills [26]).

4 Field Intensity Factors and Energy Release Rate

Now, as an example, the solution of (44) will be derived for the case of an infinite number of collinear periodical interfacial cracks with equal length $2a$ and equal interval $2b$ $(b>a)$ as shown in Fig. 2. In order to get the solution of (44) for an other system of collinear cracks we refer to the discussion in paragraph 6 of [23].

Equation (44) can be transformed into

$$\mathbf{g}_{\sigma}^{+}=\boldsymbol{\Lambda}^{-1}\overline{\boldsymbol{\Lambda}}\mathbf{g}_{\sigma}^{-}=-\boldsymbol{\Lambda}^{-1}\boldsymbol{\sigma}_2^{\infty},\quad x_1\in L_c. \tag{47}$$

For convenience, we introduce

$$\mathbf{Q}^{-1}\boldsymbol{\Lambda}^{-1}\overline{\boldsymbol{\Lambda}}\mathbf{Q}=\mathrm{diag}(e^{2\pi\varepsilon}\quad e^{-2\pi\varepsilon}\quad 1)=\mathrm{diag}(e^{2\pi\varepsilon_a}) \tag{48}$$

where $\mathbf{Q}$ is a constant matrix. The positive and real constant ε depends on the material properties and characterizes the oscillatory singularities [27].

The general solution of (47) is

$$\mathbf{Q}^{-1}\mathbf{g}_{\sigma}(z)=-\mathrm{diag}\left(\frac{1}{1+e^{2\pi\varepsilon_a}}\right)\mathbf{R}_0+\mathrm{diag}\left(\frac{X_{\alpha}(z)X_{\alpha}^{-1}(\infty)}{1+e^{2\pi\varepsilon_{\alpha}}}\right)\mathbf{R}_0 \tag{49}$$

where

$$X_a^{-1}(\infty)X_{\alpha}(z)=\frac{\sin\left[\frac{\pi}{2b}(z-2ia\varepsilon_{\alpha})\right]\sec\frac{\pi a}{2b}\sec\frac{\pi z}{2b}}{\sqrt{tg^2\frac{\pi z}{2b}-tg^2\frac{\pi a}{2b}}}\times\left[\frac{tg\frac{\pi z}{2b}-tg\frac{\pi a}{2b}}{tg\frac{\pi z}{2b}+tg\frac{\pi a}{2b}}\right]^{-i\varepsilon_{\alpha}},$$

$$\mathbf{R}_0=[\boldsymbol{\Lambda}\mathbf{Q}]^{-1}\boldsymbol{\sigma}_2^{\infty}.$$

The derivation of the formula (49) is given in the Appendix.

In order to calculate the field intensity factors we define a vector as, [28],

$$\mathbf{k}_{\sigma}=\lim_{r\to 0}\sqrt{2\pi r}[\boldsymbol{\Lambda}\mathbf{Q}]\mathrm{diag}(r^{i\varepsilon_{\alpha}})[\boldsymbol{\Lambda}\mathbf{Q}]^{-1}\boldsymbol{\sigma}_2(r) \tag{50}$$

where r means the distance from the crack-tip and $\boldsymbol{\sigma}_2(r)$ stands for the stress vector $\boldsymbol{\sigma}_2(x_1)$ ahead of the crack tip for a coordinate system moved to any right crack tip.

Noting (34), it results from (42) that

$$[\boldsymbol{\Lambda}\mathbf{Q}]^{-1}\sigma_2(r)=\mathrm{diag}(1+e^{2\pi\varepsilon_{\alpha}})\mathbf{Q}^{-1}\mathbf{g}_{\sigma}(r)+\mathbf{R}_0,\quad r=x_1-a-2nb\in L_b. \tag{51}$$

Substituting (51) with respect to (49) into (50) yields

$$\mathbf{k}_{\sigma}=\sqrt{\pi a}[\boldsymbol{\Lambda}\mathbf{Q}]\mathrm{diag}((2a)^{i\varepsilon_{\alpha}}\omega_{\alpha})[\boldsymbol{\Lambda}\mathbf{Q}]^{-1}\boldsymbol{\sigma}_2^{\infty} \tag{52}$$

where

$$\omega_{\alpha}=\sec\frac{\pi a}{2b}\sin\left[\frac{\pi a}{2b}(1-2i\varepsilon_{\alpha})\right]\left(\frac{\pi a}{2b}\right)^{-i\varepsilon_{\alpha}-1/2}\left(tg\frac{\pi a}{2b}\right)^{i\varepsilon_{\alpha}-1/2}. \tag{53}$$

The bimaterial intensity factors are following from Eq. (52) by multiplying it with $[\boldsymbol{\Lambda}\mathbf{Q}]^{-1}$ (cf. [13,23]):

$$[\boldsymbol{\Lambda}\mathbf{Q}]^{-1}\mathbf{k}_{\sigma}=\mathbf{k}=\sqrt{\pi a}\,\mathrm{diag}((2a)^{i\varepsilon_{\alpha}}\omega_{\alpha})\mathbf{R}_0,\quad \mathbf{R}_0=[\boldsymbol{\Lambda}\mathbf{Q}]^{-1}\boldsymbol{\sigma}_2^{\infty}. \tag{54}$$

The relations between the components of the vectors $\mathbf{k}=[K_1\ K_2\ K_3]^T$ and $\mathbf{R}_0=[R_{01}\ R_{02}\ R_{03}]^T$, respectively, will be given below.

If $b\to\infty$, from (53) it follows that

$$\omega_{\alpha}=1-2i\varepsilon_{\alpha} \tag{55}$$

and thus (54) becomes

$$\mathbf{k}=\sqrt{\pi a}\,\mathrm{diag}((2a)^{i\varepsilon_{\alpha}}(1-2i\varepsilon_{\alpha}))\mathbf{R}_0 \tag{56}$$

which is the corresponding solution of a single electrode-ceramic interface crack. In addition, (54) and (56) show that $\mathbf{k}$ is independent of the applied electric load.

If $b\to a$, from (53) by setting $b=a+c$ and $c\ll a$ we have

$$\mathbf{k}=\sqrt{\pi a}\,\mathrm{diag}\left((2a)^{i\varepsilon_{\alpha}}\left(\frac{\pi}{2}\right)^{-(1+2i\varepsilon_{\alpha})}e^{i\varepsilon_{\alpha}\ln a/c}\sqrt{\frac{a}{c}}\right)\mathbf{R}_0. \tag{57}$$

Here we additionally assumed $\varepsilon\ll 1$, which holds normally. It can be seen that the amount of the intensity factors increases rapidly by $\sqrt{a/c}$ while the complex intensity factors are rotated by the angle $\varepsilon_{\alpha}\ln a/c$ if the crack tips get closer to each other.

Further, for $\varepsilon \to 0$ (as in a homogeneous material) we obtain

$$\mathbf{k} = \sqrt{\frac{2b}{\pi a} tg \frac{\pi a}{2b}} \sqrt{\pi a}\mathbf{R}_0 \quad \text{and} \quad \mathbf{k}_\sigma = \sqrt{\frac{2b}{\pi a} tg \frac{\pi a}{2b}} \sqrt{\pi a}\boldsymbol{\sigma}_2^\infty ,$$

respectively. The latter is the known solution of a periodical array of cracks in a homogeneous material.

The intensity factors are related as follows. From $[\boldsymbol{\Lambda}\mathbf{Q}]\mathbf{R}_0 = \boldsymbol{\sigma}_2^\infty$ and (48) it can be seen that $\mathbf{Q}\mathbf{R}_0 = \overline{\mathbf{Q}}\,\mathrm{diag}(e^{-2\pi\varepsilon_\alpha})\overline{\mathbf{R}}_0$. Since $\mathbf{Q} = [\overline{\mathbf{q}}\ \mathbf{q}\ \mathbf{q}_3]$, where $\mathbf{q}_3$ can be chosen real, [23],

$$\overline{\mathbf{Q}} = \mathbf{Q}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

applies and thus

$$R_{01} = e^{2\pi\varepsilon}\overline{R}_{02} \quad R_{03} = \overline{R}_{03} \quad \text{and} \quad K_1 = e^{2\pi\varepsilon}\overline{K}_2 \quad K_3 = \overline{K}_3 . \quad (58)$$

Because the cracks are permeable, the energy release rate follows from the purely mechanical crack closure integral, [23]. From it we obtain

$$G = \frac{e^{2\pi\varepsilon}\mathbf{q}^T(\boldsymbol{\Lambda} + \overline{\boldsymbol{\Lambda}})\overline{\mathbf{q}}K\overline{K}}{4\cosh^2 \pi\varepsilon} + \frac{1}{8}\mathbf{q}_3^T(\boldsymbol{\Lambda} + \overline{\boldsymbol{\Lambda}})\mathbf{q}_3 K_3^2 \quad (59)$$

where $K = K_2$. The Eq. (58) shows that G is always greater than zero except for $K = K_3 = 0$.

Below a short outlook is given to interfacial fracture mechanics. In consideration of (58) we can unambiguously define a mixed mode angle ψ as

$$\psi = \tan^{-1}\left(\frac{\mathrm{Im}(Kl^{i\varepsilon})}{\mathrm{Re}(Kl^{i\varepsilon})}\right), \quad (60)$$

where l is a reference length, whose choice is discussed in [29]. A second mixed mode angle, introduced by Suo [[25], Section 11], can be defined as

$$\phi = \tan^{-1}\left(\frac{|Kl^{i\varepsilon}|}{K_3}\right). \quad (61)$$

According to [25], in terms of ψ, l, and ϕ we propose a criterion for the initiation of crack advance in the interface in form of $G = \Gamma(\psi, l, \phi)$, where Γ is the toughness of the interface.

For the most important situation in practical applications, where the poling directions of the (transversal isotropic) piezoelectric materials are in plane and no antiplane load exists ($\sigma_{23}^\infty = 0$), $K_3 = 0$ applies. For it the energy release rate is given by

$$G = \frac{e^{2\pi\varepsilon}\mathbf{q}^T(\boldsymbol{\Lambda} + \overline{\boldsymbol{\Lambda}})\overline{\mathbf{q}}K\overline{K}}{4\cosh^2 \pi\varepsilon}. \quad (62)$$

Consequently, a criterion for initiation of crack advance in the interface can be stated as

$$G = \Lambda(\psi, l) \quad (63)$$

analogous to the criterion for the initiation of crack advance in the interface of two (isotropic) elastic materials, [29].

The toughness of the interface $\Gamma(\psi, l, \phi)$ and $\Gamma(\psi, l)$, respectively, must be determined by experiments. As our results show for permeable electrode-ceramic interfacial cracks, these experiments can be purely mechanical experiments. Nevertheless, in a finite structural component the piezoelectric coupling could produce stress due to the influence of the boundaries or due to an applied inhomogeneous electric field. This stress could cause field intensity factors. For this reason the electrical loads can't be neglected a priori in a finite structural component.

Table 1 Material properties of PZT4 and PZT5H (c_{ij} in 10^4 N/mm^2, e_{ij} in 10^3 nAs/mm^2, and ε_{ij} in 10^3 nAs/MV mm)

	c_{11}	c_{12}	c_{13}	c_{33}	c_{44}
PZT4	13.9	7.78	7.74	11.3	2.56
PZT5H	12.6	5.50	5.30	11.7	3.53
	e_{31}	e_{33}	e_{15}	ε_{11}	ε_{33}
PZT4	−6.98	13.84	13.44	6.00	5.47
PZT5H	−6.50	23.30	17.00	15.1	13.0

5 Numerical Examples

In the following numerical examples transversely isotropic piezoelectric materials are considered. The poling direction is assumed to be parallel to the x_2-axis so that the matrices of Eq. (5) yield to (note that Voigt notation is used here)

$$\mathbf{T} = \begin{bmatrix} c_{11} & 0 & 0 & 0 \\ 0 & c_{44} & 0 & e_{15} \\ 0 & 0 & \dfrac{c_{11} - c_{12}}{2} & 0 \\ 0 & e_{15} & 0 & -\varepsilon_{11} \end{bmatrix}, \quad (64)$$

$$\mathbf{R} = \begin{bmatrix} 0 & c_{13} & 0 & e_{31} \\ c_{44} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ e_{15} & 0 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{W} = \begin{bmatrix} c_{44} & 0 & 0 & 0 \\ 0 & c_{33} & 0 & e_{33} \\ 0 & 0 & c_{44} & 0 \\ 0 & e_{33} & 0 & -\varepsilon_{33} \end{bmatrix}.$$

In the numerical calculation we use for the material in the upper half-plane PZT4 and for the material in the lower half-plane PZT5H. The material properties are given in Table 1.

The numerical solution of (5) with respect to (9), (14), (32), (41), and (48) provides

$$\boldsymbol{\Lambda} = \begin{bmatrix} 24425.0 & 1312.7i & 0 \\ -1312.7i & 23189.0 & 0 \\ 0 & 0 & 15630.5 \end{bmatrix} \frac{\mathrm{N}}{\mathrm{mm}^2},$$

$$\mathbf{Q} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.513152i & -0.513152i & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathrm{diag}(c_1\ \ c_2\ \ c_3), \quad (65)$$

$$\mathbf{Q}^{-1}\boldsymbol{\Lambda}^{-1}\overline{\boldsymbol{\Lambda}}\mathbf{Q} = \mathrm{diag}(1.11675\ \ 0.895455\ \ 1)$$

with $\varepsilon = 0.0175745$. The constants c_i $(i = 1,2,3)$ are arbitrary except zero. As mentioned above, because of the in-plane-poling of the transversely piezoelectric materials, the fields in x_3-direction decouple from the in-plane fields.

For comparison, without an internal electrode layer we would have

$$\boldsymbol{\Lambda} = \begin{bmatrix} 27748.8 & 1629.7i & 0 \\ -1629.7i & 23219.3 & 0 \\ 0 & 0 & 15630.5 \end{bmatrix} \frac{\mathrm{N}}{\mathrm{mm}^2},$$

$$\mathbf{Q} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.546598i & -0.546598i & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathrm{diag}(c_1\ \ c_2\ \ c_3), \quad (66)$$

$$\mathbf{Q}^{-1}\boldsymbol{\Lambda}^{-1}\overline{\boldsymbol{\Lambda}}\mathbf{Q} = \mathrm{diag}(1.13722\ \ 0.879339\ \ 1)$$

with $\varepsilon = 0.0204649$ as shown in [30]. Note, although the structure of the results for interfacial cracks with and without an internal

electrode layer are identical (cf. [[30], Eq. (34)] and (44)), the ways of solving the problems and the involved matrices are different.

On the other hand, without an piezoelectric effect by setting $e_{ij}=0$, as for (permeable) interfacial cracks in anisotropic elastic-dielectric bimaterials with or without an internal electrode layer, the matrices are given by

$$\mathbf{\Lambda}=\begin{bmatrix} 22939.0 & 391.15i & 0 \\ -391.15i & 21303.4 & 0 \\ 0 & 0 & 15630.5 \end{bmatrix}\frac{\mathrm{N}}{\mathrm{mm}^2},$$

$$\mathbf{Q}=\begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.518839i & -0.518839i & 0 \\ 0 & 0 & 1 \end{bmatrix}\mathrm{diag}(c_1\ \ c_2\ \ c_3), \quad (67)$$

$$\mathbf{Q}^{-1}\mathbf{\Lambda}^{-1}\overline{\mathbf{\Lambda}}\mathbf{Q}=\mathrm{diag}(1.03603\ \ 0.965227\ \ 1)$$

with $\varepsilon=0.0056328$.

For a single interface crack ($b\to\infty$) the field intensity factor $K_2=K$ is given by Eq. (56). With the use of the matrices from Eqs. (65), (66), and (67) ($c_1,c_2=1$) and only considering the in-plane loads $\sigma_{21}^{\infty}=\tilde{\sigma}_{21}\mathrm{N/mm}^2$ and $\sigma_{22}^{\infty}=\tilde{\sigma}_{22}\mathrm{N/mm}^2$, it follows that

$$\begin{aligned} &\text{PZT4/EI./PZT5H} && K=\sqrt{\pi a}(2a)^{-i\varepsilon}(1+2i\varepsilon)(0.38802\tilde{\sigma}_{21} \\ &&& \quad +i0.39822\tilde{\sigma}_{22}10^{-4} \\ &\text{PZT4/PZT5H} && K=\sqrt{\pi a}(2a)^{-i\varepsilon}(1+2i\varepsilon)(0.33863\tilde{\sigma}_{21} \\ &&& \quad +i0.37019\hat{\sigma}_{22}10^{-4} \\ &e_{ij}=0 && K=\sqrt{\pi a}(2a)^{-i\varepsilon}(1+2i\varepsilon)(0.42836\tilde{\sigma}_{21} \\ &&& \quad +i0.44450\tilde{\sigma}_{22}10^{-4}, \end{aligned} \quad (68)$$

respectively. Setting $a=\tilde{a}$ mm the corresponding energy release rates are given by (62)

$$\text{PZT4/EI./PZT5H}\quad G=\tilde{a}(0.35954\tilde{\sigma}_{21}^2+0.37870\tilde{\sigma}_{22}^2)10^{-4}\frac{\mathrm{N}}{\mathrm{mm}}$$

$$\text{PZT4/PZT5H}\quad G=\tilde{a}(0.32242\tilde{\sigma}_{21}^2+0.38531\tilde{\sigma}_{22}^2)10^{-4}\frac{\mathrm{N}}{\mathrm{mm}} \quad (69)$$

$$e_{ij}=0\quad G=\tilde{a}(0.35476\tilde{\sigma}_{21}^2+0.38200\tilde{\sigma}_{22}^2)10^{-4}\frac{\mathrm{N}}{\mathrm{mm}}.$$

It can be seen that the electrical boundary condition in the interface as well as the piezoelectric performance of the materials influences the field intensity factors and the energy release rates. The energy release rates with and without an internal electrode layer differ about 12%, considering the terms of (69) which contain $\tilde{\sigma}_{21}^2$. However, the difference between the energy release rates is only about 1% for the case of an internal electrode layer compared to the case $e_{ij}=0$.

6 Conclusions

We studied a generalized two-dimensional problem of collinear (periodic) electrode-ceramic interfacial cracks in piezoelectric bimaterials. Based on the Stroh formalism for mixed boundary conditions, the problem is reduced to an interfacial crack problem equivalent to that in purely elastic media. The solutions of the field intensity factors are obtained in very explicit form. It is found that the field intensity factors and the energy release rate are dependent on the material properties and the mechanical fields, but not directly on the electric fields.

Acknowledgments

The authors would like to express their gratitude for the support of the German Research Foundation and the Alexander von Humboldt Foundation (Germany).

Appendix

The Derivation of Eq. (49). Multiplying both sides of (47) by $\mathbf{Q}^{-1}$ and using (48) we have

$$\mathbf{R}^{+}(x_1)+\mathrm{diag}(e^{2\pi\varepsilon_\alpha})\mathbf{R}^{-}(x_1)=-\mathbf{R}_0 \quad (A1)$$

where $\mathbf{R}(z)$ is a function vector and $\mathbf{R}_0$ is a constant vector, which are defined as

$$\mathbf{R}(z)=\mathbf{Q}^{-1}\mathbf{g}_\sigma(z),\quad \mathbf{R}_0=\mathbf{Q}^{-1}\mathbf{\Lambda}^{-1}\boldsymbol{\sigma}_2^{\infty}.$$

Writing the components of $(A1)$ in separate equations results in

$$R_\alpha^{+}(x_1)+e^{2\pi\varepsilon_\alpha}R_\alpha^{-}(x_1)=-R_{\alpha 0},\quad \alpha=1,2,3. \quad (A2)$$

The general solution of $(A2)$ is, [24],

$$R_\alpha(z)=\frac{-R_{\alpha 0}}{1+e^{2\pi\varepsilon_\alpha}}+\left[\frac{X_\alpha(z)X_\alpha^{-1}(\infty)}{1+e^{2\pi\varepsilon_\alpha}}\right]R_{\alpha 0}+X_\alpha(z)P_\alpha(z) \quad (A3)$$

where $X_\alpha^{-1}(\infty)$ stands for the principle part of $X_\alpha^{-1}(z)$ at infinity, $P_\alpha(z)$ is a polynomial, and

$$X_\alpha(z)=\prod_{n=-\infty}^{\infty}[(z-2nb)-a]^{-1/2-i\varepsilon_\alpha}[(z-2nb)+a]^{-1/2+i\varepsilon_\alpha}, \quad (A4)$$

$$X_\alpha^{-1}(\infty)=\prod_{n=-\infty}^{\infty}[z-2nb-2ia\varepsilon_\alpha].$$

For the present case it can be shown that

$$\mathrm{P}_\alpha(z)=0. \quad (A5)$$

Thus, $(A3)$ reads in vector notation as

$$\mathbf{Q}^{-1}\mathbf{g}_\sigma(z)=-\mathrm{diag}\left(\frac{1}{1+e^{2\pi\varepsilon_\alpha}}\right)\mathbf{R}_0+\mathrm{diag}\left(\frac{X_\alpha(z)X_\alpha^{-1}(\infty)}{1+e^{2\pi\varepsilon_\alpha}}\right)\mathbf{R}_0. \quad (A6)$$

Using the identity

$$\sin\pi z=\pi z\prod_{n=1}^{\infty}\left(1-\frac{z^2}{n^2}\right)$$

we have

$$\begin{aligned} &X_\alpha^{-1}(\infty)X_\alpha(z) \\ &\quad=\frac{\Pi_{n=-\infty}^{\infty}[z-2nb-2ia\varepsilon_\alpha]\Pi_{n=-\infty}^{\infty}[(z-2nb)-a]^{-1/2-i\varepsilon_\alpha}}{\Pi_{n=-\infty}^{\infty}[(z-2nb)+a]^{1/2-i\varepsilon_\alpha}} \\ &\quad=\sin\left[\frac{\pi(z-2ia\varepsilon_\alpha)}{2b}\right]\left\{\sin\left[\frac{\pi(z-a)}{2b}\right]\right\}^{-1/2-i\varepsilon_\alpha} \\ &\qquad\times\left\{\sin\left[\frac{\pi(z+a)}{2b}\right]\right\}^{-1/2+i\varepsilon_\alpha}. \end{aligned} \quad (A7)$$

Using

$$\sin\left[\frac{\pi(z\pm a)}{2b}\right]=\sin\frac{\pi z}{2b}\cos\frac{\pi a}{2b}\pm\cos\frac{\pi z}{2b}\sin\frac{\pi a}{2b}$$

Eq. $(A7)$ can be transformed into

$$X_\alpha^{-1}(\infty)X_\alpha(z)=\frac{\sin\left[\frac{\pi}{2b}(z-2ia\varepsilon_\alpha)\right]\sec\frac{\pi a}{2b}\sec\frac{\pi z}{2b}}{\sqrt{tg^2\frac{\pi z}{2b}-tg^2\frac{\pi a}{2b}}}\times\left[\frac{tg\frac{\pi z}{2b}-tg\frac{\pi a}{2b}}{tg\frac{\pi z}{2b}+tg\frac{\pi a}{2b}}\right]^{-i\varepsilon_\alpha}. \quad (A8)$$

This concludes the derivation of (49).

References

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

[2] Yang, W., and Suo, Z., 1994, ''Cracking in Ceramic Actuators Caused by Electrostriction,'' J. Mech. Phys. Solids, **42**, pp. 649–663.

[3] Hao, T. H., Gong, X., and Suo, Z., 1996, ''Fracture Mechanics for the Design of Ceramic Multilayer Actuators,'' J. Mech. Phys. Solids, **44**, pp. 23–48.

[4] Shindo, Y., Narita, F., and Sosa, H., 1998, ''Electroelastic Analysis of Piezoelectric Ceramics With Surface Electrodes,'' Int. J. Eng. Sci., **36**, pp. 1001–1009.

[5] Ye, R. Q., and He, L. H., 2001, ''Electric Field and Stresses Concentrations at the Edge of Parallel Electrodes in Piezoelectric Ceramics,'' Int. J. Solids Struct., **38**, pp. 6941–6951.

[6] Dos Santos e Lucato, S. L., Lupascu, D. C., Kamlah, M., Rödel, J., and Lynch, C. S., 2001, ''Constraint-Induced Crack Initiation at Electrode Edges in Piezoelectric Ceramics,'' Acta Mater., **49**, pp. 2751–2759.

[7] Deng, D., and Meguid, S. A., 1998, ''Analysis of Conducting Rigid Inclusion at the Interface of Two Dissimilar Piezoelectric Materials,'' ASME J. Appl. Mech., **65**, pp. 76–84.

[8] Ru, C. Q., 2000, ''Exact Solution for Finite Electrode Layers Embedded at the Interface of Two Piezoelectric Half-Planes,'' J. Mech. Phys. Solids, **48**, pp. 693–708.

[9] Weitzing, H., 2000, ''Phenomena of Damage and Fracture-Mechanical Characterization of Piezoelectric Multilayer Actuators,'' dissertation Technische Universität Hamburg-Harburg (in German).

[10] Zickgraf, B., 1995, ''Fatigue of Piezoelectric Multilayer Actuators,'' dissertation Universität Stuttgart (in German).

[11] Ru, C. Q., 2000, ''Electrode-Ceramic Interfacial Cracks in Piezoelectric Multilayer Materials,'' ASME J. Appl. Mech., **67**, pp. 255–261.

[12] Wang, X., and Shen, Y. P., 2002, ''Exact Solution for Mixed Boundary Value Problems at Anisotropic Piezoelectric Bimaterial Interface and Unification of Various Interface Defects,'' Int. J. Solids Struct., **39**, pp. 1591–1619.

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

[14] Ting, T. C. T., and Wang, M. Z., 1992, ''Generalized Stroh Formalism for Anisotropic Elasticity for General Boundary Conditions,'' Acta Mech. Sin., **8**, pp. 193–207.

[15] Wang, M. Z., Ting, T. C. T., and Yan, Gongpu, 1993, ''The Anisotropic Elastic Semi-Infinite Strip,'' Q. Appl. Math., **51**, pp. 283–297.

[16] Homulka, T. A., and Keer, L. M., 1995, ''A Mathematical Solution of a Special Mixed Boundary Value Problem of Anisotropic Elasticity,'' Q. J. Mech. Appl. Math., **48**, pp. 635–658.

[17] Häusler, C., and Balke, H., 2001, ''An Interface Crack Between a Piezoelectric Medium and an Electric Conductor,'' *Contributions to Modeling and Identification*, P. Haupt, T. Kersten, and V. Ulbricht, eds., Gesamthochschule, Kassel, pp. 57–66 (in German).

[18] Parton, V. Z., 1976, ''Fracture Mechanics of Piezoelectric Materials,'' Acta Astronaut., **3**, pp. 671–683.

[19] Balke, H., Kemmer, G., and Drescher, J., 1997, ''Some Remarks on Fracture Mechanics of Piezoelectric Solids,'' *Proceedings of the International Conference and Exhibition of Micro Materials '97*, B. Michel, and T. Winkler, eds., Berlin, pp. 398–401.

[20] Park, S., and Sun, C.-T., 1995, ''Fracture Criteria for Piezoelectric Ceramics,'' J. Am. Ceram. Soc., **78**(6), pp. 1475–1480.

[21] Hao, T.-H., and Shen, Z.-Y., 1994, ''A New Electric Boundary Condition of Electric Fracture Mechanics and Its Applications,'' Eng. Fract. Mech., **47**(6), pp. 793–802.

[22] Schneider, G. A., Felten, F., and McMeeking, R. M., 2003, ''The Electrical Potential Difference Across Cracks in PZT Measured by Kelvin Probe Microscopy and the Implications for Fracture,'' Acta Mater., **51**, pp. 2235–2241.

[23] Suo, Z., 1990, ''Singularities, Interfaces and Cracks in Dissimilar Anisotropic Media,'' Proc. R. Soc. London, Ser. A, **A427**(1873), pp. 331–358.

[24] Muskhelishvili, N. I., 1975, *Some Basic Problems of Mathematical Theory of Elasticity*, Noordhoff, Leyden.

[25] Suo, Z., 1993, ''Models for Breakdown-Resistant Dielectric and Ferroelectric Ceramics,'' J. Mech. Phys. Solids, **47**, pp. 1155–1176.

[26] Boniface, V., and Banks-Sills, L., 2002, ''Stress Intensity Factors for Finite Interface Cracks Between a Special Pair of Transversely Isotropic Materials,'' ASME J. Appl. Mech., **69**, pp. 230–239.

[27] Ting, T. C. T., 1986, ''Explicit Solution and Invariance of the Singularities at an Interface Crack in Anisotropic Composites,'' Int. J. Solids Struct., **22**, pp. 965–983.

[28] Wu, K. C., 1990, ''Stress Intensity Factors and Energy Release Rate for Interfacial Cracks Between Dissimilar Anisotropic Materials,'' ASME J. Appl. Mech., **57**, pp. 882–886.

[29] Hutchinson, J.-W., and Suo, Z., 1992, ''Mixed Mode Cracking in Layered Materials,'' Adv. Appl. Mech., **29**, pp. 63–191.

[30] Gao, C.-F., Häusler, C., and Balke, H., 2004, ''Periodic Permeable Interface Cracks in Piezoelectric Materials,'' Int. J. Solids Struct., **41**(2), pp. 323–335.

Katalin Bagi
HAS Research Group for Computational Structural Mechanics,
Department of Structural Mechanics,
Budapest University of Technology and Economics,
Muegyetem rkp. 3. K.mf.35,
H-1521 Budapest, Hungary
e-mail: kbagi@mail.bme.hu

Matthew R. Kuhn
Department of Civil and Environmental Engineering,
University of Portland,
5000 N. Willamette Boulevard,
Portland, OR 97203
e-mail: kuhn@up.edu

A Definition of Particle Rolling in a Granular Assembly in Terms of Particle Translations and Rotations

The paper presents a definition of particle rolling for the interactions of two and three-dimensional particles of arbitrary shape, in case of infinitesimal particle translations and rotations. The definition is based on a purely kinematical analysis, and it is shown to satisfy the objectivity condition. [DOI: 10.1115/1.1755693]

1 Introduction

Granular assemblies consist of individual particles, each having its own translational and rotational degrees-of-freedom. In addition to elastic deformations, the particles are able to slide as well as to roll across each other, and both mechanisms have their role in the global deformation of the assembly. The meaning of contact deformation and sliding is unambiguous, but this is not the case for the concept of *rolling*. The aim of this paper is to give a general definition of rolling between particles.

We shall consider particles whose stiffness is large enough to ensure that the particles retain their shapes except in the small vicinity of their contacts. It will also be assumed that the contact surfaces are so small that in their geometrical description, the contacts can be modeled as points.

When two particles are in contact, and move relative to each other, a wide range of mechanical phenomena may occur at the contact point. If the two contacting particles move in such a way that the distance between any arbitrarily chosen points on the two particles remain constant during the motion, then the particle pair performs a *rigid-body* motion. Otherwise, two types of phenomena may occur in the contact: *deformation* and *rolling*.

(1.) Contact deformation
The neighborhoods of the two particles around the contact point may be deformed. From a kinematical point of view, contact deformations, whether elastic or dissipative, are usually separated into *relative translations* (both normal and tangential) and *relative rotations* (both bending and twisting).

These deformation types are well known and are widely applied in distinct element simulations for the description of contact behavior.

(2.) Rolling
Two particles may *roll* across each other without energy dissipation or storage. This possibility provides for extra degrees of freedom for a granular assembly. In extreme situations assemblies may be deformed without any (or, at least, with very small) energy investment, due to the ability of particles to roll along each other. The effect is also present in general cases, making the overall response much softer than it would be without the possibility of rolling motions (Iwashita and Oda [1]). Consequently, particle rolling plays a fundamental role in the deformation of granular assemblies.

Most of the literature related to particle rotations and to rolling phenomena focuses on assemblies of ideal circles or spheres. But since the extension of the results is questionable for real situations with nonideal particle shape, we only consider those studies that apply to more realistic particle shapes.

The effects of particle rotations and contact rolling—taking into account noncircular or nonspherical particle shapes—have also been analyzed by several authors. These studies have been performed on two-dimensional assemblies, through physical experiments (Oda, Konishi, and Nemat-Nasser [2]) as well as by computer simulations (Rothenburg and Bathurst [3], Ting et al. [4], and Matsushima and Konagai [5]). The literature also includes a few studies of three-dimensional numerical simulation results (for example, Lin and Ng [6]). The importance of rolling as a deformation mechanism has been clearly identified in these investigations, and the fact that the particle shape strongly influences the material behavior has also become clear.

In spite of all these experimental or numerical results, there has been no general definition of what is meant by a ''rolling displacement'' of two arbitrarily shaped, possibly three-dimensional particles. Different researchers may think of different phenomena when using the expression ''rolling displacement.'' Sometimes it is interpreted as the relocation of the contact point on the two touching surfaces, which requires principles of differential geometry to describe the local surface shapes (Montana [7]). Sometimes it is simply understood as the tangential component of the difference between the rotation vectors of the two particles. The rolling displacement can also be that part of the common translation of the contact point which originates from the particle rotations (Iwashita and Oda [1]).

We do not think that there is a unique way to define the rolling displacement. On the contrary: all of these (or other) approaches to define a rolling displacement may be correct and useful for certain applications, provided that the definition is given in an exact, mathematically clear manner.

We present a version of rolling displacement that can directly be expressed in terms of the translations of particle centroids and of the rotations of the particles about their centroids. In this way we derive a microlevel kinematical variable—assigned to the contacts—that can easily be applied in numerical or real experiments to analyze deformation mechanisms, and that can be the basis for defining a global kinematical state variable for rolling.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, October 22, 2002; final revision, November 25, 2003. Editor: R. M. Meeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

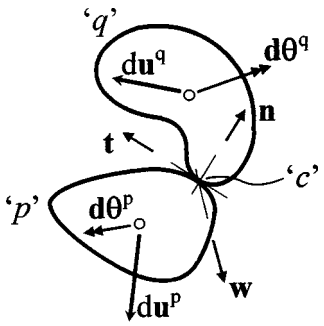


Fig. 1

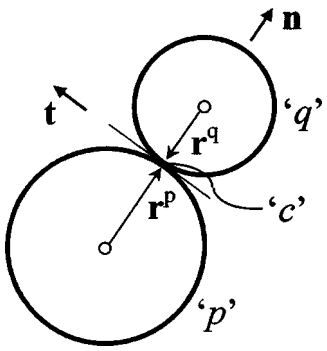


Fig. 2

To find such a definition of the rolling displacement, we extend the concept that was suggested by Oda et al. [2] and by Iwashita and Oda [1] for special situations, and develop a general definition for particles having any smooth convex or concave shape in both two dimensions and three dimensions. This general definition of rolling displacement will have the following three properties:

1. It is objective, and as a result of objectivity, its magnitude is unaffected by the common rigid-body-like motion of the analyzed particle pair.
2. It produces no tangential or normal deformation (sliding or indentation) at the contact in the sense that the two particles can move in a manner that produces rolling but no relative translation at the contact.
3. It reduces to the definition of rolling given by Iwashita and Oda in [1] when the two particles are circular discs.

We apply an incremental approach: the incremental rolling displacements belonging to a contact will be expressed in terms of the incremental rotations of the particles about their centroids and their incremental translations. In Section 2 we analyze the special two-dimensional case of circular particles and suggest a definition of what will be termed ''Iwashita-Oda rolling.'' A generalization for arbitrary two-dimensional particles will follow in Section 3. Finally, in Section 4, we introduce a definition for the general three-dimensional situation.

The usual Cartesian coordinates will be used. All formulas will be written in tensorial notations, with vectors and higher-order tensors denoted by bold characters, and scalars denoted by normal-weight characters. The scalar product of two vectors will be denoted by '$\cdot$' and the vector product by '$\times$':

$$c=\mathbf{a}\cdot\mathbf{b} \qquad \text{for the scalar product,}$$

and

$$\mathbf{c}=\mathbf{a}\times\mathbf{b} \qquad \text{for the vector product.}$$

Figure 1 illustrates the two particles, ''p'' and ''q'' forming a point-like contact ''c'' in a general three-dimensional situation. The incremental translations of the particle centroids are $d\mathbf{u}^p$ and $d\mathbf{u}^q$, and the incremental rotations of the particles about their centroids are $d\boldsymbol{\theta}^p$ and $d\boldsymbol{\theta}^q$. Vector $\mathbf{n}$ denotes the unit normal of the common tangent plane at ''c,'' directed outward from particle ''p.'' The triplet of unit vectors $(\mathbf{n},\mathbf{t},\mathbf{w})$ is assigned to the contact and form a right-hand system, with

$$\mathbf{n}\times\mathbf{t}=\mathbf{w}; \quad \mathbf{t}\times\mathbf{w}=\mathbf{n}; \quad \mathbf{w}\times\mathbf{n}=\mathbf{t}. \tag{1}$$

In the case of two-dimensional analysis, the tangent vector $\mathbf{t}$ will lie in the plane of the particles and is directed counterclockwise from $\mathbf{n}$ around ''c,'' and $\mathbf{w}$ is perpendicular to the same plane. For three-dimensional analysis, $\mathbf{t}$ and $\mathbf{w}$ are arbitrarily specified unit vectors satisfying (1).

In the derivations and proofs the following identities will be used:

Identity 1 $\quad \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})=\mathbf{b}\cdot(\mathbf{c}\times\mathbf{a})=\mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$

Identity 2 $\quad (\mathbf{a}\times\mathbf{b})\times\mathbf{c}=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-\mathbf{a}(\mathbf{b}\cdot\mathbf{c})$

Identity 3 $\quad (\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d})=(\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d})-(\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c})$

Identity 4 $\quad (\mathbf{w}\times\mathbf{b})(\mathbf{c}\cdot\mathbf{n})-(\mathbf{w}\times\mathbf{c})(\mathbf{b}\cdot\mathbf{n})$

$$=\mathbf{n}((\mathbf{b}\cdot\mathbf{n})(\mathbf{c}\cdot\mathbf{t})-(\mathbf{c}\cdot\mathbf{n})(\mathbf{b}\cdot\mathbf{t})),$$

assuming that the triplet of unit vectors $(\mathbf{n},\mathbf{t},\mathbf{w})$ satisfies (1).

2 Two Circular Particles

2.1 The Definition of Rolling Displacement. In this section we analyze the rolling displacements of two circular particles, ''p'' and ''q,'' forming a contact ''c'' (Fig. 2). The infinitesimally small neighborhood of ''p'' around ''c'' will be denoted as ''pc'' (a material point attached to ''p'' and moving together with it). Similarly, the material point ''qc'' denotes a small neighborhood on particle ''q'' around ''c.'' Vectors $\mathbf{r}^p$ and $\mathbf{r}^q$ point from the centroids of ''p'' and ''q'' to the contact, and the unit vectors $(\mathbf{n},\mathbf{t},\mathbf{w})$ are the same as those defined in (1): $\mathbf{n}$ and $\mathbf{t}$ are shown, and $\mathbf{w}$ is perpendicular to the plane of analysis. The translations of the particle centroids, $d\mathbf{u}^p$ and $d\mathbf{u}^q$, take place in the $(\mathbf{n},\mathbf{t})$ plane, and the particle rotation vectors, $d\boldsymbol{\theta}^p$ and $d\boldsymbol{\theta}^q$, are perpendicular to that plane and are parallel to $\mathbf{w}$. The displacement of any point on either particle can be computed from the translation of the particle centroid and its rotation. In this section we consider the $\mathbf{t}$-directional translations only.

The $\mathbf{t}$-directional components of the particle translations will be denoted by $du^{p,t}$ and $du^{q,t}$, with direction $\mathbf{t}$ and magnitudes

$$du^{p,t}=d\mathbf{u}^p\cdot\mathbf{t} \quad \text{and} \quad du^{q,t}=d\mathbf{u}^q\cdot\mathbf{t}. \tag{2}$$

The magnitudes of particle rotations will be denoted by $d\theta^p$ and $d\theta^q$, and since in two dimensions the particle rotation vectors are parallel with the unit vector $\mathbf{w}$,

$$d\boldsymbol{\theta}^p=d\theta^p\cdot\mathbf{w} \quad \text{and} \quad d\boldsymbol{\theta}^q=d\theta^q\cdot\mathbf{w}. \tag{3}$$

Though before the incremental motions of the two particles the two material points ''pc'' and ''qc'' have the same position, this might not be the case after the particle motions. Using the notations defined in (2) and (3), their $\mathbf{t}$-directional translations can be expressed as

$$d\mathbf{u}^{pc}\cdot\mathbf{t}=d\mathbf{u}^p\cdot\mathbf{t}+(d\boldsymbol{\theta}^p\times\mathbf{r}^p)\cdot\mathbf{t}=du^{p,t}+d\theta^p(\mathbf{r}^p\cdot\mathbf{n}) \tag{4a}$$

$$d\mathbf{u}^{qc}\cdot\mathbf{t}=d\mathbf{u}^q\cdot\mathbf{t}+(d\boldsymbol{\theta}^q\times\mathbf{r}^q)\cdot\mathbf{t}=du^{q,t}+d\theta^q(\mathbf{r}^q\cdot\mathbf{n}). \tag{4b}$$

The four scalars $du^{p,t}$, $du^{q,t}$, $d\theta^p$, $d\theta^q$ uniquely determine the $\mathbf{t}$-directional translations of any point. But the following four alternative quantities will be more useful in characterizing the tangential contact motions:

(1) the $\mathbf{t}$-directional rigid-body translation of the contact point, defined as

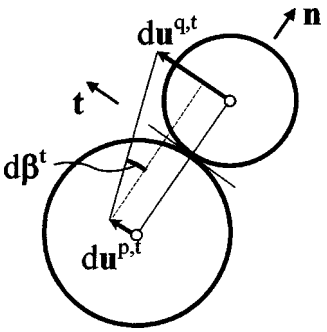


Fig. 3

$$du^{\text{c,t-rigid}} := \frac{\mathbf{r}^{\text{p}}\cdot\mathbf{n}}{(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{n}}\,du^{\text{q,t}} - \frac{\mathbf{r}^{\text{q}}\cdot\mathbf{n}}{(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{n}}\,du^{\text{p,t}}. \tag{5}$$

(2) the rigid-body **t**-rotation of the particle pair:

$$d\beta^{\text{t}} := \frac{du^{\text{q,t}} - du^{\text{p,t}}}{(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{n}} \quad \text{and} \quad d\boldsymbol{\beta}^{\text{t}} := d\beta^{\text{t}}\cdot\mathbf{w} \tag{6}$$

(3) the ''excess'' rotation of particle ''p'':

$$d\theta^{\text{p,excess}} := d\theta^{\text{p}} - d\beta^{\text{t}} \quad \text{and} \quad d\boldsymbol{\theta}^{\text{p,excess}} := d\boldsymbol{\theta}^{\text{p}} - d\boldsymbol{\beta}^{\text{t}} \tag{7a}$$

(4) the ''excess'' rotation of particle ''q'':

$$d\theta^{\text{q,excess}} := d\theta^{\text{q}} - d\beta^{\text{t}} \quad \text{and} \quad d\boldsymbol{\theta}^{\text{q,excess}} := d\boldsymbol{\theta}^{\text{q}} - d\boldsymbol{\beta}^{\text{t}}. \tag{7b}$$

The $du^{\text{p,t}}$, $du^{\text{q,t}}$, $d\theta^{\text{p}}$, $d\theta^{\text{q}}$ displacements can be uniquely calculated from these four displacements in the following way:

$$du^{\text{p,t}} = du^{\text{c,t-rigid}} - d\beta^{\text{t}}(\mathbf{r}^{\text{p}}\cdot\mathbf{n}) \equiv du^{\text{c,t-rigid}} - (d\boldsymbol{\beta}^{\text{t}}\times\mathbf{r}^{\text{p}})\cdot\mathbf{t} \tag{8a}$$

$$du^{\text{q,t}} = du^{\text{c,t-rigid}} - d\beta^{\text{t}}(\mathbf{r}^{\text{q}}\cdot\mathbf{n}) \equiv du^{\text{c,t-rigid}} - (d\boldsymbol{\beta}^{\text{t}}\times\mathbf{r}^{\text{q}})\cdot\mathbf{t} \tag{8b}$$

$$d\theta^{\text{p}} = d\beta^{\text{t}} + d\theta^{\text{p,excess}} \tag{9a}$$

$$d\theta^{\text{q}} = d\beta^{\text{t}} + d\theta^{\text{q,excess}}, \tag{9b}$$

so that we can change the variables in (2*a*) and (2*b*), to give an alternative expression for the **t**-directional displacements of the contact point:

$$d\mathbf{u}^{\text{pc}}\cdot\mathbf{t} = du^{\text{c,t-rigid}} + d\theta^{\text{p,excess}}(\mathbf{r}^{\text{p}}\cdot\mathbf{n}) \equiv du^{\text{c,t-rigid}} + (d\boldsymbol{\theta}^{\text{p,excess}}\times\mathbf{r}^{\text{p}})\cdot\mathbf{t} \tag{10a}$$

$$d\mathbf{u}^{\text{qc}}\cdot\mathbf{t} = du^{\text{c,t-rigid}} + d\theta^{\text{q,excess}}(\mathbf{r}^{\text{q}}\cdot\mathbf{n}) \equiv du^{\text{c,t-rigid}} + (d\boldsymbol{\theta}^{\text{q,excess}}\times\mathbf{r}^{\text{q}})\cdot\mathbf{t}. \tag{10b}$$

To illustrate the meaning of the displacements $du^{\text{c,t-rigid}}$, $d\beta^{\text{t}}$, $d\theta^{\text{p,excess}}$ and $d\theta^{\text{q,excess}}$, two elementary cases are now discussed (Figs. 3 and 4).

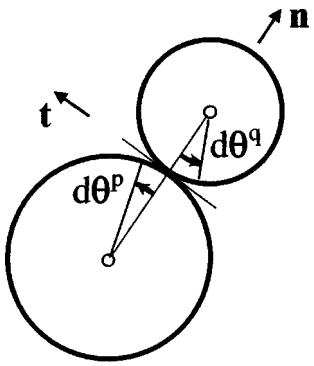


Fig. 4

Elementary case 1: rigid-body displacements in regard to **t**-directional translations

If the excess rotations $d\theta^{\text{p,excess}}$ and $d\theta^{\text{q,excess}}$ are both zero, the two particles translate and rotate together as a single rigid body, at least in regard to their **t**-directional translations (Fig. 3). In this special case the **t**-directional translations of the contact point on the two particles are equal:

$$d\mathbf{u}^{\text{pc}}\cdot\mathbf{t} = d\mathbf{u}^{\text{qc}}\cdot\mathbf{t} = du^{\text{c,t-rigid}}. \tag{11}$$

Elementary case 2: pure particle rotations

If the translation $du^{\text{c,t-rigid}}$ and the rotation $d\beta^{\text{t}}$ are both zero, the particles rotate about their centroids without translating in the **t**-direction (Fig. 4)

$$d\theta^{\text{p}} = d\theta^{\text{p,excess}} \tag{12a}$$

$$d\theta^{\text{q}} = d\theta^{\text{q,excess}}, \tag{12b}$$

and the **t**-directional translations of the contact point are

$$d\mathbf{u}^{\text{pc}}\cdot\mathbf{t} = d\theta^{\text{p,excess}}(\mathbf{r}^{\text{p}}\cdot\mathbf{n}) \equiv (d\boldsymbol{\theta}^{\text{p,excess}}\times\mathbf{r}^{\text{p}})\cdot\mathbf{t} \tag{13a}$$

$$d\mathbf{u}^{\text{qc}}\cdot\mathbf{t} = d\theta^{\text{q,excess}}(\mathbf{r}^{\text{q}}\cdot\mathbf{n}) \equiv (d\boldsymbol{\theta}^{\text{q,excess}}\times\mathbf{r}^{\text{q}})\cdot\mathbf{t}. \tag{13b}$$

If these two translations, $d\mathbf{u}^{\text{pc}}\cdot\mathbf{t}$ and $d\mathbf{u}^{\text{qc}}\cdot\mathbf{t}$, are equal and nonzero, the two particles move together in a gear-like fashion at contact ''c,'' a situation of pure rolling. However, if the two translations are not equal, we can calculate their difference, $du^{\text{c,t-def}}$:

$$du^{\text{c,t-def}} = d\mathbf{u}^{\text{qc}}\cdot\mathbf{t} - d\mathbf{u}^{\text{pc}}\cdot\mathbf{t} \tag{14}$$

which is the tangential deformation of the contact; and their average, $du^{\text{c,t-avr}}$:

$$du^{\text{c,t-avr}} = \frac{1}{2}(d\mathbf{u}^{\text{qc}}\cdot\mathbf{t} + d\mathbf{u}^{\text{pc}}\cdot\mathbf{t}). \tag{15}$$

This average is the ''common'' motion of the contact point on ''p'' and ''q,'' hence this motion can be considered as a rolling displacement:

$$du^{\text{c,t-roll}} = \frac{1}{2}((d\boldsymbol{\theta}^{\text{q,excess}}\times\mathbf{r}^{\text{q}})\cdot\mathbf{t} + (d\boldsymbol{\theta}^{\text{p,excess}}\times\mathbf{r}^{\text{p}})\cdot\mathbf{t}). \tag{16}$$

The general case:

A general system of particle displacements $du^{\text{p,t}}$, $du^{\text{q,t}}$, $d\theta^{\text{p}}$, $d\theta^{\text{q}}$ can be uniquely expressed in terms of $du^{\text{c,t-rigid}}$, $d\beta^{\text{t}}$, $d\theta^{\text{p,excess}}$, and $d\theta^{\text{q,excess}}$. Any general system of particle displacements can, therefore, be considered as the superposition of a rigid-body displacement in regard to the **t**-directional translations plus the ''excess'' rotations of the particles. To define the **t**-directional rolling displacement for the general case of two circles, we release the restrictions of elementary case 2, and give the following definition for the general case:

$$du^{\text{c,t-roll}} := \frac{1}{2}(d\boldsymbol{\theta}^{\text{p,excess}}\times\mathbf{r}^{\text{p}} + d\boldsymbol{\theta}^{\text{q,excess}}\times\mathbf{r}^{\text{q}})\cdot\mathbf{t} \tag{17}$$

which can also be expressed using (8) as

$$du^{\text{c,t-roll}} = du^{\text{c,t-avr}} - du^{\text{c,t-rigid}} \tag{18}$$

or in terms of the particle displacements, applying (4) and (5):

$$du^{\text{c,t-roll}} = \frac{1}{2}\left[d\boldsymbol{\theta}^{\text{p}}\times\mathbf{r}^{\text{p}} + d\boldsymbol{\theta}^{\text{q}}\times\mathbf{r}^{\text{q}} - \frac{(\mathbf{r}^{\text{p}}+\mathbf{r}^{\text{q}})\cdot\mathbf{n}}{(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{n}}(d\mathbf{u}^{\text{q}} - d\mathbf{u}^{\text{p}})\right]\cdot\mathbf{t}. \tag{19}$$

We now show that our suggested expression of rolling is equivalent to the Iwashita-Oda [1] rolling. Denote the particle radii with the positive scalars R^{p} and R^{q}; evidently, $R^{\text{p}} = \mathbf{r}^{\text{p}}\cdot\mathbf{n}$ and $R^{\text{q}} = -\mathbf{r}^{\text{q}}\cdot\mathbf{n}$. Now (19) can be written as

$$du^{\text{c,t-roll}}=\frac{1}{2}\left[(d\theta^{\text{p}}\mathbf{w})\times(R^{\text{p}}\mathbf{n})-(d\theta^{\text{q}}\mathbf{w})\times(R^{\text{q}}\mathbf{n})\right.$$
$$\left.-\frac{R^{\text{p}}-R^{\text{q}}}{R^{\text{p}}+R^{\text{q}}}(d\mathbf{u}^{\text{q}}-d\mathbf{u}^{\text{p}})\right]\cdot\mathbf{t} \tag{20}$$

and because of (1), this is equal to

$$du^{\text{c,t-roll}}=\frac{1}{2}\left[d\theta^{\text{p}}R^{\text{p}}-d\theta^{\text{q}}R^{\text{q}}-\frac{R^{\text{p}}-R^{\text{q}}}{R^{\text{p}}+R^{\text{q}}}(d\mathbf{u}^{\text{q}}-d\mathbf{u}^{\text{p}})\cdot\mathbf{t}\right] \tag{21}$$

which is the same as the Iwashita-Oda rolling.

2.2 Objectivity. We now show that the **t**-directional rolling displacement is *objective*, meaning that

(a) If the observer '*O*'—located at the origin of the coordinate system—undergoes an incremental movement $d\mathbf{u}^O$ and $d\boldsymbol{\theta}^O$ (where $d\boldsymbol{\theta}^O=d\theta^O\mathbf{w}$), the observer will measure the same **t**-directional rolling displacement of the contact as if the observer was stationary, regardless of the displacements $d\mathbf{u}^O$ and $d\boldsymbol{\theta}^O$.

(b) If the coordinate system is relocated (with the observer resting in its origin) anywhere in the plane by a finite translation and rotation, the observer will measure the same **t**-directional rolling displacement of the contact.

Objectivity (a)

An arbitrary point '*A*', located at $\mathbf{x}^A$, whose displacements are $d\mathbf{u}^A$ and $d\boldsymbol{\theta}^A$ ($d\boldsymbol{\theta}^A=d\theta^A\mathbf{w}$), is seen by the moving observer as making the displacements

$$d\hat{\mathbf{u}}^A=d\mathbf{u}^A-d\mathbf{u}^O-d\boldsymbol{\theta}^O\times\mathbf{x}^A \tag{22}$$

$$d\hat{\boldsymbol{\theta}}^A=d\boldsymbol{\theta}^A-d\boldsymbol{\theta}^O. \tag{23}$$

The **t**-directional rolling displacement in (19), as seen by the moving observer, is

$$d\hat{u}^{\text{c,t-roll}}=\frac{1}{2}\left[d\hat{\boldsymbol{\theta}}^{\text{p}}\times\mathbf{r}^{\text{p}}+d\hat{\boldsymbol{\theta}}^{\text{q}}\times\mathbf{r}^{\text{q}}-\frac{(\mathbf{r}^{\text{p}}+\mathbf{r}^{\text{q}})\cdot\mathbf{n}}{(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{n}}(d\hat{\mathbf{u}}^{\text{q}}-d\hat{\mathbf{u}}^{\text{p}})\right]\cdot\mathbf{t}. \tag{24}$$

Substituting (22) and (23) into (24), we find that

$$d\hat{u}^{\text{c,t-roll}}=\frac{1}{2}\left[(d\boldsymbol{\theta}^{\text{p}}-d\boldsymbol{\theta}^O)\times\mathbf{r}^{\text{p}}+(d\boldsymbol{\theta}^{\text{q}}-d\boldsymbol{\theta}^O)\times\mathbf{r}^{\text{q}}\right.$$
$$-\frac{(\mathbf{r}^{\text{p}}+\mathbf{r}^{\text{q}})\cdot\mathbf{n}}{(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{n}}((d\mathbf{u}^{\text{q}}-d\mathbf{u}^O-d\boldsymbol{\theta}^O\times\mathbf{x}^{\text{q}})$$
$$\left.-(d\mathbf{u}^{\text{p}}-d\mathbf{u}^O-d\boldsymbol{\theta}^O\times\mathbf{x}^{\text{p}}))\right]\cdot\mathbf{t}. \tag{25}$$

Noting that $\mathbf{x}^{\text{q}}-\mathbf{x}^{\text{p}}=\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}}$, and substituting (19),

$$d\hat{u}^{\text{c,t-roll}}=du^{\text{c,t-roll}}-\frac{1}{2}\frac{d\theta^O}{(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{n}}[((\mathbf{w}\times(\mathbf{r}^{\text{p}}+\mathbf{r}^{\text{q}}))((\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{n}))$$
$$-((\mathbf{w}\times(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}}))((\mathbf{r}^{\text{p}}+\mathbf{r}^{\text{q}})\cdot\mathbf{n}))]\cdot\mathbf{t}. \tag{26}$$

The expression in brackets is a vector parallel to **n** (see Identity 4), and hence it is perpendicular to **t**. Consequently,

$$d\hat{u}^{\text{c,t-roll}}=du^{\text{c,t-roll}}$$

and the objectivity (a) is shown.

Objectivity (b)

According to Identity 1, the definition (19) can be written as

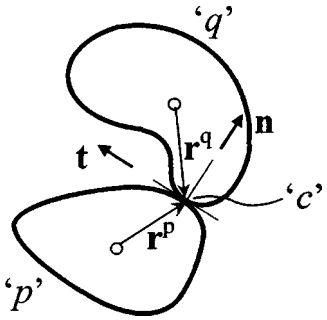


Fig. 5

$$du^{\text{c,t-roll}}=\frac{1}{2}\left[d\theta^{\text{p}}(\mathbf{r}^{\text{p}}\cdot\mathbf{n})+d\theta^{\text{q}}(\mathbf{r}^{\text{q}}\cdot\mathbf{n})\right.$$
$$\left.-\frac{(\mathbf{r}^{\text{p}}+\mathbf{r}^{\text{q}})\cdot\mathbf{n}}{(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{n}}((d\mathbf{u}^{\text{q}}-d\mathbf{u}^{\text{p}})\cdot\mathbf{t})\right]. \tag{27}$$

On the right side of Eq. (27) we find scalar products of vectors whose lengths and whose angle with each other do not change during the rigid-body-like translations and rotations of the coordinate frame. Consequently, the whole right side remains unchanged. Objectivity (b) is shown.

3 Two Arbitrary Two-Dimensional Particles

In this section we extend the definition of rolling given in Section 2 to arbitrarily shaped (but smooth) two-dimensional particles (see the illustration in Fig. 5).

The **t**-directional rolling displacement in Eqs. (17), (18), and (19) apply to any pair of two-dimensional particles, provided that

$$(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{n}\neq 0. \tag{28}$$

A counter-example is shown in Fig. 6, where the $(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})$ vector is perpendicular to **n**, so that the **t**-directional rolling displacement does not exist (see also Section 5).

As with circular particles, the scalar value $du^{\text{c,t-roll}}$ is objective for noncircular particles, since the derivations in Section 2.2 did not impose any relationship among the two-dimensional vectors $\mathbf{r}^{\text{p}}$, $\mathbf{r}^{\text{q}}$, and **n**.

In the case of generally shaped particles we have no reason to restrict ourselves to the analysis of **t**-directional translations. Assuming that

$$(\mathbf{r}^{\text{p}}-\mathbf{r}^{\text{q}})\cdot\mathbf{t}\neq 0, \tag{29}$$

we can define an **n**-directional rolling displacement, $du^{\text{c,n-roll}}$, in an analogous way:

$$du^{\text{c,n-roll}}=\tfrac{1}{2}(d\mathbf{u}^{\text{pc}}+d\mathbf{u}^{\text{qc}})\cdot\mathbf{n}-du^{\text{c,n-rigid}}, \tag{30}$$

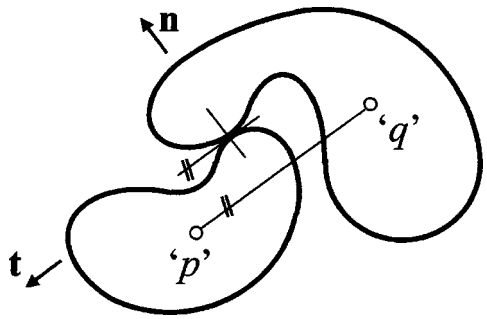


Fig. 6

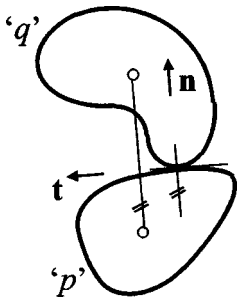


Fig. 7

where

$$du^{c,n\text{-rigid}} := \frac{\mathbf{r}^p \cdot \mathbf{t}}{(\mathbf{r}^p - \mathbf{r}^q)\cdot \mathbf{t}} (d\mathbf{u}^q \cdot \mathbf{n}) - \frac{\mathbf{r}^q \cdot \mathbf{t}}{(\mathbf{r}^p - \mathbf{r}^q)\cdot \mathbf{t}} (d\mathbf{u}^p \cdot \mathbf{n}). \quad (31)$$

This rolling motion can also be expressed in terms of particle displacements:

$$du^{c,n\text{-roll}} = \frac{1}{2}\left[d\boldsymbol{\theta}^p \times \mathbf{r}^p + d\boldsymbol{\theta}^q \times \mathbf{r}^q - \frac{(\mathbf{r}^p + \mathbf{r}^q)\cdot \mathbf{t}}{(\mathbf{r}^p - \mathbf{r}^q)\cdot \mathbf{t}} (d\mathbf{u}^q - d\mathbf{u}^p) \right] \cdot \mathbf{n}. \quad (32)$$

If the branch vector $(\mathbf{r}^p - \mathbf{r}^q)$ is perpendicular to $\mathbf{t}$ (for instance, if the two particles are circles), the $\mathbf{n}$-directional rolling displacement does not exist. Another example of this situation is shown in Fig. 7.

The objectivity of $du^{c,n\text{-roll}}$ can similarly be checked in the manner of Section 2.2, so we shall not repeat the arguments here.

4 Two Arbitrary Three-Dimensional Particles

4.1 Definition of t-Directional Rolling Displacement in Three Dimensions. In this section we focus on the displacements of two arbitrary three-dimensional particles, "p" and "q," forming contact "c." The triplet of unit vectors $(\mathbf{n},\mathbf{t},\mathbf{w})$ is assigned to the pair, as was explained in Sections 1 and 2.1 (Fig. 8). We shall first analyze the $\mathbf{t}$-directional translations of the contact point on "p" and "q," leaving the definition of $\mathbf{n}$ and $\mathbf{w}$-directional rolling to Section 4.3.

In the two-dimensional case (Section 2.1, and Eq. (20)), we separated the average $\mathbf{t}$-directional contact displacement $du^{c,t\text{-avr}}$ into a part $du^{c,t\text{-roll}}$ corresponding to rolling, and another part $du^{c,t\text{-rigid}}$ corresponding to a rigid-body-like motion. A similar line of thought will be followed here. Since the displacements in three dimensions are more complex, to define the $\mathbf{t}$-directional rolling displacement, we first introduce a triplet of unit vectors that will help to separate those displacement components that produce rolling and those that do not. This triplet will be denoted as $(\boldsymbol{\lambda}^t,\mathbf{t},\mathbf{z}^t)$ and defined with the help of the branch vector, $\mathbf{l} = \mathbf{r}^p - \mathbf{r}^q$ (Fig. 9).

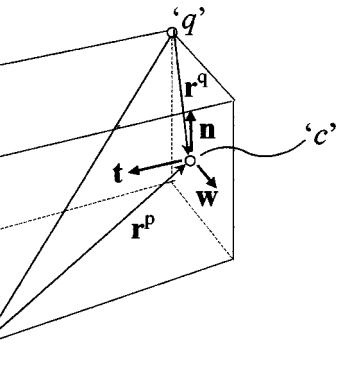


Fig. 8

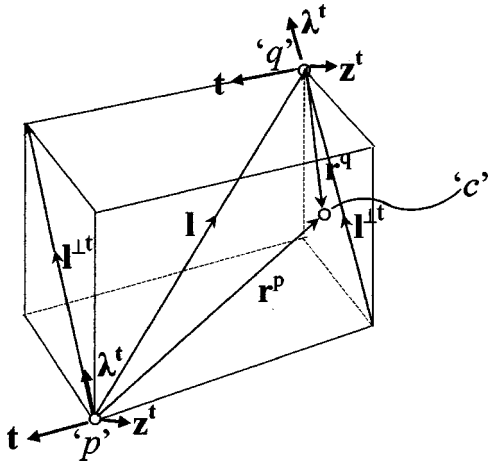


Fig. 9

Now we separate the branch vector into two components: one parallel to $\mathbf{t}$:

$$\mathbf{l}^t = ((\mathbf{r}^p - \mathbf{r}^q)\cdot \mathbf{t})\mathbf{t} \quad (33)$$

and the other one perpendicular to $\mathbf{t}$ (see Fig. 9):

$$\mathbf{l}^{\perp t} = \mathbf{l} - \mathbf{l}^t \equiv (\mathbf{r}^p - \mathbf{r}^q) - ((\mathbf{r}^p - \mathbf{r}^q)\cdot \mathbf{t})\mathbf{t}, \quad (34)$$

whose magnitude is

$$l^{\perp t} = \sqrt{\mathbf{l}^{\perp t}\cdot \mathbf{l}^{\perp t}}. \quad (35)$$

Provided that $l^{\perp t} \neq 0$, the vector $\boldsymbol{\lambda}^t$ is defined as the unit vector pointing in the direction of $\mathbf{l}^{\perp t}$:

$$\boldsymbol{\lambda}^t := \frac{\mathbf{l}^{\perp t}}{l^{\perp t}} \quad (36)$$

(obviously, $l^{\perp t} = \mathbf{l}\cdot\boldsymbol{\lambda}^t$), and the vector $\mathbf{z}^t$ is

$$\mathbf{z}^t := \boldsymbol{\lambda}^t \times \mathbf{t}. \quad (37)$$

The vectors $(\boldsymbol{\lambda}^t,\mathbf{t},\mathbf{z}^t)$ form a right-handed system with

$$\mathbf{t} := \mathbf{z}^t \times \boldsymbol{\lambda}^t \quad \text{and} \quad \boldsymbol{\lambda}^t := \mathbf{t} \times \mathbf{z}^t, \quad (38)$$

as illustrated in Fig. 9.

We can decompose the particle translations into their $\mathbf{t}$, $\mathbf{z}^t$, and $\boldsymbol{\lambda}^t$-directional components. Since the $\mathbf{z}^t$ and $\boldsymbol{\lambda}^t$-directional translations do not lead to $\mathbf{t}$-directional translations of the contact point, they will be excluded from the analysis. We need to consider the $\mathbf{t}$-directional translations only:

$$du^{p,t} = d\mathbf{u}^p \cdot \mathbf{t} \quad \text{and} \quad du^{q,t} = d\mathbf{u}^q \cdot \mathbf{t}. \quad (39)$$

The particle rotation vectors $d\boldsymbol{\theta}^p$ and $d\boldsymbol{\theta}^q$ can also be decomposed into their $\mathbf{t}$, $\mathbf{z}^t$, and $\boldsymbol{\lambda}^t$-directional components:

$$d\boldsymbol{\theta}^p = d\theta^{p,t}\mathbf{t} + d\theta^{p,z^t}\mathbf{z}^t + d\theta^{p,\lambda^t}\boldsymbol{\lambda}^t \quad \text{and}$$

$$d\boldsymbol{\theta}^q = d\theta^{q,t}\mathbf{t} + d\theta^{q,z^t}\mathbf{z}^t + d\theta^{q,\lambda^t}\boldsymbol{\lambda}^t, \quad (40)$$

where

$$d\theta^{p,t} = d\boldsymbol{\theta}^p \cdot \mathbf{t} \quad d\theta^{q,t} = d\boldsymbol{\theta}^q \cdot \mathbf{t}$$

$$d\theta^{p,z^t} = d\boldsymbol{\theta}^p \cdot \mathbf{z}^t \quad \text{and} \quad d\theta^{q,z^t} = d\boldsymbol{\theta}^q \cdot \mathbf{z}^t \quad (41)$$

$$d\theta^{p,\lambda^t} = d\boldsymbol{\theta}^p \cdot \boldsymbol{\lambda}^t \quad d\theta^{q,\lambda^t} = d\boldsymbol{\theta}^q \cdot \boldsymbol{\lambda}^t.$$

We now analyze their contributions to the average $\mathbf{t}$-directional displacement of the contact point.

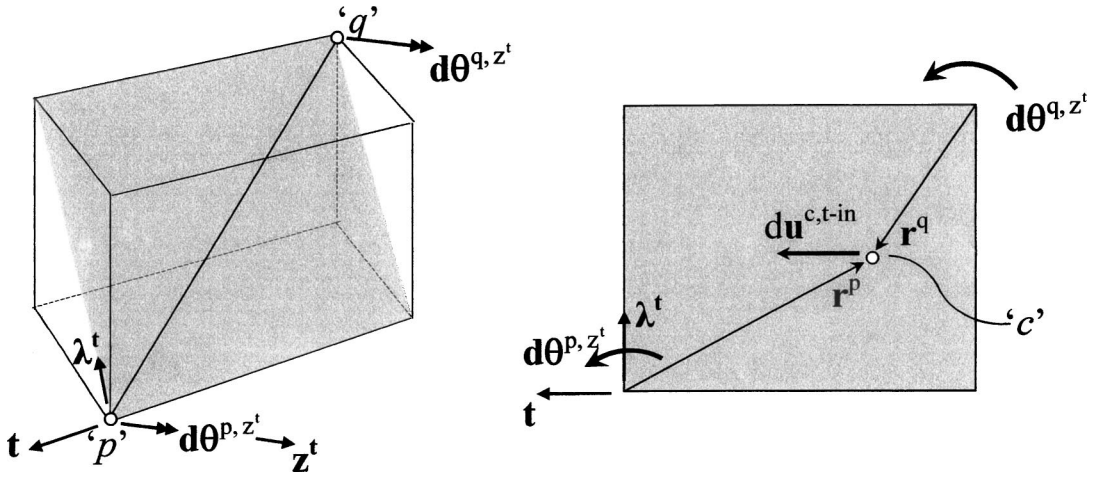


Fig. 10

1 Particle rotations about **t**:

The rotation vectors $(\mathrm{d}\theta^{\mathrm{p,t}}\mathbf{t})$ and $(\mathrm{d}\theta^{\mathrm{q,t}}\mathbf{t})$ do not produce any **t**-directional translation at "c," so we can exclude them from the analysis.

2 Particle rotations about $\mathbf{z}^{\mathrm{t}}$:

The rotation vectors $(\mathrm{d}\theta^{\mathrm{p,z^t}}\mathbf{z}^{\mathrm{t}})$ and $(\mathrm{d}\theta^{\mathrm{q,z^t}}\mathbf{z}^{\mathrm{t}})$ produce translations whose direction is parallel to the $\mathbf{t}$-$\boldsymbol{\lambda}^{\mathrm{t}}$–plane (shaded in Fig. 10), and the corresponding **t**-directional translations of the material points are

$$\mathrm{d}u^{\mathrm{pc,t\text{-}in}} := ((\mathrm{d}\theta^{\mathrm{p,z^t}}\mathbf{z}^{\mathrm{t}}) \times \mathbf{r}^{\mathrm{p}}) \cdot \mathbf{t} = \mathrm{d}\theta^{\mathrm{p,z^t}}(\mathbf{r}^{\mathrm{p}} \cdot \boldsymbol{\lambda}^{\mathrm{t}}) \tag{42a}$$

and

$$\mathrm{d}u^{\mathrm{qc,t\text{-}in}} = ((\mathrm{d}\theta^{\mathrm{q,z^t}}\mathbf{z}^{\mathrm{t}}) \times \mathbf{r}^{\mathrm{q}}) \cdot \mathbf{t} = \mathrm{d}\theta^{\mathrm{q,z^t}}(\mathbf{r}^{\mathrm{q}} \cdot \boldsymbol{\lambda}^{\mathrm{t}}). \tag{42b}$$

Their average, caused by these "in-plane rotations," will be denoted as $\mathrm{d}u^{\mathrm{c,t\text{-}in}}$ and calculated as

$$\mathrm{d}u^{\mathrm{t,in}} = \tfrac{1}{2}(\mathrm{d}u^{\mathrm{pc,t\text{-}in}} + \mathrm{d}u^{\mathrm{qc,t\text{-}in}}) = \tfrac{1}{2}(\mathrm{d}\theta^{\mathrm{p,z^t}}(\mathbf{r}^{\mathrm{p}} \cdot \boldsymbol{\lambda}^{\mathrm{t}}) + \mathrm{d}\theta^{\mathrm{q,z^t}}(\mathbf{r}^{\mathrm{q}} \cdot \boldsymbol{\lambda}^{\mathrm{t}})). \tag{43}$$

Note that if the other rotation components and the particle translations are both zero, this situation is analogous to elementary case 2. in Section 2, and $\mathrm{d}u^{\mathrm{c,t\text{-}in}}$ is the same as the **t**-directional rolling displacement of that case.

3 Particle rotations about $\boldsymbol{\lambda}^{\mathrm{t}}$:

The rotation vectors $(\mathrm{d}\theta^{\mathrm{p},\boldsymbol{\lambda}^{\mathrm{t}}}\boldsymbol{\lambda}^{\mathrm{t}})$ and $(\mathrm{d}\theta^{\mathrm{q},\boldsymbol{\lambda}^{\mathrm{t}}}\boldsymbol{\lambda}^{\mathrm{t}})$ are parallel to the shaded $\mathbf{t}$-$\boldsymbol{\lambda}^{\mathrm{t}}$–plane in Fig. 11 and rotate the particles about the direction $\boldsymbol{\lambda}^{\mathrm{t}}$. If the contact point is not located in the shaded plane in Fig. 11(*a*), these rotations cause **t**-directional translations of the material points:

$$\mathrm{d}u^{\mathrm{pc,t\text{-}out}} = ((\mathrm{d}\theta^{\mathrm{p},\boldsymbol{\lambda}^{\mathrm{t}}}\boldsymbol{\lambda}^{\mathrm{t}}) \times \mathbf{r}^{\mathrm{p}}) \cdot \mathbf{t} = -\mathrm{d}\theta^{\mathrm{p},\boldsymbol{\lambda}^{\mathrm{t}}}(\mathbf{r}^{\mathrm{p}} \cdot \mathbf{z}^{\mathrm{t}}) \tag{44a}$$

and

$$\mathrm{d}u^{\mathrm{qc,t\text{-}out}} = ((\mathrm{d}\theta^{\mathrm{q},\boldsymbol{\lambda}^{\mathrm{t}}}\boldsymbol{\lambda}^{\mathrm{t}}) \times \mathbf{r}^{\mathrm{q}}) \cdot \mathbf{t} = -\mathrm{d}\theta^{\mathrm{q},\boldsymbol{\lambda}^{\mathrm{t}}}(\mathbf{r}^{\mathrm{q}} \cdot \mathbf{z}^{\mathrm{t}}). \tag{44b}$$

The average of these two translations will be denoted as $\mathrm{d}u^{\mathrm{c,t\text{-}out}}$:

$$\mathrm{d}u^{\mathrm{c,t\text{-}out}} = \tfrac{1}{2}(\mathrm{d}u^{\mathrm{pc,t\text{-}out}} + \mathrm{d}u^{\mathrm{qc,t\text{-}out}}) = \tfrac{1}{2}(-\mathrm{d}\theta^{\mathrm{p},\boldsymbol{\lambda}^{\mathrm{t}}}(\mathbf{r}^{\mathrm{p}} \cdot \mathbf{z}^{\mathrm{t}}) - \mathrm{d}\theta^{\mathrm{q},\boldsymbol{\lambda}^{\mathrm{t}}}(\mathbf{r}^{\mathrm{q}} \cdot \mathbf{z}^{\mathrm{t}})). \tag{45}$$

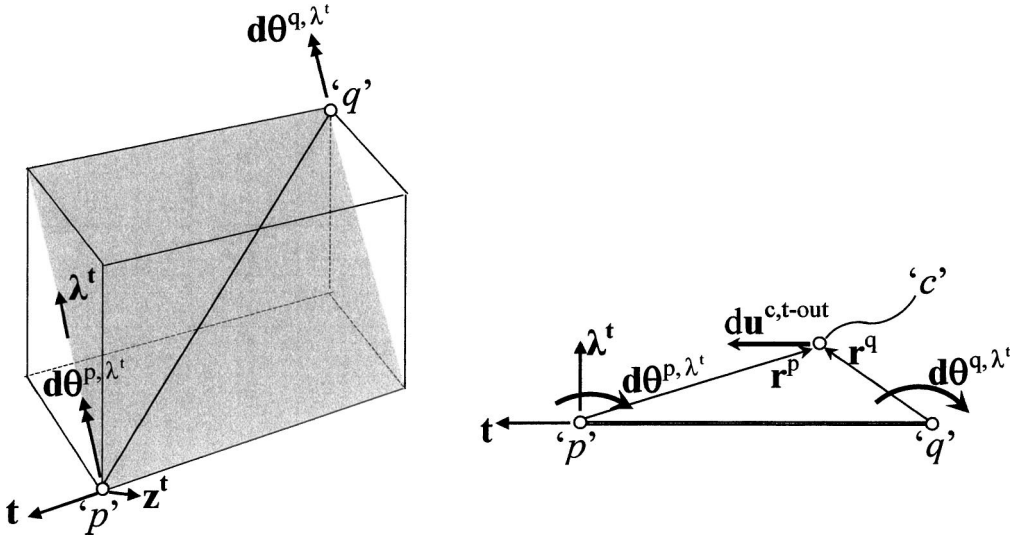


Fig. 11

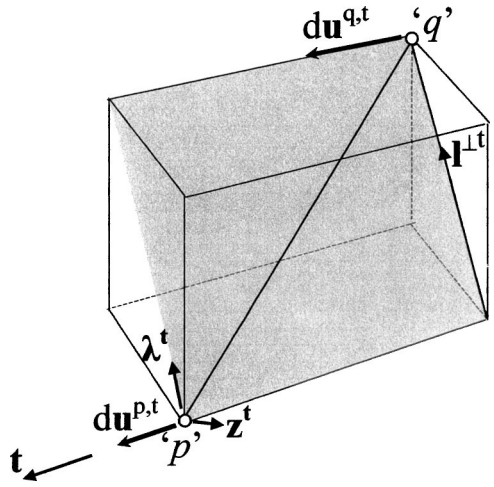


Fig. 12

This component does not exist in a two-dimensional analysis. The displacement $du^{\text{c,t-out}}$ is not a rolling motion, but as is seen in Fig. 11(*b*), is the effect of the common rotation of the two particles, and it is made possible only by the out-of-the plane location of the contact point.

Following the approach in Section 2, we identify the part of the in-plane displacements that corresponds to a rigid-body-like motion of the pair. This rotation, denoted as $d\beta^t\mathbf{z}^t$ or $d\boldsymbol{\beta}^t$, can be calculated from the **t**-directional translations of the two particles $du^{p,t}$ and $du^{q,t}$, where $d\mathbf{u}^{p,t}=du^{p,t}\mathbf{t}$ and $d\mathbf{u}^{q,t}=du^{q,t}\mathbf{t}$ (Fig. 12):

$$d\beta^t := \frac{du^{q,t}-du^{p,t}}{l^{\perp t}}. \tag{46}$$

The **t**-directional translation caused by this rotation in the contact point is

$$du^{\text{c,t-rigid}} := (d\mathbf{u}^p + d\boldsymbol{\beta}^t\times\mathbf{r}^p)\cdot\mathbf{t} \equiv (d\mathbf{u}^q + d\boldsymbol{\beta}^t\times\mathbf{r}^q)\cdot\mathbf{t}. \tag{47}$$

The displacement $du^{\text{c,t-rigid}}$ can also be expressed in terms of the particle translations:

$$du^{\text{c,t-rigid}} := \frac{1}{l^{\perp t}}((\mathbf{r}^p\cdot\boldsymbol{\lambda}^t)du^{q,t}-(\mathbf{r}^q\cdot\boldsymbol{\lambda}^t)du^{p,t}). \tag{48}$$

We now give the definition of **t**-directional rolling displacement by considering the excess in-plane rotations of the particles:

$$du^{\text{c,t-roll}} := \tfrac{1}{2}((d\theta^{p,z^t}-d\beta^t)(\mathbf{r}^p\cdot\boldsymbol{\lambda}^t)+(d\theta^{q,z^t}-d\beta^t)(\mathbf{r}^q\cdot\boldsymbol{\lambda}^t)) \tag{49}$$

which equals

$$du^{\text{c,t-roll}} = du^{\text{c,t-in}} - \frac{1}{2}d\beta^t((\mathbf{r}^p+\mathbf{r}^q)\cdot\boldsymbol{\lambda}^t). \tag{50}$$

With the use of Identities 1 and 2 and Eq. (15), the **t**-directional rolling displacement can also be written as

$$du^{\text{c,t-roll}} = du^{\text{c,t-avr}} - du^{\text{c,t-out}} - du^{\text{c,t-rigid}}, \tag{51}$$

an expression similar to that in Eq. (18) for the two-dimensional situation, but which contains the new term $du^{\text{c,t-out}}$ that does not exist in two dimensions. In terms of the particles' displacements, the **t**-directional rolling displacement is

$$du^{\text{c,t-roll}} = \frac{1}{2}\left[(d\boldsymbol{\theta}^p\cdot\mathbf{z}^t)(\mathbf{r}^p\cdot\boldsymbol{\lambda}^t)+(d\boldsymbol{\theta}^q\cdot\mathbf{z}^t)(\mathbf{r}^q\cdot\boldsymbol{\lambda}^t) - \frac{((\mathbf{r}^p+\mathbf{r}^q)\cdot\boldsymbol{\lambda}^t)}{((\mathbf{r}^p-\mathbf{r}^q)\cdot\boldsymbol{\lambda}^t)}((d\mathbf{u}^q-d\mathbf{u}^p)\cdot\mathbf{t})\right], \tag{52}$$

provided that $l^{\perp t}\neq 0$. If $l^{\perp t}$ is zero, then the **t**-directional rolling displacement does not exist (a situation analogous to the one illustrated in Fig. 6 and discussed in Section 5).

4.2 Definition of the n and w-Directional Rolling Displacement in Three Dimensions. The **n** and **w**-directional rolling displacements can be analogously defined. For **n**-directional rolling, we decompose the branch vector **l** into two components: one that is aligned in the **n**-direction, and one (denoted by $\mathbf{l}^{\perp n}$) that is perpendicular to **n**. We can then identify the unit vectors $\boldsymbol{\lambda}^n$ and $\mathbf{z}^n$ that are analogous to (36) and (37). The triplet $(\boldsymbol{\lambda}^n,\mathbf{n},\mathbf{z}^n)$ forms a right-handed system. The **n**-directional rolling displacement (assuming that $l^{\perp n}\neq 0$) is

$$du^{\text{c,n-roll}} = du^{\text{c,n-avr}} - du^{\text{c,n-out}} - du^{\text{c,n-rigid}}$$
$$= \frac{1}{2}\left[(d\boldsymbol{\theta}^p\cdot\mathbf{z}^n)(\mathbf{r}^p\cdot\boldsymbol{\lambda}^n)+(d\boldsymbol{\theta}^q\cdot\mathbf{z}^n)(\mathbf{r}^q\cdot\boldsymbol{\lambda}^n) - \frac{((\mathbf{r}^p+\mathbf{r}^q)\cdot\boldsymbol{\lambda}^n)}{((\mathbf{r}^p-\mathbf{r}^q)\cdot\boldsymbol{\lambda}^n)}((d\mathbf{u}^q-d\mathbf{u}^p)\cdot\mathbf{n})\right]. \tag{53}$$

We can also determine a vector $\mathbf{l}^{\perp w}$ and, if $l^{\perp w}\neq 0$, the associated unit vectors $\boldsymbol{\lambda}^w$ and $\mathbf{z}^w$. The **w**-directional rolling displacement is

$$du^{\text{c,w-roll}} = du^{\text{c,w-avr}} - du^{\text{c,w-out}} - du^{\text{c,w-rigid}}$$
$$= \frac{1}{2}\left[(d\boldsymbol{\theta}^p\cdot\mathbf{z}^w)(\mathbf{r}^p\cdot\boldsymbol{\lambda}^w)+(d\boldsymbol{\theta}^q\cdot\mathbf{z}^w)(\mathbf{r}^q\cdot\boldsymbol{\lambda}^w) - \frac{((\mathbf{r}^p+\mathbf{r}^q)\cdot\boldsymbol{\lambda}^w)}{((\mathbf{r}^p-\mathbf{r}^q)\cdot\boldsymbol{\lambda}^w)}((d\mathbf{u}^q-d\mathbf{u}^p)\cdot\mathbf{w})\right]. \tag{54}$$

4.3 Objectivity. To demonstrate the objectivity of $du^{\text{c,t-roll}}$, we must show that (a) an observer that is located at the origin of the coordinate system and undergoes arbitrary $d\mathbf{u}^O$ and $d\boldsymbol{\theta}^O$ incremental displacements will measure the same rolling displacement as if it was stationary; and that (b) regardless of where the coordinate frame is located, the observer at its origin will always measure the same rolling displacement. For simplicity, we shall only deal with the **t**-directional rolling displacement, since the other rolling displacements can analogously be checked.

Objectivity (a)

An observer with displacements $d\mathbf{u}^O$ and $d\boldsymbol{\theta}^O$ measures the following particle displacements:

$$d\hat{\mathbf{u}}^p = d\mathbf{u}^p - d\mathbf{u}^O - d\boldsymbol{\theta}^O\times\mathbf{x}^p; \quad d\hat{\boldsymbol{\theta}}^p = d\boldsymbol{\theta}^p - d\boldsymbol{\theta}^O \tag{55}$$

$$d\hat{\mathbf{u}}^q = d\mathbf{u}^q - d\mathbf{u}^O - d\boldsymbol{\theta}^O\times\mathbf{x}^q; \quad d\hat{\boldsymbol{\theta}}^q = d\boldsymbol{\theta}^q - d\boldsymbol{\theta}^O. \tag{56}$$

After substituting these expressions into (52), the **t**-directional rolling displacement measured by the observer is

$$d\hat{u}^{\text{c,t-roll}} = du^{\text{c,t-roll}} + \frac{1}{2}[-(d\boldsymbol{\theta}^O\cdot\mathbf{z}^t)(\mathbf{r}^p\cdot\boldsymbol{\lambda}^t)-(d\boldsymbol{\theta}^O\cdot\mathbf{z}^t)(\mathbf{r}^q\cdot\boldsymbol{\lambda}^t)]-$$
$$-\frac{1}{2}\frac{1}{l^{\perp t}}[((-d\boldsymbol{\theta}^O\times(\mathbf{x}^q-\mathbf{x}^p))\cdot\mathbf{t})((\mathbf{r}^p+\mathbf{r}^q)\cdot\boldsymbol{\lambda}^t)], \tag{57}$$

where $du^{\text{c,t-roll}}$ is the "reference" rolling that would be measured by a stationary observer. Identity (1) gives

$$(d\boldsymbol{\theta}^O\times(\mathbf{x}^q-\mathbf{x}^p))\cdot\mathbf{t} = d\boldsymbol{\theta}^O\cdot((\mathbf{x}^q-\mathbf{x}^p)\times\mathbf{t}) \tag{58}$$

and since $\mathbf{x}^q-\mathbf{x}^p=\mathbf{l}$, Eqs. (36) and (37) imply that

$$(\mathbf{x}^q-\mathbf{x}^p)\times\mathbf{t}=l^{\perp t}\mathbf{z}^t, \tag{59}$$

and hence

$$(\mathrm{d}\boldsymbol{\theta}^O\times(\mathbf{x}^q-\mathbf{x}^p))\cdot\mathbf{t}=l^{\perp t}\mathrm{d}\boldsymbol{\theta}^O\cdot\mathbf{z}^t. \tag{60}$$

Applying these substitutions to (57) gives

$$\mathrm{d}\hat{u}^{\mathrm{c,t\text{-}roll}}=\mathrm{d}u^{\mathrm{c,t\text{-}roll}}+\frac{1}{2}[-(\mathrm{d}\boldsymbol{\theta}^O\cdot\mathbf{z}^t)((\mathbf{r}^p+\mathbf{r}^q)\cdot\boldsymbol{\lambda}^t)]$$

$$+\frac{1}{2}\frac{1}{l^{\perp t}}[l^{\perp t}(\mathrm{d}\boldsymbol{\theta}^O\cdot\mathbf{z}^t)((\mathbf{r}^p+\mathbf{r}^q)\cdot\boldsymbol{\lambda}^t)] \tag{61}$$

or

$$\mathrm{d}\hat{u}^{\mathrm{c,t\text{-}roll}}=\mathrm{d}u^{\mathrm{c,t\text{-}roll}}, \tag{62}$$

and the objectivity (a) is shown.

Objectivity (b)

Since the scalar products in (49) do not change with finite rigid-body translations and rotations of the coordinate system, the **t**-directional rolling displacement remains unchanged.

5 Discussion: Special Cases

In this section we shall discuss a few special situations, which should give a clearer understanding of this definition of rolling displacements.

Situations when the rolling displacement in a certain direction does not exist

The vectors $\mathbf{l}^{\perp t}$, $\mathbf{l}^{\perp n}$, or $\mathbf{l}^{\perp w}$ (generally denoted as $\mathbf{l}^{\perp(\cdot)}$) are used for determining the unit vectors $\boldsymbol{\lambda}^{(\cdot)}$ in (31). If the branch vector **l** is aligned with any one of **n**, **t**, or **w**, then the corresponding length $l^{\perp(\cdot)}$ will be zero, and the $(\cdot)$-directional rolling displacement cannot be defined. The most likely instance of such a situation is when **l** is aligned with the contact normal, **n**. This will happen, for instance, in case of two spherical particles; since for them $l^{\perp n}$ is always zero, and the **n**-directional rolling displacement does not exist.

If $l^{\perp(\cdot)}$ is zero for a certain direction from among **n**, **t**, or **w**, it cannot be zero for the other two directions (the branch vector can be aligned only with one of the three perpendicular directions). Consequently we cannot find more than one direction for which the rolling displacement does not exist.

The unit vectors **t** and **w** were chosen at the beginning of the analysis to be perpendicular to each other and **n**, satisfying (1) but otherwise arbitrarily. If the branch vector is perpendicular to **n** (see, for example, the situation in Fig. 6), we can always define **t** and **w** such as one of them (let us say **w**) would not exist, by simply aligning **w** with the branch vector, so that $l^{\perp w}$ will be zero.

Two-dimensional analysis as a special case of the general three-dimensional definition

In two dimensions the unit vector **w** is perpendicular to the plane of analysis (Section 1). The $l^{\perp(\cdot)}$ lengths and the $\boldsymbol{\lambda}^{(\cdot)}$ and $\mathbf{z}^{(\cdot)}$ vectors for determining the **n**, **t**, and **w**-directional rolling displacements are

$$l^{\perp n}=-\mathbf{t}\cdot\mathbf{l};\quad \boldsymbol{\lambda}^n=-\mathbf{t};\quad \mathbf{z}^n=\mathbf{w} \tag{63}$$

$$l^{\perp t}=\mathbf{n}\cdot\mathbf{l};\quad \boldsymbol{\lambda}^t=\mathbf{n};\quad \mathbf{z}^t=\mathbf{w} \tag{64}$$

$$l^{\perp w}=l;\quad \boldsymbol{\lambda}^w=\frac{\mathbf{l}}{l};\quad \mathbf{z}^w=\boldsymbol{\lambda}^w\times\mathbf{w}, \tag{65}$$

where l is the length of the branch vector: $l=\sqrt{\mathbf{l}\cdot\mathbf{l}}$. If all the displacements take place in the (**n**,**t**) plane, then by applying (52), (53), and (54) the rolling displacements are

$$\mathrm{d}u^{\mathrm{c,t\text{-}roll}}=\frac{1}{2}\left[(\mathrm{d}\boldsymbol{\theta}^p\cdot\mathbf{w})(\mathbf{r}^p\cdot\mathbf{n})+(\mathrm{d}\boldsymbol{\theta}^q\cdot\mathbf{w})(\mathbf{r}^q\cdot\mathbf{n})\right.$$
$$\left.-\frac{(\mathbf{r}^p+\mathbf{r}^q)\cdot\mathbf{n}}{\mathbf{l}\cdot\mathbf{n}}((\mathrm{d}\mathbf{u}^q-\mathrm{d}\mathbf{u}^p)\cdot\mathbf{t})\right] \tag{66}$$

$$\mathrm{d}u^{\mathrm{c,n\text{-}roll}}=\frac{1}{2}\left[(\mathrm{d}\boldsymbol{\theta}^p\cdot\mathbf{w})(\mathbf{r}^p\cdot(-\mathbf{t}))+(\mathrm{d}\boldsymbol{\theta}^q\cdot\mathbf{w})(\mathbf{r}^q\cdot(-\mathbf{t}))\right.$$
$$\left.-\frac{(\mathbf{r}^p+\mathbf{r}^q)\cdot(-\mathbf{t})}{\mathbf{l}\cdot(-\mathbf{t})}((\mathrm{d}\mathbf{u}^q-\mathrm{d}\mathbf{u}^p)\cdot\mathbf{n})\right] \tag{67}$$

$$\mathrm{d}u^{\mathrm{c,w\text{-}roll}}=0. \tag{68}$$

Because of Identity 3, expression (66) is identical to (19), and (67) is the same as (32), which means that *the general two-dimensional situation is a special case of three-dimensional analysis*. Furthermore, if the two particles are circles with radii $R^p=\mathbf{r}^p\cdot\mathbf{n}$ and $R^q=-\mathbf{r}^q\cdot\mathbf{n}$, the n-directional rolling displacement does not exist, and (66) becomes identical to (21): the Iwashita-Oda rolling is a special case of the general three-dimensional definition.

Situations when a rolling displacement vector can be defined

The definitions of the scalar quantities $\mathrm{d}u^{\mathrm{c,t\text{-}roll}}$, $\mathrm{d}u^{\mathrm{c,n\text{-}roll}}$, and $\mathrm{d}u^{\mathrm{c,w\text{-}roll}}$ may have given the reader an impression that these scalars are the **t**, **n** and **w**-coordinates of a single vector $\mathrm{d}\mathbf{u}^{\mathrm{c,roll}}$. This is not true, as such a $\mathrm{d}\mathbf{u}^{\mathrm{c,roll}}$ vector does not exist in general. Instead of (52), (53), and (54), we can write the general form of a **v**-directional rolling displacement as

$$\mathrm{d}u^{\mathrm{c,v\text{-}roll}}=\frac{1}{2}\left[(\mathrm{d}\boldsymbol{\theta}^p\times\mathbf{r}^p+\mathrm{d}\boldsymbol{\theta}^q\times\mathbf{r}^q)-((\mathrm{d}\boldsymbol{\theta}^p\cdot\boldsymbol{\lambda}^v)(\mathbf{r}^p\times\boldsymbol{\lambda}^v)\right.$$
$$\left.+(\mathrm{d}\boldsymbol{\theta}^q\cdot\boldsymbol{\lambda}^v)(\mathbf{r}^q\times\boldsymbol{\lambda}^v))-\frac{(\mathbf{r}^p+\mathbf{r}^q)\cdot\boldsymbol{\lambda}^v}{(\mathbf{r}^p-\mathbf{r}^q)\cdot\boldsymbol{\lambda}^v}(\mathrm{d}\mathbf{u}^q-\mathrm{d}\mathbf{u}^p)\right]\cdot\mathbf{v}. \tag{69}$$

Both the second and third terms in the brackets will change with a changing direction of **v**, so the quantity in the brackets is not independent of direction **v**. This means that (in general) the scalar $\mathrm{d}u^{\mathrm{c,v\text{-}roll}}$ cannot be considered as the **v**-directional projection of a single vector $\mathrm{d}\mathbf{u}^{\mathrm{c\text{-}roll}}$.

But as a special case, if the branch vector $\mathbf{l}=(\mathbf{r}^p-\mathbf{r}^q)$ is aligned with either of **t**, **n**, or **w**, the situation is somewhat different. Let us consider, for instance, the $\mathbf{l}\|\mathbf{n}$ case (the other two cases, i.e., $\mathbf{l}\|\mathbf{t}$ and $\mathbf{l}\|\mathbf{w}$, can analogously be analyzed). Now the $\mathrm{d}u^{\mathrm{c,n\text{-}roll}}$ rolling displacement does not exist, and

$$\boldsymbol{\lambda}^t=\boldsymbol{\lambda}^w=\mathbf{n} \tag{70}$$

$$\mathbf{z}^t=\mathbf{w};\quad \mathbf{z}^w=-\mathbf{t}. \tag{71}$$

After some algebra, (52) and (54) gives

$$\mathrm{d}u^{\mathrm{c,t\text{-}roll}}=\frac{1}{2}\left[(\mathrm{d}\boldsymbol{\theta}^p\times\mathbf{n})(\mathbf{r}^p\cdot\mathbf{n})+(\mathrm{d}\boldsymbol{\theta}^q\times\mathbf{n})(\mathbf{r}^q\cdot\mathbf{n})\right.$$
$$\left.-\frac{((\mathbf{r}^p+\mathbf{r}^q)\cdot\mathbf{n})}{((\mathbf{r}^p-\mathbf{r}^q)\cdot\mathbf{n})}(\mathrm{d}\mathbf{u}^q-\mathrm{d}\mathbf{u}^p)\right]\cdot\mathbf{t} \tag{72}$$

$$\mathrm{d}u^{\mathrm{c,w\text{-}roll}}=\frac{1}{2}\left[(\mathrm{d}\boldsymbol{\theta}^p\times\mathbf{n})(\mathbf{r}^p\cdot\mathbf{n})+(\mathrm{d}\boldsymbol{\theta}^q\times\mathbf{n})(\mathbf{r}^q\cdot\mathbf{n})\right.$$
$$\left.-\frac{((\mathbf{r}^p+\mathbf{r}^q)\cdot\mathbf{n})}{((\mathbf{r}^p-\mathbf{r}^q)\cdot\mathbf{n})}(\mathrm{d}\mathbf{u}^q-\mathrm{d}\mathbf{u}^p)\right]\cdot\mathbf{w}. \tag{73}$$

The expression in brackets in (72) is the same as in (73). More generally, we can consider a unit vector **v** in the (**t**,**w**) plane (**v** is perpendicular to the branch vector, but otherwise arbitrary). The **v**-directional rolling displacement is

$$du^{c,v\text{-}roll}=\frac{1}{2}\left[(d\boldsymbol{\theta}^p\times\mathbf{n})(\mathbf{r}^p\cdot\mathbf{n})+(d\boldsymbol{\theta}^q\times\mathbf{n})(\mathbf{r}^q\cdot\mathbf{n})-\frac{((\mathbf{r}^p+\mathbf{r}^q)\cdot\mathbf{n})}{((\mathbf{r}^p-\mathbf{r}^q)\cdot\mathbf{n})}(d\mathbf{u}^q-d\mathbf{u}^p)\right]\cdot\mathbf{v}. \quad (74)$$

The bracket contains the same vector as in (72) or (73), independently of the direction of $\mathbf{v}$. To summarize, in the special $\mathbf{l}\|\mathbf{n}$ case when the $\mathbf{n}$-directional rolling displacement does not exist, a tangential rolling displacement vector can be defined as

$$d\hat{\mathbf{u}}^{c,roll}=\frac{1}{2}\left[(d\boldsymbol{\theta}^p\times\mathbf{n})(\mathbf{r}^p\cdot\mathbf{n})+(d\boldsymbol{\theta}^q\times\mathbf{n})(\mathbf{r}^q\cdot\mathbf{n})-\frac{((\mathbf{r}^p+\mathbf{r}^q)\cdot\mathbf{n})}{((\mathbf{r}^p-\mathbf{r}^q)\cdot\mathbf{n})}(d\mathbf{u}^q-d\mathbf{u}^p)\right], \quad (75)$$

whose $\mathbf{t}$ and $\mathbf{w}$-directional components are $du^{c,t\text{-}roll}$ and $du^{c,w\text{-}roll}$:

$$du^{c,t\text{-}roll}=d\hat{\mathbf{u}}^{c,roll}\cdot\mathbf{t};\quad du^{c,w\text{-}roll}=d\hat{\mathbf{u}}^{c,roll}\cdot\mathbf{w}; \quad (76)$$

and

$$d\hat{\mathbf{u}}^{c,roll}=du^{c,t\text{-}roll}\cdot\mathbf{t}+du^{c,w\text{-}roll}\cdot\mathbf{w}. \quad (77)$$

The special case of $\mathbf{r}^p\|\mathbf{n}$ and $\mathbf{r}^q\|\mathbf{n}$

In this case it is easy to derive an expression for the rolling displacement vector $d\hat{\mathbf{u}}^{c,roll}$ in terms of the relative translations and rotations at the contact. We can write $\mathbf{r}^p$ and $\mathbf{r}^q$ as

$$\mathbf{r}^p=R^p\mathbf{n}\quad\text{and}\quad\mathbf{r}^q=-R^q\mathbf{n} \quad (78)$$

where R^p and R^q are the distance of the contact point from the reference points of particles "p" and "q." Since $\mathbf{l}\|\mathbf{n}$, the relations (70) and (71) are still valid, and

$$l^{\perp t}=l^{\perp w}=R^p+R^q. \quad (79)$$

The tangential rolling displacement vector in (75) can be written as

$$d\hat{\mathbf{u}}^{c,roll}=\frac{1}{2}\left[-2\frac{R^pR^q}{R^p+R^q}(d\boldsymbol{\theta}^q-d\boldsymbol{\theta}^p)\times\mathbf{n})-\frac{R^p-R^q}{R^p+R^q}(d\mathbf{u}^{qc}-d\mathbf{u}^{pc})\right]. \quad (80)$$

From this expression, $du^{c,t\text{-}roll}$ and $du^{c,w\text{-}roll}$ are calculated by (76).

Conclusions

We have presented a general definition of rolling displacement for the interactions of two and three-dimensional particles of arbitrary shape. The defined rolling displacement is shown to be objective, and reduces to that of Iwashita and Oda [1] when the two particles are circular disks. The type of rolling in this paper is not based on following the trail that the contact point would trace across the surfaces of the two particles, but instead provides an independent definition of rolling by analyzing the displacements of the material points that form the contact. The proposed definition can either be easily implemented in particle simulation computer codes to investigate the relative importance and effect of rolling, or serve as a microlevel kinematical state variable in order to develop a global state variable for the characterization of rolling phenomena during the deformation of granular materials.

Acknowledgments

This research was supported by the OTKA F-31889 grant.

References

[1] Iwashita, K., and Oda, M., 1998, "Rolling Resistance at Contacts in Simulations of Shear Band Development by DEM," J. Eng. Mech., **124(3)**, pp. 285–292.
[2] Oda, M., Konishi, J., and Nemat-Nasser, S., 1982, "Experimental Micromechanical Evaluation of Strength of Granular Materials: Effects of Particle Rolling," Mech. Mater., **1**, pp. 269–283.
[3] Rothenburg, L., and Bathurst, R. J., 1993, "Influence of Particle Eccentricity on the Micromechanical Behavior of Granular Materials," Mech. Mater., **16**(1–2), pp. 141–152.
[4] Ting, J. M., Meachum, L., and Rowell, J. D., 1995, "Effect of Particle Shape on the Strength and Deformation Mechanisms of Ellipse-Shaped Granular Assemblages," Eng. Comput., **12**, pp. 99–108.
[5] Matsushima, T., and Konagai, K., 2001, "Grain-Shape Effect on Peak Strength of Granular Materials," *Proc. 10 IACMAG: Computer Methods and Advances in Geomechanics*, Desai et al., eds., Balkema, pp. 361–366.
[6] Lin, X., and Ng, T.-T., 1997, "A Three-Dimensional Discrete Element Model Using Arrays of Ellipsoids," Geotechnique, **47**(2), pp. 319–329.
[7] Montana, D. J., 1988, "The Kinematics of Contact and Grasp," Int. J. Robot. Res., **7**(3), pp. 17–32.

Three-Dimensional Vibration Analysis of Thick, Complete Conical Shells

Jae-Hoon Kang
Assistant Professor
Department of Architectural Engineering,
Chung-Ang University,
Seoul 156–756, Korea
e-mail: jhkang@cau.ac.kr

Arthur W. Leissa
Adjunct Professor
Colorado State University,
Fort Collins, CO 80523-1374

A three-dimensional (3D) method of analysis is presented for determining the free vibration frequencies and mode shapes of thick, complete (not truncated) conical shells of revolution. Unlike conventional shell theories, which are mathematically two-dimensional (2D), the present method is based upon the 3D dynamic equations of elasticity. Displacement components u_r, u_z, and u_θ in the radial, axial, and circumferential directions, respectively, are taken to be sinusoidal in time, periodic in θ, and algebraic polynomials in the r and z-directions. Potential (strain) and kinetic energies of the conical shells are formulated, the Ritz method is used to solve the eigenvalue problem, thus yielding upper bound values of the frequencies by minimizing the frequencies. As the degree of the polynomials is increased, frequencies converge to the exact values. Convergence to four-digit exactitude is demonstrated for the first five frequencies of the conical shells. Novel numerical results are presented for thick, complete conical shells of revolution based upon the 3D theory. Comparisons are also made between the frequencies from the present 3D Ritz method and a 2D thin shell theory. [DOI: 10.1115/1.1767843]

1 Introduction

The literature that addresses the free vibration of *thick* conical shells based upon 3D analyses is quite limited. Most of the existing literature describes the vibration analysis for *thin* conical shells and based upon a thin shell or membrane type of 2D shell theory, [1]. Among the literature on the thin conical shells, the vast majority of numerical results for the free vibrations deal with *truncated* conical shells mainly because of singularities occurring at the vertex of a *complete* (not truncated) conical shell.

The first contribution to the 3D analysis of conical shells was by Leissa and So [2] applying the Ritz method. Buchanan [3] and Buchanan and Wong [4] analyzed conical shells by a 3D finite element method. The present authors, [5,6], gave numerical results for shells having linearly varying wall thickness. However, all the above literature related to 3D analyses is limited to *truncated* conical shells. A search of the literature has revealed no 3D analysis of *complete* conical shells of revolution.

In recent research, [2–6], on vibrations of thick *truncated* conical shells using 3D methods a conical coordinate system has been used, whose original introduction was made by Leissa and So [2]. In the present analysis, a cylindrical coordinate system is applied, instead of the conical one, to avoid some complications in evaluating energy integrals in the vicinity of the vertex.

In the present 3D analysis, the Ritz method is used to obtain accurate frequencies. Although the method itself does not yield exact solutions, proper use of displacement components in the form of algebraic polynomials permits one to obtain frequency upper bounds that are as close to the exact values as desired. Frequencies presented in this work are thus obtained that are very accurate, being exact to four significant figures.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, December 4, 2002; final revision, October 16, 2003. Associate Editor: O. O'Reilly. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

2 Method of Analysis

A representative cross section of a thick, complete conical shell of revolution with the vertex half-angle α, the radius of bottom circle of the cone R, and the uniform wall thickness h ($=H\sin\alpha$), where H is the vertical slant thickness, is shown in Fig. 1. The cylindrical coordinate system (r,z,θ), also shown in the figure, is used in the analysis. The axes of r, z, and θ are the radial, axial, and circumferential coordinates, respectively. The origin of the (r,z) coordinates is located at the vertex of the midplane of the conical shell. Thus the ranges of the coordinates are given for the conical shell by

$$0\leqslant r\leqslant R,\quad r\cot\alpha-H/2\leqslant z\leqslant r\cot\alpha+H/2,\quad 0\leqslant\theta\leqslant 2\pi. \tag{1}$$

Utilizing tensor analysis, [7], the three equations of motion in the coordinate system (r,z,θ) are found to be (see also [8], p. 184):

$$\sigma_{rr,r}+\sigma_{rz,z}+\frac{1}{r}(\sigma_{rr}-\sigma_{\theta\theta}+\sigma_{r\theta,\theta})=\rho\ddot{u}_r, \tag{2a}$$

$$\sigma_{rz,r}+\sigma_{zz,z}+\frac{1}{r}(\sigma_{rz}+\sigma_{z\theta,\theta})=\rho\ddot{u}_z, \tag{2b}$$

$$\sigma_{r\theta,r}+\sigma_{z\theta,z}+\frac{1}{r}(2\sigma_{r\theta}+\sigma_{\theta\theta,\theta})=\rho\ddot{u}_\theta, \tag{2c}$$

where the σ_{ij} are the normal ($i=j$) and shear ($i\neq j$) stress components; u_r, u_z, and u_θ are the displacement components in the r, z, and θ directions, respectively; ρ is mass density per unit volume; the commas indicate spatial derivatives; and the dots denote time derivatives.

The well-known relationships between the tensorial stress (σ_{ij}) and tensorial strain (ε_{ij}) of isotropic, linear elasticity are

$$\sigma_{ij}=\lambda\varepsilon\,\delta_{ij}+2G\varepsilon_{ij}, \tag{3}$$

where λ and G are the Lamé parameters, expressed in terms of Young's modulus (E) and Poisson's ratio (ν) for an isotropic solid as

$$\lambda=\frac{E\nu}{(1+\nu)(1-2\nu)},\quad G=\frac{E}{2(1+\nu)}, \tag{4}$$

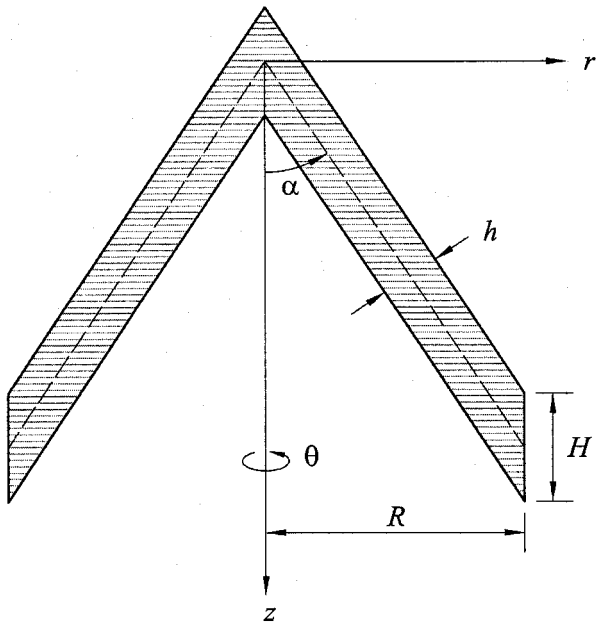


Fig. 1 A cross section of thick conical shell and the cylindrical coordinate system (r, z, θ)

$\varepsilon \equiv \varepsilon_{rr}+\varepsilon_{zz}+\varepsilon_{\theta\theta}$ is the trace of the strain tensor, and δ_{ij} is Kronecker's delta.

The three-dimensional tensorial strains are found to be related to the three displacements u_r, u_z, and u_θ, by, [7,8],

$$\varepsilon_{rr}=u_{r,r}, \quad \varepsilon_{zz}=u_{z,z}, \quad \varepsilon_{\theta\theta}=\frac{u_r+u_{\theta,\theta}}{r}, \tag{5a}$$

$$\varepsilon_{rz}=\frac{1}{2}(u_{r,z}+u_{z,r}), \quad \varepsilon_{r\theta}=\frac{1}{2}\left[u_{\theta,r}+\frac{u_{r,\theta}-u_\theta}{r}\right],$$

$$\varepsilon_{z\theta}=\frac{1}{2}\left[u_{\theta,z}+\frac{u_{z,\theta}}{r}\right]. \tag{5b}$$

Substituting Eqs. (3) and (5) into Eqs. (2), one obtains a set of second-order partial differential equations in u_r, u_z, and u_θ governing free vibrations. However, exact solutions of these equations are intractable because the independent variables are not all separable (that is, θ and t separate out, but r and z remain coupled), and also because of the variable coefficients that appear in many terms. Alternatively, one may approach the problem from an energy perspective.

Because the strains are related to the displacement components by Eqs. (5), unacceptable strain singularities may be encountered exactly at $r=0$ due to the term $1/r$. Such singularities may be avoided by numerically integrating at Gauss-Legendre points within the volume of the body such that $r\neq 0$ (for example, $10^{-5}\leqslant r/R\leqslant 1$ instead of $0\leqslant r/R\leqslant 1$).

During vibratory deformation of the body, its strain (potential) energy (V) is the integral over the domain (Ω):

$$V=\frac{1}{2}\int_\Omega (\sigma_{rr}\varepsilon_{rr}+\sigma_{zz}\varepsilon_{zz}+\sigma_{\theta\theta}\varepsilon_{\theta\theta}+2\sigma_{rz}\varepsilon_{rz}+2\sigma_{r\theta}\varepsilon_{r\theta} +2\sigma_{z\theta}\varepsilon_{z\theta})r\ dr\ dz\ d\theta. \tag{6}$$

Substituting Eqs. (3) and (5) into Eq. (6) yields the strain energy in terms of the three displacements:

$$V=\frac{1}{2}\int_\Omega [\lambda(\varepsilon_{rr}+\varepsilon_{zz}+\varepsilon_{\theta\theta})^2+2G\{\varepsilon_{rr}^2+\varepsilon_{zz}^2+\varepsilon_{\theta\theta}^2 +2(\varepsilon_{rz}^2+\varepsilon_{z\theta}^2+\varepsilon_{r\theta}^2)\}]r\ dr\ dz\ d\theta, \tag{7}$$

where the tensorial strains ε_{ij} are expressed in terms of the three displacements by Eqs. (5).

The kinetic energy (T) is simply

$$T=\frac{1}{2}\int_\Omega \rho(\dot{u}_r^2+\dot{u}_z^2+\dot{u}_\theta^2)r\,dr\,dz\,d\theta. \tag{8}$$

For convenience, the radial r and axial z coordinates are made dimensionless as $\psi\equiv r/R$ and $\zeta\equiv z/H$. Thus the ranges of the nondimensional cylindrical coordinates (ψ,ζ,θ) are given by

$$0<\psi\leqslant 1, \quad \delta_1\leqslant\zeta\leqslant\delta_2, \quad 0\leqslant\theta\leqslant 2\pi, \tag{9}$$

where

$$\delta_1\equiv\frac{R}{H}\psi\cot\alpha-\frac{1}{2}, \quad \delta_2\equiv\frac{R}{H}\psi\cot\alpha+\frac{1}{2}. \tag{10}$$

For the free, undamped vibration, the time (t) response of the three displacements is sinusoidal and, moreover, the circular symmetry of the conical shell allows the displacements to be expressed by

$$u_r(\psi,\zeta,\theta,t)=U_r(\psi,\zeta)\cos n\theta\sin(\omega t+\alpha), \tag{11a}$$

$$u_z(\psi,\zeta,\theta,t)=U_z(\psi,\zeta)\cos n\theta\sin(\omega t+\alpha), \tag{11b}$$

$$u_\theta(\psi,\zeta,\theta,t)=U_\theta(\psi,\zeta)\sin n\theta\sin(\omega t+\alpha), \tag{11c}$$

where U_r, U_z, and U_θ are displacement functions of ψ and ζ, ω is a natural frequency, and α is an arbitrary phase angle determined by the initial conditions. The circumferential wave number is taken to be an integer ($n=0,1,2,3,\ldots,\infty$), to ensure periodicity in θ. Then Eqs. (11) account for all free vibration modes except for the torsional ones. Torsional modes arise from an alternative set of solutions which are the same as Eqs. (11), except that $\cos n\theta$ and $\sin n\theta$ are interchanged. For $n>0$, this set duplicates the solutions of Eqs. (11), with the symmetry axes of the mode shapes being rotated. But for $n=0$ the alternative set reduces to $u_r=u_z=0$, $u_\theta=U_\theta^*\sin(\omega t+\alpha)$, which corresponds to the torsional modes. The displacements uncouple by circumferential wave number (n), leaving only coupling in r and z.

The Ritz method uses the maximum potential (strain) energy ($V_{\max}$) and the maximum kinetic energy ($T_{\max}$) functionals in a cycle of vibratory motion. The functionals are obtained by setting $\sin^2(\omega t+\alpha)$ and $\cos^2(\omega t+\alpha)$ equal to unity in Eqs. (7) and (8) after the displacements (11) are substituted, and by using the nondimensional coordinates ψ and ζ as follows:

$$V_{\max}=\frac{GR^2}{2H}\int_0^1\int_{\delta_1}^{\delta_2}\left[\left\{\frac{\lambda}{G}(\kappa_1+\kappa_2+\kappa_3)^2 +2(\kappa_1^2+\kappa_2^2+\kappa_3^2)+\kappa_4^2\right\}\Gamma_1+(\kappa_5^2+\kappa_6^2)\Gamma_2\right]\psi d\zeta d\psi, \tag{12}$$

$$T_{\max}=\frac{\rho\omega^2R^2H}{2}\int_0^1\int_{\delta_1}^{\delta_2}[(U_r^2+U_z^2)\Gamma_1+U_\theta^2\Gamma_2]\psi d\zeta d\psi, \tag{13}$$

where

$$\kappa_1\equiv\frac{H}{R}\frac{U_r+nU_\theta}{\psi}, \quad \kappa_2\equiv\frac{H}{R}U_{r,\psi}, \quad \kappa_3\equiv U_{z,\zeta}, \tag{14a}$$

$$\kappa_4 \equiv U_{r,\zeta} + \frac{H}{R} U_{z,\psi}, \quad \kappa_5 \equiv U_{\theta,\zeta} - \frac{H}{R}\frac{nU_z}{\psi},$$

$$\kappa_6 \equiv \frac{H}{R}\left[U_{\theta,\psi} - \frac{nU_r + U_\theta}{\psi}\right], \tag{14b}$$

and Γ_1 and Γ_2 are constants, defined by

$$\Gamma_1 \equiv \int_0^{2\pi} \cos^2 n\theta d\theta = \begin{cases} 2\pi & \text{if } n=0 \\ \pi & \text{if } n \geqslant 1 \end{cases},$$

$$\Gamma_2 \equiv \int_0^{2\pi} \sin^2 n\theta d\theta = \begin{cases} 0 & \text{if } n=0 \\ \pi & \text{if } n \geqslant 1 \end{cases}. \tag{15}$$

It should be mentioned here that the *constant* limits of integration δ_1 and δ_2 in Eqs. (12) and (13) correspond to the cone edges ($r=R$) shown in Fig. 1. For other edges, including ones normal to the cone surfaces, δ_1 and δ_2 would be variables. From Eqs. (4) it is seen that the the nondimensional constant λ/G in Eq. (12) involves only ν; i.e.,

$$\frac{\lambda}{G} = \frac{2\nu}{1-2\nu}. \tag{16}$$

The displacement functions U_r, U_z, and U_θ in Eqs. (11) are further assumed as algebraic polynomials,

$$U_r(\psi,\zeta) = \eta_r \sum_{i=0}^{I}\sum_{j=0}^{J} A_{ij}\psi^i\zeta^j \tag{17a}$$

$$U_z(\psi,\zeta) = \eta_z \sum_{k=0}^{K}\sum_{l=0}^{L} B_{kl}\psi^k\zeta^l \tag{17b}$$

$$U_\theta(\psi,\zeta) = \eta_\theta \sum_{m=0}^{M}\sum_{n=0}^{N} C_{mn}\psi^m\zeta^n \tag{17c}$$

and similarly for U_θ^*, where i, j, k, l, m, and n are integers; I, J, K, L, M, and N are the highest degrees taken in the polynomial terms; A_{ij}, B_{kl}, and C_{mn} are arbitrary coefficients to be determined, and the η are functions depending upon the geometric boundary conditions to be enforced. For example:

1. completely free: $\eta_r = \eta_z = \eta_\theta = 1$,
2. the bottom edge ($r=R$ or $\psi=1$) fixed: $\eta_r = \eta_z = \eta_\theta = (\psi-1)^2$.

The functions of η shown above, impose only the necessary geometric constraints such as displacement and slope boundary conditions. Together with the algebraic polynomials in Eqs. (17), they form function sets which are mathematically complete (Kantorovich and Krylov [9], pp. 266–268). Thus, the function sets are capable of representing any 3D motion of the plate with increasing accuracy as the indices $I, J, \ldots, N$ are increased. In the limit, as sufficient terms are taken, all internal kinematic constraints vanish, and the functions (17) will approach the exact solution as closely as desired.

The eigenvalue problem is formulated by minimizing the free vibration frequencies with respect to the arbitrary coefficients A_{ij}, B_{kl}, and C_{mn}, thereby minimizing the effects of the internal constraints present, when the function sets are finite. This corresponds to the equations [10]:

$$\frac{\partial}{\partial A_{ij}}(V_{\max} - T_{\max}) = 0, \quad (i=0,1,2,\ldots,I; j=0,1,2,\ldots,J) \tag{18a}$$

$$\frac{\partial}{\partial B_{kl}}(V_{\max} - T_{\max}) = 0, \quad (k=0,1,2,\ldots,K; l=0,1,2,\ldots,L) \tag{18b}$$

$$\frac{\partial}{\partial C_{mn}}(V_{\max} - T_{\max}) = 0, \quad (m=0,1,2,\ldots,M; n=0,1,2,\ldots,N). \tag{18c}$$

Equations (18) yield a set of $(I+1)(J+1)+(K+1)(L+1)+(M+1)(N+1)$ linear, homogeneous, algebraic equations in the unknowns A_{ij}, B_{kl}, and C_{mn}. For a nontrivial solution, the determinant of the coefficient matrix is set equal to zero, which yields the frequencies (eigenvalues). These frequencies are upper bounds on the exact values. The mode shape (eigenfunction) corresponding to each frequency is obtained, in the usual manner, by substituting each ω back into the set of algebraic equations, and solving for the ratios of coefficients.

3 Convergence Studies

To guarantee the accuracy of frequencies obtained by the procedure described above, it is necessary to conduct some convergence studies to determine the number of terms required in the power series of Eqs. (17). A convergence study is based upon the fact that all the frequencies obtained by the Ritz method should converge to their exact values, when the displacement functions (17) are used, and in an upper bound manner. If the results do not converge properly, or converge too slowly, it is likely that the assumed displacements may be poor ones, or be missing some functions from a minimal complete set of polynomials.

Table 1 is such a study for a completely free, conical shell of revolution with a *small* vertex half-angle ($\alpha=15$ deg) and h/R

Table 1 Convergence of frequencies in $\omega R\sqrt{\rho/G}$ of a completely free, complete conical shell of revolution for the five lowest axisymmetric modes ($n=0$) with $\alpha=15$ deg and $h/R=0.3$ for $\nu=0.3$

TZ	*TR*	*DET*	1	2	3	4	5
2	2	8	1.515	2.740	4.383	6.664	12.75
2	3	12	1.474	2.360	3.529	6.183	7.782
2	4	16	1.469	2.325	3.308	5.331	6.316
2	5	20	1.466	2.301	3.234	4.768	5.629
2	6	24	1.463	2.299	3.219	4.752	5.324
2	7	28	1.463	2.298	3.209	4.735	5.284
3	2	12	1.445	2.007	3.017	3.934	5.503
3	3	18	1.439	1.883	2.919	3.165	4.626
3	4	24	1.433	1.827	2.796	2.893	4.323
3	5	30	1.431	1.813	2.643	2.850	4.157
3	6	32	1.431	1.795	2.554	2.831	4.066
4	2	16	1.434	1.815	2.636	2.897	4.248
4	3	24	1.431	1.758	2.356	2.840	3.968
4	4	32	1.430	1.738	2.238	2.815	3.552
4	5	40	1.429	1.725	2.114	2.809	3.131
4	6	48	1.428	1.722	2.073	2.807	2.939
4	7	56	1.428	1.720	2.051	2.795	2.845
5	2	20	1.431	1.745	2.251	2.827	3.364
5	3	30	1.429	1.722	2.073	2.810	3.029
5	4	40	1.428	1.717	2.020	2.779	2.809
5	5	50	1.428	1.715	1.975	2.571	2.806
6	2	24	1.430	1.729	2.045	2.818	2.869
6	3	36	1.428	1.716	1.968	2.539	2.808
6	4	48	1.428	1.715	1.948	2.430	2.803
7	2	28	1.429	1.728	1.977	2.505	2.816
7	3	42	1.428	1.715	1.943	2.340	2.800
7	4	56	**1.427**	**1.714**	1.937	2.299	2.790
7	5	70	1.427	1.714	1.936	2.279	2.749
8	2	32	1.429	1.727	1.971	2.360	2.810
8	3	48	1.427	1.715	1.939	2.288	2.760
8	4	64	1.427	1.714	**1.936**	2.273	2.716
8	5	80	1.427	1.714	1.936	**2.269**	**2.678**

Notes:
TZ = Total numbers of polynomial terms used in the z (or ζ) direction
TR = Total numbers of polynomial terms used in the r (or ψ) direction
DET = Frequency determinant order

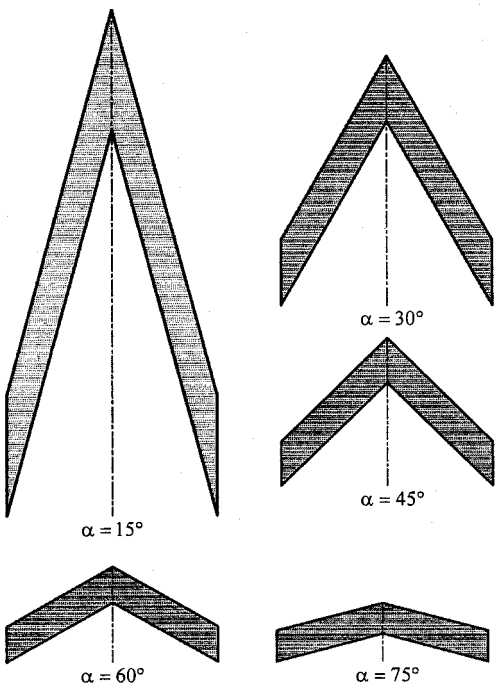


Fig. 2 Cross sections of conical shells with $h/R=0.3$

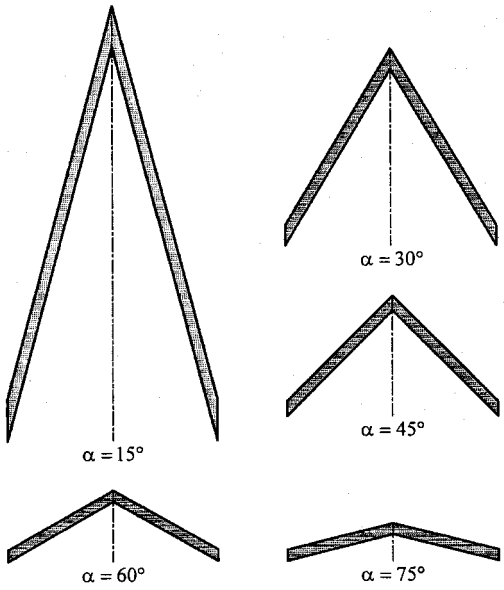


Fig. 3 Cross sections of conical shells with $h/R=0.1$

Table 2 Convergence of frequencies in $\omega R\sqrt{\rho/G}$ of a completely free, complete conical shell of revolution for the five lowest bending modes for $n=2$ with $\alpha=75$ deg and $h/R=0.1$ for $\nu=0.3$

TZ	*TR*	*DET*	1	2	3	4	5
2	2	12	0.7713	2.651	5.267	10.01	13.89
2	4	24	0.2679	2.070	2.467	4.236	6.660
2	6	36	0.2670	1.773	2.404	4.093	4.277
2	8	48	0.2667	1.764	2.401	3.999	4.252
2	10	60	0.2665	1.764	2.401	3.998	4.252
2	12	72	0.2665	1.764	2.401	3.997	4.252
3	2	18	0.5671	2.558	4.863	8.714	10.83
3	4	36	0.2587	1.808	2.409	4.217	5.161
3	6	54	0.2580	1.575	2.383	3.605	4.238
3	8	72	0.2578	1.569	2.381	3.533	4.235
3	10	90	0.2578	1.569	2.381	3.532	4.235
4	2	24	0.5259	2.550	4.732	7.918	9.857
4	4	48	0.2581	1.694	2.401	4.215	4.852
4	6	72	0.2577	1.565	2.381	3.569	4.237
4	8	96	**0.2575**	1.563	**2.380**	3.505	**4.234**
4	10	120	0.2575	**1.562**	2.380	**3.503**	4.234
5	2	30	0.5071	2.546	4.711	7.854	9.830
5	4	60	0.2579	1.628	2.394	4.201	4.527
5	6	90	0.2576	1.563	2.381	3.541	4.236
5	8	120	0.2575	1.562	2.380	3.503	4.234
5	10	150	0.2575	1.562	2.380	3.503	4.234
6	2	36	0.4897	2.541	4.682	7.734	9.751
6	4	72	0.2578	1.619	2.393	4.134	4.308
6	6	108	0.2576	1.563	2.381	3.525	4.235
6	8	144	0.2575	1.562	2.380	3.503	4.234
6	9	162	0.2575	1.562	2.380	3.503	4.234

Notes:
TZ = Total numbers of polynomial terms used in the z (or ζ) direction
TR = Total numbers of polynomial terms used in the r (or ψ) direction
DET = Frequency determinant order

$=0.3$ (see Fig. 2) for $\nu=0.3$. The table lists the first five nondimensional frequencies in $\omega R\sqrt{\rho/G}$ for the axisymmetric modes $(n=0)$.

To make the study of convergence less complicated, equal numbers of polynomial terms were taken in both the r (or ψ) coordinate (i.e., $I=K=M$) and z (or ζ) coordinate (i.e., $J=L=N$), although some computational optimization could be obtained for some configurations and some mode shapes by using unequal number of polynomial terms. The symbols **TZ** and **TR** in the table indicate the total numbers of polynomial terms used in the z (or ζ) and r (or ψ) directions, respectively. Note that the frequency determinant order **DET** is related to **TZ** and **TR** as follows:

$$\mathbf{DET}=\left\{\begin{array}{ll}\mathbf{TZ}\times\mathbf{TR} & \text{for torsional modes } (n=0)\\ 2\times\mathbf{TZ}\times\mathbf{TR} & \text{for axisymmetric modes } (n=0)\\ 3\times\mathbf{TZ}\times\mathbf{TR} & \text{for general modes } (n\geqslant 1)\end{array}\right\}. \tag{19}$$

Table 1 shows the monotonic convergence of all five frequencies as **TZ** $(=J+1, L+1,$ and $N+1$ in Eqs. (17)) are increased, as well as **TS** $(=I+1, K+1,$ and $M+1$ in Eqs. (17)). One sees, for example, that the fundamental (i.e., lowest) nondimensional frequency $\omega R\sqrt{\rho/G}$ converges to four digits (1.427) when $2\times(7\times4)=56$ terms are used, which results in **DET**$=56$. Moreover, this accuracy requires using at least seven terms in the axial coordinate (**TZ**$=7$) and four in the radial coordinate (**TR**$=4$). Numbers in underlined, boldface type in Table 1 are the most accurate values (i.e., least upper bounds) achieved with the smallest determinant sizes.

Table 2 is a similar convergence study for a completely free, conical shell of revolution with a *large* vertex half-angle ($\alpha=75$ deg) and $h/R=0.1$ (see Fig. 3) for $\nu=0.3$. One sees that the fundamental frequency (0.2575) requires using at least (**TZ**,**TR**) $=(4,8)$ for the exactitude to four significant figures. It was noticed that as the vertex half-angle (α) becomes larger, more terms in r (or ψ) are needed to yield the converged values.

4 Numerical Results and Discussion

Tables 3 and 4 present the nondimensional frequencies in $\omega R\sqrt{\rho/G}$ of completely free, very thick ($h/R=0.3$) and moderately thick ($h/R=0.1$), complete conical shells of revolution, respectively, with the vertex half-angles of $\alpha=15$ deg, 30 deg, 45

Table 3 Nondimensional frequencies in $\omega R\sqrt{\rho/G}$ of completely free, conical shells with $h/R=0.3$ for $\nu=0.3$

n	s	$\alpha=15°$	$\alpha=30°$	$\alpha=45°$	$\alpha=60°$	$\alpha=75°$
	1	1.267	2.466	3.526	4.377	4.938
	2	1.978	3.865	5.573	7.007	8.027
0^T	3	2.597	5.081	7.333	9.206	10.31
	4	3.214	6.274	8.982	10.70	11.20
	5	3.875	7.566	10.73	11.38	11.68
	1	1.427	**1.655(4)**	**1.608(3)**	**1.481(3)**	**1.289(2)**
	2	1.714	2.296	2.824	3.210	3.413
0^A	3	1.935	3.233	3.787	3.926	3.904
	4	2.267	3.301	4.700	5.911	6.706
	5	2.665	4.257	6.197	7.699	8.563
	1	**0.6938(2)**	1.718	**2.024(4)**	**2.203(4)**	**2.317(4)**
	2	1.189	2.029	2.520	2.662	2.718
1	3	1.558	2.557	3.648	4.541	5.151
	4	1.602	2.832	3.797	4.908	5.654
	5	1.910	3.458	5.334	6.130	6.633
	1	**0.3949(1)**	**0.5155(1)**	**0.5999(1)**	**0.6626(1)**	**0.7032(1)**
	2	**1.008(4)**	**1.651(3)**	2.106	2.294	**2.339(5)**
2	3	1.603	2.583	2.873	3.220	3.502
	4	2.098	3.129	4.059	4.177	4.223
	5	2.452	3.903	4.640	5.755	6.503
	1	**0.7627(3)**	**1.101(2)**	**1.299(2)**	**1.427(2)**	**1.503(3)**
	2	1.563	2.591	3.295	3.528	3.587
3	3	2.200	3.519	3.835	4.328	4.721
	4	2.817	4.004	5.512	5.738	5.797
	5	3.454	5.088	5.777	6.979	7.804
	1	**1.061(5)**	**1.711(5)**	**2.075(5)**	**2.292(5)**	2.414
	2	2.059	3.432	4.417	4.630	4.675
4	3	2.893	4.516	4.786	5.439	5.933
	4	3.775	4.899	6.661	7.282	7.373
	5	4.477	6.209	7.105	8.138	9.032

Notes:
T=Torsional mode; A=Axisymmetric mode.
Numbers in parentheses identify frequency sequence.

Table 4 Nondimensional frequencies in $\omega R\sqrt{\rho/G}$ of completely free, conical shells with $h/R=0.1$ for $\nu=0.3$

n	s	$\alpha=15°$	$\alpha=30°$	$\alpha=45°$	$\alpha=60°$	$\alpha=75°$
	1	1.323	2.559	3.623	4.442	4.959
	2	2.161	4.181	5.925	7.273	8.125
0^T	3	2.971	5.751	8.159	10.03	11.21
	4	3.765	7.296	10.36	12.75	14.27
	5	4.556	8.842	12.54	15.45	17.31
	1	1.443	1.535	**1.394(5)**	**1.149(4)**	**0.7804(3)**
	2	1.677	2.012	2.135	2.121	1.937
0^A	3	1.945	2.713	3.164	3.384	3.451
	4	2.287	3.244	3.872	4.071	3.914
	5	2.664	3.565	4.576	5.487	6.035
	1	**0.7495(5)**	1.558	1.550	**1.370(5)**	**1.110(5)**
	2	1.276	2.040	2.438	2.617	2.654
1	3	1.540	2.134	2.550	2.665	2.720
	4	1.658	2.818	3.560	4.278	4.746
	5	1.954	2.872	4.057	5.095	5.744
	1	**0.1831(1)**	**0.2209(1)**	**0.2434(1)**	**0.2545(1)**	**0.2575(1)**
	2	**0.5495(3)**	**0.9244(4)**	**1.216(4)**	1.424	1.562
2	3	1.085	1.868	2.329	2.427	2.380
	4	1.525	2.557	2.798	3.183	3.503
	5	1.972	2.722	3.893	4.190	4.234
	1	**0.4151(2)**	**0.4931(2)**	**0.5390(2)**	**0.5678(2)**	**0.5839(2)**
	2	0.8984	**1.362(5)**	1.759	2.075	2.285
3	3	1.425	2.385	3.209	3.570	3.603
	4	1.942	3.442	3.648	4.046	4.484
	5	2.525	3.475	4.886	5.753	5.816
	1	**0.7001(4)**	**0.8487(3)**	**0.9283(3)**	**0.9782(3)**	**1.007(4)**
	2	1.346	1.933	2.433	2.829	3.089
4	3	1.934	3.087	4.101	4.620	4.678
	4	2.531	4.336	4.594	4.999	5.519
	5	3.144	4.424	5.932	7.248	7.411

Notes:
T=Torsional mode; A=Axisymmetric mode.
Numbers in parentheses identify frequency sequence.

deg, 60 deg, and 75 deg for $\nu=0.3$. Tables 3 and 4 correspond to Figs. 2 and 3, respectively. Thirty frequencies are given for each configuration, which arise from six circumferential wave numbers ($n=0^T,0^A,1,2,3,4$) and the first five modes ($s=1,2,3,4,5$) for each value of n, where the superscripts T and A indicate torsional and axisymmetric modes, respectively. The numbers in parentheses identify the first five frequencies for each configuration. The zero frequencies of rigid body modes are omitted from the tables.

It is seen from Tables 3 and 4 that the fundamental (lowest) frequencies of all the configurations are for modes having two ($n=2$) circumferential waves irrespective of the thickness ratio (h/R) and the vertex half-angle (α). Tables 3 and 4 show that the torsional mode ($n=0^T$) frequencies are all higher ones.

The numbers of polynomial terms of (**TZ**,**TR**) used for Tables 3 and 4 were (8,6), (7,7), (6,8), (5,10), and (4,12) for $\alpha=15$ deg, 30 deg, 45 deg, 60 deg, and 75 deg, respectively.

Numerically integrating the energy integrals (12) and (13) over $10^{-5}\leqslant\psi\leqslant 1$, instead $0\leqslant\psi\leqslant 1$, to avoid singularities encountered exactly at $\psi=0$, corresponds to leaving an extremely small cylindrical hole at the vertex. Table 5 shows the change in the lowest frequencies with five digits for each n for the shell ($\alpha=15$ deg, $h/R=0.3$) of Table 3, as the hole radius (a) is diminished ($a/R=0.2,10^{-1},10^{-2},10^{-3},10^{-4},10^{-5}$) to essentially vanish. That is, a very small hole has essentially no effect upon the frequencies.

5 Comparison With 2D Shell Theory

Dreher and Leissa [11,12] used the Donnell-Mushtari thin shell theory and the exact solution procedure involving expansion of the displacements in terms of power series to study the axisymmetric ($n=0$) free vibrations of completely free, complete conical shells of revolution.

Comparison of the present 3D Ritz method (3DR) is made in Table 6 for the nondimensional frequencies in $\omega R\sqrt{\rho/G}$ with the 2D shell theory (2DS) for the lowest four axisymmetric modes ($n=0$) with the stiffness parameter $K\equiv 12(1-\nu^2)\times(h/R)^{-2}/(\tan^4\alpha\cdot\cos^2\alpha)=10000$ and $\nu=0.3$. The parameter K is converted to the thickness ratio (h/R) for each of the α shown in Table 6. This shows that the corresponding shells range from being moderately thick ($h/R=0.151$, for $\alpha=15$ deg) to very thin ($h/R=0.00290$, for $\alpha=75$ deg). The percent difference in frequencies obtained by the two analyses is given by

$$\text{difference}(\%)=\frac{\text{2DS}-\text{3DR}}{\text{3DR}}\times 100. \tag{20}$$

It is observed that the 3D Ritz method yields lower frequencies than the 2D thin shell results in all the frequencies irrespective of

Table 5 Variation of $\omega R\sqrt{\rho/G}$ for the shell ($\alpha=15$ deg, $h/R=0.3$) of Table 3 with a small cylindrical hole of radius $r=a$ at its vertex (lowest frequencies for each n)

	a/R					
n	0.2	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
0^T	1.2870	1.2688	1.2671	1.2671	1.2671	1.2671
0^A	1.5191	1.4550	1.4275	1.4272	1.4271	1.4271
1	0.90685	0.76936	0.69509	0.69385	0.69384	0.69384
2	0.39479	0.39487	0.39488	0.39489	0.39489	0.39489
3	0.76265	0.76267	0.76272	0.76273	0.76273	0.76273
4	1.0611	1.0612	1.0613	1.0613	1.0613	1.0613

Table 6 Comparisons of nondimensional frequencies in $\omega R\sqrt{\rho/G}$ from the 3D and 2D methods of completely free, complete conical shells of revolution for the first four axisymmetric modes ($n=0$) with the stiffness parameter $K\equiv 12(1-\nu^2)(h/R)^{-2}/(\tan^4\alpha\cdot\cos^2\alpha)=100{,}000$ for $\nu=0.3$

		$[\alpha, h/R]$				
s	Method	[15°, 0.151]	[30°, 0.0362]	[45°, 0.0148]	[60°, 0.00697]	[75°, 0.00290]
1	2DS	1.742	1.562	1.275	0.9017	0.4668
	3DR	1.442	1.481	1.240	0.8916	0.4655
	(Difference %)	(20.8%)	(5.47%)	(2.82%)	(1.13%)	(0.28%)
2	2DS	2.190	1.963	1.603	1.134	0.5868
	3DR	1.697	1.795	1.527	1.111	0.5841
	(Difference %)	(29.1%)	(9.36%)	(4.98%)	(2.07%)	(0.46%)
3	2DS	2.686	2.408	1.966	1.390	0.7196
	3DR	1.978	2.245	1.888	1.365	0.7166
	(Difference %)	(35.8%)	(7.26%)	(4.13%)	(1.83%)	(0.42%)
4	2DS	3.315	2.882	2.353	1.664	0.8614
	3DR	2.322	2.792	2.352	1.663	0.8617
	(difference %)	(42.8%)	(3.22%)	(0.04%)	(0.06%)	(-0.03%)

Notes:
2DS = 2-D frequencies by the Donnell-Mushtari thin shell theory
3DR = 3-D frequencies by the Ritz method
Difference % = (2DS-3DR)/3DR×100

thickness parameter (h/R) and half-angle of vertex (α) as expected, except for the fourth ($s=4$) mode for ($\alpha, h/R$)=(75 deg,0.00290). An accurate 3D analysis should typically yield lower frequencies than those 2D thin shell theory, mainly because shear deformation and rotary inertia effects are accounted for in a 3D analysis, but not in 2D, thin shell theory. It is noticed that the frequencies by the 2D shell theory are accurate within 10% errors when the thickness parameter (h/R) is less than 0.04. The results show also the great inaccuracy of using a thin shell theory for the moderately thick shell ($h/R=0.151$).

6 Conclusions

A three-dimensional method of analysis has been presented for determining the free vibration frequencies and mode shapes of complete, conical shells of revolution. The analysis is based upon the equations of circular cylindrical shells, not conical ones. The 3D equations of the theory of elasticity are used in their general forms for isotropic, homogeneous materials. They are only limited to small strains. No other constraints are placed upon the displacements.

Numerical results for frequencies were presented for several completely free conical shells of revolution. The frequencies are exact up to four significant figures for the lower frequencies, but if greater accuracy is needed, especially for the higher frequencies, larger determinants would be required. Nevertheless, these determinant sizes are at least one order of magnitude less than those that would be required for equivalent accuracy by finite element analysis.

Results were presented for completely free shells. However, as described in Sec. 2, the procedure could also be used for the shells having the bottom ($r=R$) fixed.

References

[1] Leissa, A. W., 1973, *Vibration of Shells*, U.S. Government Printing Office, Washington DC.; reprinted by The Acoustical Society of America.
[2] Leissa, A. W., and So, J., 1995, "Three-Dimensional Vibrations of Truncated Hollow Cones," J. Vib. Control, **1**, pp. 145–158.
[3] Buchanan, G. R., 2000, "Vibration of Truncated Conical Cylinders of Crystal Class 6/mmm," J. Vib. Control, **6**, pp. 985–998.
[4] Buchanan, G. R., and Wong, F. T.-I., 2001, "Frequencies and Mode Shapes for Thick Truncated Hollow Cones," Int. J. Mech. Sci., **43**, pp. 2815–2832.
[5] Leissa, A. W., and Kang, J.-H., 1999, "Three-Dimensional Vibration Analysis of Thick Shells of Revolution," J. Eng. Mech., **125**(12), pp. 1365–1371.
[6] Kang, J.-H., and Leissa, A. W., 1999, "Three-Dimensional Vibrations of Hollow Cones and Cylinders With Linear Thickness Variations," J. Acoust. Soc. Am., **106**(2), pp. 748–755.
[7] Kang, J.-H., 1997, "Three-Dimensional Vibration Analysis of Thick Shells of Revolution With Arbitrary Curvature and Variable Thickness," Ph.D. dissertation, The Ohio State University.
[8] Sokolnikoff, I. S., 1956, *Mathematical Theory of Elasticity*, Second Ed., McGraw-Hill, New York.
[9] Kantorovich, L. V., and Krylov, V. I., 1958, *Approximate Methods in Higher Analysis*, Noordhoff, Gronigen, The Netherlands.
[10] Ritz, W., 1909, "Über eine neue Methode zur Lösung Gewisser Variationsprobleme der Mathematischen Physik," J. Reine Angew. Math., **135**, pp. 1–61.
[11] Dreher, J. F., 1966, "Axisymmetic Vibration of Thin Conical Shells," Ph.D. dissertation, The Ohio State University.
[12] Dreher, J. F., and Leissa, A. W., 1968, "Axisymmetic Vibration of Thin Conical Shells," *Proc. 4th Southwestern Conf. on Theoretical and Appl. Mech*, New Orleans, La., Feb. 29–Mar. 1, Pergamon Press, New York, pp. 163–181.

On the Acoustic Nonlinearity of Solid-Solid Contact With Pressure-Dependent Interface Stiffness

S. Biwa[1]
e-mail: biwa@energy.kyoto-u.ac.jp

S. Nakajima[2]

N. Ohno

Department of Micro System Engineering,
Nagoya University,
Nagoya 464-8603, Japan

Nonlinear interaction between elastic wave and contact interface, known to result in the so-called contact acoustic nonlinearity, is examined in a one-dimensional theoretical framework. The present analysis is based on a nonlinear interface stiffness model where the stiffness property of the contact interface is described as a function of the nominal contact pressure. The transmission/reflection coefficients for a normally incident harmonic wave, and the amplitudes of second harmonics as well as DC components arising at the contact interface are derived in terms of the interface stiffness properties and other relevant acoustic parameters. Implications of power-law relations between the linear interface stiffness and the contact pressure are examined in detail regarding the linear and nonlinear acoustic responses of the contact interface. Also, a plausible range of the relevant power-law exponent is provided from considerations based on the rough-surface contact mechanics. The analysis clarifies the qualitative contact-pressure dependence of various nonlinearity parameters based on different definitions. A particular power law is identified from existing experimental data for aluminum-aluminum contact, for which some explicit nonlinear characteristics are demonstrated. The theoretical contact-pressure dependence of the second harmonic generation at the contact interface is found to be in qualitative agreement with previous measurements. [DOI: 10.1115/1.1767169]

1 Introduction

Transmission and reflection characteristics of ultrasonic waves propagating through contact interface between solid bodies have long been known to be sensitive to the contact conditions such as the contact pressure and the true area of contact, and used as an experimental tool in contact mechanics and adhesion technology, [1–5]. Such effects can be explained by the fact that at an interface between two solids in contact, the load is supported by surface asperities. With increased load, more asperities come into contact while each asperity undergoes flattening deformation. Due to the change of the contact asperity configuration with varying contact pressure, the mechanical response of the contact interface involves certain nonlinear behavior. As a consequence, second or higher-order harmonics are generated when a wave interacts with the contact interface, which constitutes a relatively new area of study called contact acoustic nonlinearity, [6–8]. Several investigators, [6–10], have experimentally demonstrated the contact-pressure dependence of the second or higher-order harmonics in ultrasonic waves transmitted through or reflected from various contact interfaces between solids. Contact acoustic nonlinearity has been shown to exhibit much greater harmonic generation than classical material nonlinearities, [11,12], and is expected to offer valuable information of the contact conditions, with possible application to evaluation of bond interfaces, [13], and detection of partially closed defects in materials, [14–16].

For elastic waves with wavelength much larger than the typical length-scale of surface topography, it is known, [17], that such contact interface can be treated as a spring interface with the equivalent stiffness property. In this case, theoretical expressions for the reflection and transmission coefficients have been derived for normally incident elastic waves in terms of the interface stiffness and other acoustic parameters, [1,2].

Importantly, the interface stiffness introduced above depends on the contact condition due to the aforementioned effects, and becomes a source of nonlinearity in wave propagation. Theoretically, Richardson [18] has analyzed one-dimensional nonlinear wave propagation through a unilateral contact interface, where either the displacement gap or the traction vanishes. In Richardson's analysis, however, realistic features of finite and nonlinear interface stiffness, in the sense that it may vary continuously with the contact condition, are not accounted for. Such effects were considered by Rudenko and Chin An Vu [19], who used a distributed-microasperity model to analyze nonlinear wave reflection from rough-surface contact, and interpreted existing experimental data. Nevertheless, it appears that implications of such nonlinear stiffness properties for the acoustic response of the contact interface are yet to be fully explored in a general circumstance.

The aim of the present paper is to demonstrate a concise theoretical analysis to yield a basic insight into the nonlinear behavior of elastic waves propagating through contact interface, as a basis of ultrasonic evaluation of contact conditions. To this end, a nonlinear contact stiffness model is incorporated into a one-dimensional framework as analyzed by Richardson [18], and the transmission/reflection coefficients as well as the second harmonic amplitudes are derived in terms of the interface stiffness property and other acoustic parameters. Furthermore, it is shown that the nonlinear gap-pressure relationship can be extracted from the contact-pressure dependence of the linear interface stiffness. This is demonstrated using the pressure-stiffness relation for aluminum blocks recently reported by Drinkwater et al. [3]. The connection between the contact stiffness parameters and the transmission/

[1]To whom correspondence should be addressed. Present address: Department of Energy Conversion Science, Graduate School of Energy Science, Kyoto University, Kyoto 606-8501, Japan.

[2]Presently with Toyoda Automatic Loom Works, Ltd.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, December 11, 2002; final revision, January 24, 2004. Associate Editor: K. T. Ramesh. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

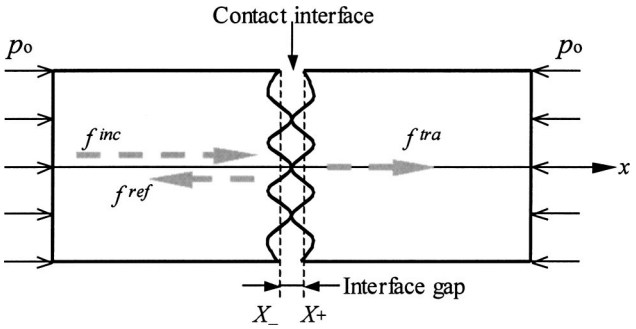


Fig. 1 One-dimensional wave propagation through contact interface.

reflection as well as the harmonic generation behavior is discussed in detail. The analysis is shown to reproduce certain qualitative features of nonlinear acoustic response of contact interface observed experimentally in previous studies, e.g., qualitative contact-pressure dependence of the second-harmonic amplitudes, and the appearance of "DC" displacement components in transmitted/reflected waves.

2 Theoretical Formulation

2.1 Governing Equations. One-dimensional elastic longitudinal wave propagation along the x-axis is considered here as schematically shown in Fig. 1. Two linear elastic solids with identical properties (mass density ρ and the longitudinal stiffness E) are assumed to occupy the regions $x<X_-$ and $x>X_+$, respectively, where $X_-<X_+$. The mated surfaces of these solids are nominally flat but assumed to have certain roughness. The reference planes of both surfaces, e.g., planes of average height, [20], are identified by $x=X_-$ and $x=X_+$, respectively. The quantity X_+-X_- denotes the gap distance between the reference planes. The formulation described in this section essentially follows a line similar to Richardson [18], except that the present analysis accounts for a realistic relation between the gap distance and the contact pressure. For the sake of completeness, the formulation will be recapitulated below including a straightforward outcome of Richardson's analysis.

In the absence of elastic waves, the two solids are assumed to be at equilibrium under nominal contact pressure p_0 with the equilibrium gap distance h_0. For longitudinal wave propagation into the positive x-direction, the equation of motion as well as the stress-displacement relation are

$$\rho\frac{\partial^2 u}{\partial t^2}=\frac{\partial \sigma}{\partial x},\quad \sigma+p_0=E\frac{\partial u}{\partial x}, \tag{1}$$

where $u(x,t)$ is the displacement in the x-direction accompanying the wave motion, $\sigma(x,t)$ is the stress, and t denotes time. The linear kinematics is assumed throughout the analysis. As a solution to Eq. (1), the following forms are considered:

$$u(x,t)=f^{inc}(x-ct)+f^{ref}(x+ct),\quad x<X_-, \tag{2a}$$

$$u(x,t)=f^{tra}(x-ct),\quad x>X_+, \tag{2b}$$

where $f^{inc}(x-ct)$, $f^{ref}(x+ct)$, $f^{tra}(x-ct)$ represent the incident, reflected, and transmitted waves, respectively, and $c=(E/\rho)^{1/2}$ is the wave velocity.

When the wave interacts with the contact interface, the gap distance varies with time, i.e.,

$$h(t)=h_0+u(X_+,t)-u(X_-,t). \tag{3}$$

The boundary condition at the interface is given as

$$\sigma(X_-,t)=\sigma(X_+,t)=-p(h(t)),\quad p(h_0)=p_0. \tag{4}$$

Namely, the stress is continuous across the interface and obeys a functional relationship between the gap distance h and the contact pressure $p(h)$, which is in general a nonlinear one.

2.2 Reduced Formulation. By introducing the following change of variables, [18],

$$X(t)=\{u(X_-,t)+u(X_+,t)\}/2, \tag{5a}$$

$$Y(t)=u(X_+,t)-u(X_-,t)=h(t)-h_0, \tag{5b}$$

the governing equations reduce to

$$\dot{X}=-cf^{inc\prime}(X_--ct), \tag{6a}$$

$$\dot{Y}=2cf^{inc\prime}(X_--ct)+\frac{2}{\rho c}\{p(h_0+Y)-p_0\}, \tag{6b}$$

where dots denote time derivatives and primes denote differentiation with respect to the argument of functions. The variables X and Y represent the translational motion of the contact interface and the time-dependent change of the gap distance, respectively. The transmitted and reflected waves are in turn expressed as

$$f^{tra}(x-ct)=f^{inc}(x-h_0-ct)+\frac{1}{2}Y(t-(x-X_+)/c), \tag{7a}$$

$$f^{ref}(x+ct)=-\frac{1}{2}Y(t+(x-X_-)/c). \tag{7b}$$

Therefore, the problem is reduced to obtaining the solution to the first-order nonlinear ordinary differential equation in Eq. (6*b*).

For a demonstrative purpose, Fig. 2 shows examples of transmitted and reflected waveforms for an incident wave of Gaussian-modulated profile with the center frequency 10 MHz, obtained by direct numerical analysis of Eqs. (6) and (7), with $\rho=2700$ [kg/m^3] and $c=6420$ [m/s] pertinent to aluminum and by assuming $X_+\cong X_-$. The nonlinear function $p(h)$ is characterized by the parameters as $p_0=10$ [MPa], $m=0.5$ and $C=6.0\times10^{10}$ [Pa$^{-1/2}$ m^{-1}], which definitions will be clarified in later discussions. Figure 2(*a*) is for the incident wave packet with peak amplitude of 5 nm, and Figs. 2(*b*) and (*c*) are the resulting transmitted and reflected waveforms, respectively. Figures 2(*d*)–(*f*) show the corresponding results in the case of peak amplitude 20 nm, which is a typical level in recent nonlinear ultrasonic measurements, [10,16]. In the results in Figs. 2(*e*) and (*f*), it can be observed that the transmitted and reflected waveforms are somewhat distorted due to the nonlinear nature of the interface, while the distortions are more or less negligible for the incident amplitude of 5 nm as in Figs. 2(*b*) and (*c*). Furthermore, in the spectral amplitudes of these waveforms shown in Fig. 3, the frequency components around 0 MHz as well as the 20 MHz (second harmonics) have appeared (higher harmonics are also present but fall below the plot range). As discussed below, the harmonic generation and the appearance of the zero-frequency component are more evident in the case of larger incident amplitude.

3 Analysis for Harmonic Wave Incidence

3.1 A Perturbation Analysis. In order to propel the analysis further, the incident wave is now assumed in the form of a monochromatic wave of the form

$$f^{inc}(x-ct)=A\cos\left\{\frac{\omega}{c}(x-X_--ct)\right\}, \tag{8}$$

where A is the wave amplitude and ω is the angular frequency. When the displacement accompanying the wave is small, the gap distance changes only by a small amount. In this case the function $p(h)$ can be replaced by its Taylor expansion near $h=h_0$ up to the second-order term, i.e.,

$$p(h)=p(h_0+Y)=p_0-K_1Y+K_2Y^2, \tag{9}$$

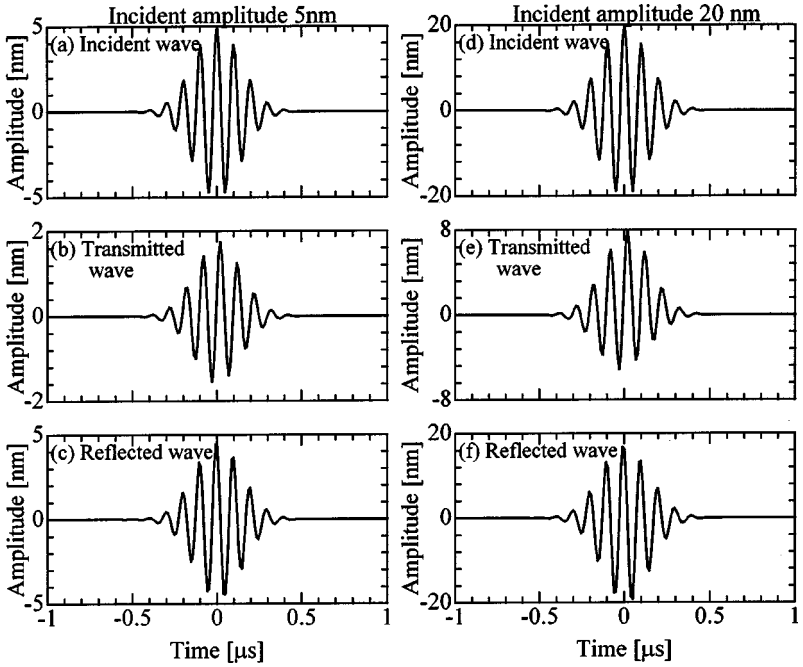


Fig. 2 Gaussian-modulated incident wave with center frequency 10 MHz, and the transmitted and reflected waves at the aluminum-aluminum contact interface characterized by the parameters $p_0=10$ [MPa], $m=0.5$, and $C=6.0\times10^{10}$ [$\mathrm{Pa}^{-1/2}\,\mathrm{m}^{-1}$]. (*a*), (*b*), (*c*): Incident wave amplitude 5 nm; (*d*), (*e*), (*f*): incident wave amplitude 20 nm. The time scales are taken with the origin at the center of each wave packet.

$$K_1=-\left.\frac{dp}{dh}\right|_{h=h_0},\quad K_2=\left.\frac{1}{2}\frac{d^2p}{dh^2}\right|_{h=h_0}. \tag{10}$$

In the above expression, K_1 denotes the linear stiffness, and K_2 the second-order stiffness of the contact interface. Substitution of Eqs. (9) and (10) in Eq. (6*b*) yields

$$\dot{Y}+\frac{2K_1}{\rho c}Y-\frac{2K_2}{\rho c}Y^2=2A\omega\sin\omega t. \tag{11}$$

To obtain an explicit approximate solution to Eq. (11), Y is expressed as the sum of a solution of the linearized equation Y_1 and the perturbation Y_2, i.e., $Y=Y_1+Y_2$. Then Y_1 satisfies the linearized form of Eq. (11),

$$\dot{Y}_1+\frac{2K_1}{\rho c}Y_1=2A\omega\sin\omega t, \tag{12}$$

which solution corresponding to steady-state oscillation is readily obtained as

$$Y_1(t)=\frac{2A}{\sqrt{1+a^2/\omega^2}}\sin(\omega t-\delta_1), \tag{13}$$

where

$$a=\frac{2K_1}{\rho c},\quad \delta_1=\arctan(\omega/a). \tag{14}$$

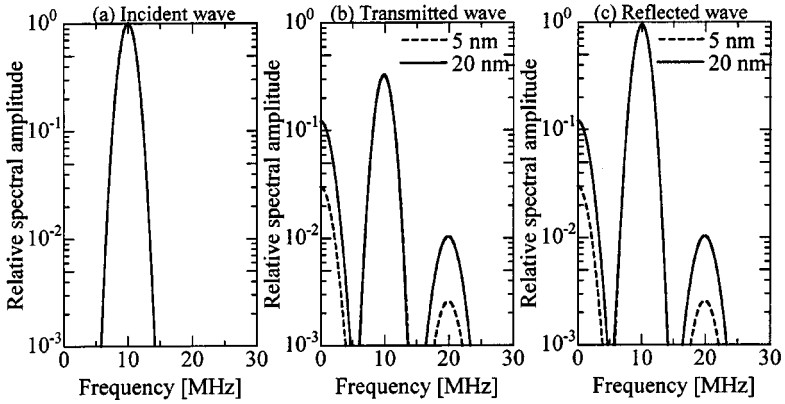


Fig. 3 Amplitude spectra of the waves shown in Fig. 2. The spectral amplitudes are normalized with respect to the amplitude of the incident wave at the fundamental frequency (10 MHz).

Then, assuming $Y_2 \ll Y_1$, the solution for Y_2 can be obtained, again ignoring the transient component, as

$$Y_2(t) = \frac{2K_2A^2}{K_1(1+a^2/\omega^2)}\left\{1 - \frac{\sin(2\omega t - 2\delta_1 + \delta_2)}{\sqrt{1+4\omega^2/a^2}}\right\}, \quad (15)$$

where

$$\delta_2 = \arctan(a/(2\omega)). \quad (16)$$

Substituting $Y = Y_1 + Y_2$ and Eqs. (13) and (15) in Eqs. (7*a*) and (7*b*), the transmitted and reflected waves can be obtained as

$$f^{tra}(x-ct) = \frac{K_2A^2}{K_1\{1+4K_1^2/(\rho^2c^2\omega^2)\}} + \frac{2K_1A}{\rho c\omega\sqrt{1+4K_1^2/(\rho^2c^2\omega^2)}} \times \cos\left\{\omega t - \frac{\omega}{c}(x-X_+) - \delta_1\right\} - \frac{K_2A^2}{\rho c\omega\{1+4K_1^2/(\rho^2c^2\omega^2)\}\sqrt{1+K_1^2/(\rho^2c^2\omega^2)}} \times \sin\left\{2\omega t - \frac{2\omega}{c}(x-X_+) - 2\delta_1 + \delta_2\right\}. \quad (17a)$$

$$f^{ref}(x+ct) = -\frac{K_2A^2}{K_1\{1+4K_1^2/(\rho^2c^2\omega^2)\}} - \frac{A}{\sqrt{1+4K_1^2/(\rho^2c^2\omega^2)}}\sin\left\{\omega t + \frac{\omega}{c}(x-X_-) - \delta_1\right\} + \frac{K_2A^2}{\rho c\omega\{1+4K_1^2/(\rho^2c^2\omega^2)\}\sqrt{1+K_1^2/(\rho^2c^2\omega^2)}} \times \sin\left\{2\omega t + \frac{2\omega}{c}(x-X_-) - 2\delta_1 + \delta_2\right\}. \quad (17b)$$

It is seen clearly that both transmitted and reflected waves contain a term with the angular frequency 2ω, which is the second harmonic generated by the nonlinear nature of the contact stiffness. Further perturbation analysis would yield higher-order harmonics, although it is kept out of the present discussion.

Furthermore, static displacement components are contained in both waves. These zero-frequency components are similar in nature to the "DC" signal observed experimentally by Korshak et al. [21] recently for polished glass-piezoceramic interface subjected to surface wave incidence. The appearance of these components has been derived naturally as the outcome of the present analysis. When $K_2/K_1 > 0$, the DC component brings about the displacement corresponding to gap opening, which is proportional to the second-order stiffness K_2 and the square of the input wave amplitude.

3.2 Reflection/Transmission Coefficients and Nonlinearity Parameters. When the amplitude transmission coefficient T and the amplitude reflection coefficient R are defined as the ratio between the absolute amplitude of the fundamental-frequency components of the transmitted/reflected waves and the incident wave, they read

$$T = \frac{2K_1}{\rho c\omega\sqrt{1+\frac{4K_1^2}{\rho^2c^2\omega^2}}}, \quad R = \frac{1}{\sqrt{1+\frac{4K_1^2}{\rho^2c^2\omega^2}}}. \quad (18)$$

These results coincide with those derived previously from the linear analysis, [1,2].

Next, a parameter of nonlinearity β_1 is defined as the ratio between the absolute amplitudes of the second harmonic (2ω) and the fundamental (ω) component in the transmitted wave, i.e.,

$$\beta_1 = \frac{K_2A}{2K_1\sqrt{1+\frac{4K_1^2}{\rho^2c^2\omega^2}}\sqrt{1+\frac{K_1^2}{\rho^2c^2\omega^2}}}. \quad (19)$$

Likewise, for the reflected wave, a similar parameter of nonlinearity γ_1 is defined as

$$\gamma_1 = \frac{K_2A}{\rho c\omega\sqrt{1+\frac{4K_1^2}{\rho^2c^2\omega^2}}\sqrt{1+\frac{K_1^2}{\rho^2c^2\omega^2}}}, \quad (20)$$

which is the ratio between the absolute amplitudes of the second harmonic and the fundamental component in the reflected wave. In contrast to the transmission and reflection coefficients, the above nonlinearity parameters are proportional to the second-order stiffness of the interface. These parameters are also proportional to the incident wave amplitude, so the nonlinearity is higher for waves of larger amplitude. Since the contact stiffness parameters K_1 and K_2 vary with the applied contact pressure p_0, the transmission/reflection coefficients and nonlinearity parameters are dependent on p_0. These features are demonstrated below for a particular nonlinear model of the contact interface.

4 Power-Law Pressure Dependence of Contact Stiffness

4.1 Derivation of Pressure-Gap Relation. The functional relationship between the gap distance and the contact pressure, written as $p(h)$, can be modeled analytically based on various statistical models of rough surfaces. Rudenko and Chin An Vu [19], and recently Drinkwater et al. [3] and Baltazar et al. [5] attempted to link the roughness topography to the contact stiffness or the transmission/reflection coefficients measured ultrasonically. However, it is in general difficult to take into account all detailed information of roughness features as well as deformation properties of real surfaces. As an alternative approach, in this section it is shown how the desired relation can be extracted from the relation between the linear stiffness of the contact and the applied contact pressure, which can be easily obtained experimentally using ultrasonic methods.

In their detailed investigation, Drinkwater et al. [3] measured the reflection coefficients of the normally incident ultrasonic wave at the contact interface between two aluminum blocks, which were made rough by grit blasting and subjected to contact loading/unloading cycles. By using the relation in Eq. (18), they connected the reflection coefficients to the linear contact stiffness K_1 as a function of the applied contact pressure p_0. According to their results, K_1 showed more or less linear dependence on p_0, i.e., $K_1 \propto p_0$, at the first compressive loading up to about 250 MPa. The relation between K_1 and p_0 showed some hysteresis for the first loading/unloading cycle. For the unloading after the first loading, and for the subsequent loading/unloading cycles, K_1 varied nearly in proportion to the square root of p_0, i.e., $K_1 \propto p_0^{0.5}$, without significant hysteresis. The linear stiffness-pressure relationship shown in the first loading is attributed to the successive flattening of contact asperities on the rough interface, and the following square-root behavior is attributed to the response of the flattened (partially conformed) surfaces.

In our discussion, the above experimental findings are generalized and it is assumed that a simple power-law relation holds between the linear stiffness and the applied pressure, that is

$$K_1 = Cp_0^m, \quad (21)$$

where C and m are positive constants. Hereafter, this particular relation is employed to model the contact interface stiffness. Combined with Eq. (10), Eq. (21) can be regarded as a differential equation for $p(h)$,

$$\frac{dp}{dh} = -Cp^m, \tag{22}$$

which can be solved with Eq. (4) as

$$p(h) = \begin{cases} \{p_0^{1-m} - (1-m)C(h-h_0)\}^{1/(1-m)} & (m \neq 1), \\ p_0 \exp\{-C(h-h_0)\} & (m=1). \end{cases} \tag{23}$$

Hence, the relation between the contact pressure and the gap distance can be extracted from the observed linear stiffness-pressure relation. The same procedure may be applied to the situation where the linear contact stiffness K_1 is expressed as an arbitrary function of the contact pressure p_0.

When both of the results in Eq. (23) are expanded near $h = h_0$, one obtains

$$p = p_0 - Cp_0^m(h-h_0) + \frac{1}{2}mC^2p_0^{2m-1}(h-h_0)^2, \tag{24}$$

up to second-order terms, which gives the linear and second-order stiffness as functions of the equilibrium contact pressure, i.e.,

$$K_1(p_0) = Cp_0^m, \quad K_2(p_0) = \frac{1}{2}mC^2p_0^{2m-1}. \tag{25}$$

4.2 Contact Mechanics Considerations. Prior to proceeding to the discussion of nonlinear acoustic response of contact interface, some remarks are noted for the power-law dependence in Eq. (21) from a viewpoint of contact mechanics of rough surfaces. A great number of models have been proposed to model the contact behavior of real surfaces having various roughness features. Among them, for noninteracting elastic hemispherical asperities with constant curvature and an exponential height distribution, Greenwood and Williamson [22] showed that the relation between the nominal pressure and the gap distance takes on an exponential form. Recently, Larsson et al. [23] obtained a similar conclusion for exponential distributions of asperity height and constant asperity curvatures but for nonlinear stress-strain behavior of power-law type. Such exponential pressure-gap relation coincides with the case of $m=1$ in Eq. (23). For Gaussian asperity height distributions, similar results have been obtained, [22,23].

On the other hand, when all hemispherical asperities have the same height and curvature, the nominal pressure follows $p(h) \propto (\bar{h}-h)^{1+M/2}$ for the materials obeying the equivalent stress-strain relation $\sigma \propto \varepsilon^M$ (M: positive material constant), when the gap distance at which the asperities first come into contact is denoted by $\bar{h}$, [23]. As a special case, for linear elastic asperities, $p \propto (\bar{h}-h)^{3/2}$ and $K_1 \propto p_0^{1/3}$, i.e., $m=1/3$. According to the similarity principle of contact, [24,25], the same dependence can be expected not only for hemispherical asperities but also for asperities with local shapes described by homogeneous functions.

From these considerations, it may be conjectured that rough surfaces with statistical asperity-height distributions tend to have m values close to 1, while m values are lower for surfaces with more uniform asperity heights. Coming back to the experimental finding by Drinkwater et al. [3], it may now be reasonably interpreted that the grit-blasted aluminum surfaces revealed the exponent $m=1$ for the first loading, as they ought to have had some degrees of roughness: after the compression, many asperities were likely conformed to similar heights due to plastic deformation, and resulted in lower exponent of $m=0.5$.

5 Implications of Power-Law Pressure-Dependent Stiffness

5.1 Qualitative Pressure Dependence of Contact Acoustic Properties. Substituting Eq. (25) into Eqs. (18)–(20), the transmission/reflection coefficients as well as the nonlinearity parameters are readily obtained as function of the applied pressure. Explicitly,

$$T = \frac{2Cp_0^m/(\rho c\omega)}{\sqrt{1+4C^2p_0^{2m}/(\rho c\omega)^2}}, \quad R = \frac{1}{\sqrt{1+4C^2p_0^{2m}/(\rho c\omega)^2}}, \tag{26}$$

$$\beta_1 = \frac{mCp_0^{m-1}A}{4\sqrt{1+4C^2p_0^{2m}/(\rho c\omega)^2}\sqrt{1+C^2p_0^{2m}/(\rho c\omega)^2}}, \tag{27a}$$

$$\gamma_1 = \frac{mC^2p_0^{2m-1}A}{2\rho c\omega\sqrt{1+4C^2p_0^{2m}/(\rho c\omega)^2}\sqrt{1+C^2p_0^{2m}/(\rho c\omega)^2}}. \tag{27b}$$

From these results it is found that the transmission (reflection) coefficient is a monotonically increasing (decreasing) function of the contact pressure.

It is also seen in Eqs. (27*a*) and (27*b*) that the nonlinearity parameters for the transmitted and reflected waves are decreasing with the contact pressure, when the pressure becomes sufficiently large. However, the qualitative contact-pressure dependence of the nonlinearity parameters in a low pressure range depends on the power exponent m. Namely, when $Cp_0^m/(\rho c\omega) \to 0$, three types of behaviors are observed, i.e.,

$$\beta_1 \approx \frac{mCA}{4}p_0^{m-1} \to \begin{cases} 0 & (m>1) \\ CA/4 & (m=1), \\ \infty & (m<1) \end{cases} \tag{28a}$$

$$\gamma_1 \approx \frac{mC^2A}{2\rho c\omega}p_0^{2m-1} \to \begin{cases} 0 & (m>1/2) \\ C^2A/(4\rho c\omega) & (m=1/2). \\ \infty & (m<1/2) \end{cases} \tag{28b}$$

There have been several experimental investigations for second harmonic amplitudes generated at contact interface. Among others, the second harmonics in the reflected wave measured by Severin and Solodov [6] and Ko Sel Len et al. [7] for polished glass interfaces show a peak value at a certain pressure (0.2–1.5 MPa) and tends to zero as the pressure becomes smaller, which is qualitatively akin to the nonmonotonic behavior with $m>1/2$. Recent results by Kawashima et al. [10] for the transmitted wave through a buff-polished aluminum-aluminum interface show monotonically decreasing pressure dependence for an extremely low-pressure range below 0.03 MPa, which corresponds to the feature for $m<1$. These qualitative features do not contradict the discussion noted in 4.2, where a plausible range for m was given as 1/3 to 1 from simple contact mechanics arguments. At the present, there appear to be no detailed experimental reports available on correlated quantitative data linking the stiffness-pressure relation and the harmonic amplitude-pressure relation. It remains as an intriguing problem if this theoretical observation can be supported from an experimental point of view.

5.2 Results for Particular Model Parameters. Drinkwater et al. [3] used two aluminum blocks to measure the linear interface stiffness as function of the frequency and the contact pressure. As mentioned above, a more or less linear pressure-dependence of the stiffness was observed in the first loading of rough aluminum surfaces. For the subsequent unloading and the following cycles, the overall stiffness-pressure relation was best represented using a square-root relation with $m=0.5$ in Eq. (22). To fit to the data for the latter behavior, the parameters were chosen as $m=0.5$ and $C=6.0\times10^{10}$ [$\mathrm{Pa}^{-1/2}\,\mathrm{m}^{-1}$]. Figure 4(*a*) shows the relation between K_1 and p_0 thus determined, together with the experimental data by Drinkwater et al. [3], while Fig. 4(*b*) illustrates the corresponding function $p(h)$ in Eq. (23).

To simulate the measurements by Drinkwater et al. [3], the acoustic properties of aluminum are employed, i.e., $c = 6420$ [m/s] and $\rho = 2700$ [$\mathrm{kg/m^3}$]. These were the parameters used to generate the results in Fig. 2 and Fig. 3. The transmission and reflection coefficients are calculated for several fundamental

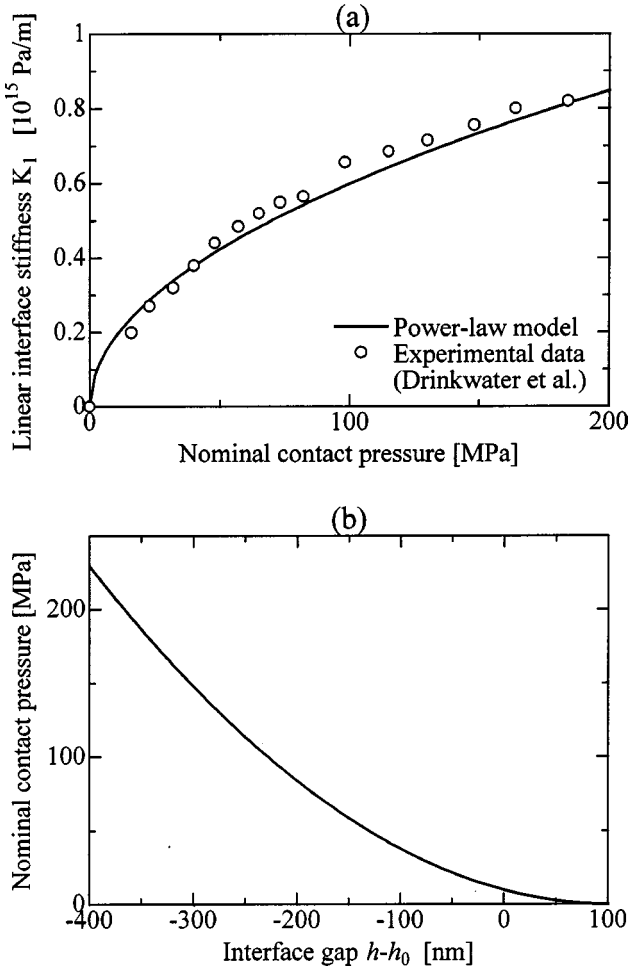


Fig. 4 (*a*) The power-law relation between the linear interface stiffness and the nominal contact pressure, for $m=0.5$ and $C=6\times10^{10}$ [$Pa^{-1/2}\,m^{-1}$], with the experimental data by Drinkwater et al. [3] for Al-Al interface at unloading from initial loading (circles); (*b*) the corresponding pressure-gap relation derived thereof, where h_0 is taken as the gap at 10 MPa

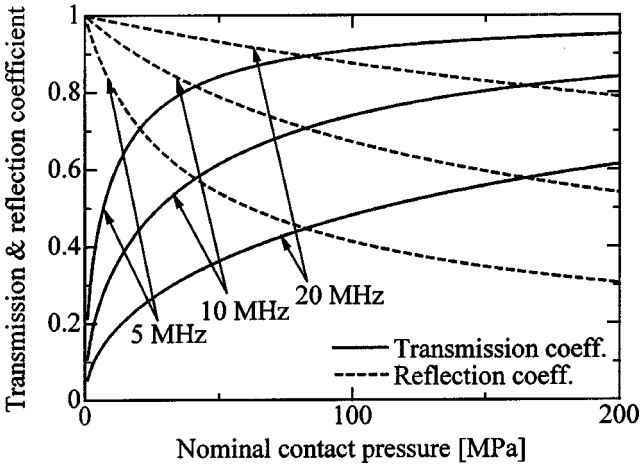


Fig. 5 Variation of the transmission and reflection coefficients with the nominal contact pressure, for different fundamental frequencies

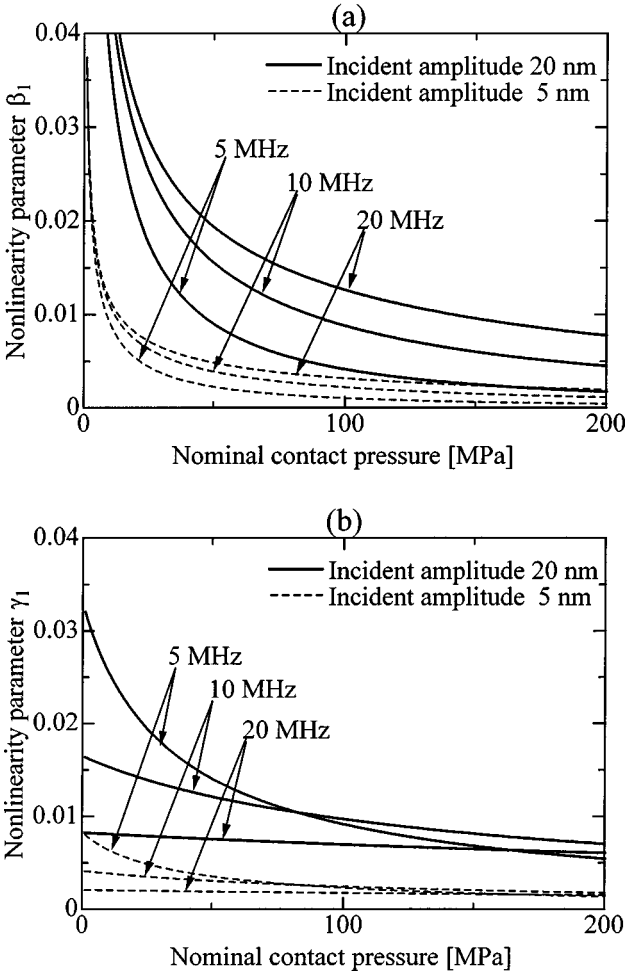


Fig. 6 Variation of (*a*) nonlinearity parameter β_1 and (*b*) nonlinearity parameter γ_1 with the nominal contact pressure, for different fundamental frequencies

frequencies as function of the applied nominal pressure in Fig. 5. As the contact pressure increases from zero, the transmission coefficient increases from zero and the reflection coefficient decreases from unity. For these linear-acoustic characteristics, readers are referred to the experimental results by Drinkwater et al. [3], Dwyer-Joyce and Drinkwater [4], and the discussions therein.

The dependence of the nonlinearity parameters on the nominal contact pressure is shown in Fig. 6 for different fundamental frequencies and two different amplitudes. As mentioned before, the nonlinearity is more profound in the case of higher amplitude. Also, the nonlinearity parameters decrease with the applied pressure, as has been clarified above for $m=1/2$. The influence of the frequency on the pressure-sensitivity of the nonlinearity parameters is, however, different for β_1 and γ_1. In the transmitted wave, β_1 is larger for higher frequency, while in the reflected wave, γ_1 is more profoundly dependent on the nominal pressure in the case of lower frequency. As an explanation for this, it is noted that the transmitted and reflected waves contain the same second-harmonic amplitude for any frequency, while their fundamental-component amplitudes are separately dependent on the frequency. This feature indicates the importance of selecting an appropriate frequency level to carry out a sensitive nonlinear ultrasonic measurement for contact conditions.

In the context of finite-amplitude ultrasonics, [11], an alternative measure of the nonlinearity is often used that is expressed as the ratio between the harmonic amplitude and the square of the

fundamental-frequency amplitude. If such parameters are denoted by β_2 and γ_2 for transmitted and reflected waves, respectively, they read

$$\beta_2 = \frac{\rho c \omega K_2}{4K_1^2 \sqrt{1+\dfrac{K_1^2}{\rho^2 c^2 \omega^2}}} = \frac{mC\rho c\omega}{2p_0\sqrt{1+C^2 p_0^{2m}/(\rho c\omega)^2}}, \tag{29a}$$

$$\gamma_2 = \frac{K_2}{\rho c \omega \sqrt{1+\dfrac{K_1^2}{\rho^2 c^2 \omega^2}}} = \frac{mC^2 p_0^{2m-1}}{\rho c\omega\sqrt{1+C^2 p_0^{2m}/(\rho c\omega)^2}}. \tag{29b}$$

These nonlinearity parameters have a feature that they are independent of the amplitude of the incident wave. For the power-law pressure stiffness relation, the asymptotic behavior for these parameters follows as

$$\beta_2 \approx \frac{mC\rho c\omega}{2}\frac{1}{p_0}, \quad \gamma_2 \approx \frac{mC^2}{\rho c\omega} p_0^{2m-1}, \tag{30}$$

in the limit of $Cp_0^m/(\rho c\omega) \to 0$. It is noteworthy that β_2 scales as the inverse of p_0 irrespective of the value of the exponent m, while the asymptotic dependence of γ_2 on p_0 is qualitatively the same as that of γ_1 shown in Eq. (28*b*).

Buck et al. [9] measured the harmonic generation at the contact interface between two aluminum blocks, and recorded β_2 as a function of the applied pressure when the fundamental frequency was 5 MHz. The experimental conditions of Buck et al. [9] were somewhat similar to those of Drinkwater et al. [3], in that aluminum blocks were examined for a similar range of applied contact pressure, although the surface roughness of their specimens are not known to be similar. In spite of this ambiguity, the almost negligible hysteresis for loading and unloading in the experimental data of Buck et al. [9] is indicative of the features of surfaces that are characterized by $m=0.5$ as above. Therefore, it is of some interest to attempt interpretation of the parameter β_2 in their experiment based on the stiffness parameters employed here.

The parameter β_2 has a dimension of the inverse of length. Since no unit for the β_2 values is recorded by Buck et al. [9], their data and the present theoretical predictions are both normalized by their values at the contact pressure of 10 MPa, say, to compare their qualitative features of the pressure dependence. In the theoretical formula in Eq. (29*a*), the fundamental frequency is set as 5 MHz to simulate the measurement. Figure 7 shows the measured and theoretical nonlinearity parameter β_2 as function of the applied pressure. When normalized in the fashion explained above, the theoretical result fits the measured data fairly well. Furthermore, in the log-log plot, the asymptotic relation given in Eq. (30) for $Cp_0^m/(\rho c\omega) \to 0$ is represented as a straight line with slope -1, as closely approached by the theoretical result of Eq. (29*a*). This inverse proportionality of β_2 to p_0, as predicted universally for arbitrary power-law relations, is supported here by the experimental data.

6 Concluding Remarks

According to the nonlinear interface model described in this paper, the linear stiffness of the contact interface is expressed by a power-law function of the contact pressure. This model has yielded the relation between the contact pressure and the interface gap distance, which then gives the second-order stiffness of the interface as a function of the contact pressure. For an incident harmonic wave, the contact-pressure dependence of the fundamental and second harmonic amplitudes of the transmitted as well as reflected waves have been derived in terms of the power-law parameters and other acoustic properties of the solid. Also, the appearance of the zero-frequency component at the interface has been indicated as the outcome of the analysis. The analysis has been applied to interpret some of the existing experimental results, and the qualitative contact-pressure dependence of the harmonic amplitudes (nonlinearity parameters) has been illustrated theoretically. The present analysis has shown that the quantitative second-harmonic measurement can yield the second-order contact stiffness, and may be used as a supplementary parameter to characterize the state of contact.

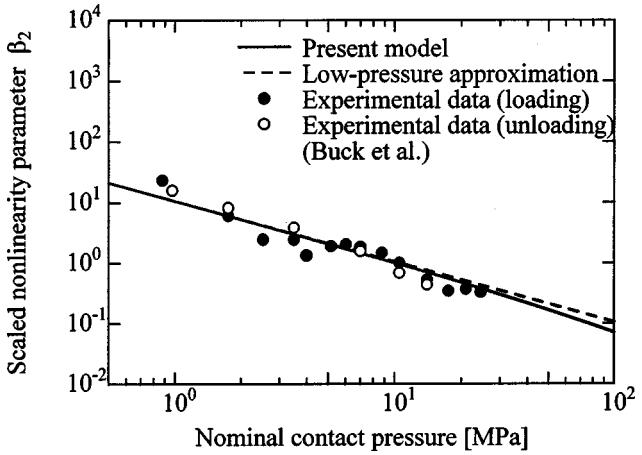


Fig. 7 Variation of the nonlinearity parameter β_2 with the nominal contact pressure, at 5 MHz fundamental frequency. Solid line: present model, broken line: low-pressure asymptotic form, circles: experimental data by Buck et al. [9]. The parameters are scaled with respect to the value at 10 MPa.

Based on some contact-mechanics considerations, certain explanation has been given to the power-law nature of the stiffness-pressure relation of the contact interface. It has been indicated that the power-law exponent may vary depending on the roughness character of the interface. It has been also shown in the present analysis that this power-law exponent has a direct bearing on the pressure-dependence of the second harmonics appearing at the interface. While the nonlinearity parameters defined as the ratio between the second-harmonic and the fundamental amplitudes show various contact-pressure sensitivities depending on the power-law exponent, an analogous parameter defined as the ratio between the transmitted second-harmonic and the squared fundamental amplitudes shows an inverse proportional dependence on the contact pressure for arbitrary power-law behavior. Such implications ought to be studied more closely in a coordinated theoretical and experimental study, and constitutes a subject of continuing interest.

Acknowledgments

The authors are thankful to Prof. K. Kawashima, Dr. I. Yu. Solodov, and Dr. R. S. Dwyer-Joyce for bringing their relevant papers into the authors' attention. The investigation is supported under the Grant-in-Aid for Encouragement of Young Scientists No. 14750057 to one of the authors (S. B.).

References

[1] Kendall, K., and Tabor, D., 1971, "An Ultrasonic Study of the Area of Contact Between Stationary and Sliding Surfaces," Proc. R. Soc. London, Ser. A, **323**, pp. 321–340.
[2] Tattersall, H. G., 1973, "The Ultrasonic Pulse-Echo Technique as Applied to Adhesion Testing," J. Phys. D, **6**, pp. 819–832.
[3] Drinkwater, B. W., Dwyer-Joyce, R. S., and Cawley, P., 1996, "A Study of the Interaction Between Ultrasound and a Partially Contacting Solid-Solid Interface," Proc. R. Soc. London, Ser. A, **452**, pp. 2613–2628.
[4] Dwyer-Joyce, R. S., and Drinkwater, B. W., 1998, "Analysis of Contact Pressure Using Ultrasonic Reflection," *Proc. 11th Int. Conf. Experimental Mechanics*, Balkema, Rotterdam, pp. 747–753.

[5] Baltazar, A., Rokhlin, S. I., and Pecorari, C., 2002, ''On the Relationship Between Ultrasonic and Micromechanical Properties of Contacting Rough Surfaces,'' J. Mech. Phys. Solids, **50**, pp. 1397–1416.
[6] Severin, F. M., and Solodov, I. Y., 1989, ''Experimental Observation of Acoustic Demodulation in Reflection From a Solid-Solid Interface,'' Sov. Phys. Acoust., **35**, pp. 447–448.
[7] Ko, Sel Len, Severin, F. M., and Solodov, I. Y., 1991, ''Experimental Observation of the Influence of Contact Nonlinearity on the Reflection of Bulk Acoustic Waves and the Propagation of Surface Acoustic Waves,'' Sov. Phys. Acoust., **37**, pp. 610–612.
[8] Solodov, I. Y., 1998, ''Ultrasonics of Non-Linear Contacts: Propagation, Reflection and NDE-Applications,'' Ultrasonics, **36**, pp. 383–390.
[9] Buck, O., Morris, W. L., and Richardson, J. M., 1978, ''Acoustic Harmonic Generation at Unbonded Interfaces and Fatigue Cracks,'' Appl. Phys. Lett., **33**, pp. 371–373.
[10] Kawashima, K., Nawa, K., and Hattori, Y., 1999, ''Detection of Higher Harmonics With Large Amplitude Ultrasonics,'' *Proc. 6th Symp. Ultrasonic Testing*, Japan Society of Non-Destructive Testing, Tokyo, pp. 44–45 (in Japanese).
[11] Breazeale, M. A., and Thompson, D. O., 1963, ''Finite-Amplitude Ultrasonic Waves in Aluminum,'' Appl. Phys. Lett., **3**, pp. 77–78.
[12] Thompson, R. B., Buck, O., and Thompson, D. O., 1976, ''Higher Harmonics of Finite Amplitude Ultrasonic Waves in Solids,'' J. Acoust. Soc. Am., **59**, pp. 1087–1094.
[13] Hirsekorn, S., 2001, ''Nonlinear Transfer of Ultrasound by Adhesive Joints—A Theoretical Description,'' Ultrasonics, **39**, pp. 57–68.
[14] Van Den Abeele, K. E. A., Johnson, P. A., and Sutin, A., 2000, ''Nonlinear Elastic Wave Spectroscopy (NEWS) Techniques to Discern Material Damage, Part I: Nonlinear Wave Modulation Spectroscopy (NWMS),'' Res. Nondestruct. Eval., **12**, pp. 17–30.
[15] Donskoy, D., Sutin, A., and Ekimov, A., 2001, ''Nonlinear Acoustic Interaction on Contact Interfaces and Its Use for Nondestructive Testing,'' NDT & E Int., **34**, pp. 231–238.
[16] Okada, J., Ito, T., Kawashima, K., and Nishimura, N., 2001, ''Finite Element Simulation of Nonlinear Acoustic Behavior at Minute Cracks Using Singular Element,'' Jpn. J. Appl. Phys., **40**, pp. 3579–3582.
[17] Achenbach, J. D., 1987, ''Flaw Characterization by Ultrasonic Scattering Methods,'' *Solid Mechanics Research for Quantitative Non-Destructive Evaluation*, J. D. Achenbach and Y. Rajapakse, eds., Martinus Nijhoff, Dordrecht, pp. 67–81.
[18] Richardson, J. M., 1979, ''Harmonic Generation at an Unbonded Interface: I. Planar Interface Between Semi-Infinite Elastic Media,'' Int. J. Eng. Sci., **17**, pp. 73–85.
[19] Rudenko, O. V., and , , Rudenko, O. V., 1994, ''Nonlinear Acoustic Properties of a Rough Surface Contact and Acoustodiagnostics of a Roughness Height Distribution,'' Acoust. Phys., **40**, pp. 593–596.
[20] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge Univ. Press, Cambridge, UK.
[21] Korshak, B. A., Solodov, I. Y., and Ballad, E. M., 2002, ''DC-Effects, Sub-Harmonics, Stochasticity and ''Memory'' for Contact Acoustic Non-Linearity,'' Ultrasonics, **40**, pp. 707–713.
[22] Greenwood, J. A., and Williamson, J. B. P., 1966, ''Contact of Nominally Flat Surfaces,'' Proc. R. Soc. London, Ser. A, **295**, pp. 300–319.
[23] Larsson, J., Biwa, S., and Storåkers, B., 1999, ''Inelastic Flattening of Rough Surfaces,'' Mech. Mater., **31**, pp. 29–41.
[24] Storåkers, B., Biwa, S., and Larsson, P.-L., 1997, ''Similarity Analysis of Inelastic Contact,'' Int. J. Solids Struct., **34**, pp. 3061–3083.
[25] Biwa, S., Ogaki, K., and Shibata, T., 1999, ''Analytical Aspects of Cumulative Superposition Procedure for Elastic Indentation Problems,'' JSME Int. J., Ser. A, **42**, pp. 167–175.

Rayleigh Waves in Anisotropic Crystals Rotating About the Normal to a Symmetry Plane

M. Destrade
Laboratoire de Modélisation en Mécanique,
UMR 7607, CNRS,
Université Pierre et Marie Curie,
4 place Jussieu, Tour 66, case 162,
75252 Paris Cedex 05, France
e-mail: destrade@lmm.jussieu.fr

The propagation of surface acoustic waves in a rotating anisotropic crystal is studied. The crystal is monoclinic and cut along a plane containing the normal to the symmetry plane; this normal is also the axis of rotation. The secular equation is obtained explicitly using the "method of the polarization vector," and it shows that the wave is dispersive and decelerates with increasing rotation rate. The case of orthorhombic symmetry is also treated. The surface wave speed is computed for 12 monoclinic and 8 rhombic crystals, and for a large range of the rotation rate/wave frequency ratio.
[DOI: 10.1115/1.1756140]

1 Introduction

Introduced more than 30 years ago, surface acoustic wave (SAW) devices have been used with great success by the telecommunication industry: nowadays, they are produced in large quantities (several billions per year) and used in wireless transmission and reception technology for color television sets, cell phones, global positioning systems, etc. In recent years, new applications for SAW devices have emerged, namely acoustic sensors, which are passive (no power supply is needed), resistant, almost nonaging, cheap (only one photo-lithographic process is involved in the production), light (less than 1g) and can be operated remotely and wirelessly. For instance [1], SAW identification tags are used for highway toll collection in Norway and for the Munich subway system; SAW temperature sensors can achieve a resolution of 0.02°C from −196°C up to 500°C; wide ranges and fine resolutions are also achieved for pressure, torque, or current sensors, etc. Also, the automotive industry is engaged in the search for an ''intelligent tire'' which could provide direct information on its current state as the car is moving; in this context SAW sensors have been used to measure tire pressure, [2], or friction, [3], as the wheel rotates. In general, SAW devices may be used as angular rate sensors (gyroscopes) to measure frequency shifts due to the rotation, [4–6]. In the present paper, an investigation of the effect of rotation upon the speed of surface (Rayleigh) waves in an anisotropic crystal is presented.

The crystal may possess as little as a single plane of symmetry. It is cut along any plane containing the normal to the symmetry plane and is assumed to rotate at a constant rate about this normal. The surface wave is polarized in the symmetry plane. In other words, it suffices to consider the propagation of a surface wave in the x_1-direction of a monoclinic crystal with symmetry plane at $x_3=0$, cut along the $x_2=0$ plane, and rotating about the x_3-axis (see Fig. 1). The secular equation for rotating materials was obtained by others but in simpler settings: by Clarke and Burdess in an isotropic material, first for small rotation rate/wave frequency ratios, [4], then for any ratio, [5]; by Grigor'evskii, Gulyaev, and Kozlov [7] also for isotropic materials but neglecting the centrifugal force; and by Fang, Yang, and Jiang [6] for crystals having tetragonal symmetry. Here, the analysis is fully developed for crystals with a single symmetry plane, up to the derivation of the secular equation in explicit form, that is an equation giving the Rayleigh wave speed in terms of the elastic parameters and of the rotation rate.

The equation is reached in Sec. 3, after the governing equations have been written down in Sec. 2. The secular equation turns out to be a polynomial of degree 8 for the squared wave speed and also for the squared rotation rate/wave frequency ratio. In the simpler case of orthorhombic symmetry (Sec. 4), the polynomial is of degree 6. The Rayleigh wave speed is computed numerically for 20 specific anisotropic materials (12 monoclinic, 8 orthorhombic) and for a rotation rate/wave frequency ratio varying from 0 to 10. Of course, this range is way beyond the elastic behavior limit, and is irrealistic for pratical purposes where the frequency of a SAW device is typically in the 100 kHz–10 MHz range. It is presented to show that the method of resolution is exact and not approximate, applies for any rate of rotation, and that in contrast with the nonrotating case, the secular equation is dispersive. At small rotation rates, and for certain crystals such as PZT-5, other papers, [6,8], show that the Rayleigh wave speed may at first increase slightly with the rotation frequency/wave frequency ratio. At large ratios, it is seen here that the wave speed decreases with increasing ratios. These variations are crucial to the understanding and correct design of rotating SAW sensors or SAW signal processing devices. A recent article, [9], describes the manufacturing of a 1 cm×1 cm SAW gyroscope and how the rotation rate may be measured using SAW technology. Another example that springs to mind is that of ''spinning missiles,'' [10], for which one might speculate that the communication is ensured via SAW generation and processing of high-frequency signals modified by the rotation. Finally in Sec. 5, the merits of several methods of derivation for the secular equation in non-rotating crystals are discussed. This paper aims to provide a theoretical and analytical framework for the study of surface acoustic waves in rotating crystals.

2 Basic Equations

We consider a half-space $x_2 \geqslant 0$ occupied by a homogeneous anisotropic crystal possessing one plane of symmetry at $x_3=0$, and rotating at a constant angular velocity Ω about the x_3-axis. We study the propagation of a surface (Rayleigh) wave in the x_1-direction, with attenuation in the x_2-direction. In the rotating Cartesian frame $(Ox_1, Ox_2, Ox_3) \equiv (O, \mathbf{i}, \mathbf{j}, \mathbf{k})$, the equations of motion are, [11],

$$\operatorname{div} \boldsymbol{\sigma} = \rho \mathbf{u}_{,tt} + 2\rho\Omega \mathbf{k} \times \mathbf{u}_{,t} + \rho\Omega^2 \mathbf{k} \times (\mathbf{k} \times \mathbf{u}), \qquad (1)$$

where $\boldsymbol{\sigma}$ is the Cauchy stress tensor, ρ is the constant mass density of the material, and the comma denotes differentiation. The sec-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, January 22, 2003; final revision, September 29, 2003. Associate Editor: R. C. Benson. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

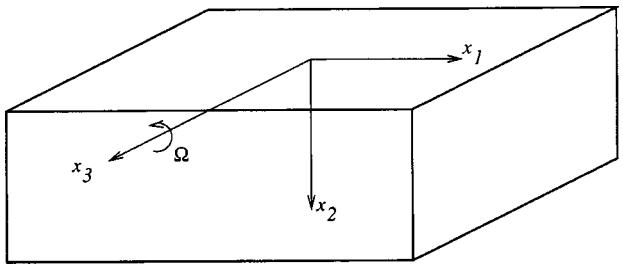


Fig. 1 Monoclinic crystal with symmetry plane at $x_3=0$, cut along $x_2=0$, and rotating about x_3 at constant angular velocity Ω

ond term in the right-hand side of Eq. (1) is due to the Coriolis acceleration, the third is due to the centrifugal acceleration. Note that Eq. (1) represents the time-dependent part of the full equations of motion. The time-independent part, namely $\mathrm{div}\,\boldsymbol{\sigma}^s = \rho\Omega^2\mathbf{k}\times[\mathbf{k}\times(\mathbf{u}^s+\mathbf{x})]$, where $\mathbf{u}^s=\mathbf{u}^s(\mathbf{x})$ and $\sigma^s_{ij}=c_{ijkl}u^s_{l,k}$ must be solved separately. The questions remain of (a) whether an actual time-independent solution exists for all Ω and if it does, of (b) whether the boundary conditions of a traction-free rotating half-space may be satisfied without perturbating the time-dependent boundary value problem. These questions do not seem to have been addressed in the literature, but some preliminary work seem however to suggest that (a) and (b) may be answered positively, at least within the framework of small amplitude waves superimposed upon a large elastic deformation.

Now, turning back to the time-dependent Eq. (1), the mechanical displacement $\mathbf{u}$ is taken in the form

$$\mathbf{u}(x_1,x_2,x_3,t)=\mathbf{U}(kx_2)e^{ik(x_1-vt)}, \tag{2}$$

showing a sinusoidal propagation with speed v and wave number k in the x_1-direction, and the possibility of an attenuation in the x_2-direction through the unknown function $\mathbf{U}(kx_2)$.

We wish to describe the influence of the frame rotation upon the speed of Rayleigh waves, and to this end, we introduce the following quantities:

$$X=\rho v^2, \quad \delta=\Omega/(kv)=\Omega/\omega,$$

where ω is the real frequency of the wave.

For two-dimensional motions ($\partial\mathbf{u}/\partial x_3=0$) such as Eq. (2), the anisotropy of a crystal possessing $x_3=0$ as a symmetry plane is described by the following strain-stress relationship, [12]:

$$\begin{bmatrix}\epsilon_{11}\\ \epsilon_{22}\\ 2\epsilon_{23}\\ 2\epsilon_{31}\\ 2\epsilon_{12}\end{bmatrix}=\begin{bmatrix}s'_{11} & s'_{12} & 0 & 0 & s'_{16}\\ & s'_{22} & 0 & 0 & s'_{26}\\ & & s'_{44} & s'_{45} & 0\\ & & & s'_{55} & 0\\ & & & & s'_{66}\end{bmatrix}\begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{23}\\ \sigma_{31}\\ \sigma_{12}\end{bmatrix},$$

where the strain components ϵ_{ij} are defined in terms of the displacement components by: $2\epsilon_{ij}=u_{i,j}+u_{j,i}$, and the s'_{ij} are the reduced compliances. Alternatively, the equivalent strain-stress relations can be used, ([12], p. 39),

$$\boldsymbol{\sigma}^\circ=\mathbf{C}^\circ\boldsymbol{\epsilon}^\circ, \quad \mathbf{C}^\circ\mathbf{s}'=\mathbf{1}, \tag{3}$$

where $\boldsymbol{\sigma}^\circ=[\sigma_{11},\sigma_{22},\sigma_{23},\sigma_{31},\sigma_{12}]^{\mathrm{T}}$, $\boldsymbol{\epsilon}^\circ=[\epsilon_{11},\ldots,2\epsilon_{12}]^{\mathrm{T}}$. The C°_{ij} are elements, in the Voigt notation, of the fourth-order elastic stiffness tensor C_{ijkl}. Table 1 shows the relevant reduced compliances of 12 different monoclinic crystals, computed from the corresponding stiffnesses as collected by Chadwick and Wilson [13]; the last column gives the corresponding Rayleigh wave speed in the nonrotating case, [14].

In view of the form Eq. (2) for the displacement, we introduce the functions t_1, t_2 for the tractions σ_{12}, σ_{22} on the planes $x_2=$ const. as

$$\sigma_{12}(x_1,x_2,x_3,t)=ikt_1(kx_2)e^{ik(x_1-vt)},$$

$$\sigma_{22}(x_1,x_2,x_3,t)=ikt_2(kx_2)e^{ik(x_1-vt)}.$$

Then, substituting Eqs. (2) and (3) into the equations of motion (1), we derive the following system of linear first-order differential equations for U_1, U_2, t_1, t_2,

$$\begin{bmatrix}\mathbf{U}'\\ \mathbf{t}'\end{bmatrix}=i\begin{bmatrix}\mathbf{N}_1 & \mathbf{N}_2\\ \check{\mathbf{N}}_3+(1+\delta^2)X\mathbf{1} & \mathbf{N}_1\end{bmatrix}\begin{bmatrix}\mathbf{U}\\ \mathbf{t}\end{bmatrix}, \tag{4}$$

where $\mathbf{U}=[U_1,U_2]^{\mathrm{T}}$, $\mathbf{t}=[t_1,t_2]^{\mathrm{T}}$, and the prime denotes differentiation with respect to kx_2. Note that, as in the static case, [15], the antiplane strain (stress) decouples from the plane strain (stress) and need not be considered for this problem. This decoupling would not occur if the crystal was rotating about the x_1-axis or the x_2-axis, [6].

The surface $x_2=0$ is free of tractions and so, the boundary conditions are

$$t_1(0)=t_2(0)=0. \tag{5}$$

In Eq. (4), $\mathbf{N}_1$ and $\mathbf{N}_2$ are the same as the 2×2 submatrices of the 6×6 fundamental elasticity matrix $\mathbf{N}$ from Ingebrigsten and Tonning [16]. Their real matrix $\mathbf{N}_3$, however, has been modified by the introduction of off-diagonal pure imaginary terms. Explicitly, we have

Table 1 Values of the reduced compliances (10^{-12} m^2/N), density (kg/m^3), and (nonrotating) surface wave speed (m/s) for 12 monoclinic crystals

Material	s'_{11}	s'_{22}	s'_{12}	s'_{16}	s'_{26}	s'_{66}	ρ	v_R
diphenyl	854	1858	−366	−698	−1.44	5049	1114	1276
tin fluoride	345	228	−59.2	−197	120	922	4875	1339
tartaric acid	343	211	−164	−223	301	1650	1760	1756
oligoclase	133	227	−108	97.0	−160	483	2638	2413
microcline	94.5	165	−35.1	47.2	1.69	446	2561	2816
gypsum	243	130	−68.6	32.9	28.1	326	2310	3011
hornblende	63.3	103	−32.7	−15.8	−2.72	320	3120	3049
aegirite-augite	53.6	78.4	−21.0	−10.6	−33.5	237	3420	3382
epidote	53.3	49.6	−11.3	17.7	−3.74	237	3400	3409
augite	54.5	64.4	−19.5	−19.0	−15.7	211	3320	3615
diopside	53.1	58.6	−20.1	24.0	6.98	186	3310	3799
diallage	49.8	69.1	−11.3	−6.88	−14.5	166	3300	4000

$$-\mathbf{N}_1=\begin{bmatrix} r_6 & 1 \\ r_2 & 0 \end{bmatrix},\quad \mathbf{N}_2=\begin{bmatrix} n_{66} & n_{26} \\ n_{26} & n_{22} \end{bmatrix},\quad -\check{\mathbf{N}}_3=\begin{bmatrix} \eta & 2i\delta X \\ -2i\delta X & 0 \end{bmatrix},$$

where the quantities r_2, r_6, n_{22}, n_{26}, n_{66}, η are given in terms of the elastic parameters as ([14,17])

$$\eta=\frac{1}{s'_{11}},\quad r_6=-\frac{s'_{16}}{s'_{11}},\quad r_2=-\frac{s'_{12}}{s'_{11}},$$

$$n_{66}=\frac{1}{s'_{11}}\begin{vmatrix} s'_{11} & s'_{16} \\ s'_{16} & s'_{66} \end{vmatrix},\quad n_{22}=\frac{1}{s'_{11}}\begin{vmatrix} s'_{11} & s'_{12} \\ s'_{12} & s'_{22} \end{vmatrix},\quad n_{26}=\frac{1}{s'_{11}}\begin{vmatrix} s'_{11} & s'_{16} \\ s'_{12} & s'_{26} \end{vmatrix}.$$

Thus the rotation of the crystal perturbs the equations of motion in three ways: the introduction of dispersion through δ; a shift of magnitude δ^2 in $X=\rho v^2$ for the lower left submatrix of $\mathbf{N}$ proportional to the 2×2 unit matrix; and the modification of $\mathbf{N}_3$, which is diagonal in the nonrotating case (note that the new matrix $\check{\mathbf{N}}_3$ is Hermitian: $\overline{\check{\mathbf{N}}_3}=\check{\mathbf{N}}_3^{\mathrm{T}}$.) Despite these modifications, the secular equation can be obtained explicitly for the surface wave speed, using a method proposed by Currie [18] and by Taziev [19] for nonrotating anisotropic crystals with and without a plane of symmetry, respectively.

3 Secular Equation

The method of the polarization vector was first presented by Currie [18] to derive the secular equation for Rayleigh waves in the symmetry plane of monoclinic nonrotating crystals. Then Taziev [19] generalized the method to triclinic (no symmetry plane) crystals. This method takes advantage of the Cayley-Hamilton theorem for the fundamental matrix $\mathbf{N}$, which implies that only n matrices $\mathbf{N}^k$ $(k=1,\ldots,n)$ are linearly independent ($n=3$ for monoclinic crystals, $n=5$ for triclinic crystals). Currie used the matrices $\mathbf{N}$, $\mathbf{N}^2$, $\mathbf{N}^3$; Taziev, the matrices $\mathbf{N}$, $\mathbf{N}^2$, $\mathbf{N}^3$, $\mathbf{N}^4$, $\mathbf{N}^5$. Recently, Ting [20] placed their results within the context of the Stroh-Barnett-Lothe formalism and improved on them by showing that the choices of $\mathbf{N}^{-1}$, $\mathbf{N}$, $\mathbf{N}^2$ for monoclinic crystals and of $\mathbf{N}^{-2}$, $\mathbf{N}^{-1}$, $\mathbf{N}$, $\mathbf{N}^2$, $\mathbf{N}^3$ for triclinic crystals lead to simpler and more explicit secular equations. His approach is now adapted to our present context of a rotating crystal with one symmetry plane. An alternative derivation, not based on the Stroh-Barnett-Lothe formalism, is available elsewhere, [21,22].

We seek solutions to the equations of motion Eq. (4) presenting exponential decay with distance

$$\mathbf{U}(kx_2)=\mathbf{a}e^{ikpx_2},\quad \mathbf{t}(kx_2)=\mathbf{b}e^{ikpx_2},\quad \mathcal{I}(p)>0,$$

where the constant vectors $\mathbf{a}$ and $\mathbf{b}$ are related through [12], p. 139: $b_i=(C_{k1i1}+pC_{i2k2})a_k$. Then the equations of motion Eq. (4) give

$$p\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix}=\check{\mathbf{N}}\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix}, \tag{6}$$

where $\check{\mathbf{N}}$ is the 4×4 matrix in Eq. (4). This eigenvalue problem yields a quartic for p. We limit our investigation to the subsonic range, defined as the greatest interval of values for v where the determinant of $\check{\mathbf{N}}-p\mathbf{1}$ possesses two roots p_1, p_2, with positive imaginary parts. We call $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{b}_1$, $\mathbf{b}_2$, the vectors $\mathbf{a}$ and $\mathbf{b}$ corresponding to each root. Then the solution is of the form ([12], p. 141)

$$\mathbf{U}=\mathbf{A}\langle e^{ikp^*}\rangle\mathbf{q},\quad \mathbf{t}=\mathbf{B}\langle e^{ikp^*}\rangle\mathbf{q},\quad \langle e^{ikp^*}\rangle=\mathrm{diag}(e^{ikp_1x_2},e^{ikp_2x_2}),$$

where $\mathbf{A}=[\mathbf{a}_1,\mathbf{a}_2]$, $\mathbf{B}=[\mathbf{b}_1,\mathbf{b}_2]$, and $\mathbf{q}$ is a constant vector. Using the boundary conditions Eq. (5), we have at the free surface $x_2=0$,

$$\mathbf{Bq}=0,\quad \text{and}\quad \mathbf{u}(x_1,0,x_3,t)=\mathbf{a}_{\mathrm{R}}e^{ik(x_1-vt)},\quad \mathbf{a}_{\mathrm{R}}=\mathbf{Aq}. \tag{7}$$

Moreover, the matrices $\mathbf{A}$ and $\mathbf{B}$ satisfy the orthogonality condition, [23],

$$\overline{\mathbf{B}}^{\mathrm{T}}\mathbf{A}+\overline{\mathbf{A}}^{\mathrm{T}}\mathbf{B}=\mathbf{0}. \tag{8}$$

Now, the eigenrelation Eq. (6) may be generalized for any positive or negative integer n as

$$p^n\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix}=\check{\mathbf{N}}^n\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix},\quad \text{where}\quad \check{\mathbf{N}}^n=\begin{bmatrix}\mathbf{N}_1^{(n)} & \mathbf{N}_2^{(n)}\\ \check{\mathbf{K}}^{(n)} & \mathbf{N}_1^{(n)\mathrm{T}}\end{bmatrix}\quad (\text{say}). \tag{9}$$

Explicitly, the elements of $\check{\mathbf{N}}^n$ are computed by multiplication of $\check{\mathbf{N}}$ or its inverse by itself. For instance, $\check{\mathbf{K}}^{(n)}$ for $n=1$, 2, -1 is given by $\check{\mathbf{K}}^{(1)}=\check{\mathbf{N}}_3+X(1+\delta^2)\mathbf{1}$,

$$\check{K}_{11}^{(2)}=-2s'_{16}[1-s'_{11}(1+\delta^2)X]/s'^2_{11},$$

$$\check{K}_{12}^{(2)}=[1-(s'_{11}-s'_{12})(1+\delta^2)X-2is'_{16}\delta X]/s'_{11},=\overline{\check{K}_{21}^{(2)}}, \tag{10}$$

$$\check{K}_{22}^{(2)}=0,$$

and

$$\check{K}_{11}^{(-1)}=-[s'_{22}(1+\delta^2)-(s'_{11}s'_{22}-s'^2_{12})(1-\delta^2)^2X]X/D,$$

$$\check{K}_{12}^{(-1)}=[s'_{26}(1+\delta^2)+2is'_{12}\delta+(s'_{12}s'_{16}-s'_{11}s'_{26})(1-\delta^2)^2X]X/D$$

$$=\overline{\check{K}_{21}^{(-1)}}, \tag{11}$$

$$\check{K}_{22}^{(-1)}=[1-(s'_{11}+s'_{66})(1+\delta^2)X+(s'_{11}s'_{66}-s'^2_{16})(1-\delta^2)^2X^2]/D,$$

where D is a real denominator common to the $K_{ij}^{(-1)}$ whose expression is too long to reproduce and which turns out to be irrelevant for the derivation of the secular equation.

Now we write in turn the second vector line of Eq. $(9)_1$ for p_1 and for p_2, and deduce

$$\check{\mathbf{K}}^{(n)}\mathbf{A}+\check{\mathbf{N}}_1^{(n)}\mathbf{B}=\mathbf{B}\,\mathrm{diag}(p_1,p_2).$$

Multiplying this equality to the left by $\overline{\mathbf{a}_{\mathrm{R}}}^{\mathrm{T}}=\overline{\mathbf{q}}^{\mathrm{T}}\overline{\mathbf{A}}^{\mathrm{T}}$ and to the right by $\mathbf{q}$, and using Eqs. (7), (8), we conclude that (see [20] for the nonrotating case)

$$\overline{\mathbf{a}_{\mathrm{R}}}^{\mathrm{T}}\check{\mathbf{K}}^{(n)}\mathbf{a}_{\mathrm{R}}=0. \tag{12}$$

At $n=-1$, 1, 2, and $\mathbf{a}_{\mathrm{R}}=[1,\alpha]^{\mathrm{T}}$ (say), three equations follow:

$$\check{K}_{12}^{(-1)}\alpha+\overline{\check{K}_{12}^{(-1)}}\overline{\alpha}+\check{K}_{22}^{(-1)}\alpha\overline{\alpha}=-\check{K}_{11}^{(-1)},$$

$$\check{K}_{12}^{(1)}\alpha+\overline{\check{K}_{12}^{(1)}}\overline{\alpha}+\check{K}_{22}^{(1)}\alpha\overline{\alpha}=-\check{K}_{11}^{(1)},$$

$$\check{K}_{12}^{(2)}\alpha+\overline{\check{K}_{12}^{(2)}}\overline{\alpha}=-\check{K}_{11}^{(2)}.$$

We rearrange this system as: $F_{ik}g_k=h_i$, by introducing the following quantities:

$$F_{11}=D\mathcal{R}(\check{K}_{12}^{(-1)}),\quad F_{12}=D\mathcal{I}(\check{K}_{12}^{(-1)}),\quad F_{13}=D\check{K}_{22}^{(-1)},$$

$$F_{21}=0,\quad F_{22}=s'_{11}\mathcal{I}(\check{K}_{12}^{(1)}),\quad F_{23}=s'_{11}\check{K}_{22}^{(1)},$$

$$F_{31}=s'_{11}\mathcal{R}(\check{K}_{12}^{(2)}),\quad F_{32}=s'_{11}\mathcal{I}(\check{K}_{12}^{(2)}),\quad F_{33}=0,$$

$$g_1=\alpha+\overline{\alpha},\quad g_2=i(\alpha-\overline{\alpha}),\quad g_3=\alpha\overline{\alpha},$$

$$h_1=-D\check{K}_{11}^{(-1)},\quad h_2=-s'_{11}\check{K}_{11}^{(1)},\quad h_3=-s'_{11}\check{K}_{11}^{(2)}.$$

Note that the explicit expressions for the nondimensional quantities F_{ik} and h_i in terms of $X=\rho v^2$, $\delta=\Omega/\omega$, and the s'_{ij} are easily read off Eqs. (10), (11). For instance, $F_{12}=2s'_{12}\delta X$, $F_{32}=-2s'_{16}\delta X$, $h_2=1-s'_{11}(1+\delta^2)X$, and so on.

The linear nonhomogeneous system $\mathbf{Fg}=\mathbf{h}$ has a unique solution for $\mathbf{g}$. Introducing $\Delta=\det\mathbf{F}$ and Δ_k $(k=1,2,3)$, the determinant of the matrix obtained from $\mathbf{F}$ by replacing its kth column with $\mathbf{h}$, we write the solution as $g_k=\Delta_k/\Delta$. But the components of $\mathbf{g}$ are related one to another through $g_3=(g_1/2)^2+(g_2/2)^2$. This relation is the *explicit secular equation for Rayleigh waves on an anisotropic crystal rotating in its plane of symmetry*,

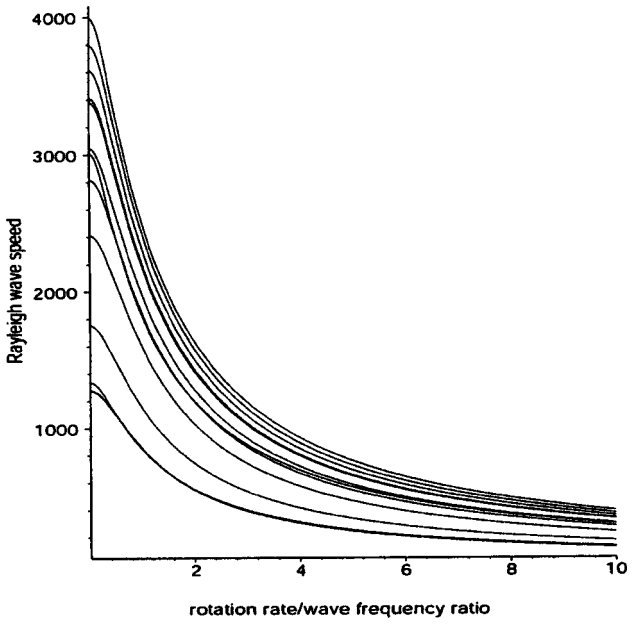


Fig. 2 Rayleigh wave speeds for 12 monoclinic crystals rotating about x_3

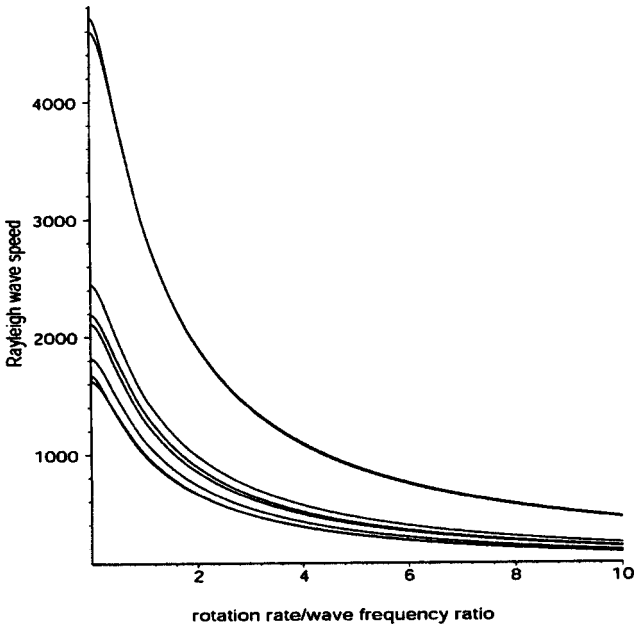


Fig. 3 Rayleigh wave speeds for 8 rhombic crystals rotating about x_3

$$\Delta_1^2+\Delta_2^2-4\Delta_3\Delta=0. \tag{13}$$

This equation is a polynomial of degree 8 in $X=\rho v^2$, and also of degree 8 in δ^2. Because $\delta=\Omega/\omega$ appears only in even powers in the secular equation, the Rayleigh speed obtained as a root of Eq. (13) does not depend on the sense of rotation. Numerically, we find that the rotation slows the Rayleigh wave down and that the speed is a monotone decreasing function of δ. We see this behavior on Fig. 2, where the dependence of the Rayleigh wave speed upon δ is shown for the 12 monoclinic crystals from Table 1. The curves are arranged in the same order as in the table, from the slowest (diphenyl, starting at 1276 m/s) to the fastest (diallage, starting at 4000 m/s).

The secular equation is valid for any crystal possessing at least one plane of symmetry, as long as the half-space is cut along a plane containing the normal to the plane of symmetry. In particular, it is also valid for orthorhombic crystals when the plane of cut contains one of the crystallographic axes. When this plane contains two crystallographic axes, the secular equation factorizes and a separate treatment is required.

4 Orthorhombic Materials

When the material possesses three orthogonal planes of symmetry and the axes $(O,\mathbf{i},\mathbf{j},\mathbf{k})$ are aligned with the crystallographic axes, some compliances vanish: $s'_{16}=s'_{26}=0$. Table 2 lists the values of the relevant reduced compliances for 8 rhombic crystals, computed from the corresponding values of the stiffnesses as collected by Shutilov [24]. The corresponding Rayleigh wave speed v_R in the nonrotating case (last column) is found from the exact secular equation, ([25])

$$(1-Xs'_{11})\sqrt{1-Xs'_{66}}-X\sqrt{s'_{11}[s'_{22}-X(s'_{11}s'_{22}-s'^2_{12})]}=0.$$

When the frame is rotating, the system $\mathbf{Fg}=\mathbf{h}$ reduces to

$$\begin{bmatrix} 0 & F_{12} & F_{13} \\ 0 & F_{22} & F_{23} \\ F_{31} & 0 & 0 \end{bmatrix}\begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix}=\begin{bmatrix} h_1 \\ h_2 \\ 0 \end{bmatrix},$$

where $F_{31}\neq 0$ and

$$F_{12}=2s'_{12}\delta X,$$

$$F_{13}=1-(s'_{11}+s'_{66})(1+\delta^2)X+s'_{11}s'_{66}(1-\delta^2)^2X^2,$$

$$F_{22}=-2s'_{11}\delta X,\quad F_{23}=s'_{11}(1+\delta^2)X,$$

$$h_1=[s'_{22}(1+\delta^2)-(s'_{11}s'_{22}-s'^2_{12})(1-\delta^2)^2X]X,$$

$$h_2=1-s'_{11}(1+\delta^2)X.$$

From this new system of equations, we deduce that $g_2=\hat{\Delta}_2/\hat{\Delta}$ and $g_3=\hat{\Delta}_3/\hat{\Delta}$, where

$$\hat{\Delta}=F_{12}F_{23}-F_{22}F_{13},\quad \hat{\Delta}_2=h_1F_{23}-h_2F_{13},$$

$$\hat{\Delta}_3=F_{12}h_2-F_{22}h_1,$$

and also that $g_1=\alpha+\overline{\alpha}=0$, implying that $g_2=2i\alpha$, $g_3=-\alpha^2=(g_2/2)^2$ as well. This last equality is the *explicit secular equation for Rayleigh waves on an orthorhombic crystal rotating in one plane of symmetry*,

$$\hat{\Delta}_2^2-4\hat{\Delta}_3\hat{\Delta}=0.$$

This equation is a polynomial of degree 6 in $X=\rho v^2$ and in $\delta=\Omega/\omega$. As in the monoclinic case above, the roots are even functions of δ. Numerically, the results are similar to those of the monoclinic case, as Fig. 3 shows for the eight orthorhombic crys-

Table 2 Values of the reduced compliances (10^{-12} m^2/N), density (kg/m^3), and (nonrotating) surface wave speed (m/s) for 8 orthorhombic crystals

Material	s'_{11}	s'_{22}	s'_{12}	s'_{66}	ρ	v_R
sulfur	65.1	76.2	−42.2	132	2070	1628
iodic acid	36.1	20.1	−7.88	57.5	4630	1678
α-uranium	4.89	5.29	−1.13	13.5	19000	1819
rochelle salt	49.3	33.0	−18.2	102	1775	2114
sodium-tartrate	32.1	27.1	−16.8	102	1818	2197
strotium formate	24.5	30.9	−7.32	58.1	2250	2451
olivine	3.26	5.34	−0.97	12.6	3324	4599
benzophenone	13.0	13.9	−7.17	27.9	1219	4723

tals of Table 2. Again, the curves are arranged in the same order as in the table, from the slowest (sulfur, starting at 1628 m/s) to the fastest (benzophenone, starting at 4723 m/s).

5 Concluding Remarks

Several methods have been proposed to derive explicitly the secular equation for surface waves in nonrotating monoclinic crystals with the plane of symmetry at $x_3=0$. This author, [14], wrote the equations of motion as a system of two second-order differential equations for the tractions **t**: $\hat{\alpha}_{ik}t_k''-i\hat{\beta}_{ik}t_k'-\hat{\gamma}_{ik}t_k=0$, where $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$ are 2×2 real symmetric matrices. Then the method of first integrals [26], yields the secular equation. The equations of motion Eq. (4) may also be written in a similar manner for a rotating crystal, but $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\gamma}}$ become complex and the method of first integrals is no longer applicable as such. Next, Ting [17] assumed an exponential form for $\mathbf{t}(kx_2)$ and obtained the secular equation through some simple algebraic manipulations, taking advantage of the fact that $\hat{\alpha}_{12}=\hat{\beta}_{22}=0$; in the rotating case, however, these quantities are no longer zero. Furs [27] (using the displacement field) and this author [25] (using the traction field) devised yet another method, where the secular equation is the resultant of two polynomials; again, having real quantities for the components of **N** is a crucial property, no longer true for rotating crystals.

All in all, it seems that the method of the polarization vector is the most appropriate for the case of a rotating crystal. Note that a simple derivation of its main result (12), not relying on the Stroh formalism, was presented recently, [21,22].

Acknowledgments

I thank T. C. T. Ting, G. Saccomandi, and A. Shuvalov for being so generous with their time and for helping me address some fundamental questions related to this material.

References

[1] Reindl, L., Scholl, G., Ostertag, T., Scherr, H., Wolff, U., and Schmidt, F., 1998, ''Theory and Application of Passive SWA Radio Transponders as Sensors,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control, **45**, pp. 1281–1292.
[2] Pohl, A., Ostermayer, G., Reindl, L., and Seifert, F., 1997, ''Monitoring the Tire Pressure at Cars Using Passive SWA Sensors,'' IEEE Ultrasonics Symposium, **1**, pp. 471–474.
[3] Pohl, A., Steindl, R., and Reindl, L., 1999, ''The 'Intelligent Tire' Utilizing Passive SWA Sensors—Measurement of Tire Friction,'' IEEE Trans. Instrum. Meas., **48**, pp. 1041–1046.
[4] Clarke, N. S., and Burdess, J. S., 1994, ''A Rotation Rate Sensor Based Upon a Rayleigh Resonator,'' ASME J. Appl. Mech., **61**, pp. 139–143.
[5] Clarke, N. S., and Burdess, J. S., 1994, ''Rayleigh Waves on a Rotating Surface,'' ASME J. Appl. Mech., **61**, pp. 724–726.
[6] Fang, H., Yang, J., and Jiang, Q., 2000, ''Rotation Perturbed Surface Acoustic Waves Propagating in Piezoelectric Crystals,'' Int. J. Solids Struct., **37**, pp. 4933–4947.
[7] Grigor'evskiĭ, V. I., Gulyaev, Yu. V., and Kozlov, A. I., 2000, ''Acoustic Waves in a Rotating Elastic Medium,'' Acoust. Phys., **46**, pp. 236–238.
[8] Collet, B., 2003, ''Gyroscopic Effect on Surface Acoustic Waves in Anisotropic Solid Media,'' *Proceedings of the 5th World Congress on Ultrasonics*, pp. 991–995.
[9] Jose, K. A., Suh, W. D., Xavier, P. B., Varadan, V. K., and Varadan, V. V., 2002, ''Surface Acoustic Wave MEMS Gyroscope,'' Wave Motion, **36**, pp. 367–381.
[10] Jahangir, E., and Howe, R. M., 1993, ''Time-Optimal Attitude Control Scheme for a Spinning Missile,'' J. Guid. Control Dyn., **16**, pp. 346–353.
[11] Schoenberg, M., and Censor, D., 1973, ''Elastic Waves in Rotating Media,'' Q. Appl. Math., **31**, pp. 115–125.
[12] Ting, T. C. T., 1996, *Anisotropic Elasticity: Theory and Applications*, Oxford University Press, New York.
[13] Chadwick, P., and Wilson, N. J., 1992, ''The Behavior of Elastic Surface Waves Polarized in a Plane of Material Symmetry, II. Monoclinic Media,'' Proc. R. Soc. London, Ser. A, **438**, pp. 207–223.
[14] Destrade, M., 2001, ''The Explicit Secular Equation for Surface Acoustic Waves in Monoclinic Elastic Crystals,'' J. Acoust. Soc. Am., **109**, pp. 1398–1402.
[15] Stroh, A. N., 1962, ''Steady State Problems in Anisotropic Elasticity,'' J. Math. Phys., **41**, pp. 77–103.
[16] Ingebrigsten, K. A., and Tonning, A., 1969, ''Elastic Surface Waves in Crystal,'' Phys. Rev., **184**, pp. 942–951.
[17] Ting, T. C. T., 2002, ''Explicit Secular Equations for Surface Waves in Monoclinic Materials With the Symmetry Plane at $x_1=0$, $x_2=0$ or $x_3=0$,'' Proc. R. Soc. London, Ser. A, **A458**, pp. 1017–1031.
[18] Currie, P. K., 1979, ''The Secular Equation for Rayleigh Waves on Elastic Crystals,'' Q. J. Mech. Appl. Math., **32**, pp. 163–173.
[19] Taziev, R. M., 1989, ''Dispersion Relation for Acoustic Waves in an Anisotropic Elastic Half-Space,'' Sov. Phys. Acoust., **35**, pp. 535–538.
[20] Ting, T. C. T., 2004, ''The Polarization Vector and Secular Equation for Surface Waves in an Anisotropic Elastic Half-Space,'' Int. J. Solids Struct., **41**, pp. 2065–2083.
[21] Destrade, M., 2004, ''Surface Waves in Rotating Rhombic Crystals,'' Proc. R. Soc. London, Ser. A, **460**, pp. 653–665.
[22] Destrade, M., 2004, ''Explicit Secular Equation for Scholte Waves Over a Monoclinic Crystal,'' J. Sound Vib., to appear.
[23] Barnett, D. M., and Lothe, J., 1973, ''Synthesis of the Sextic and the Integral Formalism for Dislocations, Green's Function, and Surface Wave (Rayleigh Wave) Solutions in Anisotropic Elastic Solids,'' Phys. Norv., **7**, pp. 13–19.
[24] Shutilov, V., 1988, *Fundamental Physics of Ultrasound*, Gordon and Breach, New York.
[25] Destrade, M., 2003, ''Rayleigh Waves in Symmetry Planes of Crystals: Explicit Secular Equations and Some Explicit Wave Speeds,'' Mech. Mater., **35**, 931–939.
[26] Mozhaev, V. G., 1995, ''Some New Ideas in the Theory of Surface Acoustic Waves in Anisotropic Media,'' *IUTAM Symposium on Anisotropy, Inhomogeneity and Nonlinearity in Solids*, D. F. Parker and A. H. England, eds., Kluwer, Dordrecht, The Netherlands, pp. 455–462.
[27] Furs, A. N., 1997, ''Covariant Form of the Dispersion Equation for Surface Acoustic Waves in Symmetry Planes of Crystals,'' Crystallogr. Rep., **4**, pp. 196–201.

Modeling the Rotation of Orthotropic Axes of Sheet Metals Subjected to Off-Axis Uniaxial Tension

Wei Tong

Hong Tao

Xiquan Jiang

Department of Mechanical Engineering,
Yale University,
219 Becton Center,
New Haven, CT 06520-8284

A simplified version of a newly developed anisotropic plasticity theory is presented to describe the anisotropic flow behavior of orthotropic polycrystalline sheet metals under uniaxial tension. The theory is formulated in terms of the intrinsic variables of principal stresses and a loading orientation angle and its uniaxial tension version requires a non-quadratic stress exponent and up to five anisotropic material functions of the loading orientation angle to specify a flow condition, a flow rule for plastic strain rates, a flow rule for macroscopic plastic spin, and an evolution law of isotropic hardening. In this investigation, the proper analytical form and the associated parameter identification of the anisotropic material functions defining the flow rule of macroscopic plastic spin are discussed for sheet metals with persistent but rotated orthotropic symmetry axes under off-axis uniaxial tension. It is shown that the proposed flow rule of macroscopic plastic spin can successfully model the experimental data on the rotation of orthotropic symmetry axes in the three sheet metals reported, respectively, by Boehler et al. (Boehler and Koss, 1991, Advances in Continuum Mechanics, O. Bruller et al., eds., Springer, Heidelberg, pp. 143–158; Losilla, Boehler, and Zheng, 2000, Acta Mech. **144***, pp. 169–183); Kim and Yin (1997, J. Mech. Phys. Solids* **45***, pp. 841–851); and Bunge and Nielsen (1997 Int. J. Plasticity* **13***, pp. 435–446).* [DOI: 10.1115/1.1755694]

1 Introduction

The microstructure of a polycrystalline sheet metal generally evolves as it undergoes some finite plastic deformation. The plasticity-induced microstructural evolution occurs at least at two levels: the crystallographic texture evolution of grains and the dislocation substructure texture evolution within the grains. There have been continued efforts on improving phenomenological macroscopic plasticity theories by incorporating some constitutive modeling capabilities of material microstructural evolution using scalar and tensorial internal state variables. Isotropic strain or work hardening characterized by an effective plastic strain or equivalent specific plastic work, [1], is perhaps the best-known single scalar state variable model of material microstructural evolution (it basically accounts for the increase of the average dislocation density in a metal due to plastic flow). The kinematic hardening model with a backstress tensor developed for isotropic plasticity theories, [2–4], can be regarded as the phenomenological description of anisotropic hardening behavior due to the evolution of the dislocation substructure towards some preferred spatial orientations that are aligned with current plastic straining directions.

On the other hand, metal products manufactured by rolling (sheet metals), drawing (wires), and extrusion (plates) are typically anisotropic (primarily due to the resulting crystallographic texture, i.e., grains packed with some preferred orientations) and so the use of anisotropic plasticity theories is more appropriate in engineering design and analysis of these materials, [5–8]. While many macroscopic anisotropic plasticity theories proposed in the literature have incorporated isotropic hardening and even kinematic hardening models developed originally for isotropic plasticity theories, almost all of them have explicitly or implicitly assumed that the initial material texture is strong and it persists upon further plastic straining, i.e., the evolution of crystallographic texture is not considered. However, both micromechanical analyses and experimental investigations of rolled sheet metals have shown that there are noticeable and even significant changes of material orthotropic symmetry when a sheet metal is subjected to a plastic strain up to 20%–30%, [9–12]. For two rolled steel and one aluminum sheet metals that were subjected to *off-axis* uniaxial tension (i.e., *the axial loading direction is not aligned with the orthotropic axes of the sheets*), experimental observations have shown that the orthotropic symmetry of these sheet metals is more or less intact but the symmetry axes rotate relatively with respect to the sheet metal itself in the plane of the sheet, [10–12]. A flow rule for macroscopic plastic spin (accounting for the orientational evolution of the material texture frame) may thus be appropriate in an anisotropic plasticity theory to describe these experimental observations.

The concept of macroscopic plastic spin has been explicitly introduced since early 1970s into the framework of polycrystalline plasticity theories, [13–15]. Considerable attentions have been devoted to its role from theoretical considerations (such as a missing kinematics link to the material microstructural evolution) and the necessity from the standpoint of the stability of numerical simulations, [16–34]. The flow rule of macroscopic plastic spin proposed in the literature are mainly motivated and derived through either the representation theorems or some heuristic micromechanical arguments involving tensorial structure variables, identified to be either the orthotropic or other privileged material frames or back stress tensors. Only very simple analytical forms of the flow rule of macroscopic plastic spin have been suggested for illustrative purpose, mostly for either von Mises quadratic isotropic plasticity theory with tensorial backstress kinematic hardening

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, February 28, 2003; final revision, September 3, 2003. Associate Editor: M.-J. Pindera. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

or Hill's 1948 quadratic anisotropic plasticity theory with isotropic hardening and strong and persistent orthotropic symmetry. Nevertheless, the existing simple flow rules of macroscopic plastic spin for orthotropic sheet metals are found to be unable to describe consistently the rotation of orthotropic axes observed in experiments, [11,28].

In this investigation, a flow rule of macroscopic plastic spin proposed in a newly developed anisotropic plasticity theory, [35–40] is evaluated for modeling the rotation of orthotropic symmetry axes in the two steel and one aluminum sheet metals subjected to off-axis uniaxial tension. First, constitutive equations of the anisotropic plasticity theory along with the procedure on evaluating anisotropic material functions in these constitutive equations are summarized in Section 2 for polycrystalline sheet metals under uniaxial tension. The experimental investigations on the orientational evolution of orthotropic symmetry axes in sheet metals are briefly reviewed and material parameters of the anisotropic material functions defining the flow rule of macroscopic plastic spin proposed in the theory are identified in Section 3. Results of both the experimental measurements and the model descriptions are compared for the three sheet metals in Section 3 as well. A discussion on the proper formulation and the necessity of the flow rule of macroscopic plastic spin for modeling the anisotropic plastic flow of sheet metals is presented in Section 4. Conclusions drawn from this investigation on modeling macroscopic plastic spin are given in Section 5.

2 A Model of Anisotropic Plastic Flows Under Uniaxial Tension

The finite elastic-plastic deformation kinematics of a sheet metal may be expressed through the multiplicative decomposition of the macroscopic deformation gradient tensor $\mathbf{F}$ into the elastic and plastic parts $\mathbf{F}^e$ and $\mathbf{F}^p$, [31]. By neglecting the small elastic stretching in sheet metals involving finite plastic deformation, one has the commonly known results of rigid-viscoplastic deformation kinematics as follows:

$$\begin{aligned}\mathbf{F}&=\mathbf{F}^e\mathbf{F}^p\approx\mathbf{R}^*\mathbf{F}^p,\\ \mathbf{L}&=\dot{\mathbf{F}}\mathbf{F}^{-1}\approx(\dot{\mathbf{R}}^*\mathbf{F}^p+\mathbf{R}^*\dot{\mathbf{F}}^p)\mathbf{F}^{p^{-1}}\mathbf{R}^{*-1}\\ &=\dot{\mathbf{R}}^*\mathbf{R}^{*-1}+\mathbf{R}^*\dot{\mathbf{F}}^p\mathbf{F}^{p^{-1}}\mathbf{R}^{*-1},\end{aligned}\tag{1}$$

$$\begin{aligned}\mathbf{L}&=\mathbf{D}+\mathbf{W}\approx\mathbf{D}^p+\mathbf{W}^*+\mathbf{W}^p,\\ \mathbf{D}&=(\mathbf{L}+\mathbf{L}^T)/2=\mathbf{D}^e+\mathbf{D}^p\approx\mathbf{D}^p,\\ \mathbf{W}&=(\mathbf{L}-\mathbf{L}^T)/2=\mathbf{W}^*+\mathbf{W}^p,\end{aligned}\tag{2}$$

where $\mathbf{R}^*$ is the rigid body rotation of the underlying material "texture" frame (some preferred orientations such as orthotropic symmetry axes), $\mathbf{D}^p$ is the plastic rate of deformation tensor, $\mathbf{W}^p$ is the plastic spin tensor defined as the difference between the material spin $\mathbf{W}$, and the so-called constitutive spin $\mathbf{W}^*$, [17,28]. A complete macroscopic theory of plastic flow usually provides a flow condition, flow rules that define both the plastic strain rate tensor $\mathbf{D}^p$ and the plastic spin rate tensor $\mathbf{W}^p$, and isotropic and even anisotropic hardening models via a set of internal state variables and associated kinetic equations on their evolution, [32–34].

Using the principal axes of the applied stress tensor as the Cartesian coordinate system of the choice in this investigation (see Fig. 1), the expressions for the stress tensor $\boldsymbol{\sigma}$, the macroscopic plastic rate of deformation tensor $\mathbf{D}^p$, the macroscopic plastic spin tensor $\mathbf{W}^p$, and the material constitutive spin tensor $\mathbf{W}^*$ under *uniaxial tension* are

$$\boldsymbol{\sigma}=\begin{pmatrix}\sigma_\theta & & \\ & 0 & \\ & & 0\end{pmatrix},\quad \mathbf{D}^p=\begin{pmatrix}\dot{\varepsilon}_1 & \dot{\varepsilon}_{12} & 0\\ \dot{\varepsilon}_{21} & \dot{\varepsilon}_2 & 0\\ 0 & 0 & \dot{\varepsilon}_3\end{pmatrix},$$

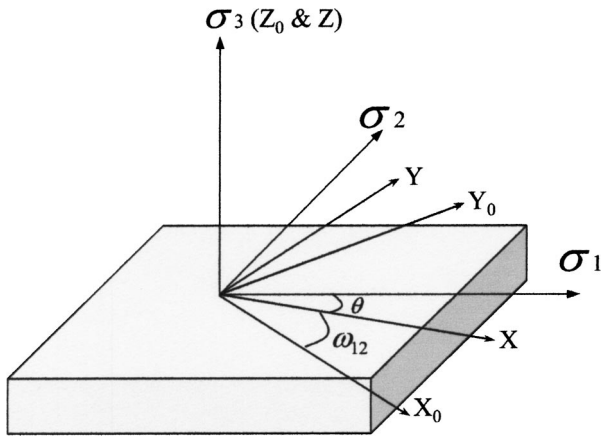


Fig. 1 Definitions of the three Cartesian coordinate systems for a monoclinic sheet metal: (*a*) the principal axes of stress ($\sigma_1,\sigma_2,\sigma_3$); (*b*) the principal axes of the current material texture frame *XYZ*; and (*c*) the sheet material coordinate system $X_0Y_0Z_0$. The principal axis of σ_3 always coincides with Z_0-axis and *Z*-axis to ensure the planar plastic flow of the sheet metal. The in-plane axes *X* and *Y* of the texture frame are defined to be the principal straining directions of the sheet metal under equal biaxial tension ($\sigma_1=\sigma_2$, $\sigma_3=0$). The material coordinate system $X_0Y_0Z_0$ undergoes the same rigid body rotation as the sheet metal itself and it may be chosen to coincide with the initial texture frame of the sheet metal (the initial texture frame of an orthotropic sheet metal is defined by its rolling (RD), transverse (TD), and normal (ND) directions). The loading orientation angle θ is defined as the angle between the principal axis of σ_1 and the *X*-axis of the material texture frame. The relative rotation ω_{12} of the texture frame with respect to the material coordinate system of the sheet metal is due to the macroscopic plastic spin $\dot{\omega}_{12}$, [28].

$$\mathbf{W}^p=\begin{pmatrix}0 & \dot{\omega}_{12} & 0\\ \dot{\omega}_{21} & 0 & 0\\ 0 & 0 & 0\end{pmatrix},\quad \mathbf{W}^*=\begin{pmatrix}0 & \dot{\omega}^*_{12} & 0\\ \dot{\omega}^*_{21} & 0 & 0\\ 0 & 0 & 0\end{pmatrix},\tag{3}$$

where θ is the *loading orientation angle* defined as the angle between the axial loading direction $\sigma_1=\sigma_\theta>0$ and the current in-plane *X*-axis of the sheet metal texture frame (see Fig. 1), and

$$\begin{aligned}&\dot{\varepsilon}_3=-\dot{\varepsilon}_1-\dot{\varepsilon}_2,\quad \dot{\varepsilon}_{21}=\dot{\varepsilon}_{12},\quad \dot{\omega}_{21}=-\dot{\omega}_{12},\\ &\dot{\omega}^*_{21}=-\dot{\omega}^*_{12},\quad \dot{W}_{12}=\dot{\omega}^*_{12}+\dot{\omega}_{12},\end{aligned}\tag{4}$$

where $\dot{W}_{12}$ is the in-plane material spin (macroscopically observable) of the sheet metal. We propose the following rate-dependent phenomenological theory to model the anisotropic plastic flow of a sheet metal under *uniaxial tension*

$$\tau=\tau_0(\xi,\dot{\gamma}),$$

$$\tau^a=\sigma^a_\theta\Phi_1(\theta)\quad\text{(the flow condition and flow function)},\tag{5}$$

$$\dot{\varepsilon}_1=\dot{\gamma}\left(\frac{\sigma_\theta}{\tau}\right)^{a-1}\Phi_1(\theta),\quad \dot{\varepsilon}_2=\dot{\gamma}\left(\frac{\sigma_\theta}{\tau}\right)^{a-1}\Phi_2(\theta),$$

$$\dot{\varepsilon}_{12}=\dot{\gamma}\left(\frac{\sigma_\theta}{\tau}\right)^{a-1}\Phi_3(\theta),\quad\text{(the flow rule for } \mathbf{D}^p)\tag{6}$$

$$\dot{\omega}_{12}=\dot{\gamma}\left(\frac{\sigma_\theta}{\tau}\right)^{a-1}\Phi_4(\theta),\quad\text{(the flow rule for } \mathbf{W}^p)\tag{7}$$

$$\dot{\xi}=\dot{\gamma}\left(\frac{\Phi_5(\theta)}{\Phi_1(\theta)}\right)^{a-1/a}$$

(the evolution law of isotropic hardening), (8)

where τ is the effective flow stress, $\dot{\gamma}$ is the work-conjugate effective plastic strain rate, ξ is a certain internal state variable characterizing the isotropic hardening state of the material, τ_0 is the effective flow strength, $a(>1)$ is the stress exponent and it is a noninteger in general, and $\Phi_1(\theta)$, $\Phi_2(\theta)$, $\Phi_3(\theta)$, $\Phi_4(\theta)$, and $\Phi_5(\theta)$ are five material functions characterizing the planar plastic anisotropy of the sheet metal under uniaxial tension. The constitutive equations Eqs. (5)–(8) are the simplified (uniaxial tension) version of a planar anisotropic plastic flow theory recently developed by Tong et al. [35–40] in terms of principal stresses and a loading orientation angle (which have been called intrinsic variables of a stress field according to Hill [41,42]). The above constitutive equations under uniaxial tension can be justified from a micromechanical point of view (see Appendix). When the associated flow rule is applied to $\dot{\varepsilon}_{12}$, [41,42], one has $\Phi_3(\theta) = \Phi_1'(\theta)/2a$. If one assumes $\Phi_5(\theta) = \Phi_1(\theta)$, then $\dot{\xi} = \dot{\gamma}$, i.e., the isotropic hardening is characterized by the cumulative effective plastic strain. On the other hand, if one assumes $\Phi_5(\theta) = \Phi_1(\theta)\tau^{a/(a-1)}$, then $\dot{\xi} = \tau\dot{\gamma}$, i.e., the isotropic hardening is characterized by the cumulative plastic work per unit volume.

Because the equivalence of the loading orientation angles of θ and $\theta \pm \pi$ due to the symmetry of mechanical loading, each of the five anisotropic material functions of a sheet metal can be represented by a Fourier series, namely,

$$\begin{aligned}
\Phi_1(\theta) &= A_0 + A_1 \sin 2\theta + A_2 \cos 2\theta + \ldots + A_{2k-1} \sin 2k\theta \\
&\quad + A_{2k} \cos 2k\theta + \ldots, \\
\Phi_2(\theta) &= B_0 + B_1 \sin 2\theta + B_2 \cos 2\theta + \ldots + B_{2k-1} \sin 2k\theta \\
&\quad + B_{2k} \cos 2k\theta + \ldots, \\
\Phi_3(\theta) &= C_0 + C_1 \sin 2\theta + C_2 \cos 2\theta + \ldots + C_{2k-1} \sin 2k\theta \\
&\quad + C_{2k} \cos 2k\theta + \ldots, \\
\Phi_4(\theta) &= D_0 + D_1 \sin 2\theta + D_2 \cos 2\theta + \ldots + D_{2k-1} \sin 2k\theta \\
&\quad + D_{2k} \cos 2k\theta + \ldots, \\
\Phi_5(\theta) &= E_0 + E_1 \sin 2\theta + E_2 \cos 2\theta + \ldots + E_{2k-1} \sin 2k\theta \\
&\quad + E_{2k} \cos 2k\theta + \ldots,
\end{aligned} \tag{9}$$

where $k = 1,2,\ldots$, and A_n, B_n, C_n, D_n, and E_n are the Fourier coefficients. The stress exponent a and the Fourier coefficients of the five anisotropic material functions in Eq. (9) may evolve with subsequent plastic deformation when anisotropic hardening due to material texture evolution is modeled. They are all assumed to be constant in this investigation, i.e., the characteristics of the material texture remains more or less the same but the whole texture frame can rotate relatively with respect to the sheet metal itself. If the sheet metal has some additional symmetry characteristics such as orthotropic, trigonal, or cubic symmetry in the plane of the sheet, one can reduce the number of terms in each Fourier series by imposing the equivalency of loading conditions between θ and $-\theta$, θ and $\theta + 2\pi/3$, and θ and $\theta + \pi/2$ respectively. Furthermore, a truncated Fourier series may be used in practice to approximate each anisotropic material function and the number of terms kept in each truncated Fourier series depends on planar anisotropy of the sheet metal. Besides flow stress-strain curves $\sigma_\theta(\varepsilon_1, \dot{\varepsilon}_1)$, plastic strain and spin ratios can be measured under uniaxial tension:

$$R_\theta \equiv \frac{\dot{\varepsilon}_2}{\dot{\varepsilon}_3} = -\frac{\Phi_2(\theta)}{\Phi_1(\theta) + \Phi_2(\theta)}, \quad \text{(the plastic axial strain ratio)} \tag{10a}$$

$$\Gamma_\theta \equiv \frac{\dot{\varepsilon}_{12}}{\dot{\varepsilon}_1} = \frac{\Phi_3(\theta)}{\Phi_1(\theta)}, \quad \text{(the plastic shear strain ratio)} \tag{10b}$$

$$\Pi_\theta \equiv \frac{\dot{\omega}_{12}}{\dot{\varepsilon}_1} = \frac{\Phi_4(\theta)}{\Phi_1(\theta)}. \quad \text{(the plastic spin ratio).} \tag{10c}$$

The five anisotropic material functions $\Phi_1(\theta)$, $\Phi_2(\theta)$, $\Phi_3(\theta)$, $\Phi_4(\theta)$, and $\Phi_5(\theta)$ in the proposed anisotropic plastic flow theory under uniaxial tension can thus be completely examined from the experimental measurements of $\sigma_\theta(\varepsilon_1, \dot{\varepsilon}_1)$, R_θ, Γ_θ, and Π_θ, [35–37,40]. Evaluation of the proper analytical form of the anisotropic material function $\Phi_4(\theta)$ that defines the flow rule of macroscopic plastic spin or the plastic spin ratio Π_θ for a given sheet metal is the focus of this investigation and will be discussed in details in the next section.

3 Macroscopic Plastic Spin in Sheet Metals Subjected to Off-Axis Uniaxial Tension

In this section, the procedures of experimental investigations on the orientational evolution of material texture frames in three sheet metals will be reviewed first and the reported experimental results will be summarized briefly. The Fourier series representation and the identification of its Fourier coefficients of the anisotropic material function $\Phi_4(\theta)$ or the plastic spin ratio Π_θ will then be detailed. The model description of the rotation of orthotropic axes of these three sheet metals under off-axis uniaxial tension due to macroscopic plastic spin will be compared with the experimental measurements. Although the theory presented in Section 2 can be applied to monoclinic sheet metals under uniaxial tension, [37,40], the sheet metals considered in the following are assumed to be initially orthotropic and remain so under off-axis uniaxial tension, [11,28]. So only the coefficients of the sine terms in the Fourier series of $\Phi_4(\theta)$ and the coefficients of the cosine terms in the Fourier series of $\Phi_1(\theta)$ are nonzero.

3.1 On the Experimental Measurements of the Rotation of Orthotropic Axes due to Macroscopic Plastic Spin. There have been rather limited experimental investigations on detecting the macroscopic plastic spin and its evolution in orthotropic sheet metals so far, [10–12,43,44]. A direct *mechanistic* evaluation of the macroscopic plastic spin in orthotropic sheet metals under uniaxial tension has been carried out using a two-step experimental technique by Boehler and Koss [10] and Kim and Yin [11]. It consists of (a) *the plastic deformation step* of uniaxial straining multiple large sheet samples up to various plastic strain levels (10%–30%) without necking at different off-axis angles and (b) *the material (texture frame) probing step* of measuring the directional dependence of uniaxial tension flow stress (more specifically, the yield stress with a big offset strain of 0.2%) of smaller tensile sheet samples cut off from the deformed large sheet with every 10 deg or 15 deg offset angle increments from the original rolling direction.

Boehler et al. [10,45] tested large sheets of an aluminum killed soft steel of size 1000×360 mm^2 under uniaxial tension with initial off-axis loading orientation angles of 30 deg, 45 deg, and 60 deg for various plastic strain levels up to 20% and above. Twelve smaller specimens cut off from each of the deformed large sheets with angles of 0 deg, 15 deg, 30 deg, 45 deg, 60 deg, 75 deg, 90 deg, 105 deg, 120 deg, 135 deg, 150 deg, 165 deg offset from the rolling direction of the sheets were then tested for yield stress measurements. Kim and Yin [11] carried out very similar tests on an automotive low carbon steel sheet using the same three initial off-axis loading orientation angles for various strain levels up to 10%. They used a total of 18 smaller specimens cut off from each of the deformed large sheets in the material probing step with each specimen at an offset angle of every 10 deg increment from the rolling direction of the large sheets. To enhance the degree of anisotropy of the steel sheets that were nearly isotropic initially, Kim and Yin [11] pre-strained the steel sheets along the rolling direction up to strains of 3% and 6%, respectively. By examining the directional dependence of flow stress measured from 12 and 18 small tensile specimens, respectively, both Boehler et al.

[10,45] and Kim and Yin [11] concluded that their sheet metals remain approximately orthotropic but there exists a large in-plane rotation of the orthotropic symmetry axes relative to the sheet metal itself under off-axis uniaxial tension. Both positive (counterclockwise) and negative (clockwise) rotations as defined in Fig. 1 were observed for initial loading orientation angles of 30 deg and 60 deg, respectively, and the orthotropic symmetry axes become completely aligned with the external axial loading direction within 5–10% uniaxial plastic strain. That is, the texture axis that is coincided with the initial rolling direction rotates towards the axial loading direction with an initial loading orientation angle of 30 deg but the texture axis that is coincided with the initial transverse direction rotates towards the axial loading direction with an initial loading orientation angle of 60 deg. A rotation of the orthotropic symmetry axes with an initial loading orientation angle of 45 deg was detected in both investigations as well but a positive rotation was reported by Boehler and Koss [10] while a negative rotation was found by Kim and Yin [11].

An experimental determination of macroscopic plastic spin of polycrystals based on material crystallographic texture measurements has been proposed by Bunge and Nielsen [12]. They divided the crystallographic texture change of a polycrystal undergoing plastic deformation into an average rotation of some common texture reference axes characteristic for the whole polycrystal material element and a ''spreading'' of the individual crystal orientations away from the common (rotated) reference texture frame. When the texture ''spreading'' is neglected, the texture evolution can thus be characterized approximately by the texture rotation or texture spin. Bunge and Nielsen [12] measured the orientation distribution function (ODF) of an annealed polycrystalline aluminum sheet of 1 mm thickness before and after being subjected to off-axis uniaxial tension to a total plastic strain of 20% at 11 different initial loading orientation angles. They analyzed the rotation of a characteristic reference system formed by the symmetric elements of the texture with an accuracy of $\sim$0.5 deg using an autocorrelation function of ODF and considered the texture rotation or spin being related to the macroscopic plastic spin in the macroscopic theories of plasticity. They found that the amount of texture rotation at a uniaxial plastic strain of 20% depends on the initial loading orientation angle and the maximum plastic spin ratio is about 5 deg/20% occurred around the off-axis loading angle of 22.5 deg (i.e., $\Pi_{22.5\text{deg}} \approx 5$ deg/20% if a constant plastic spin is assumed). The macroscopic plastic spin of the aluminum sheet defined by the crystallographic texture spin according to Bunge and Nielsen [12] is much smaller than that of low-carbon steel sheets defined by the symmetry characteristics of the directional dependence of flow stress according to Boehler and Koss [10] and Kim and Yin [11]. Measured pole figures of a steel sheet investigated by Boehler and Koss [10] under 45 deg off-axis uniaxial tension showed that the symmetry part of the crystallographic texture in the steel sheet did rotate completely towards the axial loading direction at a plastic strain level of about 10%.

3.2 An Analysis of the Rotation of Orthotropic Axes due to Macroscopic Plastic Spin. As there is very little relative rotation of the sheet metal with respect to the fixed laboratory loading frame under uniaxial tension, [10–12,43,44], i.e., $\dot{W}_{12} \approx 0$, one has (see Fig. 1)

$$\omega_{12} + \theta \approx \theta_0, \quad \text{and} \quad \dot{\omega}_{12} + \dot{\theta} \approx 0, \tag{11}$$

where θ_0 and θ are, respectively, the initial and current loading orientation angles, and ω_{12} is the rotation of the sheet metal texture frame due to plastic spin $\dot{\omega}_{12}$. One can thus rewrite Eq. (10*c*) as

$$\Pi_\theta = -\frac{\dot{\theta}}{\dot{\varepsilon}_1} \quad \text{or} \quad \varepsilon_1 = -\int_{\theta_0}^{\theta} \frac{d\theta}{\Pi_\theta}. \tag{12}$$

The directional dependence of flow stress under uniaxial tension is usually much milder than that of plastic strain and spin ratios.

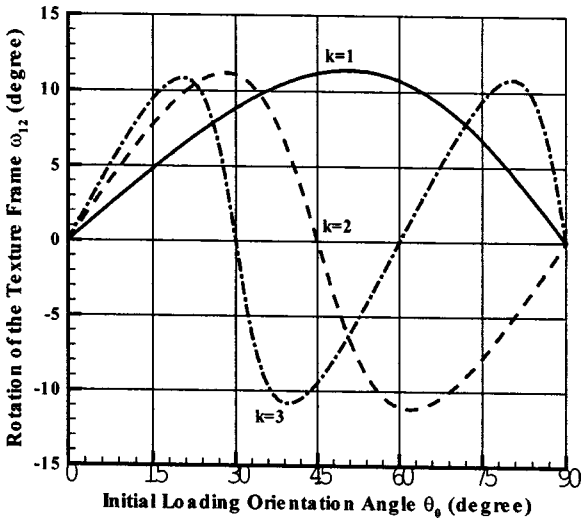


Fig. 2 The amount of rotation ω_{12} of the material texture frame due to plastic spin at a fixed uniaxial plastic strain ε_1 of 20% as a function of the initial loading orientation angle θ_0 with three different k values according to Eq. (13*b*) ($D=1$ is used for all data points)

To the first approximation, one may set $\Phi_1(\theta) \approx 1$ (assuming $\tau(\xi, \dot{\gamma}) = \sigma_0$). When only one of coefficients in the Fourier series of $\Phi_4(\theta)$ is nonzero, i.e., $\Pi_\theta \approx \Phi_4(\theta) \approx D \sin 2k\theta (k=1,2,\ldots)$, one can obtain an analytical expression of Eq. (12) as

$$\theta = \frac{1}{k} \arctan[\{\tan(k\theta_0) e^{-2kD\varepsilon_1}\}], \tag{13a}$$

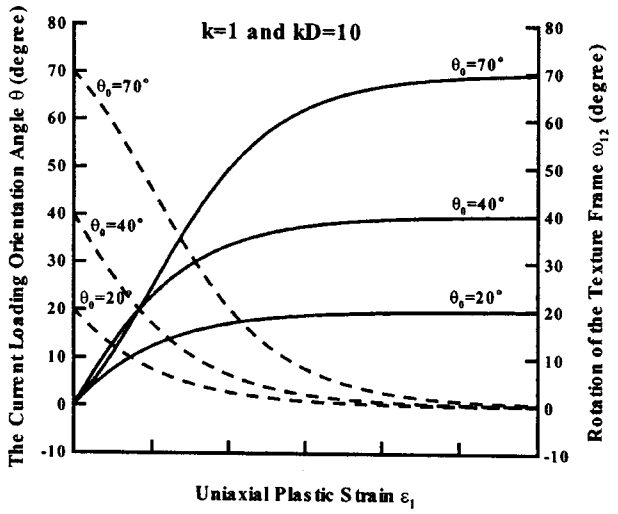
so

$$\omega_{12} = \theta_0 - \frac{1}{k} \arctan[\{\tan(k\theta_0) e^{-2kD\varepsilon_1}\}]. \tag{13b}$$

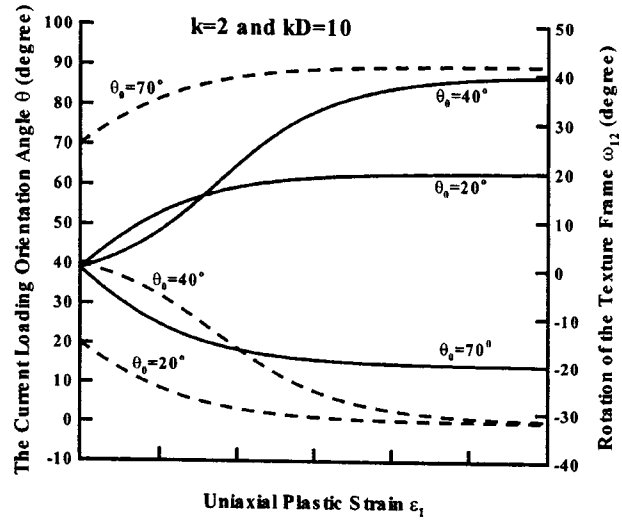
The rotation ω_{12} of the material texture frame with respect to the sheet metal itself or the current loading orientation angle θ as a function of the initial loading orientation angle θ_0 and the uniaxial plastic strain $\varepsilon_1 (\geq 0)$ are shown in Fig. 2 and Fig. 3, respectively, using Eq. (13). The material texture frame will eventually stop spinning at certain loading orientation angles at sufficiently large plastic strains and these loading orientation angles are the *equilibrium* orientations of the material texture frame. Possible equilibrium orientations of the material texture frame of a sheet metal are the loading orientation angles that satisfy the conditions $\Pi_\theta = 0$ and $\Pi'_\theta \geq 0$. Loading orientation angles that satisfy the conditions $\Pi_\theta = 0$ and $\Pi'_\theta < 0$ are metastable orientations and are not true equilibrium orientations (i.e., any small disturbance can cause the material texture frame to rotate away from those orientations). When D>0, the possible equilibrium orientations of the material texture frame are $\theta = 0$ deg for $k=1$, $\theta = 0$ deg and 90 deg for $k=2$, and $\theta = 0$ deg, 60 deg and 120 deg for $k=3$. When D<0, the possible equilibrium orientations of the material texture frame are $\theta = 90$ deg for $k=1$, $\theta = 45$ deg and 135 deg for $k=2$, and $\theta = 30$ deg, 90 deg, and 150 deg for $k=3$. In general, Eq. (12) may also be rewritten using the Fourier series expansion of $\Pi_\theta = \Phi_4(\theta)/\Phi_1(\theta)$ as

$$\varepsilon_1 = -\int_{\theta_0}^{\theta} \frac{d\theta}{(d_1 \sin 2\theta + d_2 \sin 4\theta + d_3 \sin 6\theta + \ldots)}. \tag{14}$$

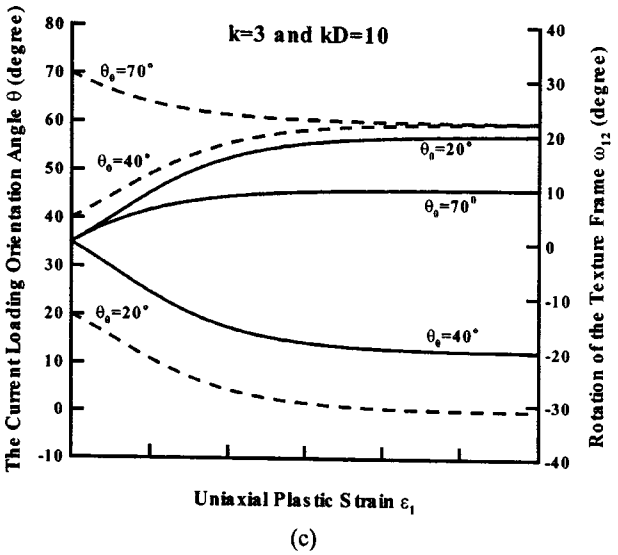
The relation between the current loading orientation angle θ or the rotation of the texture frame ω_{12} and the uniaxial plastic strain ε_1 can be obtained by integrating Eq. (14) numerically.



(a)



(b)



(c)

Fig. 3 The current loading orientation angle θ and the amount of rotation ω_{12} of the material texture frame due to plastic spin as a function of uniaxial plastic strain ε_1 with three different initial loading orientation angles θ_0 and three different k values according to Eq. (13) ($kD=10$ is used for all data points)

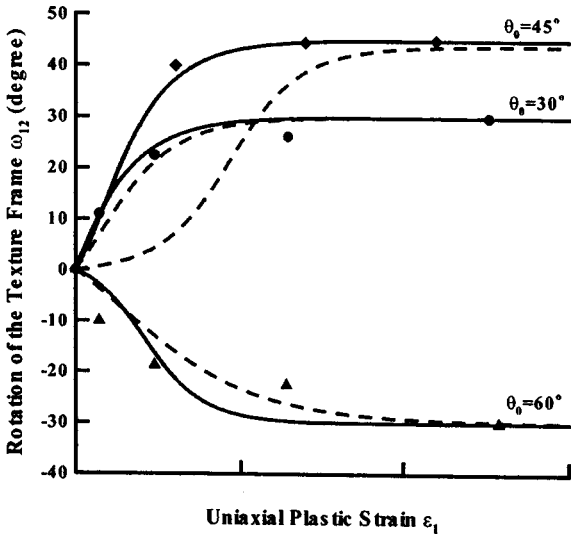


Fig. 4 Comparison of the model description (solid and dashed lines) and the experimental data (filled symbols) of a steel sheet reported by Boehler and Koss [10] and Losilla et al. [45] on the rotation ω_{12} of the material texture frame due to plastic spin as a function of uniaxial plastic strain ε_1 with different initial loading orientation angles $\theta_0=30deg$, 45 deg, and 60 deg. The solid lines are given by Eq. (14) with $d_1=7$, $d_2=10$ and $d_3=-3$ (all other coefficients are zero). The dashed lines are given by Eq. (13*b*) with $k=2$ and $D=9$ (the initial loading orientation angles of 30 deg, 46 deg, and 60 deg were used).

3.3 Comparison Between the Model Description and Experimental Measurements on the Rotation of Orthotropic Axes. Unlike plastic axial and shear strain ratios R_θ and Γ_θ (see Eqs. (10*a*) and (10*b*)) that can be directly determined from incremental plastic strain measurements in each uniaxial tension test, the plastic spin ratio has to be determined from the measurements of both axial plastic strain increments and rotations of the material symmetry axes (by a separate mechanical or material texture measurement, see Section 3.1). Only limited experimental data are reported for a given sheet metal either in terms of the rotation of orthotropic axes as a function of uniaxial plastic strain with selected initial loading orientation angles (see Fig. 4 and Fig. 5) or in terms of the rotation of crystallographic texture symmetry axes as a function of initial loading orientation angles at a fixed uniaxial plastic strain (see Fig. 6). The plastic spin ratio can be obtained in principle by curve-fitting and numerical differentiation of the experimental data shown in Fig. 4 and Fig. 5, the material parameters and coefficients in Eq. (13*b*) and/or Eq. (14) can then be determined. The following trial-and-error procedure is used instead for parameter identification:

1. Use the results shown in Fig. 2 and Fig. 3 (which are given by Eq. (13) using different k values) as the basis to determine the dominant term (the value of k) in the Fourier series of Π_θ measured in experiments.
2. Adjust the value of the coefficient D (both its sign and magnitude) to best describe the experimental data. If Eq. (13) with the selected values of k and D can model all of the experimental measurements reasonably well, then the plastic spin ratio is determined as $\Pi_\theta = D \sin 2k\theta$.
3. Add one or more sine terms for the Fourier series of Π_θ if Eq. (13) cannot describe the experimental measurements satisfactorily. Estimate the sign and magnitude of the Fourier coefficient of each new term by comparing the results shown in Fig. 2 and Fig. 3 with the experimental measurements. Adjust the Fourier coefficients d_i in Eq. (14) iteratively until

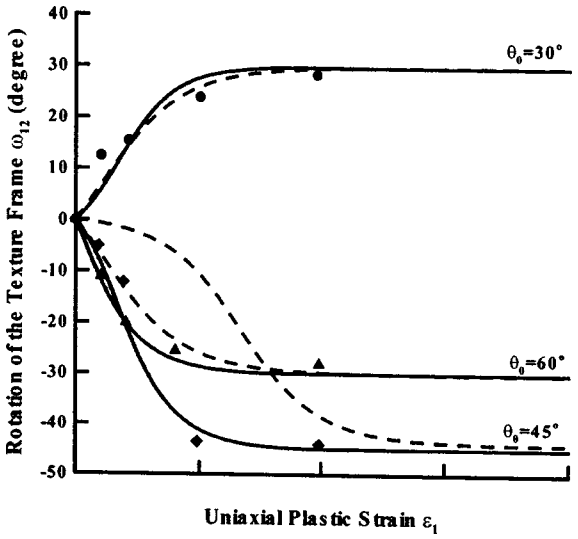


Fig. 5 Comparison of the model description (solid and dashed lines) and the experimental data (filled symbols) of a steel sheet reported by Kim and Yin [11] on the rotation ω_{12} of the material texture frame due to plastic spin as a function of uniaxial plastic strain ε_1 with different initial loading orientation angles $\theta_0=30deg$, 45 deg, and 60 deg. The solid lines are given by Eq. (14) with $d_1=-8$, $d_2=17$ and $d_3=3$ (all other coefficients are zero). The dashed lines are given by Eq. (13*b*) with $k=2$ and $D=12.5$ (the initial loading orientation angles of 30 deg, 46 deg, and 60 deg were used).

the model description matches closely the experimental measurements. The plastic spin ratio is then given as $\Pi_\theta = d_1 \sin 2\theta + d_2 \sin 4\theta + d_3 \sin 6\theta + \ldots$.

The above procedure was applied to model the experimental data on the three sheet metals reported by Boehler et al. [10,45], Kim and Yin [11], and Bunge and Nielsen [12] and plastic spin ratios of these three sheet metals were determined as follows:

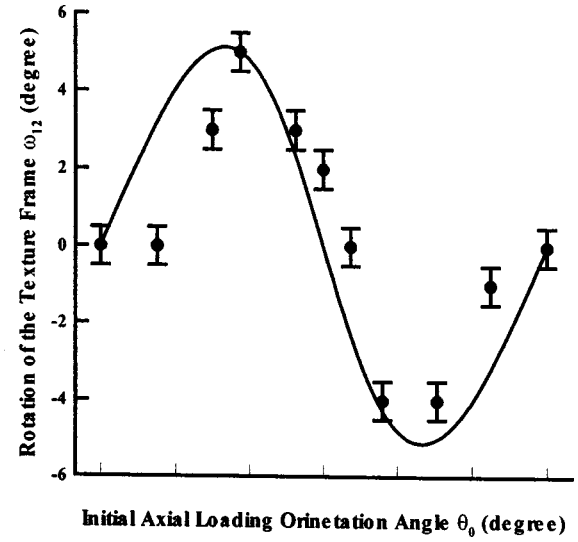


Fig. 6 Comparison of the model description (solid and dashed lines) and the experimental data (filled symbols) of an aluminum sheet reported by Bunge and Nielsen [12] on the amount of rotation ω_{12} of the material texture frame due to plastic spin at a fixed uniaxial plastic strain ε_1 of 20% with 11 different initial loading orientation angles θ_0. The solid line is given by Eq. (13*b*) with $k=2$ and $D=0.45$.

$$\Pi_\theta = 9 \sin 4\theta, \quad \text{and} \quad \Pi_\theta = 7 \sin 2\theta + 10 \sin 4\theta - 3 \sin 6\theta, \ \text{(Boehler and Koss)} \tag{15a}$$

$$\Pi_\theta = 12.5 \sin 4\theta, \quad \text{and} \quad \Pi_\theta = -8 \sin 2\theta + 17 \sin 4\theta + 3 \sin 6\theta, \ \text{(Kim and Yin)} \tag{15b}$$

$$\Pi_\theta = 0.45 \sin 4\theta \ \text{(Bunge and Nielsen)}. \tag{15c}$$

As shown in Fig. 4 and Fig. 5, the plastic spin ratio using a single sine term with $k=2$ can only describe some of the experimental data on the two steel sheets reported by Boehler and Koss [10] and Kim and Yin [11] (for initial loading orientation angles of 30 deg and 60 deg). Actually, no rotation of the orthotropic axes is predicted by $\Pi_\theta = D \sin 4\theta$ for the initial loading orientation angle of 45 deg at all. However, $\theta_0=45deg$ is not one of the true equilibrium orientations of the material texture frame with such a plastic spin ratio (see Section 3.2). If one assumes the initial loading orientation angle to be 44 deg and 46 deg, respectively, for each investigation (say, there were some experimental errors due to some slight misalignments), then the rotation of orthotropic axes occurs in both cases and matches the experimental observations at large strains. The predictions at small strains are however inconsistent with the experimental data. Indeed, the plastic spin ratios with three sine terms given in Eq. (15*a*) and Eq. (15*b*) are needed to adequately model the experimental data reported by both Boehler and Koss [10] and Kim and Yin [11] for all three different initial loading orientation angles. On the other hand, the plastic spin ratio using a single sine term with $k=2$ describe reasonably well the directional dependence of the rotation of texture symmetry axes at a fixed uniaxial plastic strain of 20% as observed by Bunge and Nielsen [12] for an annealed aluminum sheet. The magnitude of the plastic spin ratio of the aluminum sheet is, however, about 1/20 to 1/30 of that of the steel sheets. The difference in the magnitude of plastic spin ratios between steel and aluminum sheets may be due to the difference in the characteristics of their initial anisotropy and a further micromechanical investigation is warranted to elucidate its physical origin. According to the plastic spin ratios given in Eq. (15), the equilibrium orientation of the material texture frames in these three sheet metals is either $\theta=0$ deg (the *X*-axis will eventually coincide with the external axial loading direction) or $\theta=90$ deg (the *Y*-axis will eventually coincide with the external axial loading direction).

4 Discussion

A plane-stress anisotropic plasticity theory of sheet metals is often formulated using the Cartesian stress components σ_x, σ_y, and σ_{xy} projected onto the principal axes of the material texture (symmetry) frame *XYZ* as shown in Fig. 1. Hill [41,42] has recently advocated the use of the so-called intrinsic variables of principal stresses (σ_1, σ_2) and a loading orientation angle θ for developing anisotropic plasticity theories. He has argued that the resulting plasticity theories should be more appealing to both theoreticians and experimentalists. A new anisotropic plasticity theory has indeed been proposed using these intrinsic variables by Tong et al. [35–40] and the uniaxial tension version of the theory is presented in this investigation. In both theoretical analyses and experimental evaluations of an anisotropic plasticity theory using either formulation, one needs to know the initial orientation of the material texture frame and its subsequent evolution with respect to the sheet metal itself during a plastic deformation process. In other words, an explicit mechanistic definition (i.e., an experimental procedure for its determination by a mechanical test) of the material texture frame at a given plastic deformation stage is required by such an anisotropic plasticity theory. For orthotropic sheets, the orthotropic axes can be identified with the symmetry axes detected in the orientational dependence of mechanical properties such as flow stress under uniaxial tension, [10,11]. If plastic anisotropy of the sheet metal is solely due to the crystallographic texture, then the crystallographic texture symmetry axes may be

used as well to measure the orientation of the material texture frame in the sheet metal, [12]. While such a definition of the material texture frame works well for sheet metals with orthotropic or other higher-order symmetries, it cannot however be extended to monoclinic sheets. A different and more general definition of the in-plane axes X and Y of the sheet material texture frame coordinate system has been given in the anisotropic plasticity theory presented here, that is, the X and Y-axes are defined to be the principal axes of in-plane plastic strain rates of the sheet metal under equal biaxial tension ($\sigma_1=\sigma_2$). In practice, an out-of-plane uniaxial compression test may be used to experimentally determine the material texture axes X and Y if plastic flow is unaffected by hydrostatic loading. This definition of the material texture axes is equivalent to the one based on the symmetry characteristics of flow stress under uniaxial tension for orthotropic sheets with isotropic hardening but requires much less experimental efforts, [35].

Most of the existing anisotropic plasticity theories assume that a polycrystalline sheet metal is orthotropic initially and the initial orthotropy symmetry is strong and persists during subsequent plastic deformation, [6,28]. When the sheet metal deforms plastically under on-axis loading conditions (i.e., the principal axes of stress coincide with the orthotropic axes of the sheet), there is no ambiguity on the orientation of the current orthotropic texture frame with respect to the sheet metal itself as there is no plastic spin of the material texture frame. However, when the loading is off-axis in uniaxial tension or shear tests, the original RD and TD directions of the sheet metal are no longer orthogonal and the current orientation of the material texture axes X and Y cannot be clearly identified without additional theoretical hypotheses or experimental characterization. Few existing anisotropic plasticity theories offer a flow rule for macroscopic plastic spin at all so most of them assume explicitly or implicitly that the macroscopic plastic spin is always zero and the material texture frame rotates along with the sheet metal itself. The experimental measurements on the current orientation of the material texture frame after a sheet metal is subjected to off-axis uniaxial tension have shown that there exist detectable and even significant relative rotations between the material texture frame and the sheet metal, [10–12]. A robust and flexible flow rule of macroscopic plastic spin as proposed in this investigation should be incorporated into an anisotropic plasticity theory to improve its modeling capabilities. A comparative evaluation of the proposed flow rule of macroscopic plastic spin with some of the specific analytical forms of the flow rule of macroscopic plastic spin appeared in the literature is in order. As pointed out in the Section 1, the explicit forms of the flow rule of macroscopic plastic spin have been motivated largely by invoking the representation theorems for isotropic functions in conjunction with the concept of tensorial structure variables mostly for quadratic plastic flow theories, [13–30]. Only two analytical expressions of the flow rule of macroscopic plastic spin have often been cited in the literature and they have the following forms for an orthotropic sheet metal under off-axis uniaxial tension, [28],

$$\dot{\omega}_{12}=\eta_a\dot{\varepsilon}_{xy},$$

$$\text{or}\quad \Pi_\theta\equiv\frac{\dot{\omega}_{12}}{\dot{\varepsilon}_1}=\eta_a\frac{(\dot{\varepsilon}_1-\dot{\varepsilon}_2)\sin\theta\cos\theta+\dot{\varepsilon}_{12}\cos 2\theta}{\dot{\varepsilon}_1}=\eta_a\left[\left(1+\frac{R_\theta}{1+R_\theta}\right)\sin\theta\cos\theta+\Gamma_\theta\cos 2\theta\right], \quad (16a)$$

$$\dot{\omega}_{12}=\eta_b\sigma_\theta\dot{\varepsilon}_{12},\ \text{or}\ \Pi_\theta\equiv\frac{\dot{\omega}_{12}}{\dot{\varepsilon}_1}=\eta_b\sigma_\theta\Gamma_\theta, \quad (16b)$$

where η_a and η_b are plastic spin coefficients and can be in general a function of the isotropic invariants of a given loading stress tensor and the material symmetry orientations, R_θ, Γ_θ, and σ_θ are, respectively, the plastic axial strain ratio, the plastic shear strain ratio, and the flow stress under uniaxial tension (see Section 2). The flow rule Eq. (16*b*) is the uniaxial tension version of the constitutive equation for plastic spin $\mathbf{W}^p=\eta(\boldsymbol{\sigma}\mathbf{D}^p-\mathbf{D}^p\boldsymbol{\sigma})$ that has been adapted widely in the literature, [26–28]. In many actual applications of the above two flow rules reported in the literature, a *constant* plastic spin coefficient is usually assumed, [18,28]. The flow rule given by Eq. (16*a*) has been used extensively for investigating the plastic spin effect but they were mostly illustrative without experimental corroboration, [15,18]. The flow rule given by Eq. (16*b*) has been employed by Kuroda [26] to simulate the inverse Swift effect in free-end torsion experiments assuming that the material is orthotropic prior to torsion. Kim and Yin [11] and Dafalias [28] have used the same expression for the plastic spin along with Hill's 1948 quadratic anisotropic plasticity theory, [5], to simulate the orientational evolution of orthotropic symmetry axes in steel sheets upon off-axis uniaxial tensile deformation. While qualitative agreements were found in their analyses, the orientational evolution of the orthotropic axes with increasing plastic deformation is not described with great accuracy and significantly different values of the plastic spin coefficient η_b (which also has to be opposite in sign when comparing with the one used by Kuroda [26]) were needed for the best description of each of the off-axis tensile tests with the initial loading orientation angles of 30 deg, 45 deg, and 60 deg. In the light of the anisotropic plastic flow theory proposed here (see Section 2 and Section 3), the two widely used analytical expressions for the flow rule of macroscopic plastic spin as given in Eqs. (16*a*) and (16*b*) for uniaxial tension are indeed both overly restrictive (linking directly the plastic spin $\dot{\omega}_{12}$ with the plastic shear rate $\dot{\varepsilon}_{xy}$ or $\dot{\varepsilon}_{12}$) and overly simplistic (no dependence on the loading orientation angle θ is given for the plastic spin coefficients η_a and η_b at all in their actual application examples). The micromechanical analysis of the plastic flow of single crystals with a regularized Schmid law under uniaxial tension given in the Appendix also shows that the general validity of the plastic spin equation given by either Eq. (16*a*) or Eq. (16*b*) is indeed questionable.

As mentioned in the Introduction, the other major area of interest of incorporating a flow rule of macroscopic plastic spin is in isotropic plasticity theories with kinematic hardening [23,46]. One can define the material texture frame as the principal axes of the backstress tensor, then a flow rule of macroscopic plastic spin becomes basically a part of the evolution law of kinematic hardening that describes the orientational evolution of the principal axes of the backstress tensor (the other part covers the evolution of the strength of the backstress in terms of its principal components). However, direct measurements of the backstress tensor (and hence its principal axes) using tension-compression or similar tests at different loading orientation angles are required to properly evaluate any specific form of the flow rule of plastic spin for kinematic hardening. Experimental inference of the form of the flow rule of plastic spin for isotropic plasticity theories with kinematic hardening by simulating finite deformation simple shear tests is problematic [46] as the crystallographic aspect of the material texture evolution may become significant and even dominant. As observed early in Section 3, the magnitude of the plastic spin ratio of the aluminum sheet, [12], is only about 1/20 to 1/30 of that of the steel sheets, [10,11]. Such a difference may be attributed to the two different plastic spin detection methods used: the plastic spin determined by Bunge and Nielsen [12] is mainly related to the material crystallographic texture evolution while the plastic spin determined by Boehler and Koss [10] and Kim and Yin [11] may be related primarily to the evolution of dislocation substructures in steel sheet metals (especially at small plastic strains). In this investigation, it was assumed that all three sheet metals have a pre-existing orthotropic symmetry and the orienta-

tional dependence of flow stress and plastic strain ratio of the sheet metals under off-axis uniaxial tension remains orthotropic with the *same* symmetry axes, [28]. Consequently, a single macroscopic flow rule of macroscopic plastic spin was used to characterize satisfactorily the relative rotation of the material symmetry axes regardless of its origin. The validity of the persistent orthotropic symmetry assumption cannot be examined directly for these three sheet metals due to lack of experimental data. If the experimental data on the orientational dependence of plastic strain ratio of the two steel sheets upon off-axis uniaxial tension were also made available and they would show indeed that the symmetry axes of plastic strain ratio do not coincide with the symmetry axes of flow stress, the sheet metals should then be treated as monoclinic instead of orthotropic. Incorporations of both a backstress kinematic hardening model and a flow rule of macroscopic plastic spin in an anisotropic plastic flow theory may be a possible modeling approach, [23,43,44]. Direct experimental evaluation of the backstress tensor is however very challenging if not impossible for sheet metals (as in-plane uniaxial compression tests are rather difficult to carry out for thin sheets) and flow rules of two plastic spins are required to describe the orientation evolution of the principal axes of both the backstress (related primarily to the dislocation substructure) and the crystallographic texture frame, respectively, [23]. When only one plastic spin is used, its experimental evaluation becomes ambiguous unless further clarification on the definition of the material texture frame that is associated with the plastic spin is provided and extensive experimental data are made available. For example, Truong Qui and Lippmann [43,44] have proposed a quadratic anisotropic plasticity theory that generalizes Hill's orthotropic theory, [5], for monoclinic sheets with combined isotropic and kinematic hardening and a plastic spin. Their theory is formulated using the Cartesian stress components σ_x, σ_y, and σ_{xy} on the axes of the material texture coordinate system which is associated with their plastic spin. However, as no experimental data on in-plane uniaxial compression flow stress are available for the steel and aluminum sheets investigated, respectively, by Boehler and Koss [10] and by Truong Qui and Lippmann [43,44], the evaluation of *both* the backstress and the rotation of the material texture frame due to plastic spin in their theory is impossible using solely the experimental data on the directional dependence of uniaxial tensile flow stress. Truong Qui and Lippmann [43,44] used a least-square fitting parameter identification procedure that lumps together all material parameters plus a rotation angle ω_{12} due to plastic spin. Such an indirect approach in evaluating the effect of plastic spin and other aspects of a highly nonlinear anisotropic plastic flow behavior is very questionable as noted by McDowell et al. [46]. Alternatively, one may invoke an anisotropic plasticity theory with a nonassociated flow rule, [47], and uses a yield surface to model the anisotropy of flow stress and a separate flow surface to model the anisotropy of plastic strain ratios. Two plastic spins associated with the evolution of symmetry axes of yield and flow surfaces can thus be in principle evaluated independently based on the experimental data on the orientational dependence of flow stress and plastic strain ratio respectively following the methodology give in this investigation. Under this context, the plastic spin ratio obtained here for the two steel sheets should perhaps be limited to the evolution of the yield surface and its applicability to the evolution of the flow surface cannot be assessed without additional experimental data on the orientational dependence of plastic strain ratio.

We take the viewpoint that the purpose of a macroscopic anisotropic plasticity theory is mainly to provide a mathematically more compact but still physically sound description of the plastic flow behavior of a sheet metal so that engineering analyses and designs of sheet metal forming processes can be carried out in an efficient way with accepted accuracy. One basically calibrates the material parameters in the theory through a set of mechanical tests under simple loading conditions and then applies the theory to analyze problems under more general loading conditions. Our anisotropic plasticity theory formulated in terms of the intrinsic variables of principal stresses and a loading orientation angle intends to strike a proper balance between the mathematical compactness and the descriptive robustness through the truncated Fourier series representation of each of anisotropic material functions characterizing plastic anisotropy of a sheet metal. Depending on the extent of the experimental data made available, the degree of plastic anisotropy, and the accuracy required for an analysis, a flexible and adaptive anisotropic plasticity theory can thus be established for practical engineering applications, [35–40]. There have been some debates in recent years on the role and necessity of macroscopic plastic spin within the framework of a macroscopic polycrystalline plasticity theory [24,46,48]. We suggest that a flow rule of macroscopic plastic spin should be considered if it improves the mathematical formulation (say, the compactness) and the constitutive modeling quality of a phenomenological theory of a sheet metal and if there is a clear physical basis (such as the relative rotation of the material texture frame against the material itself) and an associated experimental procedure for its evaluation. This investigation showed that the proposed flow rule of macroscopic plastic spin can be used to describe effectively the orientational evolution of the material texture frame in three orthotropic sheet metals subjected to off-axis uniaxial tension and thus their plasticity anisotropy without invoking the use of other mathematically more complicated anisotropic hardening models, [43–45]. When other aspects of material texture evolution such as texture spreading or texture sharpening have significant effects on the anisotropic plastic flow behavior of a sheet metal, constitutive anisotropic hardening equations in addition to the flow rule of plastic spin may have to be added to characterize the evolution of texture intensity.

5 Conclusions

A new flow rule of macroscopic plastic spin has been proposed for modeling the orientational evolution of the material texture frame of a sheet metal subjected to off-axis uniaxial tension. When a sheet metal has a pre-existing and persisting orthotropic symmetry, the anisotropic material function in the flow rule can be approximated by a truncated Fourier sine series of the loading orientation angle and its Fourier coefficients can be identified using the experimental data on the rotation of the material texture frame relative to the sheet metal itself. The flow rule of macroscopic plastic spin is found to provide a consistent description of the experimental data on the orientational evolution of the material texture frame of three sheet metals reported in the literature. Such a flow rule of macroscopic plastic spin should be incorporated into an anisotropic plasticity theory for finite plastic deformation applications when it can improve both the mathematical formulation (the compactness) and the descriptive quality of the theory and when it can be unambiguously evaluated experimentally based on an explicit mechanistic definition of the material texture frame. Additional theoretical and experimental investigations are needed to clarify the definition of macroscopic plastic spin and its evaluation for a monoclinic sheet metal.

Acknowledgments

The work reported here was supported in part by a CAREER award to WT from the National Science Foundation (Grant No. CMS-973397, Program Director: Dr. K. Chong). WT would like to acknowledge Profs. John Hutchinson and Jim Rice of Harvard University for their comments that helped to clarify our definition of the material texture frame in a monoclinic sheet.

Appendix

On the Micromechanical Basis of the Proposed Anisotropic Plastic Flow Theory. At ambient conditions, the plastic flow of a single crystal is primarily due to *crystallographic* slips on se-

lected slip systems. The plastic rate of deformation tensor $\mathbf{D}^p$ and the plastic spin tensor $\mathbf{W}^p$ of a single crystal can be represented by slip rates associated with slip modes, [49,50]:

$$\mathbf{D}^p = \mathrm{sym}(\mathbf{R}^*\dot{\mathbf{F}}^p\mathbf{F}^{p^{-1}}\mathbf{R}^{*-1}) = \sum_{i=1}^{N^*} \dot{\gamma}_i^*\mathbf{P}_i\,,$$

$$\mathbf{W}^p = \mathrm{skew}(\mathbf{R}^*\dot{\mathbf{F}}^p\mathbf{F}^{p^{-1}}\mathbf{R}^{*-1}) = \sum_{i=1}^{N^*} \dot{\gamma}_i^*\mathbf{Q}_i\,, \qquad (A1)$$

where $\mathbf{R}^*$ is the lattice rigid body rotation tensor, $\mathbf{F}^p$ is the plastic deformation gradient tensor, $\dot{\gamma}_i^*$ is the absolute value of the rate of change of integrated shear strain for the *i*-th crystallographic slip system, and N^* is the total number of the activated crystallographic slip modes. Each slip mode is composed of a slip direction and a slip plane. The tensors $\mathbf{P}_i$ and $\mathbf{Q}_i$ for the *i*th slip mode are defined by

$$\mathbf{P}_i = \mathrm{sym}\,\mathbf{T}_i = \frac{1}{2}(\mathbf{s}_i\otimes\mathbf{m}_i + \mathbf{m}_i\otimes\mathbf{s}_i),$$

$$\mathbf{Q}_i = \mathrm{skew}\,\mathbf{T}_i = \frac{1}{2}(\mathbf{s}_i\otimes\mathbf{m}_i - \mathbf{m}_i\otimes\mathbf{s}_i), \qquad (A2)$$

where the Schmid tensor $\mathbf{T}_i$ is defined by $\mathbf{T}_i = \mathbf{s}_i\otimes\mathbf{m}_i$, and unit vectors $\mathbf{s}_i$ and $\mathbf{m}_i$ are the slip direction and normal to the slip plane associated with the *i*th slip mode in the deformed configuration, respectively. Activation of the selected slip systems can be prescribed by a certain slip condition. The driving force to activate the *i*-th slip system is the resolved shear stress τ_i^* along the slip direction on the crystallographic slip plane of the slip system in the current configuration, which can be obtained by $\tau_i^* = \mathbf{P}_i : \boldsymbol{\sigma}$, where $\boldsymbol{\sigma}$ is the Cauchy stress tensor.

We assume the existence of *a rate-dependent slip potential* τ^* (the effective resolved shear stress) for plastic deformation of single crystal grains in a polycrystalline aggregate with a stress exponent $b(>1)$

$$\tau^*(\tau_1^*, \tau_2^*, \ldots, \tau_{N^*}^*) = [\alpha_1^*|\tau_1^*|^b + \alpha_2^*|\tau_2^*|^b + \ldots + \alpha_{N^*}^*|\tau_{N^*}^*|^b]^{1/b} = \left[\sum_{i=1}^{N^*} \alpha_i^*|\tau_i^*|^b\right]^{1/b}, \qquad (A3)$$

where τ_i^* are resolved shear stresses on the available slip systems of the crystal and α_i^* are the weight coefficients related to the relative strength of the slip systems. A work-conjugate effective shear rate $\dot{\gamma}^*$ and the slip surface can be defined as

$$\tau^*\dot{\gamma}^* = \tau_1^*\dot{\gamma}_1^* + \tau_2^*\dot{\gamma}_2^* + \ldots + \tau_{N^*}^*\dot{\gamma}_{N^*}^* = \sum_{i=1}^{N^*} \tau_i^*\dot{\gamma}_i^*\,, \qquad (A4)$$

$$\tau^*(\tau_1^*, \tau_2^*, \ldots, \tau_{N^*}^*) - \tau_0^*(\xi^*, \dot{\gamma}^*) = 0, \qquad (A5)$$

where ξ^* is a scalar characterizing the overall isotropic hardening of the crystal and $\tau_0^*(\xi^*, \dot{\gamma}^*)$ is the effective slip strength of the crystal. *The associated flow rule* results the slip rate on the each slip system as

$$\dot{\gamma}_i^* = \lambda^*\frac{\partial\tau^*(\tau_1^*, \tau_2^*, \ldots, \tau_{N^*}^*)}{\partial\tau_i^*} = \lambda^*\alpha_i^*\left(\frac{|\tau_i^*|}{\tau^*}\right)^{b-2}\frac{\tau_i^*}{\tau^*} = \dot{\gamma}^*\alpha_i^*\left(\frac{|\tau_i^*|}{\tau^*}\right)^{b-2}\frac{\tau_i^*}{\tau^*} = \dot{\gamma}^*\left(\frac{|\tau_i^*|}{\tau_{i0}^*}\right)^{b-2}\frac{\tau_i^*}{\tau_{i0}^*}, \qquad (A6)$$

where $\lambda^* = \dot{\gamma}^*$ from the plastic work-equivalency requirement (see Eq. (*A*4)), $\alpha_i^* = (\tau_0^*/\tau_{i0}^*)^{b-1}$, and $\tau_{i0}^* = \tau_{i0}^*(g_i^*, \dot{\gamma}^*)$ is the slip strength on each slip system. Evolution laws of isotropic hardening of the single crystal with N^*+1 scalar state variables are given as

$$\dot{\xi}^* = \dot{\xi}^*(\dot{\gamma}_1^*, \dot{\gamma}_2^*, \ldots, \dot{\gamma}_{N^*}^*), \quad \dot{g}_i^* = \dot{g}_i^*(\dot{\gamma}_1^*, \dot{\gamma}_2^*, \ldots, \dot{\gamma}_{N^*}^*). \qquad (A7)$$

When one assumes $b = 1+m$ (with $m>0$) and

$$\tau_0^*(\xi^*, \dot{\gamma}^*) = g^*(\xi^*)\left(\frac{\dot{\gamma}^*}{\dot{\gamma}_0^*}\right)^{1/m} \quad \text{and} \quad \tau_{i0}^*(g_i^*, \dot{\gamma}^*) = g_i^*\left(\frac{\dot{\gamma}^*}{\dot{\gamma}_{i0}^*}\right)^{1/m}, \quad \text{so} \quad \alpha_i^* = \left(\frac{\tau_0^*}{\tau_{i0}^*}\right)^m = \left(\frac{\dot{\gamma}_{i0}^*}{\dot{\gamma}_0^*}\right)\left(\frac{g^*}{g_i^*}\right)^m, \qquad (A8)$$

one can show that

$$\dot{\gamma}_i^* = \dot{\gamma}^*\left(\frac{|\tau_i^*|}{\tau_{i0}^*}\right)^{m-1}\frac{\tau_i^*}{\tau_{i0}^*} = \dot{\gamma}_{i0}^*\left(\frac{|\tau_i^*|}{g_i^*}\right)^{m-1}\frac{\tau_i^*}{g_i^*}. \qquad (A9)$$

This is exactly the rate-dependent slip rule that has been proposed by Hutchinson [51], Asaro [49], and Asaro and Needleman [52]. A combined self-hardening and latent hardening model can be adapted here with $\xi^* = \gamma_T^*$

$$\dot{g}^* = h^*(\gamma_T^*)\dot{\gamma}_T^*\,, \quad \dot{g}_i^* = \sum_{j=1}^{N^*} h_{ij}^*\dot{\gamma}_j^*\,, \qquad (A10)$$

where the Taylor strain γ_T^* and the hardening moduli matrix h_{ij}^* are defined by

$$\dot{\gamma}_T^* = \sum_{i=1}^{N^*} |\dot{\gamma}_i^*|, \quad h_{ij}^* = qh^* + (1-q)h^*\delta_{ij}\,, \qquad (A11)$$

where q is a parameter characterizing the latent hardening. When the latent hardening is equal to the self-hardening ($q=1$), one has the Taylor isotropic hardening model of single crystals with $h_{ij}^*(\gamma_T^*) = h^*(\gamma_T^*)$, $\alpha_i^* = 1$, $\dot{\xi}^* = \dot{\gamma}_T^*$, and $g_i^* = g^*$. On the other hand, the current crystal plasticity model is a rate-dependent extension of the rate-independent models proposed by Gambin [53] and Darrieulat and Piot [54]. The stress exponent b has been identified by them respectively as either the interaction exponent of slip systems in a single crystal or the texture dispersion exponent in a polycrystal. This crystal plasticity model with an associated rate-dependent slip potential is more flexible as in general $b \neq 1+m$ or $m \neq \infty$ (where m is a parameter in a simple power-law rate-dependence model, see Eq. (*A*8)).

Under uniaxial tension, the resolved shear stresses in terms of the uniaxial tensile stress $\sigma_\theta > 0$ and the slip system vectors (slip direction and the slip plane normal in terms of the in-plane loading orientation angle θ) are given as

$$\tau_i^* = s_{i1}^* m_{i1}^* \sigma_\theta = L_{i1}^*\sigma_\theta = [s_{i1}^{0*} m_{i1}^{0*}\cos^2\theta + (s_{i1}^{0*} m_{i2}^{0*} + s_{i2}^{0*} m_{i1}^{0*})\sin\theta\cos\theta + s_{i2}^{0*} m_{i2}^{0*}\sin^2\theta]\sigma_\theta\,, \qquad (A12)$$

where $(s_{i1}^{0*}, s_{i2}^{0*}, s_{i3}^{0*})$ and $(m_{i1}^{0*}, m_{i2}^{0*}, m_{i3}^{0*})$ are the Cartesian components of the slip system vectors defined in the material texture coordinate system XYZ. The slip potential Eq. (*A*3) can be expressed in terms of the uniaxial tensile stress and the in-plane loading orientation angle as

$$[\tau^*(\sigma_\theta,\theta)]^b=\sigma_\theta^b\sum_{i=1}^{N^*}\alpha_i^*|L_{i1}^*|^b$$

$$=\sigma_\theta^b\sum_{i=1}^{N^*}\alpha_i^*|s_{i1}^{0*}m_{i1}^{0*}\cos^2\theta+(s_{i1}^{0*}m_{i2}^{0*}+s_{i2}^{0*}m_{i1}^{0*})\times\sin\theta\cos\theta+s_{i2}^{0*}m_{i2}^{0*}\sin^2\theta|^b$$

$$\equiv\sigma_\theta^b\Phi_1^*(\theta). \tag{A13}$$

The in-plane components of the plastic rate of deformation tensor $\mathbf{D}^p$ and the plastic spin tensor $\mathbf{W}^p$ of a single crystal can be obtained using Eqs. (*A*1), (*A*2), (*A*6), (*A*12), and (A13)

$$\dot{\varepsilon}_1=\sum_{i=1}^{N^*}s_{i1}^*m_{i1}^*\dot{\gamma}_i^*=\dot{\gamma}^*\left(\frac{\sigma_\theta}{\tau_0}\right)^{b-1}\sum_{i=1}^{N^*}\alpha_i^*|L_{i1}^*|^b$$

$$\equiv\dot{\gamma}^*\left(\frac{\sigma_\theta}{\tau_0}\right)^{b-1}\Phi_1^*(\theta), \tag{A14a}$$

$$\dot{\varepsilon}_2=\sum_{i=1}^{N^*}s_{i2}^*m_{i2}^*\dot{\gamma}_i^*=\dot{\gamma}^*\left(\frac{\sigma_\theta}{\tau_0}\right)^{b-1}\sum_{i=1}^{N^*}\alpha_i^*|L_{i1}^*|^{b-2}L_{i1}^*L_{i2}^*$$

$$\equiv\dot{\gamma}^*\left(\frac{\sigma_\theta}{\tau_0}\right)^{b-1}\Phi_2^*(\theta), \tag{A14b}$$

$$\dot{\varepsilon}_{12}=\sum_{i=1}^{N^*}\frac{s_{i1}^*m_{i2}^*+s_{i2}^*m_{i1}^*}{2}\dot{\gamma}_i^*=\dot{\gamma}^*\left(\frac{\sigma_\theta}{\tau_0}\right)^{b-1}\sum_{i=1}^{N^*}\alpha_i^*|L_{i1}^*|^{b-2}L_{i1}^*V_{i1}^*$$

$$\equiv\dot{\gamma}^*\left(\frac{\sigma_\theta}{\tau_0}\right)^{b-1}\Phi_3^*(\theta), \tag{A14c}$$

$$\dot{\omega}_{12}=\sum_{i=1}^{N^*}\frac{s_{i1}^*m_{i2}^*-s_{i2}^*m_{i1}^*}{2}\dot{\gamma}_i^*=\dot{\gamma}^*\left(\frac{\sigma_\theta}{\tau_0}\right)^{b-1}\sum_{i=1}^{N^*}\alpha_i^*|L_{i1}^*|^{b-2}L_{i1}^*V_{i2}^*$$

$$\equiv\dot{\gamma}^*\left(\frac{\sigma_\theta}{\tau_0}\right)^{b-1}\Phi_4^*(\theta), \tag{A14d}$$

where

$$L_{i1}^*=s_{i1}^*m_{i1}^*=s_{i1}^{0*}m_{i1}^{0*}\cos^2\theta+(s_{i1}^{0*}m_{i2}^{0*}+s_{i2}^{0*}m_{i1}^{0*})\sin\theta\cos\theta+s_{i2}^{0*}m_{i2}^{0*}\sin^2\theta,$$

$$L_{i2}^*=s_{i2}^*m_{i2}^*=s_{i1}^{0*}m_{i1}^{0*}\sin^2\theta-(s_{i1}^{0*}m_{i2}^{0*}+s_{i2}^{0*}m_{i1}^{0*})\sin\theta\cos\theta+s_{i2}^{0*}m_{i2}^{0*}\cos^2\theta,$$

$$V_{i1}^*=\frac{s_{i1}^*m_{i2}^*+s_{i2}^*m_{i1}^*}{2}=\frac{s_{i2}^{0*}m_{i2}^{0*}-s_{i1}^{0*}m_{i1}^{0*}}{2}\sin 2\theta+\frac{s_{i1}^{0*}m_{i2}^{0*}+s_{i2}^{0*}m_{i1}^{0*}}{2}\cos 2\theta,$$

$$V_{i2}^*=\frac{s_{i1}^*m_{i2}^*-s_{i2}^*m_{i1}^*}{2}=\frac{s_{i1}^{0*}m_{i2}^{0*}-s_{i2}^{0*}m_{i1}^{0*}}{2},\quad\text{and}$$

$$\Phi_3^*(\theta)=\frac{1}{2b}\frac{d\Phi_1^*(\theta)}{d\theta}\ \text{(the associated flow rule)}.$$

The material anisotropic functions $\Phi_1^*(\theta)$, $\Phi_2^*(\theta)$, and $\Phi_4^*(\theta)$ of the single crystal under uniaxial tension are related only to its slip system vectors $\mathbf{s}_i$ and $\mathbf{m}_i$, the slip system weight coefficient α_i^*, and the stress exponent b. By using the Sachs assumption of the uniform stress field in each single crystal grain of a polycrystal, the macroscopic flow potential and the associated flow rules for plastic strain rates and the flow rule for plastic spin can be obtained approximately via simple volume averaging. The resulting mathematical formulation of these constitutive equations is identical to the one presented in Section 2 if $a=b$.

The plastic spin appeared in the *micromechanical* theory of single crystal plasticity (see Eq. (*A*1) in general and Eq. (14*d*) under uniaxial tension) is the natural consequence of the kinematics of crystallographic slips. When one does not directly measure the crystal orientations using either X-ray or electron diffraction techniques, one does not know explicitly the slip system vectors $\mathbf{s}_i$ and $\mathbf{m}_i$. One has to rely on the mechanical tests instead to evaluate the material anisotropic functions $\Phi_1^*(\theta)$, $\Phi_2^*(\theta)$, and $\Phi_4^*(\theta)$ of the single crystal under uniaxial tension. In other words, this results in a *macroscopic* theory of single crystal plasticity using these material anisotropic functions for uniaxial tensile loading. One can show that in general the material anisotropic function $\Phi_4^*(\theta)$ cannot be deduced from the knowledge of other two anisotropic material anisotropic functions $\Phi_1^*(\theta)$ and $\Phi_2^*(\theta)$ for single crystals: that is, the knowledge of the orientational dependence of flow stress and plastic strain ratio under uniaxial tension will not provide any prediction on the orientational dependence of plastic spin ratio at all. Indeed, this result directly contradicts one of the commonly cited expressions for macroscopic plastic spin, [26–28]:

$$\mathbf{W}^p=\eta(\boldsymbol{\sigma}\mathbf{D}^p-\mathbf{D}^p\boldsymbol{\sigma}),\quad\text{or}\quad\dot{\omega}_{12}=\eta\sigma_\theta\dot{\varepsilon}_{12}\ \text{(under uniaxial tension)}, \tag{A15}$$

where η is the plastic spin coefficient (which has been assumed to be a constant in its application examples). Kurroda [26] and Dafalias [28] found that the plastic spin coefficient have to be a negative constant for modeling certain experimental data while it is required to be non-negative according to Levitas [27]. One can show that the relation $\dot{\omega}_{12}=\eta\sigma_\theta\dot{\varepsilon}_{12}$ (uniaxial tension) does not hold for single crystals in general according to the micromechanical crystal plasticity model presented above (numerical simulations of single crystals have revealed that there exist cases that $\dot{\omega}_{12}$ is nonzero when $\dot{\varepsilon}_{12}=0$!) and the general validity of the constitutive equation for macroscopic plastic spin Eq. (*A*15) is thus questionable. A similar conclusion can also be reached about another proposed simple expression for the macroscopic plastic spin, [28], $\dot{\omega}_{12}=\eta\dot{\varepsilon}_{xy}=\eta[(\dot{\varepsilon}_1-\dot{\varepsilon}_2)\sin\theta\cos\theta+\dot{\varepsilon}_{12}\cos 2\theta]$ as it holds only strictly for single slips for single crystals.

References

[1] Hill, R., 1950, *The Mathematical Theory of Plasticity*, Clarendon Press, Oxford, UK.

[2] Prager, W., 1955, "The Theory of Plasticity: A Survey of Recent Achievements," Proc. Inst. Electr. Eng., **169**, pp. 41–57.

[3] Eisenberg, M. A., and Yen, C. F., 1981, "A Theory of Multiaxial Anisotropic Viscoplasticity," ASME J. Appl. Mech., **48**, pp. 276–284.

[4] McDowell, D. L., 1987, "An Evaluation of Recent Developments in Hardening and Flow Rules for Rate-Independent, Non-Proportional Cyclic Plasticity," ASME J. Appl. Mech., **54(2)**, pp. 323–334.

[5] Hill, R., 1948, "A Theory of the Yielding and Plastic Flow of Anisotropic Metals," Proc. R. Soc. London, Ser. A, **193**, pp. 281–297.

[6] Hill, R., 1979, "Theoretical Plasticity of Textured Aggregates," Math. Proc. Cambridge Philos. Soc., **85**, pp. 179–191.

[7] Karafillis, A. P., and Boyce, M. C., 1993, "A General Anisotropic Yield Criterion Using Bounds and a Transformation Weighting Tensor," J. Mech. Phys. Solids, **41**, pp. 1859–1886.

[8] Wu, H. C., Hong, H. K., and Shiao, Y. P., 1999, "Anisotropic Plasticity With Application to Sheet Metals," Int. J. Mech. Sci., **41**, pp. 703–724.

[9] Dawson, P., and Marin, E. B., 1998, "Computational Mechanics for Metal Deformation Process Using Polycrystal Plasticity," Adv. Appl. Mech., **34**, pp. 77–169.

[10] Boehler, J. P., and Koss, S., 1991, "Evolution of Anisotropy in Sheet-Steels Subjected to Off-Axes Large Deformation," *Advances in Continuum Mechanics*, O. Bruller, V. Mannl, and J. Najar, eds., Springer, Berlin, pp. 143–158.

[11] Kim, K. H., and Yin, J. J., 1997, "Evolution of Anisotropy Under Plane Stress," J. Mech. Phys. Solids, **45**(5), pp. 841–851.

[12] Bunge, H. J., and Nielsen, I., 1997, "Experimental Determination of Plastic Spin in Polycrystalline Materials," Int. J. Plast., **13**(5), pp. 435–446.

[13] Mandel, J., 1974, "Thermodynamics and Plasticity," *Foundations of Continuum Thermodynamics*. J. J. Delgado et al., eds., MacMillan, New York.

[14] Loret, B., 1983, "On the Effects of Plastic Rotation in the Finite Deformation

of Anisotropic Elastoplastic Materials,'' Mech. Mater., **2**, pp. 287–304.
[15] Dafalias, Y. F., 1984, ''The Plastic Spin Concept and a Simple Illustration of Its Role in Finite Plastic Transformations,'' Mech. Mater., **3**, pp. 223–233.
[16] Onat, E. T., 1984, ''Shear Flow of Kinematically Hardening Rigid-Plastic Materials,'' *Mechanics of Material Behavior*, G. J. Dvorak and R. T. Shield, eds., Elsevier, New York, pp. 311–324.
[17] Dafalias, Y. F., 1985, ''The Plastic Spin,'' ASME J. Appl. Mech., **52**, pp. 865–871.
[18] Dafalias, Y. F., and Rashid, M. M., 1989, ''The Effect of Plastic Spin on Anisotropic Material Behavior,'' Int. J. Plast., **5**, pp. 227–246.
[19] Dafalias, Y. F., and Aifantis, E. C., 1990, ''On the Microscopic Origin of the Plastic Spin,'' Acta Mech., **82**, pp. 31–48.
[20] Aravas, E. C., and Aifantis, E. C., 1991, ''On the Geometry of Slip and Spin in Finite Plastic Deformation,'' Int. J. Plast., **7**, pp. 141–160.
[21] van der Giessen, E., 1992, ''A 2D Analytical Multiple Slip Model for Continuum Texture Development and Plastic Spin,'' Mech. Mater., **13**, pp. 93–115.
[22] Prantil, V. C., Jenkins, J. T., and Dawson, P. R., 1993, ''An Analysis of Texture and Plastic Spin for Planar Polycrystals,'' J. Mech. Phys. Solids, **41**(8), pp. 1357–1382.
[23] Dafalias, Y. F., 1993, ''On Multiple Spins and Texture Development. Case Study: Kinematic and Orthotropic Hardening,'' Acta Mech., **100**, pp. 171–194.
[24] Lubarda, V. A., and Shih, C. F., 1994, ''Plastic Spin and Related Issues in Phenomenological Plasticity,'' ASME J. Appl. Mech., **61**, pp. 524–529.
[25] Schieck, B., and Stumpf, H., 1995, ''The Appropriate Corotational Rate, Exact Formula for the Plastic Spin and Constitutive Model for Finite Elastoplasticity,'' Int. J. Solids Struct., **32**(24), pp. 3643–3667.
[26] Kuroda, M., 1997, ''Interpretation of the Behavior of Metals Under Large Plastic Shear Deformations: A Macroscopic Approach,'' Int. J. Plast., **13**, pp. 359–383.
[27] Levitas, V. I., 1998, ''A New Look at the Problem of Plastic Spin Based on Stability Analysis,'' J. Mech. Phys. Solids, **46**(3), pp. 557–590.
[28] Dafalias, Y. F., 2000, ''Orientational Evolution of Plastic Orthotropy in Sheet Metals,'' J. Mech. Phys. Solids, **48**, pp. 2231–2255.
[29] Sidoroff, F., and Dogui, A., 2001, ''Some Issues About Anisotropic Elastic-Plastic Models at Finite Strain,'' Int. J. Solids Struct., **38**, pp. 9569–9578.
[30] Kuroda, M., and Tvergaard, V., 2001, ''Plastic Spin Associated With a Non-Normality Theory of Plasticity,'' Eur. J. Mech. A/Solids, **20**, pp. 893–905.
[31] Lee, E. H., 1969, ''Elastic-Plastic Deformations at Finite Strains,'' ASME J. Appl. Mech., **36**, pp. 1–6.
[32] Rice, J. R., 1970, ''On the Structure of Stress-Strain Relations for Time-Dependent Plastic Deformation in Metals,'' ASME J. Appl. Mech., **37**, pp. 728–737.
[33] Rice, J. R., 1971, ''Inelastic Constitutive Relations for Solids: An Internal-Variable Theory and Its Application to Metal Plasticity,'' J. Mech. Phys. Solids, **19**, pp. 433–455.
[34] Rogers, T. G., 1990, ''Yield Criteria, Flow Rules, and Hardening in Anisotropic Plasticity,'' *Yielding, Damage, and Failure of Anisotropic Solids*, J. P. Boehler, ed., Mechanical Engineering Publications Limited, London, pp. 53–79.
[35] Tong, W., 2002, ''A Planar Plastic Flow Theory of Orthotropic Sheets and the Experimental Procedure for Its Evaluations,'' Proc. R. Soc. London, Ser. A, submitted for publication.
[36] Tong, W., 2002, ''A Plane Stress Anisotropic Plastic Flow Theory for Orthotropic Aluminum Sheet Metals,'' Int. J. Plast., accepted for publication.
[37] Tong, W., 2003, ''A Planar Plastic Flow Theory for Monoclinic Sheet Metals,'' Int. J. Mech. Sci., submitted for publication.
[38] Tong, W., Zhang, N., and Xie, C., 2003, ''Modeling of the Anisotropic Plastic Flows of Automotive Sheet Metals,'' *Aluminum 2003*, S. K. Das, ed., The Minerals, Metals & Materials Society, to appear.
[39] Tong, W., Xie, C., and Zhang, N., 2003, ''Micromechanical and Macroscopic Modeling of Anisotropic Plastic Flows of Textured Polycrystalline Sheets,'' Modell. Simul. Mater. Sci. Eng., submitted for publication.
[40] Tong, W., Xie, C., and Zhang, N., 2003, ''Modeling the Anisotropic Plastic Flow of Textured Polycrystalline Sheets Using the Generalized Hill's 1979 Non-Quadratic Flow Potential,'' Int. J. Plast., submitted for publication.
[41] Hill, R., 1980, ''Basic Stress Analysis of Hyperbolic Regimes in Plastic Media,'' Math. Proc. Cambridge Philos. Soc., **88**, pp. 359–369.
[42] Hill, R., 1990, ''Constitutive Modeling of Orthotropic Plasticity in Sheet Metals,'' J. Mech. Phys. Solids, **38**, pp. 405–417.
[43] Truong Qui, H. P., and Lippmann, H., 2001, ''Plastic Spin and Evolution of an Anisotropic Yield Condition,'' Int. J. Mech. Sci., **43**, pp. 1969–1983.
[44] Truong Qui, H. P., and Lippmann, H., 2001, ''On the Impact of Local Rotation on the Evolution of an Anisotropic Plastic Yield Condition,'' J. Mech. Phys. Solids, **49**, pp. 2577–2591.
[45] Losilla, G., Boehler, J. P., and Zheng, Q. S., 2000, ''A Generally Anisotropic Hardening Model for Big Offset-Strain Yield Stresses,'' Acta Mech., **144**, pp. 169–183.
[46] McDowell, D. L., Miller, M. P., and Bammann, D. J., 1993, ''Some Additional Considerations for Coupling of Material and Geometric Nonlinearities for Polycrystalline Metals,'' *Large Plastic Deformations: Fundamental Aspects and Applications to Metal Forming* (MECAMAT'91), C. Teodosiu, J. L. Raphanel, and F. Sidoroff, eds., Balkema, Rotterdam, pp. 319–327.
[47] Stoughton, T. B., 2002, ''A Non-Associated Flow Rule for Sheet Metal Forming,'' Int. J. Plast., **18**, pp. 687–714.
[48] Dafalias, Y. F., 1998, ''The Plastic Spin: Necessity or Redundancy?'' Int. J. Plast., **14**, pp. 909–931.
[49] Asaro, R. J., 1983, ''Micromechanics of Crystals and Polycrystals,'' Advances in Mechanics, **23**, p. 1.
[50] Bassani, J. L., 1994, ''Plastic Flow of Crystals,'' Advances in Appl. Mech., **30**, pp. 191–258.
[51] Hutchinson, J. W., 1976, ''Bounds and Self-Consistent Estimates for Creep of Polycrystalline Materials,'' Proc. R. Soc. London, Ser. A, **348**, pp. 101–127.
[52] Asaro, R. J., and Needleman, A., 1985, ''Texture Development and Strain Hardening in Rate-Dependent Polycrystals,'' Acta Metall., **33**(6), pp. 923–953.
[53] Gambin, W., 1992, ''Refined Analysis of Elastic-Plastic Crystals,'' Int. J. Solids Struct., **29**(16), pp. 2013–2021.
[54] Darrieulat, M., and Piot, D., 1996, ''A Method of Generating Analytical Yield Surfaces of Polycrystalline Materials,'' Int. J. Plast., **12**(5), pp. 575–610.

T. Iwasa
Graduate Student of the University of Tokyo,
Department of Aeronautics and Astronautics
e-mail: taka@taurus.eng.isas.jaxa.jp

M. C. Natori
Professor,
Mem. ASME

K. Higuchi
Associate Professor,
Mem. ASME

Division of Spacecraft Engineering,
The Institute of Space and Astronautical Science,
3-1-1 Yoshinodai Sagamihara,
Kanagawa 229-8510, Japan

Evaluation of Tension Field Theory for Wrinkling Analysis With Respect to the Post-Buckling Study

A comparative study of two representative wrinkling theories, a bifurcation theory and a tension field theory, is carried out for quantitative evaluation of the tension field theory regarding wrinkling analysis. Results obtained from the bifurcation theory show the limitations of tension field theory on the wrinkling analysis. Existence of compressive stresses caused by wrinkling phenomena, which is not accounted for tension field theory, is quantitatively presented. Considering strain energy due to these compressive stresses and geometrical boundary constraints, it is clarified that there are regions, in which the tension field theory is not properly applied. [DOI: 10.1115/1.1767171]

Introduction

Large membrane structures are proposed for future space missions, [1–3], because they are lightweight and have excellent encapsulation characteristics. Solar sails, sunshields, and solar power arrays are representative of these future space structures. However, since a membrane has negligible bending stiffness, wrinkling occurs easily once the membrane is subjected to compressive stresses. These phenomena deteriorate the operating abilities of membrane space structures. For that reason, many studies on wrinkling phenomena for these structures have been reported, [4–18].

Currently, conventional studies on wrinkling phenomena use tension field theory, in which it is assumed that the membrane has no bending stiffness and can carry no compressive stress. From these assumptions, wrinkling of initially flat membranes are treated as in-plane problems through the special definition in which one principal stress must be zero and the other nonzero in wrinkled regions. Therefore, in tension field theory, only wrinkled regions and directions of wrinkles are obtained, which are indicators of the overall behaviors of wrinkled membranes.

However, in actual situations, wrinkling phenomena are due to bifurcation and are intensely affected by small bending stiffness. For that reason, if wrinkling phenomena are to be investigated in detail, a geometrically nonlinear analysis based on the bifurcation theory must be carried out. In bifurcation theory, detailed wrinkling behaviors, whose indices are number and amplitude of wrinkles, can be clarified, while in tension field theory this kind of quantitative evaluation cannot be obtained. This is because the effects of structural scale (ratio of length to thickness) on wrinkling phenomena cannot be treated in tension field theory. Therefore, clarifying limitations for the use of tension field theory for wrinkling analysis becomes important in order to accurately understand the meaning of conventional solutions calculated using the tension field theory.

In this research, a comparative study on wrinkling analysis using bifurcation theory and tension field theory is carried out. The wrinkling of a stretched circular membrane with a center rotation hub is investigated. The effect of the structural scale (ratio of length to thickness) on the wrinkled membrane and the limitations for use of the tension field theory in wrinkling analysis are investigated with respect to the post-buckling study. The regions in which tension field theory is not applicable are discussed considering the actual stress state of the wrinkled membrane.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, May 1, 2003; final revision, Jan. 30, 2004. Associate Editor: Z. Suo. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California-Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

Numerical Approach

Analytical Model. Figure 1 shows an analytical model used in this paper. Wrinkling phenomena, which appear after rotating a center hub attached to the stretched circular membrane with some torque force, are investigated. In this analysis the torque on the rotation hub is applied after the circular membrane is stretched by an initial uniform tensile stress. This analytical modeling procedure is typical of other wrinkling studies, [11–18]. In this paper, the ratio of the radius of the circular membrane to that of the hub is set to constant $(b/a=6)$, although the deformation modes of the wrinkled membranes are affected by this ratio, [19,20]. We wanted to investigate the effects of the structural scale (length of radial cross section/membrane thickness) on the wrinkled membranes focusing on the membrane thickness. They are the same for the different values of b/a. It is assumed that the membrane is made from a polyester film, and is linearly elastic. Table 1 lists the numerical parameters used in the analysis.

Tension Field Theory. The tension field theory treated in this paper is based on Stein-Hedgepeth theory [11], in which the surface of a wrinkled membrane is divided into three states indicated by

$$\sigma_1>0, \quad \sigma_2>0\cdots\text{taut region} \tag{1}$$

$$\sigma_1>0, \quad \sigma_2\leqslant 0\cdots\text{wrinkle region} \tag{2}$$

$$\sigma_1\leqslant 0, \quad \sigma_2\leqslant 0\cdots\text{slack region} \tag{3}$$

where σ_1 and σ_2 are the principal stresses. From these equations, the following equation is always satisfied in the wrinkled regions.

$$\sigma_x\cdot\sigma_y=\tau_{xy}^2 \tag{4}$$

On the other hand, the equilibrium equations in taut regions are the same as those on treated in conventional plane stress problems. Therefore, the solutions based on tension field theory are given by solving two equations, which are Eq. (4) in the wrinkled regions and the conventional equilibrium equation in taut regions.

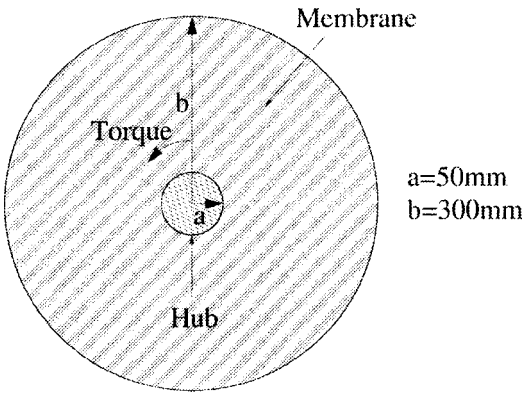


Fig. 1 Analytical model

At the same time, an appropriate material parameter, which is Young's modulus or Poisson's ratio, is defined to be variable in the wrinkled regions.

For analysis based on Stein-Hedgepeth theory, a variable Poisson's ratio is used in wrinkled regions, and partial wrinkles on the stretched circular membrane are investigated, [11]. The validity of this wrinkling theory is confirmed by Mikulas' experimental study, [12]. As a result, analytical solutions based on tension field theory presented in this paper are referred to ''Mikulas' solution.''

Geometrically Nonlinear Analysis

To investigate wrinkling behavior in detail, a geometrically nonlinear analysis for studying on post-buckling phenomena is carried out. Figure 2 shows a finite element mesh model treated in this analysis. This mesh model has equivalent radial and circumferential divisions. In order to include the effect of small bending stiffness in membranes, MITC (mixed interpolation tensorial components) shell elements, [21–25], are used. This element has a proven predictive capability for numerical results related to distortion of shell elements associated with finite deformation in nonlinear analysis. We consider that this element is most efficient for wrinkling analysis when displacements become very large compared to the thickness.

Table 1 Numerical parameters

Material	Unit	Polyester Film
Young's modulus	[MPa]	5723
Radius of a membrane	[mm]	300
Radius of a hub	[mm]	50
Membrane thickness	[μm]	75, 100, 175, 300, 500, 1000, 2500

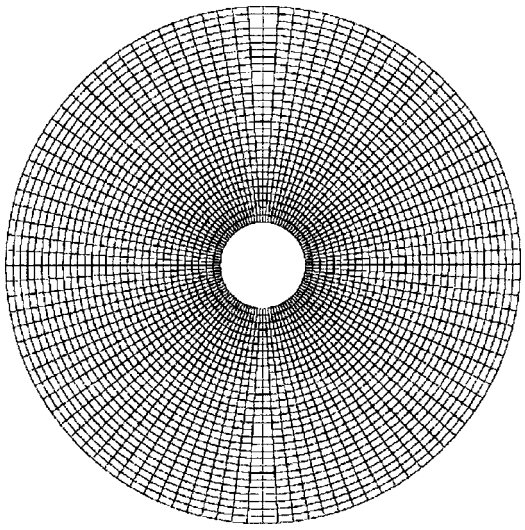


Fig. 2 Finite element mesh

Table 2 Initial uniform tensile stress and rotation angle

Initial uniform tensile stress	2.938 Mpa (0.108 mm)
Rotating angle	1.145 deg (1.0 mm)

The geometrically nonlinear analysis in this study is performed using the freely available finite element analysis program FEAPpv (finite element analysis program personal version), [26]. However, since FEAPpv does not include MITC shell elements, the authors have incorporated MITC shell elements into FEAPpv through the user-defined element function of the program. In the formulation of this element, in order to improve the accuracy of numerical results, a finite rotation increment of directors is considered in the calculation of tangent stiffness matrices, [27].

In the case of a geometrically nonlinear analysis using shell elements for the post-buckling study in wrinkling phenomena, numerical results are intensely affected by mesh aspect ratio, step ratio of numerical analysis, and shape functions of the shell element. Therefore, we have investigated these effects on wrinkling analysis focusing on the deformation modes of a wrinkled membrane, [16]. From the study, it turned out that the deformation modes converged if a radial mesh aspect ratio (radial mesh width/membrane thickness)$<$50, the circumferential mesh aspect ratio (circumferential mesh width/membrane thickness)$<$20, and the step ratio of a numerical analysis (tangential increment displacement of the center rotation hub/circumferential mesh width) $<$0.001 were satisfied. And the effect of the shape function of shell elements, which are those of 4-nodes and 9-nodes MITC shell elements, on wrinkling analysis is inconsequential. In this paper, all results are given through numerical calculations, which use the appropriate mesh aspect ratio and a step ratio.

For the post-buckling study regarding wrinkling phenomena, initial imperfections with normal random numbers in the out-of-plane direction are installed at all nodal points on the circular membrane, with magnitude ranging from 0.1 to 0.01 of the membrane thickness. Using these initial imperfections we did the post-buckling analysis without doing any complicated perfect bifurcation analysis. Since wrinkling phenomena are affected by the initial imperfections, the effects of the initial imperfections on wrinkling behavior have been investigated by the authors, [16]. From the investigation, it is confirmed that the effects of initial imperfections with normal random number on wrinkling behavior do not appear when these magnitude are within 0.1 of the membrane thickness.

The initial uniform tensile stress is provided by a given radial displacement, which is applied at all nodal points on the outer circumference of the circular membrane. A given tangential displacement is applied at all nodal points on the inner circumference. Table 2 shows the initial uniform tensile stress and the rotation angle of the hub. Values in parentheses indicate the displacements used in this analysis. These are the same values used in previous studies, [14–18]. And also linear stress-strain relations are considered, since the rotation angle used in the analysis is assumed to be small.

Results and Discussion

Effect of Structural Scale on Wrinkling Phenomena. Figures 3 and 4 show bird-eye views of wrinkled membranes magnified by ten times in the out-of-plane direction. These figures show the effects of structural scale (length of radial cross section/membrane thickness) on the deformation modes of wrinkled membranes. Figure 5 shows dependency of the number of wrinkles on the structural scale. From this figure, the number of winkles gradually increases as the structural scale becomes large.

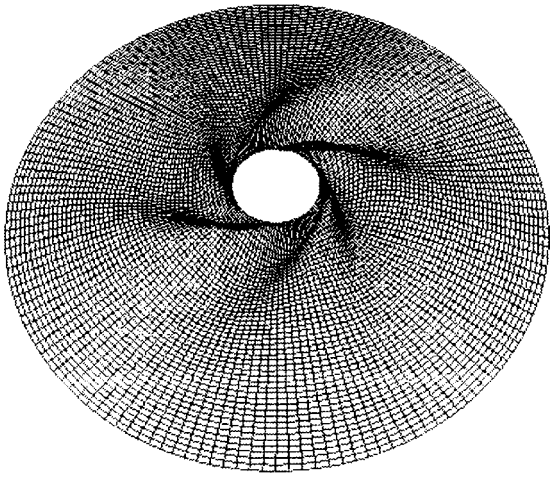


Fig. 3 Bird-eye view of a wrinkled membrane (structural scale: 500)

This tendency occurs because the bending stiffness becomes small with increasing structural scale. In the membrane with initial tensile stresses, the deformation mode is intensely affected by the bending stiffness, although this tendency do not appear in the case of a plate problem, which does not include the effects of initial tensile stresses, [19,20].

Figure 6 shows nondimensional loading curves up to the rotation angle of 1.145 deg. In this figure, the solution obtained from tension field theory (Mikulas' solution) is also presented. From these results, it is seen that torsional rigidity of the membrane decreases as the structural scale increases. This occurs because these results are nondimensionalized with respect to membrane thickness, but not with respect to bending stiffness. When the structural scale becomes large, bifurcation points approach the solution obtained from tension field theory. On the other hand, the gradient of the curves (torsional rigidity) do not coincide with those of tension field theory. The reason for this is clarified later in the subsection titled ''Note on the Actual Mechanism of Wrinkling Behavior.'' Distribution of the out-of-plane displacement ratio (out-of-plane displacement/membrane thickness) is shown in Fig. 7. In this figure, the abscissa indicates the dimensionless radial location, and the ordinate indicates the out-of-plane displacement ratio, which is the average value of the absolute out-of-plane displacement (average amplitude of wrinkles) at the same circumferential locations divided by the membrane thickness. This figure shows that the out-of-plane displacement ratio becomes large and the location having maximum out-of-plane displacement approaches the center hub with increasing structural scale. Figures 8 and 9 indicate these behaviors, quantitatively. Figure 8 shows the maximum amplitude ratio of wrinkles. In this research, maximum amplitude ratio of wrinkles is defined as the maximum value of the displacements shown in Fig. 7. From Fig. 8, it also turns out that the maximum amplitude ratio of wrinkles is increased when the structural scale becomes large. Figure 9 shows the radial location at which maximum out-of-plane displacement occurs. It approaches the rotation hub as structural scale becomes large.

The distribution of minor principal stress ratios in the wrinkled membrane is shown in Fig. 10. Here, the minor principal stress ratio indicates the average value of the minor principal stress at the same circumferential locations divided by the initial tensile stress. This figure shows that the compressive stress decreases as structural scale becomes large. Figure 11 shows the wrinkled re-

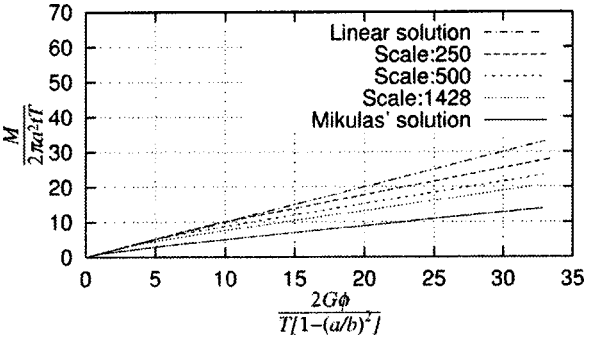


Fig. 6 Nondimensional $M-\phi$ curves

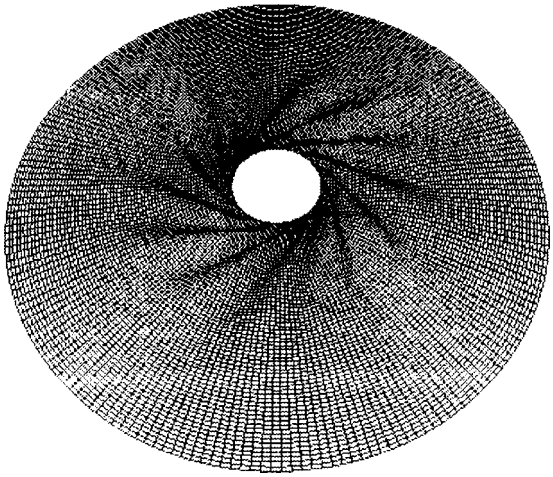


Fig. 4 Bird-eye view of a wrinkled membrane (structural scale: 2500)

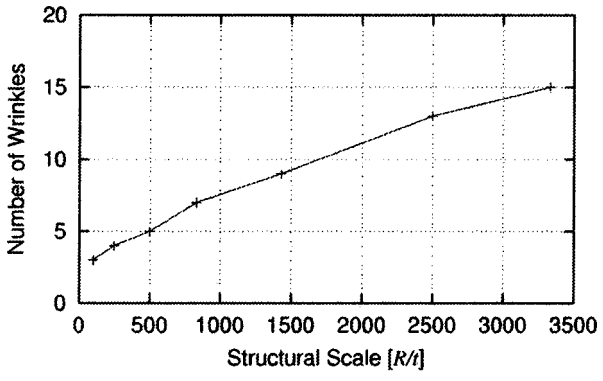


Fig. 5 Number of wrinkles

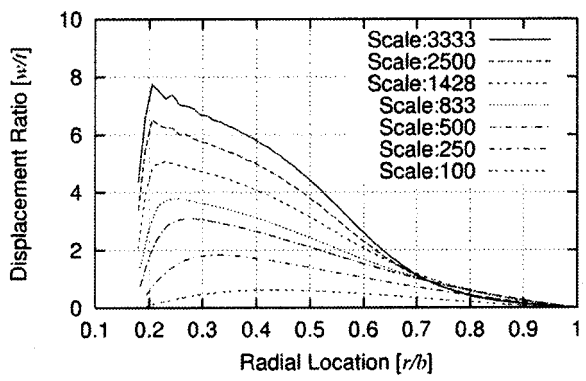


Fig. 7 Out-of-plane displacement ratio of wrinkles

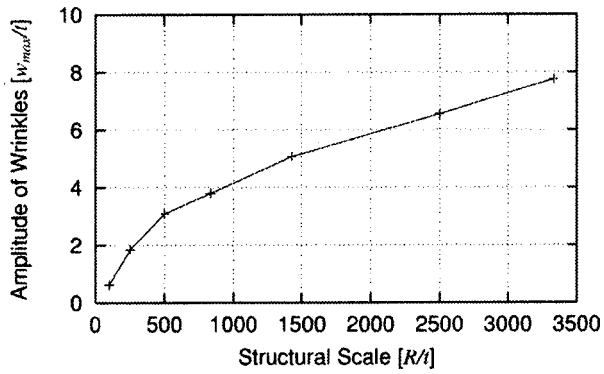


Fig. 8 Maximum amplitude ratio of wrinkles

gion of the membrane, which is indicated by locations in the radial direction having compressive stresses to the radius of membrane. In this study, we assume the wrinkled region to be that with negative minor principal stresses. In tension field theory, this ratio is a constant value because it does not depend on membrane thickness. As can be seen from this figure, numerical results from the geometrically nonlinear analysis are comparable to solutions obtained based on tension field theory, when the structural scale becomes large, or when the converged wrinkled region ratio reaches the constant value of 0.58.

In the wrinkled region shown in Fig. 10 ($r/b<0.58$), there are two wrinkled regions, one with large compressive stress and another with small compressive stress. We can assume that the location and stress, which divide these two regions, are 0.33 and -0.5 MPa, respectively, when the structural scale is larger than 500. Under these values, minor principal stress ratio begins to decrease very much.

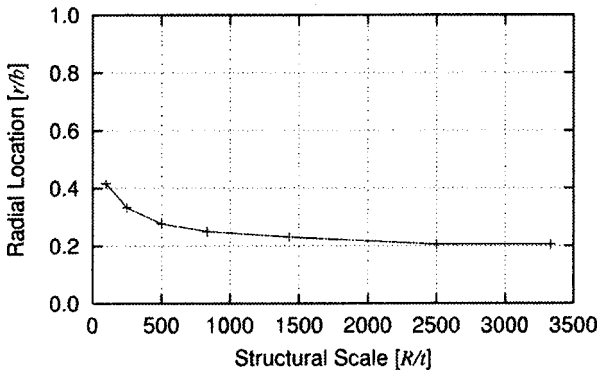


Fig. 9 Location of maximum amplitude of wrinkles

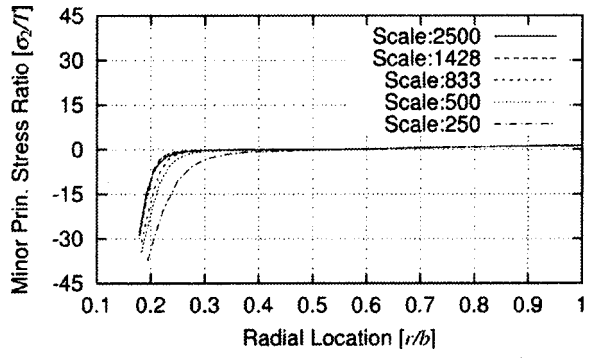


Fig. 10 Minor principal stress

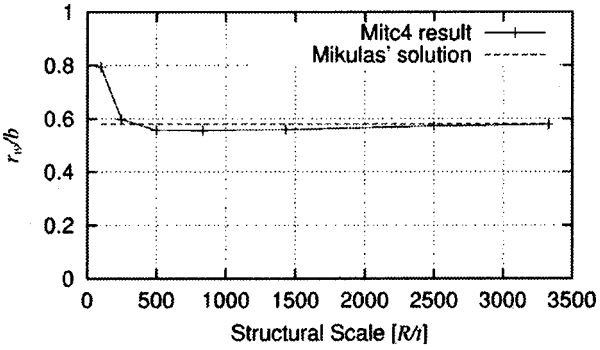


Fig. 11 Wrinkled region, $0<r<r_w$

Existence of Wrinkles in Wrinkled Membrane With No Compressive Stress

Although the location of the wrinkled region calculated based on tension field theory is 0.58, Fig. 7 shows that wrinkles also form outside of this region from 0.58 to 1.0 as shown in the abscissa of Fig. 7. This is especially true when the structural scale is 3333, and the out-of-plane displacement for $0.58<r/b<1.0$ is about three times larger than the membrane thickness. This implies that the geometrical nonlinearities in the membrane distortion appear in the unwrinkled region, although linearity is assumed for this region in tension field theory. For example, Fig. 12 shows the wave of wrinkles along the circumferential cross section $r/b=0.83$. Clearly, the wave of wrinkles exists in the supposed unwrinkled region assumed in tension field theory. These waves of wrinkles occur in compliance with the wrinkling phenomena in the wrinkled regions, which are related to bifurcation phenomena. Tension field theory cannot predict these wrinkling behavior, because it treats wrinkling phenomena only as problems of in-plane states of stress.

Applicability of Tension Field Theory

Figure 13 shows dependency of a nondimensionalized moment within the wrinkled membrane on the structural scale. The values for this nondimensionalized moment in this figure correspond to those at the final state ($\phi=1.145$ deg, Table 2) shown in Fig. 6. The solution based on tension field theory is also presented in this figure. These numerical results of the geometrically nonlinear analysis are gradually decreased simultaneously with an increase in structural scale. When the structural scale is larger than 1500, they converge to a constant value different from the solution obtained based on tension field theory. This tendency is also shown in comparisons of the angle of wrinkles (Figs. 14 and 15) and of the strain energy ratio within the wrinkled membrane (Fig. 16). The strain energy ratio is obtained by the following equation:

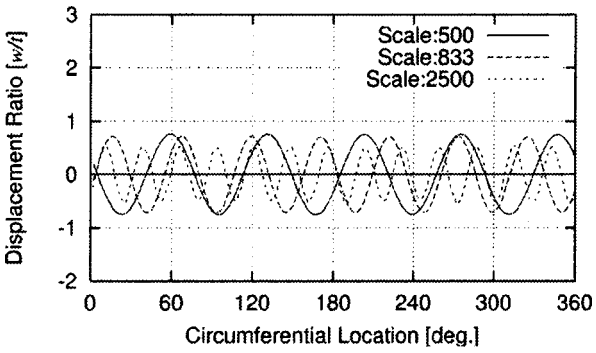


Fig. 12 Wave of wrinkles at the radial location of $r/b=0.83$

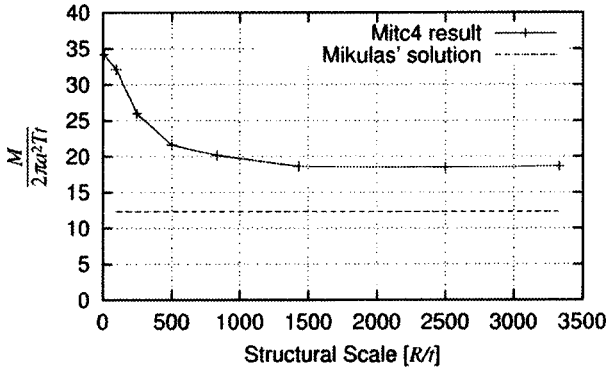


Fig. 13 Comparison of nondimensional moment

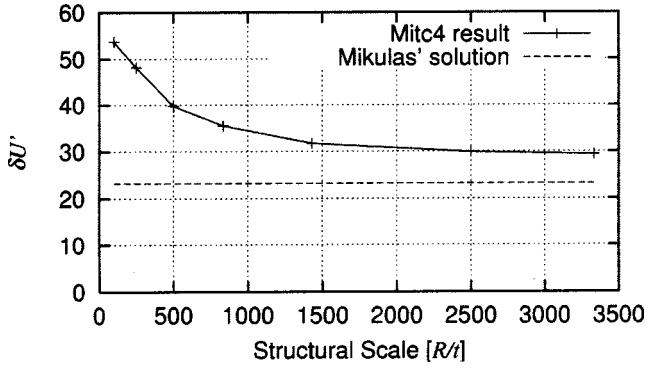


Fig. 16 Comparison of strain energy ratio

$$\delta U' = \frac{\delta U}{U_i} = \frac{U_f - U_i}{U_i} \tag{5}$$

where

U_i: strain energy stored in the stretched membrane before the torque force is applied

U_f: strain energy stored in the wrinkled membrane after the torque force is applied

and $U_{i,f}$ is given by

$$U = \frac{1}{2}\int \frac{1}{E}(\sigma_x^2 + \sigma_y^2 - 2\nu\sigma_x\sigma_y) + \frac{1}{G}(\tau_{xy}^2 + \tau_{xz}^2 + \tau_{yz}^2)t \cdot 2\pi r dr. \tag{6}$$

Here, Eq. (6) omits the notation of *i* and *f*, which indicate initial and final state of the circular membrane. Figure 16 shows that strain energy in the wrinkled membrane gradually increases when the structural scale is smaller than 1500 due to effects of the relative bending stiffness in the membrane. In general, in plate theory membranes are treated as structural materials if their structural scale is larger than 100. However, when we consider the wrinkling phenomena of a membrane using conventional tension field theory, it is better to deal with a membrane as a structural material with structural scale larger than 1000. If the structural scale is smaller than 1000, the effect of the relative bending stiffness on the strain energy increases and cannot be ignored.

Existence of Compressive Stresses Within Wrinkled Region

As can be seen in Figs. 13–16, not all numerical results asymptotically approach to the solution obtained based on tension field theory, and differences between the numerical results of this study on post-buckling phenomena and the solution from tension field theory are clearly shown. These differences remain even if the structural scale becomes large, and imply limitations potentially included in solutions based on tension field theory for actual wrinkling phenomena. Since the wrinkling phenomena are kinds of out-of-plane problems, it should be considered that the geometrically nonlinear analysis simulates the wrinkling phenomena precisely. From the viewpoint of the total strain energy ratio stored in the wrinkled membrane, differences between two solutions are investigated.

Figures 17 and 18 show the strain energy ratio stored only in the wrinkled region ($0.166 < r/b < 0.58$) and the unwrinkled region ($0.58 < r/b < 1.0$) of the membrane, respectively. The value of 0.58 indicates the boundary of the wrinkled region obtained from tension field theory. The summation of results shown in Figs. 17 and 18 gives the result shown in Fig. 16. Therefore, a following equation is given:

$$\delta U' = \delta U'_w + \delta U'_{wn} \tag{7}$$

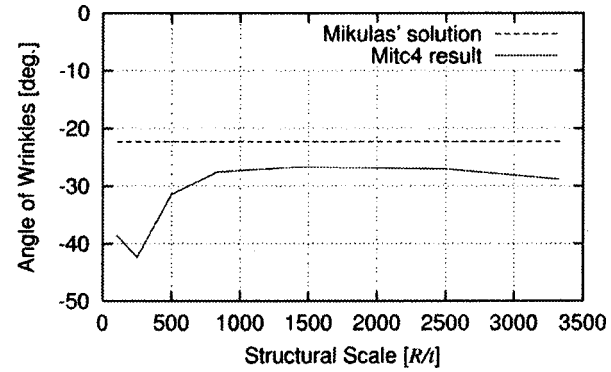


Fig. 14 Angle of wrinkles (radial location: 0.4)

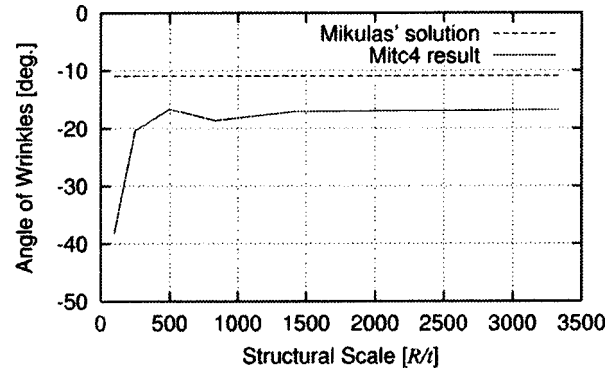


Fig. 15 Angle of wrinkles (radial location: 0.8)

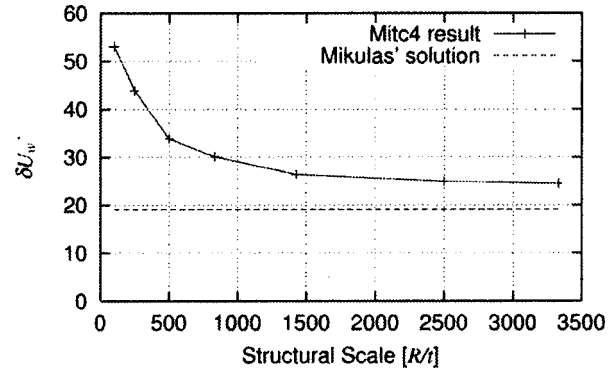


Fig. 17 Strain energy ratio in wrinkled region

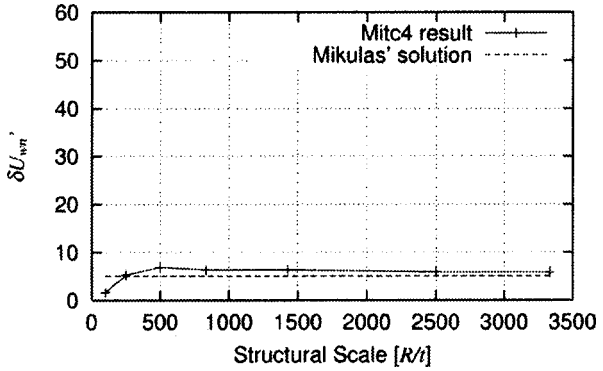


Fig. 18 Strain energy ratio in unwrinkled region

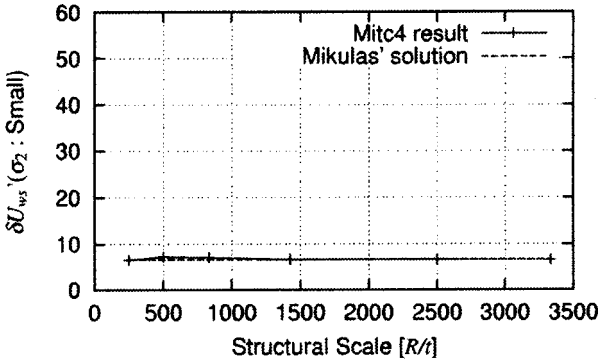


Fig. 20 Strain energy in wrinkled region with small compressive stress

From these figures, it is clear that the difference between the two representative wrinkling theories appears in the wrinkled region, but does not appear in the unwrinkled region when the structural scale becomes large. Therefore these results imply that the difference of these theories shown in Fig. 16 occurs in the wrinkled region ($0.166<r/b<0.58$).

In the post-buckling study, as can be seen from Fig. 10, there are two wrinkled regions in the wrinkled membrane, One with large compressive stress and another with small compressive stress, as stated before. Therefore, the effect of these compressive stresses on the differences of two wrinkling theories is investigated. The strain energy ratio $\delta U_{wl}'$ ($|\sigma_2|$: Large) in Fig. 19 corresponds to that in the part of the wrinkled region whose minor principal stresses correspond to large compressive stresses (wrinkled region with large compressive stresses, $|\sigma_2|>0.5$ MPa). The strain energy ratio $\delta U_{ws}'$ ($|\sigma_2|$: Small) in Fig. 20 corresponds to that in another part of the wrinkled region whose minor principal stresses correspond to small compressive stresses (wrinkled region with small compressive stresses, $|\sigma_2|<0.5$ MPa). Therefore, the summation of the strain energy ratio in Figs. 19 and 20 is equal to $\delta U_w'$ in Fig. 17. Therefore, the following equation is given:

$$\delta U_w' = \delta U_{wl}' + \delta U_{ws}' . \tag{8}$$

From the comparison regarding to the strain energy in the two wrinkled regions (Figs. 19 and 20), it turns out that the differences of these theories appear in the wrinkled region with large compressive stresses ($|\sigma_2|>0.5$ MPa), while in the region with small compressive stress the solutions based on tension field theory are in good agreement with the results of this study on post-buckling phenomena. From these results, we conclude that the difference of these theories is caused by the existence of the large compressive stress in the wrinkled region. And it turns out that the analytical method of tension field theory, in which the material parameter is adjusted in the wrinkled region, simulates the nonlinear characteristics of the wrinkles adequately in the wrinkled regions with small compressive stresses ($|\sigma_2|<0.5$ MPa).

The difference, which appears in the wrinkled region, depends on the tension field theory definition for the membrane as a structural material, which can carry no compressive stress. However, in actual situations, as can be seen in Fig. 10, the membrane must carry some compressive stress, which becomes larger than the minimum buckling stress of the membrane. The threshold value stated above should correspond to the minimum buckling stress. These phenomena occur because the wave of wrinkles is constrained by our given geometrical condition. Figure 21 shows the wave of wrinkles along the line of the minor principal stresses in the circular membrane. Points A, B, C, and D indicate the boundary of the hub, location having maximum out-of-plane displacement, boundary of the wrinkled region and outer circumference of the circular membrane, respectively. If tension field theory can be true, it is considered that inextensional theory is applied along the line of the minor principal stresses in wrinkled regions. In the case of the circular membrane, the compressive strain, which corresponds to the wrinkling strain in tension field theory, becomes large toward the center hub, because the moment along the radial direction is a constant. At that time, out-of-plane displacements increase toward the center hub, and it becomes maximum at Point A, if inextensional theory is applied. However, as can be seen from Fig. 22, which is the same figure shown in Fig. 21, the amplitude of the out-of-plane displacement in the region from A to B gradually decreases toward the hub. This means wave of wrinkles within the regions from A to B has some geometrical constraints by the hub, and is not in agreement with wave of wrinkles obtained from the conditions of inextentional theory.

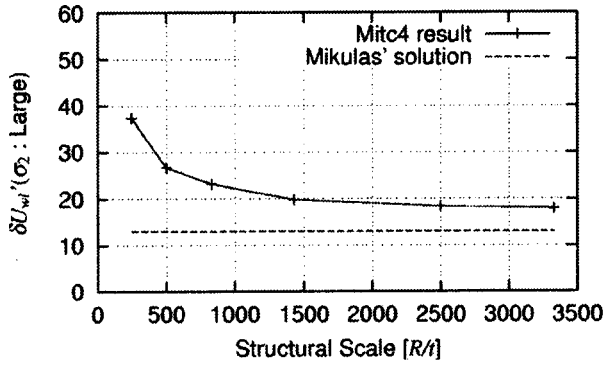


Fig. 19 Strain energy in wrinkled region with large compressive stress

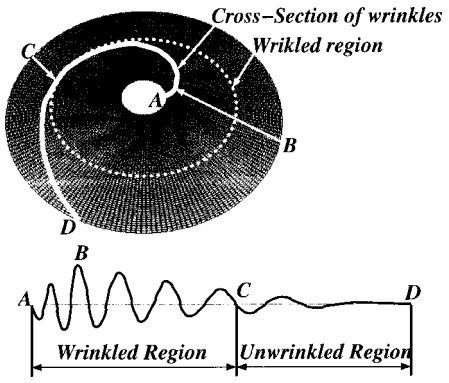


Fig. 21 Wave of wrinkles

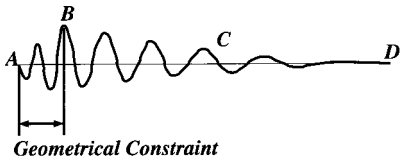


Fig. 22 Geometrical constraint

In general, it is considered that the membrane releases compressive stresses by forming wrinkles. However, in the case of the stretched circular membrane, the wave of wrinkles around the center rotation hub is constrained by the existence of the hub itself. For that reason, the wave of wrinkles, which is initially necessary to release compressive stresses in the wrinkled region, cannot occur, and the membrane must carry some compressive stresses. For the deformation mode subjected to the geometrical constraint, these compressive stresses become larger than the minimum buckling stress of the membrane, which is because the deformation mode with respect to the minimum buckling stress is not equal to the actual one. Considering this, it turns out that the membrane is a structural material that can carry compressive stresses larger than the minimum buckling stress. The region having larger compressive stresses than some threshold value, as explained before, nearly represents the region subjected to the geometrical constraint. In addition, Figs. 7, 10, and 22 show that the region having compressive stress is roughly equal to the region between the boundary of the hub and the radial location having the maximum out-of-plane displacement. These three regions defined above coincide with one another almost perfectly.

Figure 23 shows that the radial location having the maximum out-of-plane displacement changes when the structural scale becomes large. In this figure, the change of the wrinkled region and the boundary of the hub are also presented. Also, this figure shows that the change of the radial location having the maximum out-of-plane displacement roughly indicates the same tendency compared with that of the strain energy ratio shown in Fig. 19. Since the strain energy ratio shown in Fig. 19 is calculated in the region having some compressive stresses, it is also confirmed that the three regions stated before coincide with one another with respect to strain energy.

In general, tension field theory treats the ideal membrane, which has no bending stiffness and can carry no compressive stress. In the situation, it is considered that the amplitude of wrinkles is nearly zero and the number of wrinkles is infinite, [28]. Therefore, for detailed quantitative evaluation of the differences between the two representative theories, the discrepancy in strain energy ratio for the case where the structural scale is infinite must be explained. From this viewpoint, it must be clarified whether the wrinkles exist when the structural scale becomes infinite. From the previous considerations, if the wave of wrinkles does not converge to a flat plane state, some geometrical constraint caused by the hub exists in the wrinkled region, and the membrane absolutely has some compressive stresses. In the following section, using inextensional theory stated above, it is investigated whether the wave of wrinkles can exist or not when the structural scale becomes infinite.

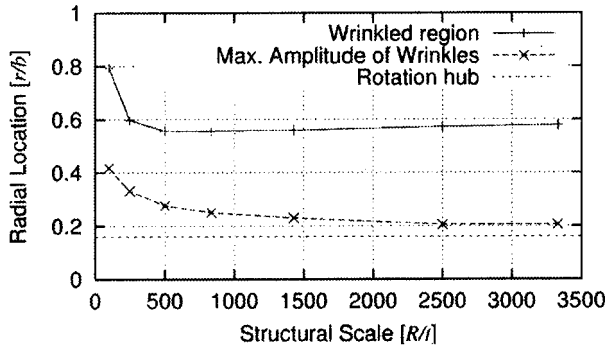


Fig. 23 Location of maximum amplitude of wrinkles and wrinkled region

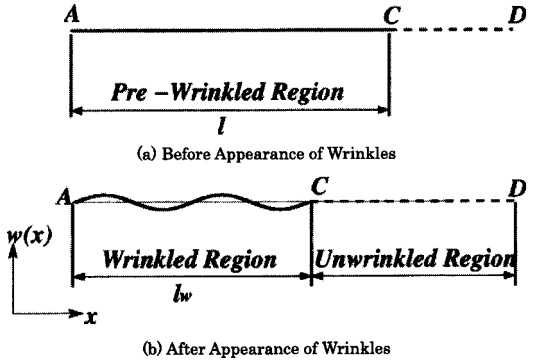


Fig. 24 Mechanism of wrinkling behavior

Note on the Actual Mechanism of Wrinkling Behavior

Figure 24 shows the mechanism of wrinkling behavior based on inextensional theory. This figure corresponds to Fig. 21. Therefore, it is assumed that inextensional theory describes the behavior of wrinkles along the line of minor principal stresses in the wrinkled region. In this subsection, focusing on whether wrinkles exist or not when the structural scale becomes infinite, a simple wave shape, sine wave, is considered for wrinkles. Then, a shape function of the cross section of wrinkles (from A to C in Fig. 24) is expressed by

$$w_{(x)} = \delta_n \sin \frac{n\pi}{l_w} x \tag{9}$$

where, δ_n, n, and l_w indicate the amplitude, number, and length of wrinkles, respectively. From inextensional theory, the following equation is given:

$$l = \int_0^{l_w} \sqrt{1 + \left(\frac{\partial w}{\partial x}\right)^2}\, dx. \tag{10}$$

Substitution of Eq. (9) into Eq. (10) yields

$$l = \int_0^{l_w} \sqrt{1 + \left(\delta_n \frac{n\pi}{l_w} \cos \frac{n\pi}{l_w} x\right)^2}\, dx. \tag{11}$$

Using the following nondimensional parameter

$$X = \frac{x}{l_w}, \quad \delta_n' = \frac{\delta_n}{l_w} \tag{12}$$

and defining α as

$$l_w = \alpha \cdot l, \tag{13}$$

we can get the following equation:

$$l = \alpha l_w \int_0^1 \sqrt{1 + (n\pi\delta_n')^2 \cos^2(n\pi X)}\, dX. \tag{14}$$

Now, we assume the following relations

$$(n\pi\delta_n')^2 \cos^2(n\pi X) \ll 1 \tag{15}$$

Then Eq. (14) becomes

$$l=\alpha l_w\int_0^1\left[1+\frac{1}{2}(n\pi\delta_n')^2\cos^2(n\pi X)\right]dX$$

$$=\alpha l_w\left[1+\frac{1}{4}(n\pi\delta_n')^2\right]. \tag{16}$$

Here, Eq. (15) indicates the following relation:

$$0.8<\alpha<1. \tag{17}$$

From Eq. (16), the nondimensional amplitude of wrinkles is expressed by

$$\delta_n'=\frac{2}{n\pi}\sqrt{\frac{1-\alpha}{\alpha}} \tag{18}$$

and the amplitude of wrinkles is given by

$$\delta_n=\frac{2l}{n\pi}\sqrt{\alpha(1-\alpha)}. \tag{19}$$

From this equation, the shape function of wrinkles is finally obtained by

$$w_{(x)}=\delta_n\sin\frac{n\pi}{l_w}x=\frac{2l}{n\pi}\sqrt{\alpha(1-\alpha)}\sin\frac{n\pi}{l_w}x \tag{20}$$

and this gradient is expressed by

$$\frac{\partial w_{(x)}}{\partial x}=\delta_n\frac{n\pi}{l_w}\cos\frac{n\pi}{l_w}x$$

$$=\frac{2l}{l_w}\sqrt{\alpha(1-\alpha)}\cos\frac{n\pi}{l_w}x$$

$$=2\sqrt{\frac{(1-\alpha)}{\alpha}}\cos\frac{n\pi}{l_w}x$$

$$=\overline{\delta}_n\cos\frac{n\pi}{l_w}x \tag{21}$$

where

$$\overline{\delta}_n=2\sqrt{\frac{(1-\alpha)}{\alpha}}\cdots\text{ gradient factor.} \tag{22}$$

As can be seen from Eqs. (19) and (20), if the structural scale becomes large (the relative bending stiffness is decreased), the amplitude of the wrinkles approaches zero because the number of wrinkles becomes infinite. However, Eq. (22) shows that the gradient factor is constant, even if the structural scale becomes infinite. This implies that the gradient factor of wrinkles does not depend on the number of wrinkles, and remains non-zero except if $\alpha=1$. When $\alpha=1$, no wrinkling behavior will occur. Therefore, the wave of wrinkles also exists absolutely when the structural scale becomes infinite, because the gradient factor of wrinkles does not depend on the number of wrinkles (which is affected by the structural scale). The existence of the wave of wrinkles implies the existence of the region having the geometrical constraint shown in Fig. 22. This also represents the existence of the region having some compressive stress, which is larger than the minimum buckling stress of the membrane. For these results, the wrinkled regions, in which tension field theory is not applied, exist absolutely in the membrane even if the structural scale becomes infinite. This implies that the rigidity of the wrinkled membrane calculated based on tension field theory does not approximate the exact solution as $t\rightarrow 0$.

Furthermore, in the case of the circular membrane, the compressive strain within the wrinkled regions, which corresponds to the wrinkling strain in tension field theory, becomes large when the rotation angle of the hub increases. As stated before, if inextentional theory is applied, out-of plane displacements around the rotation hub become large with increasing compressive strain. However, since the region around the hub has a geometrical constraint, effects of the constraint become large with increasing rotation angle of the hub. This implies that as the wrinkling phenomena progress, the behavior increasingly differs from the solutions of tension field theory. This implication appears in the difference of the values for torsional rigidity after bifurcation, shown in Fig. 6, in which the torsional rigidity based on our post-buckling study is apparently larger than that of tension field theory. This is confirmed by the results shown in Fig. 13, in which the difference of the nondimensional moment at the final states exists when the structural scale becomes large.

Conclusions

A comparative study of two representative theories, bifurcation theory and tension field theory is carried out quantitatively. The effect of the structural scale on the wrinkling behavior of a stretched circular membrane with the center rotation hub is investigated, and the limitation of conventional tension field theory is quantitatively clarified. The detailed conclusions obtained in this study are summarized below.

1. Effect of the structural scale on the wrinkling behavior

The number and amplitude ratio of wrinkles gradually increases, and the radial location having the maximum amplitude of wrinkles approaches the edge of the rotation hub when the structural scale becomes large.

2. Existence of wrinkles in the wrinkled membrane with no compressive stress

It turns out that the wrinkles do exist in the supposed unwrinkled region provided by the tension field theory. These wrinkling phenomena are linked to the wrinkling phenomena in wrinkled regions, which are related to the bifurcation. The tension field theory cannot predict this wrinkling behavior, since it treats wrinkling phenomena only as problems of in-plane states of stress.

3. Applicability of tension field theory

We conclude that it is better to deal with the membrane as a structural material with a structural scale larger than 1000, when tension field theory is applied to wrinkling analysis, although it is considered in plate theory that the membrane is treated as a structural material with a structural scale larger than 100. If the structural scale is smaller than 1000, the effect of relative bending stiffness on the strain energy stored in the wrinkled membrane increases and cannot be ignored.

4. Existence of compressive stresses within the wrinkled region

There are wrinkled regions having larger compressive stress than a given threshold value with respect to strain energy stored in the wrinkled membrane. These regions represent the transition between the boundary of the hub and the radial location having the maximum amplitude of wrinkles. Conventionally, tension field theory is not properly applied to these regions. The wrinkled regions having some compressive stresses appear because the wave of wrinkles is constrained by the geometrical condition.

5. Note on the actual mechanism of wrinkling behavior

It is clarified using inextensional theory that there are the wrinkled regions, in which tension field theory is not properly applied, even as the structural scale becomes infinite. Actual torsional rigidity of a wrinkled membrane does not coincide with the solution of tension field theory, and the difference of the solutions based on two wrinkling theories, bifurcation theory and tension field theory, becomes large when wrinkling progresses.

Acknowledgments

We thank Prof. H. Noguchi for providing support with respect to incorporating the MITC shell element into FEAPpv.

Nomenclature

a = radius of hub
b = radius of circular membrane

E = Young's modulus
G = shear modulus
l = length of cross section of wrinkles
l_w = length of wrinkles
M = moment
n = number of wrinkles
R = radial cross-section length of circular membrane $(b-a)$
r = radial location
T = initial uniform tensile stress
t = membrane thickness
w = out-of-plane displacement
R/t = structural scale
α = factor of wrinkles (Eq. (10))
δ_n = amplitude of wrinkles
$\underline{\delta_n'}$ = nondimensional amplitude of wrinkles (Eq. (9))
δ_n = gradient factor (Eq. (19))
φ = rotation angle of hub
θ = circumferential location
σ_x = stress in X component
σ_y = stress in Y component
σ_1 = major principal stress
σ_2 = minor principal stress
τ_{xy} = shear stress
ν = Poisson's ratio

References

[1] Johnston, J., and Lienard, S., 2001, "Modeling and Analysis of Structural Dynamics for a One-Tenth Scale Model NSGT Sunshield," Paper No. AIAA-2001-1407.
[2] Murphey, T. W., Murphey, D. W., Mikulas, M. M., and Adler, A. L., 2002, "A Method to Quantify the Thrust Degradation Effects of Structural Wrinkles in Solar Sails," Paper No. AIAA-2002-1560.
[3] Johnston, J., 2002, "Finite Element Analysis of Wrinkled Membrane Structures for Sunshield Applications," Paper No. AIAA-2002-1456.
[4] Blandino, J. R., Johnston, J. D., Miles, J. J., and Soplop, J. S., 2001, "Thin Film Membrane Wrinkling due to Mechanical and Thermal Loads," Paper No. AIAA-2001-1345.
[5] Roddeman, D. G., Drukker, J., Oomens, C. W. J., and Janssen, J. D., 1987, "Wrinkling of Thin Membranes: Part I—Theory; Part II—Numerical Analysis," ASME J. Appl. Mech., **54**, pp. 884–892.
[6] Adler, A. L., Mikulas, M. M., and Hedgepeth, J. M., 2000, "Static and Dynamic Analysis of Partially Wrinkled Membrane Structures," Paper No. AIAA-2000-1810.
[7] Kang, S., and Im, S., 1997, "Finite Element Analysis of Wrinkling Membrane," ASME J. Appl. Mech., **64**, pp. 263–269.
[8] Kang, S., and Im, S., 1999, "Finite Element Analysis of Dynamic Response of Wrinkling Membranes," ASME J. Appl. Mech., **173**, pp. 227–240.
[9] Yang, B., and Ding, H., 2002, "A Two-Viable Parameter Membrane Model for Wrinkling Analysis of Membrane Structures," Paper No. AIAA-2002-1460.
[10] Wu, C. H., 1974, "The Wrinkled Axisymmetric Air Bags Made of Inextensible Membrane," ASME J. Appl. Mech., **41**, pp. 963–968.
[11] Stein, M. S., and Hedgepeth, J. M., 1961, "Analysis of Partly Wrinkled Membrane," NASA TN D-813.
[12] Mikulas, M. M., 1964, "Behavior of a Flat Stretched Membrane Wrinkled by the Rotation of an Attached Hub," NASA TN D-2456.
[13] Li, X. Li, and Steigmann, D. J., 1993, "Finite Plane Twist of an Annular Membrane," Q. J. Mech. Appl. Math., **46**, pp. 601–625.
[14] Miyamura, T., 2000, "Wrinkling of Stretched Circular Membrane Under In-Plane Torsion: Bifurcation Analysis and Experiments," Eng. Struct., **20**, pp. 1407–1425.
[15] Miyamura, T., 1995, "Bifurcation Analysis and Experiment for Wrinkling on Stretched Membrane," Doctoral dissertation, Univ. of Tokyo (in Japanese).
[16] Iwasa, T., Natori, M. C., Noguchi, H., and Higuchi, K., 2003, "Geometrically Nonlinear Analysis on Wrinkling Phenomena of a Circular Membrane," The Membrane Structures Association of Japan, Research Report on Membrane Structures, (16), pp. 7–14 (in Japanese); also International Conference on Computational Experimental Engineering Sciences (ICCES'03), Corfu, Greece.
[17] Iwasa, T., Natori, M. C., and Higuchi, K., 2003, "Numerical Study on Wrinkling Properties of a Circular Membrane," Journal of JSASS, **51**(591) (in Japanese).
[18] Iwasa, T., Natori, M. C., and Higuchi, K., 2003, "Comparative Study on Bifurcation Theory and Tension Field Theory for Wrinkling Analysis," J Struct. Eng., **49B**, pp. 319–326 (in Japanese).
[19] Dean, W. R., 1924, "The Elastic Stability of an Annular Plate," Proc. R. Soc. London, Ser. A, **106**, pp. 268–284.
[20] Bucciarelli, L. L., 1969, "Torsional Stability of Shallow Shells of Revolution," AIAA J., **7**, pp. 648–653.
[21] Bathe, K.-J., and Dvorkin, E. N., 1986, "A Formation of General Shell Elements–The Use of Mixed Interpolation of Tensorial Components," Int. J. Numer. Methods Eng., **22**, pp. 697–722.
[22] Bathe, K.-J., and Dvorkin, E. N., 1985, "Short Communication a Four-Node Plate Bending Element Based on Mindlin/Reissner Plate Theory and a Mixed Interpolation," Int. J. Numer. Methods Eng., **21**, pp. 367–383.
[23] Bucalem, M. L., and Bathe, K.-J., 1993, "Higher-Order MITC General Shell Elements," Int. J. Numer. Methods Eng., **36**, pp. 3729–3754.
[24] Chapell, D., and Bathe, K.-J., 1998, "Fundamental Considerations for the Finite Element Analysis of Shell Structures," Comput. Struct., **66**(1), pp. 19–36.
[25] Bathe, K.-J., Iosilevich, A., and Chapelle, D., 2000, "An Evaluation of the MITC Shell Elements," Comput. Struct., **75**.
[26] Taylor, R. L., "FEAP-A Finite Element Analysis Program Version 7.3 User Manual," http://www.ce.berkeley.edu/~rlt/.
[27] Noguchi, H., and Hisada, T., 1992, "An Efficient Formulation for a Shell Element Considering Finite Rotation Increments and Its Assessment," Trans. Jpn. Soc. Mech. Eng., **58**(550) (in Japanese).
[28] Libai, A., and Simmonds, J. G., 1998, *The Nonlinear Theory of Elastic Shells*, 2nd Ed., Cambridge University Press, New York, pp. 276–280, 431–441.

Oliver M. O'Reilly[1]
Professor
Department of Mechanical Engineering,
University of California,
Berkeley, CA 94720-1740
e-mail: o'reilly@me.Berkeley.edu

Peter C. Varadi
Engineering Consultant
1995 University Avenue, Suite 375,
Berkeley, CA 94704
Mem. ASME

On Some Peculiar Aspects of Axial Motions of Closed Loops of String in the Presence of a Singular Supply of Momentum

We consider the dynamics of a closed loop of inextensible string which is undergoing an axial motion. At each instant, one material point of the string is in contact with a singular supply of linear momentum (also known as an external constraint). Several peculiar features of this problem which have not been previously discussed are presented. These include the possible presence of an arbitrary number of kinks, the vanishing nature of the singular supply of momentum, and the critical nature of the tension in the string. When the linear momentum is supplied by a mass-spring-dashpot system, we are also able to establish an exact expression for the frequency of the resulting vibrations, prove that dissipation cannot be present, show that these vibrations only occur for discrete speeds of axial motion, and establish that Coulomb friction is absent. [DOI: 10.1115/1.1756139]

1 Introduction

The subject of axially moving media has attracted considerable attention. Earlier works on this subject considered the linear vibration about a steady axial motion of a segment of string where no body forces or surface tractions were present. A recent trend in this research area has been to examine the influence of discrete mechanical systems which contact the moving string at a single point. The present paper contributes to this topic by examining the effects of a point source (or singular supply) of linear momentum, denoted by $\mathbf{F}$, on the linear vibrations of a closed loop of axially moving string. We point out several interesting degeneracies in this problem and the consequences that they have on the dynamics of the string. Among our observations are the possible presence of an arbitrary number of kinks, the vanishing nature of $\mathbf{F}$, and the critical nature of the tension n_0 in the string. When the linear momentum is supplied by a mass-spring-dashpot system, we establish, among other results, an exact expression for the discrete frequency of the resulting vibrations.

Our work is a contribution to understanding the dynamics of a system described in a 1984 paper by Schajer [1]. He considered the effects of a contact (or external constraint) at a single point of a closed circular loop of inextensible string (see Fig. 1). This system, albeit idealized, was regarded as a prototype system to improve understanding of the more difficult problems associated with vibrations induced by contact in spinning flexible disks. Subsequent papers by Perkins and Mote [2] and Hutton and his coworkers, [3–6] generalized Schajer's work. Other authors, such as Chen [7] and Cheng and Perkins [8], considered the related problem where the steady motion of the string described a straight line (the so-called moving threadline problem) and one point of the string was in contact with a mass-spring-dashpot system.

In Refs. [1–6], the equations of motion of the string are linearized about a steady axial motion. During this motion, the string forms a closed circular loop of radius R. The tension in the string during the steady motion is assumed to be a constant n_0, as is the axial speed $V=R\Omega$. One of the assumptions employed in the vibration analyses in [1–6] is that n_0 is independent of V (or Ω). In this paper, we use a classical result (see, e.g., Lamb [9]) to show that if there are no surface tractions or body forces on the string, then this is not the case. Indeed, we show that n_0 is always equal to its critical value $\rho_0 V^2$. This result has several interesting consequences when the string is in contact with a mass-spring-dashpot system. However, only in Yang and Hutton [4] has this critical case been analyzed. In this paper, we supplement their results by showing that vibrations of the string which were dismissed in [4] are indeed possible. In fact, the frequency of these vibrations is the natural frequency $\sqrt{k/m}$ of the mass-spring system. In the closing section of the paper we show how an external force can be introduced to achieve the steady motions discussed in [1–6].

We note that for the moving threadline problem, discussed in, e.g., [7,10,11], n_0 can be varied independently of $\rho_0 V^2$ and the degeneracies we report here don't apply for this problem. Similarly, for the axially moving string in a gravitational field discussed by O'Reilly [12], $n_0 \neq \rho_0 V^2$. Finally, $n_0=\rho_0 V^2$ for a steady motion of a circular loop of string in the absence of body forces and singular supplies of momentum, and Healey [13] has proven that this motion is nonlinearly stable.

An outline of this paper is as follows: We first formulate the problem of a closed loop of inextensible string that is in motion with one point in contact with a supply of linear momentum. This supply exerts a force $\mathbf{F}$ on the string at the material point $\xi=\gamma$. The presence of the supply serves to model the effect of a mass-spring-dashpot in contact with the string. It also introduces several algebraic jump conditions into the equations of motion. We then turn to solving for the steady motion of the string, and first show that $n_0=\rho_0 V^2$ even in the presence of the source of linear momentum. This result implies that during a steady motion $\mathbf{F}=\mathbf{0}$ and moreover that the string can have an arbitrary finite number of points where the tangent vector $\mathbf{e}_t$ is not continuous (i.e., the closed loop of steadily moving string can have an arbitrary finite number of kinks). We then discuss the equations governing the small amplitude linear vibrations of the string about the steady motion. Because $n_0=\rho_0 V^2$, these equations also exhibit several

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, May 12, 2003; final revision, October 23, 2003. Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

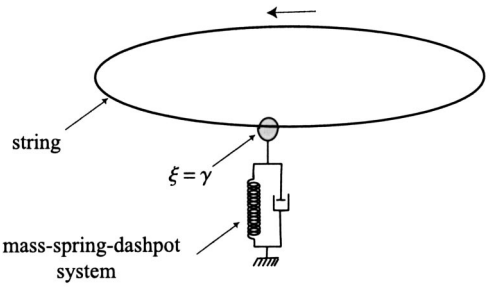


Fig. 1 Schematic of a closed loop of inextensible string which is in contact with a mass-spring-dashpot system. The contact of the system and the string occurs at the material point $\xi=\gamma(t)$ of the string. The system shown in this figure is similar to those discussed in [1–6].

degeneracies. Among the degeneracies we find are the existence of vibrations at a fixed frequency which only occur for certain discrete speeds V. The closing section of the paper gives a physical interpretation of the necessary conditions we found for these vibrations. There, we also explicitly calculate the external force field needed to achieve the steady motions discussed in [1–6].

2 The Governing Equations

For an inextensible string, we denote the position vector of each material point by the vector $\mathbf{r}$. This vector can be described as a function of an arc length parameter ξ in an undeformed configuration and the time t. The arc length parameter of the string in its present configuration is denoted by s, and the (uniform) mass density of the string is denoted by ρ_0.

The contact force $\mathbf{n}$ in the string has the representation $\mathbf{n}=n\mathbf{e}_t$ where $\mathbf{e}_t=\partial\mathbf{r}/\partial s$ is the unit tangent vector to the string, and n is the tension in the string. We assume that at the point $s=s_\gamma$, which corresponds to the material point $\xi=\gamma(t)$, the string experiences a singular supply of linear momentum $\mathbf{F}$ (see Fig. 2). This force represents the force exerted on the string by the mass-spring-dashpot systems discussed in [1–6]. It is assumed that no other body forces or surface tractions act on the string.

The nontrivial equations governing the motion of the string are the inextensibility constraint and compatibility conditions,

$$\frac{\partial\mathbf{r}}{\partial\xi}\cdot\frac{\partial\mathbf{r}}{\partial\xi}=1,\quad [\![\mathbf{r}]\!]=\mathbf{0},\quad \left[\!\left[\dot{\mathbf{r}}+\dot{\gamma}\frac{\partial\mathbf{r}}{\partial\xi}\right]\!\right]=\mathbf{0}, \tag{1}$$

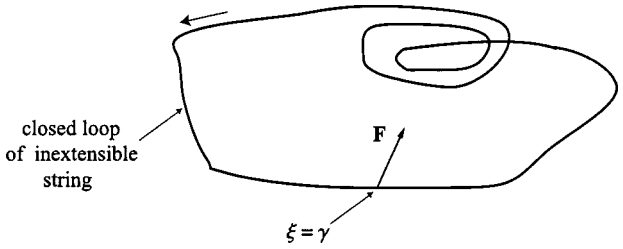


Fig. 2 Schematic of a closed loop of inextensible string which experiences a singular supply of momentum F at the point $\xi=\gamma$. The system shown in this figure can be considered a generalization of that shown in Fig. 1.

and the local balance law and jump condition for linear momentum,[2]

$$\frac{\partial\mathbf{n}}{\partial s}=\rho_0\ddot{\mathbf{r}},\quad [\![\mathbf{n}]\!]+[\![\rho_0\dot{\mathbf{r}}]\!]\dot{\gamma}+\mathbf{F}=\mathbf{0}. \tag{2}$$

In these equations, the superposed dot denotes the material time derivative and $[\![\mathbf{f}]\!]=\mathbf{f}(\gamma^+,t)-\mathbf{f}(\gamma^-,t)$ denotes the jump in $\mathbf{f}(\xi,t)$ across the material point $\xi=\gamma(t)$.

3 The Steady Motion

The steady motion of the string of interest is an axial motion where the string moves at a constant speed V along its length and describes a fixed curve $\mathbf{r}=\mathbf{r}_0$ in space:

$$\dot{\gamma}=\dot{\gamma}_0,\quad \dot{\mathbf{r}}=V\frac{\partial\mathbf{r}_0}{\partial s}=V\mathbf{e}_{t0},\quad \ddot{\mathbf{r}}=V^2\frac{\partial\mathbf{e}_{t0}}{\partial s},$$
$$\mathbf{n}=n_0\mathbf{e}_{t0},\quad \mathbf{F}=\mathbf{F}_0, \tag{3}$$

and $\mathbf{e}_{t0}=\partial\mathbf{r}_0/\partial s$ is the unit tangent vector to the string. Here, and henceforth, the subscript 0 denotes quantities associated with the steady motion.

From the compatibility condition Eq. (1),[3] we find that

$$\dot{\gamma}_0=-V. \tag{4}$$

Substituting Eq. (3) and $\dot{\gamma}=-V$ into Eq. (2), we find that

$$\frac{\partial}{\partial s}((n_0-\rho_0V^2)\mathbf{e}_{t0})=\mathbf{0},$$
$$[[(n_0-\rho_0V^2)\mathbf{e}_{t0}]]+\mathbf{F}_0=\mathbf{0}. \tag{5}$$

The other two equations, Eqs. (1),[1,2] are trivially satisfied.

Equation (5) provide six equations to solve for the six unknowns n_0, $\mathbf{e}_{t0}$ and $\mathbf{F}=\mathbf{F}_0$. First, because $\mathbf{e}_{t0}$ changes along the length of the closed loop of string, Eq. (5) implies that

$$n_0=\rho_0V^2, \tag{6}$$

throughout the entire length of string. With this solution for n_0, it is easy to see that any $\mathbf{e}_{t0}$ satisfies Eq. $(5)_1$. As a result, the closed loop of string can take any spatial form provided $n_0=\rho_0V^2$. This result is classical (see, e.g., Section 50 of Lamb [9]).[3] In addition, because $n_0=\rho_0V^2$, we find from Eq. $(5)^2$ that the singular supply of linear momentum is

$$\mathbf{F}_0=\mathbf{0}. \tag{7}$$

In other words, in order for the string to execute a steady axial motion, the singular supply of momentum must vanish. We also observe that $\mathbf{e}_{t0}$ is not necessarily continuous at $\xi=\gamma$. Indeed, because $n_0=\rho_0V^2$ we have the unusual feature that the steady motion of the string can have an arbitrary finite number of points where $\mathbf{e}_{t0}$ is not continuous. This situation is analogous to the possibility of an arbitrary number of kinks in a slack string discussed by Reeken [17]. It is also similar to the discontinuities with a discontinuous tangent vector discussed by Purohit and Bhattacharya [18].[4]

For the circular loop of string of radius R considered in [1–6] it is convenient to define the rotational speed $\Omega=V/R$. Our analysis implies that an assumption used in these papers, which dates to Schajer's paper [1], namely that the tension n_0 is independent of Ω, is, as noted in [4], artificial.[5]

[2]These equations can be obtained from the thermomechanical theory for a string discussed in O'Reilly and Varadi [14] (see also [15]).

[3]In 1990, it was extended to closed loops of nonlinearly elastic strings by Healey and Papadopoulos [16].

[4]We are grateful to an anonymous reviewer for bringing this paper to our attention.

[5]Schajer's paper has an error in the jump condition at the contact point (see Eq. (7) of [1]). However, this error was corrected in the papers by Hutton et al. [3–6].

4 Linearized Equations of Motion

It is interesting to establish the equations governing the small amplitude vibrations of the string about the steady motion. To assist with these equations, we follow standard practice and define the coordinate θ:

$$\theta = \xi + Vt \bmod(L), \tag{8}$$

where L is the length of the closed loop of string. Next, we consider asymptotic expansions of the displacement and force variables:

$$\begin{aligned} \mathbf{r} &= \mathbf{r}_0 + \epsilon \mathbf{u} + O(\epsilon^2), \\ \gamma &= \gamma_0 + \epsilon \gamma_1 + O(\epsilon^2), \\ n &= n_0 + \epsilon n_1 + O(\epsilon^2), \\ \mathbf{F} &= \mathbf{F}_0 + \epsilon \mathbf{F}_1 + O(\epsilon^2). \end{aligned} \tag{9}$$

Substituting Eq. (9) into Eq. (1) and Eq. (2), using Eq. (5) and ignoring terms of $O(\epsilon^2)$, we find that

$$\begin{gathered} \frac{\partial \mathbf{r}_0}{\partial \theta} \cdot \frac{\partial \mathbf{u}}{\partial \theta} = 0, \\ \frac{\partial}{\partial \theta}\left(n_1 \frac{\partial \mathbf{r}_0}{\partial \theta} + n_0 \frac{\partial \mathbf{u}}{\partial \theta} \right) = \rho_0 \left(V^2 \frac{\partial^2 \mathbf{u}}{\partial \theta^2} + 2V \frac{\partial^2 \mathbf{u}}{\partial \theta \partial t} + \frac{\partial^2 \mathbf{u}}{\partial t^2} \right), \\ [\![\mathbf{u}]\!] = \mathbf{0}, \\ \left[\!\left[\frac{\partial \mathbf{u}}{\partial t} + \dot{\gamma}_1 \frac{\partial \mathbf{r}_0}{\partial \theta} \right]\!\right] = \mathbf{0}, \\ \left[\!\left[(n_1 + \rho_0 V \dot{\gamma}_1) \frac{\partial \mathbf{r}_0}{\partial \theta} + (n_0 - \rho_0 V^2) \frac{\partial \mathbf{u}}{\partial \theta} - \rho_0 V \frac{\partial \mathbf{u}}{\partial t} \right]\!\right] + \mathbf{F}_1 = \mathbf{0}. \end{gathered} \tag{10}$$

Notice that we have not yet imposed the condition $n_0 = \rho_0 V^2$ in Eq. (10).

Equation (10) constitutes a system of equations for the unknowns $\mathbf{u}$, n_1, and γ_1. For closure in the present circumstances, we require that $\mathbf{F}_1$ is prescribed as a function of the kinematics of the string. It is also prudent to note the well known result that the first of Eq. (10) implies that the displacement $\mathbf{u}$ is perpendicular to the steady motion $\mathbf{r}_0$.

5 Solutions of the Linearized Equations

We henceforth restrict our attention to the case where $\mathbf{r}_0$ lies entirely on a plane and $\mathbf{u}$ is perpendicular to this plane: $\mathbf{u} = u\mathbf{E}$ where $\mathbf{E}$ is a constant unit vector.[6] After first presenting some general results for a supply of momentum, we shall turn to the specific case of a mass-spring-dashpot system.

To start, we note that a consequence of our planarity assumption, the linearized constraint Eq. $(10)^1$ is identically satisfied. Using Eq. (6), Eq. $(10)^2$ can be divided into two equations:

$$\begin{aligned} \frac{\partial}{\partial \theta}(n_1 \mathbf{e}_{t0}) &= \mathbf{0}, \\ \rho_0 \left(2V \frac{\partial^2 \mathbf{u}}{\partial \theta \partial t} + \frac{\partial^2 \mathbf{u}}{\partial t^2} \right) &= \mathbf{0}. \end{aligned} \tag{11}$$

It is interesting to note that Eq. $(11)^2$ implies that a (continuous) function of θ can be added to $\mathbf{u}$ and Eq. $(11)^2$ will still be satisfied.[7] We can split each of the jump conditions Eqs. $(10)^{4,5}$ into two equations in a similar manner with the goal of solving Eq. (10). It is to this task that we now turn.

First, we notice that Eq. $(11)_1$ is easily solved to determine the tension n_1:

$$n_1 = 0. \tag{12}$$

The method of solution parallels the argument we used earlier to determine n_0. This result implies that $\mathbf{n} = n_0 \mathbf{e}_{t0} + O(\epsilon^2)$. In words, $\mathbf{n}$ is equal to its critical value to second order.

Before solving the partial differential equation Eq. (11),[2] we first turn to the jump conditions. There, we find from Eqs. $(10)^{3,4}$ that $\mathbf{u}$ and $\partial \mathbf{u}/\partial t$ are continuous at $\xi = \gamma$. As $n_0 = \rho_0 V^2$, the remaining jump conditions simplify to

$$\begin{aligned} [\![\dot{\gamma}_1 \mathbf{e}_{t0}]\!] &= \mathbf{0}, \\ [\![n_1 \mathbf{e}_{t0}]\!] + \mathbf{F}_1 &= \mathbf{0}. \end{aligned} \tag{13}$$

However, we already showed that $n_1 = 0$, so we can conclude from Eq. $(13)^2$ that

$$\mathbf{F}_1 = \mathbf{0}. \tag{14}$$

Turning to Eq. (13),[1] we find that if the tangent vector $\mathbf{e}_{t0}$ to the steady motion of the string is continuous at $\xi = \gamma$, then this jump condition is trivially satisfied and the speed $\dot{\gamma}_1$ is indeterminate. On the other hand, if $\mathbf{e}_{t0}$ is not continuous at $\xi = \gamma$, then $\dot{\gamma}_1 = 0$.

The solution of Eq. $(11)^2$ which satisfies the continuity condition $[\![\mathbf{u}]\!] = \mathbf{0}$ is easily found:

$$\mathbf{u}(\theta, t) = (g_1(\theta - 2Vt) + g_2(\theta))\mathbf{E}. \tag{15}$$

Here, g_1 and g_2 are functions, determined from the initial conditions on the string's motion, which satisfy the periodicity condition

$$\mathbf{u}(\theta, t) = \mathbf{u}(\theta + L, t). \tag{16}$$

The displacement at $\xi = \gamma$ can be found from Eq. (15) after noting that this point corresponds to $\theta = \gamma + Vt$:

$$\mathbf{u}_\gamma = \mathbf{u}(\gamma + Vt, t) = (g_1(\gamma - Vt) + g_2(\gamma + Vt))\mathbf{E}. \tag{17}$$

For completeness, we also note that

$$\gamma(t) = \gamma(t_0) - V(t - t_0) + \epsilon \int_{t_0}^{t} \dot{\gamma}_1(z)\, dz + O(\epsilon^2), \tag{18}$$

where t_0 and $\gamma(t_0)$ are constants. It is not possible to conclude from Eq. (17) that $\mathbf{u}_\gamma$ is zero.

Thus far we have shown that $\mathbf{F} = \mathbf{0} + O(\epsilon^2)$ and $\mathbf{n} = \rho_0 V^2 \mathbf{e}_{t0} + O(\epsilon^2)$. In addition, $\dot{\gamma}_1$ may be indeterminate.[8] To obtain concrete results on $\mathbf{u}$ we need to consider a specific supply of linear momentum. We choose the prototypical situation where $\mathbf{F}$ is supplied by a mass-spring-dashpot system. Our results for this system have an overlap with the critical case discussed in Section 3.2 of Yang and Hutton [4]. However, these authors did not consider the possibility for the vibrations that we find below. Further, our results apply when the steady motion of the string lies on a plane—this motion does not have to be circular. We also note that because $\mathbf{F} = \mathbf{0}$ to second order in ϵ, the mass-spring-dashpot system does not exert a normal force on the string. Consequently, when assuming a friction force which is proportional to a normal load, there is no friction force acting on the string.

5.1 A Mass-Spring-Dashpot System. We now restrict our attention to the case where $\mathbf{F}$ is supplied by a mass-spring-dashpot system aligned in the $\mathbf{E}$ direction. For the mass-spring-dashpot system, the mass is denoted by m, the linear spring has a stiffness k and the dashpot has a viscous coefficient d. We denote by $\epsilon x \mathbf{E}$ the displacement of the mass particle and assume that the system only moves in the $\mathbf{E}$ direction. It is not too difficult to see that

$$x\mathbf{E} = \mathbf{u}_\gamma, \quad \mathbf{F} = \mathbf{0} - \epsilon(m\ddot{x} + d\dot{x} + kx)\mathbf{E}. \tag{19}$$

[6] This encompasses the situations discussed in [1–6].

[7] This function is the function g_2 discussed below.

[8] A similar indeterminacy arises in studies on phase transformations (see, e.g., [19,20]) and folded strings (see, e.g., [14,17,18]). There, it is removed by prescribing a singular supply of energy and then using the jump condition from the energy equation (see, for instance, Eq. (5.5) of [14]). However, for the problem at hand, such an approach does not resolve the indeterminacy.

However, our previous results showed that $\mathbf{F}_1=\mathbf{0}$. Consequently,

$$m\ddot{x}+d\dot{x}+kx=0. \tag{20}$$

Substituting Eq. (17) and Eq. (18) into Eq. (20) and dropping terms of $O(\epsilon)$, we find that

$$4V^2mg_1''(\tau)-2Vdg_1'(\tau)+kg_1(\tau)+kg_2(\tau+2Vt)=0, \tag{21}$$

where

$$\tau=\gamma_0(t)-Vt=\gamma(0)-2Vt+O(\epsilon), \tag{22}$$

and the prime denotes the derivative of a function $g=g(\tau)$ with respect to τ.

After noting that $\tau+2Vt=\gamma(0)+O(\epsilon)$, it is easy to solve Eq. (21):

$$g_1=\hat{g}_1(\tau)=e^{\delta\tau}(a_1\sin(\chi\tau)+a_2\cos(\chi\tau))-g_2(\gamma(0)). \tag{23}$$

In this equation, $a_{1,2}$ are real-valued constants which are determined by the initial conditions, and

$$\lambda=\delta+i\chi=\frac{d}{4mV}\left(-1+\sqrt{1-\frac{4km}{d^2}}\right). \tag{24}$$

In Eq. (24), χ and δ are the imaginary and real parts, respectively, of λ.

Returning to the vibration of the string, we now find from Eq. (15) that

$$\begin{aligned}\mathbf{u}(\theta,t)&=(\hat{g}_1(\theta-2Vt)+g_2(\theta))\mathbf{E}\\&=a_1e^{\delta(\gamma(0)-2Vt)}\sin(\chi(\gamma(0)-2Vt))\mathbf{E}+a_2e^{\delta(\gamma(0)-2Vt)}\\&\quad\times\cos(\chi(\gamma(0)-2Vt))\mathbf{E}+(g_2(\theta)-g_2(\gamma(0)))\mathbf{E}.\end{aligned} \tag{25}$$

As mentioned earlier, this solution must satisfy the periodicity conditions Eq. (16), and a_1, a_2, and g_2 are determined from the initial conditions. Choosing a_1 and a_2 to be zero, we find that it is necessary that $g_2(\theta)=g_2(\theta+L)$. Setting $g_2=0$, it is not too difficult to see from Eq. (23) and Eq. (25) that the following restrictions are necessary:

$$\sin(\chi L)=0,\quad e^{\delta L}\cos(\chi L)=1. \tag{26}$$

The parameters χ and δ in this equation are specified by Eq. (24). We noted earlier that the solution of $\mathbf{u}$ was determined from Eq. (11)[2] to an arbitrary added function of θ. Consequently, in Eq. (25) we can, without loss in generality, set $g_2=0$. As a result, the restrictions Eq. (26) become both necessary and sufficient.

Clearly Eq. (26) can only be satisfied if $\delta=\mathrm{Re}(\lambda)=0$. In other words, if the dashpot is absent and $\chi=1/2V\sqrt{k/m}$. Even in this case, Eq. (26) requires the parameters of the mass-spring system to be related to the length of the string in order for a time varying vibration to be present:

$$\frac{2N\pi}{L}=\frac{1}{2V}\sqrt{\frac{k}{m}}. \tag{27}$$

We can interpret this relation to be an equation for the speeds V such that for a given k, m, and L, a time-dependent vibration $\mathbf{u}$ exists in the string. The resulting vibration, which can also be interpreted as a forward travelling wave[9] superposed on a standing wave $(g_2(\theta)-g_2(\gamma(0)))\mathbf{E}$. The explicit form of $\mathbf{u}$ is found from Eq. (23), Eq. (25), and Eq. (27):

$$\begin{aligned}\mathbf{u}(\theta,t)=a_1\sin\left(\frac{1}{2V}\sqrt{\frac{k}{m}}(\theta-2Vt)\right)\mathbf{E}+a_2\cos\left(\frac{1}{2V}\sqrt{\frac{k}{m}}(\theta\right.\\\left.-2Vt)\right)\mathbf{E}+(g_2(\theta)-g_2(\gamma(0)))\mathbf{E},\end{aligned} \tag{28}$$

[9]Because $\theta=\xi+Vt$, this is a forward travelling wave both for an observer who is moving with the steady axial motion of the string and a fixed observer.

where V is a solution of Eq. (27) for some integer n. It is also interesting to note that the frequency of these vibrations is the natural frequency of the mass-spring system.

When the steady motion is circular of radius R with an axial speed $V=\Omega R$, then Eq. (27) simplifies to

$$\Omega=\frac{1}{2N}\sqrt{\frac{k}{m}}. \tag{29}$$

From this relation we conclude that, given the appropriate initial conditions for $\mathbf{u}$, a nontrivial vibration of the string is possible provided the natural frequency of the mass-spring system is an even integer multiple of the speed Ω.

6 Closing Comments and Interpretations

For the vibrations of the string in contact with the mass-spring system, it is easy to give interpretations of the conditions Eq. (27) and Eq. (29). If we imagine the mass-spring system vibrating with a frequency $\sqrt{k/m}$, this will induce a vibration in the string which is carried by the axial motion. For the vibration to continue to exist, it must, after traversing the length of the string with a speed V, have the same magnitude as the vibration of the mass spring at the instant it returns to the contact point $\xi=\gamma$. The conditions Eq. (27) and Eq. (29) ensure that this is the case.

Ultimately, because $\mathbf{F}=\mathbf{0}$, the mass-spring system oscillates at its resonant frequency and the axial speed V must be synchronized with this frequency. By extension, and by virtue of linear superposition, if the spring is replaced with any elastic structure with a discrete number of natural frequencies ω_k then a relation similar to Eq. (27) would hold with ω_k replacing $\sqrt{k/m}$.

Throughout this paper we have argued that, and explored the consequences of, $n_0=\rho_0V^2$. We also commented that [1–6] assume that for closed loops of string, $n_0\neq\rho_0V^2$ in general. To achieve a tension n_0 in the string which satisfies this assumption, a distributed force $\rho_0\mathbf{f}$ needs to be applied to the lateral surface of the string. This force is determined by the balance of linear momentum for the string (cf. Eq. (5)):

$$\rho_0\mathbf{f}=-\frac{\partial}{\partial s}((n_0-\rho_0V^2)\mathbf{e}_{t0}). \tag{30}$$

We interpret this equation as one where prescribing the steady state motion's parameters $n_0-\rho V^2$ and $\mathbf{e}_{t0}$, we can find the force $\rho_0\mathbf{f}$ needed to achieve this motion. For the circular loops of string considered in [1–6], $n_0-\rho_0V^2$ is assumed to be a (nonzero) constant and $\mathbf{e}_{t0}$ is the unit tangent vector to the circular motion. Consequently it is easy to calculate the force $\rho_0\mathbf{f}$ needed to sustain the steady motions discussed in [1–6]. Indeed, a mechanism to achieve this force is discussed in [3].

Acknowledgments

The work of Oliver M. O'Reilly on this paper was supported by the National Science Foundation under Grant No. CMS0084808. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

The authors are also grateful to Prof. Eric Mockensturm and three anonymous reviewers for their helpful comments on an earlier draft of this paper.

References

[1] Schajer, G. S., 1984, "The Vibration of a Rotating Circular String Subject to a Fixed Elastic Restraint," J. Sound Vib., **92**(1), pp. 11–19.

[2] Perkins, N. C., and Mote, Jr., C. D., 1986, "Comments on Curve Veering in Eigenvalue Problems," J. Sound Vib., **106**(3), pp. 451–463.

[3] Xiong, Y., and Hutton, S. G., 1994, "Vibration and Stability Analysis of a Multi-Guided Rotating String," J. Sound Vib., **169**(5), pp. 669–683.

[4] Yang, L., and Hutton, S. G., 1995, "Interaction Between an Idealized Rotating String and Stationary Constraints," J. Sound Vib., **185**(1), pp. 139–154.

[5] Tian, J., and Hutton, S. G., 1998, "Lateral Vibration Instability Mechanisms in a Constrained Rotating String," ASME J. Appl. Mech., **65**(3), pp. 774–776.

[6] Tian, J., and Hutton, S. G., 1999, ''On the Mechanisms of Vibrational Instability in a Constrained Rotating String,'' J. Sound Vib., **225**(1), pp. 111–126.
[7] Chen, J.-S., 1997, ''Natural Frequencies and Stability of an Axially-Travelling String in Contact With a Stationary Load System,'' ASME J. Vibr. Acoust., **119**(2), pp. 152–157.
[8] Cheng, S.-P., and Perkins, N. C., 1991, ''The Vibration and Stability of a Friction-Guided, Translating String,'' J. Sound Vib., **144**(2), pp. 281–292.
[9] Lamb, H., 1929, *Dynamics*, 2nd Reprinted Ed., Cambridge University Press, Cambridge, UK.
[10] Miranker, W. L., 1960, ''The Wave Equation in a Medium in Motion,'' IBM J. Res. Dev., **4**(1), pp. 36–42.
[11] Wickert, J. A., and Mote, Jr., C. D., 1990, ''Classical Vibration Analysis of Axially Moving Continua,'' ASME J. Appl. Mech., **57**(3), pp. 738–744.
[12] O'Reilly, O. M., 1996, ''On Steady Motions of a Drawn Cable,'' ASME J. Appl. Mech., **63**(1), pp. 180–189.
[13] Healey, T. J., 1990, ''Stability and Bifurcation of Rotating Nonlinearly Elastic Loops,'' Q. Appl. Math., **48**(4), pp. 679–698.
[14] O'Reilly, O. M., and Varadi, P. C., 1999, ''A Treatment of Shocks in One-Dimensional Thermomechanical Media,'' Continuum Mech. Thermodyn., **11**(6), pp. 339–352.
[15] O'Reilly, O. M., and Varadi, P. C., 2003, ''On Energetics and Conservations for Strings in the Presence of Singular Sources of Momentum and Energy,'' Acta Mech., **126**, pp. 27–45.
[16] Healey, T. J., and Papadopoulos, J. N., 1990, ''Steady Axial Motions of Strings,'' ASME J. Appl. Mech., **57**(3), pp. 785–787.
[17] Reeken, M., 1977, ''On the Equation of Motion of a Chain,'' Math. Z., **155**(3), pp. 219–237.
[18] Purohit, P. K., and Bhattacharya, K., 2003, ''Dynamics of Strings Made of Phase-Transforming Materials,'' J. Mech. Phys. Solids, **51**, pp. 393–424.
[19] Abeyaratne, R., and Knowles, J. K., 1993, ''Nucleation, Kinetics and Admissibility Criteria for Propagating Phase Boundaries,'' *Shock Induced Transitions and Phase Structures in Generalized Media*, J. E. Dunn, R. Fosdick, and M. Slemrod, eds., Springer-Verlag, Berlin, pp. 1–33.
[20] Truskinovsky, L. M., 1993, ''Shocks and Kinks,'' *Shock Induced Transitions and Phase Structures in Generalized Media*, J. E. Dunn, R. Fosdick, and M. Slemrod, eds., Springer-Verlag, Berlin, pp. 185–229.

M. Ravi Shankar

Srinivasan Chandrasekar

School of Industrial Engineering,
Purdue University,
315 N. Grant Street,
West Lafayette, IN 47907-2023

Thomas N. Farris

School of Aeronautics and Astronautics,
Purdue University,
315 N. Grant Street,
West Lafayette, IN 47907-2023

Interaction Between Dislocations in a Couple Stress Medium

Taylor's theory of crystal plasticity is reformulated for dislocations in a couple stress medium. The divergence between Taylor's approach and an approach that includes the effects of couple stresses on dislocation interactions is demonstrated. It is shown that dislocations separated by a distance that is comparable to a characteristic material length scale, have mutual interaction somewhat weaker than that predicted by classical elasticity. [DOI: 10.1115/1.1767172]

1 Introduction

Classical elasticity has been traditionally used to model the stress field in the vicinity of a dislocation. This stress field has provided a basic understanding of the interaction between dislocations. Taylor [1], starting from the classical elasticity theory of dislocations, showed that the flow stress of a material is proportional to the square root of the dislocation density. A finite dislocation density may result either from a random trapping of dislocations (statistically stored dislocations) or due to an accumulation of dislocations for ensuring geometrical compatibility (geometrically necessary dislocations), [2]. The latter phenomenon is important when large strain-gradients occur in a material, [3]. In problems involving a large strain gradient and/or large dislocation densities, the effect of couple stresses on elastic properties of dislocations may be significant.

Mindlin [4,5] showed that the inclusion of couple stresses can introduce significant changes in the elastic stress state of a problem, especially at length scales on the same order as a characteristic ''material length scale.'' For instance, the elastic stress concentration factor of a circular hole in an infinite sheet under uniaxial tension is somewhat lower in a couple stress mediun than that predicted otherwise. This divergence becomes particularly conspicuous when the size of the hole is comparable to the characteristic length scale of the medium, [5]. Dislocations, being an atomic scale phenomenon should, hence, be expected to be affected by the presence of couple stresses.

Mindlin [4,5] suggested a number of possible sources of the characteristic material length scale depending on the type of the problem, with values of the length scale ranging any where from the lattice spacing in a perfect crystalline lattice to the grain size in a polycrystalline material. In this study, the length scale (l) is assumed to be equal to 500 nm (Table 1), which is typical of the cell size in a highly work hardened material. While the characteristic length scale enters the calculations as a parameter that decides the range and strength of the couple stress terms, its specific value is not likely to affect the general conclusions that are arrived at in the ensuing analysis.

Here, couple stresses are included in the formulation of the problem of an edge dislocation. It is shown that the presence of couple stresses introduces a so-called ''weak-interaction'' parameter in the elastic stress field that is operative only at very small distances and whose principal effect is to depress the stress field around the dislocation. On the basis of this observation, it is inferred that the stress required to operate a Frank-Read source in a couple stress medium is generally smaller than that required otherwise. This is not to mean that there is a softening of the material at large dislocation densities. What is suggested is that as the dislocation density increases, strain gradient effects appear to reduce the ability of dislocations to pin each other vis-a-vis the classical solution.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, May 22, 2003; final revision; December 17, 2003. Associate Editor: M.-J. Pindera. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

2 Couple Stress Theory

The stresses and couple stresses on an element in equilibrium are shown in Fig. 1. Couple stresses or couple per unit area are denoted by μ_r and μ_θ. The remaining stresses have the usual meaning. For this two-dimensional formulation, these stresses can be derived from stress functions Φ and Ψ satisfying [5]:

$$\nabla^4\Phi = 0 \tag{1}$$

$$\frac{\partial}{\partial r}(\Psi - l^2\nabla^2\Psi) = -2(1-\nu)l^2\frac{1}{r}\frac{\partial}{\partial\theta}\nabla^2\Phi \tag{2}$$

$$\frac{1}{r}\frac{\partial}{\partial\theta}(\Psi - l^2\nabla^2\Psi) = 2(1-\nu)l^2\frac{1}{r}\frac{\partial}{\partial\theta}\nabla^2\Phi \tag{3}$$

$$\nabla^2\Psi - l^2\nabla^4\Psi = 0 \tag{4}$$

where, ν is the Poisson's ratio, and l is a length scale characteristic of the material defined as [5]:

$$l^2 = \frac{M}{G}. \tag{5}$$

G and M, respectively, are the shear modulus and bending modulus of the material.

The stresses corresponding to Ψ and Φ are

$$\sigma_r = \frac{1}{r}\frac{\partial\Phi}{\partial r} + \frac{1}{r^2}\frac{\partial^2\Phi}{\partial\theta^2} - \frac{1}{r}\frac{\partial^2\Psi}{\partial r\partial\theta} + \frac{1}{r^2}\frac{\partial\Psi}{\partial\theta} \tag{6}$$

$$\sigma_\theta = \frac{\partial^2\Phi}{\partial r^2} + \frac{1}{r}\frac{\partial^2\Psi}{\partial r\partial\theta} - \frac{1}{r^2}\frac{\partial\Psi}{\partial\theta} \tag{7}$$

$$\tau_{r\theta} = -\frac{1}{r}\frac{\partial^2\Phi}{\partial r\partial\theta} + \frac{1}{r^2}\frac{\partial\Phi}{\partial\theta} - \frac{1}{r}\frac{\partial\Psi}{\partial r} - \frac{1}{r^2}\frac{\partial^2\Psi}{\partial\theta^2} \tag{8}$$

$$\tau_{\theta r} = -\frac{1}{r}\frac{\partial^2\Phi}{\partial r\partial\theta} + \frac{1}{r^2}\frac{\partial\Phi}{\partial\theta} + \frac{\partial^2\Psi}{\partial r^2} \tag{9}$$

$$\mu_r = \frac{\partial\Psi}{\partial r} \tag{10}$$

Table 1 Material parameters assumed for copper

Shear modulus (G)	44 GPa
Poisson's ratio (ν)	0.34
Length of the dislocation (λ)	0.256 nm
Material length scale (l)	500 nm

$$\mu_\theta = \frac{1}{r}\frac{\partial \Psi}{\partial \theta}. \tag{11}$$

Mindlin [4,5] has provided a detailed derivation of Eqs. (1)–(11).

3 Stress Functions for a Dislocation in a Couple Stress Medium

Consider the problem of a dislocation similar to the one considered by Taylor [1], i.e., an elastic isotropic cylinder with a dislocation as shown in Fig. 2.

If the radius R of the cylinder is large, then the Airy stress function (Φ) is

$$\Phi = -Dr \log r \sin\theta - \frac{B \sin\theta}{r}. \tag{12}$$

Couple stress effects can be included by introducing Ψ:

$$\Psi = \frac{E}{r}\cos\theta + FK_1\left(\frac{r}{l}\right)\cos\theta \tag{13}$$

where, D, B, E, and F are constants that can be calculated subject to the boundary conditions. $K_n(x)$ is a modified Bessel function of the second kind and the nth order.

The stress function in Eq. (13) is similar, if not the same as that postulated by Mindlin [5] for the problem of a circular hole in an infinite sheet subjected to uniaxial tension in a couple stress medium.

At $r = a$, the following boundary conditions hold:

1. $\mu_r|_{r=a} = 0$: It may be shown using Eqs. (13), (12), and (10) that in the limit $a \to 0$, $F \to -E/l$.
2. $\tau_{r\theta}|_{r=a} = 0$: It may be shown using Eqs. (13), (12), and (8) that this boundary condition and the following boundary condition lead to the same criterion.
3. $\sigma_r|_{r=a} = 0$: It may be shown using Eqs. (13), (12), and (6) that in the limit $a \to 0$, $B \to 0$.

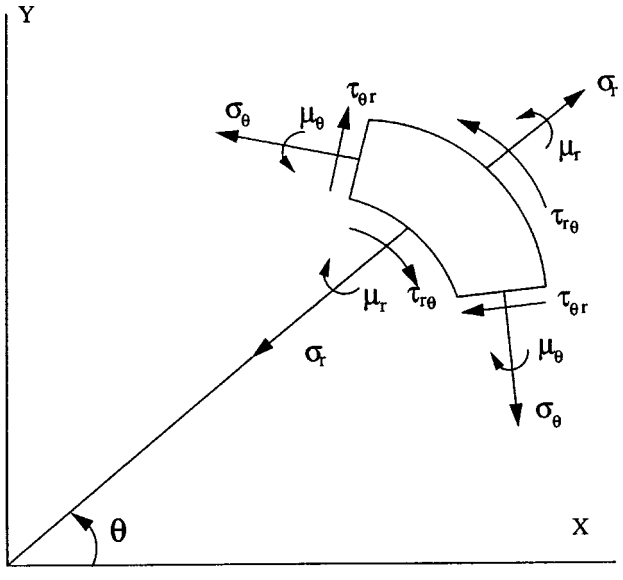


Fig. 1 Stress and couple stress in polar coordinates

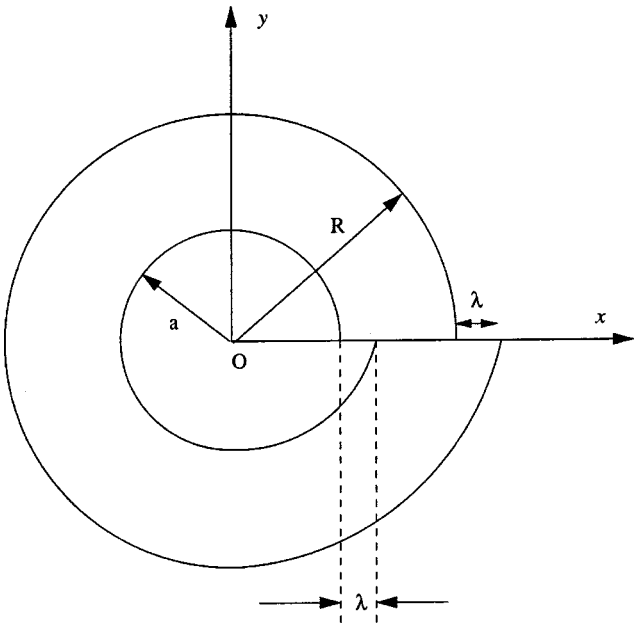


Fig. 2 Elastic cylinder with dislocation

Equations (2) and (3) lead to a 4th relationship that is intrinsic to the ensemble:

$$E = -4(1-\nu)l^2 D. \tag{14}$$

D may be evaluated from the fact that at $\theta = 2\pi$, the relative displacement in the r-direction, $u_r|_{\theta=2\pi} - u_r|_{\theta=0} = \lambda$ (Fig. 2). Hence,

$$D = \frac{G\lambda}{2\pi(1-\nu)}. \tag{15}$$

The coefficients D, E, and F were calculated for material parameters typical of copper (Table 1) and are given in Table 2. Note that the coefficients D and F are positive while E is negative.

4 Taylor's Theory in a Couple-Stress Medium

Consider two edge dislocations of opposite signs at O and A in an infinite medium devoid of couple stress, whose glide-planes are spaced a distance h apart as shown in Fig. 3. Let us assume the dislocations are kept apart under the action of a remote shear stress S acting along the glide plane. It was suggested by Taylor that this shear stress would pull the two dislocations apart a distance x (Fig. 3) such that

$$S = \frac{Dx}{(x^2+h^2)}. \tag{16}$$

It was further argued that the maximum (or critical) value of S is $D/(2h)$ and if S were less than this critical value, then the two centers of dislocation cannot escape their mutual attraction, [1].

A similar definition of the critical shear stress needed to pull apart two dislocations of opposite signs may be attempted for a couple stress medium as follows. Since $\tau_{\theta r}$ and $\tau_{r\theta}$ are not equal in a couple-stress medium, we define two stress parameters τ_s and τ_a as

$$\tau_s = (\tau_{\theta r} + \tau_{r\theta})/2$$

Table 2 Coefficients for copper

D	2.7162 N/m
E	-1.79×10^{-12} Nm
F	3.58×10^{-6} N

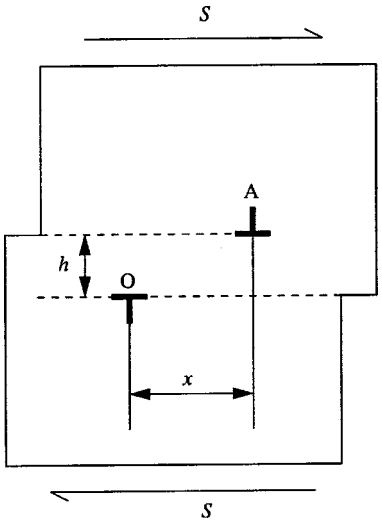


Fig. 3 Positive and negative dislocations in the presence of a remote shear stress

and

$$\tau_a = (\tau_{\theta r} - \tau_{r\theta})/2.$$

Then, using the stress functions in Eqs. (12) and (13), we get

$$\tau_s = \frac{1}{2}\left(\frac{2(2E+Dr^2)}{r^3} + \frac{FK_1(r/l)}{l^2} + \frac{2FK_2(r/l)}{lr}\right)\cos\theta \quad (17)$$

$$\tau_a = \frac{FK_1(r/l)\cos\theta}{2l^2}. \quad (18)$$

The Bessel function contribution to the shear stresses in Eq. (17) is negligible compared to the contribution from the other two terms involving D and E even at a distance of $r \sim 3l$ from the dislocation. Hence, we will not consider the contribution from terms involving the Bessel function $K_n(r/l)$.

We take τ_s to be the analog of S in Fig. 3. This analogy arises directly as a result of Mindlin's formulation, [4]. In this formulation, the symmetric part of the shear stress (τ_s) produces the usual shear strain ($\gamma_{r\theta}$), that in classical elasticity is produced by $\sigma_{r\theta}$ (of course, in classical elasticity $\sigma_{r\theta} = \sigma_{\theta r}$).

That is

$$\gamma_{r\theta} = \gamma_{\theta r} = \frac{\tau_s}{G}. \quad (19)$$

Thus, to a first approximation, if a comparison is to be made between the classical result of Taylor and one involving couple stresses, it has to be between S and τ_s. In this approximation we have, of course, neglected the terms arising from the interaction of the curvatures in the vicinity of the dislocation at A with the couple stresses produced by the dislocation at O in Fig. 3, since both of them are rapidly decaying quantities.

By discarding the terms involving the Bessel functions in Eq. (17) we get

$$\tau_s = \frac{D\cos\theta}{r} + \frac{2E\cos\theta}{r^3}. \quad (20)$$

It can be immediately seen using Eq. (20) and Table 2 that, between two dislocations of opposite signs, the term involving E in Eq. (20) is a repulsive interaction. This term leads to a ''weak interaction'' that gains significance at large dislocation densities when the distances between dislocations are small. The term involving D in Eq. (20) is the ''Taylor interaction'' which causes two dislocations of opposite signs to always attract each other.

We can rewrite τ_s (Eq. (20)) in Cartesian coordinates as

$$\tau_s = \frac{Dx}{x^2+y^2} + \frac{2Ex}{(x^2+y^2)^2}. \quad (21)$$

The equivalent of Eq. (16) for Fig. 3 in a couple stress medium can be obtained by setting $y = h$ in Eq. (21):

$$S_{\text{couple-stress}} = \frac{Dx}{x^2+h^2} + \frac{2Ex}{(x^2+h^2)^2}. \quad (22)$$

5 Weak Interaction Between Dislocations

Equation (22) can be broken into two parts, namely the Taylor interaction and the weak interaction. That is

$$S_{\text{couple-stress}} = T(x) + R(x) \quad (23)$$

where $T(x) = Dx/(x^2+h^2)$, is the Taylor interaction (same as Eq. (16)), and $R(x) = 2Ex/(x^2+h^2)^2$, is the repulsive weak interaction.

S' and $S'_{\text{couple-stress}}$, the nondimensional interaction shear stresses, are defined as the nondimensionalized counterparts of S and $S_{\text{couple-stress}}$ such that

$$S' = \frac{h}{G\lambda}S \quad (24)$$

$$S'_{\text{couple-stress}} = \frac{h}{G\lambda}S_{\text{couple-stress}}. \quad (25)$$

If the distance h, between the glide planes is very large compared to the material length scale parameter l, i.e., say $h > 10l$, then $R(x)$ is negligible and the solution for $S'_{\text{couple-stress}}$ practically coincides with that obtained from Taylor's theory (S'). This is apparent from Fig. 4 where the variation of the nondimensional interaction shear stresses (S' and $S'_{\text{couple-stress}}$) with (x/h) is plotted for $h = 10l$. However, when the distance h, between the glide planes becomes comparable to l, i.e., say $h = 3l$, then the contribution of the repulsive weak interaction term, $R(x)$, in Eq. (23) is seen to be significant (Fig. 5).

The effect of the $R(x)$ term, hence, appears to be one of reducing the interaction stress between dislocations. For $h = 3l$ this effect is seen to be $\sim$25% based on the maximum value of the nondimensional interaction stress (Fig. 5).

The variation of the maximum value of S' and $S'_{\text{couple-stress}}$ with (h/l) is plotted in Fig. 6. Since Taylor's theory does not involve a characteristic material length scale, the maximum value of S' has no dependence on the value of h/l, and hence, is a constant equal to $\sim$0.12. However, in the presence of couple stresses it can be seen (Fig. 6) that the maximum value of $S'_{\text{couple-stress}}$ decreases as h/l decreases.

6 Consequences of Weak Interaction for Strain-Gradient Plasticity

Classically, the line tension of a dislocation line is proportional to the elastic self energy of the dislocation line. The stronger the elastic stress field around a dislocation, the stronger it interacts with other dislocations and the larger is the stress required to bend a dislocation line to the critical extent. Cottrell [6] has calculated the elastic self energy of an edge dislocation by treating it as the work done in displacing the faces of a cut (made as in Fig. 2) by λ against the resistance of the shear stress field of the dislocation. While such an approach yields the correct result in classical elasticity, similar simplistic calculations in the presence of couple stresses will lead to grossly misleading results. Since exact solutions for the self energy and line tension are extremely complicated, we will only attempt to qualitatively gauge the effect of couple stresses on dislocation interactions.

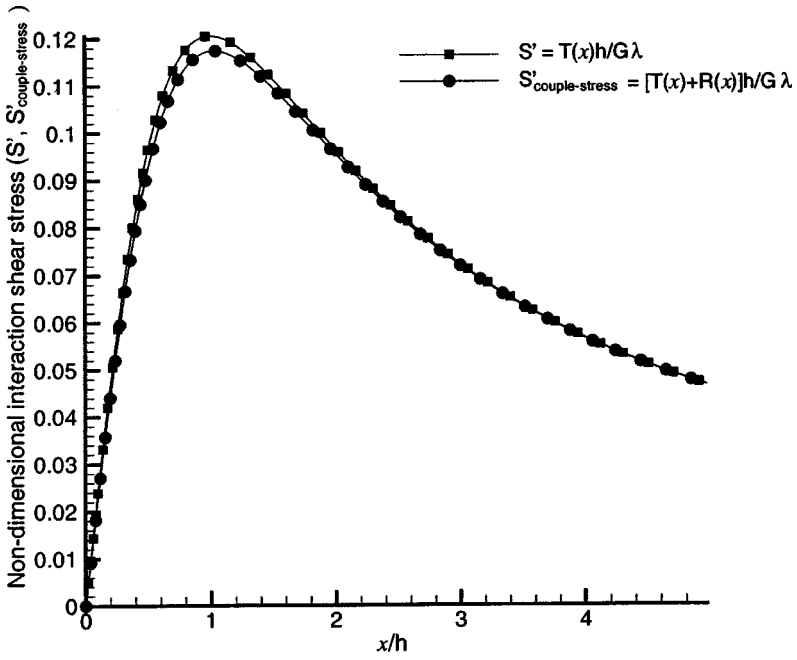


Fig. 4 Variation of nondimensional interaction shear stress with x/h for $h=10l$. The material parameters used are those for copper given in Tables 1 and 2.

The maximum value of the interaction shear stress is representative of the stress required to disentangle a pair of dissimilar dislocations. As illustrated in Fig. 6, the maximum value of $S'_{\text{couple-stress}}$ becomes smaller as h/l decreases. Hence, the stress required to disentangle a pair of dissimilar dislocations is generally smaller than that predicted by classical elasticity when the spacing between the dislocations becomes comparable to the characteristic material length scale (l).

This would imply that in the case of a large dislocation density, as is typical of highly strained materials or of many problems involving a significant strain gradient, when the distances between dislocations become comparable to the characteristic length scale of the material, the presence of couple stresses would tend to reduce the interaction between dislocations and reduce the ability of dislocations to pin each other. However, it is not likely that this ''softening'' produced by the reduced dislocation interactions will be comparable in magnitude to the hardening produced by the large dislocation density.

While there have been attempts at understanding the effect of couple stresses close to the dislocation core by Eringen [7], Gutkin and Aifantis [8], and others, the importance of the corrections resulting from strain gradient effects vis-a-vis anharmonic effects

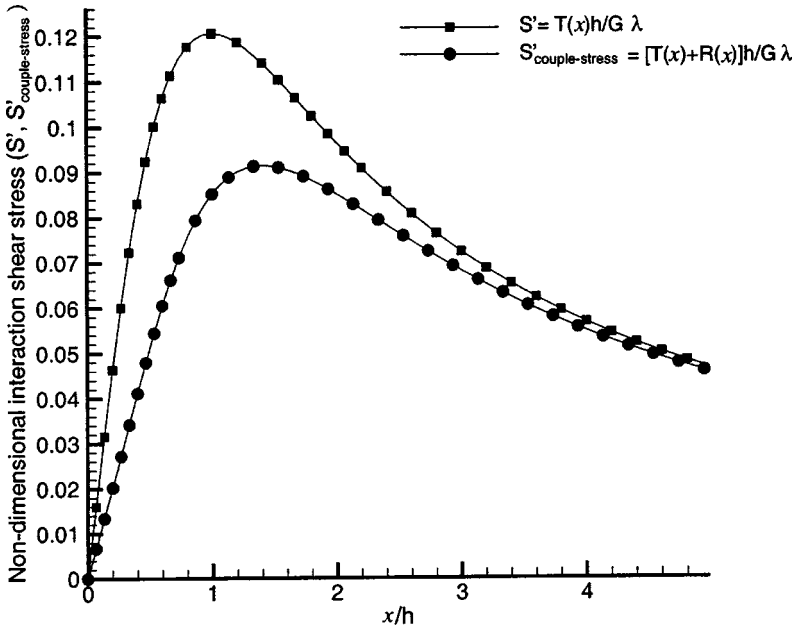


Fig. 5 Variation of nondimensional interaction shear stress with x/h for $h=3l$. The material parameters used are those for copper in Tables 1 and 2.

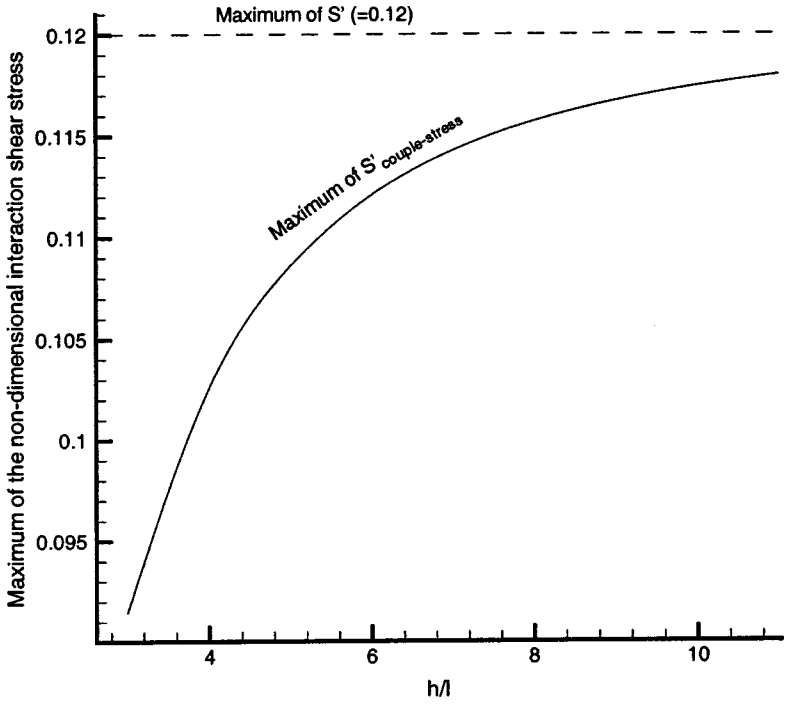


Fig. 6 Variation of the maximum nondimensional interaction shear stress with *h*/*l*. The material properties used are those for copper given in Tables 1 and 2.

close to the core remain questionable. Hence, we have confined ourselves here to trying to understand the effect of couple stresses on dislocation interactions occuring in regions far from the dislocation core where simplifying assumptions can be made to make the problem tractable.

Due to the mathematical complexity involved, it is not attempted in this study to explicitly derive the expression for line tension in a couple stress medium. However, the results derived are sufficient to predict the nature of the change in interactions between dislocations in a couple stress medium at large dislocation densities.

7 Conclusions

Taylor's theory of plasticity has been reformulated for a couple-stress medium with a characteristic material length scale. The analysis has revealed the following:

1. The shear stress required to disentangle a pair of dissimilar dislocations is estimated to be smaller when couple stresses are considered. This reduction is attributed to the presence of a repulsive element $R(x)$ (Eq. (20)) in the interaction shear stress.
2. $R(x)$ becomes significant only at large dislocation densities ($h<10l$ where l is a material length scale).
3. The proposed model degenerates to the Taylor model at low dislocation densities where the average distance between dislocations is much greater than the material length scale.
4. At large dislocation densities, the ability of dislocations to pin each other is estimated to be smaller than that predicted by classical elasticity. This by no means predicts any significant softening of materials at large dislocation densities. But, materials will behave somewhat softer than expected at large dislocation densities.

Acknowledgment

We would like to thank the reviewers for their very constructive comments and suggestions. The support of the National Science Foundation through grants CMS 0200509 and DMI 0115467 is gratefully acknowledged.

References

[1] Taylor, G. I., 1934, "The Mechanism of Plastic Deformation of Crystals," Proc. R. Soc. London, Ser. A, **145**, pp. 362–404.
[2] Ashby, M. F., 1970, "The Deformation of Plastically Non-Homogeneous Alloys," Philos. Mag., **21**, pp. 399–424.
[3] Fleck, N. A., Muller, G. M., Ashby, M. F., and Hutchinson, J. W., 1994, "Strain Gradient Plasticity: Theory and Experiment," Acta Metall. Mater., **42**(2), pp. 475–487.
[4] Mindlin, R. D., and Tiersten, H. F., 1962, "Effects of Couple Stresses in Linear Elasticity," Arch. Ration. Mech. Anal., **11**, pp. 415–448.
[5] Mindlin, R. D., 1963, "Influence of Couple-Stresses on Stress Concentrations," Exp. Mech., **3**, pp. 1–7.
[6] Cottrell, A. H., 1956, *Dislocations and Plastic Flow in Crystals*, Oxford University Press, Oxford, UK.
[7] Eringen, A. C., 1983, "On Differential Equations of Nonlocal Elasticity and Solutions of Screw Dislocations and Surface Waves," J. Appl. Phys., **54**(9), pp. 4703–4710.
[8] Gutkin, M. Y., and Aifantis, E. C., 1999, "Dislocations in the Theory of Gradient Elasticity," Scr. Mater., **40**, pp. 559–566.

Zhongping Bao
Graduate Student
e-mail: zb22@cornell.edu

Subrata Mukherjee
Professor
Fellow ASME
e-mail: sm85@cornell.edu

Department of Theoretical and Applied Mechanics,
Cornell University,
Ithaca, NY 14853

Max Roman
Graduate Student
e-mail: mxr6074@njit.edu

Nadine Aubry
Professor
Mem. ASME
e-mail: aubry@njit.edu

Department of Mechanical Engineering,
New Jersey Institute of Technology,
Newark, NJ 07102

Nonlinear Vibrations of Beams, Strings, Plates, and Membranes Without Initial Tension

The subject of this paper is nonlinear vibrations of beams, strings (defined as beams with very thin uniform cross sections), plates and membranes (defined as very thin plates) without initial tension. Such problems are of great current interest in minute structures with some dimensions in the range of nanometers (nm) to micrometers (μm). A general discussion of these problems is followed by finite element method (FEM) analyses of beams and square plates with different boundary conditions. It is shown that the common practice of neglecting the bending stiffness of strings and membranes, while permissible in the presence of significant initial tension, is not appropriate in the case of nonlinear vibrations of such objects, with no initial tension, and with moderately large amplitude (of the order of the diameter of a string or the thickness of a plate). Approximate, but accurate analytical expressions are presented in this paper for the ratio of the nonlinear to the linear natural fundamental frequency of beams and plates, as functions of the ratio of amplitude to radius of gyration for beams, or the ratio of amplitude to thickness for square plates, for various boundary conditions. These expressions are independent of system parameters—the Young's modulus, density, length, and radius of gyration for beams; the Young's modulus, density, length of side, and thickness for square plates. (The plate formula exhibits explicit dependence on the Poisson's ratio.) It is expected that these results will prove to be useful for the design of macro as well as micro and nano structures. [DOI: 10.1115/1.1767167]

1 Introduction

Vibrations of structures with some dimensions in the range of nanometers (nm) to micrometers (μm), is of great current interest in applications in nano and micro-electro-mechanical systems (NEMS and MEMS). Examples are vibrations of carbon nanotubes, [1], and nanowires; and vibrations of very thin (usually silicon or polysilicon) membranes, [2]. Applications of the latter include microspeakers, [3], and synthetic microjets (e.g., for mixing, cooling of electronic components, micropropulsion, and flow control).

Figure 1 (from McEuen [4]) shows (*a*) a schematic and (*b*) a scanning electron microscope (SEM) image of a carbon nanotube suspended over a pre-existing trench in the substrate. Typically, a single-walled carbon nanotube (SWNT) is a hollow tube that is about 2 nm in diameter and .335 nm in thickness. Also, it generally has very small (if any) initial tension. The typical thickness of a $3\,\text{mm}\times 3\,\text{mm}$ square polysilicon plate, in an electrostatically actuated synthetic microjet, is of the order of 2 μm, [2]. Therefore, this object is really a thin membrane. Also, it generally has no initial tension. Finally, vibrations of such strings and membranes without initial tension, in MEMS and NEMS applications, are generally nonlinear (moderately large amplitude) since the amplitudes are of the order of the diameter (for strings) or of the thickness (for membranes).

Consider a macroscale beam made of a linearly elastic material, with uniform cross section, and let its radius of gyration be very small relative to its length. Typically, such an object is called a string in the literature, and its bending stiffness is neglected. It is shown in the present paper that neglecting of the bending stiffness, while permissible in the presence of significant initial tension, is not appropriate when nonlinear vibrations of a thin beam, with no initial tension, and with moderately large amplitude (of the order of its radius of gyration), is considered. Similarly, a thin plate is called a membrane in the literature and the bending stiffness of a membrane is typically neglected. It is also shown in the present paper that neglecting of the bending stiffness is not appropriate when nonlinear vibrations of a thin plate, with no initial tension, and with moderately large amplitude (of the order of its thickness), is considered

The primary objective of the present paper is to obtain natural frequencies of such vibrating strings and membranes, with no initial tension, as functions of its amplitude of vibration. Continuum mechanics modeling is used throughout and it is understood that the results presented here are only valid within this assumption. Typically, results from continuum mechanics compare well with experiments for objects with some dimensions of the order of microns (μm) or larger, and start deviating from experiments for lower values. Under some circumstances, however, continuum theory (surprisingly) performs well even for objects with diameters in the order of 1–2 nm (such as carbon nanotubes), [1].

Nonlinear vibrations of standard beams and plates is of course, a well-studied subject, and numerous papers and books exist in this area in the literature (see, e.g., Nayfeh and Mook [5]). Only a few representative samples are cited here. A very nice critical review of studies of large amplitude free vibrations of beams is presented by Singh et al. [6]. Researchers have employed various models such as inclusion or exclusion of axial displacement, and linear or nonlinear kinematic (strain-displacement) relations. Starting with the work of Woinosky-Krieger [7] who used elliptic integrals, researchers have used various methods such as perturbation, Ritz-Galerkin and the finite element method (FEM, see, e.g., [8]) to study this problem. The reader is referred to [6] for appropriate references. A good general reference for the nonlinear vibration of plates is Chia [9], while a nice study of nonlinear

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, August 12, 2003; final revision, January 21, 2004. Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

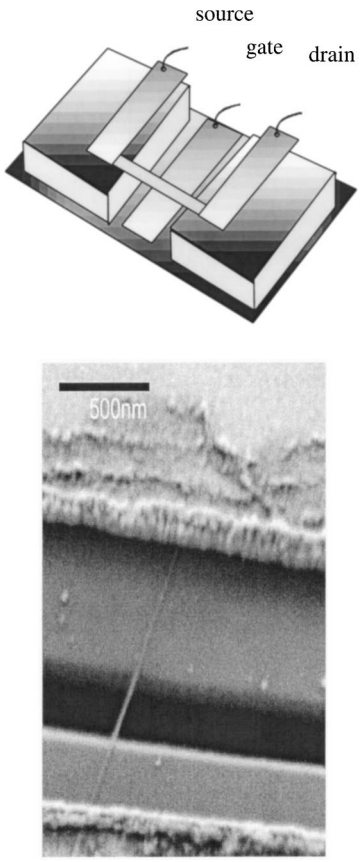


Fig. 1 A suspended nanotube between two electrical contacts. (*a*) Schematic (*b*) SEM image.

vibrations of plates with the FEM is Han and Petyt [10]. The FEM is the primary tool that is employed in order to carry out the numerical analyses in the present paper.

Section 2 of the present paper discusses nonlinear vibrations of beams and strings. A brief description of the adopted model, a well-known analytical solution for a beam that is simply-supported at both ends, and a FEM formulation for a beam with general boundary conditions follow. It is shown that the corresponding eigenvalue problem leads to an unique variation of ω/ω_0, as a function of A/ξ, that is independent of system parameters, for various boundary conditions. Here, ω is the nonlinear (amplitude dependent) fundamental natural frequency, ω_0 is the corresponding natural frequency from the linear theory, A is the vibration amplitude (at the center of the beam for a beam with symmetric boundary conditions) and ξ is the radius of gyration of the beam. FEM results are presented and discussed for beams with immovable end-points that are either simply-supported at both ends or clamped at both ends.

Section 3 presents nonlinear vibrations of square plates and membranes of rectangular cross sections. The outline here is quite similar to the previous section on beams and strings. The variation of ω/ω_0 as a function of A/h is considered this time. (Here h is the plate thickness and A is the vibration amplitude at the center for a plate with symmetric boundary conditions.) FEM results are presented and discussed for three kinds of immovable boundary conditions—clamped on all edges (C), simply-supported on all edges (S), or simply-supported on two opposite parallel edges and clamped on the other two (SC).

Approximate, but accurate analytical expressions are presented in this paper for ω/ω_0 as a function of A/ξ (for beams) or A/h (for square plates), for various boundary conditions. These expressions are independent of system parameters—the Young's modulus, density, length, and radius of gyration for beams; the Young's modulus, density, length of side and thickness for square plates. The Poisson's ratio appears explicitly in the plate formula. It is expected that these results will prove to be useful for the design of macro as well as micro and nano structures.

2 Beams and Strings

Nonlinear vibration of beams, without initial tension, are discussed in this section. The beams are linearly elastic, have immovable ends and are of uniform cross section. The cross section is symmetric such that there is no twisting of the beam under applied bending moments.

2.1 The Model. Following [6], the kinematic equations adopted here, for a beam with immovable ends, are

$$\epsilon_{xx} = u_{,x} + (1/2)(w_{,x})^2 \tag{1}$$

$$\kappa_x = -w_{,xx} \tag{2}$$

where $u(x,t)$ and $w(x,t)$ are the axial and transverse displacements of the beam, respectively, ϵ_{xx} is the midplane axial strain and κ_x is the curvature. Here, $,x$ denotes a derivative with respect to the axial coordinate x.

The strain energy $\mathcal{E}^{(s)}$ and the kinetic energy $\mathcal{E}^{(k)}$ of an uniform beam of length L are

$$\mathcal{E}^{(s)} = \frac{ES}{2}\int_0^L [(u_{,x})^2 + u_{,x}(w_{,x})^2 + (1/4)(w_{,x})^4]dx + \frac{EI}{2}\int_0^L (w_{,xx})^2 dx \tag{3}$$

$$\mathcal{E}^{(k)} = \frac{\rho S}{2}\int_0^L [(\dot{u})^2 + (\dot{w})^2]dx \tag{4}$$

where E, ρ, L, S, and I are the Young's modulus, density (mass per unit volume), length, area of cross section, and area moment of inertia of the cross section of the beam, respectively, and a superposed dot denotes differentiation with respect to time t.

2.2 Analytical Solution for Simply-Supported Beam. A very important result relevant to this work is the analytical solution for a beam of uniform cross section, that is *simply-supported at both ends*, [11,6]:

$$\frac{\omega}{\omega_0} = \sqrt{1 + \alpha\left(\frac{A}{\xi}\right)^2} \tag{5}$$

where $\alpha = 3/16$, and $\xi = \sqrt{I/S}$ is the radius of gyration of the beam cross-section.

Also, in this case, [12],

$$\omega_0 = \frac{\pi^2 \xi}{L^2}\sqrt{\frac{E}{\rho}}. \tag{6}$$

It is important to note, for what follows, that Eq. (5) provides a "master curve" that is valid for all simply-supported beams with parameters E, L, ρ, and ξ. Thus, it is applicable even for very small ξ (e.g., a nanotube or a nanowire) provided, of course, that the assumptions used to construct this theoretical model remain valid. It is shown later in this paper that the formula (5), with an appropriate value of α, remains valid (within a very good approximation) for clamped-clamped beams as well. A similar discussion for plates follows in Section 3.

2.3 FEM Model for Beams With Immovable Ends. The procedure followed in this section, for FEM discretization of vibrating beams, is standard (see, e.g., Zienkiewicz and Taylor [13]). Some details are necessary, however, for the discussion of eigenvalue problems that follows.

FEM Discretization. Each beam element has two end nodes with three degrees-of-freedom at each node. These are u, w, and $\theta = w_{,x}$. For each element, one has

$$\begin{bmatrix} u \\ w \end{bmatrix} = \begin{bmatrix} N^{(I)} & 0 \\ 0 & N^{(O)} \end{bmatrix} \begin{bmatrix} q^{(I)} \\ q^{(O)} \end{bmatrix} \tag{7}$$

with

$$[N^{(I)}(x)] = [N_1, N_2], \quad [N^{(O)}(x)] = [P_1, P_2, P_3, P_4] \tag{8}$$

$$[q^{(I)}(t)] = [u_1, u_2]^T, \quad [q^{(O)}(t)] = [w_1, \theta_1, w_2, \theta_2]^T. \tag{9}$$

Here N_k and P_k are the usual linear and cubic (Hermite polynomial) interpolation functions, and $[q^{(I)}]$ and $[q^{(O)}]$ contain the appropriate nodal degrees-of-freedom.

Define

$$D = w,x, \quad [G] = [N^{(O)}_{,x}], \quad [B^{(I)}] = [N^{(I)}_{,x}], \quad [B^{(O)}] = -[N^{(O)}_{,xx}]. \tag{10}$$

Substituting the interpolations (7) into the energy expressions (3), (4), and applying Hamilton's principle, results in the element level equations:

$$\begin{bmatrix} M^{(I)} & 0 \\ 0 & M^{(O)} \end{bmatrix} \begin{bmatrix} \ddot{q}^{(I)} \\ \ddot{q}^{(O)} \end{bmatrix} + \begin{bmatrix} K^{(I)} & 0 \\ 0 & K^{(O)} \end{bmatrix} \begin{bmatrix} q^{(I)} \\ q^{(O)} \end{bmatrix} + \begin{bmatrix} 0 & K^{(IO)} \\ 2K^{(IO)T} & K^{(NI)} \end{bmatrix} \times \begin{bmatrix} q^{(I)} \\ q^{(O)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{11}$$

The various submatrices in (11) are (here ℓ the length of an element):

$$[M^{(I)}] = \frac{\rho S}{2} \int_0^\ell [N^{(I)}]^T [N^{(I)}] dx,$$

$$[M^{(O)}] = \frac{\rho S}{2} \int_0^\ell [N^{(O)}]^T [N^{(O)}] dx \tag{12}$$

$$[K^{(I)}] = ES \int_0^\ell [B^{(I)}]^T [B^{(I)}] dx,$$

$$[K^{(O)}] = EI \int_0^\ell [B^{(O)}]^T [B^{(O)}] dx \tag{13}$$

$$[K^{(IO)}] = \frac{ES}{2} \int_0^\ell [B^{(I)}]^T [DG] dx,$$

$$[K^{(NI)}] = \frac{ES}{2} \int_0^\ell [DG]^T [DG] dx. \tag{14}$$

Note that the in-plane (axial) and out-of-plane (bending) matrices $[K^{(I)}]$ and $[K^{(O)}]$ are $\propto S$ and $S\xi^2$, respectively, the matrix $[K^{(IO)}] \propto AS$ represents coupling between the axial and bending displacements, and the matrix $[K^{(NI)}] \propto A^2 S$ arises purely from the nonlinear axial strains.

It is well known that for the linear theory $[K^{(O)}] \ll [K^{(I)}]$ as $\xi \to 0$. It is very interesting, however, to note that if A/ξ remains $\mathcal{O}(1)$, the bending matrix $[K^{(O)}]$, which arises from the linear theory, and the matrix $[K^{(NI)}]$ from the nonlinear theory, remain of the same order as $\xi \to 0$. This fact has important consequences for the modeling of nanotubes and nanowires, as can be seen in the following section.

Eigenvalue Problem. Following the procedure detailed in [10] (this approach neglects axial inertia, see the Appendix), the following eigenvalue equations are obtained:

$$([K^{(O)}] + [K^{(NL)}] - \omega^2 [M^{(O)}])[q^{(O)}] = [0] \tag{15}$$

where

$$[K^{(NL)}] = (3/4)([K^{(NI)}] - 2[K^{(IO)}]^T [K^{(I)}]^{-1} [K^{(IO)}]). \tag{16}$$

The next task is to investigate the dependence of the fundamental eigenvalue of (15) on system parameters. To this end, the system of Eqs. (15) is modified such that each element of a given global matrix has the same physical dimension. This is done as follows.

The vector $[q^{(O)}] = [w_1, \theta_1, w_2, \theta_2]^T$ is not dimensionally homogeneous. It can, however, be written in homogeneous form as $[w_1, L\theta_1, w_2, L\theta_2]$ and the appropriate columns of $[K^{(O)}], [K^{(NL)}], [M^{(O)}]$, suitably divided by L. Also, certain rows of the entire system (15) are divided by L. One now has, for example,

$$\hat{K}^{(O)}_{11} = \frac{12EI}{\ell^3}, \quad \hat{K}^{(O)}_{12} = \hat{K}^{(O)}_{21} = \frac{6EI}{L\ell^2}. \tag{17}$$

Next, each element of these matrices are further scaled (multiplied, appropriately, by suitable powers of L/L), so that one gets, for example,

$$\hat{K}^{(O)}_{11} = \frac{12EI}{L^3} \left(\frac{L}{\ell} \right)^3, \quad \hat{K}^{(O)}_{12} = \hat{K}^{(O)}_{21} = \frac{6EI}{L^3} \left(\frac{L}{\ell} \right)^2. \tag{18}$$

It is assumed that the ratio ℓ_k/L (where ℓ_k is the length of the k_{th} beam element) is kept fixed, even if the length L of the beam is changed (isotropic expansion or contraction). Next, the global matrices are formed and appropriate boundary conditions (simply-supported or clamped are considered in this work, together with $u = 0$ at both ends of the beam) are imposed. One now has

$$[\hat{K}^{(I)}]^{(G)} \propto \frac{ES}{L}, \quad [\hat{K}^{(IO)}]^{(G)} \propto \frac{EAS}{L^2}$$

$$[\hat{K}^{(O)}]^{(G)} \propto \frac{ES\xi^2}{L^3}, \quad [\hat{K}^{(NL)}]^{(G)} \propto \frac{EA^2S}{L^3}, \quad [\hat{M}^{(O)}]^{(G)} \propto \rho SL. \tag{19}$$

Multiply the new global eigenvalue equation by $L^3/(ES\xi^2)$, note the forms of the matrices $[\hat{K}^{(O)}]^{(G)}, [\hat{K}^{(NL)}]^{(G)}, [\hat{M}^{(O)}]^{(G)}$, and that $1/\omega_0^2 \propto (\rho L^4)/(E\xi^2)$ for beams with many different boundary conditions, [12]. Also, let $A/\xi = \gamma$, a fixed number. Now one gets the equation

$$([M_1] - (\omega^2/\omega_0^2)[M_2])[v] = [0] \tag{20}$$

in which the matrices $[M_1], [M_2]$ are *independent* of the system (beam) parameters E, ρ, L, ξ; the dependence on these parameters being wholly captured in ω_0^2 (provided, of course, that A/ξ remains fixed.) Therefore, for any values of the system parameters (within the limits of validity of the beam model employed in this work), and with a fixed value of A/ξ, the eigenvalue problem (20) yields a *single unique fundamental eigenvalue*, here called $\hat{\omega}^2/\omega_0^2$. In other words, a point on a plot of ω/ω_0 versus A/ξ, for a fixed value of A/ξ, is *invariant* with respect to variations of the system parameters. Equivalently, a plot of ω/ω_0, as a function of A/ξ, is *invariant* with respect to the system parameters E, ρ, L, ξ, for many different boundary conditions, including beams that are simply-supported at both ends (of course, this is a known fact—see (5)), or *clamped at both ends*.

Please note that $[M_1]$ is a function of $[v]$ with $\max_k |v_k| = 1$, so that iterations are needed to solve (20).

It is now conjectured that the form of Eq. (5) remains valid, of course with a different value of α, for clamped-clamped beams as well. This conjecture is numerically investigated in the next section. (Such an equation possibly remains valid for other boundary conditions as well—but this has not been investigated in the present work).

2.4 Numerical Results. The eigenvalue problem (20) is solved by employing the iterative algorithm proposed in [10]. The

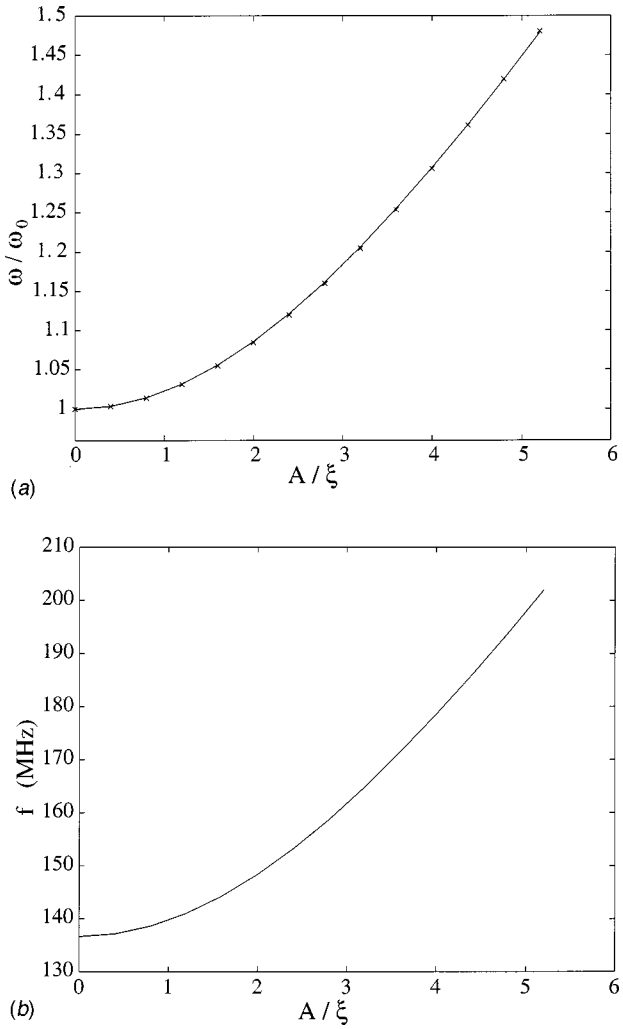


Fig. 2 Vibration of a clamped-clamped beam—fundamental mode. (*a*) Master curve—FEM 50 elements; xx Eq. (5) with $\alpha = 0.04406$; (*b*) SWNT: Eq. (5) with $\alpha = 0.04406$, $E = 705$ GPa, $\rho = 1330$ Kg/m^3, $L = 600$ nm, $d_o = 2$ nm, $h = 0.335$ nm.

FEM calculations are first verified against the analytical solution (5) for a simply-supported beam with circular cross section (here $\xi = d/4$, where d is the diameter of the beam cross section). The match between the analytical solution and the FEM solution with 50 elements is excellent. It is next verified that the plot of ω/ω_0 as a function of A/ξ, obtained from the FEM, is invariant with respect to the system parameters E, ρ, L, ξ.

Results for a clamped beam appear in Fig. 2(*a*). The FEM results, obtained here for a beam of solid circular cross section, match very well with Eq. (5) with $\xi = d/4$ and $\alpha = 0.04406$ (as obtained from a least-squares fit). Table 1 shows the data for simply-supported and clamped beams. The root mean square (RMS) error is defined here as

Table 1 The parameter α and the RMS error in the fit of Eq. (5) for vibrations of a beam under different boundary conditions

Boundary Condition	α	ε
Simply-supported	0.1876	6.379×10^{-4}
Clamped	0.04406	1.3257×10^{-3}

$$\varepsilon = \sqrt{\frac{\sum_{k=1}^{n} (\varepsilon_k)^2}{n}} \tag{21}$$

where ε_k is the difference between the FEM value and the value from Eq. (5), at a point on Fig. 2(*a*), and n is the number of points. With $\omega/\omega_0 \sim \mathcal{O}(1)$, the RMS errors in Table 1 are of the order of 0.1%. (Of course, the simply-supported case basically serves to verify the computer code).

Results for a practical example, vibration of a clamped single-walled carbon nanotube (SWNT), from Eq. (5), with $\alpha = 0.04406$, are presented in Fig. 2(*b*). The SWNT is a hollow tube. Its parameters (shown in the caption of Fig. 2) have been obtained from [14]. For this tube, the radius of gyration $\xi = 0.6$ nm and $\omega_0 = 22.37\sqrt{E/\rho}\,\xi/L^2$, [12], $f_0 = \omega_0/2\pi = 136.6$ MHz. Note that the bending stiffness $EI \propto ES\xi^2$ while $\omega_0 \propto \sqrt{E/\rho}\,\xi/L^2$, so that even though the bending stiffness of the nanotube is small, the fundamental natural frequency ω_0, obtained here from the standard linear beam theory, is a large number.

3 Plates and Membranes

Nonlinear vibration of plates, without initial in-plane forces, are discussed in this section. The plates are square (side$=L$), linearly elastic, are of uniform rectangular cross section (thickness h), and have immovable edges. The three boundary conditions considered here are all edges clamped (C), all edges simply-supported (S) and two opposite edges simply-supported and the other two clamped (SC). Also, the edges are immovable, i.e., $u = v = 0$ on all edges of the plate.

3.1 The Model. Following [6], the kinematic equations adopted here, for a plate with immovable edges, are, [10],

$$\begin{bmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \gamma_{xy} \end{bmatrix} = \begin{bmatrix} u_{,x} + \frac{1}{2}(w_{,x})^2 \\ v_{,y} + \frac{1}{2}(w_{,y})^2 \\ u_{,y} + v_{,x} + w_{,x}w_{,y} \end{bmatrix}, \quad \begin{bmatrix} \kappa_{xx} \\ \kappa_{yy} \\ \kappa_{xy} \end{bmatrix} = \begin{bmatrix} -w_{,xx} \\ -w_{,yy} \\ -2w_{,xy} \end{bmatrix} \tag{22}$$

where $u(x,y,t), v(x,y,t)$ are the in-plane and $w(x,y,t)$ out-of-plane displacements, respectively, $[\epsilon] = [\epsilon_{xx}, \epsilon_{yy}, \gamma_{xy}]^T$ is the in-plane strain (measured at the mid-plane), and $[\kappa] = [\kappa_{xx}, \kappa_{yy}, \kappa_{xy}]^T$ are the curvatures.

The constitutive equations are

$$[N] = H^{(I)}[C][\epsilon], \quad [M] = H^{(O)}[C][\kappa] \tag{23}$$

where $[N] = [N_{xx}, N_{yy}, N_{xy}]^T = h[\sigma_{xx}, \sigma_{yy}, \sigma_{xy}]^T$ are the in-plane forces per unit length, σ_{ij} are the components of stress and $[M] = [M_{xx}, M_{yy}, M_{xy}]^T$ are the bending and twisting moments. Also

$$C = \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{1}{2}(1-\nu) \end{bmatrix}, \quad H^{(I)} = \frac{Eh}{1-\nu^2}, \quad H^{(O)} = \frac{Eh^3}{12(1-\nu^2)} \tag{24}$$

with ν the Poisson's ratio of the plate material.

The membrane strain energy $\mathcal{E}^{(I)}$, the bending strain energy $\mathcal{E}^{(O)}$, and the kinetic energy $\mathcal{E}^{(k)}$ are, [15],

$$\mathcal{E}^{(I)} = \frac{1}{2}\int_{\Delta} [N_x\epsilon_{xx} + N_y\epsilon_{yy} + N_{xy}\gamma_{xy}]\,dx\,dy \tag{25}$$

$$\mathcal{E}^{(O)} = \frac{1}{2}\int_{\Delta} [M_x\kappa_{xx} + M_y\kappa_{yy} + M_{xy}\kappa_{xy}]\,dx\,dy \tag{26}$$

$$\mathcal{E}^{(k)}=\frac{\rho h}{2}\int_{\Delta}[\dot{u}^2+\dot{v}^2+\dot{w}^2]dxdy \tag{27}$$

where Δ is the area of the plate surface.

Using Eqs. (22)–(24), the energy expressions (25)–(26) can be written in terms of the plate parameters E, ν, h and the displacement derivatives. These expressions are available in [15] on pages 313 and 95, respectively.

3.2 FEM Model for Plates With Immovable Edges. The procedure followed in this section is quite analogous to that for beams described before in Section 2.3.

FEM Discretization. Each plate element has four corner nodes with six degrees-of-freedom at each node. These are u, v, w, $w_{,x}$, $w_{,y}$, $w_{,xy}$. For each element, one has

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix}=\begin{bmatrix} N^{(I)} & 0 \\ 0 & N^{(O)} \end{bmatrix}\begin{bmatrix} q^{(I)} \\ q^{(O)} \end{bmatrix} \tag{28}$$

with

$$[N^{(I)}(x,y)]=\begin{bmatrix} N_1 & 0 & N_2 & 0 & N_3 & 0 & N_4 & 0 \\ 0 & N_1 & 0 & N_2 & 0 & N_3 & 0 & N_4 \end{bmatrix},$$

$$[N^{(O)}(x,y)]=[P_1,P_2,\ldots,P_{16}] \tag{29}$$

$$[q^{(I)}(t)]=[u_1,v_1,\ldots,u_4,v_4]^T,$$

$$[q^{(O)}(t)]=[w_1,(w_{,x})_1,(w_{,y})_1,(w_{,xy})_1,\ldots,w_4,(w_{,x})_4,$$

$$(w_{,y})_4,(w_{,xy})_4]^T. \tag{30}$$

Here N_k and P_k are bilinear interpolation functions, [16], and $[q^{(I)}]$ and $[q^{(O)}]$ contain the appropriate nodal degrees-of-freedom.

Define

$$[D]=\begin{bmatrix} w_{,x} & 0 \\ 0 & w_{,y} \\ w_{,y} & w_{,x} \end{bmatrix},\quad [G]=\begin{bmatrix} N^{(O)}_{,x} \\ N^{(O)}_{,y} \end{bmatrix} \tag{31}$$

$$[B^{(I)}]=\begin{bmatrix} N_{1,x} & 0 & N_{2,x} & 0 & N_{3,x} & 0 & N_{4,x} & 0 \\ 0 & N_{1,y} & 0 & N_{2,y} & 0 & N_{3,y} & 0 & N_{4,y} \\ N_{1,y} & N_{1,x} & N_{2,y} & N_{2,x} & N_{3,y} & N_{3,x} & N_{4,y} & N_{4,x} \end{bmatrix},$$

$$[B^{(O)}]=-\begin{bmatrix} N^{(O)}_{,xx} \\ N^{(O)}_{,yy} \\ 2N^{(O)}_{,xy} \end{bmatrix}. \tag{32}$$

Substituting the interpolations (28) into the energy expressions (25)–(27), and applying Hamilton's principle, results in the element level equations:

$$\begin{bmatrix} M^{(I)} & 0 \\ 0 & M^{(O)} \end{bmatrix}\begin{bmatrix} \ddot{q}^{(I)} \\ \ddot{q}^{(O)} \end{bmatrix}+\begin{bmatrix} K^{(I)} & 0 \\ 0 & K^{(O)} \end{bmatrix}\begin{bmatrix} q^{(I)} \\ q^{(O)} \end{bmatrix}+\begin{bmatrix} 0 & K^{(IO)} \\ 2K^{(IO)T} & K^{(NI)} \end{bmatrix}$$

$$\times\begin{bmatrix} q^{(I)} \\ q^{(O)} \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{33}$$

The various submatrices in (33) are

$$[M^{(I)}]=\frac{\rho h}{2}\int_{\Delta^{(e)}}[N^{(I)}]^T[N^{(I)}]dxdy,$$

$$[M^{(O)}]=\frac{\rho h}{2}\int_{\Delta^{(e)}}[N^{(O)}]^T[N^{(O)}]dxdy \tag{34}$$

$$[K^{(I)}]=H^{(I)}\int_{\Delta^{(e)}}[B^{(I)}]^T[C][B^{(I)}]dxdy,$$

$$[K^{(O)}]=H^{(O)}\int_{\Delta^{(e)}}[B^{(O)}]^T[C][B^{(O)}]dxdy \tag{35}$$

$$[K^{(IO)}]=\frac{H^{(I)}}{2}\int_{\Delta^{(e)}}[B^{(I)}]^T[C][D][G]dxdy,$$

$$[K^{(NI)}]=\frac{H^{(I)}}{2}\int_{\Delta^{(e)}}([D][G])^T[C][D][G]dxdy \tag{36}$$

where $\Delta^{(e)}$ is the area of a finite element. For the purpose of the discussion to follow, only an uniform mesh of identical square elements, of side ℓ, is considered here. Of course, the FEM code allows elements of very general shape.

Note that the in-plane and out-of-plane (bending) matrices $[K^{(I)}]$ and $[K^{(O)}]$ are $\propto h$ and h^3, respectively, the matrix $[K^{(IO)}]\propto Ah$ represents coupling between the in-plane and out-of-plane displacements, and the matrix $[K^{(NI)}]\propto A^2h$ arises purely from the nonlinear axial strains.

It is well known that for the linear theory $[K^{(O)}]\ll[K^{(I)}]$ as $h\to 0$. It is very interesting, however, to note that if A/h remains $\mathcal{O}(1)$, the bending matrix $[K^{(O)}]$, which arises from the linear theory, and the matrix $[K^{(NI)}]$ from the nonlinear theory, remain of the same order as $h\to 0$. This fact has important consequences for the modeling of very thin plates, as can be seen in the following section.

Eigenvalue Problem. Following the procedure detailed in [10] (this approach neglects in-plane inertia, see the Appendix), the following eigenvalue equations are obtained:

$$([K^{(O)}]+[K^{(NL)}]-\omega^2[M^{(O)}])[q^{(O)}]=[0] \tag{37}$$

where

$$[K^{(NL)}]=(3/4)([K^{(NI)}]-2[K^{(IO)}]^T[K^{(I)}]^{-1}[K^{(IO)}]). \tag{38}$$

The next task is to investigate the dependence of the fundamental eigenvalue of (37) on system parameters. The procedure followed now is analogous to that for beams outlined in Section 2.3.

The vector $[q^{(O)}]$ is first written in dimensionally homogeneous form as

$$[w_1,L(w_{,x})_1,L(w_{,y})_1,L^2(w_{,xy})_1,\ldots,w_4,L(w_{,x})_4,$$

$$L(w_{,y})_4,L^2(w_{,xy})_4]^T \tag{39}$$

and the appropriate columns of $[K^{(O)}]$, $[K^{(NL)}]$, $[M^{(O)}]$ are suitably divided by L or L^2. Also, as for beams, certain rows of the entire system (37) are divided by suitable powers of L.

Each element of these matrices are now further scaled (multiplied, appropriately, by suitable powers of L/L). As in the case of beams, it is assumed that ℓ/L (where ℓ is the length of the side of a plate element) is kept fixed, even if L is changed (isotropic expansion or contraction). Next, the global matrices are formed and appropriate immovable boundary conditions (S, C, or SC) are imposed. One now has

$$[\hat{K}^{(O)}]^{(G)}\propto\frac{Eh^3}{L^2},\quad [\hat{K}^{(NL)}]^{(G)}\propto\frac{EA^2h}{L^2},\quad [\hat{M}^{(O)}]^{(G)}\propto\rho L^2h. \tag{40}$$

Note that the first two matrices above also depend on the Poisson's ratio ν but that this dependency is complicated in view of the form of the constitutive matrix $[C]$ in Eq. (24*a*). It is, of course, well known that ω_0 is obtained by solving the linearized problem ((37) with $[K^{(NL)}]=[0]$). It is interesting to point out that for a square S plate, one has the exact solution for the fundamental mode, [17]:

$$\omega_0=\frac{\sqrt{3}\pi^2h}{3L^2}\frac{1}{\sqrt{(1-\nu^2)}}\sqrt{\frac{E}{\rho}} \tag{41}$$

for which the dependence on ν is the same as that of $\sqrt{H^{(O)}}$ (see (24*c*)). However, ω_0 can, for certain boundary conditions, have a more complicated dependence on the Poisson's ratio ([17], p 250).

Multiply the new global eigenvalue equation by $L^2/(Eh^3)$, note the forms of the matrices $[\hat{K}^{(O)}]^{(G)}$, $[\hat{K}^{(NL)}]^{(G)}$, $[\hat{M}^{(O)}]^{(G)}$, and that $1/\omega_0^2 \propto (\rho L^4)/(Eh^2)$ for many different boundary conditions that include S and SC but not proved for C plates ([17], pp 249–250)—this fact can be deduced from the discussion on plane waves in ([17], p. 245).

One now gets the equation

$$([M_1]-(\omega^2/\omega_0^2)[M_2])[v]=[0] \tag{42}$$

in which, for a fixed value of A/h, the matrices $[M_1]$, $[M_2]$ are *independent* of the system (plate) parameters E, ρ, L, h; the dependence on these parameters being wholly captured in ω_0^2 (provided, of course, that A/h remains fixed.) Therefore, for any values of the system parameters (within the limits of validity of the plate model employed in this work), and with a fixed value of A/h, the eigenvalue problem (42) yields a *single unique fundamental eigenvalue*, here called $\hat{\omega}^2/\omega_0^2$. In other words, a point on a plot of ω/ω_0 versus A/h, for a fixed value of A/h, is *invariant* with respect to variations of these system parameters. Equivalently, a plot of ω/ω_0, as a function of A/h, for a square plate with many different boundary conditions, including S and SC, is *invariant* with respect to the system parameters E, ρ, L, h. Note that the Poisson's ration ν is not included in this list of system parameters, i.e. that this ''master curve'' is not, in general, independent of the Poisson's ratio.

Please note that, as for beams, $[M_1]$ is a function of $[v]$ with $\max_k |v_k|=1$, so that iterations are needed to solve (42).

3.3 Numerical Results. The eigenvalue problem (42) is solved by employing the iterative algorithm proposed in [10]. The FEM calculations are verified against the FEM solution for ω/ω_0 as a function of A/h, for a fully clamped (C) square plate, as presented in Fig. 3 in [10]. The agreement between the two numerical solutions, with an 8×8 array of finite elements, is excellent. It is next observed numerically that the plot of ω/ω_0 as a function of A/h, for a square C plate, obtained from the FEM, is invariant with respect to the system parameters E, ρ, L, h. As stated before, proof of existence of this master curve is only possible for square S and SC plates, but this fact appears to be true for C plates as well.

Based on the existence of this master curve (of the form $\omega/\omega_0=g(A/h,\nu)$ for a specified boundary condition—S or SC), and inspired by the form of Eq. (5) for beams, it is conjectured that the variation of ω/ω_0, as a function of A/h, for nonlinear vibrations of square S, SC, as well as C plates, has the form

$$\frac{\omega}{\omega_0}=\sqrt{1+f(\alpha_0,\nu)\left(\frac{A}{h}\right)^2} \tag{43}$$

where the parameter α_0 depends only on the imposed boundary conditions (and not on any of the system parameters E, ν, ρ, h, or L). Next, the choice $f(\alpha_0,\nu)=\alpha_0/\sqrt{1-\nu^2}$ leads to

$$\frac{\omega}{\omega_0}=\sqrt{1+\frac{\alpha_0}{\sqrt{(1-\nu^2)}}\left(\frac{A}{h}\right)^2}. \tag{44}$$

Equation (44) is an intuitive conjecture, inspired by the dependence of the flexural rigidity D on the Poisson's ratio ν, that is tested a posteriori. First, as seen in Fig. 3(*a*), the FEM results for a clamped square plate compare very well with Eq. (44) for $\nu=0.22$ with $\alpha_0=0.3670$. (The value of α_0 is obtained from a least-squares fit). Next, FEM calculations are carried out for various boundary conditions and for various values of Poisson's ratio, and the results are compared to those from Eq. (44). Table 2 shows the data for S,C and SC square plates. Eq. (44) is seen to describe the FEM results in excellent fashion-the maximum RMS error (see Eq. (21)) being less than 1% in all cases.

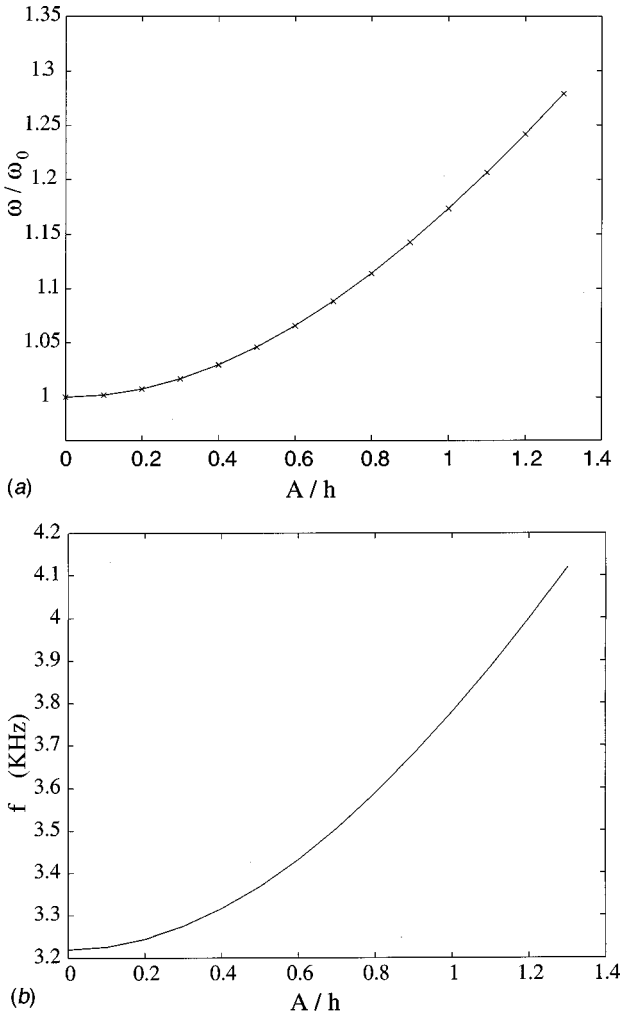


Fig. 3 Vibration of a square plate clamped on all sides—fundamental mode. (*a*) Master curve—FEM 8×8 elements, $\nu=0.22$; xx Eq. (44) with $\alpha_0=0.3670$; (*b*) MEMS plate: Eq. (44) with $\alpha_0=0.3670$, $E=169$ GPa, $\nu=0.22$, $\rho=2300$ Kg/m^3, $L=3$ mm, $h=2$ μm.

Finally, a practical situation, that of a vibrating thin MEMS plate (really a membrane), made of polysilicon, appears in Fig. 3(*b*). Here, $f_0=3.2194$ KHz, is obtained from the FEM, and then the full curve is obtained from (44). (A FEM calculation is verified to yield the same results within plotting accuracy). The material properties of polysilicon are obtained from [18,19] and the membrane dimensions from [2].

4 Conclusions

The primary contributions of the present paper are as follows.

Inclusion or Exclusion of Bending Stiffness in the Vibration of Strings and Membranes. Start with a beam with immovable supports and let T_0 (Newtons) be the initial tension in the beam. The contribution of T_0 to $\mathcal{E}^{(s)}$ in Eq. (3) is $(T_0/2)\int_0^L (w_{,x})^2 dx$, so that, in this case the eigenvalue equation (15) is modified to read

$$([K^{(O)}]+[K^{(F)}]+[K^{(NL)}]-\omega^2[M^{(O)}])[q^{(O)}]=[0] \tag{45}$$

where the new element matrix due to the initial tension is

$$[K^{(F)}]=T_0\int_0^\ell [G]^T[G]dx \tag{46}$$

Table 2 The parameter α_0 and the RMS error in the fit of Eq. (44) for vibrations of a square plate under different boundary conditions

Boundary Condition	α_0	ν	ε
All edges simply-supported	0.9621	0.12	7.766×10^{-3}
		0.22	1.104×10^{-3}
		0.32	2.690×10^{-3}
All edges clamped	0.3670	0.12	9.707×10^{-4}
		0.22	2.302×10^{-4}
		0.32	9.707×10^{-4}
Two opposite edges clamped and the other two simply-supported	0.6369	0.12	4.410×10^{-3}
		0.22	2.524×10^{-3}
		0.32	2.524×10^{-3}

with $[G]$ defined in Eq. (10*b*).

It can be shown that $[\hat{K}^{(F)}]^{(G)} \propto T_0/L$. Therefore, with a sufficiently large value of T_0 ($T_0 \gg T^{(c)} \equiv EI/L^2$), $[\hat{K}^{(O)}]^{(G)}$ and $[\hat{K}^{(NL)}]^{(G)}$ become very small (see (19)) relative to $[\hat{K}^{(F)}]^{(G)}$, as $\xi \to 0$ with $A/\xi \sim \mathcal{O}(1)$; and can be neglected, as is commonly done for strings. The remaining linear equation yields the usual result, [12]:

$$f_0 = \frac{1}{2L}\sqrt{\frac{T_0}{\hat{\rho}}} \tag{47}$$

where $\hat{\rho}$ is the mass per unit length of the string. It is clear, however, that with $T_0 \to 0$, and with $A/\xi \sim \mathcal{O}(1)$, $[\hat{K}^{(O)}]^{(G)}$ and $[\hat{K}^{(NL)}]^{(G)}$ remain of the same order as $\xi \to 0$ (see (19)), and the bending stiffness matrix $[\hat{K}^{(O)}]^{(G)}$ cannot be neglected in this case.

The above discussion applies to vibrations of strings. A very similar argument can be made for membranes as well.

It is interesting to compare the above discussion with that in a recently published nice paper by Sapmaz et al. [20]. These authors *correctly* include both the bending and nonlinear effects (but not the axial displacement) in their analysis. They compare the *induced tension*, (called T in their work and given by their Eq. (2) which is $T = (ES/(2L))\int_0^L (w_{,x})^2 dx$) with EI/L^2. (Note that $T^{(c)}$ is defined as EI/L^2 in the present work). It should also be observed that their induced tension is $T \propto EA^2S/L^2 = EIA^2/(\xi^2 L^2)$. Therefore, (case one) $T \ll T^{(c)}$ for $A \ll \xi$ (i.e., in the linear vibration limit) and of the same order (case two) for $A/\xi \sim \mathcal{O}(1)$. Also, (case three) $T \gg T^{(c)}$ for $A \gg \xi$. The present authors feel that the last regime is inappropriate in [20] since their model (i.e., their Eq. (1)) is only good for small or moderately large vibration amplitude, i.e., for cases one and two only.

Sapmaz et al. [20] also conclude (see their Eq. (19)) that a carbon nanotube vibrates like a ''loose string'' for *induced tension* $T \ll T^{(c)}$. This statement, however, is somewhat misleading, in that a ''loose string'' conjures up images of a low natural frequency of vibration. In fact, Eq. (19) of [20] without the correction term (this equation, as expected, being the same as the standard formula for ω_0 for a clamped beam, [12], p. 223), yields a value of $f_0 = 136.6$ MHz for the carbon nanotube example (Fig. 2(*b*)) in the present paper. This is, indeed, a large number, and calling this object a ''wire'' rather than a ''loose string'' is more appropriate.

It should be mentioned again that the present authors feel that [20] presents a nice analytical study of this problem. The above two paragraphs are meant to clarify some issues in Sapmaz et al. [20], not to criticize their work.

Summary of Various Cases of Vibrations of Beams, Strings, Plates, and Membranes. Table 3 summarizes various cases which are defined as:

1. nonlinear vibrations of beams or plates with $A \approx \xi$ or $A \approx h$,
2. linear vibrations of beams or plates with $A \ll \xi$ or $A \ll h$,
3. nonlinear vibrations of strings or membranes with $A \approx \xi$ or $A \approx h$, and
4. linear vibrations of strings or membranes with $A \ll \xi$ or $A \ll h$.

Please refer to Eqs. (19), (40), and (45) for the following discussion and consider the interesting cases 3 and 4. For case 3 with sufficiently high initial tension, one can perhaps argue that $K^{(O)}$ and $K^{(NL)}$ are both $\ll K^{(F)}$ and one can neglect $K^{(O)}$ but keep $K^{(NL)}$ without having significant impact on the numerical results. This is not allowed, however, when T_0 is small. Also, case 4 clearly cannot be modeled if $T_0 = 0$.

Proof of the Existence of Certain "Master Curves". It is proved that for beams that are clamped at both ends, the variation of ω/ω_0, as a function of A/ξ, is *invariant* with respect to the system parameters E, ρ, L, ξ. It is also proved for square S and SC plates, and conjectured for C plates, that the variation of ω/ω_0, as a function of A/h, is *invariant* with respect to the system parameters E, ρ, L, h. Examples of these ''master curves'' appear in Figs. 2(*a*) and 3(*a*).

It is important to note here that the above proof for square plates has been carried out for a uniform FEM mesh with identical square finite elements. The FEM solution for this problem, however, is expected to converge to the exact solution in the limit of mesh refinement, for any reasonable starting mesh. Therefore, the master curves are a feature of the exact solution of the problem, and the proof given in this paper is not limited by the particular choice of mesh, but should be valid in general.

Approximate Formulas for ω/ω_0 as a Function of A/ξ or A/h. Certain approximate Eqs. (5) and (44), have been proposed for the variation of ω/ω_0, as a function of A/ξ, for beams; and for ω/ω_0, as a function of A/h, for plates, respectively. These equations are shown to be very accurate for the cases considered—clamped beams (Eq. (5) is well known for simply-supported beams), and S, C, and SC plates. These equations should prove to be useful for the design and analysis of macro as well as micro and nano structures that include beams or plates.

Table 3 Summary of various cases

Case	Initial Tension	Bending	Nonlinearity	Usual Model OK?
1	high	yes	yes	yes
2	high	yes	no	yes
3	high	no	yes	?
4	high	no	no	yes
1	low	yes	yes	yes
2	low	yes	no	yes
3	low	no	yes	no
4	0	no	no	no

Acknowledgments

This research has been partially supported by Grant #EEC-0303674 of the National Science Foundation to Cornell University, and by Grant #01-2042-007-25 of the New Jersey Commission on Science and Technology to New Jersey Institute of Technology through the New Jersey Center for Micro-Flow Control. Sincere thanks are expressed to Prof. Paul McEuen for providing Fig. 1 and for various discussions related to carbon nanotubes, to Dr. Yu Xie Mukherjee for many valuable discussions related to mechanics modeling, and to one of the reviewers and Prof. Richard Rand for their ideas related to nondimensionalized governing differential equations for beams, and for plates, respectively.

Appendix

Beams. One of the reviewers of this paper has kindly pointed out an interesting alternative approach to proving the existence of master curves. This is based on appropriate nondimensionalization of the governing differential equations of the problem. For the beam problem, the analysis, *as provided by this reviewer*, proceeds as follows.

The governing equations for a vibrating beam (for the model adopted in the present paper) are:

$$\begin{aligned} &\text{axial motion:} \quad Eu_{,xx}=\rho u_{,tt} \\ &\text{transverse motion:} \quad EIw_{,xxxx}-(Nw_{,x})_{,x}+\rho Sw_{,tt}=0 \\ &\text{constitutive and kinematic laws:} \quad N=ES\epsilon_{xx}=ES[u_{,x}+(1/2)\times(w_{,x})^2] \end{aligned} \tag{A1}$$

where the only new symbol is N, the axial force.

The following nondimensionalization is used:

$$x=L\hat{x}, \quad w=A\hat{w}, \quad u=(A^2/L)\hat{u}, \quad N=ES(A^2/L^2)\hat{N}, \quad t=\hat{t}/\overline{\omega} \tag{A2}$$

with $\overline{\omega}=(\xi/L^2)\sqrt{(E/\rho)}$.

Applying Eqs. (A2) to (A1) results in

$$\begin{aligned} &\text{axial motion:} \quad \frac{\partial \hat{N}}{\partial \hat{x}}=\left(\frac{\xi}{L}\right)^2\frac{\partial^2 \hat{u}}{\partial^2 \hat{t}^2} \\ &\text{transverse motion:} \quad \frac{\partial^4 \hat{w}}{\partial \hat{x}^4}-\left(\frac{A}{\xi}\right)^2\frac{\partial}{\partial \hat{x}}\left(\hat{N}\frac{\partial \hat{w}}{\partial \hat{x}}\right)+\frac{\partial^2 \hat{w}}{\partial \hat{t}^2}=0 \\ &\text{constitutive and kinematic laws:} \quad \hat{N}=\frac{\partial \hat{u}}{\partial \hat{x}}+\frac{1}{2}\left(\frac{\partial \hat{w}}{\partial \hat{x}}\right)^2 \end{aligned} \tag{A3}$$

It can be seen from $(A3)_1$ that for $\xi/L \ll 1$ (in this case $u \approx \mathcal{O}(w^2)$), axial inertia, i.e., $u_{,tt}$, can be neglected. Therefore, $\hat{N}=\hat{N}(t)$. Now integrating $(A3)_3$ from $\hat{x}=0$ to $\hat{x}=1$ and using the fact that the ends of the beam are immovable (i.e., $\hat{u}(0,\hat{t})=\hat{u}(1,\hat{t})=0$), yields

$$\hat{N}=\frac{1}{2}\int_0^1\left(\frac{\partial \hat{w}}{\partial \hat{x}}\right)^2 d\hat{x}. \tag{A4}$$

Finally, use of (A4) in $(A3)_2$ leads to a single partial differential equation governing $\hat{w}$:

$$\frac{\partial^4 \hat{w}}{\partial \hat{x}^4}-\frac{1}{2}\left(\frac{A}{\xi}\right)^2\frac{\partial^2 \hat{w}}{\partial \hat{x}^2}\int_0^1\left(\frac{\partial \hat{w}}{\partial \hat{x}}\right)^2 d\hat{x}+\frac{\partial^2 \hat{w}}{\partial \hat{t}^2}=0. \tag{A5}$$

It can be seen that A/ξ is the only relevant parameter in Eq. (A5). It is clear at this stage that master curves exist for ω/ω_0 as functions of A/ξ for clamped and simply supported beams.

For completeness, the dimensional form of (A5) is

$$EIw_{,xxxx}-\frac{SE}{2L}w_{,xx}\int_0^L(w_{,x})^2dx+\rho Sw_{,tt}=0. \tag{A6}$$

Equation (A6) is available in many references (see, for example, Nayfeh and Mook [5], p. 453, where it appears in a dimensionless form different from (A5)).

Plates. A nice account of nonlinear vibrations of plates is available in [5], pp. 500–507. The model employed in the present work is the dynamic von Karman model ([5], pp. 502–503) without in-plane inertia, i.e., with $u_{,tt}=v_{,tt}=0$. A further simplification can be obtained with the Berger assumption $\epsilon_{xx}\epsilon_{yy}-(1/4)\gamma_{xy}^2=0$ (with the strain components defined in terms of displacements in Eq. (22)). The Berger simplification, together with neglecting in-plane inertia, leads to a single governing partial differential equation for the transverse displacement $w(x,y,t)$ ([5], p. 507):

$$D\nabla^4 w-\frac{hEe}{1-\nu^2}\nabla^2 w+\rho h w_{,tt}=0 \tag{A7}$$

with

$$e(t)=\frac{1}{2\Delta}\int_\Delta \nabla^2 w dA. \tag{A8}$$

Rand [21], using the nondimensionalization (A2) in (A7), has shown that the Berger model, without in-plane inertia, leads to existence of master curves for ω/ω_0 as functions of A/h. Furthermore, these curves are *independent* of Poisson's ratio for plates with boundary conditions that do not involve ν. As pointed out in [5], however, the Berger assumption sometimes produces erroneous results and the von Karman model is more reliable. *Interestingly, the von Karman model without in-plane inertia, employed in the present work, leads to dependence of ω/ω_0 on ν, as seen in Eq. (44)!*

References

[1] Qian, D., Wagner, G. J., Liu, W. K., Yu, M.-F., and Ruoff, R. S., 2002, "Mechanics of Carbon Nanotubes," Appl. Mech. Rev., **55**, pp. 495–533.
[2] Roman, M., and Aubry, N., 2003, "Design and Fabrication of Electrostatically Actuated Synthetic Microjets," ASME Paper No. IMECE2003-41579.
[3] Ko, S. C., Kim, Y. C., Lee, S. S., Choi, S. H., and Kim, S. R., 2003, "Micromachined Piezoelectric Membrane Acoustic Device," Sens. Actuators, **103**, pp. 130–134.
[4] McEuen, P. L., 2003, private communication, Department of Physics, Cornell University, Ithaca, NY.
[5] Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillations*, John Wiley and Sons, New York.
[6] Singh, G., Sharma, A. K., and Rao, G. V., 1990, "Large-Amplitude Free Vibrations of Beams—A Discussion of Various Formulations and Assumptions," J. Sound Vib., **142**, pp. 77–85.
[7] Woinowsky-Krieger, S., 1950, "The Effect of an Axial Force on the Vibration of Hinged Bars," Trans. ASME, **72**, pp. 35–36.
[8] Singh, G., Rao, G. V., and Iyengar, N. G. R., 1990, "Re-investigation of Large-Amplitude Free Vibrations of Beams Using Finite Elements," J. Sound Vib., **143**, pp. 351–355.
[9] Chia, C. Y., 1980, *Nonlinear Analysis of Plates*, McGraw-Hill, New York.
[10] Han, W., and Petyt, M., 1997, "Geometrically Nonlinear Vibration Analysis of Thin Rectangular Plates Using the Hierarchical Finite Element Method—I: The Fundamental Mode of Isotropic Plates," Comput. Struct., **63**, pp. 295–308.
[11] Dumir, P. C., and Bhaskar, A., 1988, "Some Erroneous Finite Element Formulations of Nonlinear Vibrations of Beams and Plates," J. Sound Vib., **123**, pp. 517–527.
[12] Thomson, W. T., 1988, *Theory of Vibration With Application*, 3rd. Ed., Prentice-Hall, Englewood Cliffs, NJ.
[13] Zienkiewicz, O. C., and Taylor, R. L., 1994, *The Finite Element Method*, Vol. 2, 4th, Ed., McGraw-Hill, Maidenhead, Berkshire, UK.
[14] Zhang, P., Huang, Y., Geubelle, P. H., Klein, P. A., and Hwang, K. C., 2002, "The Elastic Modulus of Single-Wall Carbon Nanotubes: A Continuum Analysis Incorporating Interatomic Potentials," Int. J. Solids Struct., **39**, pp. 3893–3906.
[15] Ugural, A. C., 1999, *Stresses in Plates and Shells*, 2nd. Ed., McGraw-Hill, New York.
[16] Yang, T. Y., 1986, *Finite Element Structural Analysis*, Prentice-Hall, Englewood Cliffs, NJ.

[17] Graff, K. F., 1991, *Wave Motion in Elastic Solids*, Dover, New York.
[18] Kurt Petersen, K. E., 1982, ''Silicon as a Mechanical Material,'' Proc. IEEE, **70**, pp. 420–455.
[19] Sharpe, W. N., Jr., 2001, ''Mechanical Properties of MEMS Materials,'' *The MEMS Handbook*, CRC Press, Boca Raton, FL, pp. 3–33.
[20] Sapmaz, S., Blanter, Y. M., Gurevich, L., and van der Zant, H. S. J., 2003, ''Carbon Nanotubes as Nanoelectromechanical Systems,'' Phys. Rev. B, **67**, pp. 235414:1–7.
[21] Rand, R. H. 2004, private communication, Department of Theoretical and Applied Mechanics, Cornell University, Ithaca, NY.

Effect of Residual Stress on Cavitation Instabilities in Constrained Metal Wires

Viggo Tvergaard
Department of Mechanical Engineering,
Solid Mechanics,
Technical University of Denmark,
DK-2800 Kgs. Lyngby, Denmark

Numerical calculations are carried out for a test specimen geometry used by Ashby et al. to study effects of ductile reinforcements in brittle materials. A lead wire well bonded to an outer glass cylinder was used to investigate crack bridging by ductile particles. The main focus here is on a single void growing very large in the metal wire, in the crack plane, perhaps involving a cavitation instability. Therefore, full finite strain elastic-plastic theory is used for the analyses, and remeshing procedures are applied to avoid unacceptable mesh distortion. Residual stresses induced by thermal contraction mismatch during cooling from the processing temperature can have a noticeable influence on the results, and this is quantified by the analyses. [DOI: 10.1115/1.1767845]

1 Introduction

For a single void in an infinite elastic-plastic solid cavitation instabilities have been predicted in cases where the stress levels are sufficiently high such that the work released in the field surrounding the expanding void is enough to drive continued expansion. At this critical stress level the void grows without bound, while the remote strain remains stationary. The existence of such instabilities has been recognized by Bishop et al. [1] for spherically symmetric conditions, and analyses by Huang et al. [2] and Tvergaard et al. [3] have shown similar cavitation instabilities for spherical voids subject to axisymmetric stress conditions, as long as the ratio of the transverse stress and the axial tensile stress is near unity. There has also been a number of related spherically symmetric studies in the context of nonlinear elasticity (e.g., Ball [4], Horgan and Abeyaratne [5], and Chou-Wang and Horgan [6]), where the occurrence of a cavitation instability has been interpreted either as a bifurcation from a homogeneously stressed solid to a solid containing a void, or as the growth of a preexisting void.

In experiments by Flinn et al. [7] for Al_2O_3 reinforced by Al particles scanning electron microscopy has been used to obtain quantitative information about the plastic stretch of particles bridging a crack in the brittle matrix, and the extent of debonding from the matrix during bridging. It was observed that the failure of axisymmetric reinforcement zones often involves the nucleation of a single hole at the center of the neck, which rapidly expands to final failure. This type of behavior has been analyzed in some detail by Tvergaard [8]. A set of experiments by Ashby et al. [9] were designed to obtain insight in the toughening of brittle matrix materials obtained by ductile reinforcements bridging a crack. In these experiments lead wires were well bonded to a thick outer glass cylinder and the glass was cracked circumferentially, so that during a tensile test all load in the crack plane was carried in the highly constrained metal wire. Several of these specimens failed by the enlargement of a single void in the wire near the crack plane, and this failure mechanism will be analyzed in the present paper. Experiments by Dalgleish et al. [10] for a thin ductile metal layer used to bond two ceramic blocks together have indicated that also here the constraint on plastic flow during tension normal to the layer is high enough to make cavitation a likely failure mechanism, and analyses by Tvergaard [11–13] have confirmed this mechanism in thin ductile bond layers.

In photographs, [9] of the final fracture surface for well bonded metal wires show that the diameter of the single void in the wire has reached values up to 0.53 times the diameter of the wire. To better understand these results a full numerical analysis of the tensile tests is carried out here, accounting for the interaction of the elastic-plastic metal wire with the surrounding elastic glass. In producing the test specimens slow cooling from the melting temperature of lead was applied to minimize thermal stresses, resulting from the contraction mismatch. However, as the observed final void size seems large, the effect of residual stresses in the test specimen is investigated. Furthermore, to be able to rather accurately evaluate the final void size predicted by the computations, remeshing procedures are used (Pedersen [14] and Tvergaard [12]) to be able to resolve the huge strains developing around the void without unacceptable mesh distortion.

2 Problem Formulation

Some of the earlier analyses of unstable cavity growth have focussed on a single void in an infinite elastic-plastic solid, but the present analyses consider directly the specimen geometry used by Ashby et al. [9], as shown in Fig. 1. The purpose of these experiments was to get some understanding of the toughening effect of ductile reinforcements in brittle materials such as ceramics, by a crack bridging mechanism. Lead wires were well bonded to a thick outer glass cylinder, then the glass was cracked circumferentially so that all load at the crack plane would be carried by the bridging metal wire, and the specimen was pulled in tension to failure. In many of the tests the bond strength was high enough so that no debonding occurred, leading to highly constrained plastic flow in the wire near the crack plane, and for several specimens the fracture surface showed large growth of a single void. In the present analyses the glass is modeled as linearly elastic, while the metal wire is described as elastic-plastic.

A Lagrangian convected coordinate formulation of the field equations is used for the finite strain analyses. A material point is identified by cylindrical coordinates x^i in the reference frame, and the corresponding displacement components are u^i. The Lagrangian strain tensor is given by

$$\eta_{ij}=\frac{1}{2}\left(u_{i,j}+u_{j,i}+u^k_{,i}u_{k,j}\right) \tag{2.1}$$

where $(\,)_{,i}$ denotes covariant differentiation in the reference frame, Latin indices range from 1 to 3, and the summation con-

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the Applied Mechanics Division, November 24, 2003; final revision, March 8, 2004. Editor: R. M. McMeeking. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

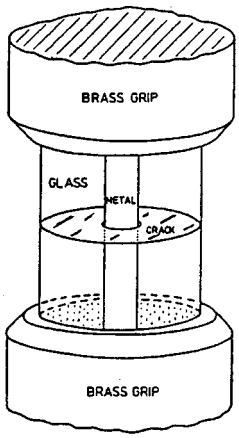


Fig. 1 Sketch of the tensile test specimen used by Ashby et al. [9]

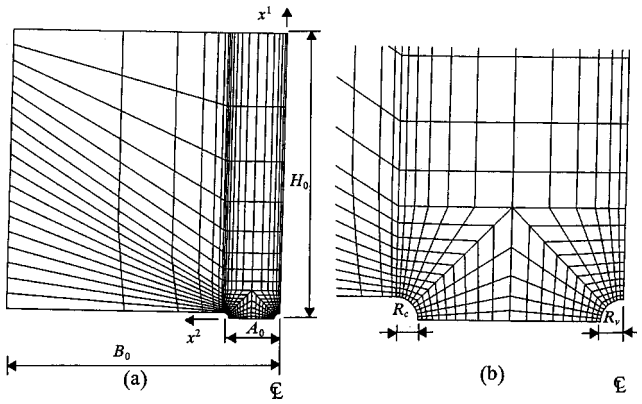


Fig. 2 Axisymmetric model of the half test specimen analyzed, showing the initial dimensions, the coordinate system and a finite element mesh. The ductile wire occupies the region $0 \leq x^2 \leq A_0$.

vention is used for repeated indices. The metric tensors in the current configuration and the reference configuration are denoted by G_{ij} and g_{ij}, respectively, with determinants G and g, and the contravariant components of the Kirchhoff stress tensor τ^{ij} and the Cauchy stress tensor σ^{ij} are related by $\tau^{ij}=\sqrt{G/g}\sigma^{ij}$. Equilibrium is expressed in terms of the principle of virtual work

$$\int_V \tau^{ij}\delta\eta_{ij}dV=\int_S T^i\delta u_i dS \tag{2.2}$$

where V and S are the volume and surface, respectively, in the reference configuration, and T^i are the nominal traction components on the surface.

The deformations of the metal wire are described in terms of a finite strain generalization of the classical J_2 flow theory of the form $\dot{\tau}^{ij}=L^{ijkl}\dot{\eta}_{kl}$ (see Hutchinson [15]), making use of Young's modulus E, Poisson's ratio ν, the initial yield stress σ_y, and the strain hardening exponent N. The effective Mises stress is $\sigma_e=(3s_{ij}s^{ij}/2)^{1/2}$, where $s^{ij}=\tau^{ij}-G^{ij}\tau^k_k/3$ is the stress deviator. The tangent modulus E_t is the slope of the uniaxial stress strain curve, which is here taken to be represented by a power law

$$\varepsilon=\begin{cases}\dfrac{\sigma}{E} & \text{for } \sigma\leq\sigma_y\\[2ex] \dfrac{\sigma_y}{E}\left[\dfrac{\sigma}{\sigma_y}\right]^{1/N} & \text{for } \sigma>\sigma_y\end{cases}. \tag{2.3}$$

For the numerical solution a linear incremental finite element method is used, based on an incremental version of the principle of virtual work (2.2). In each increment equilibrium correction terms are applied to prevent drifting away from the true equilibrium path (e.g., see [11]). The displacement fields are approximated in terms of axisymmetric eight-noded isoparametric elements, as in [16]. The volume integral in the incremental version of (2.2) is carried out by using 2×2 point Gauss integration (i.e., "reduced integration") within each element.

The initial geometry of the axisymmetric problem analyzed is shown in Fig. 2, with a mesh drawn in. The radius of the metal wire is A_0, the outer radius of the glass cylinder is B_0, and the half-length of the test specimen is H_0. At the center of the metal wire, in the crack plane, the presence of a void with initial radius R_v is assumed. Furthermore, at the location where the crack in the glass meets the surface of the metal wire, an initial rounding of the crack-tip with radius R_c is assumed, to be able to represent the large plastic strains that develop at this point as soon as the crack starts to open. The geometry considered in all analyses here is specified by the values $B_0/A_0=4.8$ and $H_0/A_0=5.0$. In the cylindrical coordinate system used as reference here (Fig. 2), x_1 is the axial coordinate, x_2 is the radius and x_3 is the circumferential angle.

In the metal wire, symmetry conditions are assumed at the plane ahead of the crack

$$u_1=0, \quad T_2=0, \quad \text{for } x_1=0. \tag{2.4}$$

At the top, the specimen is pulled by the grip, and the boundary conditions are taken to be

$$u_1=U, \quad T_2=0, \quad \text{for } x_1=H_0 \tag{2.5}$$

where U is the prescribed uniform end displacement. The cylindrical sides of the specimen, the crack surfaces and the void surface are load free, so that the remaining boundary conditions are

$$T_1=T_2=0, \quad \text{on all other surfaces.} \tag{2.6}$$

The average nominal traction T_a on the wire in the crack plane is obtained from the expression

$$T_a=\frac{1}{2A_0^2}\int_0^{A_0}[-T^1]_{x^1=0}2\pi x^2dx^2. \tag{2.7}$$

The variation of this nominal traction with the end displacement U will be used to illustrate the test behavior.

When residual stresses due to thermal contraction mismatch between glass and metal are included here, they are introduced in the following manner. In the metal initial principal true stresses σ^* are applied in both axial, radial and circumferential directions, while the glass is stress free. This is only possible with additional forces applied on all internal and external surfaces of the metal wire. First these additional forces are stepped down to zero over many increments, while $U=0$, to allow for corresponding plastic yielding to develop around the void and the crack-tip. Subsequently, the overall tensile loading is applied by incrementally prescribing an increasing value of U. At a later stage, the increment $\dot{U}$ may become zero or negative while the void grows rapidly. To avoid numerical instabilities in such situations, a mixed Rayleigh Ritz–finite element method is applied with the possibility of prescribing node displacements on the void surface, [17].

3 Remeshing

Previous numerical studies of spherical ductile particles bridging a crack in a brittle matrix, [8,18], have approximated the matrix as rigid, but are otherwise somewhat similar to the present investigation, and these earlier studies have been limited by strong

mesh distortions during bridging. Therefore, the present investigation makes use of a remeshing procedure to be able to follow the void growth further into the failure process.

The remeshing procedure has been incorporated into the program applied here by Pedersen [14]. In this procedure, also used by Tvergaard [12,13], the values of field quantities in the integration points of the new mesh are determined by interpolation in the old mesh. First, a bilinear surface in terms of the local element coordinates ξ and η is used to extrapolate values (such as stress components) in the old mesh from integration points to nodal points, where the region of the element is specified by $-1 \leqslant \xi \leqslant 1$ and $-1 \leqslant \eta \leqslant 1$. Then, values of the field quantities in the nodal points of the new mesh are determined by interpolation in the old mesh, using the shape functions. Finally, the values in the integration points of the new mesh are determined by interpolation from the new nodal points, using the shape functions. To do this, it is necessary to determine the location of each new nodal point in the old mesh, i.e., the element number and the appropriate values of the local coordinates ξ and η inside that element.

The old coordinates of a new nodal point are determined by the procedure [12] that first the nearest old nodal point representing an element corner is found, and then a Newton-Raphson iteration is used to determine the old values of ξ and η. This is done for each adjacent element in the old mesh, until the element is found that contains the new point. If none of these adjacent elements contain the new point, the results of the iterations are used to select another corner point for one of these elements, and the iterations are repeated for elements adjacent to that point.

As in [12] a remeshing is carried out when $\Delta\varepsilon_e \geqslant (\Delta\varepsilon_e)_{\max}$ in any integration point, where $\varepsilon_e = \int (2\,\dot{\eta}_{ij}\dot{\eta}^{ij}/3)^{1/2} dt$ is an effective strain, and $\Delta\varepsilon_e$ denotes the amount of this strain accumulated since last remeshing. Here, the limiting value $(\Delta\varepsilon_e)_{\max}$ is mostly taken to be 0.2, but sometimes as low as 0.1, depending on how much mesh distortion is tolerated by the solution.

In a new mesh the points on the void surface and on the blunted crack-tip are calculated by interpolation between old points on these surfaces. The new mesh is stretched such that elements are relatively finer in the regions where very large strains develop.

4 Results

For the glass the elastic material parameters are Young's modulus E_g and Poisson's ratio ν_g, while for the metal wire the corresponding elastic material parameters are denoted E and ν, the initial yield stress is σ_y and the power hardening exponent is N. With a lead wire in glass the ratio of the elastic moduli is taken to be $E_g/E = 4.3$ and $\nu_g = 0.23$ is used. For the metal wire two sets of material parameters are used, first a model material with $\sigma_y/E = 0.003$, $\nu = 0.3$ and $N = 0.1$, which makes the cavitation predictions comparable to those in Tvergaard et al. [3]. The second set of elastic-plastic material parameters for the metal are based on material tables for lead and on the specifications of Ashby et al. [9], $\sigma_y/E = 0.0004$, $\nu = 0.42$ and $N = 0.2$.

When residual stresses are applied by the procedure described in Section 2, just below Eq. (2.7), the residual stresses in the wire, after that the additional forces have been stepped down, can be calculated by simple formulas for long elastic tubes, making the approximation that the specimen is long compared to its radius and neglecting plasticity. For the model material parameters ($\nu = 0.3$) this gives the following values of the residual principal stress components in the wire $\sigma_1^R = 0.7777\sigma^*$, $\sigma_2^R = \sigma_3^R = 0.6295\sigma^*$, and thus the mean residual stress $\sigma_m^R = 0.6789\sigma^*$ (where σ^* is the hydrostatic stress level first applied in the wire). A purely elastic computation for the specimen geometry and the boundary conditions considered here has shown that these residual stress values are obtained with very good approximation, even though the specimen is not much longer than its radius.

For the second set of material parameters ($\nu = 0.42$) the same procedure leads to the following residual principal stresses $\sigma_1^R = 0.5179\sigma^*$, $\sigma_2^R = \sigma_3^R = 0.4261\sigma^*$, and thus the mean residual stress $\sigma_m^R = 0.4567\sigma^*$. In the following figures the level of residual stress is specified by the value of the ratio σ_m^R/σ_y.

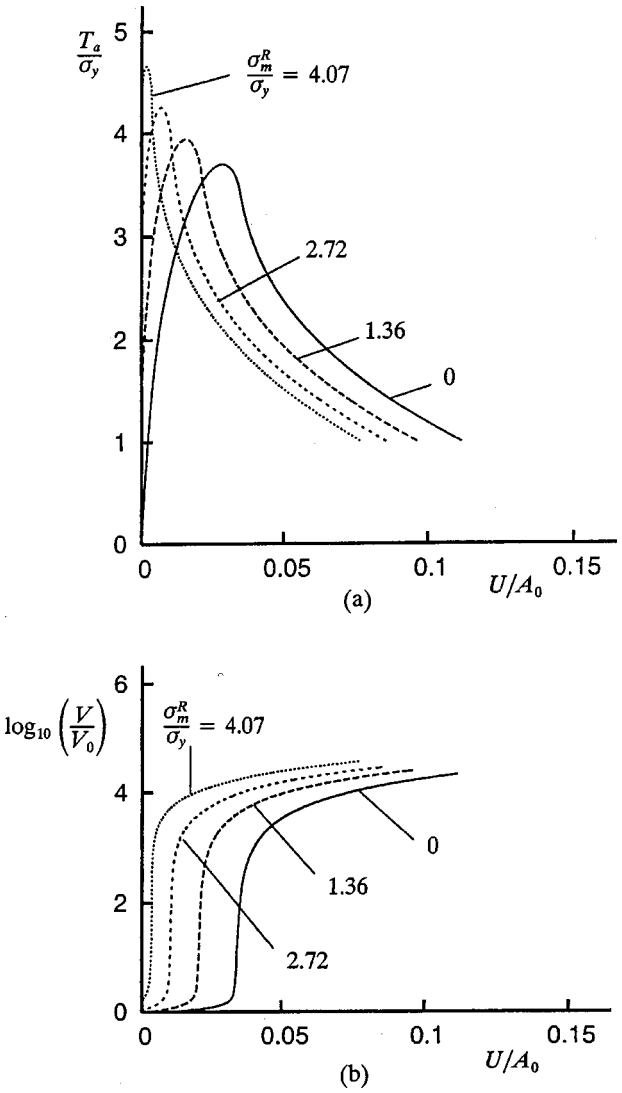


Fig. 3 Effect of different residual stress levels for initial void size $R_v/A_0 = 0.01$ and crack-tip radius $R_c/A_0 = 0.01$, with metal wire material parameters $\sigma_y/E = 0.003$, $\nu = 0.3$ and $N = 0.1$. (*a*) Average nominal stress versus end displacement. (*b*) Void volume growth.

Figure 3 shows results using the first set of material parameters for the metal wire, and using $R_v/A_0 = R_c/A_0 = 0.01$. The figure shows the evolution of the average nominal traction T_a in the metal wire and the ratio V/V_0 of the current and the initial void volumes, versus the end displacement U, where $2U$ gives a rather good estimate of the crack opening. One set of curves corresponds to an initially stress free wire, while the other three sets of curves are obtained for various levels of tensile residual stresses. Thus, in one of cases, where $\sigma^*/\sigma_y = 4$ was applied initially, the expressions given above resulted in $\sigma_m^R/\sigma_y = 2.72$. The curves for $\sigma_m^R = 0$ in Fig. 3 show that at the maximum traction the value of the ratio V/V_0 grows by two or three orders of magnitude while U increases only very slowly. Thus, the behavior is close to a cavitation instability, but a real cavitation instability would occur at constant U, corresponding to a constant remote stress and strain state. When a residual stress is applied before the tensile loading Fig. 3(*a*) shows that the maximum traction reached increases for

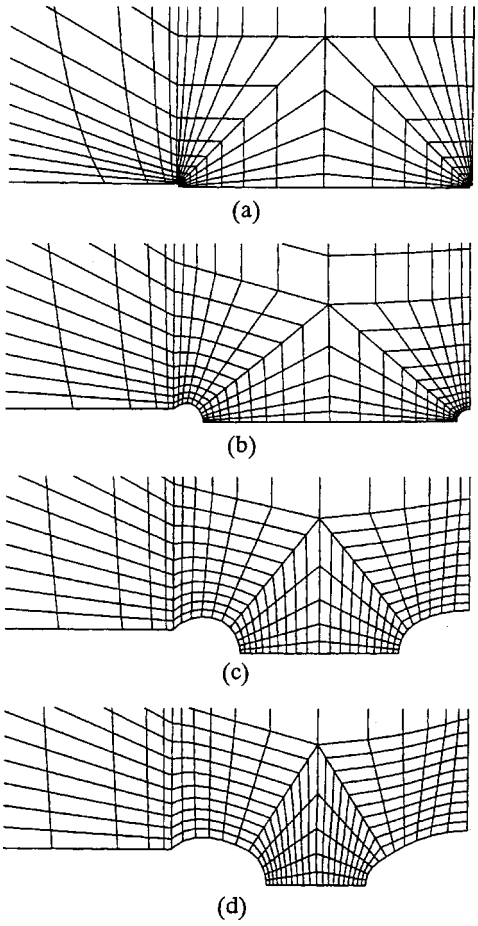


Fig. 4 Deformed meshes at four different stages for $\sigma_m^R=0$. The initial void size and crack tip radius are $R_v/A_0=0.01$ and $R_c/A_0=0.01$, and parameters for the metal wire are $\sigma_y/E=0.003$, $\nu=0.3$, and $N=0.1$. (*a*) Initial mesh. (*b*) Stage where $V/V_0=100$. (*c*) Stage where $V/V_0=10^4$. (*d*) End of computation.

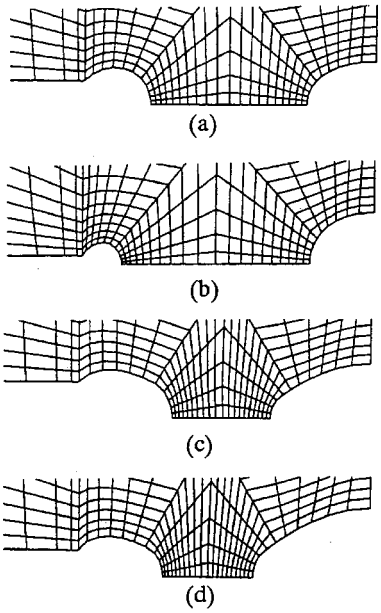


Fig. 5 Comparison of deformed meshes for two different values of σ_m^R. The initial void size and crack-tip radius are $R_v/A_0=0.01$ and $R_c/A_0=0.01$, and parameters for the metal wire are $\sigma_y/E=0.003$, $\nu=0.3$ and $N=0.1$. (*a*) Stage where $V/V_0=10^4$ for $\sigma_m^R=0$. (*b*) Stage where $V/V_0=10^4$ for $\sigma_m^R/\sigma_y=4.07$. (*c*) End of computation for $\sigma_m^R=0$. (*d*) End of computation for $\sigma_m^R/\sigma_y=4.07$.

increasing level of the tensile residual stress, and Fig. 3(*b*) shows that the rapid increase of the void volume occurs at a smaller value of U. Also, the value of U is closer to being constant during the rapid void growth, when the residual stress level is higher. The model material used here for the metal wire is identical to one considered by Tvergaard et al. [3], so it is known that a cavitation instability occurs at a mean stress $\sigma_m=5.2\sigma_y$. Therefore, it is not surprising that the critical value of U/A_0 decreases towards zero in Fig. 3(*b*) when the mean value of the applied residual stress, σ_m^R, approaches the value that would give a cavitation instability before applying any tensile load.

For the calculation in Fig. 3 with no residual stress Fig. 4 shows the initial mesh and three stages of the deformed mesh, when the void volume has increased by a factor 100, when the void volume has increased by a factor 10^4, and at the end of the computation defined here as the stage where the average nominal traction has decayed to the level $T_a/\sigma_y=1$. It is seen that while the void grows in the center of the metal wire, the blunted crack tip develops into something like a toroidal void, as has also been found for a ductile particle bridging a crack in a rigid matrix, [18].

Figure 5 shows comparisons of the deformed meshes for $\sigma_m^R=0$ and $\sigma_m^R/\sigma_y=4.07$, at the stage where the void volume has increased by a factor 10^4 and at the end of the computations. Thus, Figs. 5(*a*) and 5(*c*) are identical to Figs. 4(*c*) and 4(*d*), respectively. Figure 5(*b*) shows clearly that when $V/V_0=10^4$ has been reached, the toroidal void at the crack-tip has grown much less than that in Fig. 5(*a*), in agreement with the much smaller value of the crack opening $2U$. The same tendency is visible when comparing Figs. 5(*c*) and 5(*d*), but more interesting is the fact that the void radius in Fig. 5(*d*) has reached the value $0.41A_0$, while in Fig. 5(*c*) the void radius has reached the smaller value $0.34A_0$.

The sensitivity to the initial void radius is considered in Fig. 6 by comparing results for three different values of R_v/A_0. Thus, one set of curves in Fig. 6 is identical to the curves for $\sigma_m^R/\sigma_y=2.72$ in Fig. 3, and comparison is made with results for a smaller and a larger initial void size. Figure 6(*a*) shows hardly any sensitivity of the average traction versus end displacement curve, while Fig. 6(*b*) shows significant differences between the values reached for V/V_0. These differences are expected, as the deformed meshes at the end of the computation are quite similar to Fig. 5(*d*) in all three cases, but reaching the same final void size requires a much larger increase of the void size when the void was initially smaller. It is also seen in Fig. 6(*b*) that the value of U during rapid void growth is closer to being constant when the void is initially smaller, which reflects the fact that the smaller void can grow more before the remote stress state is affected by the void size.

Figure 7 keeps the initial void radius fixed at $R_v/A_0=0.01$, but studies the sensitivity to the assumed initial rounding of the crack-tip by comparing predictions for three different values of R_c/A_0. As in Fig. 6 the comparison is made for $\sigma_m^R/\sigma_y=2.72$, and thus the set of curves for $R_c/A_0=0.01$ in Fig. 7 appears also in Figs. 6 and 3. The peak value of the average traction T_a shows some sensitivity in Fig. 7(*a*), with rather little difference between the peak values for $R_c/A_0=0.003$ and $R_c/A_0=0.01$, but with a noticeably lower peak for $R_c/A_0=0.03$. The void size vs. end displacement curves in Fig. 7(*b*) show practically no sensitivity to the value of R_c/A_0, although the curve corresponding to the lowest peak in Fig. 7(*a*) differs a little.

The second set of material parameters for the metal wire, more realistic for lead based on material tables, is considered in Fig. 8. As in Fig. 3, the comparison here is carried out for R_v/A_0

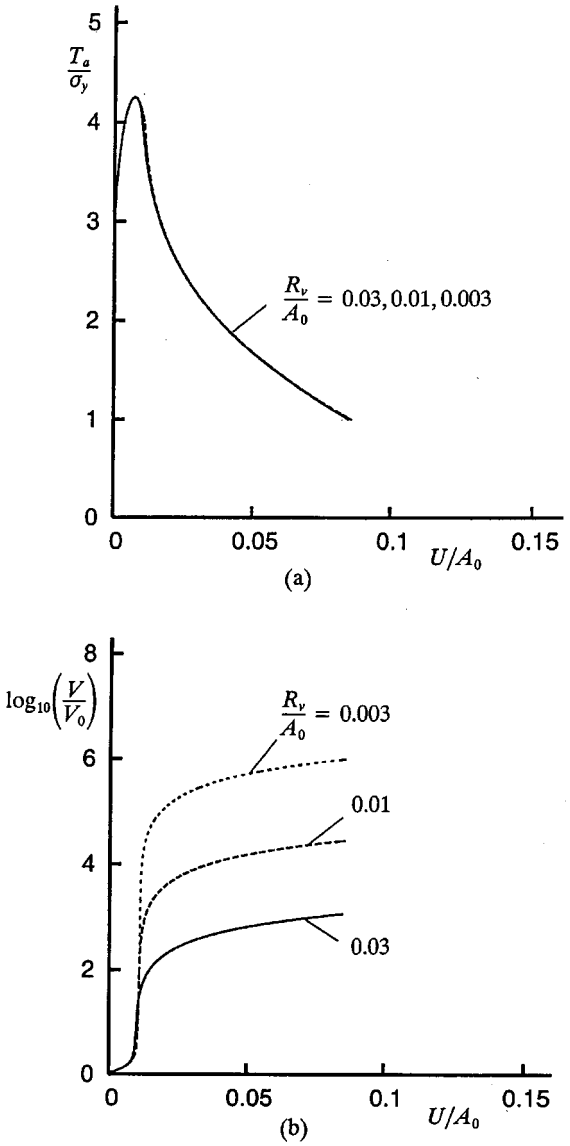


Fig. 6 Effect of different values of the initial void size, for $\sigma_m^R/\sigma_y=2.72$ and $R_c/A_0=0.01$. Parameters for the metal wire are $\sigma_y/E=0.003$, $\nu=0.3$ and $N=0.1$. (*a*) Average nominal stress versus end displacement. (*b*) Void volume growth.

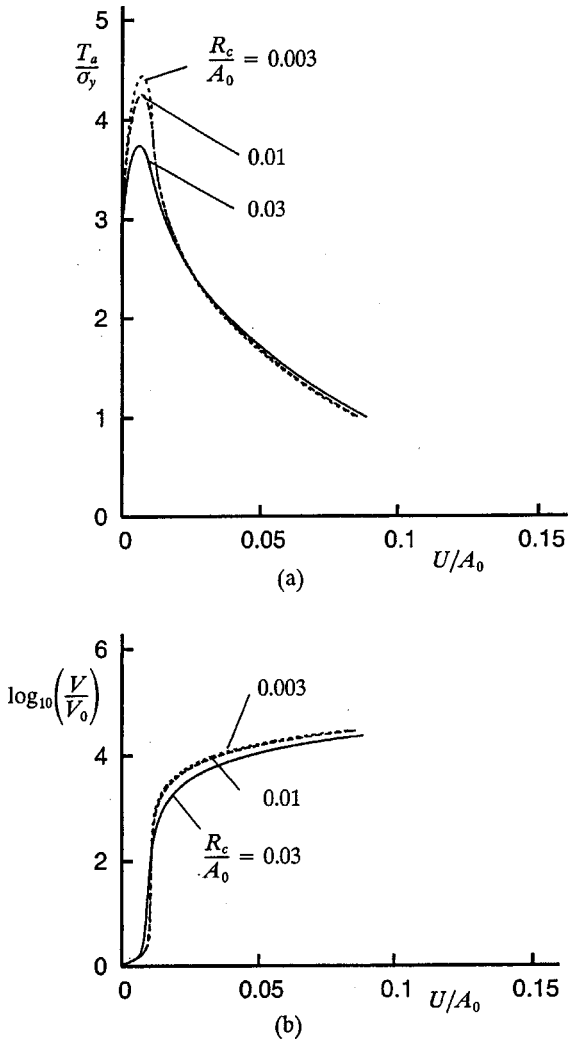


Fig. 7 Effect of different values of the initial crack-tip radius, for $\sigma_m^R/\sigma_y=2.72$ and $R_v/A_0=0.01$. Parameters for the metal wire are $\sigma_y/E=0.003$, $\nu=0.3$ and $N=0.1$. (*a*) Average nominal stress versus end displacement. (*b*) Void volume growth.

$=R_c/A_0=0.01$, and the different sets of curves correspond to different levels of the initial residual stress field. The initial values of σ^*/σ_y applied here are much larger than those applied for the reference material in Fig. 3, but this relates to the finding of Huang et al. [2] that a lower value of σ_y/E gives a higher value of the critical stress level σ_m/σ_y for the occurrence of a cavitation instability, and also the increased value of the power hardening exponent N gives a significant increase of the critical stress level. The highest residual stress level, $\sigma_m^R/\sigma_y=8.22$, considered here is obtained by first applying the hydrostatic stress level $\sigma^*/\sigma_y=18$ and then stepping down additional forces to obtain equilibrium, as described above. This highest level of σ_m^R/σ_y is only slightly below the critical cavitation level, and therefore the nearly vertical part of the corresponding growth curve in Fig. 8(*b*) occurs at a very small value of U/A_0. The difference between the four different sets of curves in Fig. 8 is analogous to that found in Fig. 3. Thus, the peak value of the average nominal traction T_a in the metal wire increases for increasing level of the residual stress, and the higher peaks occur at lower values of U/A_0. However, the differences between the levels of the peaks are not quite as pronounced in Fig. 8 as in Fig. 3. The curves in Fig. 8(*b*) for the three lower values of σ_m^R/σ_y are not vertical at any stage, but for $\sigma_m^R/\sigma_y=8.22$ an early part, below $\log_{10}(V/V_0)=2$, shows void growth at a constant value of U, so here a cavitation instability has occurred.

Figure 9 shows deformed meshes analogous to those in Fig. 5, comparing results for $\sigma_m^R=0$ and $\sigma_m^R/\sigma_y=8.22$, with the initial geometry identical to that considered in Fig. 8. Figures 9(*a*) and 9(*b*) illustrate stages where $V/V_0=10^4$, which is well beyond the rapid void growth range and well beyond the traction peaks, as is seen in Fig. 8. For the high residual stress (Fig. 8(*b*)) this void size is reached at a stage where the toroidal void at the crack-tip has grown much less than found in the absence of residual stress. Some of this tendency is also seen at the end of the computations, for $T_a/\sigma_y=1$, where Fig. 9(*d*) shows that the void radius on the symmetry plane has reached the value $0.49A_0$, while in Fig. 9(*c*) this void radius has only reached the value $0.46A_0$.

The comparison in Fig. 10 is somewhat analogous to that in Fig. 7. Thus, the initial void radius is kept fixed at $R_v/A_0=0.01$, while two values 0.01 and 0.0002 are considered for the initial rounding of the crack-tip, R_c/A_0. It is seen that the smaller value of R_c/A_0 does give a higher peak value of the average

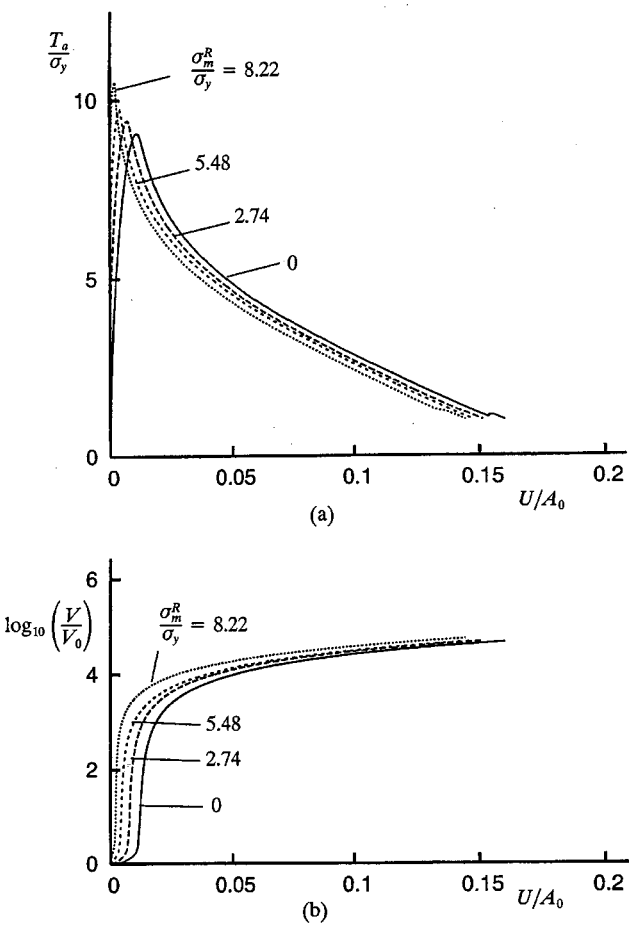


Fig. 8 Effect of different residual stress levels for initial void size $R_v/A_0=0.01$ and crack-tip radius $R_c/A_0=0.01$, with metal wire material parameters $\sigma_y/E=0.0004$, $\nu=0.42$ and $N=0.2$. (*a*) Average nominal stress versus end displacement. (*b*) Void volume growth.

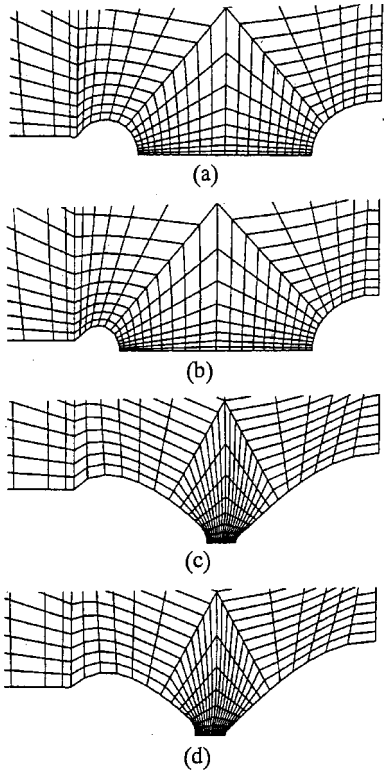


Fig. 9 Comparison of deformed meshes for two different values of σ_m^R. The initial void size and crack-tip radius are $R_v/A_0=0.01$ and $R_c/A_0=0.01$, and parameters for the metal wire are $\sigma_y/E=0.0004$, $\nu=0.42$ and $N=0.2$. (*a*) Stage where $V/V_0=10^4$ for $\sigma_m^R=0$. (*b*) Stage where $V/V_0=10^4$ for $\sigma_m^R/\sigma_y=8.22$. (*c*) End of computation for $\sigma_m^R=0$. (*d*) End of computation for $\sigma_m^R/\sigma_y=8.22$.

traction T_a, and also part of the corresponding void growth curve in Fig. 10(*b*) is practically vertical. Thus, a sharper crack-tip increases the tendency to observe a real cavitation instability in the center of the metal wire.

5 Discussion

An earlier investigation of cavity growth in ductile particles bridging a brittle matrix crack (Tvergaard [8]) had some features in common with the present study, including that effects of residual tensile stresses in the ductile particles were accounted for. However, the elastic deformations of the surrounding ceramic were not incorporated in these previous analyses and at large void expansions the accuracy of the results was limited by strong mesh distortion, as had also been true in other early studies of cavitation instabilities. In the analyses here, the ability to evaluate the final size and shape of the void in the special tensile test specimen of Ashby et al. [9] is entirely dependent on the application of a remeshing procedure.

In the experiments modelled here the specimens were cooled slowly from the processing temperature to minimize the residual stresses due to thermal contraction mismatch between the lead and the glass. However, some level of tensile residual stresses may have been left in the metal wire, and such residual stresses have been incorporated in the present analyses to get a parametric understanding of their influence on the growth of a single void in the center of the wire, in the crack plane. It has been found that a tensile residual stress increases the peak value of the average nominal traction in the metal wire, and reduces the end displacement needed to reach a cavitation instability, since the residual stress may be high enough to be nearly at the cavitation limit when the tensile test starts.

Perhaps more interesting is the fact that the experiments have shown a final diameter of the single void in the wire more than half of the wire diameter, and that the parametric studies show an increasing value of the final void diameter for increasing values of the tensile residual stress. Also, the predicted final void diameters approach half the wire diameter at the highest levels of residual stress considered here. It is noted that the end of the computations is rather arbitrarily defined as the stage where the average traction has dropped to the value $T_a=\sigma_y$, but if a smaller value had been chosen, the final void diameters would have been larger. The final geometries shown in Figs. 5(*c*), 5(*d*), 9(*c*), and 9(*d*) do not exactly agree with the photographs of the final fracture surfaces shown by Ashby et al. [9]. Thus, in the experiments the void has not grown exactly in the center of the wire, and the remaining ligament appears to have failed by shearing off. But parts of the observed fracture surfaces do indicate the type of crack growth by a blunting mechanism shown in Figs. 5 and 9 by a toroidal void growing from the initially sharp crack-tip.

The cavitation instability or near cavitation instability behavior occurs at the vertical or near vertical parts of the curves in Figs. 3(*b*) and 8(*b*), and it is noted that the calculated final void shapes with large diameters occur far beyond these vertical parts. In fact, even the stages illustrated in Figs. 5(*a*), 5(*b*), 9(*a*), and 9(*b*) occur well beyond the vertical parts of the void growth curves. Thus, when the large voids in these lead wires are seen as examples of experimentally observed cavitation instabilities, the instability

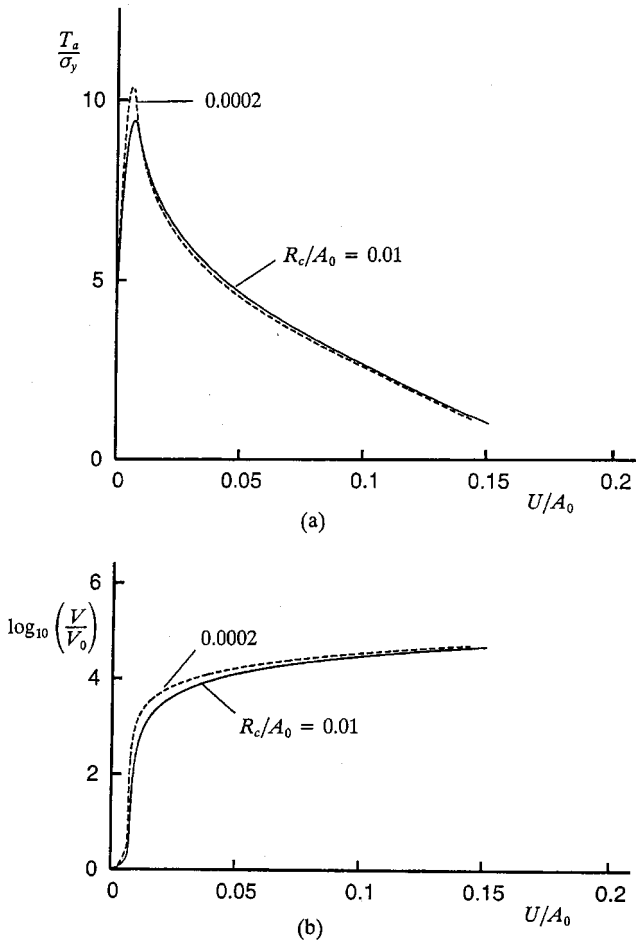


Fig. 10 Effect of different crack-tip radii, for the initial void size $R_v/A_0=0.01$, with metal wire material parameters $\sigma_y/E=0.0004$, $\nu=0.42$ and $N=0.2$. (*a*) Average nominal stress versus end displacement. (*b*) Void volume growth.

mechanism is really only active in the beginning of cavity growth, until the void diameter has reached about 10% of the metal wire diameter.

Two sets of material parameters for the elastic-plastic behavior of the metal wire are considered here. The first set of material parameters are those also used by Tvergaard et al. [3], and indeed the peak values of the average nominal traction T_a reached in Fig. 3 are in good agreement with the levels of the cavitation instability limits predicted in the previous study. The second set of material parameters are based on material tables for lead and on the specifications of Ashby et al. [9], and here the peak values of T_a predicted in Fig. 8 are a bit on the high side of most of the experimental results.

The level of tensile residual stress to be expected in a ductile reinforcing particle in a ceramic, due to cooling from the processing temperature, has been estimated in [8], by an elastic analysis for a spherical well bonded Al particle in an Al_2O_3 matrix. When the cooling range is $\Delta T=500°C$ it was found that $\sigma_R/\sigma_y=2.7$ if $\sigma_y/E=0.005$, while $\sigma_R/\sigma_y=13.5$ if $\sigma_y/E=0.001$. Thus, the levels of residual stresses considered in the present paper are not unrealistic. In addition to the effect on early occurrence of a cavitation instability the tensile residual stresses increase the peak value of the nominal traction for the crack-bridging mechanism (Figs. 3 and 8). But the figures also show that the area under the traction separation curve is reduced, so that the energy needed to break the bridging particle is smaller.

References

[1] Bishop, R. F., Hill, R., and Mott, N. F., 1945, ''The Theory of Indentation and Hardness Tests,'' Proc. Phys. Soc. London, **57**, pp. 147–159.

[2] Huang, Y., Hutchinson, J. W., and Tvergaard, V., 1991, ''Cavitation Instabilities in Elastic-Plastic Solids,'' J. Mech. Phys. Solids, **39**, pp. 223–241.

[3] Tvergaard, V., Huang, Y., and Hutchinson, J. W., 1992, ''Cavitation Instabilities in a Power Hardening Elastic-Plastic Solid,'' Eur. J. Mech. A/Solids, **11**, pp. 215–231.

[4] Ball, J. M., 1982, ''Discontinuous Equilibrium Solutions and Cavitation in Nonlinear Elasticity,'' Philos. Trans. R. Soc. London, Ser. A, **A306**, pp. 557–610.

[5] Horgan, C. O., and Abeyaratne, R., 1986, ''A Bifurcation Problem for a Compressible Nonlinearly Elastic Medium: Growth of a Microvoid,'' J. Elast., **16**, pp. 189–200.

[6] Chou-Wang, M.-S., and Horgan, C. O., 1989, ''Void Nucleation and Growth for a Class of Incompressible Nonlinearly Elastic Materials,'' Int. J. Solids Struct., **25**, pp. 1239–1254.

[7] Flinn, B., Rühle, M., and Evans, A. G., 1989, ''Toughening in Composites of Al_2O_3 Reinforced With Al,'' Materials Department Report, University of California, Santa Barbara, CA.

[8] Tvergaard, V., 1995, ''Cavity Growth in Ductile Particles Bridging a Brittle Matrix Crack,'' Int. J. Fract., **72**, pp. 277–292.

[9] Ashby, M. F., Blunt, F. J., and Bannister, M., 1989, ''Flow Characteristics of Highly Constrained Metal Wires,'' Acta Metall., **37**, pp. 1847–1857.

[10] Dalgleish, B. J., Trumble, K. P., and Evans, A. G., 1989, ''The Strength and Fracture of Alumina Bonded With Aluminum Alloys,'' Acta Metall., **37**, pp. 1923–1931.

[11] Tvergaard, V., 1991, ''Failure by Ductile Cavity Growth at a Metal-Ceramic Interface,'' Acta Metall. Mater., **39**, pp. 419–426.

[12] Tvergaard, V., 1997, ''Studies of Void Growth in a Thin Ductile Layer Between Ceramics,'' Comput. Mech., **20**, pp. 186–191.

[13] Tvergaard, V., 2000, ''Interface Failure by Cavity Growth to Coalescence,'' Int. J. Mech. Sci., **42**, pp. 381–395.

[14] Pedersen, T. Ø., 1998, ''Remeshing in Analysis of Large Plastic Deformations,'' Comput. Struct., **67**, pp. 279–288.

[15] Hutchinson, J. W., 1973, ''Finite Strain Analysis of Elastic-Plastic Solids and Structures,'' *Numerical Solution of Nonlinear Structural Problems*, R. F. Hartung, ed., **17**, ASME, New York.

[16] Tvergaard, V., 1982, ''On Localization in Ductile Materials Containing Spherical Voids,'' Int. J. Fract., **18**, pp. 237–252.

[17] Tvergaard, V., 1976, ''Effect of Thickness Inhomogeneities in Internally Pressurized Elastic-Plastic Spherical Shells,'' J. Mech. Phys. Solids, **24**, pp. 291–304.

[18] Tvergaard, V., 1992, ''Effect of Ductile Particle Debonding During Crack-Bridging in Ceramics,'' Int. J. Mech. Sci., **34**, pp. 635–649.

Journal of Applied Mechanics

Brief Notes

A Brief Note is a short paper that presents a specific solution of technical interest in mechanics but which does not necessarily contain new general methods or results. A Brief Note should not exceed 2500 words *or equivalent* (a typical one-column figure or table is equivalent to 250 words; a one line equation to 30 words). Brief Notes will be subject to the usual review procedures prior to publication. After approval such Notes will be published as soon as possible. The Notes should be submitted to the Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME International, Three Park Avenue, New York, NY 10016-5990, or to the Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on Brief Notes appearing in this issue will be accepted until two months after publication. Readers who need more time to prepare a Discussion should request an extension of the deadline from the Editorial Department.

Nonaxisymmetric Dynamic Problem of a Penny-Shaped Crack in a Three-Dimensional Piezoelectric Strip Under Normal Impact Loads

Ji Hyuck Yang

Kang Yong Lee
Professor

Department of Mechanical Engineering, Yonsei University, Seoul 120-749, Korea

[DOI: 10.1115/1.1636793]

1 Introduction

The dynamic response of a penny-shaped crack in a three-dimensional piezoelectric ceramic strip under nonaxisymmetric normal mechanical and electrical impact loads is analyzed based on the continuous electric boundary conditions on the crack surface. The potential theory and Hankel and Laplace transforms are used to obtain the systems of dual integral equations, which are then expressed to Fredholm integral equations. The singular mechanical and electric fields and all sorts of dynamic field intensity factors of Mode I are obtained, and the numerical values of various field intensity factors for PZT-6B piezoelectric ceramic are graphically shown for transverse bending impact loads.

2 Problem Statements and Method of Solution

Consider a piezoelectric strip of thickness $2h$ containing a center penny-shaped crack of diameter $2a$ subjected to the nonaxisymmetric combined mechanical and electrical impact loads. The cylindrical coordinates (r,θ,z) is set at the center of the crack. The piezoelectric strip is transversely isotropic with hexagonal symmetry, and the z-axis is oriented in the poling direction. The strip is subjected to a nonaxisymmetric normal impact stress or strain at the edges, and the electrical boundary condition of a nonaxisymmetric electric displacement or electric field for the piezoelectric layer is considered. In the absence of body forces, the governing equations for nonaxisymmetric loads are

$$n_i\frac{\partial^2\tilde{\Phi}_i}{\partial z^2}-\xi^2\tilde{\Phi}_i=0,\quad (i=1,2,3)$$
$$c_{44}\frac{\partial^2\tilde{\Phi}_4}{\partial z^2}-\xi^2\left[\frac{c_{11}-c_{12}}{2}-\rho q^2\right]\tilde{\Phi}_4=0 \tag{1}$$

where

$$n_i(\xi,p)=\frac{e_{33}k_{1i}+d_{33}k_{2i}}{e_{15}k_{1i}+e_{15}+e_{31}+d_{11}k_{2i}}$$
$$=\frac{c_{33}k_{1i}-e_{33}k_{2i}}{c_{44}k_{1i}+c_{13}+c_{44}-e_{15}k_{2i}+\rho q^2k_{1i}}$$
$$=\frac{c_{44}+(c_{13}+c_{44})k_{1i}-(e_{15}+e_{31})k_{2i}}{c_{11}+\rho q^2}, \tag{2}$$

$q^2=p^2/\xi^2$, $(c_{11},c_{12},c_{13},c_{33},c_{44})$ are the elastic moduli measured in a constant electric field, (d_{11},d_{33}) are the dielectric permittivities measured at a constant strain, (e_{15},e_{31},e_{33}) are the piezoelectric constants, k_{ij} $(i=1,2,\ j=1,2,3)$ are the unknown constants to be determined from Eq. (2), $\tilde{\Phi}_i(\xi,z,p)$ $(i=1,2,3,4)$ is the function transformed from $\underline{\phi_i}(r,z,t)$ $(i=1,2,3,4)$ by Hankel and Laplace transforms, and ϕ_i is related to the potential functions $\phi_k(r,\theta,z)$ $(k=1,2,3,4)$ as follows:

$$\phi_k=\overline{\phi}_k\exp(in\theta),\quad k=1,2,3,4,\quad n=0,1,2,\ldots. \tag{3}$$

According to Gao and Fan [1], we set up the following boundary conditions:

$$\sigma_z^*(r,\theta,0,p)=0\quad (0\leqslant r<a),$$
$$u_z^*(r,\theta,0,p)=0\quad (a\leqslant r<\infty), \tag{4}$$
$$D_z^*(r,\theta,0^+,p)=D_z^*(r,\theta,0^-,p)\quad (0\leqslant r<a),$$
$$E_r^*(r,\theta,0^+,p)=E_r^*(r,\theta,0^-,p)\quad (0\leqslant r<a), \tag{5}$$
$$\phi^*(r,\theta,0,p)=0\quad (a\leqslant r<\infty),$$
$$\sigma_{rz}^*(r,\theta,0,p)=0, \tag{6}$$
$$\sigma_{\theta z}^*(r,\theta,0,p)=0, \tag{7}$$
$$\sigma_{rz}^*(r,\theta,h,p)=0, \tag{8}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Dec. 13, 2000; final revision, Dec. 12, 2001. Associate Editor: K. Ravi-Chandar.

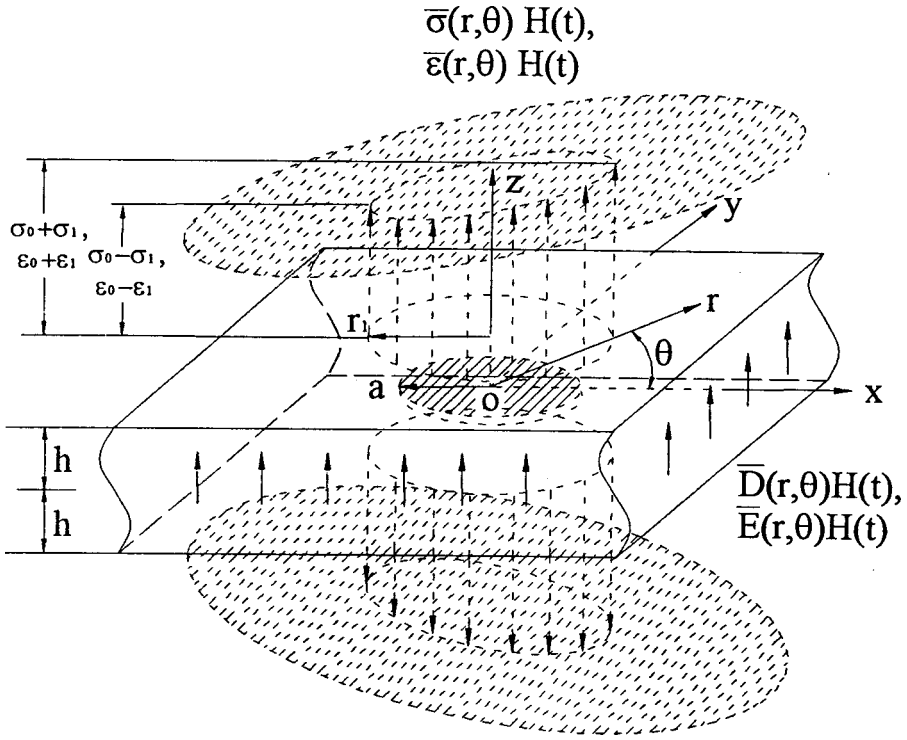


Fig. 1 Infinite piezoelectric strip with a penny-shaped crack subjected to combined mechanical transverse bending and electrical impact loads

$$\sigma_{\theta z}^*(r,\theta,h,p)=0, \tag{9}$$

Case 1: $$\sigma_z^*(r,\theta,h,p)=\frac{\overline{\sigma}(r,\theta)}{p}, \quad D_z^*(r,\theta,h,p)=\frac{\overline{D}(r,\theta)}{p}, \tag{10}$$

Case 2: $$\varepsilon_z^*(r,\theta,h,p)=\frac{\overline{\varepsilon}(r,\theta)}{p}, \quad E_z^*(r,\theta,h,p)=\frac{\overline{E}(r,\theta)}{p}, \tag{11}$$

Case 3: $$\sigma_z^*(r,\theta,h,p)=\frac{\overline{\sigma}(r,\theta)}{p}, \quad E_z^*(r,\theta,h,p)=\frac{\overline{E}(r,\theta)}{p}, \tag{12}$$

Case 4: $$\varepsilon_z^*(r,\theta,h,p)=\frac{\overline{\varepsilon}(r,\theta)}{p}, \quad D_z^*(r,\theta,h,p)=\frac{\overline{D}(r,\theta)}{p}, \tag{13}$$

where σ_r, σ_θ, σ_z are normal stresses, $\sigma_{r\theta}$, σ_{rz}, $\sigma_{\theta z}$ are shear stresses, $D_k\,(k=r,\theta,z)$ are electric displacements, $u_k\,(k=r,\theta,z)$ are displacements, ϕ is electric potential, superscript * means the Laplace domain, $\overline{\sigma}(r,\theta)$, $\overline{\varepsilon}(r,\theta)$, $D(r,\theta)$, and $E(r,\theta)$ are the magnitudes of applied stress, strain, electric displacement, and electric field, respectively. The field equations are obtained from Eqs. (1) as follows:

$$u_z^*=\sum_{n=0}^{\infty}\left\{H_n^{-1}\left(\frac{1}{\xi}\sum_{i=1}^{3}k_{1i}s_i[A_{in}\sinh(s_i\xi z)+B_{in}\cosh(s_i\xi z)]\right)\right\}\cos n\theta+\frac{\overline{a}(r,\theta)}{p}z, \tag{14}$$

$$\phi^*=-\sum_{n=0}^{\infty}\left\{H_n^{-1}\left(\frac{1}{\xi}\sum_{i=1}^{3}k_{2i}s_i[A_{in}\sinh(s_i\xi z)+B_{in}\cosh(s_i\xi z)]\right)\right\}\cos n\theta-\frac{\overline{b}(r,\theta)}{p}z, \tag{15}$$

where

$$s_i(\xi,p)=\frac{1}{\sqrt{n_i(\xi,p)}}, \quad (i=1,2,3), \tag{16}$$

A_{in}, B_{in} $(i=1,2,3,\ n=1,2,3,\ldots)$ are the unknowns to be solved, $\overline{a}$ and b are real constants, which will be determined from the edge loading conditions, H_n^{-1} is the inverse Hankel transform of

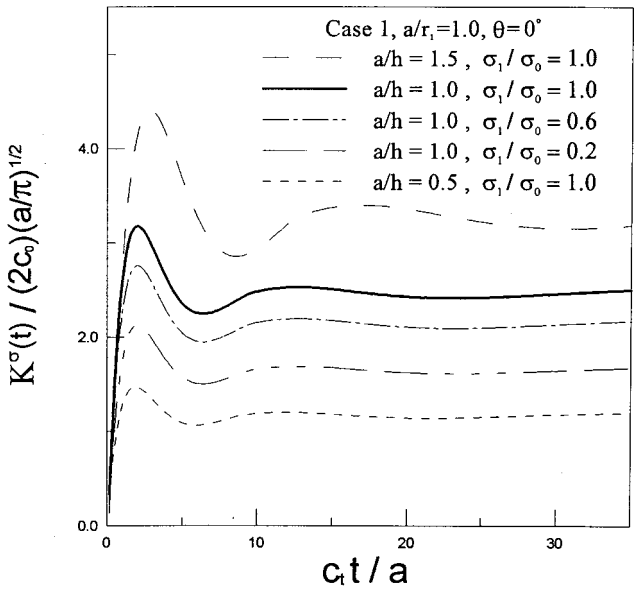


Fig. 2 Change of the normalized dynamic stress intensity factor with the normalized time, the ratio of crack radius to layer thickness, and the ratio of σ_1/σ_0 for PZT-6B ceramic under transverse bending loads of Case 1

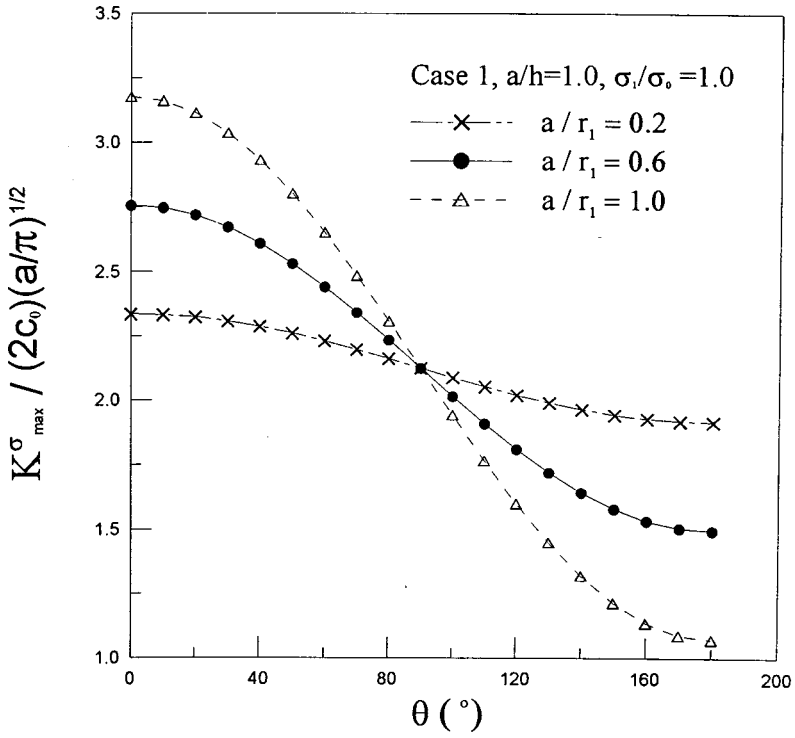


Fig. 3 Change of the maximum normalized dynamic stress intensity factor for the time with the polar angle θ and the ratio a/r_1 for PZT-6B ceramic under transverse bending loads of Case 1

order *n*. Applying Eqs. (3)–(13), we obtain *n*th systems of dual integral equations, which are then expressed to the following *n*th Fredholm integral equations of the second kind:

$$\phi_n(\alpha,p)+\int_0^a \phi_n(\beta,p)l_n(\alpha,\beta)d\beta = \sqrt{\frac{2}{\pi}}\left(-\frac{1}{pm_0}\right)\alpha^{1/2-n}\int_0^\alpha \frac{r^{1+n}}{\sqrt{\alpha^2-r^2}}c_n(r)dr \tag{17}$$

where

$$l_n(\alpha,\beta)=\alpha\int_0^\infty \xi[N(\xi,p)-1]J_{n+1/2}(\alpha\xi)J_{n+1/2}(\beta\xi)d\xi \tag{18}$$

$$N(\xi,p)=[F_{11}(\xi,p)M_{41}(\xi,p)+F_{12}(\xi,p)M_{42}(\xi,p)+F_{13}(\xi,p)M_{43}(\xi,p)]/m_0, \tag{19}$$

$$m_0=\lim_{\xi\to\infty}\left[\sum_{i=1}^{3} F_{1i}(\xi,p)M_{4i}(\xi,p)\right], \tag{20}$$

$$B_{1n}(\xi,p)=\xi^{1/2}\int_0^a \psi_n(\alpha,p)J_{n+1/2}(\xi\alpha)d\alpha, \tag{21}$$

$J_n(\,)$ is the Bessel function of order *n* of the first kind, $F_{ij}\,(i,j=1,2,3)$ and $M_{4j}\,(j=1,2,3)$ have the same forms as those of Yang and Lee [2]. Here, we assume that the applied loads can be expanded with Fourier series as follows:

$$\overline{\sigma}(r,\theta)=\sum_{n=0}^{\infty}\sigma_n(r)\cos n\theta,\quad \overline{\varepsilon}(r,\theta)=\sum_{n=0}^{\infty}\varepsilon_n(r)\cos n\theta,$$
$$\overline{D}(r,\theta)=\sum_{n=0}^{\infty}D_n(r)\cos n\theta,\quad \overline{E}(r,\theta)=\sum_{n=0}^{\infty}E_n(r)\cos n\theta. \tag{22}$$

3 Numerical Results and Discussion

Material properties of PZT-6B, [3], and the normalized time is defined as follows:

$$T=tc_t/a, \tag{23}$$

where *t* is time, *a* is crack radius, and $c_t(\equiv\sqrt{(c_{44}+e_{15}^2/d_{11})/\rho})$ is the shear wave velocity defined from the governing equation.

Example 1: Nonaxisymmetric Transverse Bending Impact Loads. Let the mechanical loads be nonaxisymmetric transverse bending impact and the electrical loads be axisymmetric impact as shown in Fig. 1. Then, the applied loads can be expressed as follows:

$$\sigma_z(r,\theta,h,t)=\left(\sigma_0+\sigma_1\frac{r}{r_1}\cos\theta\right)H(t),\quad \text{(for Case 1,3)},$$
$$\varepsilon_z(r,\theta,h,t)=\left(\varepsilon_0+\varepsilon_1\frac{r}{r_1}\cos\theta\right)H(t),\quad \text{(for Case 2,4)},$$
$$D_z(r,\theta,h,t)=D_0H(t),\quad \text{(for Case 1,4)},$$
$$E_z(r,\theta,h,t)=E_0H(t),\quad \text{(for Case 2,3)}, \tag{24}$$

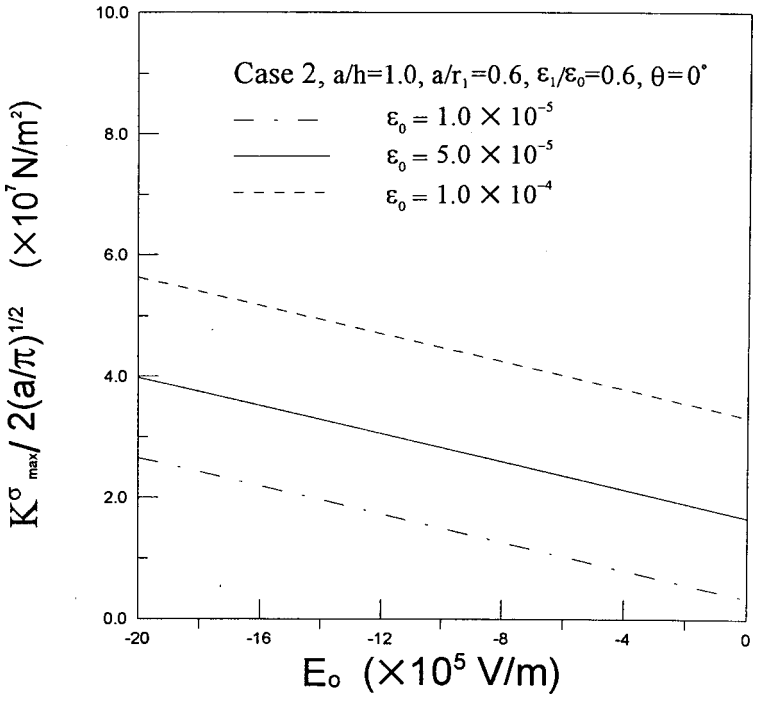


Fig. 4 Change of the maximum normalized dynamic stress intensity factor for the time with the electric field E_0 for PZT-6B ceramic under transverse bending loads of Case 2

where $H(t)$ is the Heaviside unit step function, σ_0 and σ_1 are the magnitudes of applied stress, D_0 is the magnitude of applied electric displacement at initial time, and r_1 is shown in Fig. 1. From Eqs. (22) by using the inversion of Laplace transform for Eq. (24), the first and second terms in the Fourier series are remained and two systems of Fredholm integral equations are obtained for $n=0$ and $n=1$. The Fredholm integral equation for $n=0$ is the same form as that of Yang and Lee [2] and the Fredholm integral equation for $n=1$ is in the form

$$\Psi_1(\Xi,p)+\int_0^1 \Psi_1(H,p)L_1(\Xi,H,p)dH=\Xi^2, \tag{25}$$

where

$$L_1(\Xi,H,p)=\int_0^\infty \sqrt{\Xi H}S\left[N\left(\frac{S}{a},p\right)-1\right]J_{3/2}(\Xi S)J_{3/2}(HS)dS, \tag{26}$$

$$N\left(\frac{S}{a},p\right)=\frac{F_{11}\left(\frac{S}{a},p\right)M_{41}\left(\frac{S}{a},p\right)+F_{12}\left(\frac{S}{a},p\right)M_{42}\left(\frac{S}{a},p\right)+F_{13}\left(\frac{S}{a},p\right)M_{43}\left(\frac{S}{a},p\right)}{m_0}, \tag{27}$$

$$S=\xi a,\quad \Xi=\frac{\alpha}{a},\quad H=\frac{\beta}{a}, \tag{28}$$

$$\Psi_1(\Xi,p)=-\sqrt{\frac{\pi}{2a\Xi}}\frac{3pm_0r_1}{2a^2c_1}\psi_1(\alpha,p),$$

$$\Psi_1(H,p)=-\sqrt{\frac{\pi}{2aH}}\frac{3pm_0r_1}{2a^2c_1}\psi_1(\beta,p),$$

$$c_1=\sigma_1,\quad \text{(Case 1,3)},$$

$$=c_{33}\varepsilon_1,\quad \text{(Case 2)}, \tag{29}$$

$$=\frac{(c_{33}d_{33}+e_{33}^2)\varepsilon_1}{d_{33}},\quad \text{(Case 4)}.$$

The field intensity factors in the Laplace domain become as follows:

$$K^{\sigma *}=\frac{2}{\pi}\sqrt{\pi a}\frac{1}{p}\left[c_0\Psi_0(1,p)+\frac{2ac_1}{3r_1}\Psi_1(1,p)\cos\theta\right], \tag{30}$$

$$K^{D*}=\frac{2}{\pi}\sqrt{\pi a}\frac{1}{p}\frac{m_1}{m_0}\left[c_0\Psi_0(1,p)+\frac{2ac_1}{3r_1}\Psi_1(1,p)\cos\theta\right], \tag{31}$$

$$K^{\varepsilon *}=\frac{2}{\pi}\sqrt{\pi a}\frac{1}{p}\frac{m_2}{m_0}\left[c_0\Psi_0(1,p)+\frac{2ac_1}{3r_1}\Psi_1(1,p)\cos\theta\right], \tag{32}$$

$$K^{E*}=\frac{2}{\pi}\sqrt{\pi a}\frac{1}{p}\frac{m_3}{m_0}\left[c_0\Psi_0(1,p)+\frac{2ac_1}{3r_1}\Psi_1(1,p)\cos\theta\right], \tag{33}$$

where

$$\begin{aligned} c_0&=\sigma_0, \quad \text{(Case 1,3)},\\ &=c_{33}\varepsilon_0-e_{33}E_0, \quad \text{(Case 2)},\\ &=\frac{(c_{33}d_{33}+e_{33}^2)\varepsilon_0-e_{33}D_0}{d_{33}}, \quad \text{(Case 4)}. \end{aligned} \tag{34}$$

These shows that both the mechanical and electrical loads effect the field intensity factors, respectively.

Two systems of Fredholm integral equations are solved numerically using Gaussian quadrature formula. Then dynamic field intensity factors are obtained for time from Eqs. (30)–(33). The accuracy of this numerical dynamic field intensity factor values are affected by the numerical inversion parameters such as N and δ, and Gauss Legendre and Laguerre integrating points. We use quasi-static field intensity factor values as a criterion to choose the value of N and δ. In this paper, we used values of $N=6$ and $\delta=0.4\sim1.8$ to fit the quasi-static field intensity factor values.

The change of the normalized dynamic stress intensity factor, $K^{\sigma}(t)/(2c_0)(a/\pi)^{1/2}$, with the normalized time, the ratio a/h and the ratio σ_1/σ_0 under transverse bending loads in Case 1 is shown in Fig. 2. It is shown that the normalized dynamic stress intensity factor increases with increase of the ratio a/h and the time of peak value changes a little with the ratio a/h. And as the ratio σ_1/σ_0 increases, the normalized dynamic stress intensity factor increases gradually.

The change of the peak value of the normalized dynamic stress intensity factor for the time with the polar angle θ and the ratio a/r_1 is shown in Fig. 3. The peak value increases as the ratio a/r_1 increases in the range 0 deg$<\theta<$90 deg, but the tendency is opposite in the range 90 deg$<\theta<$180 deg.

The change of the peak value of the normalized dynamic stress intensity factor for the time with the electric field E_0 for PZT-6B ceramic in Case 2 is shown in Fig. 4. The dynamic stress intensity factor increases as the magnitude of the electric field increases and it is concluded that the dynamic stress intensity factor changes with the direction of the electric field. And this also shows that the electrical load effects the mechanical field strongly. In Case 4, the tendency of the changes of the field intensity factors with the electric displacement is similar as that of the electric field in Case 2.

4 Conclusions

The normalized dynamic field intensity factors increase with increase of the ratio of crack radius to the strip thickness. For the case of constant stress loading, the dynamic electric field intensity factor and the dynamic electric displacement intensity factor depend on the material constants and the applied mechanical load, but not on the applied electrical load. For the case of constant strain loading, the dynamic field intensity factors depend on the applied mechanical and electrical loads.

References

[1] Gao, C. F., and Fan, W. X., 1999, ''A General Solution for the Plane Problem in Piezoelectric Media With Collinear Cracks,'' Int. J. Eng. Sci., **37**, pp. 347–363.

[2] Yang, J. H., and Lee, K. Y., 2001, ''Penny Shaped Crack in Three Dimensional Piezoelectric Strip Under In-Plane Normal Loadings,'' Acta Mech., **148**, pp. 187–197.

[3] Wang, Z. K., and Huang, S. H., 1995, ''Fields Near Elliptical Crack Tip in Piezoelectric Ceramics,'' Eng. Fract. Mech., **51**, pp. 447–456.

Analysis of the *M*-Integral in Plane Elasticity

Y. Z. Chen
Division of Engineering Mechanics, Jiangsu University, Zhenjiang, Jiangsu 212013, P. R. China

Kang Yong Lee
Department of Mechanical Engineering, Yonsei University, Seoul 120-749, South Korea

In this paper, analysis of the M-integral in plane elasticity is carried out. An infinite plate with any number of inclusions and cracks and with any applied forces and remote tractions is considered. To study the problem, the mutual work difference integral (abbreviated as MWDI) is introduced, which is defined by the difference of works done by each other stress field on a large circle. The concept of the derivative stress field is also introduced, which is a real elasticity solution and is derived from the physical stress field. It is found that the M-integral on a large circle is equal to a MWDI from the physical stress field and a derivative stress field. Finally, the expression for M-integral on a large circle is obtained. The variation for the M-integral with respect to the coordinate transformation is addressed. An illustrative example for the use of M-integral is presented. [DOI: 10.1115/1.1748271]

1 Introduction

In plane elasticity, many path independent integrals were investigated, [1–10]. The integrals include the *J*-integrals, *L*-integral, *M*-integral, and some others. These integrals have a general property that the values of the mentioned integrals do not depend on the path, provided there is no singularity between two integration paths. Naturally, if the closed path encloses some singularity points or cavities, these integrals must not vanish.

Some particular problems for the *M*-integral were investigated, [11–13]. It was felt that in some complicated cases solutions for the *M*-integral were still unknown. In this paper, analysis of the *M*-integral in plane elasticity is carried out. An infinite plate with any number of inclusions and cracks and with any applied forces and remote tractions is considered. After some manipulations, the expression for *M*-integral on a large circle is obtained. The variation for the *M*-integral with respect to the coordinate transformation is addressed. An illustrative example for the use of *M*-integral is present.

2 Evaluation of *M*-integral on a Large Circle

The *M*-integral in plane elasticity is a path independent integral, which was introduced previously ([3])

$$M(L)=\int_{(x_0,y_0),(L)}^{(x,y)}(Wx_in_i-u_{i,k}x_k\sigma_{ij}n_j)ds \tag{1}$$

$$M(CH)=\oint_{(CH)}(Wx_in_i-u_{i,k}x_k\sigma_{ij}n_j)ds \tag{2}$$

where $W=u_{i,j}\sigma_{ij}/2$ denotes the strain energy density, u_i the displacements, σ_{ij} the stresses, n_j the direction cosines, and $u_{i,j}=\partial u_i/\partial x_j$. In Eq. (1), the path "$L$" is generally defined as a path with the starting point (x_o,y_o) and the end point (x, y) (Fig. 1). In Eq. (2), a closed path "CH" is defined such that it may encloses some holes, cracks or inclusions (Fig. 1).

The following analysis depends on the complex variable function method in plane elasticity, [14]. In the method, the stresses $(\sigma_x,\sigma_y,\sigma_{xy})$, the resultant forces (X, Y) and the displacements (u,v) are expressed in terms of two complex potentials $\phi(z)$ and $\psi(z)$ such that

$$\sigma_x+\sigma_y=4\,\mathrm{Re}\,\phi'(z)$$

$$\sigma_y-\sigma_x+2i\sigma_{xy}=2[\bar{z}\phi''(z)+\psi'(z)] \tag{3}$$

$$f=-Y+iX=\phi(z)+z\overline{\phi'(z)}+\overline{\psi(z)} \tag{4}$$

$$2G(u+iv)=\kappa\phi(z)-z\overline{\phi'(z)}-\overline{\psi(z)} \tag{5}$$

where G is the shear modulus of elasticity, $\kappa=(3-\nu)/(1+\nu)$ is for the plane stress problem, $\kappa=3-4\nu$ is for the plane strain problem, and ν is the Poisson's ratio.

In the following analysis, two stress fields are introduced. The first field is the physical field, which is defined from the geometry and the loading condition shown in Fig. 1, and it is called the α-field hereafter. Clearly, for the α-field we can write

$$u_{i(\alpha)}=u_i,\quad \sigma_{ij(\alpha)}=\sigma_{ij}. \tag{6}$$

Meantime, the relevant complex potentials for the α-field can be expressed in the form ([14])

$$\phi_{(\alpha)}(z)=\phi(z)=A_1z+A_2\log z+a_0+\sum_{k=1}^{\infty}\frac{a_k}{z^k} \tag{7}$$

$$\psi_{(\alpha)}(z)=\psi(z)=B_1z+B_2\log z+b_0+\sum_{k=1}^{\infty}\frac{b_k}{z^k} \tag{8}$$

where

$$A_1=\frac{\sigma_x^\infty+\sigma_y^\infty}{4},\quad A_2=-\frac{F_x+iF_y}{2\pi(\kappa+1)} \tag{9}$$

$$B_1=\frac{\sigma_y^\infty-\sigma_x^\infty}{2}+i\sigma_{xy}^\infty,\quad B_2=-\kappa\overline{A_2}=\frac{\kappa(F_x-iF_y)}{2\pi(\kappa+1)}. \tag{10}$$

In Eqs. (9) and (10), σ_x^∞, σ_y^∞, and σ_{xy}^∞ are the remote stresses, and F_x and F_y are the resultant forces applied on the finite region of the infinite plate. Also, the coefficients a_k and b_k $(k=1,2,..)$ in Eqs. (9) and (10) will be determined from a concrete solution.

The second field is defined as a derivative field to the physical field, which is called the β-field hereafter, [7,15,16]. The complex potentials for the β-field is defined such that

$$\phi_{(\beta)}(z)=z\phi'(z)=A_1z+A_2-\sum_{k=1}^{\infty}\frac{ka_k}{z^k} \tag{11}$$

$$\psi_{(\beta)}(z)=z\psi'(z)=B_1z+B_2-\sum_{k=1}^{\infty}\frac{kb_k}{z^k}. \tag{12}$$

Note that the β-field is not independent and is derived from the α-field. Thus, by using Eqs. (3) and (5) the displacements and stresses for the β-field can be expressed as

$$u_{i(\beta)}=x\frac{\partial u_i}{\partial x}+y\frac{\partial u_i}{\partial y}\quad(\text{or}\quad u_{i(\beta)}=u_{i,j}x_j) \tag{13}$$

$$\sigma_{ij(\beta)}=\sigma_{ij}+x\frac{\partial\sigma_{ij}}{\partial x}+y\frac{\partial\sigma_{ij}}{\partial y}\quad(\text{or}\quad \sigma_{ij(\beta)}=\sigma_{ij}+\sigma_{ij,k}x_k). \tag{14}$$

In addition, one may introduce the following integral (Fig. 1):

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 12, 2001; final revision, December 13, 2003. Associate Editor: B. M. Moran.

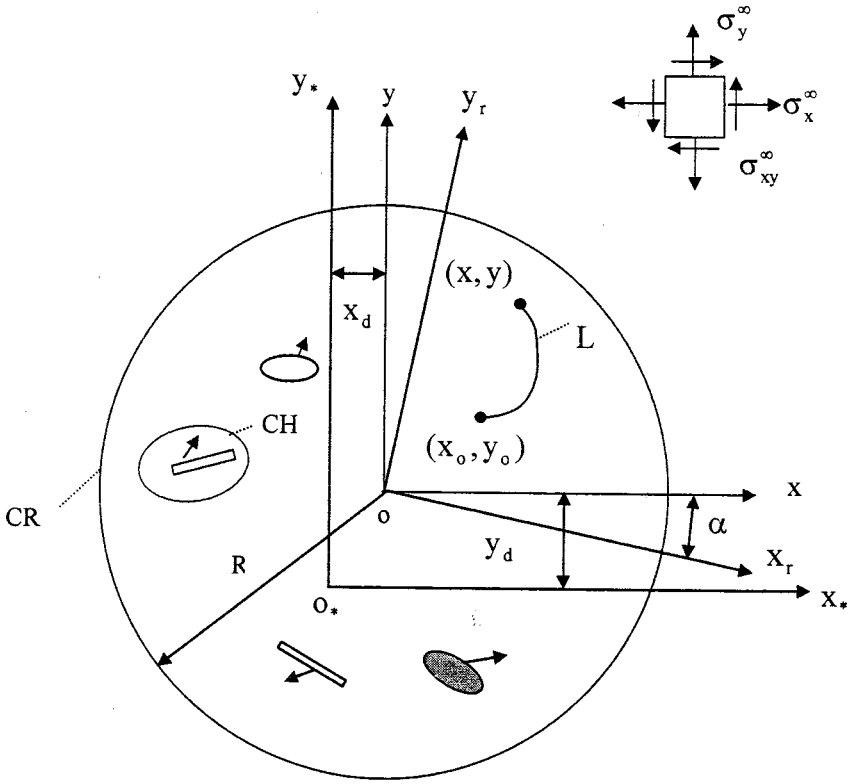


Fig. 1 An infinite plate containing cracks, holes, and inclusions

$$N(CH)=\oint_{(CH)}\frac{1}{2}(u_{i(\alpha)}\sigma_{ij(\beta)}-u_{i(\beta)}\sigma_{ij(\alpha)})n_j ds$$
$$=\oint_{(CH)}\frac{1}{2}(u_i(\sigma_{ij}+\sigma_{ij,k}x_k)n_j-u_{i,k}x_k\sigma_{ij}n_j)ds. \tag{15}$$

From the Betti's reciprocal theorem in elasticity, one can conclude that the integral $N(CH)$ takes the same value if the closed path "*CH*" includes the same defects. Clearly, $N(CH)$ shown by Eq. (15) is a kind of a mutual work difference integral (*MWDI*) for two physical stress fields (α-field and β-field).

A relation between $M(CH)$-integral and the $N(CH)$-integral has been found, and it reads

$$M(CH)=N(CH). \tag{16}$$

The equality (16) can be proved in a similar manner as shown previously, [15,16].

Since the equality (16) is proved, and the closed path "*CR*" is a particular type of "*CH*" (Fig. 1), thus, we have

$$M(CR)=N(CR) \tag{17}$$

where

$$M(CR)=\oint_{(CR)}(Wx_in_i-u_{i,k}x_k\sigma_{ij}n_j)ds \tag{18}$$

$$N(CR)=\oint_{(CR)}\frac{1}{2}(u_{i(\alpha)}\sigma_{ij(\beta)}-u_{i(\beta)}\sigma_{ij(\alpha)})n_j ds$$
$$=\oint_{(CR)}\frac{1}{2}(u_i(\sigma_{ij}+\sigma_{ij,k}x_k)n_j-u_{i,k}x_k\sigma_{ij}n_j)ds. \tag{19}$$

In Eqs. (18) and (19) the circle "*CR*" is sufficiently large such that all the defects are included in the circle. Equation (17) shows that instead of evaluating the integral $M(CR)$ we can evaluate its equivalent value $N(CR)$. This is the aim of the mentioned derivation.

Since the complex potentials for the α-field and β-field were shown by Eqs. (7), (8), (11), and (12), the relevant *MWDI* shown by Eq. (19) can be evaluated. Using a known result, [15,16], we have

$$M(CR)=N(CR)=\frac{\pi(\kappa+1)}{G}\mathrm{Re}[-A_1b_1-B_1a_1+A_2B_2]. \tag{20}$$

In Eq. (20), A_1, A_2, a_1 (or B_1, B_2, b_1) have the dimensions of stresses, resultant forces and moments, respectively.

3 Relation of the $M(CR)$ Integral With the Coordinate Transformation

The $M(CR)$ shown by Eq. (18) is evaluated in the *xoy*-coordinates (Fig. 1). Consider now how the *M*-integral corresponding to a given state of stress for the infinite plate changes under transformation from one system of rectangular coordinates to another. Let (x, y) and (x_*, y_*) be the coordinates of the same point in the (xoy) and $(x_* o_* y_*)$ systems and let

$$z_*=z+z_d \quad (\text{with } z_*=x_*+iy_*,\ z=x+iy,\ z_d=x_d+iy_d) \tag{21}$$

where z_d represents a translation of the coordinate system (Fig. 1).

In addition, the complex potentials $\phi_*(z_*)$, $\psi_*(z_*)$ in the $(x_* o_* y_*)$ system can be obtained from the known complex potentials $\phi(z)$, $\psi(z)$ in the (xoy) system, [14]. Using the obtained complex potentials $\phi_*(z_*)$, $\psi_*(z_*)$ and Eq. (20), it is found

$$M_*(CR)=M(CR)+M_a \tag{22}$$

where

$$M_a=\frac{1}{8G}\{[(\kappa+1)\sigma_x^\infty+(\kappa-3)\sigma_y^\infty]F_x x_d+[(\kappa+1)\sigma_y^\infty +(\kappa-3)\sigma_x^\infty]F_y y_d+4\sigma_{xy}^\infty(F_y x_d+F_x y_d)\}. \quad (23)$$

In Eq. (23) the term M_a is caused by the translation of the coordinates. From Eqs. (22) and (23) we see that only if the resultant forces are equal to zero ($F_x=F_y=0$) the $M(CR)$ is an invariant with respect to the translation of coordinates.

Secondly, consider now how the *M*-integral changes under rotational transformation from one system of rectangular coordinates to another (Fig. 1). Let (x, y) and (x_r, y_r) be the coordinates of the same point in the (xoy) and $(x_r o y_r)$ systems and let

$$z_r=z\exp(i\alpha) \quad (\text{with } z_r=x_r+iy_r,\ z=x+iy,) \quad (24)$$

where the angle α represents a rotation of the coordinate system (Fig. 1).

A similar derivation will give the following result ([14]):

$$M_{(r)}(CR)=M(CR)=\frac{\pi(\kappa+1)}{G}\mathrm{Re}[-A_1 b_1-B_1 a_1+A_2 B_2]. \quad (25)$$

In Eq. (25), $M_{(r)}(CR)$ denotes the value of *M*-integral evaluated in the (x_r, y_r) coordinates. Also, from Eq. (25) we see that the $M(CR)$ value is an invariant with respect to the rotation of the coordinates.

4 Discussion

It is found that the value of $M(CR)$ under the remote loading $\sigma_x^\infty=\sigma_y^\infty=p$ is a suitable measure to evaluate the included damage. An example is introduced below. In the example, an elliptic hole in an infinite plate is subjected to the remote loading $\sigma_x^\infty=\sigma_y^\infty=p$. Three cases: (a) a circle hole with a radius a_o, (b) a elliptical hole with major and minor axes "*a*" and "*b*," and (c) a crack with the half-length radius a_{cr}, are considered for comparison. After some manipulation, for the three cases we have

$$M(CR)_{(a)}=\frac{\pi(\kappa+1)}{8G}[4a_o^2p^2] \quad (26)$$

$$M(CR)_{(b)}=\frac{\pi(\kappa+1)}{8G}\left[\frac{4(1+m^2)}{(1+m)^2}a^2p^2\right], \quad (m=(a-b)/(a+b)) \quad (27)$$

$$M(CR)_{(c)}=\frac{\pi(\kappa+1)}{8G}[2a_{cr}^2p^2]. \quad (28)$$

It is assumed that, for two damaged mediums, if the values of $M(CR)$ are the same they are said to be in the equivalent damage situation. Therefore, after letting $M(CR)_{(a)}=M(CR)_{(b)}=M(CR)_{(c)}$, it is found that

$$a=\frac{1+m}{\sqrt{1+m^2}}a_o, \quad a_{cr}=\sqrt{2}a_o. \quad (29)$$

Equation (29) reveals that a crack with half-length a_{cr} $(=\sqrt{2}a_o)$ is equivalent to a circle with radius a_o on the basis of the above-mentioned assumption. Similarly, an elliptical notch with major and minor axes a and b $(a=((1+m)a_o)/\sqrt{1+m^2}, b=((1-m)a)/(1+m))$ is equivalent to a circle with radius a_o.

Acknowledgments

The project was supported by the National Natural Science Foundation of China for the first author (Y. Z. Chen), and by the Brain Pool Program in Korea for the second author (K. Y. Lee).

References

[1] Rice, J. R., 1968, "A Path-Independent Integral and the Approximation Analysis of Strain Concentration by Notches and Cracks," ASME J. Appl. Mech., **35**, pp. 379–386.

[2] Knowles, J. K., and Sternberg, E., 1972, "On a Class of Conservation Laws in Linearized and Finite Elastostatics," Arch. Ration. Mech. Anal., **44**, pp. 187–211.

[3] Budiansky, B., and Rice, J. R., 1973, "Conservation Laws and Energy-Release Rates," ASME J. Appl. Mech., **40**, pp. 201–203.

[4] Fruend, L. B., 1978, "Stress Intensity Factor Calculation Based on a Conservation Integral," Int. J. Solids Struct., **14**, pp. 241–250.

[5] Herrmann, A. G., and Hermann, G., 1981, "On Energy-Release Rates for a Plane Crack," ASME J. Appl. Mech., **48**, pp. 525–528.

[6] Bui, H. D., 1974, "Dual Path Independent Integrals in the Boundary-Value Problems of Cracks," Eng. Fract. Mech., **6**, pp. 287–296.

[7] Chen, Y. Z., 1985, "New Path Independent Integrals in Linear Elastic Fracture Mechanics," Eng. Fract. Mech., **22**, pp. 673–686.

[8] Cherepanov, G. P., 1979, *Mechanics of Brittle Fracture*, McGraw-Hill, New York.

[9] Kanninen, M. F., and Popelar, C. H., 1985, *Advanced Fracture Mechanics*, Oxford University Press, Oxford, UK.

[10] Eshelby, J. D., 1951, "The Force on an Elastic Singularity," Philos. Trans. R. Soc. London, Ser. A, **A244**, pp. 87–112.

[11] Chen, Y. Z., 1985, "A Technique for Evaluating the Stress Intensity Factors by Means of the M-Integral," Eng. Fract. Mech., **23**, pp. 777–780.

[12] Suo, Z., 2000, "Zener's Crack and the M-integral," ASME J. Appl. Mech., **67**, pp. 417–418.

[13] Chen, Y. H., 2001, "M-Integral Analysis for Two Dimensional Solids With Strongly Interacting Microcracks. Part I: In an Infinite Brittle Solid," Int. J. Solids Struct., **38**, pp. 3193–3212.

[14] Muskhelishvili, N. I., 1953, *Some Basic Problems of Mathematical Theory of Elasticity*, Noordhoff, Netherlands.

[15] Chen, Y. Z., and Lee, K. Y., 2002, "Some Properties of J-Integral in Plane Elasticity," ASME J. Appl. Mech., **69**, pp. 195–198.

[16] Chen, Y. Z., 2003, "Analysis of L-Integral and Theory of the Derivative Stress Field in Plane Elasticity," Int. J. Solids Struct., **40**, pp. 3589–3602.

Impermeable Crack and Permeable Crack Assumptions, Which One is More Realistic?

Bao-Lin Wang
e-mail: baolin.wang@aeromech.usyd.edu.au

Yiu-Wing Mai

Centre for Advanced Materials Technology (CAMT),
Department of Mechanical and Mechatronic Engineering,
Mechanical Engineering Building J07, The
University of Sydney, Sydney, NSW 2006, Australia

This paper investigates the applicability and effect of the crack-free electrical boundary conditions in piezoelectric fracture. By treating flaws in a medium as notches with a finite width, the results from different electrical boundary condition assumptions on the crack faces are compared. It is found that the electrically impermeable boundary is a reasonable one for engineering problems. Unless the flaw interior is filled with conductive media, the permeable crack assumption may not be directly applied to the fracture of piezoelectric materials in engineering applications.
[DOI: 10.1115/1.1748294]

1 Introduction

Due to the rapid growth in applications, the mechanical and fracture properties of piezoelectric ceramics are becoming more and more important. The material is brittle. In order to address the issues concerning durability and reliability of piezoelectric materials, fracture behaviors of those materials should be investigated and understood thoroughly. The effect of electromechanical field on the fracture behavior of piezoelectric ceramics is highly complex. The understanding of the cracking under mechanical and electrical loads is of both academic and practical importance.

An important issue in studying fracture mechanics of piezoelectric materials is the crack-face electric boundary conditions. There are two idealized crack-face boundary conditions that are extensively used in the literatures. One commonly used boundary condition is the specification that the normal component of electric displacement along the crack faces equals to zero (see, e.g., [1–11]). This boundary condition ignores the permittivity in the medium interior to the crack. The other commonly used boundary condition treats the crack as being electrically permeable, [12–15]. For permeable crack, the electric field and the electric displacement intensity factors are found to depend only on the applied mechanical loads. The stress intensity factor and the energy release rate become independent of the electrical loads. This is in contrast to the test results, [8]. Other approach towards resolving the piezoelectric crack problems considers the deformation of the crack, [9,16–18]. This approach is based on the assumption that under applied mechanical and electrical loads, the crack will open and there is an electrical potential difference on the upper surface and the lower surface of the crack.

In this paper, we reconsider the usually used two electrical boundary condition assumptions on the crack faces. Considering the fact that flaws in experiments are not like cleavage cracks of zero gap width, we treat the crack as a notch with a finite thickness. The results from different electrical boundary condition assumptions on the crack faces are compared. Some conclusions are made.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, June 4, 2002; final revision, September 18, 2003. Associate Editor: K. Ravi-Chandar.

2 Piezoelectric Crack Problem

We consider a plane-strain problem shown in Fig. 1. Assume that all field variables are functions of x and y only. Constitutive equations for piezoelectric materials polarized along y direction subjected to mechanical and electrical fields can be written as

$$(\sigma_{xx}\ \ \sigma_{yy}\ \ \tau_{xy}\ \ D_x\ \ D_y)^T=\mathbf{C}(\varepsilon_{xx}\ \ \varepsilon_{yy}\ \ 2\varepsilon_{xy}\ \ E_x\ \ E_y), \quad (1)$$

where

$$\mathbf{C}=\begin{bmatrix} c_{11} & c_{13} & 0 & 0 & -e_{31} \\ c_{13} & c_{33} & 0 & 0 & -e_{33} \\ 0 & 0 & c_{44} & -e_{15} & 0 \\ 0 & 0 & e_{15} & \epsilon_{11} & 0 \\ e_{31} & e_{33} & 0 & 0 & \epsilon_{33} \end{bmatrix}, \quad (2)$$

σ_{ij} and D_i are stresses and electric displacements, respectively; c_{ij}, e_{ij}, and ϵ_{ii} are elastic constants, piezoelectric constants, and dielectric permittivities, respectively. The strain ε_{ij} is related to the mechanical displacement u_i by $\varepsilon_{ij}=(u_{i,j}+u_{j,i})/2$, where a comma indicates partial derivative. The electric field E_i is related to the electric potential ϕ by $E_i=-\phi_{,i}$.

The governing equations for displacements and electric potential are

$$\left.\begin{aligned} c_{11}\frac{\partial^2 u}{\partial x^2}+c_{44}\frac{\partial^2 u}{\partial y^2}+(c_{13}+c_{44})\frac{\partial^2 v}{\partial x\partial y}+(e_{31}+e_{15})\frac{\partial^2 \phi}{\partial x\partial y}=0 \\ (c_{13}+c_{44})\frac{\partial^2 u}{\partial x\partial y}+c_{44}\frac{\partial^2 v}{\partial x^2}+c_{33}\frac{\partial^2 v}{\partial y^2}+e_{15}\frac{\partial^2 \phi}{\partial x^2}+e_{33}\frac{\partial^2 \phi}{\partial y^2}=0 \\ (e_{31}+e_{15})\frac{\partial^2 u}{\partial x\partial y}+e_{15}\frac{\partial^2 v}{\partial x^2}+e_{33}\frac{\partial^2 v}{\partial y^2}-\epsilon_{11}\frac{\partial^2 \phi}{\partial x^2}-\epsilon_{33}\frac{\partial^2 \phi}{\partial y^2}=0 \end{aligned}\right\}, \quad (3)$$

where u and v are displacement in x and y direction, respectively.

Assume that the piezoelectric medium is loaded by a remote uniform stresses $\sigma_y=\sigma_\infty$ and a uniform electric displacement $D_y=D_\infty$. Denote $\boldsymbol{D}_0$ as the normal component of the electric displacement inside the crack ($\boldsymbol{D}_0$ is unknown). The boundary conditions along the cracked plane $y=0$ is stated as follows:

$$v(x,0)=0, \quad \phi(x,0)=0, \quad |x|\geqslant a, \quad (4)$$

$$\sigma_{yy}(x,0)=0, \quad |x|\leqslant a, \quad (5)$$

$$D_y(x,0)=\boldsymbol{D}_0, \quad |x|\leqslant a. \quad (6)$$

For this problem, the solutions can be obtained by means of complex variables as outlines in [5],

$$(u,\ v,\ \phi)^T=2\ \mathrm{Re}(\mathbf{AF}), \quad (7)$$

$$(\sigma_{xx},\ D_x)^T=2\ \mathrm{Re}(\mathbf{B}_1\mathbf{F}), \quad (8)$$

$$(\sigma_{xy},\ \sigma_{yy},\ D_y)^T=2\ \mathrm{Re}(\mathbf{BF}), \quad (9)$$

where $\mathbf{A}=[A_{j\alpha}]$ and $\mathbf{B}=[B_{j\alpha}]$ are (3×3) matrices, $\mathbf{B}_1$ is a (2×3) matrix,

$$\mathbf{F}=[F_\alpha(z_\alpha)]=(F_1(z_1),\ F_2(z_2),\ F_3(z_3))^T, \quad (10)$$

$$\mathbf{f}=[f_\alpha(z_\alpha)]=(F_1'(z_1),\ F_2'(z_2),\ F_3'(z_3))^T, \quad (11)$$

$$z_\alpha=x+p_\alpha y, \quad \alpha=1,2,3. \quad (12)$$

The vectors $\mathbf{f}$ and $\mathbf{F}$ are determined by

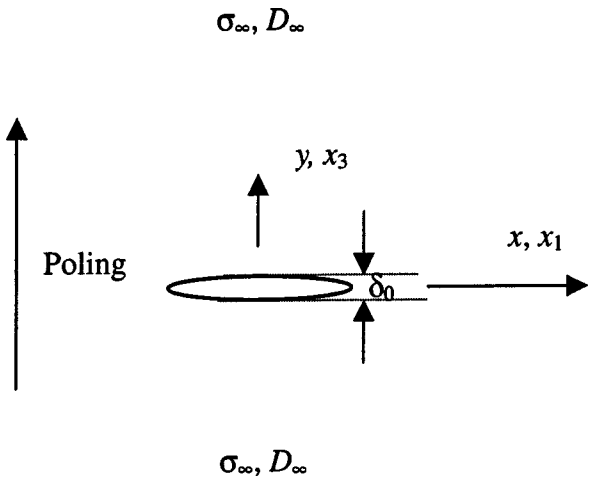


Fig. 1 A piezoelectric medium with a notch of finite thickness

$$f_\alpha(z_\alpha)=\mathbf{H}_\alpha\mathbf{t}_0\left(\frac{z_\alpha}{\sqrt{z_\alpha^2-a^2}}-1\right), \tag{13}$$

$$F_\alpha(z_\alpha)=\mathbf{H}_\alpha\mathbf{t}_0(\sqrt{z_\alpha^2-a^2}-z_\alpha), \tag{14}$$

in which $\mathbf{H}$ is a (3×2) matrix, and $\mathbf{H}_\alpha$ is the αth row of matrix $\mathbf{H}$,

$$\mathbf{t}_0=(\sigma_\infty,\ D_\infty-\boldsymbol{D}_0). \tag{15}$$

Since the vector $\mathbf{t}_0$ contains an unknown $\boldsymbol{D}_0$, additional assumptions are needed to solve Eq. (13). This will be discussed in Secs. 3 and 4 of this paper. Once $f_\alpha(z_\alpha)$ and $F_\alpha(z_\alpha)$ are evaluated, the full field solution of the cracked piezoelectric material can be determined. In particular, the electric potential jump $\Delta\phi$ across the crack are obtained as follows:

$$\Delta v=-4\sqrt{a^2-x_1^2}\,\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{2\alpha}\mathbf{H}_\alpha\mathbf{t}_0\right), \tag{16}$$

$$\Delta\phi=-4\sqrt{a^2-x_1^2}\,\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}\mathbf{H}_\alpha\mathbf{t}_0\right). \tag{17}$$

In order to obtain the stress and electric displacement distributions near the crack front, we introducing the polar coordinates r and θ such that

$$x=a+r\cos\theta,\quad y=r\sin\theta, \tag{18}$$

then the function $f_\alpha(z_\alpha)$ may be expanded for small values of r:

$$f_\alpha(z_\alpha)=\mathbf{H}_\alpha\mathbf{t}_0\frac{\sqrt{a}}{\sqrt{2r}}\frac{1}{\sqrt{\cos\theta+p_\alpha\sin\theta}}. \tag{19}$$

Using the conventionally defined stress intensity factors and electric displacement intensity factor, stress and electric displacement at the crack tip in the crack plane, $\theta=0$, are given by

$$(\sigma_{yy},\ D_y)=\frac{(K_\mathrm{I},\ K_\mathrm{IV})}{\sqrt{2r}}, \tag{20}$$

where K_I is the mode I stress intensity factor, K_IV is the electric displacement intensity factor. They take the following values

$$(K_\mathrm{I},\ K_\mathrm{IV})=(\sigma_\infty,\ D_\infty-\boldsymbol{D}_0)\sqrt{a}. \tag{21}$$

It is clear that K_I depends only on the applied stress. The electrical boundary conditions on the crack surfaces have no effect on the stress intensity factors.

From Eq. (8), the stress vector σ_{xx} and D_x at the crack tip takes in $\theta=0$ plane takes the following values:

$$(\sigma_{xx},\ D_x)^T=2\,\mathrm{Re}(\mathbf{B}_1\mathbf{H})\mathbf{t}_0\frac{\sqrt{a}}{\sqrt{2r}}. \tag{22}$$

The energy release rate can be calculated by the crack closure integral to yield

$$G=-\pi a\,\mathrm{Im}\left[\left(\sigma_\infty\sum_{\alpha=1}^{3}A_{2\alpha}+(D_\infty-\boldsymbol{D}_0)\sum_{\alpha=1}^{3}A_{3\alpha}\right)\mathbf{H}_\alpha\mathbf{t}_0\right]. \tag{23}$$

3 Solutions Based on Electrically Impermeable and Electrically Permeable Crack Assumptions

Conventionally, there are two kinds of electrical boundary condition assumptions used to the fracture of piezoelectric ceramics, i.e., the impermeable crack assumption and the permeable crack assumption. The simplest way to solve the problem is that the crack is assumed to be electrical impermeable and the electric field inside the crack is zero. Therefore, in Eq. (15) $\boldsymbol{D}_0=0$, f_α and F_α can be determined immediately from Eqs. (13) and (14).

On the other hand, the permeable crack assumption treats the crack as being electrically permeable and there is no electric potential jump across the crack. It can be shown from Eq. (17) that

$$\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}\mathbf{H}_\alpha\right)\mathbf{t}_0=0. \tag{24}$$

Denote $H_{\alpha j}$ as the αth row and the jth column element of matrix $\mathbf{H}$. The unknown electric displacement $\boldsymbol{D}_0$ inside the crack can be determined from Eq. (24). The result is

$$\boldsymbol{D}_0=D_\infty+\frac{\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha1}\right)\sigma_\infty}{\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha2}\right)}. \tag{25}$$

Once $\boldsymbol{D}_0$ is determined, Eqs. (13) and (14) can be used to evaluate the unknown constants f_α and F_α. The electric displacement intensity factor is

$$K_\mathrm{IV}=\frac{-\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha1}\right)\sigma_\infty}{\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha2}\right)}\sqrt{a}. \tag{26}$$

It can be shown from Eqs. (13) to (15) that the full field solution depends on the electrical load only through the term $D_\infty-\boldsymbol{D}_0$. It is clear from Eq. (25) that $D_\infty-\boldsymbol{D}_0$ can be expressed in terms of applied mechanical loads. Therefore, solutions for an electrically permeable crack depend only on the mechanical loads, but not on the applied electrical load.

4 A Notch of Finite Thickness

In the last section, we considered a flaw in the piezoelectric media as a cleavage crack of zero gap width. In fracture mechanics experiments, however, a flaw in a specimen is usually cut with a tool of finite thickness. Thus, flaws in the experiments are not like cleavage cracks of zero gap width, but rather like notches with a finite width.

Suppose the thickness δ of the notch is sufficiently small, except near the tip of the notch, the gradient of crack opening along the crack is small. Along the y direction the electric field $\boldsymbol{E}$ and the electric displacement $\boldsymbol{D}_0$ on the upper notch surface can be written as ([19])

$$\boldsymbol{E}(x)=-\frac{\Delta\phi(x,0)}{\delta(x)},\quad \boldsymbol{D}_0(x)=-\epsilon_0\frac{\Delta\phi(x,0)}{\delta(x)},\tag{27}$$

where ϵ_0 is the dielectric permittivity of the medium inside the notch, which is usually the air or vacuum. By substituting from Eq. (17) into Eq. (27), we obtain

$$\boldsymbol{D}_0(x)=4\sqrt{a^2-x_1^2}\,\epsilon_0 \times\frac{\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha1}\right)\sigma_\infty+\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha2}\right)(D_\infty-\boldsymbol{D}_0)}{\delta(x)}.\tag{28}$$

This equation can be used to determine the unknown $\boldsymbol{D}_0$. If we consider an elliptic notch such that

$$\delta(x)=\sqrt{a^2-x_1^2}(\delta_0/a),\tag{29}$$

from Eqs. (28) and (29), it follows that the electric displacement inside the notch is

$$\boldsymbol{D}_0=\frac{\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha1}\right)\sigma_\infty+\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha2}\right)D_\infty}{\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha2}\right)+(\delta_0/a)/(4\epsilon_0)}.\tag{30}$$

It follows from (21) that the electric displacement intensity factor at the notch tip is

$$K_{\mathrm{IV}}=\frac{(\delta_0/a)(4\epsilon_0)D_\infty-\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha1}\right)\sigma_\infty}{\mathrm{Im}\left(\sum_{\alpha=1}^{3}A_{3\alpha}H_{\alpha2}\right)+(\delta_0/a)/(4\epsilon_0)}\sqrt{a}.\tag{31}$$

It is clear that if the notch interior is filled by conductive medium such that ϵ_0 equals infinity, Eqs. (30) and (31) reduce to the permeable crack solution, Eqs. (25) and (26), respectively. On the other hand, If the permittivity of the notch is ignored such that ϵ_0 equals zero, Eqs. (30) and (31) reduce, respectively, to

$$\boldsymbol{D}_0=0,\quad \text{and}\quad K_{\mathrm{IV}}=D_\infty\sqrt{a},\tag{32}$$

which are the impermeable crack solutions.

5 Applications

A PZT-4 piezoelectric ceramic loaded by a remote stress $\sigma_y=\sigma_\infty$ and an electric displacement load $D_y=D_\infty$ is considered as an example. The properties of PZT-4 piezoelectric ceramic are same as those given in [8]. The dielectric permittivity of the air or vacuum is $\epsilon_0=0.0885\times10^{-10}$ C/Vm.

As discussed above, the stress intensity factor K_{I} rely the applied mechanical load only. The notch thickness and the crack face electrical boundary condition assumption have no effect on K_{I}. The electric intensity factor for an impermeable crack and a permeable crack are

$$K_{\mathrm{IV}}=D_\infty\sqrt{a},$$

and

$$K_{\mathrm{IV}}=2.534\times10^{-10}\sigma_0\sqrt{a},$$

respectively.

On the other hand, if the flaw is a notch of finite thickness rather than a slit crack, the electric field $\boldsymbol{D}_0$ inside the flaw is

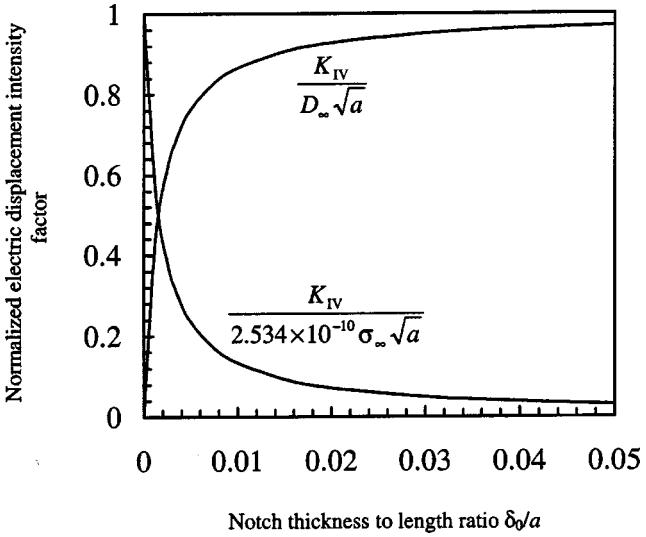


Fig. 2 Electric displacement intensity factors caused by a uniform electric displacement load D_∞ or a uniform mechanical load σ_∞

$$\boldsymbol{D}_0=\frac{-0.01108\times10^{-10}\sigma_\infty+0.004372D_\infty}{0.004372+2.825(\delta_0/a)}.$$

The electric displacement intensity factor at the notch tip is

$$K_{\mathrm{IV}}=\frac{0.01108\times10^{-10}\sigma_\infty+2.825(\delta_0/a)D_\infty}{0.004372+2.825(\delta_0/a)}\sqrt{a}.$$

The dependence of the electric displacement intensity factor on the notch thickness to length ratio is plotted in Fig. 2, for an electric displacement load or a mechanical load. The gap width is known from the numerical computations to have a major influence on the electric field intensity factors. Therefore, the effect of the finite flaw gap in a realistic structure must be assessed. Since it is impossible to make a notch of zero width so that it behaviors as a slit crack, the permeable crack assumption may not be directly applied to the piezoelectric fracture problems. The idealization of an electrically impermeable boundary is a reasonable one for notch aspect ratios of 0.01 or greater. Some authors studied this issue and reached a similar conclusion, [9,12].

It is informative to consider some experimental date. Park and Sun [8] used a 0.46 mm thick diamond wheel to cut a flaw of 11.5 mm length in a PZT-4 piezoelectric ceramic. The notch thickness to length ratio in their test is

$$\delta_0/a\approx0.46/(11.5/2)=0.08.$$

The normalized electric displacement intensity factor is

$$\frac{K_{\mathrm{IV}}}{2.534\times10^{-10}\sigma_\infty\sqrt{a}}=0.0190$$

for a pure mechanical load, and

$$\frac{K_{\mathrm{IV}}}{D_\infty\sqrt{a}}=0.981$$

for a pure electric displacement load. The results are very near to those obtained from impermeable crack assumption (zero for mechanical load, and one for electrical displacement load). However, for the permeable crack model, the normalized electric displacement intensity is one (for mechanical load) or zero (for electrical displacement load). Therefore, the permeable crack assumption gives a completely erroneous result.

6 The Antiplane Crack Problems

If the piezoelectric medium is poled perpendicular to *x-y* plane, and subjected to an antiplane shear stress τ_∞ and an inplane electric displacement D_∞, according to the method outlined by Pak [20], it can be easily shown that the electric potential on the upper crack surface is

$$\phi = \frac{e_{15}\tau_\infty - c_{44}(D_\infty - \boldsymbol{D}_0)}{c_{44}\epsilon_{11} + e_{15}^2}\sqrt{a^2 - x^2}. \tag{33}$$

If the thickness of the notch is expressed by (29), then the electric displacement inside the crack is obtained as

$$\boldsymbol{D}_0 = \frac{c_{44}D_\infty - e_{15}\tau_\infty}{c_{44} + (c_{44}\epsilon_{11} + e_{15}^2)(\delta_0/a)/(2\epsilon_0)}. \tag{34}$$

The electric displacement intensity factor is obtained in closed form:

$$K_{\mathrm{IV}} = \left[D_\infty - \frac{c_{44}D_\infty - e_{15}\tau_\infty}{c_{44} + (c_{44}\epsilon_{11} + e_{15}^2)(\delta_0/a)/(2\epsilon_0)} \right]\sqrt{a}. \tag{35}$$

The interpretations of the results for an antiplane crack problem are similar to those for an inplane crack problem.

In conclusion, we know that

1. the application of an electric field does not change the stress intensity factors.
2. the ratio of notch thickness to length has no influence on the stress intensity factors.
3. the electric displacement intensity factor depends strongly on the ratio of notch thickness to length.
4. since a flaw in engineering materials is always a notch of finite thickness rather than a slit crack, the electrically impermeable boundary is a reasonable one for engineering problems.
5. unless the flaw interior is filled with conductive media, the permeable crack assumption may not be directly applied to the fracture of piezoelectric materials in engineering applications. Here, a crack is defined as a notch of zero thickness.

References

[1] Deeg, W., 1980, ''The Analysis of Dislocation, Crack and Inclusion in Piezoelectric Solids,'' Ph.D. thesis, Stanford University, Stanford, CA.

[2] Sosa, H. A., and Pak, Y. E., 1990, ''Three-Dimensional Eigenfunction Analysis of a Crack in a Piezoelectric Material,'' Int. J. Solids Struct., **26**, pp. 1–15.

[3] Sosa, H. A., 1991, ''Plane Problems in Piezoelectric Media With Defects,'' Int. J. Solids Struct., **28**, pp. 491–505.

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

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

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

[7] Park, S. B., and Sun, C. T., 1995, ''Effect of Electric Field on Fracture of Piezoelectric Ceramics,'' Int. J. Fract., **70**, pp. 203–216.

[8] Park, S. B., Sun, C. T., 1995, ''Fracture Criteria for Piezoelectric Ceramics,'' J. Am. Ceram. Soc., **78**, pp. 1475–1480.

[9] Fulton, C. C., and Gao, H., 2001, ''Effect of Local Polarization Switching on Piezoelectric Fracture,'' J. Mech. Phys. Solids, **49**, pp. 927–952.

[10] Gao, H., and Barnett, D. M., 1996, ''An Invariance Property of Local Energy Release Rate in a Strip Saturation Model of Piezoelectric Fracture,'' Int. J. Fract., **79**, pp. R25–R29.

[11] Gao, H., Zhang, T.-Y., and Tong, P., 1997, ''Local and Global Energy Release Rates for an Electrically Yielded Crack in a Piezoelectric Ceramic,'' J. Mech. Phys. Solids, **45**, pp. 491–510.

[12] Dunn, M. L., 1994, ''The Effects of Crack Face Boundary Conditions on the Fracture of Piezoelectric Solids,'' Eng. Fract. Mech., **48**, pp. 25–39.

[13] Shindo, Y., Ozawa, E., and Nowacki, J. P., 1990, ''Singular Stress and Electric Fields of a Cracked Piezoelectric Strip,'' Int. J. Appl. Electromagn. Mater., **1**, pp. 77–87.

[14] Shindo, Y., Watanabe, K., and Narita, F., 2000, ''Electroelastic Analysis of a Piezoelectric Ceramic Strip With a Central Crack,'' Int. J. Eng. Sci., **38**, pp. 1–19.

[15] Mikahailov, G. K., and Parton, V. Z., 1990, *Electromagnetoelasticity*, Hemisphere, Washington, DC.

[16] Hao, T. H., and Shen, Z. Y., 1994, ''A New Electric Boundary Condition of Electric Fracture Mechanics and Its Applications,'' Eng. Fract. Mech., **47**, pp. 793–802.

[17] McMeeking, R. M., 2001, ''Towards a Fracture Mechanics for Brittle Piezoelectric and Dielectric Materials,'' Int. J. Fract., **108**, pp. 25–41.

[18] Zhang, T. Y., Qian, C. F., Tong, P., 1998, ''Linear Electro-Elastic Analysis of a Cavity or a Crack in a Piezoelectric Material,'' Int. J. Solids Struct., **35**, pp. 2122–2149.

[19] McMeeking, R. M., 1999, ''Crack Tip Energy Release Rate for a Piezoelectric Compact Tension Specimen,'' Eng. Fract. Mech., **64**, pp. 217–244.

[20] Pak, Y. E., 1990, ''Crack Extension Force in a Piezoelectric Material,'' ASME J. Appl. Mech., **57**, pp. 647–653.

Onset of Degenerate Hopf Bifurcation of a Vibro-Impact Oscillator

GuiLin Wen
Mem. ASME,
School of Mechanical and Production Engineering,
Nanyang Technological University, Singapore 639798
and Department of Applied Mechanics,
Southwest Jiaotong University, Chengdu 610031,
P.R. China

JianHua Xie
Department of Applied Mechanics, Southwest Jiaotong University, Chengdu 610031, P.R. China

Daolin Xu
School of Mechanical and Production Engineering,
Nanyang Technological University, Singapore 639798

An analytical method of the degenerate Hopf bifurcation is proposed for vibro-impact systems. The phenomenon of the bifurcation and its complicated dynamics are observed. This type of bifurcation originates multi-coexisting solutions dependent of the initial state of the system. [DOI: 10.1115/1.1767163]

A two-degree-of-freedom impact oscillator is shown in Fig. 1. The mass M_1 with forced excitation impacts against a rigid wall A when its displacement X_1 reaches to the gap B. The impact causes the discontinuity of velocity of the impacting mass by grazing incidence thus inducing a variety of complicated dynamics. The vibration of the oscillator consists of the nonimpact motions separated by impacts. The non-impact motion between two successive impacts ($X_1<B$) can be expressed by the linear differential Eqs. (1*a*) in a nondimensional form. For the impact motion ($X_1=B$), the relationship between the velocities before and after an impact is described by a coefficient of restitution R in (1*b*).

$$\begin{bmatrix} 1 & 0 \\ 0 & \mu_m \end{bmatrix} \begin{Bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{Bmatrix} + \begin{bmatrix} 1 & -1 \\ -1 & 1+\mu_k \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} = \begin{Bmatrix} \sin(\omega t+\tau) \\ 0 \end{Bmatrix}, \quad (x_1<b) \tag{1a}$$

$$\dot{x}_{1+} = -R\dot{x}_{1-}, \quad (x_1=b) \tag{1b}$$

where $\dot{x}_{1+}$ and $\dot{x}_{1-}$ represent the velocities of mass M_1 before and after an impact, respectively, and τ is the phase angle. The nondimensional quantities are defined as $\mu_m=M_2/M_1$, $\mu_k=K_2/K_1$, $\omega=\Omega\sqrt{M_1/K_1}$, $t=T\sqrt{K_1/M_1}$, $b=BK_1/P_1$, $x_i=X_iK_1/P_1$, $\dot{x}_{1+}=\dot{X}_{1+}K_1/P_1$, $\dot{x}_{1-}=\dot{X}_{1-}K_1/P_1$.

The vibro-impact oscillator (1) is a piecewise linear system that may exhibit very rich dynamical behaviors. Quasi-periodic motions and its torus-doubling bifurcations were observed, [1]. In strong resonance cases, the system may directly bifurcate into unstable 3-3-periodic impact motions or stable 4-4-periodic impact motion, [2], from a Hopf bifurcation. With an additional proportional damping of the Rayleigh type, a verity of periodic and chaotic behaviors was reported, [3]. The grazing bifurcation resulted from the piecewise properties and singularities of the impact were also investigated, [4], in a perfectly plastic vibro-impact case.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 4, 2002, final revision, January 27, 2004. Associate Editor: A. A. Ferri.

In this brief note, the dynamical phenomenon of a class of degenerate Hopf bifurcation is reported for the vibro-impact oscillator (1). Degenerate Hopf bifurcation may lead to two-coexisting quasi-periodic impact motions, usually observed in continuous or discrete-time systems, [5]. This dynamical behavior is interesting since it allows changing the system behavior qualitatively without altering system bifurcation parameters. An analytical method is proposed for analyzing the occurrence of the degenerate Hopf bifurcation in vibro-impact systems. To cope with the discontinuity of velocity of impact that gives rise to difficulties in solving the dynamics of the differential equations (1), we establish a four-dimensional map in a Poincaré section for the vibro-impact system. Consequently, the center manifold reduction, [6], and the normal form technique, [7,8], can be employed to reduce the Poincaré map into a two-dimensional normal form. Thus, Chenciner's theory, [5], of degenerate Hopf bifurcation of maps in R^2 can be applied to theoretically describe the phenomenon of the degenerate Hopf bifurcation in the vibro-impact system (1).

To establish the Poincaré map of the vibro-impact system (1), we first look into the general solution of the piecewise linear system, between two successive impacts, expressed as

$$x_i = \sum_{j=1}^{2} \psi_{ij}(a_j \cos\omega_j t + b_j \sin\omega_j t + A_j \sin(\omega t+\tau)), \tag{2a}$$

and

$$\dot{x}_i = \sum_{j=1}^{2} \psi_{ij}(-a_j\omega_j \sin\omega_j t + b_j\omega_j \cos\omega_j t + A_j\omega \cos(\omega t+\tau))$$

$$(i=1,2) \tag{2b}$$

where ψ_{ij} are the elements of the canonical modal matrix ψ of Eq. (1*a*), ω_j the eigenfrequencies of the system, a_j and b_j the integration constants, A_j the amplitude parameters, and τ the phase angle. These coefficients can be determined by the modal parameters and the initial conditions: $x_1(0)=b$, $x_1(2\pi/\omega)=b$, $\dot{x}_{1+}(0)=-R\dot{x}_{1-}(2\pi/\omega)$, $x_2(0)=x_2(2\pi/\omega)$, $\dot{x}_2(0)=\dot{x}_2(2\pi/\omega)$. According to the general solution (2) and its initial conditions, the existence of periodic impacts must satisfy with the following condition;

$$\left| \frac{b\tan\tau_0 \pm \sqrt{(\tan^2\tau_0+1)d^2-b^2}}{(\tan^2\tau_0+1)d} \right| \leqslant 1 \tag{3}$$

where $d=-(\psi_{11}A_1+\psi_{12}A_2)$. Let $s_i=\sin 2\pi\omega_i/\omega$, $c_i=\cos 2\pi\omega_i/\omega$. The initial phase angle τ_0 satisfies

$$\tau_0=\tilde{\tau}_0 \quad \text{if} \quad b=0; \tag{4a}$$

$$\text{or} \quad \tau_0=\cos^{-1}\left(\frac{b\tan\tilde{\tau}_0 \pm \sqrt{(\tan^2\tilde{\tau}_0+1)d^2-b^2}}{(\tan^2\tilde{\tau}_0+1)d} \right) \quad \text{if} \quad b\neq 0. \tag{4b}$$

where $\tilde{\tau}_0=\tan^{-1}((\psi_{22}\psi_{11}s_1(1-c_2)\omega_2-\psi_{12}\psi_{21}s_2(1-c_1)\omega_1)(1+R)\omega/|\psi|\omega_1\omega_2(1-c_1)(1-c_2)(1-R))$. The impact response of the system (1) can be described as follows. Given the state $(x_1,\dot{x}_1,x_2,\dot{x}_2,\tau)$ of the system at the instant set as $t=0$, the motion of the system is determined by the general solution (2) of Eq. (1*a*) until the next impact. When impact occurs, for $x_1=b$, the velocity of the impacting mass is changed by the impact law, $\dot{x}_{1+}=-R\dot{x}_{1-}$. After an impact, shift the phase $\tau=\omega t+\tau$, and reset the time $t=0$, and renew the initial conditions in Eq. (1). With the new initial state, the motion of the system (1) may be described by the general solution (2) until the next impact. In this way, the motion of the vibro-impact system can be traced over an arbitrary number of impacts.

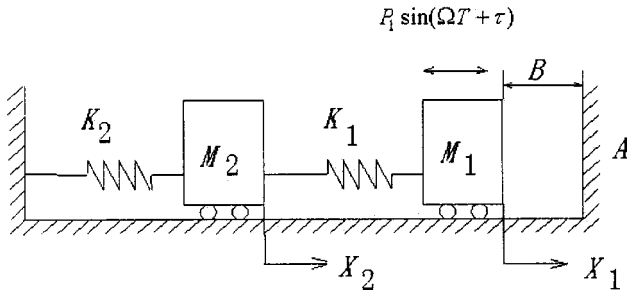


Fig. 1 A two-degree-of-freedom impact oscillator

Choosing a Poincaré section, [1], $\sigma \subset R^4 \times S$, where $\sigma = \{(x_1, \dot{x}_1, x_2, \dot{x}_2, \theta) \in R^4 \times S, x_1 = b, \dot{x}_1 = \dot{x}_{1+}\}$ and $\theta = \omega t$, we can establish a four-dimensional Poincaré map for the impact system (1). The map can be expressed in the brief form

$$X' = \tilde{f}(\omega; X) \tag{5}$$

where the state vector $X = (\dot{x}_{1+}, x_2, \dot{x}_2, \tau)^T$ and ω is chosen as the bifurcation parameter. Let X^* denotes a fixed point of Poincaré map (5) in the hyperplane σ, which corresponds to the 1-1 periodic impact motion of the system (1), [1,3]. Supposed that the degenerate Hopf bifurcation of the system (1) occurs at (X^*, ω_c), the disturbed system is given as

$$\Delta X' = f(\mu; \Delta X) \tag{6}$$

where ΔX is a disturbed vector at fixed point X^* and $\mu = \omega_c - \omega$. The Jacobian matrix $\mathbf{D}f_{\Delta X}(0;0)$ of map (6) satisfies the following conditions for a Hopf bifurcation, [1]:

(C1) $f(\mu;0) = 0$ for all μ;

(C2) $\mathbf{D}f_{\Delta X}(\mu;0)$ has a pair complex conjugate eigenvalues $\lambda_1(\mu)$, $\lambda_1(\mu)$ satisfying $|\lambda_1(0)| = 1$, and the other eigenvalues $\lambda_3(\mu)$, $\lambda_4(\mu)$ satisfying $|\lambda_{3,4}(0)| < 1$;

(C3) $\lambda_1^m(0) \neq 1$, $m = 1,2,3,4,5,6,7$;

(C4) $\mathrm{d}|\lambda_1(\mu)|/\mathrm{d}\mu|_{\mu=0} = \mathrm{Re}\,\tilde{\lambda}_1 > 0$ where $\lambda_1(\mu) = \lambda_1(0)(1 + \mu\tilde{\lambda}_1 + o(|\mu|))$.

The conditions C1, C2, and C4 are the critical conditions for the degenerate Hopf bifurcation and C3 is the nonresonance condition. Note that for the transversality condition C4, if $\mathrm{d}|\lambda_1(\mu)|/\mathrm{d}\mu|_{\mu=0} < 0$ by setting $\mu = \omega_c - \omega$, then $\mathrm{d}|\lambda_1(\mu)|/\mathrm{d}\mu|_{\mu=0} > 0$ by setting $\mu = \omega - \omega_c$.

From C1 and C2, a local center manifold, [1,6], of map (6) exists. The local dynamic behavior of the four-dimensional map, (6), can be reduced into a two-dimensional map,

$$F(\mu, z) = \lambda_1(\mu) z + \sum_{i+j=2}^{5} g_{ij}(\mu) z^i \bar{z}^j + O(|z|^5), \tag{7}$$

where the coefficients $g_{ij}(\mu)$ can be expressed by the coefficients of map (6), see [1]. With the nonresonance condition C3, through a smooth μ-dependent change of the coordinate, [7,8], we can obtain the following normal form of map (7),

$$P_\mu(\mu, z) = \lambda_1(\mu) z + c_3(\mu) z^2 \bar{z} + c_5(\mu) z^3 \bar{z}^2 + O(|z|^6) \tag{8}$$

where the exact form of the coefficients $c_3(\mu)$ and $c_5(\mu)$ in terms of $g_{ij}(\mu)$ of map (7) can be found in [8].

Let $\alpha(\mu) = \mathrm{Re}(c_3(\mu)\overline{\lambda_1(\mu)})$, $\beta(\mu) = \mathrm{Re}(c_5(\mu)\overline{\lambda_1(\mu)})$ and $\tilde{\Delta} = \alpha(\mu)^2 - 4\mu\beta(\mu)\mathrm{Re}\,\tilde{\lambda}_1$. The Chenciner's theory, [5], of degenerate Hopf bifurcations, stated in the following Lemma, can be used to check the existence of the degenerate Hopf bifurcation of the impact system (1).

Lemma, [5]. Under the transversality condition C4, a degenerate Hopf bifurcation of map (6) occurs if $\alpha(\mu) = 0$ at $\mu = 0$. Four types of solutions are illustrated in Fig. 2.

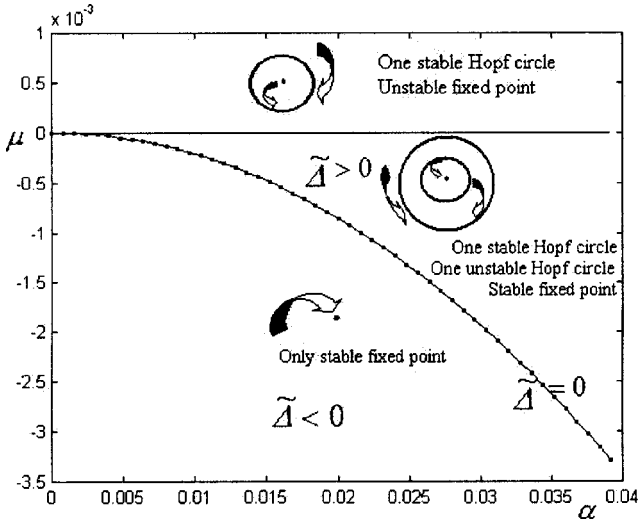


Fig. 2 Degenerate Hopf bifurcation diagram against the parameters (α, μ)

(i) In the parameter domain Θ where $\mu < 0$, $\alpha(\mu) > 0$, $\beta(\mu) < 0$ and $\tilde{\Delta} > 0$, there exist a fixed point X^* and two Hopf circles. The unstable circle separates the stable circle from the stable fixed point.

(ii) If $\mu < 0$ and $\tilde{\Delta} = 0$, as shown in Fig. 2, the coalescence phenomenon of the stable and unstable circles may occur.

(iii) If $\mu < 0$ and $\tilde{\Delta} < 0$, see Fig. 2, only a stable fixed point exists.

(iv) If $\mu > 0$, as shown in Fig. 2, the fixed point is unstable associated with a stable bifurcated Hopf circle.

Note that the proposed method is a general approach applicable to piecewise linear systems. It is possible to extend the method to nonlinear impact systems if their nonimpact solutions are available, because from the solutions we are able to establish a Poincaré map (5) based on which center manifold, [6], and Chenciner's theory, [5], can be utilized in the study of degenerate Hopf

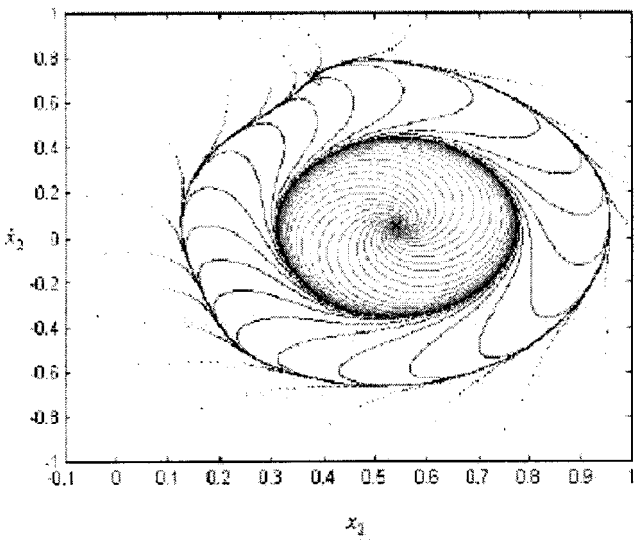


Fig. 3 Coexisting Hopf circles and a fixed point resulted from a degenerate Hopf bifurcation at $\omega = 0.7297$ ($\mu < 0$) and $\alpha = 0.0313$, where the unstable circle separates the stable circle from the stable fixed point. The symbol '×' denotes the location of the last iteration of the Poincaré map (5). A view on the projected section ($x_2, \dot{x}_2$).

bifurcation. However, for most of nonlinear systems, analytical solutions do not exist. Thus analytical approach to such systems is infeasible.

We now study the dynamical phenomena induced by the degenerate Hopf bifurcation with the aid of the analytical method proposed. Set $\mu_m=3.644$, $\mu_k=7.167787$, $b=2.037673$, $R=0.5365$ for the system (1). The critical bifurcation parameter is $\omega_c=0.72871$ in the system (1). A fixed point of map (5) is $X^*=(-0.661343,0.536408,0.0458349,1.449541)^{\mathrm{T}}$ that is a period 1-1 impact motion of the system (1). The eigenvalues of the Jacobian matrix of map (6) are $\lambda_{1,2}=0.57377\pm i0.81901$, $\lambda_{3,4}=-0.41743\pm i0.33629$, respectively, where $|\lambda_{1,2}|=1$ and $d|\lambda_1(\mu)|/d\mu|_{\mu=0}=0.22475$. With the center manifold, [6], and normal form reduction, [8], we find that $\alpha(\mu)=0$ at $\mu=0$. Thus the degenerate Hopf bifurcation occurs.

The complicated dynamics resulted from the degenerate Hopf bifurcation is summarized below. (i) Given $\mu=-0.001$ (i.e., $\omega=0.7297$) such that $\alpha(\mu)=0.03125$, $\beta(\mu)=-0.50869$ and $\tilde{\Delta}>0$, Fig. 3 shows the co-existence of three solutions of map (5). The unstable Hopf circle (inner circle) separates the stable fixed point (center point) form the stable Hopf circle (outer circle). Any initial state point starting from the regime outside the stable circle or the regime between the two circles will converge to the stable Hopf circle which corresponding to the stable quasi-periodic impact motion of the system (1). When an initial state point starts from the regime surrounded by the inner circle, the iteration of the Poincaré map eventually converges to the fixed point X^*, a 1-1 periodic impact motion. (ii) The coalescence phenomenon of the stable and unstable circles occurs at $\omega=0.73086$ ($\mu<0$ and $\tilde{\Delta}=0$). (iii) Taking $\omega=0.7318$ ($\mu<0$ and $\tilde{\Delta}<0$), all the circles vanish with the emergence of a single stable fixed point X^*. (iv) Given $\omega=0.7262$ ($\mu>0$), the bifurcated Hopf circle of the Poincaré map is attracting.

In summary, an analytical method for analyzing degenerate Hopf bifurcation is introduced to piecewise linear systems. The method could be extended to nonlinear impact systems if its non-impact solution is available. The phenomenon of multi-coexisting motions is observed from the degenerate Hopf bifurcation in a specific piecewise linear system, the vibro-impact system (1). Different motions take place dependent on the initial state of the system. This characteristic allows one to alter the system dynamics qualitatively without varying a system parameter.

References

[1] Luo, G. W., and Xie, J. H., 1998, "Hopf Bifurcation of a Two-Degree-of-freedom Vibrato Impact System," J. Sound Vib., **213**(3), pp. 391–408.

[2] Luo, G. W., and Xie, J. H., 2002, "Hopf Bifurcation and Chaos of a Two-Degree-of-Freedom Vibro-Impact System in Two Strong Resonance Cases," Int. J. Non-Linear Mech., **37**, pp. 19–34.

[3] Aidanpaa, J. O., and Gupta, R. B., 1992, "Periodic and Chaotic Behavior of a Threshold-Limited Two-Degree-of-Freedom System," J. Sound Vib., **165**(2), pp. 305–327.

[4] Luo, G. W., Xie, J. H., and Guo, S. H. L., 2001, "Periodic Motions and Global Bifurcations of a Two-degree-of-Freedom System With Plastic Vibro-Impact," J. Sound Vib., **240**(5), pp. 837–858.

[5] Chenciner, A., 1985, "Bifurcations de Points Fixes Elliptiques I-Courbes Invariantes," IHES Pub. Math., **61**, pp. 67–127.

[6] Carr, J., 1981, *Applications of Center Manifold Theory* (Applied Mathematical Sciences 35), Springer-Verlag, New York, pp. 33–36.

[7] Iooss, G., 1979, *Bifurcation of Maps and Applications* (Mathematics Studies 36), North-Holland, Amsterdam.

[8] Wen, G. L., and Xu, D., 2003, "Control of Degenerate Hopf Bifurcations in Three-Dimensional Maps," Chaos, **13**(2), pp. 486–494.

A Long Crack Penetrating a Transforming Inhomogeneity

Yuping Wang and Roberto Ballarini
Department of Civil Engineering, Case Western Reserve University, Cleveland, OH 44106-7201

This note presents the stress intensity factors of a long crack penetrating a circular transforming inhomogeneity. Using the Greens functions of dislocations interacting with a circular inhomogeneity experiencing an isotropic (free expansion) eigenstrain, the elasticity solution is reduced to a system of singular integral equations representing the traction boundary condition along the crack surfaces. The normalized stress intensity factor, obtained through a numerical solution of the integral equations, has a strong dependence on the elastic mismatch, and can be either negative or positive depending on the crack-tip location. The formulation and results generalize a previously published transformation-toughening model that assigns equal elastic moduli to the inhomogeneity and the surrounding medium.
[DOI: 10.1115/1.1767166]

Analysis

Consider the plane elastostatics problem shown in Fig. 1(*a*). A circular inhomogeneity with radius a Poisson's ratio ν_2, and shear modulus μ_2, is embedded in an infinite plate with Poisson's ratio ν_1 and shear modulus μ_1. A semi-infinite crack penetrates the inhomogeneity, which is experiencing an isotropic (free expansion) eigenstrain, $\varepsilon_{ij}^{*}=\delta_{ij}e^{*}$, where the ε_{ij} are the components of the strain tensor and δ_{ij} is the Kronecker delta. The bonding between the inhomogeneity and the surrounding matrix is perfect. As shown in Fig. 1(*a*), the origin of the coordinate system is located at the center of the inhomogeneity, and the crack tip is located at point $(w,0)$. The stress intensity factor produced by the eigenstrain within the inhomogeneity is defined as K_I^{loc}.

The solution is formulated as the superposition of two problems, as shown schematically in Figs. 1(*b*) and 1(*c*). The first involves the stresses produced along the crack line in an uncracked plate containing the expanding inhomogeneity (Fig. 1(*b*)), and the second the stresses produced along the crack line by a continuous distribution of dislocations (Fig. 1(*c*)). The stresses produced by the eigenstrain are

$$\sigma_{yy}^{1e}=\frac{4(1+\eta)\mu_1\mu_2 e^{*}}{\mu_1(\kappa_2-1)+2\mu_2}\left\{\frac{a^2}{x^2}\right\}\quad x\leqslant a \tag{1a}$$

$$\sigma_{yy}^{2e}=\frac{4(1+\eta)\mu_1\mu_2 e^{*}}{\mu_1(\kappa_2-1)+2\mu_2}\quad x\geqslant a \tag{1b}$$

while those resulting from the distribution of dislocations are represented as

$$\sigma_{yy}^{1d}=\int_0^{\infty}\frac{2b_1(t)}{1-x}dt+\int_a^{\infty}K_{11}(x,t)b_1(t)dt+\int_w^{a}K_{12}(x,t)b_2(t)dt\quad x\leqslant a \tag{2a}$$

$$\sigma_{yy}^{2d}=\int_w^{a}\frac{2b_2(t)}{t-x}dt+\int_w^{a}K_{21}(x,t)b_2(t)dt+\int_a^{\infty}K_{22}(x,t)b_1(t)dt\quad w\leqslant x\leqslant a. \tag{2b}$$

In Eq. (2), $b_i(t)=\mu_i/\pi(\kappa_i+1)\partial[\nu_y]/\partial t$ is defined as the dislocation density in region i, $[\nu_y]$ is the crack-opening displacement, $\kappa=3-4\nu$ and $\eta=\nu$ for plane strain, $\kappa=(3-\nu)/(1+\nu)$ and $\eta=0$ for plane stress, and the K_{ij} are combinations of regular and generalized Cauchy kernels that can be recovered from Ref. [1]. The zero-traction condition along the crack line is enforced by summing to zero the stress contributions from Eq. (1) and (2).

The asymptotic behavior of $b_i(t)$ was studied in detail in Ref. [1], where the loading was associated with a far-field stress consistent with a nominal stress intensity factor, rather than with an expanding inhomogeneity. Note that if one is interested in calculating the stress intensity factor produced by a far-field loading interacting with the eigenstrain within the inhomogeneity, then an appropriate superposition procedure must be performed. The dislocation densities $b_i(t)$ can be expressed as follows:

$$b_1(t)=\frac{g_1(t)}{(t-w)^{0.5-\mu}(t-a)^{\mu}} \tag{3a}$$

$$b_2(t)=\frac{g_2(t)}{(t-w)^{0.5}(a-t)^{\mu}} \tag{3b}$$

where the dominant singularity at the interface, μ, and the unknown regular functions $g_i(t)$ satisfy

$$(1-\beta^2)(1+\cos^2\mu\pi)+2[2\alpha\beta-1-(2\alpha\beta-\beta^2)\cos\mu\pi]+4\mu(2-\mu)[(\alpha-\beta)^2(1-\mu)^2-\alpha\beta+\beta(\alpha-\beta)\cos\mu\pi]=0 \tag{4a}$$

$$g_1(t)=0\ \text{ at }\ t=\infty \tag{4b}$$

$$g_2(a)/g_1(a)=\frac{(1+\alpha)\beta+(\alpha-\beta)(1-\beta)(-1+4\mu-2\mu^2)-(1-\beta^2)\cos(\mu\pi)}{(1+\alpha)(-1+2\beta-2\beta\mu)}\times\frac{\mu_2(\kappa_1+1)}{\mu_1(\kappa_2+1)}\times(a-w)^{\mu}. \tag{4c}$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Sept. 25, 2002; final revision, Feb. 6, 2004. Associate Editor: K. Ravi-Chandar.

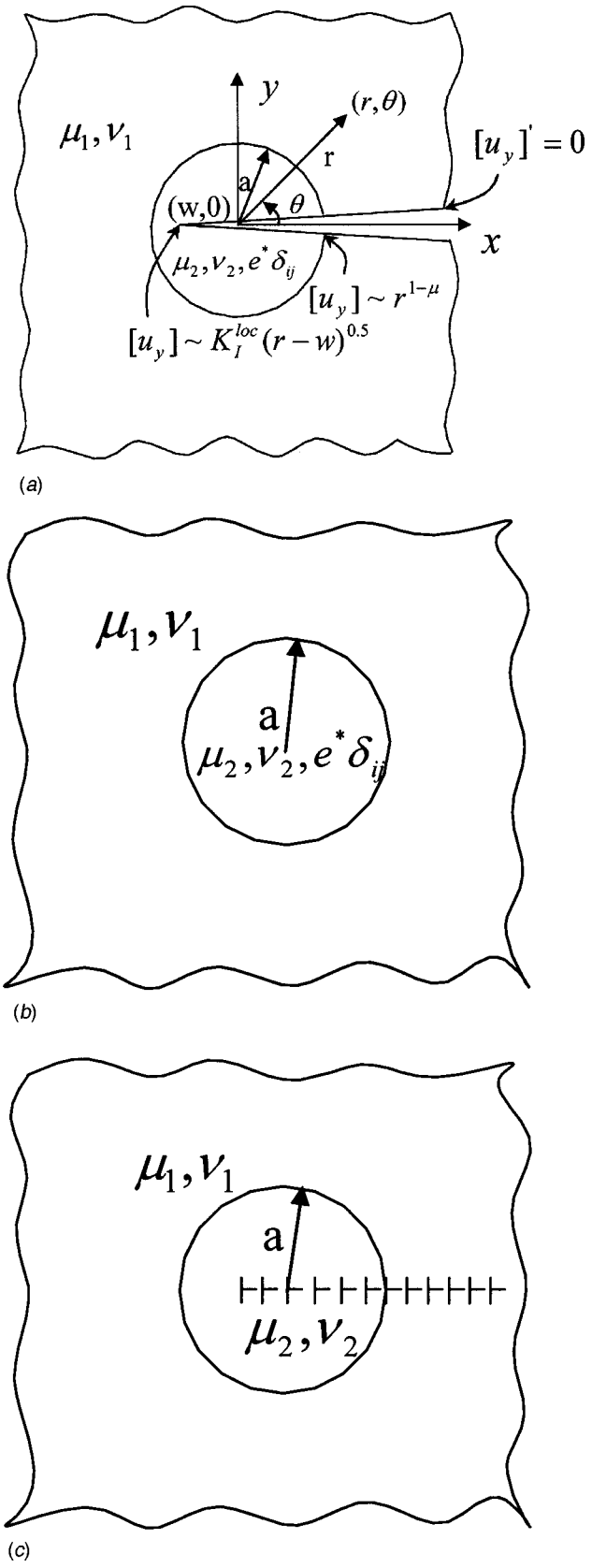


Fig. 1 (*a*) A semi-infinite crack penetrating a circular inhomogeneity which is experiencing an isotropic eigenstrain; (*b*) an uncracked infinite plane containing a transforming circular inhomogeneity; (*c*) an infinite plane containing a continuous distribution of edge dislocations

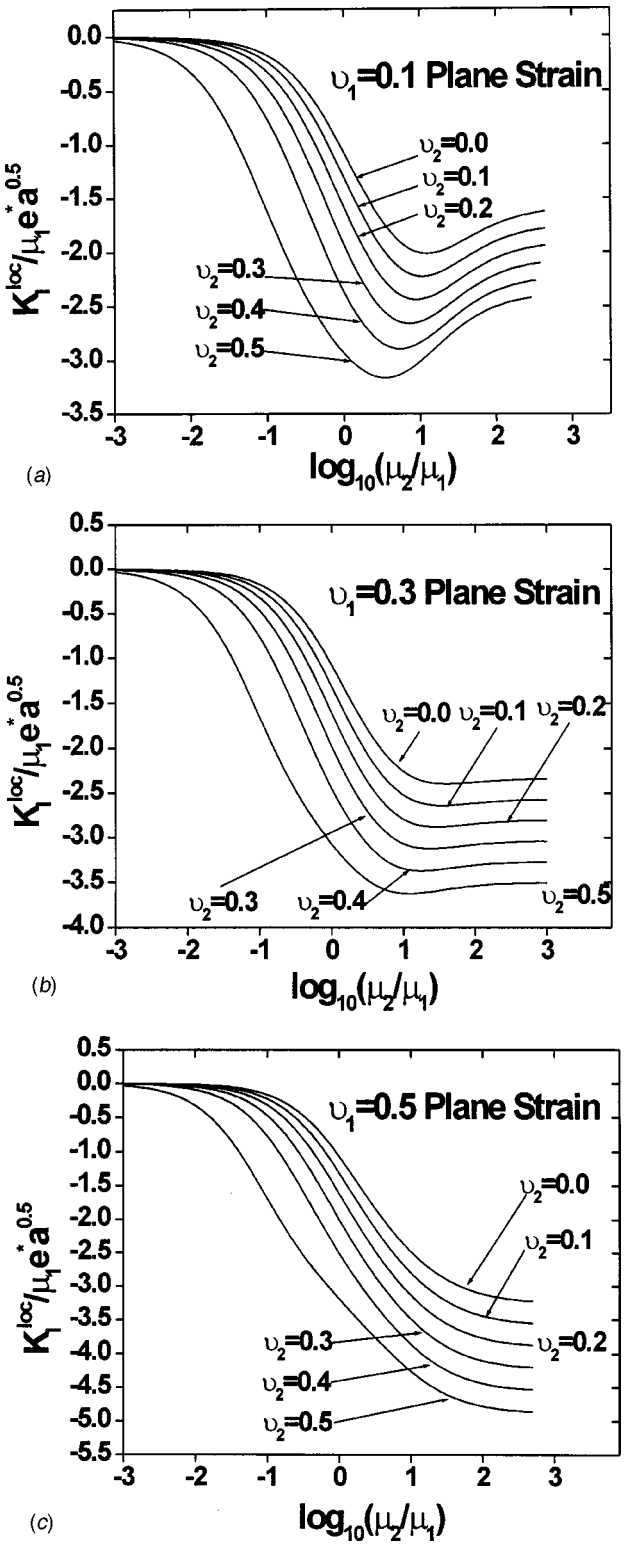
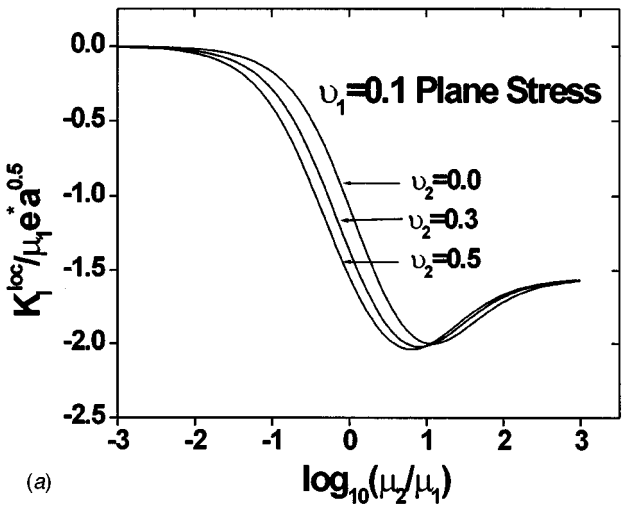
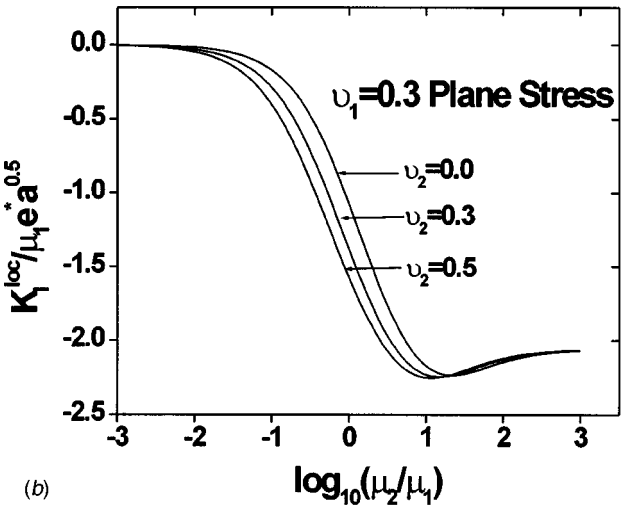


Fig. 2 Plane strain nondimensional stress intensity factor as functions of shear modulus ratio, μ_2/μ_1, for several combinations of Poisson's ratios ν_1, ν_2

Using the numerical approach developed by Erdogan et al. [2], which relies on the properties of Jocobi polynomials, the values of $g_i(t)$ are calculated at discrete points and the stress intensity factor is recovered as $K_I^{\text{loc}} = 2\pi\sqrt{2\pi}/(a-w)^{\mu} g_2(w)$. It should be noted that because the integral equations are not homogeneous, no stabilization procedure is required to calculate a unique solution.





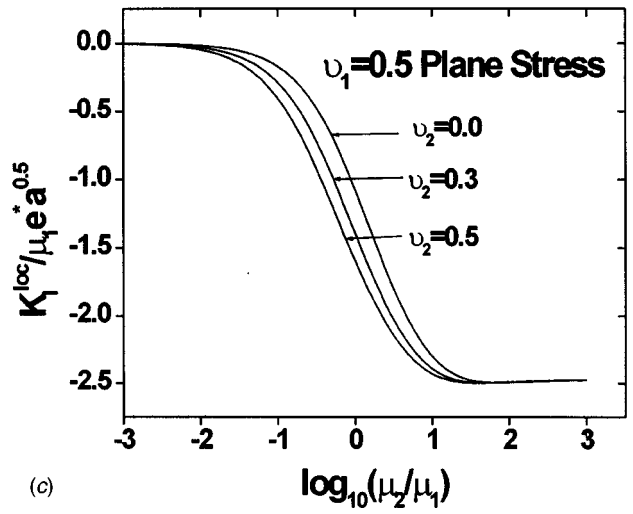
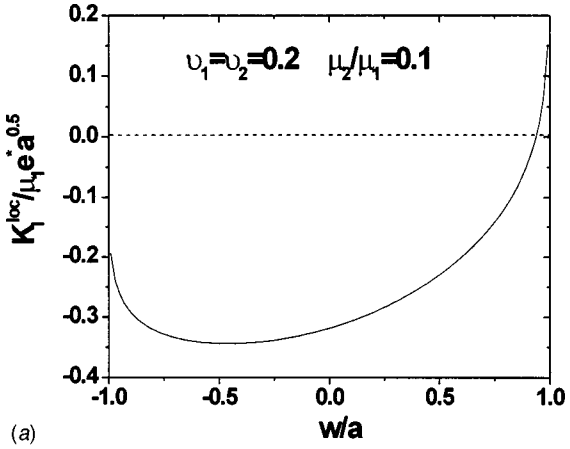
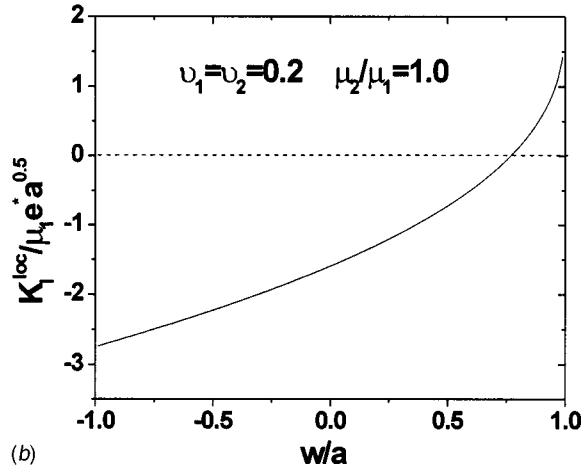


Fig. 3 Plane stress nondimensional stress intensity factor as functions of shear modulus ratio, μ_2/μ_1, for several combinations of Poisson's ratios ν_1, ν_2

However, if the loading is associated with a far-field stress intensity factor rather than an eigenstrain within the inhomogeneity, [1], then the integral equations become homogeneous, and a stabilization procedure is required for a unique solution, [3].





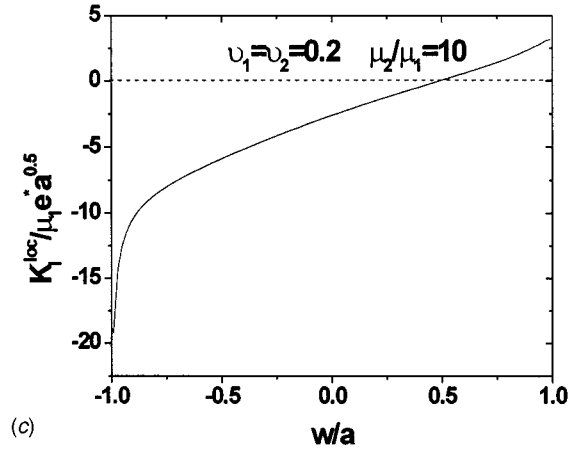


Fig. 4 Variation of the plane strain nondimensional stress intensity factor with crack-tip position, for several combinations of elastic mismatch

Results

The nondimensional stress intensity factor is defined as

$$\frac{K_I^{\text{loc}}}{\mu_1 e^* \sqrt{a}} = \frac{2\pi\sqrt{2\pi} g_2(w)}{(a-w)^{\mu} \mu_1 e^* \sqrt{a}} = h\left(\frac{\mu_2}{\mu_1}, \nu_1, \nu_2, \frac{a-w}{a}\right). \quad (5)$$

For the crack tip at the center of the inhomogeneity ($w=0$), $h(\mu_2/\mu_1)$ for various Poisson's ratios is presented in Figs. 2(a–c) for plane strain and Figs. 3(a–c) for plane stress. For relatively small levels of material mismatch, the number of integration

points required to achieve converged stress intensity factors is approximately 20. However, large levels require a significantly higher number of integration points; the converged results presented in this note were obtained using 300 points. For positive e^*, the stress intensity factor is always negative, indicating crack-tip shielding, and shows a very strong dependence on elastic mismatch, the dependence being greater for plane strain than for plane stress. As expected, K_I^{loc} approaches zero as the inhomogeneity becomes much more compliant than the matrix, and approaches a constant value indicated by dashed lines as the inhomogeneity becomes rigid.

An interesting result of this analysis is that the crack tip is not always shielded. As shown in Fig. 4(a–c), a positive stress intensity factor, indicating amplification, results for crack tips that have entered but have not reached the center of the inhomogeneity.

The results presented above generalize those calculated in Ref. [4], where a transformation toughening model is developed for an elastically homogeneous plate. For this case, the plane-strain normalized stress intensity factor reduces to

$$\frac{K_I^{\mathrm{loc}}}{\mu_1 e^* \sqrt{a}} = -\frac{16}{3\sqrt{8\pi}}\left(\frac{1+\nu}{1-\nu}\right). \tag{6}$$

The results presented in Fig. 2 corresponding to uniform elastic moduli match Eq. (6) to within three significant figures.

References

[1] Wang, Y., and Ballarini, R., 2003, "A Long Crack Penetrating a Circular Inhomogeneity," Meccanica (special issue in honor of Professor Piero Villaggio), Meccanica, **38**, pp. 579–593.

[2] Erdogan, F., Gupta, G. D., and Cook, T. S., 1973, "Numerical Solution of Singular Integral Equations," *Mechanics of Fracture*, G. C. Sih, ed., Noordhoff, Dordrecht, The Netherlands, **1**, Chap. 7, pp. 368–425.

[3] Rubinstein, A. A., 1992, "Stability of the Numerical Procedure for Solution of Singular Integral Equations on Semi-Infinite Interval. Application to Fracture Mechanics," Comput. Struct., **44**(1/2), pp. 71–74.

[4] Lambropoulos, J. C., 1986, "Effect of Nucleation on Transformation Toughening," J. Am. Ceram. Soc., **69**(3), pp. 218–222.

Upper and Lower Bounds for Incipient Failure in a Body Under Gravitational Loading

J. A. Chamberlain, D. J. Horrobin, K. A. Landman, and J. E. Sader[1]

Department of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia
e-mail: jsader@unimelb.edu.au

Recent numerical work has investigated incipient failure of yield stress materials under gravitational loading, for both the rectangular block and cylinder geometries [Chamberlain et al.; 2001, Int. J. Mech. Sci. 43(3):793-815, 2002, Int. J. Mech. Sci. 44(8):1779-1800]. While the rectangular block solution is exact, the cylinder solutions give lower bounds on the height of incipient failure. Consequently, we construct upper bound solutions for the height of incipient failure of a cylinder under gravitational loading. This closes the cylinder problem and quantifies the accuracy of the Haar-Karman hypothesis used in slip-line analysis. For completeness, we also give a simple lower bound solution for the cylinder, as well as upper and lower bound solutions for the two-dimensional rectangular block. These results have the advantage of being analytical, in contrast to the previous purely numerical results. [DOI: 10.1115/1.1767164]

1 Introduction

Flow of a cylinder of yield stress material under gravity to a lower height can be used to determine its yield stress; the relevant experimental technique is commonly termed the ''slump test'' (Murata [1], Christensen [2], and Pashias et al. [3]). While this flow phenomenon has been studied extensively, the closely related problem of *incipient* failure, where the body is on the verge of flowing, has received limited attention in the literature, with the only work to date being the slip-line analyses of Chamberlain et al. [4,5]. While an exact solution was given for the case of a rectangular block (Chamberlain et al. [5]), only a lower bound solution was derived for the complementary case of a cylinder (Chamberlain et al. [4]) with the Haar-Karman hypothesis being invoked. The principal aim of this article is therefore to close the cylinder problem by using limit analysis to construct an upper bound on the height of incipient failure. For completeness, we also present upper and lower bound analyses for the rectangular block. The analytical formulas obtained using upper and lower bound analyses are of practical value due to their simplicity.

The geometry and coordinate systems used are shown in Fig. 1. Final results are scaled by the length $2\tau_y/(\rho g)$, where τ_y is the shear yield stress, ρ is the density and g is the acceleration due to gravity, giving the scaled radius or half-width r_0 and the scaled height of incipient failure h. In the analyses we use both Tresca and von Mises yield conditions (Desai and Siriwardane [6]).

2 Lower Bound Analyses

To construct a lower bound solution, we specify a statically admissible stress field.

Plane Strain Rectangular Block. An appropriate statically admissible stress field for the rectangular block is

$$\boldsymbol{\Sigma} = -\rho g\left[\left(\frac{2\tau_y}{\rho g} - X_2\right)\hat{\mathbf{X}}_2\hat{\mathbf{X}}_2 + \left(\frac{\tau_y}{\rho g}\right)\hat{\mathbf{X}}_3\hat{\mathbf{X}}_3\right], \tag{1}$$

where the height of the block is

$$H_b = \frac{2\tau_y}{\rho g}. \tag{2}$$

Using the lower bound theorem of limit analysis (Chakrabarty [7]) it then follows that the height of incipient failure is greater than or equal to H_b, i.e., in scaled terms

$$h \geqslant 1. \tag{3}$$

Axisymmetric Cylinder. The statically admissible stress field for the cylinder case is chosen as

$$\boldsymbol{\Sigma} = -\rho g\left(\frac{\sigma_y}{\rho g} - Z\right)\hat{\mathbf{Z}}\hat{\mathbf{Z}}, \tag{4}$$

with the height of the body, H_b, equal to $\sigma_y/(\rho g)$, where σ_y is the uniaxial yield stress. In scaled terms, the lower bound results are

$$h \geqslant 1 \quad \text{(Tresca)} \quad \text{and} \quad h \geqslant \frac{\sqrt{3}}{2} \quad \text{(von Mises)}. \tag{5}$$

3 Upper Bound Analyses

To use the upper bound theorem of limit analysis we require a kinematically admissible velocity field (Chakrabarty [7]). We

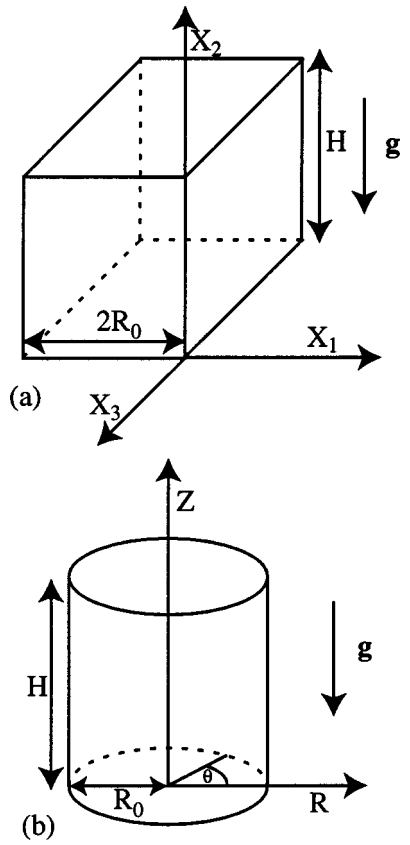


Fig. 1 Geometry and coordinate systems (*a*) plane-strain rectangular block (*b*) axisymmetric cylinder

[1]To whom correspondence should be addressed. E-mail: jsader@unimelb.edu.au

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, July 23, 2003; final revision, January 27, 2004. Associate Editor: Z. Suo.

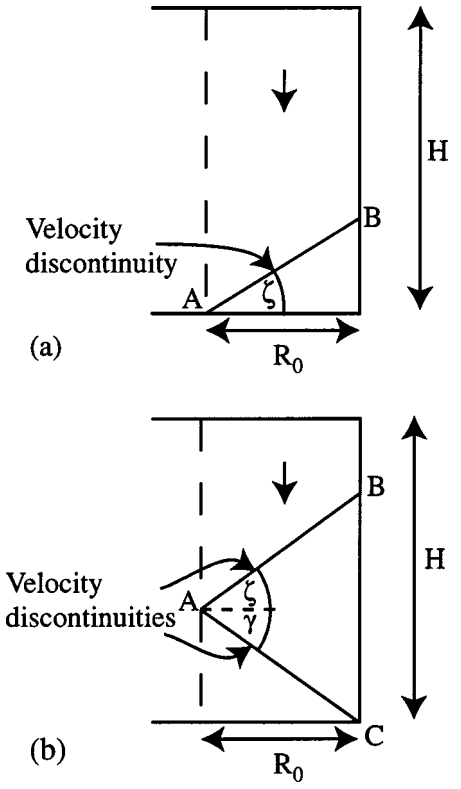


Fig. 2 Diagram of the upper bound solutions for (*a*) perfect slip on the base (single velocity discontinuity) and (*b*) perfectly rough base (double velocity discontinuity)

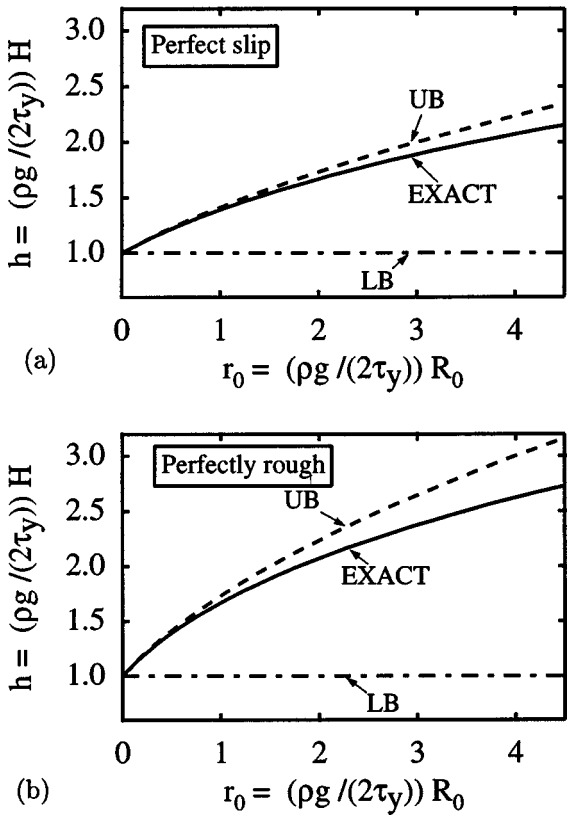


Fig. 3 Height of incipient failure for the plane strain rectangular block: (*a*) *perfect slip* upper bound (UB) from Eq. (7), exact slip-line (EXACT) see Ref. [5], lower bound (LB) from Eq. (3); (*b*) *perfectly rough base* upper bound (UB) from Eq. (10), exact slip-line (EXACT) see Ref. [5], lower bound (LB) from Eq. (3)

consider two extremes of base friction: (i) perfect slip and (ii) a perfectly rough base.

Plane-Strain Rectangular Block

Perfect Slip Base. The flow field is approximated by rigid regions separated by velocity discontinuities, as illustrated in Fig. 2(*a*). The kinematically admissible velocity field is

$$\mathbf{v}^*=\begin{cases} -\hat{\mathbf{X}}_2 & \text{above } AB, \\ \cot\zeta\hat{\mathbf{X}}_1 & \text{below } AB, \end{cases} \tag{6}$$

from which we obtain (Chakrabarty [7]) the upper bound

$$h\leqslant\sqrt{1+r_0}. \tag{7}$$

Perfectly Rough Base. To satisfy the perfectly rough condition on the base, we require two discontinuities, see Fig. 2(*b*). The kinematically admissible velocity field is given by

$$\mathbf{v}^*=\begin{cases} -\hat{\mathbf{X}}_2 & \text{above } AB, \\ u_1\hat{\mathbf{X}}_1+u_2\hat{\mathbf{X}}_2 & \text{between } AB \text{ and } AC, \\ \mathbf{0} & \text{below } AC, \end{cases} \tag{8}$$

where

$$u_1=\frac{1}{\tan\zeta+\tan\gamma},\quad u_2=\frac{-\tan\gamma}{\tan\zeta+\tan\gamma}. \tag{9}$$

The upper bound result is then

$$h\leqslant\sqrt{1+2r_0}. \tag{10}$$

Axisymmetric Cylinder. For the axisymmetric geometry, we use configurations of velocity discontinuities identical to those used in the plane strain solutions (see Fig. 2). In addition, we specify that the radial velocity is constant (Kudo [8]).

Perfect Slip Base. Referring to Fig. 2(*a*), the velocity field is specified as

$$\mathbf{v}^*=\begin{cases} -\hat{\mathbf{Z}} & \text{above } AB, \\ \dfrac{1}{2b}\left(\hat{\mathbf{R}}-\dfrac{Z\hat{\mathbf{Z}}}{R}\right) & \text{below } AB, \end{cases} \tag{11}$$

where $b=\tan\zeta$, and AB is a velocity discontinuity.

Tresca Yield Condition. Applying the upper bound inequality and minimizing with respect to b, we obtain

$$4h\leqslant\frac{1}{b}+b+2r_0b+1+\frac{1}{2}\sqrt{1+b^2}+\frac{1}{2b}\ln(\sqrt{1+b^2}+b). \tag{12}$$

where b satisfies

$$2b^2(1+2r_0)+b\sqrt{1+b^2}=2+\ln(b+\sqrt{1+b^2}). \tag{13}$$

von Mises Yield Condition. Applying the upper bound inequality and minimizing with respect to b, we obtain

$$4h\leqslant\frac{1}{b}+b+2r_0b+\sqrt{1+\frac{b^2}{4}}+\frac{2}{b}\ln\left(\sqrt{1+\frac{b^2}{4}}+\frac{b}{2}\right), \tag{14}$$

where

$$b^2(1+2r_0)+b\sqrt{1+\frac{b^2}{4}}=1+2\ln\left(\sqrt{1+\frac{b^2}{4}}+\frac{b}{2}\right). \tag{15}$$

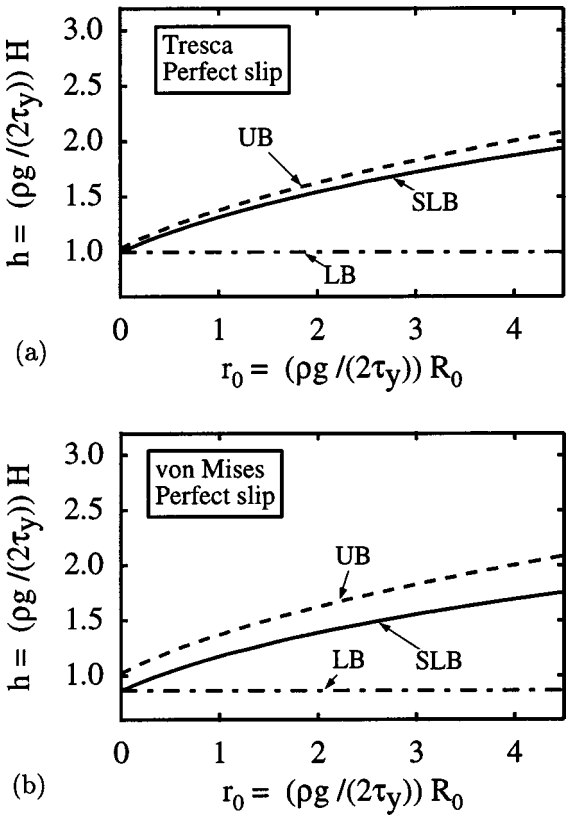


Fig. 4 Height of incipient failure for the cylinder with perfect slip on the base: (*a*) *Tresca yield condition* upper bound (UB) from Eqs. (12)–(13), slip-line lower bound (SLB) see Ref. [4], lower bound (LB) from Eq. (5); (*b*) *von Mises yield condition* upper bound (UB) from Eqs. (14)–(15), slip-line lower bound (SLB) see Ref. [4], lower bound (LB) from Eq. (5)

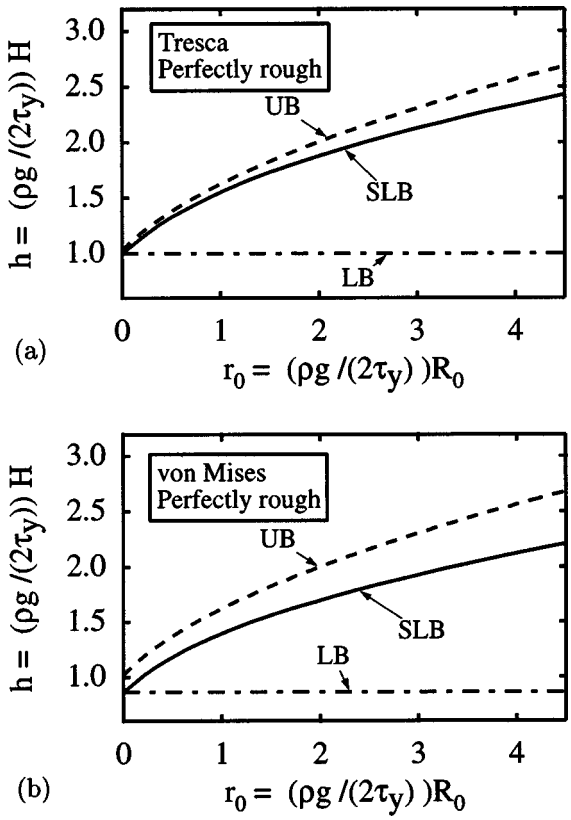


Fig. 5 Height of incipient failure for the cylinder with a perfectly rough base: (*a*) *Tresca yield condition* upper bound (UB) from Eqs. (17)–(18), slip-line lower bound (SLB) see Ref. [4], lower bound (LB) from Eq. (5); (*b*) *von Mises yield condition* upper bound (UB) from Eqs. (19)–(20), slip-line lower bound (SLB) see Ref. [4], lower bound (LB) from Eq. (5).

Perfectly Rough Base. Referring to Fig. 2(*b*), the velocity field $\mathbf{v}^*$ is given by

$$\mathbf{v}^*=\begin{cases} -\hat{\mathbf{Z}} & \text{above } AB, \\ \dfrac{1}{2t}\left\{\hat{\mathbf{R}}+\hat{\mathbf{Z}}\dfrac{1}{R}[at(R_0-2R)-Z]\right\} & \text{between } AB \text{ and } AC, \\ \mathbf{0} & \text{below } AC, \end{cases} \tag{16}$$

where a and t are defined by $at=\tan\gamma$ and $(1-a)t=\tan\zeta$.

Tresca Yield Condition. Applying the upper bound inequality, minimizing with respect to a and t, gives the required result

$$4h\leqslant\frac{2}{t}+\frac{t}{2}+2r_0t+1+\frac{1}{2}\sqrt{1+\frac{t^2}{4}}+\frac{1}{2t}\ln\left(\frac{\sqrt{4+t^2}+t}{\sqrt{4+t^2}-t}\right), \tag{17}$$

where

$$4+\ln\left(\frac{\sqrt{4+t^2}+t}{\sqrt{4+t^2}-t}\right)=(1+4r_0)t^2+\frac{t}{2}\sqrt{4+t^2}. \tag{18}$$

von Mises Yield Condition. Applying the upper bound inequality and minimizing with respect to a and t gives

$$4h\leqslant\frac{2}{t}+\frac{t}{2}+2r_0t+\sqrt{1+\frac{t^2}{16}}+\frac{2}{t}\ln\left(\frac{\sqrt{16+t^2}+t}{\sqrt{16+t^2}-t}\right), \tag{19}$$

with t given by

$$4+4\ln\left(\frac{\sqrt{16+t^2}+t}{\sqrt{16+t^2}-t}\right)=(1+4r_0)t^2+\frac{t}{2}\sqrt{16+t^2}. \tag{20}$$

4 Results and Discussion

The above results are illustrated in Figs. 3, 4, and 5. For the rectangular block we find that the upper bound solutions give an excellent approximation to the exact slip-line results, whereas the lower bound solutions are comparatively poor.

For the cylinder case, the actual solution must lie between the slip-line lower bound and the appropriate upper bound solution, possibly touching one of these curves (see Figs. 4 and 5). Consequently, the error in the height of incipient failure introduced by using the Haar-Karman hypothesis is bounded by the difference between the slip-line lower bound solution and the upper bound solution. This difference depends on the yield condition, base boundary condition, and radius, and is quantified in Figs. 3 and 4.

Acknowledgments

The authors acknowledge funding from the Particulate Fluids Processing Center, an Australian Research Council Special Research Center, and the David Hay Memorial Fund.

References

[1] Murata, J., 1984, "Flow and Deformation of Fresh Concrete," Mater. Constr. (Paris), **17**(98), pp. 117–129.

[2] Christensen, G., 1991, "Modelling the Flow of Fresh Concrete: The Slump Test," Ph.D. thesis, Princeton University, Princeton, NJ.

[3] Pashias, N., Boger, D. V., Summers, J., and Glenister, D. J., 1996, ''A Fifty Cent Rheometer for Yield Stress Measurement,'' J. Rheol., **40**(6), pp. 1179–1189.
[4] Chamberlain, J. A., Sader, J. E., Landman, K. A., Horrobin, D. J., and White, L. R., 2002, ''Incipient Failure of a Circular Cylinder Under Gravity,'' Int. J. Mech. Sci., **44**(8), pp. 1779–1800.
[5] Chamberlain, J. A., Sader, J. E., Landman, K. A., and White, L. R., 2001, ''Incipient Plane-Strain Failure of a Rectangular Block Under Gravity,'' Int. J. Mech. Sci., **43**(3), pp. 793–815.
[6] Desai, C. S., and Siriwardane, H. J., 1984, *Constitutive Laws for Engineering Materials With Emphasis on Geologic Materials*, Prentice-Hall, Englewood Cliffs, NJ.
[7] Chakrabarty, J., 2000, *Applied Plasticity*, Springer-Verlag, New York.
[8] Kudo, H., 1960, ''Some Analytical and Experimental Studies of Axisymmetric Cold Forging and Extrusion—I,'' Int. J. Mech. Sci., **2**(1), pp. 102–127.

Buckling of a Rotating Rod Under Axial Force

C. Y. Wang
Professor, Mem. ASME, Departments of Mathematics and Mechanical Engineering, Michigan State University, East Lansing, MI 48824

The difference equations governing the rotation of a segmented rod under axial force is formulated. The stability boundaries are found to be highly dependent on the number of links, the rotation rate and the compressive force. For a large number of links, the result approaches to that of the continuous elastic rod through some fractional power. The analysis is applicable to segmented drill shafts.
[DOI: 10.1115/1.1767165]

Introduction

The prediction of buckling or whirling of an axially rotating rod is important to the design of shafts and rotating machinery. The linearized stability of elastic rotating rods leads to a fourth-degree eigenvalue problem (see, for example, [1,2]). Bobisud and Christenson [3] considered the case where an axial force is applied to a rod whose one end is clamped and rotated and the other end hinged. The problem is important in the stability of drill bits.

Rotating shafts are sometimes too long to be made or transported in one piece. For example, drill shafts in mining or well drilling are composed of several segments connected together. Marine and truck engine shafts may also be jointed. Previous literature in this area includes the work by Wang [4] who studied a segmented rod rotated at one end while the other end is free. However that source cannot be applied to the drill shaft since it does not admit an axial force. The present paper studies the stability of an idealized segmented drill shaft which is composed of joined rigid links. The results are compared with those of the continuous elastic rod.

Formulation

Consider a rod with N segments joined together. The joints are strengthened by rotational springs to maintain the structure's straightness. The rod is then rotated at one end with angular velocity Ω and compressed axially by the force F'. Figure 1(*a*) shows the buckled rod under a Cartesian system rotated about the x' axis. Figure 1(*b*) shows the force balance on the nth link. Notice that the resultant centrifugal force C_n' does not act on the mid point of the link, but at a distance s_n' form the left end at (x_{n-1}', y_{n-1}'). Simple integration along the link shows

$$C_n' = m\Omega^2 (y_{n-1}' + y_n')/2 \tag{1}$$

$$s_n' = \frac{l(y_{n-1}' + 2y_n')}{3(y_{n-1}' + y_n')}. \tag{2}$$

Here m is the mass of the link and l is its length. The two ends are related by

$$x_n' = x_{n-1}' + l\cos\theta_n, \tag{3}$$

$$y_n' = y_{n-1}' + l\sin\theta_n \tag{4}$$

where θ_n is the angle of inclination. A vertical force balance gives

$$G_n' + C_n' = G_{n-1}'. \tag{5}$$

A moment balance about the left end gives

$$G_n' l\cos\theta_n + C_n' s_n'\cos\theta_n + F' l\sin\theta_n + M_n' = M_{n-1}'. \tag{6}$$

The end moment M_n' is in general a function of the angle difference of two adjacent links. For small deformations the moment is proportional to the angle difference

$$M_n' = \lambda(\theta_{n+1} - \theta_n). \tag{7}$$

where λ is the linear rotational spring constant. Normalize all lengths by the total length $L = Nl$ and all forces by λ/l and drop primes. Eqs. (1)–(6) become

$$C_n \omega^2 (y_{n-1} + y_n)/2, \tag{8}$$

$$s_n = \frac{(y_{n-1} + 2y_n)}{3N(y_{n-1} + y_n)} \tag{9}$$

$$x_n = x_{n-1} + \cos\theta_n/N, \tag{10}$$

$$y_n = y_{n-1} + \sin\theta_n/N \tag{11}$$

$$G_n = G_{n-1} - C_n \tag{12}$$

$$(G_n + C_n s_n)\cos\theta_n + F\sin\theta_n + \theta_{n+1} - 2\theta_n + \theta_{n-1} = 0. \tag{13}$$

Here

$$\omega \equiv \Omega L\sqrt{\frac{m}{\lambda N}} \tag{14}$$

is a normalized rotation rate. The boundary conditions are that the rod is ''clamped'' at one end and ''hinged'' at the other end. This is equivalent to extending with fictitious links θ_0 and θ_{N+1} such that

$$\theta_0 = 0, \tag{15}$$

$$\theta_N = \theta_{N+1}. \tag{16}$$

The other boundary conditions are

$$x_0 = 0, \tag{17}$$

$$y_0 = 0, \tag{18}$$

$$y_N = 0. \tag{19}$$

For incipient buckling, the angles of inclination are small. Equations (11)–(13) linearize to

$$G_n + F\theta_n + \frac{\omega^2}{6N}(y_{n-1} + 2y_n) + \theta_{n+1} - 2\theta_n + \theta_{n-1} = 0, \quad n = 1 \text{ to } N \tag{20}$$

$$y_n = y_{n-1} + \frac{1}{N}\theta_n \quad n = 1 \text{ to } N \tag{21}$$

$$G_n = G_{n-1} - \frac{\omega^2}{2}(y_{n-1} + y_n) \quad n = 1 \text{ to } N. \tag{22}$$

Together with Eqs. (15), (16), (18), and (19) there are $3N+4$ equations for the $3N+4$ unknowns $\theta_0 \cdots \theta_{N+1}$, $y_0 \cdots y_N$, $G_0 \cdots G_N$. For nontrivial solutions the determinant of coefficients is set to zero, giving a characteristic equation for the parameters F and ω.

Stability

The rod cannot buckle with one link. If $N = 2$, the characteristic equation gives the stability boundary

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, October 16, 2003; final revision, January 15, 2004. Associate Editor: S. Mukherjee.

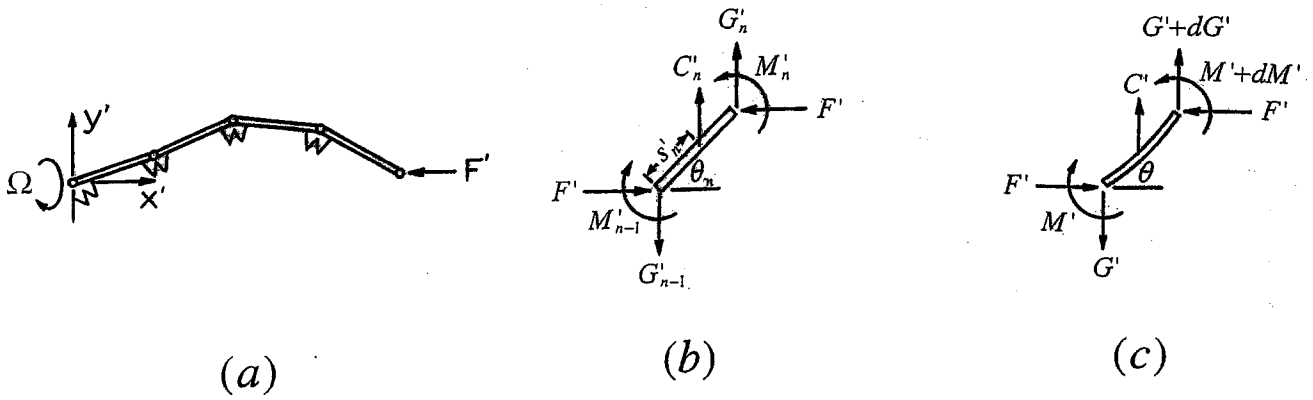


Fig. 1 (*a*) **The rotating segmented rod.** (*b*) **Force balance on the *n*th link, C'_n is the centrifugal force.** (*c*) **A segment of the continuous rod.**

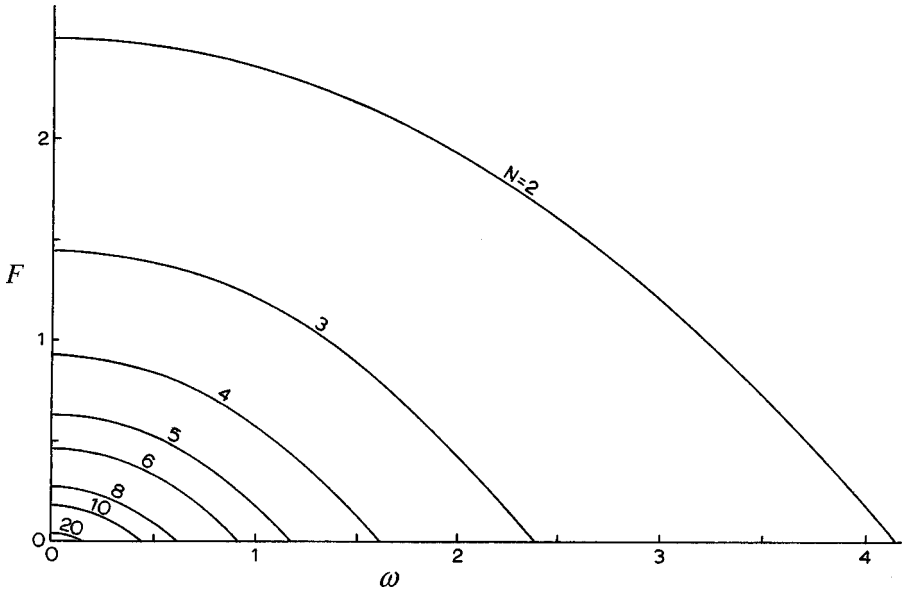


Fig. 2 Stability boundaries for the primary mode

$$F+\frac{7}{48}\omega^2-\frac{5}{2}=0. \tag{23}$$

The rod is stable if the left-hand side of Eq. (23) is less than zero. For $N=3$, after some work, we find the two solutions

$$F+\frac{1}{3}\left(\frac{4}{9}\omega^2-7\pm\sqrt{\frac{11}{108}\omega^4+\frac{5}{3}\omega^2+7}\right)=0 \tag{24}$$

corresponding to the two modes of buckling. The determinant is solved numerically for $N\geqslant 4$. Figure 2 shows the stability boundaries for the primary mode. The rod is unstable if rotation and axial force indicate a state above each curve. Of special interest are the buckling loads when rotation is absent ($\omega=0$). Table 1 shows the buckling loads of all the modes for $N\leqslant 6$.

The higher modes occur when the lateral displacement of the rod is restricted, for example a rod in a confining tube. The four modes for $N=5$ are shown in Fig. 3 for the same initial angle. On the other hand, if the rod is subjected to rotation only ($F=0$) the critical rotation speeds for $N\leqslant 6$ are given in Table 2.

The mode shapes for $N=5$ is shown in Fig. 4. In general, the rod would bend to one side for the first mode, and there are n-1 interior zeroes for the n^{th} mode. It is clear from Figs. 3 and 4 that the mode shapes due to axial force or centrifugal force are quite different.

Comparison With the Continuous Elastic Rod

The continuous elastic rod was considered by Bobisud and Christenson [3]. An alternative derivation is as follows. Figure 1(*c*) shows an elemental segment of the rod. Let s' be the arc length from the origin and θ be the local angle of inclination. A vertical force balance yields

Table 1 Buckling loads when rotation is absent

$N=2$	3	4	5	6
$F=2.5$	1.4514	0.924	0.634	0.461
	3.2153	2.306	1.677	1.258
		3.520	2.811	2.220
			3.677	3.126
				3.768

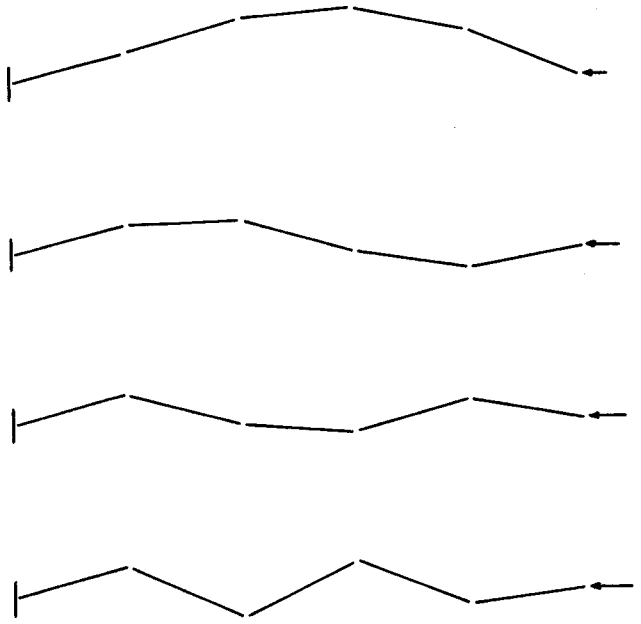


Fig. 3 The buckling modes for $N=5$ when rotation is absent. From top, first to fourth modes, respectively.

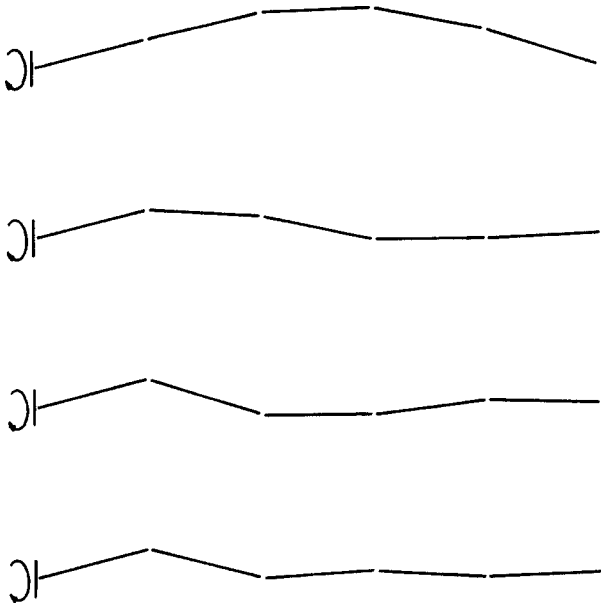


Fig. 4 The buckling modes for $N=5$ when axial force is absent. From top, first to fourth modes, respectively.

Table 2 Critical speeds when axial force is absent

$N=2$	3	4	5	6
$\omega=4.1403$	2.3918	1.6123	1.1845	0.9192
	8.7596	5.6715	4.0520	3.0932
		12.802	9.1374	6.8402
			16.117	12.442
				18.872

$$dG' + \rho ds' \Omega^2 y' = 0 \tag{25}$$

where ρ is the mass per length. Moment balance about the left end gives

$$dM' + F' ds' \sin\theta + G' ds' \cos\theta = 0. \tag{26}$$

The Euler-Bernoulli elastic law states that the local moment is proportional to the local curvature

$$M' = EI \frac{d\theta}{ds'} \tag{27}$$

where EI is the flexural rigidity. The kinematic relation is

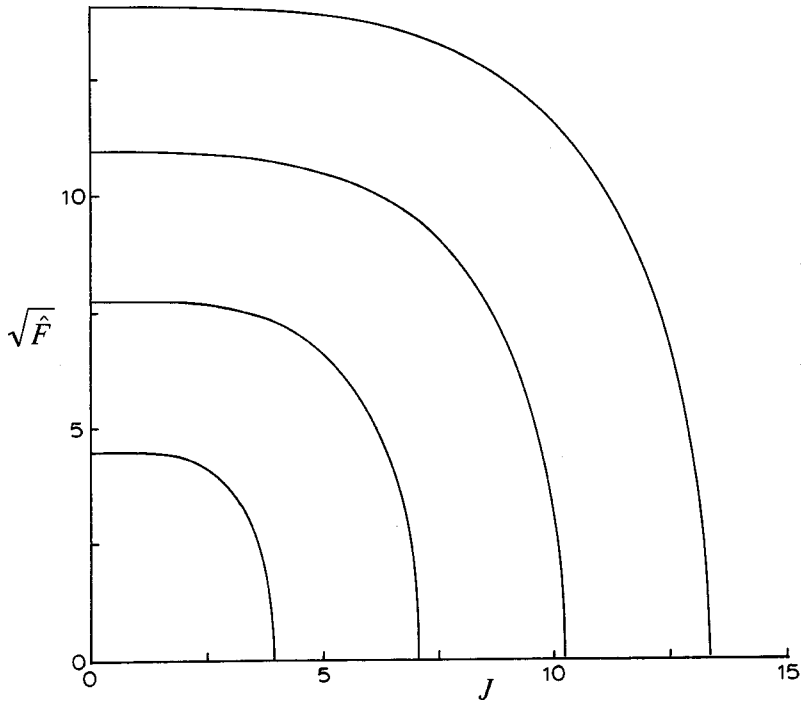


Fig. 5 Stability boundaries for the continuous rod, first four modes

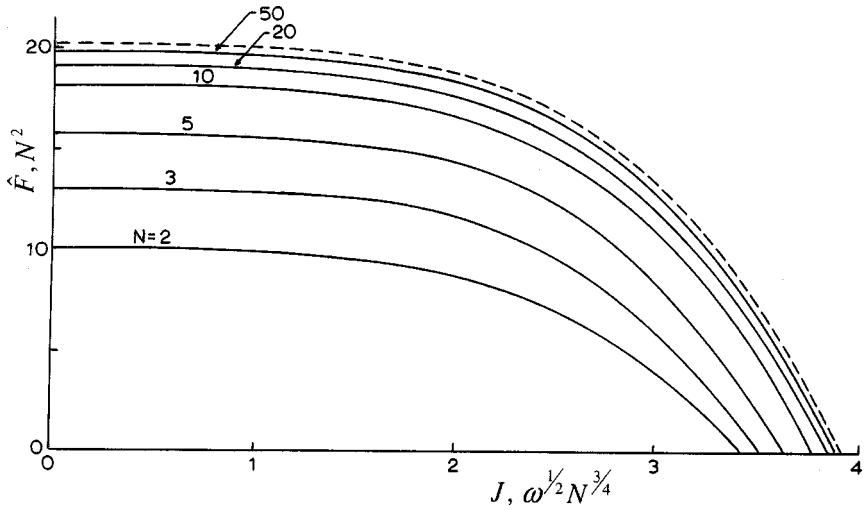


Fig. 6 Comparisons of the stability boundaries of the segmented rod (solid lines) to that of the continuous rod (dashed lines), first mode only

$$\frac{dy'}{ds'} = \sin\theta. \tag{28}$$

Normalize all lengths by the rod length L and drop primes. The forces are normalized by EI/L^2 and indicated by hats. Eqs. (25)–(28) yield the linearized equations

$$\frac{d^2\theta}{ds^2} + \hat{F}\theta + \hat{G} = 0, \tag{29}$$

$$\frac{d\hat{G}}{ds} + J^4 y = 0, \tag{30}$$

$$\frac{dy}{ds} = \theta \tag{31}$$

where $J^4 \equiv \rho\Omega^2 L^4/EI$ is the rotation parameter. Equations (29)–(31) is simplified to

$$\frac{d^4\theta}{ds^4} + \hat{F}\frac{d^2\theta}{ds^2} - J^4\theta = 0. \tag{32}$$

The solution of Eq. (32) is a linear combination of $\cosh(\alpha s)$, $\sinh(\alpha s)$, $\cos(\beta s)$, $\sin(\beta s)$ where

$$\alpha = \sqrt{(\sqrt{\hat{F}^2 + 4J^4} - \hat{F})/2}, \tag{33}$$

$$\beta = \sqrt{(\sqrt{\hat{F}^2 + 4J^4} + \hat{F})/2}. \tag{34}$$

The fixed-hinged boundary conditions are

$$\theta(0) = 0, \quad \frac{d\theta}{ds}(1) = 0. \tag{35}$$

From Eqs. (29), (30), the requirement of zero displacements at the ends gives

$$\frac{d^3\theta}{ds^3}(0) + \hat{F}\frac{d\theta}{ds}(0) = 0, \tag{36}$$

$$\frac{d^3\theta}{ds^3}(1) + \hat{F}\frac{d\theta}{ds}(1) = 0. \tag{37}$$

For nontrivial solutions of θ, the resulting characteristic equation from Eqs. (36), (37) is simplified to

$$\beta\cos\beta\sinh\alpha - \alpha\cosh\alpha\sin\beta = 0 \tag{38}$$

which is the same equation obtained in Ref. [3]. Figure 5 shows the stability boundaries for the first four modes ($\sqrt{F}$ is plotted to compress the figure).

The connection between a segmented rod and a continuous rod was delineated by Wang [4]. For large N the joint spring constant λ of the segmented rod is asymptotically related to the flexural rigidity EI of the continuous rod by

$$EI \approx \lambda L/N. \tag{39}$$

Using Eq. (39), the normalized forces and rotation rates for large N are related by

$$\hat{F} \approx FN^2, \quad J \approx \omega^{1/2} N^{3/4}. \tag{40}$$

Figure 6 shows the stability boundaries for the first mode, replotted in terms of the combinations given in Eq. (40). We see that the $N = 50$ curve is already fairly close to the continuous case. Thus the results of the continuous case can be used (within 2% error) if the number of links of the segmented rod is more than 50.

Discussions

The nonlinear difference equation governing the rotation of a segmented rod under axial force can also be used for large deformations, given the constitutive relation of the rotational spring. We find that the centrifugal force and axial force have different effects on the buckling mode, while both decreases stability. Increasing the number of links N decreases stability, and increases the number of eigenmodes. As N becomes large, the results for the continuous elastic rod is recovered, but only with the proper fractional transformation given in Eq. (40). Our Figs. 2 and 6 should be useful in the design of segmented drill shafts.

References

[1] Love, A. E. H., 1944, *A Treatise on the Mathematical Theory of Elasticity*, 4th Ed., Dover, New York.
[2] Odeh, F., and Tadjbakhsh, I., 1965, "A Nonlinear Eigenvalue Problem for Rotating Rods," Arch. Ration. Mech. Anal., **20**, pp. 81–94.
[3] Bobisud, L. E., and Christenson, C. O., 1979, "Critical Force in the Buckling of Drill Bits," ASME J. Appl. Mech., **46**, pp. 461–462.
[4] Wang, C. Y., 1993, "The Axially Rotating Segmented Rod," Int. J. Solids Struct., **30**, pp. 2437–2444.

Journal of Applied Mechanics

Discussion

Discussion: "Zeroth-Order Shear Deformation Theory for Laminated Composite Plates" (Ray, M. C., 2003 ASME J. Appl. Mech., 70, pp. 374–380)

S. Kapuria
e-mail: kapuria@am.iitd.ac.in

P. C. Dumir

Applied Mechanics Department, I.I.T. Delhi, Hauz Khas, New Delhi 110016, India

It is the contention of the authors that the ''zeroth-order'' shear deformation theory presented by Ray [1] is mathematically equivalent to Reddy's third order theory [2]. The notation of Ref. [1] is used herein. Ray's approximations for the in-plane displacements

$$u=u_0-zw_{,x}+\left(\frac{3z}{2h}-\frac{2z^3}{h^3}\right)\frac{Q_x}{\lambda_x},\quad v=v_0-zw_{,y}+\left(\frac{3z}{2h}-\frac{2z^3}{h^3}\right)\frac{Q_y}{\lambda_y} \tag{1}$$

are identical to those of the Reddy's theory, [2], with

$$\psi_x+w_{,x}=\frac{3Q_x}{2h\lambda_x},\quad \psi_y+w_{,y}=\frac{3Q_y}{2h\lambda_y}. \tag{2}$$

Hence the equations of motion and boundary conditions of Ray's theory are mathematically equivalent to those of the dynamic version of Reddy's theory. This is established explicitly by comparing the governing equations of the two theories.

For Reddy's theory, the equations of motion are

$$N_{x,x}+N_{xy,y}=I_0\ddot{u}_0-I_1\ddot{w}_{,x}+\frac{2h}{3}I_8(\ddot{\psi}_x+\ddot{w}_{,x}) \tag{3}$$

$$N_{xy,x}+N_{y,y}=I_0\ddot{v}_0-I_1\ddot{w}_{,y}+\frac{2h}{3}I_8(\ddot{\psi}_y+\ddot{w}_{,y}) \tag{4}$$

$$Q_{x,x}-\frac{4}{h^2}R_{x,x}+Q_{y,y}-\frac{4}{h^2}R_{y,y}+\frac{4}{3h^2}(P_{x,xx}+2P_{xy,xy}+P_{y,yy})+p$$
$$=I_0\ddot{w}+\frac{4}{3h^2}I_3(\ddot{u}_{0,x}+\ddot{v}_{0,y})+\frac{4}{3h^2}\left(I_4-\frac{4I_6}{3h^2}\right)(\ddot{\psi}_{x,x}+\ddot{w}_{,xx}$$
$$+\ddot{\psi}_{y,y}+\ddot{w}_{,yy})-\frac{4I_4}{3h^2}(\ddot{w}_{,xx}+\ddot{w}_{,yy}) \tag{5}$$

$$\left(M_x-\frac{4}{3h^2}P_x\right)_{,x}+\left(M_{xy}-\frac{4}{3h^2}P_{xy}\right)_{,y}-\left(Q_x-\frac{4}{h^2}R_x\right)$$
$$=\frac{2h}{3}\left[I_8\ddot{u}_0-I_9\ddot{w}_{,x}+\frac{2h}{3}I_7(\ddot{\psi}_x+\ddot{w}_{,x})\right]. \tag{6}$$

$$\left(M_y-\frac{4}{3h^2}P_y\right)_{,y}+\left(M_{xy}-\frac{4}{3h^2}P_{xy}\right)_{,x}-\left(Q_y-\frac{4}{h^2}R_y\right)$$
$$=\frac{2h}{3}\left[I_8\ddot{u}_0-I_9\ddot{w}_{,y}+\frac{2h}{3}I_7(\ddot{\psi}_y+\ddot{w}_{,y})\right]. \tag{7}$$

For Ray's theory, the equations of motion are

$$N_{x,x}+N_{xy,y}=I_0\ddot{u}_0-I_1\ddot{w}_{,x}+I_8\frac{\ddot{Q}_x}{\lambda_x} \tag{8}$$

$$N_{xy,x}+N_{y,y}=I_0\ddot{v}_0-I_1\ddot{w}_{,y}+I_8\frac{\ddot{Q}_y}{\lambda_y} \tag{9}$$

$$M_{x,xx}+2M_{xy,xy}+M_{y,yy}+p$$
$$=I_0\ddot{w}+I_1(\ddot{u}_{0,x}+\ddot{v}_{0,y})-I_2(\ddot{w}_{,xx}+\ddot{w}_{,yy})+I_9\left(\frac{\ddot{Q}_{x,x}}{\lambda_x}+\frac{\ddot{Q}_{y,y}}{\lambda_y}\right) \tag{10}$$

$$\left(M_x-\frac{4}{3h^2}P_x\right)_{,x}+\left(M_{xy}-\frac{4}{3h^2}P_{xy}\right)_{,y}-\left(Q_x-\frac{4}{h^2}R_x\right)$$
$$=\frac{2h}{3}\left[\frac{I_7}{\lambda_x}\ddot{Q}_x+I_8\ddot{u}_0-I_9\ddot{w}_{,x}\right] \tag{11}$$

$$\left(M_y-\frac{4}{3h^2}P_y\right)_{,y}+\left(M_{xy}-\frac{4}{3h^2}P_{xy}\right)_{,x}-\left(Q_y-\frac{4}{h^2}R_y\right)$$
$$=\frac{2h}{3}\left[\frac{I_7}{\lambda_y}\ddot{Q}_y+I_8\ddot{v}_0-I_9\ddot{w}_{,y}\right] \tag{12}$$

with $I_7=9I_2/4h^2-6I_4/h^4+4I_6/h^6$. Using Eq. (2), it is observed that Eqs. (8), (9), (11), (12) are identical to Eqs. (3), (4), (6), (7). Forming the combination Eq. (10)–Eq. $(11)_{,x}$–Eq. $(12)_{,y}$ yields Eq. (5).

For Reddy's theory, the boundary conditions are obtained from the following boundary integral formed after using Green's theorem in Hamilton's principle

$$\int \Big[N_n \delta u_n + N_{ns} \delta u_s + \hat{M}_n \delta \psi_n + \hat{M}_{ns} \delta \psi_s - \frac{4}{3h^2} P_n \delta w_{,n} + \Big[\hat{Q}_x n_x + \hat{Q}_y n_y + \frac{4}{3h^2} \{ (P_{x,x} + P_{xy,y}) n_x + (P_{y,y} + P_{xy,x}) n_y + P_{ns,s} \} - \frac{4I_3}{3h^2} \Big\{ \ddot{u}_0 + \Big(I_4 - \frac{4I_6}{3h^2} \Big) (\ddot{\psi}_x + \ddot{w}_{,x}) - I_4 \ddot{w}_{,x} \Big\} n_x - \frac{4I_3}{3h^2} \Big\{ \ddot{v}_0 + \Big(I_4 - \frac{4I_6}{3h^2} \Big) (\ddot{\psi}_y + \ddot{w}_{,y}) - I_4 \ddot{w}_{,y} \Big\} n_y \Big] \delta w \Big] ds - \sum_i \frac{4}{3h^2} \Delta P_{ns}(s_i) \delta w(s_i) \tag{13}$$

where s_i are locations of plate corners and

$$u_n = u_0 n_x + v_0 n_y, \quad u_s = u_0 s_x + v_0 s_y$$

$$N_n = N_x n_x^2 + N_y n_y^2 + 2N_{xy} n_x n_y,$$

$$N_{ns} = N_x n_x s_y + N_y n_y s_y + N_{xy}(n_x s_y + n_y s_x)$$

$$\hat{Q}_\alpha = Q_\alpha - \frac{4}{h^2} R_\alpha \ (\alpha = x,y), \quad \hat{M}_\beta = M_\beta - \frac{4}{3h^2} P_\beta \ (\beta = x,y,xy) \tag{14}$$

with $s_x = -n_y, s_y = n_x$. The expressions of $\hat{M}_n, \hat{M}_{ns}; M_n, M_{ns}; P_n, P_{ns}$ are similar to those of N_n, N_{ns}, and of ψ_n, ψ_s; $w_{,n}, w_{,s}$ are similar to those of u_n, u_s. For Ray's theory the corresponding boundary integral is

$$\int \Big[N_n \delta u_n + N_{ns} + \delta u_s + \frac{3\delta Q_x}{2h\lambda_x} \Big\{ \Big(M_x - \frac{4}{3h^2} P_x \Big) n_x + \Big(M_{xy} - \frac{4}{3h^2} P_{xy} \Big) n_y \Big\} + \frac{3\delta Q_y}{2h\lambda_y} \Big\{ \Big(M_y - \frac{4}{3h^2} P_y \Big) n_y + \Big(M_{xy} - \frac{4}{3h^2} P_{xy} \Big) n_x \Big\} - M_n \delta w_{,n} - M_{ns} \delta w_{,s} + \Big\{ (M_{x,x} + M_{xy,y}) n_x + (M_{xy,x} + M_{y,y}) n_y + \Big(-I_1 \ddot{u}_0 + I_2 \ddot{w}_{,x} - \frac{\ddot{Q}_x}{\lambda_x} I_9 \Big) n_x + \Big(-I_1 \ddot{v}_0 + I_2 \ddot{w}_{,y} - \frac{\ddot{Q}_y}{\lambda_y} I_9 \Big) n_y \Big\} \delta w \Big] ds. \tag{15}$$

Substituting

$$\frac{3}{2h\lambda_x} \delta Q_x = \delta \psi_x + \delta w_{,x} = (\delta \psi_n + \delta w_{,n}) n_x + (\delta \psi_s + \delta w_{,s}) s_x$$

$$\frac{3}{2h\lambda_y} \delta Q_y = \delta \psi_y + \delta w_{,y} = (\delta \psi_n + \delta w_{,n}) n_y + (\delta \psi_s + \delta w_{,s}) s_y \tag{16}$$

in Eq. (15) reduces it to

$$\int \Big[N_n \delta u_n + N_{ns} \delta u_s + \hat{M}_n \delta \psi_n + \hat{M}_{ns} \delta \psi_s - \frac{4}{3h^2} (P_n \delta w_{,n} + P_{n,s} \delta w_{,s}) + \Big\{ (M_{x,x} + M_{xy,x}) n_x + (M_{xy,y} + M_{y,y}) n_y + \Big(-I_1 \ddot{u}_0 + I_2 \ddot{w}_{,x} - \frac{\ddot{Q}_x}{\lambda_x} I_9 \Big) n_x + \Big(-I_1 \ddot{v}_0 + I_2 \ddot{w}_{,y} - \frac{\ddot{Q}_y}{\lambda_y} I_9 \Big) n_y \Big\} \delta w \Big] ds. \tag{17}$$

Substituting the expressions of $M_{x,x} + M_{xy,y}$ and $M_{y,y} + M_{xy,x}$ from equations of motion (11) and (12) into Eq. (17) and using Eq. (2), reduces it to exactly the same expression as in Eq. (13). Hence Ray's theory is not a new theory since its equations of motion and boundary conditions are mathematically equivalent to those of Reddy's theory. The results of this theory for any boundary conditions will be identical to those of Reddy's theory. The statics results of Ray's theory in Table 1 agree with Reddy's results, [2]. The difference in Table 6 from Reddy's results is due to neglect of some inertia terms by Ray while obtaining Navier's solution. Ray's theory is not a zeroth-order theory but Reddy's third order theory in disguise. Moreover, the displacement approximation of Ray's theory is valid only for the case of cross-ply and antisymmetric angle-ply laminates since for the general lay-up, the given expressions of λ_x, λ_y would not be valid.

References

[1] Ray, M. C., 2003, "Zeroth-Order Shear Deformation Theory for Laminated Composite Plates," ASME J. Appl. Mech., **70**, pp. 374–380.

[2] Reddy, J. N., 1997, *Mechanics of Laminated Composite Plates Theory and Analysis*, CRC Press, Boca Raton, FL.