Transactions of the ASME®

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

Journal of Applied Mechanics

Published Bimonthly by The American Society of Mechanical Engineers

VOLUME 69 • NUMBER 1 • JANUARY 2002

TECHNICAL PAPERS

BRIEF NOTES

BOOK REVIEW

ANNOUNCEMENTS AND SPECIAL NOTES

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.

A Theory of Fatigue: A Physical Approach With Application to Lead-Rich Solder

S. Wen
Mem. ASME

L. M. Keer
Life Fellow ASME

Department of Civil Engineering,
Northwestern University,
2145 Sheridan Road,
Evanston, IL 60201

A fatigue theory with its failure criterion based on physical damage mechanisms is presented for solders. The theory applies Mura's micromechanical fatigue model to individual grains of the solder structure. By introducing grain orientation (Schmid factor m) into the fatigue formula, an m-N curve at constant loading, similar to a fatigue S-N curve, is suggested for fatigue failure of grains with different orientations. A solder structure is defined as fatigued when the ratio of its failed grains reaches a critical threshold, since at this threshold the failed grains may form a cluster, according to percolation theory. Experimental data for 96.5Pb-3.5Sn (wt. %) solder bulk specimens showed good agreement with the theory and its associated failure criterion. The theory is anisotropic, and there is no size limitation to its application, which could be suitable for anisotropic small-scale (micron scale or smaller) solder joints. [DOI: 10.1115/1.1412453]

1 Introduction

Solder connections in electronic packaging are expected to become smaller and smaller. Today, the characteristic size of a solder joint in flip chip packaging is about a few hundred microns. According to the Semiconductor Industry Association's (SIA) roadmap ([1]), this size will decrease further into a few microns or even nanometer scale in future. It is well know that fatigue of such small joints is different from that of the bulk specimen. Therefore, the empirical fatigue formula derived from experimental data of bulk specimens, which are statistically isotropic, may not realistically predict life for such small joints. However, it is very difficult, if not impossible, to conduct fatigue tests on an individual small solder joint to derive an empirical fatigue formula. On the other hand, even if such tests are performed and an empirical fatigue formula derived as a result, the application of the formula to joints of different sizes is still questionable. A fatigue theory without size limitation is therefore needed, so that its fatigue parameters can be determined from tests of bulk specimens with the size effect built into the formula. To our knowledge, there has been no fatigue theory and its corresponding empirical formula that incorporates a size effect (e.g., see review article of the current fatigue theories for solders, Lee, et al. [2]). The theory presented here begins with fatigue behavior of an individual grain, which is caused by microcracking within its persistent slip bands (PSB). The PSB are formed by motion of slip planes, and these motions can occur only in particular directions, where such motion is related to the magnitude of local resolved shear stress on the slip planes. Therefore, resolved shear stress is the key parameter to characterize the fatigue behavior of an individual grain, and the fatigue of an individual grain is anisotropic at the micro scale. At the same time, the theory treats a bulk specimen as an agglomeration of such individually anisotropic grains. Statistically, such an agglomeration is isotropic, and the micro-scale anisotropy disappears. Therefore, the theory has the potential to connect the fatigue data from bulk specimen testing with the fatigue phenomena of meso-scale solder joints and the micro-scale fatigue of individual grains.

Solders in electronic packaging typically have low melting points, and room temperature can be a ''high'' temperature. For example, 25°C gives $T/T_m = 0.65$ for 63Sn-37Pb (wt. %) solder, where T_m is the absolute melting temperature. Fatigue phenomena of solders are thus very complicated. For example, the plastic strain based Coffin-Manson relationship does not hold for 96.5Pb-3.5Sn (wt. %) solder ([3]). Experimental results show that PSB form in the material when the device is under cyclic loading at the operating temperature range, and thereafter, microcracks form within PSB due to increment of dislocation density or along grain boundaries because of the impingement of the PSB on the grain boundaries. These microcracks are usually confined to be within the grains or along grain boundaries, and do not coalesce and form one dominant macrocrack that leads to fracture of the solder structure. Instead, the number of grains with such microcracks increases in a percolating manner, and the solder structure deteriorates, not necessarily showing a drop in load carrying ability until some critical point is reached. Figure 1 illustrates such a fatigue process. The solder specimen shows a load range increment until a certain point is reached. The specimen used in this test was made from 96.5Pb-3.5Sn (wt. %) solder and subjected to strain-controlled cyclic loading. After about 6800 cycles, the experiment was stopped and the surface of the specimen was investigated under an optical microscope. Figure 2 shows that the once very smooth surface at the onset of testing becomes full of macrocracks, microcracks, extrusions, and intrusions. Also the specimen can be seen to begin to lose its load-carrying ability during the final stage, in an accelerated way, although a certain amount of load could still be carried.

Given the experimental findings contained here and in other authors' publications ([4–8]), a fatigue theory based on dislocation and percolation damage mechanics is developed. The theory assumes the following: (1) Local resolved shear stress in the crystal slip plane causes PSB and thus the microcracks' formation within the PSB, and the magnitude of the stress differs for grains or cells of different crystallographic orientation; (2) Microcracks do not propagate but rather remain where they appear and fracture the grains or cells locally. The number of failed grains or cells increases within the solder structure; (3) The portion of such grains or cells reaches a threshold value at which point the entire structure becomes unstable. At this value the failed grains may form a large cluster or macrocrack.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, March 7, 2001; final revision, June 8, 2001. Associate Editor: M.-J. Pindera. Discussion on the paper should be addressed to the Editor, Prof. Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

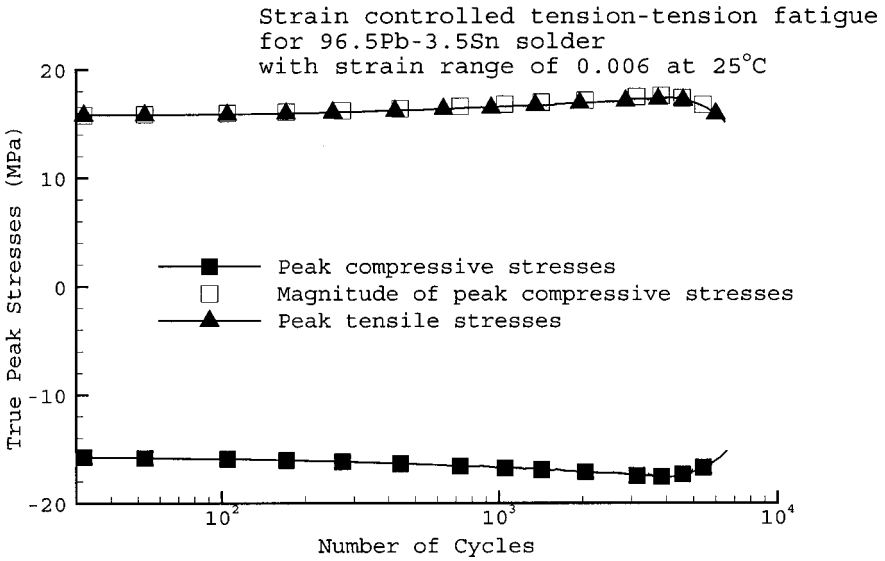


Fig. 1 Cyclic peak stresses plot for 96.5Pb-3.5Sn solder

In the following sections, fatigue of an individual grain is first discussed. A fatigue theory with its definition of fatigue criterion for solders is presented. Finally, the experimental results for 96.5Pb-3.5Sn (wt. %) solder bulk specimens are given as a verification of the theory.

2 Dislocations, PSB, and the Critical Number of Cycles for Microcrack Initiation

When ductile materials are subjected to cyclic loading, dislocations appear first and are followed by PSB formation within their grains. If the grains' facets are a free surface, so-called extrusions and intrusions will appear in the form of striation as the PSB strike the free surface. Foryth [9] first reported this phenomenon for an aluminum-copper alloy. Later, many researchers worked on a variety of other metals and have found that the phenomenon exists for most of the FCC, BCC, and some of the HCP metals. Experimental results from Vaynman [7] and Lawson [8] showed striation on the surface of fatigued 96.5Pb-3.5Sn solder specimens (Fig. 3). The striation was also found to appear on the fatigued tin-silver and tin-zinc eutectic solder specimens ([10]).

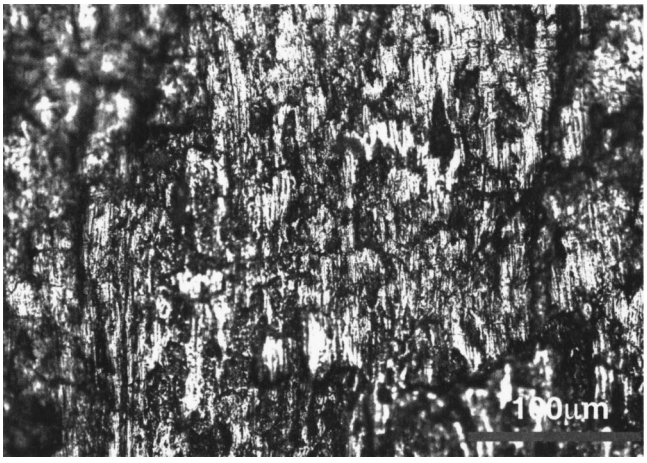


Fig. 2 Microcracks appeared on the surface of a 96.5Pb-3.5Sn solder specimen after about 6800 cycles under strain-controlled fatigue test (25°C, $\Delta\varepsilon=0.006$)

Lin and Ito suggested a ratcheting or gating mechanism for the formation of PSB ([11]). Using the same idea, Mura and coworkers ([12–14]) have further developed a micromechanical model to quantitatively analyze the microcracking process within the PSB using dislocation theory. In the Mura model, PSB formation results from dislocation density increment in two adjacent but reversely gliding slip layers during the cyclic loading. At the same time, change of the Gibb's free energy increases as the dislocation energy increases because of the dislocation increment. At a certain point, the Gibb's free energy change reaches a maximum and microcracking occurs within the PSB. This cycle number is defined as the fatigue point. The Mura model is able to produce the fatigue S-N curve, to capture the grain size effect and to incorporate material properties such as surface energy density, critical friction stresses, and others. Briefly, similar to the Griffith theory for crack initiation in linear elastic fracture mechanics (LEFM), the Mura model proposes fatigue microcrack initiation based on the Gibb's free energy change:

$$\Delta G = -W_1 - W_2 + 2c\gamma \tag{1a}$$

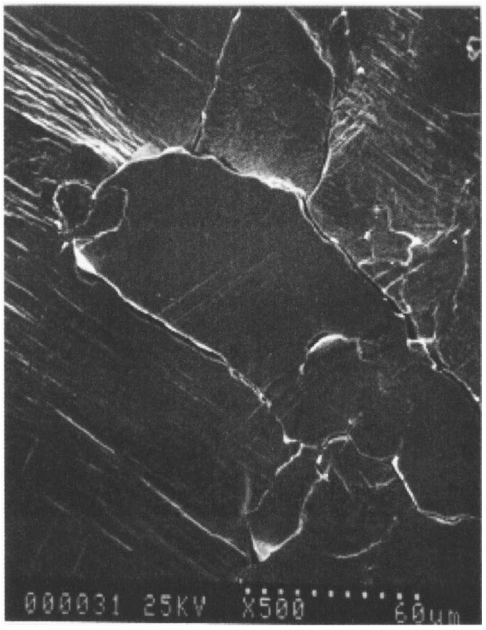
$$\frac{\partial}{\partial n}(\Delta G) = 0 \tag{1b}$$

where W_1 is the mechanical energy released; W_2 is the energy release from the loss of lattice defects—dislocations—at the nucleation site; γ is the free surface energy; c is the length of the initiated two-dimensional crack; and n is the cycle number. In ([15]), Fine expanded the above theory to a three-dimensional penny-shaped mode I crack in the study of fatigue at elevated temperature by rewriting Eq. (1*a*) as

$$\Delta G = -W_1 - W_2 + 2\gamma A \tag{1c}$$

where A is the surface area of the initiated three-dimensional crack. For solders, local unidirectional shear stress is the dominant factor to initiate a microcrack within a grain or along grain facets. The initiated microcrack is therefore a combined mode II and III type. For simplicity, assume the initiated microcrack is a penny-shaped three-dimensional microcrack. This mechanical energy released is given by integrating the energy released along with the crack growth as the following:

$$W_1 = \frac{64(1-\nu^2)\Delta\tau^2 a^3}{3(2-\nu)E} \tag{2}$$



(*a*)

(*b*)

Fig. 3 (*a*) Striations on the surface of a 96.5Pb-3.5Sn solder specimen after strain-controlled isothermal fatigue test (adopted from S. Vaynman's Ph.D. dissertation [7], Fig. 45); (*b*) microcracked grain after strain controlled thermomechanical fatigue test (adopted from L. Lawson's Ph.D. dissertation [8], Fig. 17)

where E is the Young's modulus, ν is the Poisson's ratio, $\Delta\tau$ is the shear stress range within the crack plane, and a is the radius of the formed microcrack. The energy released from the loss of lattice defects keeps Fine's form ([15]):

$$W_2 = 2\gamma_d A, \tag{3}$$

where γ_d is defined as dislocation energy per unit area and $A = \pi a^2$. Substituting (2) and (3) into (1*c*) gives

$$\Delta G = -\frac{64(1-\nu^2)\Delta\tau^2 a^3}{3(2-\nu)E} + 2\pi(\gamma-\gamma_d)a^2. \tag{4}$$

Following Mura et al. [12], the total number of dislocations measured along the diameter of the to-be-nucleated penny-shaped crack is

$$\overline{N} = \frac{2(1-\nu)(\Delta\tau - 2\tau_f)\overline{d}n}{\mu b} \tag{5a}$$

where τ_f is the friction stress, $\overline{d}$ is a characteristic grain size, b is Burger's vector, and μ is the shear modulus. The radius of the nucleated crack is therefore

$$a = \frac{1}{2}\overline{N}b = \frac{2(1-\nu^2)}{E}(\Delta\tau - 2\tau_f)\overline{d}n. \tag{5b}$$

Substitution of (5*b*) into (4) and the results into (1*b*) gives

$$n_{cr} = \frac{(2-\nu)\pi(\gamma-\gamma_d)E^2}{32(1-\nu^2)^2\overline{d}(\Delta\tau - 2\tau_f)\Delta\tau^2}. \tag{6}$$

Equation (6) gives the number of cycles at which a microcrack initiates within the PSB or on the grain facets. The dislocation energy density γ_d, grain size $\overline{d}$, and local resolved shear stress $\Delta\tau$ within the cracking plane are seen to strongly influence n_{cr}.

For solders, the critical friction stress is small compared to its operation stress loading. It is assumed henceforth that $\Delta\tau \gg 2\tau_f$ and τ_f can be neglected, i.e., $\Delta\tau - 2\tau_f \approx \Delta\tau$. Under this assumption, Eq. (6) becomes

$$n_{cr}\Delta\tau^3 \approx \frac{(2-\nu)\pi(\gamma-\gamma_d)E^2}{32(1-\nu^2)^2\overline{d}}. \tag{7a}$$

Introducing shear strength τ_y into Eq. (7*a*), we may rewrite

$$n_{cr}\left(\frac{\Delta\tau}{\tau_y}\right)^3 = \frac{(2-\nu)\pi}{32(1-\nu^2)^2}\frac{(\gamma-\gamma_d)E^2}{\overline{d}\tau_y^3}. \tag{7b}$$

The right-hand side of Eq. (7*b*) is a constant, and Eq. (7*b*) results from the assumption of a penny-shape crack loaded in unidirectional shear. Within a real solder structure, the geometry of microcracks varies and microcrack modes and energy release rate may not be as simple. To make Eq. (7*b*) more general, assume

$$n_{cr}\left(\frac{\Delta\tau}{\tau_y}\right)^{\eta} = C. \tag{7c}$$

Here, η and C are material constants. Equation (7*c*) gives the critical number of cycles for a grain to develop microcracks in its PSB or along its boundaries.

3 Local Resolved Shear Stress, Crystallographic Orientation, and Fatigue of an Individual Grain

Solders are materials with high homologous temperatures. Solder connections thus operate at a relatively high temperature range, which causes glide of the slip systems within its grains to become easier since the flow stress is low. Dislocations occur easier with the assistance of thermal activation, but at the same time so does the annihilation of dislocations and recovery. These factors help to confine the PSB to be within the grains. Lin et al. [16], studied the possibility of a fatigue band crossing a grain boundary and concluded that fatigue band is less likely to cross the grain boundary if the orientation of the neighboring grain differs by more than 5 deg. The bands were actually stopped from crossing the grain boundary when the misorientation is greater than 10 deg, an angle not uncommon between the grains within a real structure. It is also believed that at high homologous temperature, grain boundaries are obstacles to PSB, and as a result, PSB are confined to be within the grain. If microcracking occurs within

Fig. 4 The originally smooth surface of 96.5Pb-3.5Sn solder specimen now shows an agglomeration of extrusions, intrusions, striations of PSB and microcracks, with the pattern orienting at roughly 45 deg to the loading axis (vertical)

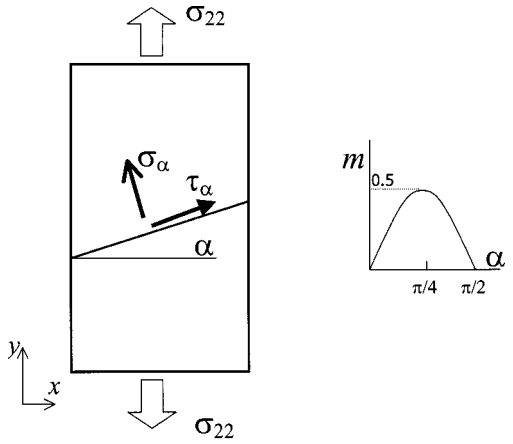


Fig. 5 Schmid factor *m* for a slip system within a crystal that undergoes uniaxial loading

the PSB, the newly created microcracks will also be confined to within the grain. If on the other hand a microcrack does not form within the PSB, the strain energy there is not released. Under this condition, the PSB impinge and produce microcracking along grain boundaries. If the grain boundaries are a free surface, then extrusion-intrusion will occur. Figure 4 shows the originally smooth surface of a high lead solder at the end of a strain-controlled fatigue test. The experiment was carried out at room temperature ($T/T_m \approx 0.5$). The surface shows an agglomeration of extrusions, intrusions, striations of PSB and microcracks. The agglomerates pattern also can be seen to orient at roughly 45 deg to the loading axis (vertical), which is the maximum shear stress direction. There are many more experimental findings ([7,8,10]), all of which suggest that microcracking within solders is a localized phenomenon, and occurs within a grain and/or along grain boundaries. According to Eq. (7*c*), the shear stress range has a strong influence on the number of cycles required to initiate a microcrack within the PSB. The local resolved shear stress range on the grain's active slip systems will then be investigated to quantify the microcracking within the PSB of an individual grain.

When a grain is subjected to a far field stress, the resolved shear stresses on its slip systems vary. For example, an FCC crystal has 12 active slip systems ([17]). Under certain loading conditions, the resolved shear stresses on these slip systems could differ substantially (see Table 1 in Lin [18] and Example (2.3.1) in Suresh [19]). A slip system consists of a glide plane and a slip direction. Two vectors can thus characterize the system: $\mathbf{n}_g$, the normal to the glide plane, and $\mathbf{n}_s$, the slip direction. Assume that the system undergoes loading $\boldsymbol{\sigma}$, the resolved shear stress τ on the system is

$$\tau = \mathbf{n}_s \cdot \boldsymbol{\sigma} \cdot \mathbf{n}_g . \tag{8}$$

It is well known that hydrostatic stress does not contribute to the resolved shear stress. Equation (8) can be rewritten as follows:

$$\tau = \mathbf{n}_s \cdot \left(\mathbf{S} + \frac{1}{3}\sigma_{kk}\mathbf{I} \right) \cdot \mathbf{n}_g = \mathbf{n}_s \cdot \mathbf{S} \cdot \mathbf{n}_g = m\|\mathbf{S}\| \tag{9}$$

where $\mathbf{S}$ is the deviatoric stress: $\mathbf{S} = \boldsymbol{\sigma} - \frac{1}{3}\mathrm{tr}(\boldsymbol{\sigma})\mathbf{I}$; m is the Schmid factor; and $\|\mathrm{o}\|$ is a form of norm. When a crystal undergoes uniaxial loading, the Schmid factor m of a slip system varies between 0.0 and 0.5. (If the direction of σ_α, τ_α, and σ_{22} lie in the same plane, then $\tau_\alpha = 0.5 \sin 2\alpha \sigma_{22} = m\sigma_{22}$, Fig. 5). Dingli [20] used a second-order Schmid tensor, $\mathbf{m}$, to get the resolved shear stress:

$$\tau = \mathbf{m} : \boldsymbol{\sigma}. \tag{10}$$

The resolved shear stress range $\Delta\tau$ under cyclic loading condition is given by the following:

$$\Delta\tau = m\|\mathbf{S}_2 - \mathbf{S}_1\|. \tag{11a}$$

Here subscripts "1" and "2" refer to the two extremes of the stress loading space. It is assumed that m does not change during cyclic loading. Under one-dimensional cyclic loading, the resolved shear stress range can be written as follows:

$$\Delta\tau = m\|\boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1\|. \tag{11b}$$

Subscripts "1" and "2" correspond to the valley and peak points of cyclic loading, respectively. If the load ratio is $R = \sigma_1/\sigma_2 \in [-1,1]$, then

$$\Delta\tau = m(1-R)|\boldsymbol{\sigma}_2|. \tag{12}$$

Substituting Eq. (11*b*) into Eq. (7*c*),

$$n_{cr} m^\eta \left(\frac{\|\boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1\|}{\tau_y} \right)^\eta = C. \tag{13a}$$

In general,

$$n_{cr} m^\eta \left(\frac{\|\mathbf{S}_2 - \mathbf{S}_1\|}{\tau_y} \right)^\eta = C. \tag{13b}$$

Further, under the same loading condition, i.e., $\|\boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1\|$ remains constant:

$$n_{cr} m^\eta = C \left(\frac{\|\boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1\|}{\tau_y} \right)^{-\eta} = \mathbf{C}_1 . \tag{14}$$

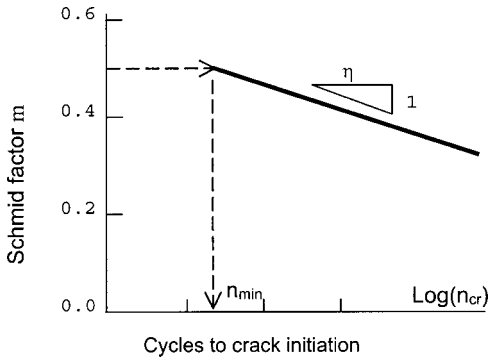


Fig. 6 The Schmid factor and the critical number of cycles to initiate a microcrack: m-N curve

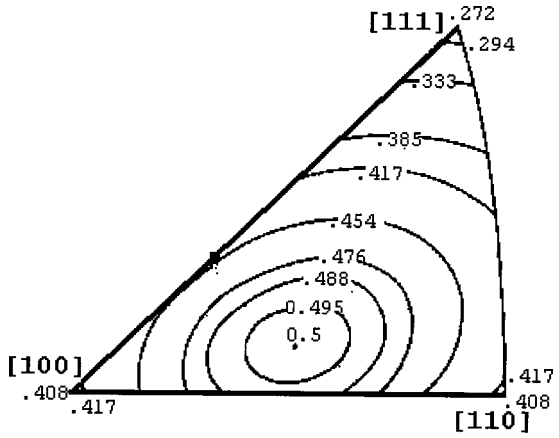


Fig. 7 Contours of a constant Schmid factor for uniaxial tension based on {111}⟨110⟩ slip (reproduced with modification from Fig. 2 ([21]))

Equation (14) indicates that the Schmid factor m has a decisive effect on the critical cycle number n_{cr} for microcrack initiation within the PSB when the loading remains constant. If plotted, the equation gives an m-N curve under constant loading (Fig. 6). The curve can be interpreted as similar to a S-N curve, except that the Schmid factor m replaces the stress. According to Fig. 6, under a given load condition, if a single crystal orients itself in such a way that its $\mathbf{n}_g$ and $\mathbf{n}_s$ for the major active slip systems make their Schmid factors m a minimum, the crystal will have the longest fatigue life. As an example, Fig. 7 shows the Schmid factor contour for a FCC crystal under uniaxial tension based on {111}⟨110⟩ slip system (reproduced with modification from Fig. 2 [21]). It can be seen that the Schmid factor varies from 0.272 to 0.5, depending on the crystal's orientation towards the loading axis. If $\eta=3$, the critical cycle number to initiate a crack within the grain could differ by as much as six times by changing its orientation. Yue [22] showed the fatigue striations, microcracks in these striations, and their dependence on the crystallographic orientations (equivalently, the Schmid factor m) by performing experiments on nickel-base single crystal superalloy specimens. Figure 3 shows two SEM photos of the surfaces of fatigued solder specimens (reproduced from Vaynman [7] and Lawson [8]). Closely packed striations on the surface are evident. Striations are traces of slip plane motion, and striation directions demonstrate the motion directions. The photos show clearly that striation directions differ for different grains. Other authors ([23–26]) also reported cyclic behavior's dependence on crystallographic orientation.

A large enough polycrystalline structure consists of numerous crystallites or grains. Individual grains are oriented in different directions and thus have a different Schmid factor m, which leads to microcracking to occur at a different number of cycles according to Eq. (14). It is assumed that the newly nucleated microcrack or family of microcracks extends immediately to the extremities of the grain. The grain is thus defined to be fatigued or failed at this point. This assumption is reasonable for solders since its application is usually associated with small size grains and a microcrack extends to the extremities very quickly.

4 Fatigue Theory for Solders

Currently, there are three types of fatigue theories for solders: strain based, energy density based, and damage based ([2,8]). The goal of these theories is to derive empirical fatigue formula from the testing of bulk specimens and to use the derived formula to predict fatigue life of small solder joints. However, micromechanical research shows that the representative size that such formulas can apply is larger than even the size of current solder joints ([27,28]), and the anisotropic nature of actual solder joints has not been included. Moreover, the fatigue criteria are more or less arbitrarily defined without a physical damage based, and thus will differ from application to application. In the following subsections, a physical damage based fatigue criterion is presented, followed by the formulation of the fatigue theory.

4.1 Fatigue Criterion for Solders. A strain-controlled test is typically used to study fatigue of solders, and the peak stress change, peak stress range change, or the area enclosed by stress-strain hysteresis loop change is recorded. The fatigue criterion is thus defined by the value of these changes. Figures 1 and 8 show the peak stress curves of two solders, and Fig. 9 sketches general cases of such curves. In Fig. 9, curve A is a representative shape for eutectic PbSn solders, while curves B and C are for Pb rich PbSn solders. It can be seen that a specific solder can exhibit either cyclic hardening or softening behavior or both under different loading conditions. If the fatigue criterion is defined as a certain value of peak stress drop, it is not applicable to the cases of cyclic hardening (Curve B and C) since no drop occurs in these

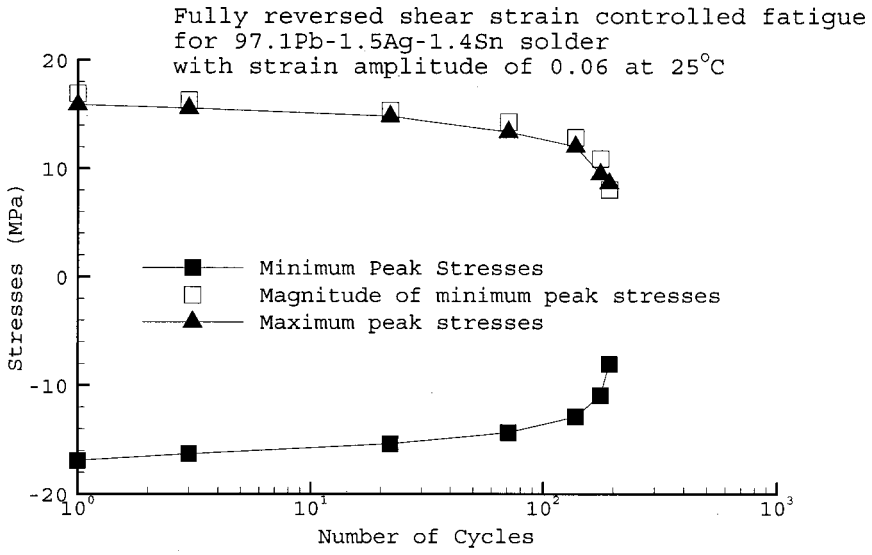


Fig. 8 Peak shear stresses change with number of cycles for silver modified PbSn solder with a strain rate of 0.003/sec (data taken from J. Liang et al., Fig. 7(*a*)([38]))

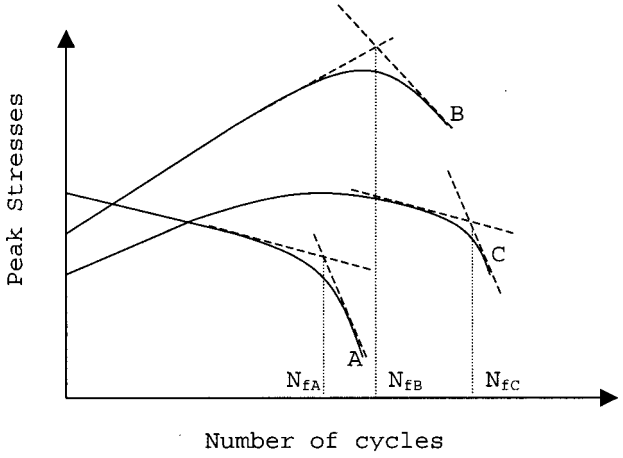


Fig. 9 Peak stresses evolution during fatigue testing and definition of fatigue point

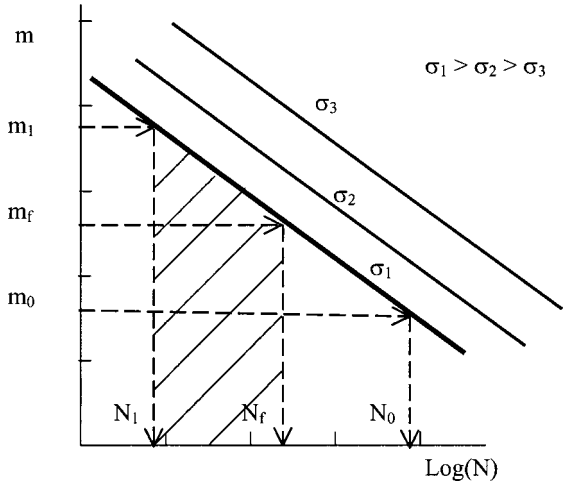


Fig. 10 Fatigue process illustration when stresses are assumed uniform throughout the structure

cases. On the other hand, if the fatigue criterion is defined as the onset point of peak stress drop during cyclic hardening, it is definitely not applicable to the cyclic softening case (Curve A) since the onset point is near the beginning of the test. Moreover, in the case of curve C, there are two onset peak stress drop points, and an arbitrary pick of one is inevitable. Clearly, a fatigue criterion that is reasonably applicable to all cases is needed. In this paper, the fatigue point is defined as the onset point where a sharp peak stress drop begins. Two straight lines tangential to the curve segments on each side of the onset point are drawn, and the abscissa of the crossing point of these two straight lines is defined as the fatigue point. According to this fatigue criterion, points N_{fA}, N_{fB}, and N_{fC} are the fatigue points for curves A, B, and C, respectively (Fig. 9). This criterion has as its physical basis on the percolation damage mechanism. As discussed in Section 1, solders develop microcracks during the cyclic process. These microcracks either appear in the PSB within the grains or along gain boundaries. According to percolation theory, a cluster of microcracked grains or macrocracks may form and the material becomes unstable when the portion of these microcracked grains or the density of microcracks reaches the percolation threshold. The onset of a sharp peak stress drop is the initiation of such an unstable state. Stolkarts et al. [6] and Guo et al. [29] applied a similar fatigue criterion to a near eutectic PbSn solder with successful results. It must be further pointed out that this criterion is conservative.

4.2 Fatigue Theory for Solders. Solder is a polycrystalline material and a solder structure is an aggregation of crystallites or grains of different sizes and orientations. A bulk solder specimen consists of numerous crystallites or grains with their orientations randomly distributed, while small-scale solder joints consist of limited number of grains and the orientations of these grains exhibit a preferred distribution. In Section 3 it was noted that the orientation difference results in the Schmid factor difference if the major loadings or stresses on these grains fall within a similar range, leading to a difference of critical number of cycles of fatigue or failure among grains. Grains with larger m fail at a lower number of cycles, while those with lower m fail at a higher number of cycles. Along with the cyclic loading process, more and more grains become fatigued or failed due to microcracking. Since the failure of grains is a highly localized phenomenon, the fatigue of a structure can be viewed as the result of a series of such localized and independent failure events.

Let $V(m)$ be the total number of those grains that have Schmid factor m, and $f(m)$ be the grain Schmid distribution function (GSDF) or grain orientation distribution function (GODF). Then $f(m)$ is defined as follows:

$$f(m)dm = \frac{dV(m)}{V} \tag{15}$$

where V is the total number of grains, and $m \in [m_0, m_1]$. Here m_0 and m_1 are the minimum and the maximum Schmid factors, respectively. If it is further assumed that each grain has the same volume, then $dV(m)$ and V can be the total volume of grains with Schmid factor m and the total volume of the structure. For a structure with numerous grains, $f(m)$ is assumed to be continuous for simplicity in analysis.

According to Eq. (13*a*) and the percolation theory based fatigue criterion defined previously, the fatigue theory can be stated as follows:

$$N_f m_f^{\eta} \left(\frac{|\sigma_2 - \sigma_1|}{\tau_y} \right)^{\eta} = C \tag{16}$$

with

$$\int_{m_f}^{m_1} f(m)dm = p_c . \tag{17}$$

Here, according to percolation theory, p_c is the percolation threshold. Figure 10 is a sketch that describes the theory, where the shaded area depicts the fatigue process. The value of p_c depends on the crystal type and on the percolation model chosen to describe the failure mode. For solders, if intergranular microcracking dominates, bond percolation would be more appropriate. However, if microcracking occurs predominantly within the grains, site percolation would be the best model to use. For example, high lead solder can be modeled as a FCC structure with site percolation, and $p_c = 0.3116$ ([30]).

If $f(m)$ is known, Eq. (17) is easy to use by integrating and solving for m_f. Bulk specimens are statistically homogenous and isotropic, and it is therefore reasonable to assume

$$f(m) = \frac{1}{m_1 - m_0} . \tag{18}$$

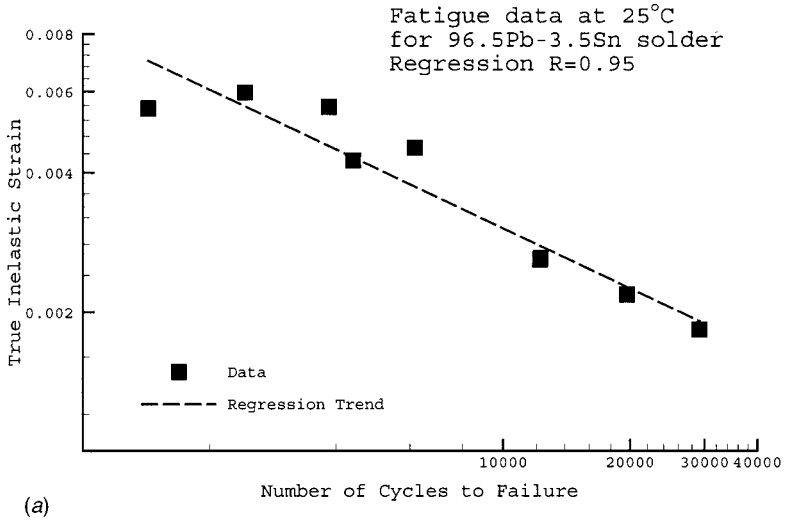
Substituting Eq. (18) into Eq. (17), integrating and solving for m_f gives

$$m_f = p_c m_0 + (1 - p_c) m_1 . \tag{19}$$

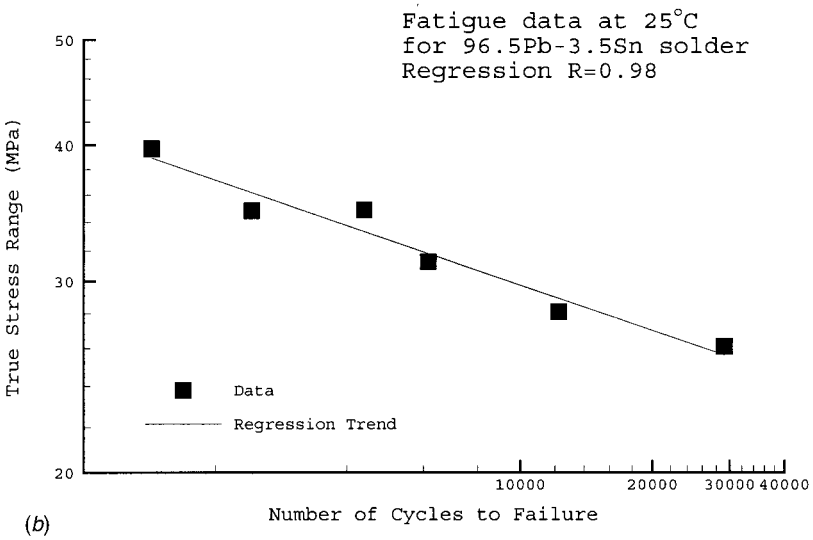
Further substitution of Eq. (19) into Eq. (16) gives the fatigue life N_f as

$$N_f = \frac{C}{(p_c m_0 + (1 - p_c) m_1)^{\eta}} \left(\frac{\tau_y}{\|\sigma_2 - \sigma_1\|} \right)^{\eta} . \tag{20}$$

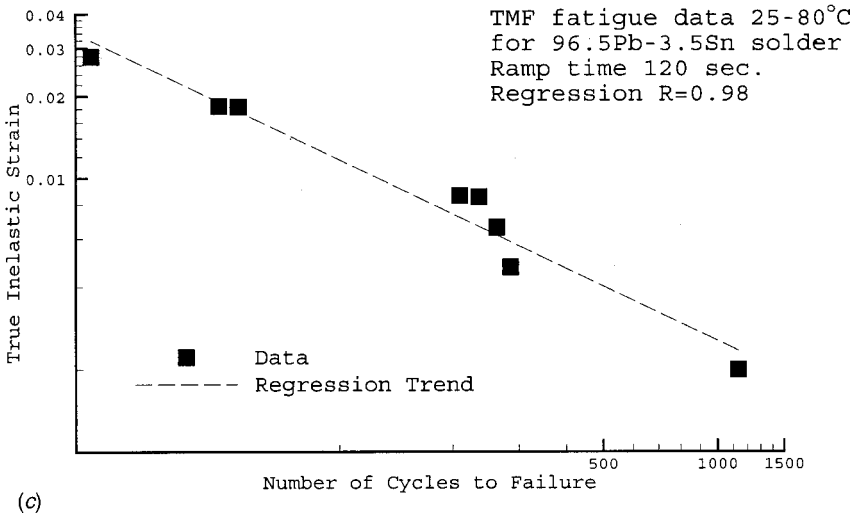
Equations (18)–(20) are not applicable for meso/micro-scale structures that are inhomogeneous and/or anisotropic, since the



(*a*)



(*b*)



(*c*)

Fig. 11 (*a*) **Fatigue and inelastic strain under strain-controlled isothermal fatigue test (ε_{min}=0, ramp=2.5 s, partial data from S. Vaynman); (*b*) fatigue and stresses under strain-controlled isothermal fatigue test (ε_{min}=0, ramp=2.5 s, partial data from S. Vaynman); (*c*) fatigue and inelastic strain under strain-controlled thermomechanical fatigue testing—conducted by L. Lawson during 1987–1989; (*d*) fatigue and stresses range under strain-controlled thermomechanical fatigue testing—conducted by L. Lawson during 1987–1989; (*e*) fatigue and inelastic strain under isothermal and thermomechanical fatigue testing—conducted by L. Lawson during 1987–1989 (temperatures are 15–60°C, 25–80°C, 60°C, 80°C, and 100°C; strain ranges from 0.3–3 percent; strain rate from $1.15\times10^{-5}\sim3.0\times10^{-3}$ s^{-1}); (*f*) fatigue and true stresses under isothermal and thermomechanical fatigue testing—conducted by L. Lawson during 1987–1989 (temperatures are 15–60°C, 25–80°C, 60°C, 80°C, and 100°C; strain ranges from 0.3–3 percent; strain rate from $1.15\times10^{-5}\sim3.0\times10^{-3}$ s^{-1})**

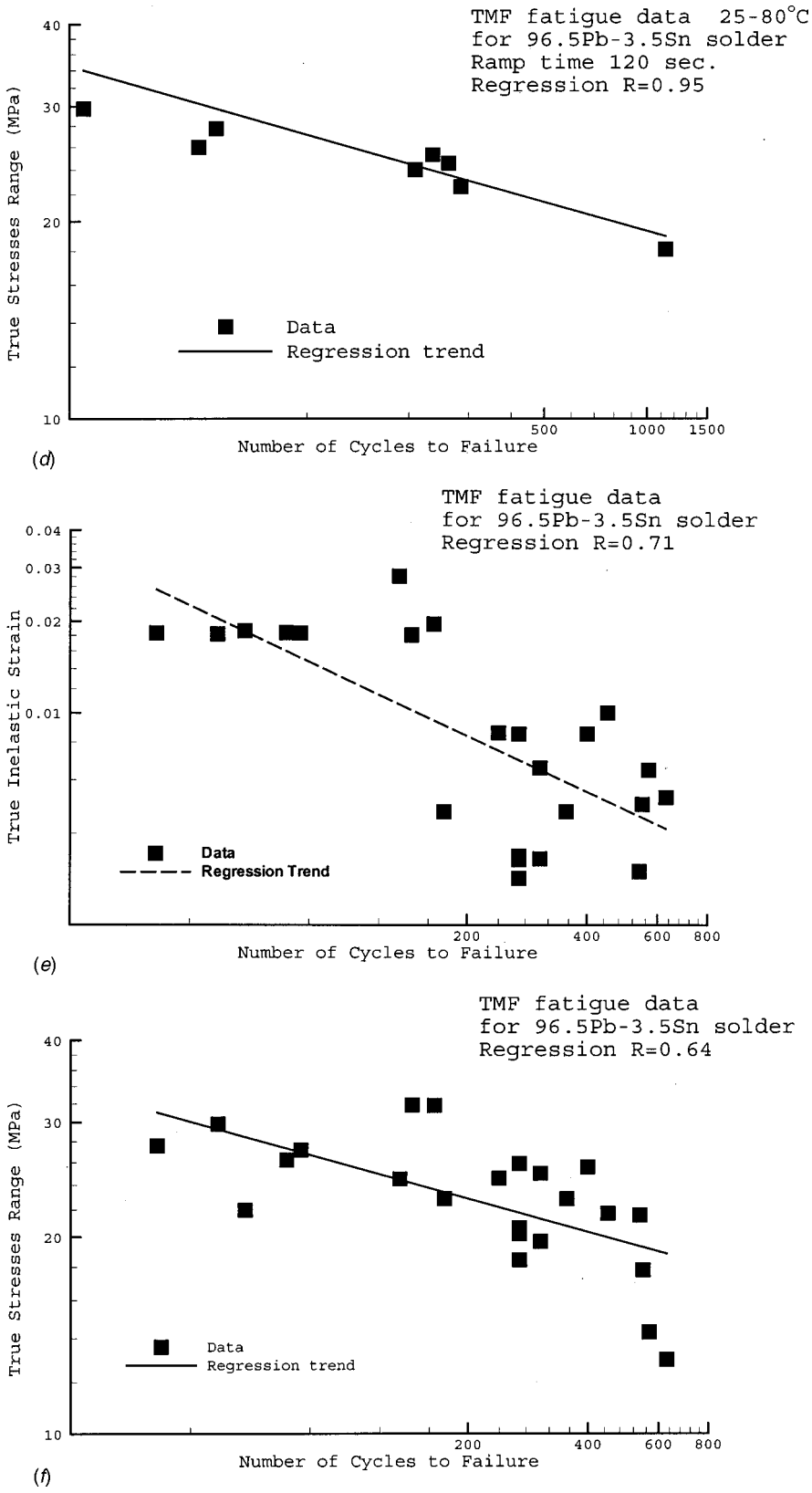


Fig. 11 Continued

structure consists of only limited number of grains and $f(m)$ is not continuous. Equation (17) must be used to solve for m_f in this case, and the anisotropy must be brought into the fatigue formula. It is not too difficult to get $f(m)$ for these applications: Theoretical study, experimental measurement reconstruction, and computational simulation have been used to obtain the grain orientation distribution function for polycrystalline structures for various applications ([31–34]).

Table 1 The coefficients of the $\log(\Delta\sigma)\sim\log N_f$ plots ($\log(\Delta\sigma)=a+b\log N_f$) and the strain-based Coffin-Manson type $\log(\varepsilon^{in})\sim\log N_f$ plots ($\log(\varepsilon^{in})=c+d\log N_f$)

Experimental Data		This Theory		Strain-based Coffin-Manson	
		a	b	c	d
S. Vaynman during 1985-1987, together with this study, for 25°C		2.03	-0.14	-0.8	-0.43
L. Lawson During 1987-1989	25-80°C	1.73	-0.15	-0.61	-0.66
	TMF[1] & ISO[2] mix	1.75	-0.17	-0.66	-0.61

[1]Thermomechanical Fatigue testing
[2]Isothermal fatigue testing

Equation (20) also shows that the fatigue life can be optimized by controlling the grain orientations of a structure during the metallurgical process. Indeed, the texture and grain orientations can be controlled for casting and deposited structures ([35]), especially in today's electronic industry where structures such as films, and solder joints are cast or deposited.

5 Application to 96.5Pb-3.5Sn Solder

The isothermal and thermomechanical fatigue data of a lead rich 96.5Pb-3.5Sn solder were used to verify the fatigue theory and its associated fatigue criterion. Bulk specimens were tested under uniaxial tension-tension strain-controlled conditions with triangular waveform strain. Most of the experiments were conducted during the period from 1985 to 1989 by S. Vaynman and L. Lawson at Northwestern University. The raw data are in analog form from paper charts taken from an X-Y recorder. Additional fatigue tests were performed using digital data acquisition during this study. The paper charts data and their digital counterparts were found to agree quite well. The specimen preparation, geometry, machining, pre-testing treatment, testing procedures, and testing set-ups are reported in publications by S. Vaynman and L. Lawson [7,8,36].

Lead is a FCC crystal, and the grains of 96.5Pb-3.5Sn solder can be approximately treated as FCC crystal. According to Fig. 7, the Schmid factor varies from 0.272 to 0.5, or $m_0=0.272$ and $m_1=0.5$, depending on the grain's orientation towards the loading axis. The grains can be approximated as having cubic shape and a simple cubic structure under site percolation damage model can be used to describe the fatigue of the 96.5Pb-3.5Sn solder bulk specimen. In this case, $P_c=0.3116$ ([30]). Substituting these coefficients into Eq. (19), m_f can be calculated as $m_f=0.429$.

As predicted by Eq. (13), the $\log(\Delta\sigma)\sim\log(N_f)$ plots should be straight lines since $m_f=0.429$ is a constant for all bulk specimens. Figures 11(*b*), (*d*), and (*f*) are these plots, and they do show that straight lines can capture the trend reasonably well under various conditions: Fig. 11(*b*) for isothermal results at 25 °C, Fig. 11(*c*) for thermomechanical results, and Fig. 11(*f*) for both isothermal and thermomechanical results. Figures 11(*b*) and (*d*) fit the experimental data quite well. As comparison, Figs. 11(*a*), (*c*), and (*e*) give the corresponding $\log\varepsilon^{in}\sim\log N_f$ plots, which would be used to fit the data if an inelastic strain-based Coffin-Manson type formula was employed. It can be seen that straight lines can fit the data as well. However, it is unclear what is the science base to apply the fitting results to conditions that are different from the experimental conditions and what the inelastic strain-based Coffin-Manson formula would behave in these conditions. Furthermore, Tables 1 gives the fitting results from Figs. 11(*a*)–(*f*). It is interesting to see that the coefficients from Figs. 11(*b*), (*d*) and (*f*) are surprisingly close even though they are from fitting the data from different experimental conditions. It is thus believed that these coefficients may represent a broad range of test conditions, which is what the theory anticipates as it is based on physical damage mechanism. However, the coefficients from Figs. 11(*a*), (*c*), and (*e*) are not as close. Especially, the slopes differ as much as 35 percent. Note that a single strain-based Coffin-Manson type formula does not fit fatigue data well for both 96.5Pb-3.5Sn and 63Sn-37Pb solders ([7,29]).

The coefficients C and η can be solved as

$$\begin{cases} C=10^{\eta(a+\log m_f)} \\ \eta=-\dfrac{1}{b} \end{cases} . \tag{21}$$

If a and b are taken as the average values from Table 1 for 96.5Pb-3.5Sn solder, $\tau_y=11.7$MPa ([37]), through Eq. (21), C and η can be calculated out: $C=4.12\times10^2$, and $\eta=6.52$. These two coefficients can then be used in Eqs. (16) and (20) to predict fatigue life for 96.5Pb-3.5Sn solder structure. Note that for different size structures, the theory uses a different percolation model to determine $f(m)$ and m_f. Thus, the theory does not have a size limitation and Eq. (17) represents the anisotropic nature that should be included in the fatigue prediction formulas for current and future small-scale structures. Metallurgical control during the manufacturing process of solder interconnects may also be used to optimize fatigue life by possibly changing the grain orientation or size distribution.

6 Conclusion

A fatigue theory for solders is presented. By adopting Mura's dislocation energy-assisted microcrack formation theory, crystallographic orientation is connected to fatigue failure through resolved shear stresses, and an m-N curve under constant loading is suggested. The theory views the fatigue process as a series of local failure events, while the fatigue of the structure is viewed as the percolation result of such local failure events forming a large cluster. The theory includes the anisotropic nature into the fatigue formula and thus can handle the anisotropic small-scale structures as well as the statistically homogeneous large structures. The fatigue data of 96.5Pb-3.5Sn solder bulk specimens under various uniaxial tension tests were analyzed. Results show that the theory gives good predictions under broad testing conditions. More importantly, the theory is materials science based so that the parameters of the fatigue formula can be worked out by testing of bulk specimens while the formula can be applicable to small structures. The theory suggests metallurgical control during the manufacturing process to optimize the fatigue life of small-scale structures.

Acknowledgments

The authors would like to thank Prof. Morris E. Fine and Dr. Sidharth Sidharth for their reading of the manuscript and very helpful discussions and James Hayward for his suggestions. We also appreciate Drs. Semyon Vaynman and Lawrence R. Lawson

for allowing us to use their experimental data. The financial support from Advanced Micro Devices, Inc., under contract No. 99-NJ-654 with the Semiconductor Research Corporation, is gratefully acknowledged.

References

[1] SEMATECH, 1999, *The International Technology Roadmap for Semiconductors: 1999 Edition*, International SEMATECH, Austin, TX.
[2] Lee, W. W., Nguyen, L. T., and Selvaduray, G. S., 2000, ''Solder Joint Fatigue Models: Review and Applicability to Chip Scale Packages,'' Microelectron. Reliab., **40**, pp. 231–244.
[3] Vaynman, S., and Zubelewicz, A., 1990, ''Fatigue Life Prediction For Low-Tin Lead-Based Solder At Low Strains,'' Weld. J. (Miami), **69**, pp. S395–S398.
[4] Frear, D. R., Grivas, D., and Morris, J. W., 1988, ''Thermal Fatigue in Solder Joints,'' Journal of Metals **40**, pp. 18–22.
[5] Rathore, H. S., Yih, R. C., and Edenfeld, A. R., 1973, ''Fatigue Behavior of Solders Used in Flip-Chip Technology,'' J. Test. Eval., **1**, pp. 170–178.
[6] Stolkarts, V., Keer, L. M., and Fine, M. E., 1999, ''Damage Evolution Governed by Microcrack Nucleation With Application to the Fatigue of 63Sn-37Pb Solder,'' J. Mech. Phys. Solids, **47**, pp. 2451–2468.
[7] Vaynman, S., 1987, ''Isothermal Fatigue of 96.5Pb-3.5Sn Solder,'' Ph.D. dissertation, Northwestern University, Evanston, IL.
[8] Lawson, L. R., 1989, ''Thermomechanical Fatigue of 97Pb-3Sn,'' Ph.D. dissertation, Northwestern University, Evanston, IL.
[9] Forsyth, P. J. E., 1953, ''Exudation of Material From Slip Bands at the Surface of Fatigued Crystals of an Aluminum-Copper Alloy,'' Nature (London), **171**, pp. 172–173.
[10] Mavoori, H., 1996, ''Mechanical Properties and Fatigue Lifetime Prediction of Solders for Electronic Applications: Tin-Silver and Tin-Zinc Eutectics,'' Ph.D. dissertation, Northwestern Universty, Evanston, IL.
[11] Lin, T. H., and Ito, Y. M., 1969, ''Micromechanics of a Fatigue Crack Nucleation Mechanism,'' J. Mech. Phys. Solids, **17**, pp. 511–523.
[12] Mura, T., and Nakasone, Y., 1990, ''A Theory of Fatigue Crack Initiation in Solids,'' ASME J. Appl. Mech., **57**, pp. 1–6.
[13] Mura, T., 1994, ''A Theory of Fatigue-Crack Initiation,'' Mater. Sci. Eng., A, **176**, pp. 61–70.
[14] Shodja, H. M., Hirose, Y., and Mura, T., 1996, ''Intergranular Crack Nucleation in Bicrystalline Materials Under Fatigue,'' ASME J. Appl. Mech., **63**, pp. 788–795.
[15] Fine, M. E., 2000, ''Phase Transformation Theory Applied to Elevated Temperature Fatigue,'' Scr. Mater., **42**, pp. 1007–1012.
[16] Lin, T. H., Wong, K. K. F., Teng, N. J., and Lin, S. R., 1998, ''Micromechanic Analysis of Fatigue Band Crossing Grain Boundary,'' Mater. Sci. Eng., A, **246**, pp. 169–179.
[17] Hirth, J. P., and Lothe, J., 1982, *Theory of Dislocations*, 2nd Ed., John Wiley and Sons, New York.
[18] Lin, T. H., 1992, ''Micromechanics of Crack Initiation in High-Cycle Fatigue,'' *Advances in Applied Mechanics*, Vol. 29, Academic Press, New York, pp. 1–62.
[19] Suresh, S., 1998, *Fatigue of Materials*, 2nd Ed., Cambridge University Press, Cambridge, U.K.
[20] Dingli, J. P., Abdul-Latif, A., and Saanouni, K., 2000, ''Predictions of the Complex Cyclic Behavior of Polycrystals Using a Self-Consistent Modeling,'' Int. J. Plast. **16**, pp. 411–437.
[21] Chin, G. Y., 1973, ''The Role of Preferred Orientation in Plastic Deformation,'' *Inhomogeneity of Plastic Deformation*, ASM, Metals Park, OH, pp. 83–111.
[22] Yue, Z. F., Tao, X. D., Ying, Z. Y., and Li, H. Y., 2000, ''A Crystallographic Model for the Orientation Dependence of Low Cyclic Fatigue Property of a Nickel-Base Single Crystal Superalloy,'' Appl. Math. Mech., **21**, pp. 415–424 (English Edition).
[23] Zhang, Z. F., and Wang, Z. G., 1998, ''Effect of Component Crystal Orientations on the Cyclic Stress-Strain Behavior of Copper Bicrystals,'' Mater. Sci. Eng., A, **255**, pp. 148–153.
[24] Henderson, M. B., and Martin, J. W., 1996, ''The Influence of Crystal Orientation on the High Temperature Fatigue Crack Growth of a Ni-Based Single Crystal Superalloy,'' Acta Mater., **44**, pp. 111–126.
[25] Li, X. W., Wang, Z. G., and Li, S. X., 1999, ''Influence of Crystallographic Orientation on Cyclic Strain-Hardening Behaviour of Copper Single Crystals,'' Philos. Mag. Lett., **79**, pp. 869–875.
[26] Tan, X., Gu, H., Laird, C., and Munroe, N. D. H., 1998, ''Cyclic Deformation Behavior of High-Purity Titanium Single Crystals: Part I. Orientation Dependence of Stress-Strain Response,'' Metall. Mater. Trans. A, **29**, pp. 507–512.
[27] Bonda, N. R., and Noyan, I. C., 1996, ''Effect of the Specimen Size in Predicting the Mechanical Properties of PbSn Solder Alloys,'' IEEE Trans. Compon. Pack. Manufact. Tech., Part A, **19**, pp. 208–212.
[28] Bonda, N. R., and Noyan, I. C., 1992, ''Deformation Inhomogeneity and Representative Volume in Pb/Sn Solder Alloys,'' Metall. Trans. A, **23**, pp. 479–484.
[29] Guo, Q., Cutiongco, E. C., Keer, L. M., and Fine, M. E., 1992, ''Thermomechanical Fatigue Life Prediction of 63Sn/37Pb Solder,'' ASME J. Electron. Packag., **114**, pp. 145–151.
[30] Stauffer, D., and Aharony, A., 1992, *Introduction to Percolation Theory*, 2nd Ed., Taylor & Francis, Washington, DC.
[31] Adams, B. L., Boehler, J. P., Guidi, M., and Onat, E. T., 1992, ''Group Theory and Representation of Microstructure and Mechanical Behavior of Polycrystals,'' J. Mech. Phys. Solids, **40**, pp. 723–737.
[32] Bunge, H. J., 1987, ''Three-Dimensional Texture Analysis,'' Int. Mater. Rev., **32**, pp. 265–291.
[33] Park, N. J., and Bunge, H. J., 1990, ''Determination of the Orientation Distribution Function of a Cuznal Shape Memory Alloy,'' Z. Metallkd., **81**, pp. 636–645.
[34] Kumar, S., Kurtz, S. K., and Agarwala, V. K., 1996, ''Micro-Stress Distribution Within Polycrystalline Aggregate,'' Acta Mech., **114**, pp. 203–216.
[35] Barrett, C. S., and Massalski, T. B., 1980, *Structure of Metals: Crystallographic Methods, Principles and Data*, 3rd Rev. Ed., Pergamon, New York.
[36] Lawson, L. R., 1987, ''Thermal Cycling Apparatus For Thermomechanical Fatigue Testing,'' Rev. Sci. Instrum., **58**, pp. 1942–1944.
[37] Marinescu, G. M., Chen, Y. Y. T., Li, H. H., Beaumont, M., Chen, J. C. F., and Ho, C. Y., 1999, ''Thermal, Mechanical, Electrical, and Physical Properties of Selected Packaging Materials,'' Report No. CINDAS Report 126, Center for Information and Numerical Data Analysis and Synthesis (CINDAS) of Purdue University, West Lafayette, IN.
[38] Liang, J., Gollhardt, N., Lee, P. S., Schroeder, S. A., and Morris, W. L., 1996, ''A Study of Fatigue and Creep Behavior of Four High Temperature Solders,'' Fatigue Fract. Eng. Mater. Struct., **19**, pp. 1401–1409.

Scaling of Sea Ice Fracture—Part I: Vertical Penetration

Z. P. Bažant
Walter P. Murphy
Professor of Civil Engineering
and Materials Science,
Northwestern University,
Evanston, IL 60208
e-mail: bazant@northwestern.edu
Fellow ASME

Based on the premise that large-scale failure of sea ice is governed by fracture mechanics, recently validated by Dempsey's in situ tests of fracture specimens of a record-breaking size, this two-part study applies fracture mechanics and asymptotic approach to obtain approximate explicit formulas for the size effect in two fundamental problems. In the present Part I, the load capacity of a floating ice plate subjected to vertical load is determined, and in Part II, which follows, the horizontal force exerted by an ice plate moving against a fixed structure is analyzed in a similar manner. The resulting formulas for vertical loading agree with previous sophisticated numerical fracture simulations as well with the limited field tests of vertical penetration that exist. The results contrast with the classical predictions of material strength or plasticity theories, which in general exhibit no size effect on the nominal strength of the structure. [DOI: 10.1115/1.1429932]

1 Introduction

Predictions of load capacity and failure of floating sea ice require good understanding of the scaling properties and size effect. Because small-scale laboratory tests of sea ice show hardly any notch sensitivity and do not exhibit fracture mechanics behavior, many studies from early to recent times have treated sea ice failure according to either plasticity or elasticity theory with a strength limit ([1–8]). Both theories exhibit no size effect. When size effects were observed in tests, they were generally attributed to randomness of material strength (e.g., [9]), captured by Weibull [10] theory stemming from the qualitative idea of Mariotte [11] and mathematically justified by extreme value statistics ([12]), see reviews in, e.g., [13–15]. However, the statistical explanation of size effect is, for the present problem, dubious because the maximum load is not reached at the initiation of fracture but only after large stable crack growth (in detail, see, e.g., [14,15]). In that case a nonlocal generalization of Weibull theory is required ([16,17]). The nonlocal probabilistic analysis shows that the statistical size effect becomes significant only of for very large structures failing at fracture initiation. Otherwise the energetic (deterministic) size effect dominates.

Many studies document the brittleness of ice (e.g., [18,19]). Various recent experiments ([20–22]) especially the remarkable in situ tests of Dempsey's team made with record-size specimens ([23–26]), indicate that on a scale exceeding about 0.5-m sea ice does follow cohesive (quasi-brittle) fracture mechanics, with a strong size effect, and on scales larger than about 10 m is very well described by linear elastic fracture mechanics (LEFM). The need for fracture mechanics approach and the presence of size effect is also suggested by the fact that the experimental load-deflection diagrams (e.g., [8]) exhibit no yield plateau but a gradual softening, i.e., a decrease of load with increasing deflection after the peak load has been reached. Analysis of acoustic observations, too, suggests a size effect ([27]).

The analysis of failure and especially the size effect must, therefore, be based on fracture mechanics. Many investigators have been applying to sea ice fracture problems the linear elastic fracture mechanics (LEFM) in which the fracture process zone at the crack tip is assumed to be infinitely small. However, as transpired from the field fracture tests of size effect by Dempsey et al. [23,24], the length of the fracture process zone of sea ice is of the order of several meters for horizontal propagation, while for vertical propagation it is roughly 25 cm. Therefore, the cohesive crack model or some of its approximations must be used. Two basic types of cohesive crack model need to be distinguished: (a) the *brittle-ductile* model, in which the stress-displacement relation has a long horizontal yield plateau, terminating by a sharp drop at a certain critical opening displacement, and (b) the *quasi-brittle* model, in which the cohesive crack-bridging stress gradually decreases according to a fixed law as a function of the opening displacement. The former was developed long ago for metals, and the latter more recently for concrete ([15]). It is the latter type which appears more appropriate for sea ice.

In view of the quasi-brittle behavior, the deterministic (energetic) size effects of quasi-brittle fracture ([14,15,28–32]) must get manifested, and must be expected to be strong, in all the problems in which large cracks grow stably prior to reaching the maximum load ([33,34]). This includes two fundamental problems to be addressed in Parts I and II of this study: (1) the vertical load capacity of floating ice plate (penetration fracture), and (2) the maximum horizontal force exerted on a fixed structure by a moving ice plate.

The vertical penetration problem has been analyzed by fracture mechanics at various levels of sophistication in several recent works. Bažant and Li [35,36] assumed that full-through bending cracks propagate radially from the loaded area, but this assumption now appears inapplicable except perhaps for very thin plates in which the horizontal forces due to dome effect nearly vanish. Dempsey with co-workers [37], in an elegant analytical solution of the problem, assumed that the radial cracks at maximum load emanating from the loaded area reach through only a part of the ice thickness. To make an analytical solution feasible, they made various simplifying assumptions, the main one being a uniform crack depth.

The aforementioned simplifications were avoided in a numerical simulation of penetration fracture in [38,39], which confirmed that indeed the cracks reach only through a part of the thickness and propagate at the maximum load stage mainly vertically, although the crack depth is not uniform. This numerical simulation indicated that for larger ice thicknesses there is a strong size effect, approaching the size effect of geometrically similar failures governed by LEFM, for which the nominal strength is proportional to $(\text{ice thickness})^{-1/2}$. This conclusion represents a sharp contrast with the classical solutions based on plasticity or elasticity with a strength limit. Such solutions inevitably imply the absence of any size effect.

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, Aug. 7, 2000; final revision, July 19, 2001. Associate Editor: A. Needleman. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

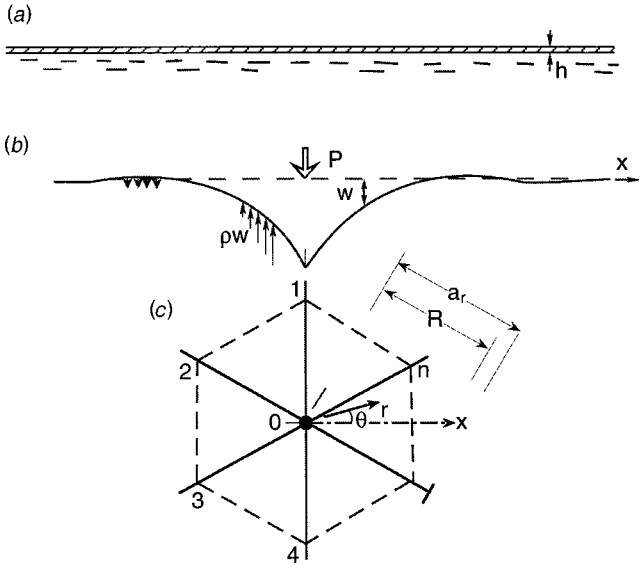


Fig. 1 Floating ice plate, its deflection under concentrated load and crack pattern

Analysis of another ice fracture problem, namely the large-scale thermal bending fracture of floating ice ([40,41]), also indicated a strong size effect, obeying, however, a different law. In this case, the critical temperature difference is not proportional to $(\text{ice thickness})^{-1/2}$, as in LEFM, but to $(\text{ice thickness})^{-3/8}$. The reason is that, at large scale, the cracks must propagate horizontally as bending cracks, rather than vertically across the thickness. A size effect following still another law was recently demonstrated for the fracture of ice subject to a line load ([42]).

As typical for all quasi-brittle materials, the size effect is very difficult to analyze for the normal sizes of interest, but becomes much simpler asymptotically for very large sizes as well as very small sizes ([14,15,43]). The philosophy of *asymptotic matching* ([44]) can then be employed to ''interpolate'' between the opposite asymptotic size effects. This furnishes an approximate solution for the size effect in the difficult intermediate range. This approach, pioneered and widely used in fluid mechanics (e.g., [45–47]), has been successfully employed in many studies of concrete and a more recently in studies of fiber composites and rock ([14,15]).

Static behavior until failure will be assumed in all of the present analysis. Situations in which the ice might acquire significant kinetic energy during a temporarily unstable fracture fracture propagation will not be considered. The creep of ice will not be explicitly considered and the elastic modulus of ice will be assumed to represent the effective modulus that approximately incorporates the effect of creep for the prevalent loading rate.

The purpose of the present two-part study, based on a recent workshop article ([43]), is to employ the asymptotic matching approach to deduce simple approximate formulas for the nominal strength of the ice plate as a function of the size as well as geometry. Such an approach helps intuitive understanding, clarifies the failure mechanism, facilitates optimization of engineering design, elucidates the role of energy release as the main source of size effect, and readily reveals how the material and geometry parameters control the size effect. Part I will deal with the vertical load, and Part II which follows with the horizontal load.

2 Problem Formulation

An ice plate floating on water behaves exactly as a plate on Winkler elastic foundation (Fig. 1(*a*,*b*)), with a foundation modulus equal to the specific weight of water, ρ. Failure under a vertical load is known to involve formation of radial bending cracks in a star pattern (shown in a plan view in Fig. 1(*c*) for the case of six cracks). As transpired from a simplified analytical study of Dempsey et al. [37] and from a detailed numerical simulation ([38,39]), these radial cracks do not reach through the full ice thickness before the maximum load is reached. Rather, they penetrate at maximum load to an average depth of about $0.8h$ and maximum depth $0.85h$ where h is the ice thickness (Fig. 2*a*). The maximum load is reached when polygonal (circumferential) cracks, needed to complete a failure mechanism, begin to form (dashed lines in Fig. 1(*c*)).

The nominal strength, which is a parameter of the maximum vertical load P, is defined for the vertical penetration problem as

$$\sigma_N = P/h^2. \tag{1}$$

In plasticity or any theory in which the material failure criterion is defined in terms of stresses and strains, the nominal strength (of a nonrandom material) is size independent for geometrically similar structures. The size effect in fracture and damage mechanics arises from the fact that the criterion of material failure (crack growth) is expressed in terms of energy (or stress-displacement relation).

Sea ice, unlike glacier ice, is not sufficiently confined to behave plastically (this is for example confirmed by the absence of yield plateau from the measured load-deflection diagrams seen, e.g., in [8]. Sea ice is a brittle material, and so the failure must be analyzed by fracture mechanics (e.g., [20–22,35,36,38–41,48]). The analysis must be based on the rate of energy dissipation at the crack front and the rate of energy release from the ice-water system. The energy release is associated with unloading, during which the ice deforms elastically, with a certain Young's modulus E (which depends on temperature and other factors).

The behavior of the ice plate may be described by the plate bending theory. Dimensional analysis, or transformation of the partial differential equation of a plate on Winkler foundation to dimensionless coordinates, shows that the behavior of the plate is fully characterized by the characteristic length

$$L = (D/\rho)^{1/4} \tag{2}$$

where $D = Eh^3/12(1-\nu^2)$ = cylindrical stiffness of the ice plate; ν = Poisson ratio of ice.

3 Energy Release and Equilibrium of Fractured Ice Plate

Superposing the expressions for the stress intensity factor K_I of the part-through radial bending crack of depth a (Fig. 3*b*,*d*) produced by bending moment M and normal force N (per unit length), one has

$$K_I = \frac{\sqrt{\pi a}}{h}\left[\frac{6M}{h}F_M(\alpha) + NF_N(\alpha)\right] \tag{3}$$

where

$$F_M(\alpha) = \sqrt{\frac{2}{\pi\alpha}\tan\frac{\pi\alpha}{2}}\left(\cos\frac{\pi\alpha}{2}\right)^{-1} \times\left[0.923 + 0.199\left(1-\sin\frac{\pi\alpha}{2}\right)^4\right] \tag{4}$$

$$F_N(\alpha) = \sqrt{\frac{2}{\pi\alpha}\tan\frac{\pi\alpha}{2}}\left(\cos\frac{\pi\alpha}{2}\right)^{-1} \times\left[0.752 + 2.02\alpha + 0.37\left(1-\sin\frac{\pi\alpha}{2}\right)^3\right] \tag{5}$$

([15,49,50]) with an error less than 0.5 percent over the entire range $\alpha \in (0,1)$. According to Irwin's relation, the energy release rate is

$$\mathcal{G} = \frac{K_I^2}{E'} = \frac{N^2}{E'h}g(\alpha) \tag{6}$$

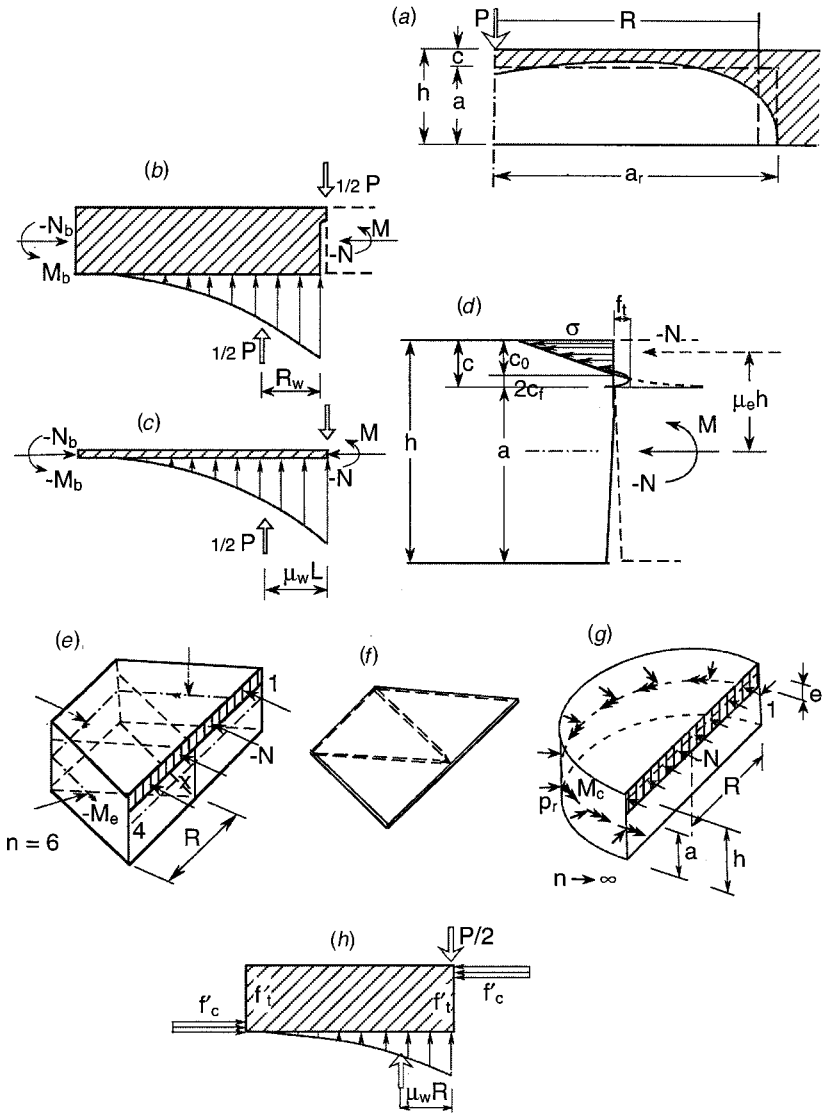


Fig. 2 Analysis of vertical penetration fracture: (*a*) crack profile and (*b–h*) forces acting on element 123401 in Fig. 1

where $E'=E/(1-\nu^2)$ and g is a dimensionless function,

$$g(\alpha)=\pi\alpha\left[\frac{6e}{h}F_M(\alpha)+F_N(\alpha)\right]^2 \quad (\alpha=a/h). \tag{7}$$

$e=-M/N=$ eccentricity of the normal force resultant in the cross section (positive when N is above the midplane).

To relate M and N to vertical load P, let us consider element 12341 of the plate (Figs. 1(*c*) and 2(*e,f,g*)), limited by a pair of opposite radial cracks and the initiating polygonal cracks. The depth to the polygonal cracks at maximum load is zero, as they just initiate, and since the cracks must form at the location of the maximum radial bending moment, the vertical shear force on the planes of these cracks is zero. The distance R of the polygonal cracks from the vertical load P may be expected to be proportional to the characteristic length L since this is the only length constant in the differential equation governing the problem, and so we may set $R=\mu_R L$ where dimensionless μ_R is assumed to be a constant.

In each narrow radial sector, the resultant of the water pressure due to deflection w (Fig. 2(*b,c*)) is located at a certain distance r_w from load P. Since r_w can be solved from the differential equation for w, and since the solution depends only on one parameter, the characteristic length L, r_w must be proportional to L. Integration over the area of a semi-circle of radius r_w yields the resultant of water pressure acting on the whole element 12341. Again, the distance of this resultant, whose magnitude it $P/2$, from load P must be proportional to L, i.e., may be written as

$$R_w=\mu_m L \tag{8}$$

where μ_w is a constant that can be solved from the differential equation of plate deflections. Of course, μ_w is a constant only as long as the behavior is elastic, which is exactly true only if the crack depth a is constant. Although the crack is growing, we will assume that its rate of growth is small enough so that μ_w would be approximately constant.

For the sake of simplicity, we assume the normal force N and bending moment M on the planes of the radial cracks and the polygonal cracks to be uniform. The condition of equilibrium of horizontal forces acting on element 12341 in the direction normal to the radial cracks is then simple; it requires the normal forces on the planes of the polygonal cracks to be equal to the normal force N acting in the radial crack planes. The axial vectors of the moments M_c acting on the polygonal sides are shown in Fig. 2(*e,g*) by double arrows. Summing the projections of these axial vectors

from all the polygonal sides of the element, one finds that their moment resultant with axis in the direction 14 is $2RM_c$, regardless of the number n of radial cracks. So, upon setting $R=\mu_R L$, the condition of equilibrium of the radial cracks with the moments about axis 14 (Fig. 2(*b,c,e,g*)) located at midthickness of the cross section may be written as

$$2(\mu_R L)M+2(\mu_R L)M_c-\frac{1}{2}P(\mu_w L)=0. \tag{9}$$

Furthermore, we must take into account condition (6) of vertical propagation of the radial bending cracks, which may be written as $\mathcal{G}=G_f$ where G_f is the fracture energy of ice. Thus, the critical value of normal force (compressive, with eccentricity e) may be written as

$$N=-\sqrt{\frac{E'G_f\ h}{g(\alpha)}}. \tag{10}$$

Depending on the energy release rate $g(\alpha_0)$ of the actual crack of length $a_0=\alpha_0 D$ (excluding the cohesive zone), there are two kinds of deterministic size effect: (a) the size effect due to energy release of a large crack, characterized by a large value of $g(\alpha_0)$, and (b) the size effect at crack initiation ($\alpha_0=0$), characterized by $g(\alpha_0)=0$. They are governed by different laws ([14,15,30,32,51]), and both must be expected to occur in ice penetration.

4 Size Effect on Flexural Strength at Initiation of Polygonal Cracks

Consider first the initiation of the polygonal cracks. Since $\alpha_0=0$ and $g(\alpha_0)=0$, the initiation criterion is that the normal stress σ reaches the tensile strength f_t' of the ice. However, the crack can begin to propagate only after a boundary layer of distributed microcracking, representing the fracture process zone, forms at the top surface of ice ([14,15,30,51,52]). The half-depth of this layer, denoted as D_b, is a material constant (which should be roughly equal to the fracture process zone length c_f introduced later). Note that the boundary layer D_b has been shown to explain the experimentally observed size effect on the modulus of rupture in the bending tests of concrete ([15,52]).

Although the crack initiation can be handled by the energy release function, it is simpler to consider the stress redistribution in the cross section caused by softening in the boundary layer ([52]). The easiest way to obtain a nominal strength formula that is correct up to the first two terms of the expansion in terms of powers of $1/h$ is to write the condition that the elastically calculated normal stress σ_e should be equal to the tensile strength of ice, f_t', at the middle of the boundary layer of thickness $2D_b$, rather than at ice surface. So the crack initiation criterion is $\sigma_e+N/h=f_t'$ where, according to the bending stress formula, $\sigma_e=M_c(h/2-D_b/2)/(h^3/12)$. This yields the crack initiation criterion:

$$\frac{6M_c}{h^2}q(h)+\frac{N}{h}=f_t' \tag{11}$$

where $q(h)=1-D_b/h$. This form of the criterion, however, becomes meaningless when $h<2D_b$, i.e., when the ice is thinner than the cracking layer thickness. It can be correct only when h is sufficiently larger than $2D_b$, i.e., asymptotically for $h/D_b\to\infty$. So it is desirable to modify function $q(h)$ so as to obtain a formula approximately applicable through the entire size range. This can be achieved by considering a range of σ_N formulas that have the same first two terms of the large-size asymptotic expansion in $1/h$ as (11), and then choosing that which gives the correct value of the small-size nominal strength. Such a kind of approach is known as asymptotic matching.

When $h=2D_b$, i.e., when the distributed cracking zone encompasses essentially the whole depth of plate, the moment at failure can be approximately determined as the plastic bending moment M_p. If f_c'/f_t', with f_c'=compression strength of ice, is about 1, then the plastic stress distribution is symmetric bi-rectangular and $M_p/M_e=1.5$, where M_e=elastically calculated bending moment for which $\sigma=f_c'/f_t'$ at ice surface. If f_c'/f_t' were very large, then the stress distribution would be a single rectangle balanced by a concentrated compression force at ice surface, and in that case M_p/M_e would be equal to 3. The real value must lie in between, but probably closer to 1.5. We will safely assume that $M_p/M_e=1.5$. So we should seek a formula for $q(h)$ that gives this ratio for $h=D_b$ and has a large-size asymptotic expansion of the form $1-D_b(1/h)+(\cdot)(1/h)^2+\ldots$. There are many such formulas but the simplest one is

$$q(h)=\frac{1+D_b/h}{1+2D_b/h}. \tag{12}$$

This is verified by the asymptotic expansion:

$$\frac{1+D_b/h}{1+2D_b/h}=\left(1+\frac{D_b}{h}\right)\left(1-\frac{2D_b}{h}+\frac{4D_b^2}{h^2}-\ldots\right)$$
$$=1-\frac{D_b}{h}+\frac{(\cdot)}{h^2}+\frac{(\cdot)}{h^3}+\ldots. \tag{13}$$

5 Size Effect on Nominal Vertical Penetration Strength

Aside from the stress redistribution at initiation of polygonal cracks ([52]), there is another deterministic source of size effect—the energy release due to vertical propagation of the radial bending cracks ([28]). The bending moment

$$M=-Ne=-N\mu_e h \tag{14}$$

may be substituted into (9); here the normal force N is defined to be positive when tensile, although the actual value of N is negative (compression); and $\mu_e=e/h$=dimensionless parameter whose value at maximum load may be assumed to be approximately constant. This assumption is indicated by the numerical simulations in [38,39], from which it further transpires that $\mu_e\approx 0.45$, as a consequence of the fact that the average crack depth a at maximum load is about $0.8h$ (in any case, $\mu_e<0.5$, and so a possible error in μ_e cannot have a large effect). The value 0.45 approximately corresponds to the correct number of cracks in the star pattern; if there were more cracks, the depth would be smaller, if fewer, larger.

After substituting (14), we may express M_c from (9) and substitute it into (11). Then, taking into account (10), we obtain after rearrangements the equation:

$$\sigma_N=\frac{2\mu_R}{3\mu_w}\left[\left(6\mu_e+\frac{1}{q(h)}\right)\sqrt{\frac{E'G_f}{hg(\alpha)}}+\frac{f_t'}{q(h)}\right] \tag{15}$$

where $q(h)$ is given by (12).

Now we need to decide how the values of α at maximum load should vary with ice thickness h. To this end, note that ice is a quasibrittle material. This is evidenced by the fact that at small laboratory scale it is notch-insensitive and exhibits no size effect while at large scale it behaves according to LEFM ([20,24]). Therefore, at the tip of the vertically propagating radial crack, there must exist a finite fracture process zone (FPZ) of a certain characteristic depth $2c_f$ which is a material property. This zone was modeled in the numerical simulations of Bažant and Kim [38,39] as a yielding zone. The tip of the equivalent LEFM crack lies approximately in the middle of the FPZ, i.e., at a distance c_f from the actual crack tip ([15]), whose location is denoted as a_0.

If the locations of the center of the FPZ in structures of different sizes were geometrically similar, i.e., if α at maximum load were the same for all h, then the size effect would be the same as in LEFM. Experience with testing of quasi-brittle materials ([15]), as well as with cohesive crack and nonlocal damage simulations,

shows the locations of the center of FPZ are usually not geometrically similar. Rather, similar locations are those of the actual crack tip. Thus the value of $\alpha_0 = a_0/h$ may be expected to be approximately constant when ice plates of different thicknesses h are compared. Denoting $g'(\alpha_0) = dg(\alpha_0)/d\alpha_0$, one may introduce the approximation

$$g(\alpha) \approx g(\alpha_0) + g'(\alpha_0)(c_f/D). \tag{16}$$

Substituting this into (15) and rearranging, one gets for the size effect the formula

$$\sigma_N = \frac{4\mu_R}{\mu_w}\left(\mu_e + \frac{1}{6q(h)}\right)\sqrt{\frac{E'G_f}{hg(\alpha_0) + c_f g'(\alpha_0)}} + \frac{\mu_R}{3\mu_w}\frac{f_t'}{q(h)}. \tag{17}$$

The results of numerical simulations in [39] were found to be quite well represented by the simple classical size effect law with large-size residual strength σ_r proposed in [53] which reads

$$\sigma_N = \sigma_0\left(1 + \frac{h}{h_0}\right)^{-1/2} + \sigma_r. \tag{18}$$

Formula (17) is now seen to reduce to this law when $q(h) \approx 1$, i.e., when D_b is negligible, in which case then

$$\sigma_0 = \frac{4\mu_R\mu_e}{\mu_w}\sqrt{\frac{E'G_f}{c_f g'(\alpha_0)}}, \quad h_0 = c_f\frac{g'(\alpha_0)}{g(\alpha_0)}, \quad \sigma_r = \frac{\mu_e}{3\mu_w}f_t'. \tag{19}$$

Furthermore, the numerical simulations in [39] indicated that $\sigma_r \approx 0$. This means that the contribution of the tensile strength f_t' governing the initiation of the polygonal cracks must be negligible, which in turn implies a negligible role for $q(h)$.

The terms in (17) containing D_b anyway decrease with increasing h much more rapidly than (18)—they decrease with increasing h as $1/h$, compared to $1/\sqrt{h}$. Consequently, they must become negligible for not too large h regardless of the value of D_b.

Same as (18), formula (17) plotted as $\log \sigma_N$ versus $\log h$ approaches for large h a downward inclined asymptote of slope $-1/2$ (Fig. 3(*g*)). This characterizes the large-size asymptote of the size effect law in (17).

How does the number n of the radial cracks enter the solution? It does not appear in the present solution for the maximum load. The reason is that the number of cracks is decided at the beginning of loading, long before the maximum load is attained.

It is interesting to contrast the size effect obtained here with that deduced for large-scale thermal bending fracture of floating ice, which was shown to be ([40])

$$\Delta T \propto h^{-3/8} \tag{20}$$

where ΔT is the temperature difference between the bottom and top of the ice plate, which is proportional to the maximum thermal stress before fracture. The large-size asymptotic size effect for fracture under vertical loads would have to follow also the $-3/8$ power law if the cracks at maximum load penetrated through the full thickness of ice and force N were negligible ([35,40,42,54]). But this turned out not to be the case ([37–39,55,56]).

6 Comments on Plasticity Approach

In contrast to the brute-force numerical simulations conducted before, the approximate analytical derivation of size effect is intuitively instructive. It clarifies the reasons why there must be a deterministic size effect in penetration of floating ice. The size effect could be absent only if the material behaved plastically.

If the sea ice were a plastic material, the stress distributions on element 12341 would be as shown in Fig. 2(*h*), where f_t' and f_c' denote the tensile and compressive yield strengths. Taking the moment equilibrium condition of this element, one can easily show that the nominal strength would in that case be expressed as

$$\sigma_N = \frac{4\mu_R}{\mu_w}(f_c'^{-1} + f_t'^{-1})^{-1} \tag{21}$$

which exhibits no size effect. Plasticity, however, requires that the material strength at all the points of the failure surface be mobilized at the same time, which is impossible for a quasi-brittle (softening) material such as sea ice.

7 Closing of Part I

The simplified asymptotic analysis of size effect in vertical penetration of the ice plate confirms the inevitability of a strong size effect for larger ice thicknesses, approaching the size effect of LEFM. This conclusion does not disagree with experiments and is supported by previous numerical studies summarized in the Appendix. Part II which follows will apply a similar approach to the problem of an ice plate moving against a fixed structure. It will be seen that size effects must again be expected, but their nature is rather different.

Appendix

Review of Previous Numerical Fracture Analysis of Size Effect. To supplement the analytical approach, it may be useful to review recent detailed numerical simulation of fracture of floating ice caused by a vertical load ([38,39]). The fracture pattern (for the case of six radial cracks) is shown in Fig. 3(*a*). The radial cracks at maximum load penetrate through only a part of ice thickness ([26,55]); Fig. 3(*b,c*). The radius of each crack is divided by nodes into vertical strips in each of which the vertical crack growth obeys Rice and Levy's [57] ''nonlinear line-spring'' model relating the normal force N and bending moment M in the cracked cross section to the relative displacement Δ and rotation θ (Fig. 3(*b*)).

The analysis is based on a simplified version of the cohesive crack model in which the vertical crack growth in each vertical strip is initiated according to a strength criterion. The cross section behavior is considered elastic-plastic until the yield envelope in the (N, M) plane is crossed by the point (N, M) corresponding to fracture mechanics. For ease of calculations, a nonassociated plastic flow rule corresponding to the vector $(d\Delta, d\theta)$ based on fracture mechanics is assumed.

The following ice characteristics have been used in calculations: tensile strength $f_t' = 0.2$ MPa, fracture toughness $K_c = 0.1\,\text{MPa}\sqrt{\text{m}}$, Poisson ratio $\nu = 0.29$, and Young' modulus $E = 1.0$ GPa, with the corresponding values: fracture energy $G_f = K_c^2/E = 10\,\text{J/m}^2$, and Irwin's fracture characteristic length $l_0 = (K_c/f_t')^2 = 0.25$ m (this value happens to be about the same as for concrete).

Figure 3(*e*) displays, with a strongly exaggerated vertical scale, the calculated crack profiles at subsequent loading stages. Fig. 3(*f*) shows the numerically calculated plot of the radial crack length a versus the ice thickness h (''fracture length'' means the radial length of open crack, and ''plastic length'' the radial length up to the tip of plastic zone). This plot reveals that, except for very thin ice, the radial crack length $a \approx c_h h$ where $c_h \approx 24$ for the typical ice properties assumed.

The data points in Fig. 3(*g*) show, in logarithmic scales, the numerically obtained size effect plot of the normalized nominal strength $\sigma_N = P/h^2$ versus the relative thickness of the ice (note that according to plasticity or elasticity with strength criterion, this plot would be a horizontal line). The initial horizontal portion, for which there is no size effect, corresponds to ice thinner than about 20 cm.

Since the model in [38,39] includes plasticity, it can reproduce the classical solutions with no size effect, depending on the input values of ice characteristics. The ice thickness at the onset of size effect depends on the ratio of ice thickness to the fracture characteristic length, h/l_0. For realistic ice thicknesses h ranging from 0.1 m to 6 m, the computer program would yield perfectly plastic

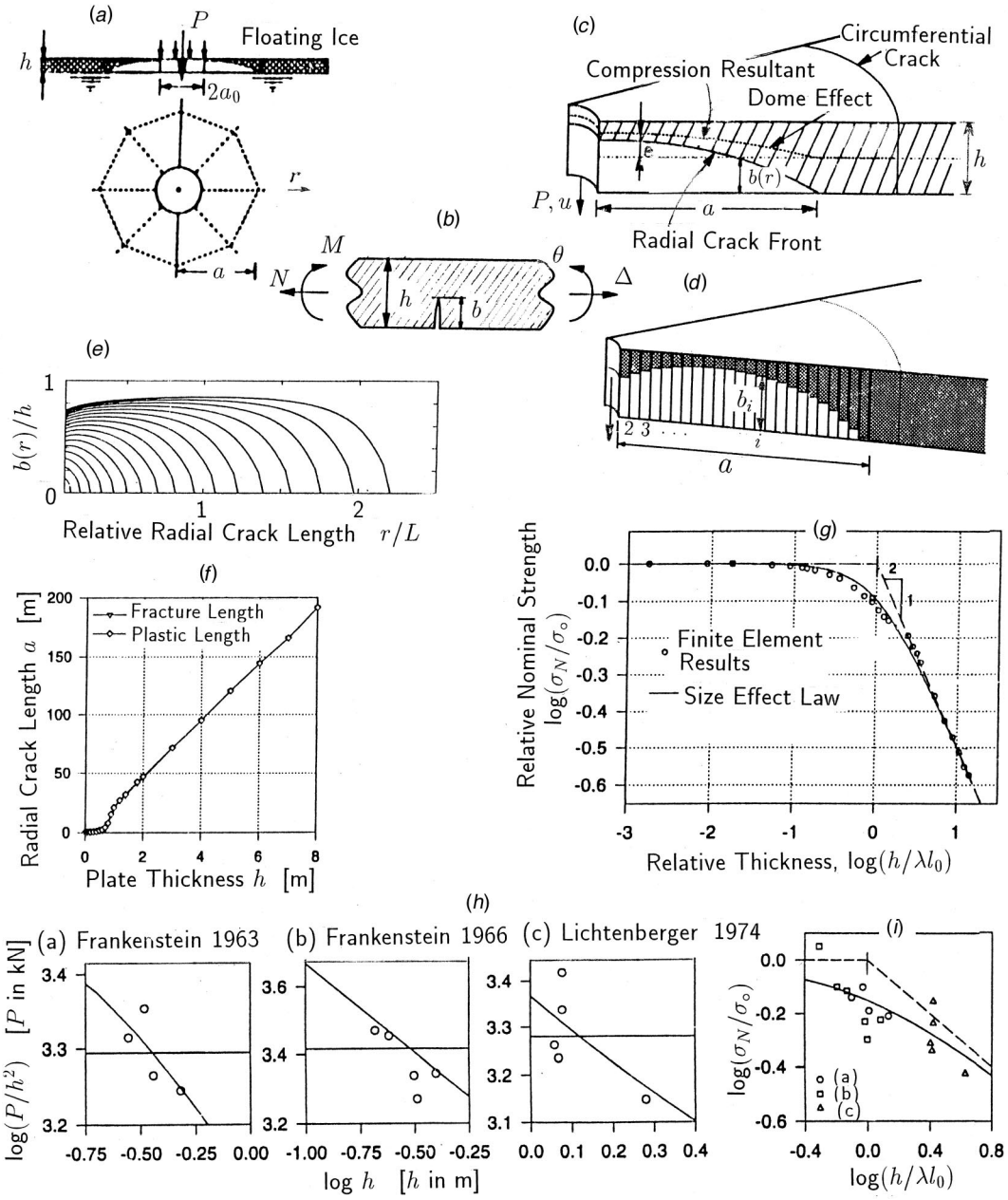


Fig. 3 Vertical penetration fracture problem analyzed by Bažant and Kim [38,39] main numerical results, and comparison with field tests of Frankenstein [59,60] and Lichtenberger [61]

response with no size effect if the fracture characteristic length l_0 were at least 100× larger, i.e., at least 25 m. This would, for instance, happen if either f_t' were at least 10× smaller ($f_t' \leqslant 0.01$ MPa or K_c at least 10× larger ($K_c \geqslant 10\ \text{MPa}\sqrt{\text{m}}$). The entire diagram in Fig. 3(*g*) would then be horizontal.

Larger values of l_0 are of course possible in view of statistical scatter, but nothing like 100× larger. For example, by fitting size effect data ([23,24]) from in situ tests at Resolute, one gets $K_c \approx 2.1\ \text{MPa}\sqrt{\text{m}}$, and with $f_t' \approx 2$ MPa one has the fracture characteristic length $l_0 = (K_c/f_t')^2 = 1$ m. But this larger value would not make much difference in the size effect plot in Fig. 3(*g*). The reason that these values were not used in the plot in Fig. 3(*g*) was that they correspond to long-distance horizontal propagation of fracture, rather than vertical growth of fracture.

The curve in Fig. 3(*g*) is the optimum fit of the numerically calculated data points by the generalized size effect law proposed in Bažant [58]. The final asymptote has slope −1/2, which means that the asymptotic size effect is $\sigma_N \propto h^{-1/2}$, the same as for LEFM with similar cracks, and not $h^{-3/8}$ as proposed by Slepyan [35,40,54]. The −3/8 power scaling would have to be true if the radial cracks at maximum load were full-through bending cracks. The −1/2 power scaling may be explained by the fact that during failure the bending cracks are not full-through and propagate mainly vertically, which is supported by the calculated crack profiles in Fig. 3(*e*).

By fitting of the data points in Fig. 3(*g*), spanning over four orders of magnitude of ice thickness *h*, the following prediction formula in the form of the generalized size effect law ([15,41]) has been calibrated (see the curve in Fig. 3(*g*)):

$$P_{\max} = \sigma_N h^2, \quad \sigma_N = B f_t' [1 + (h/\lambda_0 l_0)^r]^{-1/2r} \tag{22}$$

with $B=1.214$, $\lambda_0=2.55$, $m=1/2$, $r=1.55$, and $l_0=0.25$ m ($f'_t=0.2$ MPa in Fig. 1(*g*)).

Only very limited field test data exist. The data points in the size effect plots in Fig. 1(*h*) represent the results of the field tests by [59–61], and the curves show the optimum fits with the size effect formula verified by numerical calculations (note that if the size effect were absent, these plots of nominal strength would have to be horizontal). After optimizing the size effect law parameters by fitting the data in the three plots in Fig. 3(*h*), the data and the optimum fit are combined in the dimensionless plot in Fig. 3(*i*).

Interesting discussions of ([38,39]) were published by Dempsey [62] and Sodhi [63] and rebutted. One objection raised by Sodhi was the neglect of creep in Bažant and Kim's analysis. Intuition suggests that the influence of creep might be like that of plasticity, which tends to increase the process zone size, thereby making the response less brittle and the size effect weaker. But the opposite is true ([15]).

The influence of creep on scaling of brittle failures of concrete, which is doubtless quite similar from the mechanics viewpoint (albeit different in physical origin), was studied in depth at Northwestern University, along with the effect of the crack propagation velocity; see, e.g., [15,34,64] and especially [65,66]. The conclusion from these studies, backed by extensive fracture testing of concrete and rock at very different rates, is that creep in the material always makes the size effect due to cracks stronger (unless creep actually prevents crack initiation). In the logarithmic size effect plot of nominal strength versus structure size, it causes a shift to the right, toward the LEFM asymptote, which means that the size effect is intensified by creep. The slower the loading (or the longer its duration), the closer to LEFM is the size effect in a cracked structure.

The physical reason, clarified by numerical solutions of stress profiles with a rate-dependent cohesive crack model ([66]), is that the highest stresses in the fracture process zone get relaxed by creep, which tends to reduce the effective length of the fracture process zone. The shorter the process zone, the higher is the brittleness of response and the stronger is the size effect. This explains why experiments on notched concrete specimens consistently show the size effect to be more pronounced at a slower loading ([15]). A similar behavior might be expected for ice. It thus transpires that, in order to take the influence of creep on the size effect approximately into account, it suffices to reduce the value of fracture energy (or fracture toughness) and decrease the effective length of the fracture process zone.

References

[1] Bernstein, S., 1929, "The Railway Ice Crossing," Trudy Nauchno-Technicheskogo Komiteta Narodnogo Komissariata Putei, Soobshchenniya, Vol. **84** Moscow (in Russian).

[2] Nevel, D. E., 1958, "The Theory of Narrow Infinite Wedge on an Elastic Foundation," Trans. Eng. Inst. Canada, **2**, No. 3.

[3] Kerr, A. D., 1975, "The Bearing Capacity of Floating Ice Plates Subjected to Static or Quasi-Static Loads—A Critical Survey," Research Report 333, U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, NH.

[4] Kerr, A. D., 1996, "Bearing Capacity of Floating Ice Covers Subjected to Static, Moving, and Oscillatory Loads," Appl. Mech. Rev., **49**, No. 11, pp. 463–476.

[5] Sodhi, D. S., 1995, "Breakthrough Loads of Floating Ice Sheets," J. Cold Regions Eng. ASCE, **9**, No. 1, pp. 4–20.

[6] Sodhi, D. S., 1995, "Wedging Action During Vertical Penetration of Floating Ice Sheets," Ice Mechanics, AMD-Vol. 207, pp. 65–80.

[7] Sodhi, D. S., 1996, "Deflection Analysis of Radially Cracked Floating Ice Sheets," *17th Int. Conf. OMAE Proceedings*, ASME, New York, pp. 97–101.

[8] Sodhi, D. S., 1998, "Vertical Penetration of Floating Ice Sheets," Int. J. Solids Struct., **35**, No. 31-32, pp. 4275–4294.

[9] Butiagin, I. P., 1966, *Strength of Ice and Ice Cover*, Izdatel'stvo Nauka, Sibirskoe Otdelenie, Novosibirsk, Russia, 154 pp.

[10] Weibull, W., 1939, "The Phenomenon of Rupture in Solids," Proc., Roy. Swed. Inst. Eng. Res., **153**, Stockholm, pp. 1–55.

[11] Mariotte, E., 1686, *Traité du mouvement des eaux*, posthumously edited by M. de la Hire; Engl. Transl. by J. T. Desvaguliers, London (1718), p. 249; also *Mariotte's Collected Works*, 2nd Ed., The Hague (1740).

[12] Fisher, R. A., and Tippett, L. H. C., 1928, "Limiting Forms of the Frequency Distribution of the Largest and Smallest Member of a Sample," Proc. Cambridge Philos. Soc., **24**, pp. 180–190.

[13] Kittl, P., and Diaz, G., 1988, "Weibull's Fracture Statistics, or Probabilistic Strength of Materials: State of the Art," Res. Mech., **24**, pp. 99–207.

[14] Bažant, Z. P., and Chen, E.-P., 1997, "Scaling of Structural Failure," Appl. Mech. Rev., **50**, No. 10, pp. 593–627.

[15] Bažant, Z. P., and Planas, J., 1998, *Fracture and Size Effect in Concrete and Other Quasibrittle Materials*, CRC Press, Boca Raton, FL.

[16] Bažant, Z. P., and Xi, Y., 1991, "Statistical Size Effect in Quasi-Brittle Structures: II. Nonlocal Theory," J. Eng. Mech., **117**, No. 11, pp. 2623–2640.

[17] Bažant, Z. P., and Novák, D., 2000, "Probabilistic Nonlocal Theory for Quasibrittle Fracture Initiation and Size Effect. I. Theory," J. Eng. Mech., **126**, No. 2, pp. 166–174.

[18] Weeks, W. F., and Mellor, M., 1984, "Mechanical Properties of Ice in the Arctic Seas," *Arctic Technology & Policy*, I. Dyer and C. Chryssostomidis, eds., Hemisphere, Washington, D.C., pp. 235–259.

[19] Weeks, W. F., and Assur, A., 1972, "Fracture of Lake and Sea Ice," Fracture, H. Liebowitz, ed., **II**, pp. 879–978.

[20] Dempsey, J. P., 1991, "The Fracture Toughness of Ice," *Ice Structure Interaction*, S. J. Jones, R. F. McKenna, J. Tilotson, and I. J. Jordaan, eds., Springer-Verlag, Berlin, pp. 109–145.

[21] DeFranco, S. J., and Dempsey, J. P., 1994, "Crack Propagation and Fracture Resistance in Saline Ice," J. Glaciol., **40**, pp. 451–462.

[22] DeFranco, S. J., Wei, Y., and Dempsey, J. P., 1991, "Notch Acuity Effects on Fracture of Saline Ice," Ann. Glaciol., **15**, pp. 230–235.

[23] Dempsey, J. P., DeFranco, S. J., Adamson, R. M., and Mulmule, S. V., 1999, "Scale Effects on the in situ Tensile Strength and Fracture of Ice: Part I: Large Grained Freshwater Ice at Spray Lakes Reservoir, Alberta," Int. J. Fract., **95**, pp. 325–345.

[24] Dempsey, J. P., Adamson, R. M., and Mulmule, S. V., 1999, "Scale Effects on the in situ Tensile Strength and Fracture of Ice: Part II: First-Year Sea Ice at Resolute, N. W. T," Int. J. Fract., **95**, pp. 346–378.

[25] Mulmule, S. V., Dempsey, J. P., and Adamson, R. M., 1995, "Large-Scale in-situ Ice Fracture Experiments—Part II: Modeling Efforts," *Ice Mechanics—1995*, Vol. AMD-MD-1995, ASME, New York.

[26] Dempsey, J. P., Adamson, R. M., and Mulmule, S. V., 1995, "Large-Scale in-situ Fracture of Ice," *Proceedings of FRAMCOS-2*, F. H. Wittmann, ed., AEDIFICATIO Publishers, Freiburg, Germany.

[27] Li, Z., and Bažant, Z. P., 1998, "Acoustic Emissions in Fracturing Sea Ice Plate Simulated by Particle System," J. Eng. Mech., **124**, No. 1, pp. 69–79.

[28] Bažant, Z. P., 1984, "Size Effect in Blunt Fracture: Concrete, Rock, and Metal," J. Eng. Mech., **110**, pp. 518–535.

[29] Bažant, Z. P., 1993, "Scaling Laws in Mechanics of Failure," J. Eng. Mech., **119**, No. 9, pp. 1828–1844.

[30] Bažant, Z. P., 1997, "Scaling of Quasibrittle Fracture: Asymptotic Analysis," Int. J. Fract., **83**, No. 1, pp. 19–40.

[31] Bažant, Z. P., 1997, "Scaling of Quasibrittle Fracture: Hypotheses of Invasive and Lacunar Fractality, Their Critique and Weibull Connection," Int. J. Fract., **83**, No. 1, pp. 41–65.

[32] Bažant, Z. P., 1999, "Size Effect on Structural Strength: A Review," Arch. of Applied Mechanics **69**, pp. 703–725.

[33] Bažant, Z. P., and Kim, J.-K., 1985, "Fracture Theory for Nonhomogeneous Brittle Materials With Application to Ice," *Proc. ASCE Nat. Conf. on Civil Engineering in the Arctic Offshore—ARCTIC 85*, L. F. Bennett ed., San Francisco, CA, ASCE, Reston, VA, pp. 917–930.

[34] Bažant, Z. P., and Gettu, R., 1991, "Size Effects in the Fracture of Quasi-Brittle Materials," *Cold Regions Engineering, (Proc., 6th ASCE International Specialty Conference)*, D. S. Sodhi, ed., Hanover, NH, Feb., ASCE, Reston, VA, pp. 595–604.

[35] Bažant, Z. P., and Li, Y.-N., 1994, "Penetration Fracture of Sea Ice Plate: Simplified Analysis and Size Effect," J. Eng. Mech., **120**, No. 6, pp. 1304–1321.

[36] Li, Y.-N., and Bažant, Z. P., 1994, "Penetration Fracture of Ice Plate: 2D Analysis and Size Effect," J. Eng. Mech., **120**, No. 7, pp. 1481–1498.

[37] Dempsey, J. P., Slepyan, L. I., and Shekhtman, I. I., 1995, "Radial Cracking With Closure," Int. J. Fract., **73**, No. 3, pp. 233–261.

[38] Bažant, Z. P., and Kim, Jang-Jay H., 1998, "Size Effect in Penetration of Sea Ice Plate With Part-Through Cracks. I. Theory," J. Eng. Mech., **124**, No. 12, pp. 1310–1315; with discussions and closure in **126**, No. 4, pp. 438–442, 2000.

[39] Bažant, Z. P., and Kim, Jang-Jay H., 1998, "Size Effect in Penetration of Sea Ice Plate With Part-Through Cracks. II. Results," J. Eng. Mech., **124**, No. 12 pp. 1316–1324; with discussions and closure in **126**, No. 4, pp. 438–442, 2000.

[40] Bažant, Z. P., 1992, "Large-Scale Thermal Bending Fracture of Sea Ice Plates," J. Geophys. Res., **97**, No. C11, pp. 17739–17751.

[41] Bažant, Z. P., 1992, "Large-Scale Fracture of Sea Ice Plates," *Proc. 11th IAHR Ice Symposium*, Vol. 2, T. M. Hrudey, ed., Banff, University of Alberta, June, pp. 991–1005.

[42] Bažant, Z. P., and Guo, Z., 2002, "Size Effect on Strength of Floating Sea Ice Under Vertical Line Load," J. Eng. Mech., **128**, No. 3.

[43] Bažant, Z. P., 2000, "Scaling Laws for Brittle Failure of Sea Ice," IUTAM Symp. on Scaling Laws in Ice Mechanics, Univ. of Alaska, Fairbanks, J. P. Dempsey, H. H. Shen, and L. H. Shapiro, eds., University of Alaska, Fairbanks, Paper No. 3, pp. 1–23.

[44] Bender, M. C., and Orszag, S. A., 1978, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, New York, Chapters 9–11.

[45] Barenblatt, G. I., 1979, *Similarity, Self-Similarity and Intermediate Asymptotics*, Consultants Bureau (Plenum Press), New York (transl. from Russian original, 1978).
[46] Barenblatt, G. I., 1987, *Dimensional Analysis*, Gordon and Breach, New York.
[47] Sedov, L. I., 1959, *Similarity and Dimensional Methods in Mechanics*, Academic Press, New York.
[48] DeFranco, S. J., and Dempsey, J. P., 1992, ''Nonlinear Fracture Analysis of Saline Ice: Size, Rate and Temperature Effects,'' *Proc. of the 11th IAHR Symposium*, Vol. 3, Banff, University of Alberta, pp. 1420–1435.
[49] Tada, H., Paris, P. C., and Irwin, J. K., 1985, *The Stress Analysis of Cracks Handbook*, 2nd Ed., Paris Productions, St. Louis, MO.
[50] Murakami, Y., ed., 1987, *Stress Intensity Factors Handbook*, Vol. 1, Pergamon Press, Oxford, UK.
[51] Bažant, Z. P., 2001, *Scaling of Structural Strength*, Hermes Scientific Publications, Oxford, UK.
[52] Bažant, Z. P., and Li, Z., 1995, ''Modulus of Rupture: Size Effect due to Fracture Initiation in Boundary Layer,'' J. of Struct. Engrg. ASCE, **121**, No. 4, pp. 739–746.
[53] Bažant, Z. P., 1987, ''Fracture Energy of Heterogeneous Material and Similitude,'' *SEM-RILEM Int. Conf. on Fracture of Concrete and Rock* S. P. Shah and S. E. Swartz, eds., SEM, pp. 390–402.
[54] Slepyan, L. I., 1990, ''Modeling of Fracture of Sheet Ice,'' Mech. Solids (transl. of Izv. An SSSR Mekhanika Tverdoga Tela), 155–161.
[55] Bažant, Z. P., and Li, Y.-N., 1995, ''Penetration Fracture of Sea Ice Plate,'' Int. J. Solids Struct., **32**, No. 3/4, pp. 303–313.
[56] Bažant, Z. P., Kim, J. J., and Li, Y.-N., 1995, ''Part-Through Bending Cracks in Sea Ice Plates: Mathematical Modeling,'' *ICE MECHANICS—1995*, J. P. Dempsey and Y. Rajapakse, eds., ASME AMD-Vol. 207, pp. 97–105.
[57] Rice, J. R., and Levy, N., 1972, ''The Part-Through Surface Crack in an Elastic Plate,'' ASME J. Appl. Mech., **39**, pp. 185–194.
[58] Bažant, Z. P., 1985, ''Fracture Mechanics and Strain-Softening in Concrete,'' Preprints, U.S.-Japan Seminar on Finite Element Analysis of Reinforced Concrete Structures, Vol. 1, Tokyo, pp. 47–69.
[59] Frankenstein, E. G., 1963, ''Load Test Data for Lake Ice Sheet,'' Technical Report 89, U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, NH.
[60] Frankenstein, E. G., 1966, ''Strength of Ice Sheets,'' *Proc., Conf. on Ice Pressures Against Struct.; Tech. Memor. No. 92, NRCC No. 9851*, Laval University, Quebec, National Research Council of Canada, pp. 79–7.
[61] Lichtenberger, G. J., Jones, J., W., Stegall, R. D., and Zadow, D. W., 1974, ''Static Ice Loading Tests: Resolute Bay—Winter 1973/74.'' APOA Proj. No. 64, Rep. No. 745B-74-14, CREEL Bib #34-3095, Sunoco Sci. and Technol., Richardson, TX.
[62] Dempsey, P. P., 2000, discussion of ''Size Effect in Penetration of Ice Plate With Part-Through Cracks. I. Theory, II. Results.'' Z. P. Bažant and J. J. H. Kim, J. Eng. Mech., **126**, No. 4, p. 438.
[63] Sodhi, D. S., 2000, discussion of ''Size Effect in Penetration of Ice Plate With Part-Through Cracks. I. Theory, II. Results.'' Z. P. Bažant and J. J. H. Kim, J. Eng. Mech., **126**, No. 4, pp. 438–440.
[64] Bažant, Z. P., and Jirásek, M., 1993, ''R-Curve Modeling of Rate and Size Effects in Quasibrittle Fracture,'' Int. J. Fract., **62**, pp. 355–373.
[65] Bažant, Z. P., and Li, Yuan-Neng, 1997, ''Cohesive Crack With Rate-Dependent Opening and Viscoelasticity: I. Mathematical Model and Scaling,'' Int. J. Fract., **86**, No. 3, pp. 247–265.
[66] Li, Yuan-Neng, and Bažant, Z. P., 1997, ''Cohesive Crack With Rate-Dependent Opening and Viscoelasticity: II. Numerical Algorithm, Behavior and Size Effect,'' Int. J. Fract., **86**, No. 3, pp. 267–288.

Scaling of Sea Ice Fracture—Part II: Horizontal Load From Moving Ice

Z. P. Bažant
Walter P. Murphy Professor of Civil Engineering and Materials Science, Northwestern University, Evanston, IL 60208
e-mail: z-bazant@northwestern.edu
Fellow ASME

Continuing the analysis of fracture size effect in Part I, which was focused on the maximum force in vertical penetration of ice, Part II tackles the problem maximum force that can be applied by a moving ice plate on an obstacle presented by a fixed structure. Based on an asymptotic approach, approximate solutions for are obtained for the size effects of ice thickness, effective structure diameter and, in the case of a finite ice floe, the size of the floe. [DOI: 10.1115/1.1429933]

1 Introduction

After analyzing in Part I the vertical penetration problem, we will examine in the present Part II another fundamental problem of large-scale fracture of sea ice—the maximum force P that can be exerted by a moving ice plate of thickness h on an obstacle presented by a fixed structure of effective diameter d (imagined as a cylinder). Similar simplifications will be made and the approach of asymptotic matching will again be followed. All the definition and notations from Part I will be retained. Several possible mechanisms of breakup will be considered.

Stress analysis and fracture of floating ice plates subjected to a horizontal load has been studied by Ashton, Atkins, Goldstein and Osipenko, Lavrov, Palmer et al., Ponter and Slepyan, and others ([1–7]). These investigators used dimensional analysis to determine the scaling laws of linear elastic fracture mechanics (LEFM) and of strength theory. They did not consider cohesive fracture and did not attempt to bridge these two theories to describe the size effect transition from one to the other. Characterizing this transition is the main objective of what follows.

For the horizontal load, it is convenient to define the nominal strength as the average stress on the cross-section area hd of the structure facing the moving ice plate, i.e.,

$$\sigma_N = P/hd. \tag{1}$$

2 Global Failure due to Buckling of Ice Plate

Cylindrical buckling, in which the deflection surface is a translationary surface, can occur only if the ice plate is moving against a very long wall ($d \to \infty$). In this case the plate behaves as a beam on elastic foundation, which is a one-dimensional problem, and the critical compressive normal force per unit width of the plate is known to be (e.g., [8]) $N_{cr} = \kappa_0 \sqrt{\rho D}$ where coefficient κ_0 depends on the boundary conditions. Its minimum value occurs for a semi-infinite plate with a straight infinite free edge and is $\kappa_0 = 1$.

If the obstacle, such as the legs of an oil drilling platform, has a finite dimension d in the transverse direction, the buckling mode is two-dimensional and more complicated. In any case, however, dimensional analysis ([9,10]) suffices to determine the form of the buckling formula and the scaling ([3,7]).

There are five variables in the problem, P_{cr}, E', ρ, h, d, and the solution must have the form $F(P_{cr}, E', \rho, h, d) = 0$, where $E' = E/(1-\nu^2)$; and P_{cr} is the critical force exerted by the resisting structure on the moving ice plate (Fig. 1(*a*)). There are, however, only two independent physical dimensions in the problem, namely the length and the force. Therefore, according to Buckingham's Π theorem of dimensional analysis ([9–11]), the solution must be expressible in terms of $5-2$, i.e., 3 dimensionless parameters. They may be taken as $P_{cr}/E'hd$, $\sqrt{\rho D}/E'h$ and d/h. Because the buckling is linearly elastic, $P_{cr}/E'hd$ must be proportional to $\sqrt{\rho E'}/E'h$ and d/h. Denoting

$$\sigma_{N_{cr}} = P_{cr}/hd \tag{2}$$

which represents the nominal buckling strength (or the average critical stress applied by the face of the resisting structures on the moving ice plate), and noting that $D = E'h^3/12$ with $E' = E/(1-\nu^2)$, we conclude that the buckling solution must have the form

$$\sigma_{N_{cr}} = \kappa(d/h)\sqrt{\rho E'}\sqrt{h} \tag{3}$$

where κ is a dimensionless parameter depending on the relative diameter of the structure, d/h, as well as on the boundary conditions. For $d/h \to \infty$ (i.e., an infinite wall), this must reduce [8] to the critical stress for an infinite beam on elastic foundation loaded at the free end (vertically sliding end). Therefore, $\kappa(0)/\sqrt{12} = 1$ or $\kappa(0) = 2\sqrt{3}$, which represents the smallest possible value of κ for any d/h. This fact becomes obvious by imagining a strip of width d in the direction of movement to be separated from the rest of the ice plate; for that strip $\kappa_0 = 1$ (if the ice in contact with the structure is free to slide vertically, Fig. 1(*a*)), and re-attaching the rest of the plate cannot but increase the critical load.

An interesting property of (3) is that, for geometrically by similar structures (constant d/h), $\sigma_{N_{cr}}$ increases, rather than decreases, with ice thickness h. So there is a *reverse* size effect. Consequently, the buckling of the ice plate can be the mechanism of failure only when the plate is sufficiently thin. The reason for the reverse size effect is that the buckling wavelength (the distance between the inflexion points of the deflection profile), which is $L_{cr} = \pi(D/\rho)^{1/4}$ (as follows from dimensional analysis, or from nondimensionalization of the differential equation of plate buckling), is not proportional to h; rather

$$L_{cr}/h \propto h^{-1/4}, \tag{4}$$

i.e., L_{cr} decreases with h. This contrasts with the structural buckling problems of columns, frames, and plates, in which L_{cr} is proportional to the structure size. (Despite the analogy between axisymmetric buckling of an axially compressed cylindrical shell and a floating ice, no size effect occurs for the shell because, unlike ρ_w, the equivalent foundation modulus of the shell scales with h.)

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, Aug. 7, 2000; final revision, July 19, 2001. Associate Editor: A. Needleman. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

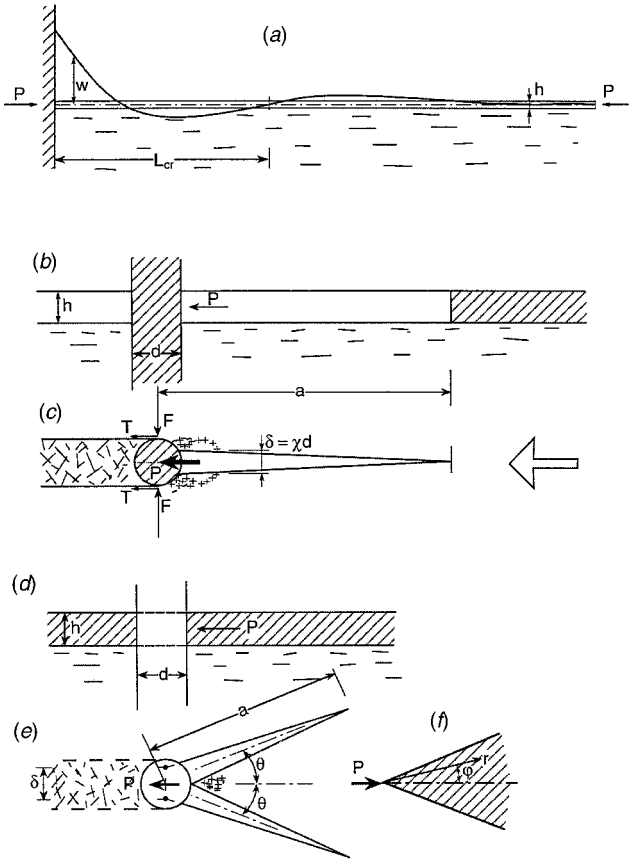


Fig. 1 (*a*) Buckling of ice plate pushing horizontally against a fixed structure, (*b,c,d*) radial cleavage crack, and (*e–g*) diverging *V*-cracks

3 Global Failure due to Radial Cleavage Fracture

Another failure mechanism consists of a long radial cleavage crack in the ice plate, propagating against the direction of ice movement (Fig. 1(*b,c*)). The resistance of the plate against being cleaved causes the ice to exert on the structure a pair of transverse force resultants F and a pair of tangential forces T in the direction of movement; $T=F\tan\varphi$ where φ may be regarded as the friction angle. Forces T have no effect on the stress intensity factor K_I at the crack tip.

First we will consider the asymptotic case of a a structure of a very large size d producing a crack of a very long length a (Fig. 1(*b,c*)). LEFM must apply in this asymptotic case. To determine a, we need to calculate the crack opening δ caused by F. Considering the ice plate as infinite, we have ([12,13])

$$K_I=\frac{F}{h}\sqrt{\frac{2}{\pi a}}. \tag{5}$$

The energy release rate is

$$\mathcal{G}=\frac{1}{h}\left[\frac{\partial\Pi^*}{\partial a}\right]_F=\frac{1}{h}\frac{\partial}{\partial a}\left[\frac{1}{2}C(a)F^2\right]=\frac{F^2}{2h}\frac{dC(a)}{da} \tag{6}$$

where $C(a)$ is the load-point compliance of forces F. Upon using (5) and Irwin's relation ([14]), we have at the same time

$$\mathcal{G}=\frac{K_I^2}{E}=\frac{2F^2}{\pi Eh^2a}. \tag{7}$$

Equating (6) and (7), we thus get

$$\frac{dC(a)}{da}=\frac{4}{\pi Eha}. \tag{8}$$

This expression may now be integrated from $a=d/2$ (the surface of the structure, considered as circular, Fig. 1(*b,c*)) to a (note that integration from $a=0$ would give infinite C but would be meaningless because a cannot be less than d). In this manner, we obtain $C(a)$, and from it the opening deflection δ:

$$\delta=C(a)F=\frac{4F}{\pi Eh}\ln\left(\frac{2a}{d}\right). \tag{9}$$

If the radial cleavage fracture were the only mode of ice breaking, we would have $\delta=d$. However, as will be discussed later, there is likely to be at least some amount of local crushing at, and ahead, of the structure. Consequently, the relative displacement between the two flanks of the crack is no doubt less than d. We denote it as χd where χ is a coefficient less than 1. Upon setting $\delta=\chi d$, (9) yields

$$a=\frac{d}{2}\exp\left(\frac{\pi Eh\chi d}{4F}\right). \tag{10}$$

(note that a/d is not constant but increases with d; hence, the fracture modes are not geometrically similar, and so the LEFM power scaling cannot be expected to apply). Substituting (10) into (5) and setting $K_I=K_c=\sqrt{EG_f}$ (Irwin's relation, $K_c=$fracture toughness of ice), we obtain

$$\frac{2F}{h\sqrt{\pi EG_fd}}=\exp\left(\frac{\pi Eh\chi d}{8F}\right). \tag{11}$$

The pair of forces F is related to load P on the structure ($P=2T$, Fig. 1(*c*)) by a friction law, which may be written as

$$P=2F\tan\varphi. \tag{12}$$

where φ is the friction angle. Substituting $F=P/2\tan\varphi$ and $P=\sigma_N hd$ into (11), we obtain, after rearrangements,

$$\frac{d}{dc}=\frac{1}{\tau^2}e^{1/\tau},\quad \tau=\frac{\sigma_N}{\eta_c} \tag{13}$$

in which τ is the dimensionless nominal strength, and d_c and σ_c are constants defined as

$$d_c=\frac{4G_f}{\pi\chi^2E},\quad \sigma_c=\frac{\pi}{2}\chi E\tan\varphi. \tag{14}$$

Equation (13), plotted in Fig. 2, represents the law of radial cleavage size effect in an inverted form. The small-size asymptotic behavior is the LEFM scaling for similar structures with similar cracks:

$$\text{for } d\ll d_c:\quad \sigma_N\approx\sqrt{d_c/d}. \tag{15}$$

The plot of (13) in Fig. 2 shows that the size effect is getting progressively weaker with increasing structure diameter d (although no horizontal asymptote is approached by the curve). The reason for this is that the crack is dissimilar, i.e., the ratio, a/d, of crack length to structure diameter is not the same for different sizes but increases according to (10) with the structure size. (In designing ocean platforms, it is nevertheless always advantageous, with respect to the radial cleavage mechanism, to use a smaller number of larger legs, which has of course been intuitively followed in practice.)

So far our radial cleavage crack analysis has been based on LEFM. In other words, the length $2c_f$ of the cohesive zone at the tip of the radial cleavage crack was considered negligible compared to a. Let us now consider the opposite asymptotic case of a very small structure diameter d and a very short crack a such that $a\ll c_f$. In that asymptotic case, the crack faces up to $x=a_0=a-d/2$ are subjected to uniform cohesive tractions f_t'. Noting that the stress intensity factor for a semi-infinite crack in an infinite

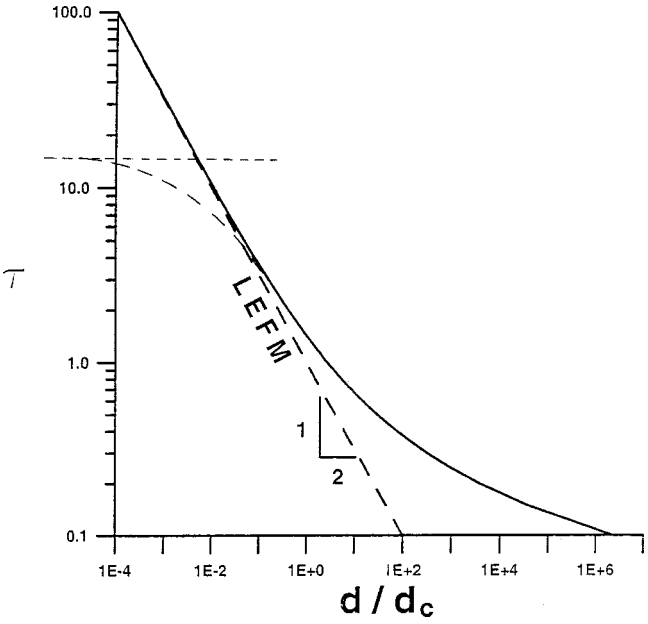


Fig. 2 Size effect associated with radial cleavage fracture (solid curve—LEFM solution, dashed curve—cohesive crack solution)

space caused by a pair of unit concentrated force acting on the crack faces at distance x from the crack tip is $K_I=(\sqrt{2\pi/x})/h$ ([12]), we find that K_I caused by uniform tractions f_t' is

$$K_{I_f}=-\int_0^{a-d/2}\sqrt{\frac{2}{\pi x}}\,f_t'\,dx=-f_t'\sqrt{\frac{8a}{\pi}}. \tag{16}$$

The stress intensity factor due to concentrated reaction F at distance a from the cohesive crack tip is, according to (5), $K_{IF}=(F/h)\sqrt{2/\pi a}$. It is necessary that the total stress intensity factor $K_{It}=K_{IF}+K_{I_f}=0$. From this condition and the friction relation (12), it follows that

$$P=4ahf_t'\tan\varphi. \tag{17}$$

To calculate the deflection δ_f due to cohesive stresses f_t', one could use Green's function. However, this leads to a complicated integral. Since a high accuracy is not needed, we prefer an approximate calculation. To this end, we imagine the cohesive crack length x to grow from 0 to $a_0=a-d/2$ while constant tractions f_t' act along the entire crack length in front of the structure (and work on the growing opening). In view of (16) and Irwin's relation, the total energy released during the imagined growth of this crack is

$$\Pi^*=\int_0^{a_0}\frac{K_I^2}{E}\,dx=\int_0^{a_0}\frac{1}{E}\left(f_t'\sqrt{\frac{8x}{\pi}}\right)^2 dx=\frac{4f_t'^2a_0^2}{\pi E} \tag{18}$$

which is a function of f_t' representing the complementary energy. According to Castigliano's theorem, differentiation of Π^* with respect to the total cohesive force a_0hf_t' provides the displacement parameter on which the cohesive stress f_t' works, which is the average crack-opening displacement $\overline{v_f}$ over the length a_0 of application of f_t';

$$\overline{v_f}=\frac{1}{a_0h}\frac{\partial\Pi^*}{\partial f_t'}=\frac{8f_t'a_0}{\pi Eh}. \tag{19}$$

Since we avoided Green's function, we now need to approximate the relationship between $\overline{v_f}$ and opening displacement v_f at the center of the structure, $x=a$ (Fig. 1(*c*)). We may assume that the face of the opened crack is approximately straight, which is a simplification widely used in materials science. Under that assumption, $v_f=\kappa_f\overline{v_f}a/a_0$, where $ka_f=2$ if the crack face remains straight. Then, from (19),

$$v_f=\frac{8f_t'}{\pi E}\kappa_f a. \tag{20}$$

The opening displacement $2v_F$ due to the pair of concentrated forces F has already been calculated in (9); $2v_F=(4F/\pi Eh)\ln(2a/d)$. Compatibility of transverse displacements at the center of structure $(x=a)$ requires that

$$2v_F-2v_f=\chi. \tag{21}$$

Substituting the foregoing expressions for v_F and v_f, and setting $F=P/2\tan\varphi=\sigma_N hd/2\tan\varphi$, one obtains, after rearrangements, the equation

$$\ln\left(\frac{\sigma_N}{2f_t'\tan\varphi}\right)=\kappa_f+\frac{\pi E\chi\tan\varphi}{2\sigma_N}. \tag{22}$$

This is a transcendental equation whose solution $\sigma_N=\sigma_N^0$, represents the average pressure applied on the area of the structure facing the moving ice. Since d and h do not appear in this equation, the σ_N^0 value is a constant, represented in Fig. 2 by the horizontal line. So, as expected, there is asymptotically no size effect if $d\to\infty$.

To obtain the approximated law of the size effect for the intermediate sizes, the small-size and large-size asymptotic behaviors must be suitably matched. Similar to many previous approximations of quasi-brittle size effect ([13,14]), the asymptotic matching may be accomplished by replacing size d in (13) with the expression $(d^r+d_0^r)^{1/r}$ where d_0 is a constant. With this replacement, (13) provides the following general approximate asymptotic matching law for the size effect:

$$(d^r+d_0^r)^{1/r}=\frac{d_c}{\tau^2}e^{1/\tau}. \tag{23}$$

Here r is an empirical constant, probably close to 1. For $d\to\infty$, this equation asymptotically approaches the LEFM Eq. (13), and for $d\to 0$ the following equation for constant d_0 is obtained:

$$d_0=\alpha_c(\sigma_c^2/\sigma_N^{0\,2})e^{\sigma_c/\sigma_N^0} \tag{24}$$

where σ_N^0 is the solution σ_N of (22). In analogy to other scaling problems, the value $r=1$ is often reasonable, and then (23) simplifies to the size effect formula:

$$d=\frac{d_c}{\tau^2}e^{1/\tau}-d_0. \tag{25}$$

Equation (23) for the quasi-brittle size effect due to a horizontal load is plotted as the dashed curve in Fig. 2. The shape of this plot documents the difficulty in deducing the size effect from small-scale experiments. If the tests are confined to the nearly horizontal intial portion of the dashed curve, there is no way to predict the size effect at large sizes unless a realistic theory is employed.

4 Compression Fracture of Ice Plate

As typically observed in the field, moving ice gets crushed in front of an obstacle, breaking up into chunks. The cause is local compression fracture of the material. Its initiation may be explained by sliding on inclined weak plains between ice crystals, which leads to axial splitting microcracks called the wing-tip cracks (for ice see, e.g., Schulson [15,16]) extending in the direction of compression for a certain finite length. This mechanism, however, explains only the generation of local compressive damage in the material but does not explain to overall failure of the plate and the size effect.

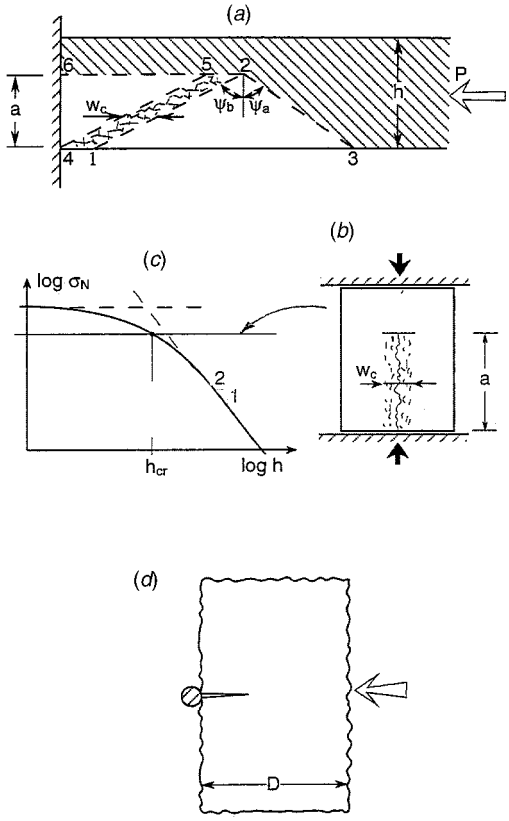


Fig. 3 (*a*) Compression fracture of ice plate, (*b*) axial splitting fracture, (*c*) size effects corresponding to (*a*) and (*b*), and (*d*) overall fracture of ice floe

To produce overall breakup of ice, the damage must propagate. As transpired in connection with studies of concrete and borehole breakout in rock ([13,17]), the propagation typically occurs in the form of a narrow band consisting predominantly of axial splitting microcracks (generated, e.g., by the wing-tip crack mechanism). The band of axial splitting microcracks can propagate either in the axial direction of the compressive stress, or laterally. The latter is shown in Fig. 3(*a*), and the former in Fig. 3(*b*).

In the spirit of fracture mechanics, one must estimate the energy release. Consider the plausible situation depicted in Fig. 3(*a*), where the band of a certain characteristic width w_c in the direction of compression has inclination ψ_b and reaches to depth a below the surface of plate. Formation of the band must evidently relieve the axial stress σ_N not only within the band area 12541, but also in the adjacent zones 1231 and 4564. The boundary of the stress relief zone is considered to have a certain characteristic inclination ψ_a, independent of the plate thickness. The combined area of the stress relief zone 43264 is $a(w_c + a/2 \tan \psi_a + a \tan \psi_b)$. Before the formation of the damage band, the initial strain energy density in this zone is $\sigma_N^2/2E$, and after the formation of the band it may be assumed as zero (more generally, one could quite easily take into account some finite residual strength σ_r of ice after crushing, see [17]; but this is omitted since no information on σ_r is available). Thus the total energy release caused by formation of the damage band per unit width is, approximately,

$$\Pi^* = \frac{\sigma_N^2}{2E} a \left(w_c + \frac{1}{2} a \tan \psi_a + a \tan \psi_b \right). \tag{26}$$

The rate of energy dissipation per unit width as the band propagates must be equal to the fracture energy of the band, G_b, which equals $G_f w_c / s_c$ where G_f is the fracture energy of the axial splitting microcracks in the band, and s_c their average spacing. Energy balance during the quasi-static extension of the band requires that the rate of energy release be equal to G_f, i.e.,

$$\frac{\partial \Pi^*}{\partial a} = \frac{\sigma_N^2}{2E} (w_c + a \tan \psi_a + 2a \tan \psi_b) = G_b . \tag{27}$$

Solving this equation for σ_N, we get, after rearrangements,

$$\sigma_N = \sigma_a \left(1 + \frac{h}{h_0} \right)^{-1/2} \tag{28}$$

in which the following notations are made:

$$h_0 = \frac{h}{a} \frac{w_c}{(\tan \psi_a + 2 \tan \psi_b)}, \quad \sigma_a = \sqrt{\frac{2EG_c}{w_c}}. \tag{29}$$

Here we deliberately introduced the plate thickness h even though it cancels out of the equation. The reason is that it appears reasonable to assume the ratio a/h for plates of various thicknesses to be approximately constant. In other words, the geometries of the damage band at failure of the plates of various thicknesses are assumed similar. This assumption is based on experience with some other fracture problems, for which it was shown to lead to realistic results. Anyway, it is intuitively clear that it would be unreasonable to assume that for thin plates the damage band at maximum σ_N penetrates through most of the thickness and for thick plates penetrates only to a very shallow depth.

Equation (28), plotted in Fig. 3(*c*), is the same as the classical size effect law proposed by Bažant [18] for quasi-brittle structures failing after a long stable growth of tensile fracture. Among the mechanisms explored here, it is the only one that can explain the size effect of ice thickness.

The ultimate cause of size effect in compressive (as well as tensile) fracture is that the volume of the energy dissipation zone, i.e., the damage band, grows linearly with the distance a of propagation while the volume of the energy release zone grows faster than linearly, having a quadratically growing term that dominates for large sizes. Thus it is intuitively clear that if the stress in these zone at failure were the same, energy balance could exist only for one size but not for other sizes ([14]). So, in a larger structure the stress in the quadratically growing zones (1231 and 4564 in Fig. 2(*a*)) must be less.

There is of course another possibility—namely that the damage band grows axially, in the direction of compression, which leads to a splitting failure (Fig. 3(*b*)). In that case the stress in the material on the sides of the crack band is not relieved, and so the energy release occurs only within the damage band itself. In that case, not only the energy dissipation but also the energy release are proportional to the length a of the band, which means that energy rates for the same failure stress σ_N can balance for any size h. So, for the axial propagation, there is no size effect.

The axial growth is more likely because no new wing-tip cracks need to be nucleated. Therefore, at small enough sizes the axial splitting of ice should prevail, which means that the splitting mechanism corresponds in the logarithmic size effect plot (Fig. 3(*c*)) to a horizontal line starting below the curve of the size effect law for lateral propagation of the damage band. However, the horizontal line must eventually cross the size effect curve at a certain critical size h_{cr}, above which the lateral propagation of damage band must prevail, and then a size effect must exist.

The present analysis is similar to that made for concrete; see [17], where various fine details are discussed (also [13,14]).

Finally, an explanation of empirical parameter χ introduced for the cleavage fracture: It is presumed that the part $(1-\chi)d$ of the cross section facing the ice movement undergoes compression crushing. This part should be governed by Eq. (28), and so the force given by that equation needs to be added to the force P based on (12) and (13).

5 Overall Fracture of Finite Ice Floe

Collision of a large ice floe with a fixed structure may cause a fracture of the whole floe. The floe is loaded by distributed inertia forces of its mass, but the problem may be treated as essentially quasi-static, owing to the low velocity of movement. Except for the loading by distributed forces, the problem is similar to fracture tests in the laboratory, especially the three-point bend beam (Fig. 3(*d*)). Dempsey's record-breaking tests on the Arctic Ocean near Resolute can be regarded as an approximate reduced-scale simulation of this kind of fracture ([19,20]). The analysis may follow similar lines as presented, for instance, in [13] for other materials. From that analogy it follows that the size L of the floe may cause one of two types of size effect:

$$(1) \quad \frac{P}{Lh}=S_0\left(1+\frac{L^r}{L_0^r}\right)^{-1/2r} \tag{30}$$

$$(2) \quad \frac{P}{Lh}=S_\infty\left(1+\frac{rL_b}{L}\right)^{1/r} \tag{31}$$

where P/Lh is the nominal strength of the whole floe; S_0, L_0, S_∞, L_b are constants that can be calculated by fracture mechanics; and r is a parameter whose value is normally between 0.5 and 2.

The first kind of size effect, which agrees very well with Dempsey et al.'s [19] field tests in the Arctic, applies when a large crack in the floe can form before the overall fracture of the floe takes place. The second kind applies to failures at fracture initiation, exemplified by the test of modulus of rupture (bending strength), and is pertinent if the maximum load is attained before a stable finite crack can develop (e.g, by means of the radial cleavage mechanism).

6 Comments on Some Periodic Failure Mechanisms

According to observations, diverging V-shaped cracks may also form ahead of an obstacle (e.g., [21], ch. 7); Fig. 1(*e,f*). To estimate in a simple manner a rough approximate value of complementary energy Π^* of an infinite ice plate after formation of such cracks, we may assume that the force P from the structure produces stress only within the wedge between the cracks (Fig. 1(*g*)). From a well-known solution ([22]),

$$\sigma_r=-Pk_\theta\cos\varphi/rh, \quad \sigma_\varphi=\sigma_{r\varphi}=0 \tag{32}$$

where σ_r, σ_φ, and $\sigma_{r\varphi}$ are the stress components in polar coordinates r, φ, and

$$k_\theta=1\Big/\left(\theta+\frac{1}{2}\sin 2\theta\right), \tag{33}$$

θ being the inclination angle of the cracks (Fig. 1(*f*)). The displacement at $r=d/2$ (structure surface) is

$$u=\int_{d/2}^{\infty}\frac{\sigma_r}{E}dr=\frac{Pk_\theta}{Eh}\ln\frac{2a}{d}. \tag{34}$$

Then $\Pi^*=Pu/2=(P^2k_\theta/2Eh)\ln(2a/d)$. The complementary energy before fracture may be estimated as the value of Π^* for $\theta=\pi$, i.e., $\Pi_0^*\approx(P^2/2\pi Eh)\ln(2a/d)$. The total energy release due to V-cracks in the ice plate is $\Delta\Pi^*=\Pi^*-\Pi_0^*$, and the derivative $\partial\Delta\Pi^*/\partial a$ at constant P must be equal to $2hG_f$. This condition yields

$$P\approx 2h\sqrt{\frac{EG_f}{\pi^{-1}-k_\theta}}\sqrt{a}. \tag{35}$$

To determine crack length a and angle θ, one may use two conditions: (a) the opening displacement at the crack mouth, δ, must be equal to $\chi d/(2\cos\theta)$, which means that the load-point displacement of force P must be $u=(\chi d/2)\tan\theta$, and (b) the expression for P should be minimized with respect to θ. These two conditions, however, make the solution quite complicated. We will not pursue it here because of this and also because of two unresolved questions: (1) An axial cleavage crack may be also present ([5]), and it may form either before or after the V-cracks. (2) Simultaneous compression crushing is very likely in the case of V-cracks, which makes the value of χ, and thus the length a of V-cracks, rather uncertain.

Unlike the cleavage fracture, the V-shaped cracks can occur only from time to time. They do not represent a steady-state mechanism that would accommodate continuous movement of the ice.

Other failure mechanisms occur in the case of an inclined face of the fixed structure, or in the case of an icebreaker ([5]). These mechanisms involve axial bending cracks as well as bending cracks normal to the direction of motion. Studying the action of an icebreaker, Goldstein and Osipenko [3] considered periodic formation of LEFM bending cracks at some distance in front of the icebreaker, normal to the direction of movement. They limited attention to one-dimensional cylindrical bending of the ice plate and did not consider simultaneous formation of axial or other cracks.

7 Conclusions (From Parts I and II)

1. The known mechanism of failure of a floating ice plate subjected to a vertical load can be used in an approximate energy analysis of quasibrittle fracture. The results do not disagree with the limited field experiments that exist. They approximately agree with previous numerical simulations and confirm that for large ice thicknesses there is a strong size effect, approaching the size effect of LEFM. Asymptotic matching leads to a simple formula for the size effect, which is similar to the size effect law proposed in 1984 by Bažant.
2. Simplified fracture analysis of the nominal strength of ice plate pushed against a fixed structure brings to light several possible mechanisms of failure with size effects due to ice thickness, the diameter of the structure and, if the size of the ice floe is finite, the size of the floe. Buckling of the floating plate causes a reverse size effect of ice thickness (i.e., the nominal strength increasing with ice thickness) and therefore plays any role only for sufficiently thin ice. Radial cleavage of the ice plate against the direction of ice movement causes a size effect of structure diameter which follows linear elastic fracture mechanics (LEFM) for small enough diameters and becomes progressively weaker with an increasing diameter. Compression fracture, with ice crushing localized into transversely propagating bands, causes a size effect of ice thickness that follows approximately the classical size effect law proposed in 1984 by Bažant. The overall fracture of a finite ice floe causes a size effect of the floe size, following again the same size effect law.
3. The present approach contrasts with the classical approach based on either plastic limit analysis or elastic analysis with a strength limit, both of which lead to no size effect.

Acknowledgment

Grateful acknowledgment is due to the National Science Foundation and the Office of Naval Research for partial financial support of this two-part study under grants CMS-9713944 and N00014-91-J-1109, respectively, to Northwestern University.

References

[1] Ashton, G., ed., 1986, *River and Lake Ice Engineering*, Water Resources Publications.

[2] Atkins, A. G., 1975, "Icebreaking Modeling," J. Ship Res., **19**, No. 1, pp. 40–43.

[3] Goldstein, R. V., and Osipenko, N. M., 1993, "Fracture Mechanics in Modeling of Icebreaking Capability of ships," J. of Cold Regions Engrg. ASCE, **7**, No. 2, pp. 33–43.

[4] Lavrov, V. V., 1958, ''The Nature of the Scale Effect in Ice and the Ice Sheet,'' Sov. Phys. Dokl., **3**, pp. 934–937 (transl. from Russian).
[5] Palmer, A. C., Goodman, D. J., Ashby, M. F., Evans, A. G., Hutchinson, J. W., and Ponter, A. R. S., 1983, ''Fracture and Its Role in Determining Ice Forces on Offshore Structures,'' Ann. Glaciol., **4**, pp. 216–221.
[6] Ponter, A. R. S., Palmer, A. C., Goodman, D. J., Ashby, M. F., Evans, A. G., and Hutchinson, J. W., 1983, ''The Force Exerted by a Moving Ice Sheet on an Offshore Structure. 1. The Creep Mode,'' Cold Regions Sci. & Tech., **8**, No. 2, pp. 109–118.
[7] Slepyan, L. I., 1990, ''Modeling of Fracture of Sheet Ice,'' Mech. Solids, 155–161.
[8] Bažant, Z. P., and Cedolin, L., 1991, *Stability of Structures: Elastic, Inelastic, Fracture and Damage Theories*, Oxford University Press, New York.
[9] Sedov, L. I., 1959, *Similarity and Dimensional Methods in Mechanics*, Academic Press, San Diego CA.
[10] Barenblatt, G. I., 1987, *Dimensional Analysis*, Gordon and Breach, New York.
[11] Barenblatt, G. I., 1979, *Similarity, Self-Similarity and Intermediate Asymptotics*, Consultants Bureau (Plenum Press), New York, (transl. from Russian original, 1978).
[12] Tada, H., Paris, P. C., and Irwin, J. K., 1985, *The Stress Analysis of Cracks Handbook*, 2nd Ed., Paris Productions, St. Louis, MO.
[13] Bažant, Z. P., and Planas, J. 1998, *Fracture and Size Effect in Concrete and Other Quasibrittle Materials*, CRC, Boca Raton, FL.
[14] Bažant, Z. P., and Chen, E.-P., 1997, ''Scaling of Structural Failure,'' Appl. Mech. Rev., **50**, No. 10, pp. 593–627.
[15] Schulson, E. M., 1990, ''The Brittle Compressive Fracture of Ice,'' Acta Metall. Mater., **38**, No. 10, pp. 1963–1976.
[16] Schulson, E. M., 2000, ''Brittle Failure of Ice,'' Eng. Fract. Mech., **68**, No. 17–18, pp. 1839–1887.
[17] Bažant, Z. P., and Xiang, Yuyin, 1997, ''Size Effect in Compression Fracture: Splitting Crack Band Propagation,'' J. Eng. Mech., **123**, No. 2, pp. 162–172.
[18] Bažant, Z. P., 1984, ''Size Effect in Blunt Fracture: Concrete, Rock, and Metal,'' J. Eng. Mech., **110**, pp. 518–535.
[19] Dempsey, J. P., Adamson, R. M., and Mulmule, S. V., 1999, ''Scale Effects on the in situ Tensile Strength and Fracture of Ice: Part II. First-Year Sea Ice at Resolute, N. W. T.,'' Int. J. Fract., **95**, pp. 346–378.
[20] Mulmule, S. V., Dempsey, J. P., and Adamson, R. M., 1995, ''Large-Scale in-situ Ice Fracture Experiments—Part II: Modeling Efforts, in Ice Mechanics—1995,'' *ASME Joint Applied Mechanics and Materials Summer Conference*, Vol. AMD - MD 1995, Los Angeles, June 28–30.
[21] Sanderson, T. J. O., 1988, *Ice Mechanics: Risks to Offshore Structures*, Graham and Trotman Ltd., London.
[22] Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*. 3rd Ed., McGraw-Hill, New York, p. 110.

Microstructural Randomness Versus Representative Volume Element in Thermomechanics

M. Ostoja-Starzewski
Department of Mechanical Engineering,
McGill University,
817 Sherbrooke Street West,
Montréal, Québec H3A 2K6, Canada
e-mail: martin.ostoja@mcgill.ca
Fellow ASME

Continuum thermomechanics hinges on the concept of a representative volume element (RVE), which is well defined in two situations only: (i) unit cell in a periodic microstructure, and (ii) statistically representative volume containing a very large (mathematically infinite) set of microscale elements (e.g., grains). Response of finite domains of material, however, displays statistical scatter and is dependent on the scale and boundary conditions. In order to accomplish stochastic homogenization of material response, scale-dependent hierarchies of bounds are extended to dissipative/irreversible phenomena within the framework of thermomechanics with internal variables. In particular, the free-energy function and the dissipation function become stochastic functionals whose scatter tends to decrease to zero as the material volume is increased. These functionals are linked to their duals via Legendre transforms either in the spaces of ensemble average velocities or ensemble-average dissipative forces. In the limit of infinite volumes (RVE limit (ii) above) all the functionals become deterministic, and classical Legendre transforms of deterministic thermomechanics hold. As an application, stochastic continuum damage mechanics of elastic-brittle solids is developed. [DOI: 10.1115/1.1410366]

1 Introduction

Presence of dissipative phenomena in mechanics of solids and fluids necessitates a formulation of continuum mechanics consistent with principles of thermodynamics; such a theory is briefly called *thermomechanics* or *continuum thermodynamics*. As lucidly and comprehensively elaborated in a recent book by Maugin [1], this challenge conventionally leads to a consideration of one of four continuum thermodynamics:

- thermodynamics of irreversible processes (TIP);
- thermodynamics with internal variables (TIV);
- rational thermodynamics (RT); and
- extended (rational) thermodynamics (ET).

A feature common to all of these approaches is a postulate of existence of a representative volume element (RVE). In other words, we are looking here at deterministic, homogeneous continuum theories, without clear account of random microstructures which are, in fact, prevalent in real materials. While we recognize here that some statistical treatments were carried out as a bridge from micro to macro levels for select variants of the above theories (e.g., [2,3]), such studies were concerned with providing foundations from the standpoint of statistical physics directly to the level of the RVE, without making clear what the size of the RVE actually was. On the other hand, homogenization procedure invoked to pass to the RVE in studies of plasticity and damage (e.g., [4,5]) always involves a periodic microstructure; see also ([6]) for other physical problems, and ([7]) for elastic/inelastic problems in composites. Some finite scale periodicity in random microstructures is also invoked in theoretical and numerical studies of the RVE size ([8,9]); in fact, this assumption allows homogenization of elastic materials on very small length scales.

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, Aug. 31, 2000; final revision, June 12, 2001. Associate Editor: J. W. Ju. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

While in plasticity the approach to RVE in a random microstructure may be rapid ([10]), this is not so for damage phenomena where presence of scatter is evident for even the largest specimens that can be handled in the laboratory, e.g., ([11,12]). In fact, the dichotomy between statistical damage models motivated by such observations and continuum damage mechanics based on the deterministic TIV formalism is perceived as one of the grand challenges of damage mechanics ([13,14]). This is one of major motivations of this paper.

As theoretical models we first consider strict-sense and wide-sense stationary random fields, possessing ergodic properties. Many models of microstructural randomness—e.g., Boolean models and tessellations—possess such homogeneity and ergodic characteristics, and they are highly desirable in stochastic homogenization. Real materials, however, often lack these nice behaviors, and, as illustrated by measurements on machine made paper, one may have to work with quasi-stationary and quasi-ergodic random fields.

As a guidance in setting up a statistical volume element (SVE) and its deterministic limit, the RVE, in thermomechanics we take the work on elastic microstructures carried out over the last decade, that relies on the Hill condition ([15]). In essence, it says that the RVE response is independent of the type of boundary conditions applied to it (i.e., uniform stress or uniform strain or their orthogonal combination). For finite-size—which we call *mesoscale*—material domains the Hill condition leads to three types of boundary conditions ([16]), and three types of apparent responses: uniform kinematic, uniform traction, and uniform mixed (orthogonal). It follows that in the case of dissipative behaviors, we must primarily consider boundary conditions of uniform dissipative force or uniform velocity.

As continuum thermodynamics setting we take TIV, and, in particular, its variant due to Ziegler [17] (also [18]) which defines a broad class of continuous media from the free energy and dissipation functions ([19]). In many cases, the uniform kinematic and uniform traction boundary conditions, respectively, bound the effective (in the macroscopic/global sense) dissipative response from above and below; the larger are the mesoscale domains of the material considered, the tighter are the bounds. These bounds, define a sequence of SVE, convergent to the RVE, and serving as a basis of statistical continuum models. We discuss the bounds for

thermal conductivity and damage phenomena. While mathematically at any finite mesoscale the bounds are distinct, the approach to RVE, with increasing window size, to the effective (macroscopic) response, depending on the dissipative process considered, may be very rapid, moderate, or very slow. Furthermore, the free energy may display a different scaling trend than the dissipation function for a given microstructure.

2 Representative Volume Element (RVE) Postulate and Structure of Random Media

2.1 Homogeneous and Ergodic Random Media. Let us first recall the classical assumption of a representative volume element (RVE) according to Hill [15]: it is "a sample that (a) is structurally entirely typical of the whole mixture on average, and (b) contains a sufficient number of inclusions for the apparent overall moduli to be effectively independent of the surface values of traction and displacement, so long as these values are 'macroscopically uniform.'" In other words, we need, respectively:

(a) *statistical homogeneity* and *ergodicity* of the material; these two properties assure the RVE to be *statistically representative* of the macroresponse (e.g., [20,21]);
(b) some scale L of the material domain, sufficiently large relative to the microscale d (inclusion size) so as to ensure the *independence of boundary conditions*.

Mechanics of random media, together with probability theory, provides a rigorous setting for study of these issues (e.g., [22]). That is, the field problem of random medium

$$\mathbf{B}=\{B(\omega);\omega\in\Omega\} \tag{2.1}$$

is governed by an equation

$$\mathcal{L}(\omega)\mathbf{u}=\mathbf{f}\quad \omega\in\Omega \tag{2.2}$$

accompanied by appropriate boundary and/or initial conditions. Here $\mathcal{L}(\omega)$ is a random field operator (with randomness caused by, say, elastic moduli being a random field), $\mathbf{u}$ is a solution field, and $\mathbf{f}$ is forcing function. Parametrization by ω (element of the sample space Ω, endowed with a probability measure P) indicates the source of uncertainty. Clearly, there are two more basic ways to introduce randomness in a mechanics problem:

- randomness in the forcing function—replacing (2.2) by $\mathcal{L}u=\mathbf{f}(\omega)$—as exemplified by problems of random vibrations;
- randomness of boundary and/or initial conditions.

As we are interested here in the case described by (2.2), we note that each of the realizations $B(\omega)$ follows laws of deterministic mechanics in that it is a specific heterogeneous material sample. The problem of setting up the RVE of volume $V=L^D$ (D is space dimension), in the sense of (a) and (b) above, in a global boundary value problem on length scales L_{macro} is illustrated with the help of Fig. 1; of course only one $B(\omega)$ is shown. In essence, we want

$$\langle\mathcal{L}^{-1}\rangle^{-1}\mathbf{u}=\mathbf{f} \tag{2.3}$$

with independence of boundary conditions as posited by Hill's condition, on length scale L satisfying

$$d\ll L\ll L_{\text{macro}}. \tag{2.4}$$

If that is the case, one can then simply deal with a deterministic continuum thermomechanics problem on scale L_{macro}. Hereinafter we assume the microstructure to be characterized by a single correlation radius l_c, such as the mean separation between the fibers in a fiber-matrix composite, or mean grain size d.

Both inequalities in (2.4) jointly ensure separation of scales in the deterministic continuum mechanics model. The first inequality may be relaxed to $d<L$ because we may be considering a microstructure with periodic (or nearly periodic) geometry, though possessing some randomness on the level of the unit cell: see relations (2.8)–(2.9) below.

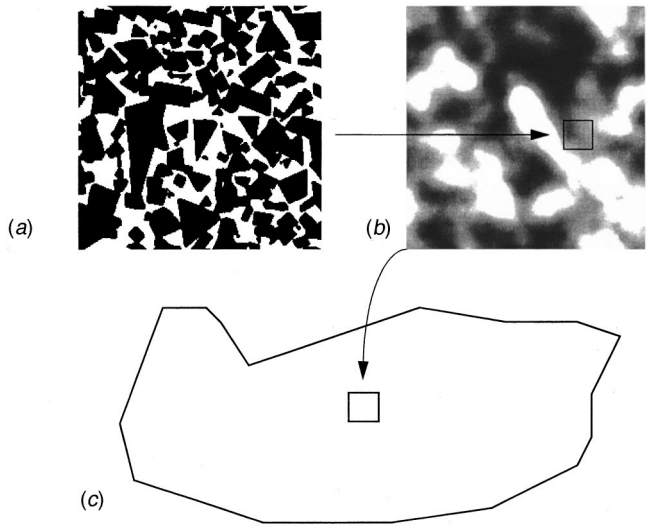


Fig. 1 Passage from a discrete system of tungsten-carbide (black) and cobalt (white) (*a*) to an intermediate continuum level (*b*) involving a mesoscale finite element, that serves as input into the macroscale model accounting for spatial nonuniformity. Figures (*a*) and (*b*) are generated by a Boolean model of Poisson polygons and a diffusion random function, respectively ([45]).

Evidently, the material property (or properties) Θ of $B(\omega)$ entering the continuum mechanics model are to be described by a random field over the D-dimensional space ($D=1,\ldots,3$)

$$\Theta:\mathbf{R}^D\times\Omega\rightarrow\mathbf{R}^1. \tag{2.5}$$

There exist two types of statistical homogeneity: strict-sense stationarity (SSS) and wide-sense stationarity (WSS). In the first case we are assured of the invariance of any finite-dimensional probability distribution of Θ with respect to arbitrary shifts

$$F_{\mathbf{x}_1,\ldots,\mathbf{x}_m}(\theta_1,\ldots,\theta_m)=F_{\mathbf{x}_1+\mathbf{x}',\ldots,\mathbf{x}_m+\mathbf{x}'}(\theta_1,\ldots,\theta_m)$$

$$\forall\mathbf{x}'\in\mathbf{R}^D. \tag{2.6}$$

This is a very restrictive property. On the other hand, in the WSS case, we have invariance of the mean only with respect to such shifts, along with the dependence of two-point correlation functions on the interpoint separations only, that is ($\langle\,\rangle$ denotes ensemble average),

- mean $\langle\Theta(\mathbf{x})\rangle=\text{const}$;
- for any two points $\mathbf{x}_1,\mathbf{x}_2\in\mathbf{R}^D$, the correlation function $K_\Theta(\mathbf{x}_1,\mathbf{x}_2)\equiv\langle[\Theta(\mathbf{x}_1)-\langle\Theta(\mathbf{x}_1)\rangle][\Theta(\mathbf{x}_2)-\langle\Theta(\mathbf{x}_2)\rangle]\rangle$ satisfies

$$K_\Theta(\mathbf{x}_1,\mathbf{x}_2)=K_\Theta(\mathbf{x}_1-\mathbf{x}_2). \tag{2.7}$$

Clearly, a much wider class of microstructures is described by WSS random fields then SSS random fields, and, as we shall see in the next sections, the former are sufficient for the RVE.

Following Eq. (2.4), we mentioned the microstructure with periodic geometry, possessing some randomness on the level of the unit cell. An appropriate model is then offered by a *strict-sense* (SS) *cyclostationary* random field, which, for a planar system of square $L\times L$ unit cells, is stated as

$$F_{\mathbf{x}_1+m\mathbf{L},\ldots,\mathbf{x}_m+m\mathbf{L}}(\theta_1,\ldots,\theta_m)=F_{\mathbf{x}_1,\ldots,\mathbf{x}_m}(\theta_1,\ldots,\theta_m) \tag{2.8}$$

where **L** is a shift vector (in any combination of directions along the coordinate axes), and *m* is any integer. Andy Warhol's 1972 creation *S&P Green Stamps* is helpful in visualizing this.

When the microstructure has an imperfectly periodic geometry, in addition to possessing some randomness on the level of the unit cell, then one should use a *wide-sense* (WS) *cyclostationary* random field, which for a system of square unit cells is stated as

$$\begin{gathered}\langle\Theta(\mathbf{x}+m\mathbf{L})\rangle=\langle\Theta(\mathbf{x})\rangle \\ K(\mathbf{x}_1+m\mathbf{L},\mathbf{x}_2+m\mathbf{L})=K(\mathbf{x}_1,\mathbf{x}_2)\end{gathered} \tag{2.9}$$

where **L** is the same as in (2.8), and *m* is any integer.

Returning to the WSS fields, we note that, to ensure the equivalence of macroscopic responses between all the realizations of the ensemble **B**, an ergodic random field is required. That is, we want any realization to be sufficient to get the ensemble average from the spatial average (denoted by $^-$)

$$\langle\Theta(\mathbf{x})\rangle=\int_\Omega\Theta(\omega)dP(\omega)=\lim_{V\to\infty}\frac{1}{V}\int_V\Theta(\omega,\mathbf{x})dV=\overline{\Theta(\omega)}. \tag{2.10}$$

In practice (2.10) holds only with some accuracy, and the limiting process $V\to\infty$ cannot truly be carried out. The latter can be discontinued in the regions whose dimension $L=V^{1/D}$ is large compared with the correlation radius l_c, so that (2.4) is rewritten as

$$l_c\ll V^{1/D}\sim L\ll L_{\text{macro}}. \tag{2.11}$$

We assume K_Θ to satisfy (with probability one) ergodic properties with respect to the mean and the correlation function, that is

$$\begin{gathered}\lim_{V\to\infty}\frac{1}{V}\int_V\Theta(\mathbf{x},\omega)dV=m=\langle\Theta(\mathbf{x},\omega)\rangle \\ \lim_{V\to\infty}\frac{1}{V}\int_V(\Theta(\mathbf{x},\omega)\Theta(\mathbf{x}+\Delta\mathbf{x},\omega))dV=K_\Theta(\Delta\mathbf{x})+m^2.\end{gathered} \tag{2.12}$$

In practice, the left and right-hand sides of $(2.12)_1$ and $(2.12)_2$ would be replaced, respectively, by a spatial average from a finite number of sampling points taken over one realization ω

$$\overline{\Theta(\omega)}=\frac{1}{N}\sum_{n=1}^{N}\Theta(\mathbf{x}_n,\theta) \tag{2.13}$$

and an ensemble average from a finite number of realizations ω taken at one sampling point

$$\langle\Theta(\mathbf{x})\rangle=\frac{1}{N}\sum_{n=1}^{N}\Theta(\mathbf{x},\omega_n). \tag{2.14}$$

The ergodicity of these estimators—i.e., $\overline{\Theta(\mathbf{x},\omega)}=\langle\Theta(\mathbf{x},\omega)\rangle$—is assured, for sufficiently large *N*, by the property of the correlation function

$$\lim_{|\Delta\mathbf{x}|\to\infty}K_\Theta(\Delta\mathbf{x})=0. \tag{2.15}$$

This, for instance, is the case with Voronoi mosaics based on a Poisson point field, both in two dimension and three dimensions ([23]); our Fig. 1(*a*) employs such a process. Many random microstructure models are set up on the basis of point fields, or their modifications. Real materials, however, oftentimes challenge us with spatially inhomogeneous patterns. The models can then easily be generalized by taking spatial inhomogeneity, but the concepts of homogeneity and ergodicity—especially from the standpoint of measurements—need to be relaxed.

2.2 Quasi-Homogeneous and Quasi-Ergodic Random Media. A typical example of inhomogeneous fluctuations in measured material properties is shown in Fig. 2. In particular, we see here elastic modulus, breaking strength, strain to failure, and tensile energy absorption of paper specimens sampled in an array in the plane of paper web ($D=2$) manufactured on a modern, high-speed paper machine. With reference to (2.5), E, $\sigma_{\max}$, $\varepsilon_{\max}$ and *TEA* form a four-component vector random field $\Theta:\mathbf{R}^2\times\Omega\to\mathbf{R}^4$. A statistical analysis of maps such as Fig. 2, in the ensemble sense, reveals spatial inhomogeneity of $\Theta_1\equiv E$ in the x_1-x_2–directions (so-called machine and cross directions of the paper web). This is not surprising given the fact that a paper web may be tens of thousands of kilometers long between maintenance intervals on a papermaking machine.

It turns out that these material inhomogeneities are too general to be described by locally homogeneous random fields, that is, fields whose variance of increments taken at different locations depends on the vector joining the locations, but not their absolute values. Of course, besides paper, numerous natural and man-made materials display such global (i.e., ''slow'' relative to ''fast'' microscale fluctuations) spatial inhomogeneities: cortical bone changing into cancelloous bone, ice fields ([24]), pig-iron cast into ingots, etc. If we let L_Θ be the characteristic space scale of the variance σ_Θ, the mean field $\langle\Theta\rangle$, and the correlation coefficient *K*, either of the following, or some intermediate situation, would apply

$$L_\Theta=L_{\text{macro}}\quad\text{or}\quad L_\Theta\ll L_{\text{macro}}. \tag{2.16}$$

In order to deal with such variations in the material one should employ (instead of WSS) a *quasi-WSS random field*, just as the so-called quasi-homogeneous fields—especially in the vertical direction—in atmospheric turbulence ([25]). Thus, we write

$$\begin{aligned}K(\mathbf{x}_1,\mathbf{x}_2)&=K_{ij}(\mathbf{x}_1,\mathbf{x}_2)=\sigma_i(\mathbf{x}_1)\sigma_j(\mathbf{x}_2)\rho_{ij}(\mathbf{x}_1,\mathbf{x}_2)\\&=\sigma_\Theta\left(R+\frac{r}{2}\right)\sigma_\Theta\left(R-\frac{r}{2}\right)\rho(r,R)\end{aligned} \tag{2.17}$$

where $\mathbf{r}=\mathbf{x}_1-\mathbf{x}_2$ and $R=(\mathbf{x}_1+\mathbf{x}_2)/2$. Keeping in mind the concept of L_Θ, for quasi-WSS fields we have

$$l_c\ll L_\Theta \tag{2.18}$$

and (2.17) is approximated by

$$K(r,R)=\sigma_\Theta^2(R)\rho(r,R). \tag{2.19}$$

If we want to estimate the properties of RVE of volume *V* from a single realization of the quasi-WSS random field, we effectively require it to be *quasi-ergodic*. The latter concept means the random field should be ergodic in volumes small as compared to the characteristic length scales L_Θ of variation of the field statistics, but (2.18) and (2.16) should still hold:

$$d\approx l_c\ll V^{1/D}\sim L\ll L_\Theta\approx L_{\text{macro}}\quad\text{or}\quad L_\Theta\ll L_{\text{macro}}. \tag{2.20}$$

We conclude that the RVE's microstructure is statistically representative if *V* is sufficiently small for fields under consideration to be statistically homogeneous and ergodic within its confines, and, at the same time, ''the volume is so large that the field Θ within *V* undergoes sufficient spatial fluctuations.'' The situation is additionally complicated by a possibility $L_\Theta\ll L_{\text{macro}}$.

Thus, the following key problem arises: the separation of scales *d* and *L* may be too large to allow the satisfaction of both strong inequalities $\ll$ in $(2.20)_1$. With reference to Fig. 1, passage from the random microstructure in Fig. 1(*a*) to a homogeneous continuum may require length scales *L* that are too large for entry into the macroscale problem of Fig. 1(*c*) occurring on scales L_{macro}. As a compromise, some intermediate random continuum approximation of Fig. 1(*b*) may have to be introduced, but a quantitative assessment of the approach to a homogeneous continuum can only be made with the help of a mechanics problem.

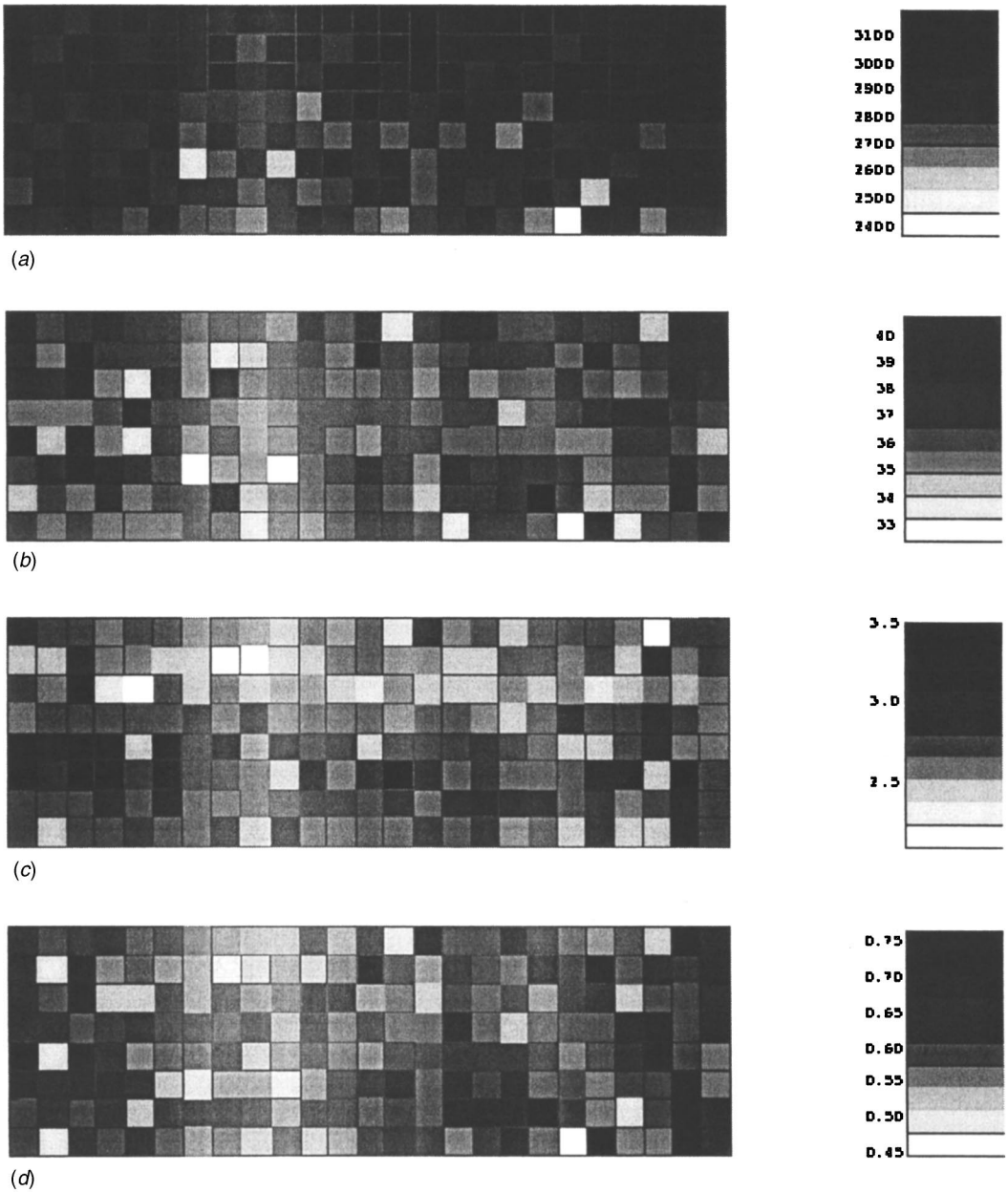


Fig. 2 Sampling of paper properties via a gray-scale plot of (*a*) elastic modulus *E* lbf/in; (*b*) breaking strength σ_{max} in lbf/in; (*c*) strain to failure ε_{max} in percentage and (*d*) tensile energy absorption *TEA* lbf/in. All data are for a 25×8 array of 1″×1″ specimens tested in the *x*-(machine) direction. The ranges and assignments of values are shown in the respective insets.

3 Hill Condition in Thermomechanics, and Mesoscale Response

The RVE response in TIV is described by the free energy Ψ and dissipation function Φ, both of which are scalar products

$$\Psi \equiv \frac{1}{2}\boldsymbol{\sigma}\cdot\boldsymbol{\varepsilon}^e \quad \Phi = \Phi_{th} + \Phi_{\text{intr}} \geqslant 0$$
$$\Phi_{th} \equiv -\mathbf{q}\cdot\nabla T/T \quad \Phi_{\text{intr}} \equiv \mathbf{Y}\cdot\overline{\dot{\mathbf{a}}} = \boldsymbol{\sigma}^v\cdot\mathbf{d}^e + \boldsymbol{\sigma}\cdot\mathbf{d}^p + A\cdot\dot{\alpha} \tag{3.1}$$

where the Clausius-Duhem inequality expresses the second law of thermodynamics with Φ_{th} being the thermal dissipation and Φ_{intr} the intrinsic dissipation. The latter quantity is a scalar product of the dissipative force $\mathbf{Y}$ with the velocity $\overline{\dot{\mathbf{a}}}$ (rate of the state variable $\mathbf{a}$). As an example, Φ_{intr} is taken to involve viscous, plastic and internal effects. Thus, $\boldsymbol{\sigma}$ is Cauchy stress, $\boldsymbol{\sigma}^v$ is viscous stress, $\mathbf{A}$ force associated to internal dissipative process, $\boldsymbol{\varepsilon}^e$ is elastic strain, $\mathbf{d}^e$ is elastic deformation rate, $\mathbf{d}^p$ is plastic deformation rate, $\dot{\alpha}$ is rate of internal parameters, $\mathbf{q}$ is heat flux, and T is temperature.

The problem we are facing is one of dependence of constitutive response on scale δ. That is, we want to be able to say something about the functionals Ψ_δ and Φ_δ for the ensemble $\mathbf{B}_\delta = \{B_\delta(\omega); \omega \in \Omega\}$ where the scale δ is finite rather than infinite—below the RVE limit; this is a particular case of (2.2). Such issues were addressed extensively for linear elastic materials (e.g., [26–31]), for nonlinear elastic materials ([32,33]) as well as for viscoelastic and damage phenomena ([34]); see also further references in these works.

We recall, with reference to these papers, that properties of an elastic body can be defined from a mechanical standpoint—i.e.,

via Hooke's law—or using energy—i.e., a Clapeyron type of formula. Both approaches are equivalent for a homogeneous material but not necessarily so for a heterogeneous one. Therefore, by analogy, in the case of a linear dissipative behavior, the mechanical approach may involve a statement like

$$\overline{\mathbf{Y}} = \mathbf{C}_m \overline{\dot{\mathbf{a}}} \tag{3.2}$$

which leads to an apparent, mechanically defined property $\mathbf{C}_m$, $\overline{\mathbf{Y}}_{ij}$ being the resultant volume average dissipative force. Alternately, it may involve

$$\overline{\dot{\mathbf{a}}} = \mathbf{S}_m \overline{\mathbf{Y}} \tag{3.3}$$

which leads to an effective property $\mathbf{S}_m$, $\overline{\dot{\mathbf{a}}}$ being the resultant volume average velocity.

On the other hand, the energy approach is stated as a volume average of the dissipation

$$\overline{\Phi} = \frac{1}{2}\,\overline{\mathbf{Y}\dot{\mathbf{a}}} \tag{3.4}$$

or a volume average of the velocity

$$\overline{\Phi} = \frac{1}{2}\,\overline{\dot{\mathbf{a}}} \cdot \mathbf{C}_e \cdot \overline{\dot{\mathbf{a}}}. \tag{3.5}$$

Clearly, when force and velocity fields are written as $\mathbf{Y} = \overline{\mathbf{Y}} + \mathbf{Y}'$ and $\dot{\mathbf{a}} = \overline{\dot{\mathbf{a}}} + \dot{\mathbf{a}}'$, where $\mathbf{Y}'$ and $\dot{\mathbf{a}}'$ are zero-mean fluctuations, and next, when (3.2) is recalled, (3.4) becomes

$$\overline{\Phi} = \frac{1}{2}\,\overline{\mathbf{Y}}\,\overline{\dot{\mathbf{a}}} + \frac{1}{2}\,\overline{\mathbf{Y}'\dot{\mathbf{a}}'} = \frac{1}{2}\,\overline{\dot{\mathbf{a}}}\mathbf{C}_m\overline{\dot{\mathbf{a}}} + \frac{1}{2}\,\overline{\mathbf{Y}'\dot{\mathbf{a}}'}. \tag{3.6}$$

A comparison of (3.5) with (3.6) shows that $\mathbf{C}_m$ is identical with $\mathbf{C}_e$ providing

$$\overline{\mathbf{Y}'\dot{\mathbf{a}}'} = 0 \tag{3.7}$$

or, equivalently,

$$\overline{\mathbf{Y}} \cdot \overline{\dot{\mathbf{a}}} = \overline{\mathbf{Y} \cdot \dot{\mathbf{a}}} = 0 \tag{3.8}$$

which may be called the *Hill condition for dissipative processes*.

For an unbounded space domain ($\delta \to \infty$), (3.8) is trivially satisfied, but for a finite body $B_\delta(\omega)$ it requires that the body be loaded in a specific way on its boundary $\partial \mathbf{B}_\delta$. Following Hazanov and Amieur [35], from (3.8), and employing the Green-Gauss theorem, we find a necessary and sufficient condition for (3.8)

$$\overline{\mathbf{Y} \cdot \dot{\mathbf{a}}} - \overline{\mathbf{Y}} \cdot \overline{\dot{\mathbf{a}}} = 0 \Leftrightarrow \int_{\partial B_\delta} (t(\mathbf{x}) - \mathbf{Y}^0 \cdot \mathbf{n}) \cdot (v(\mathbf{x}) - \dot{\mathbf{a}}^0 \cdot \mathbf{x})\, dS = 0 \tag{3.9}$$

where $\mathbf{Y}$ plays the role of stress (e.g., $\boldsymbol{\sigma}^v$ of (3.1)) and $\overline{\dot{\mathbf{a}}}$ the role of conjugate strain rate. In case of a process described by internal variables, (3.9) is a requirement of its spatial homogeneity.

Now, relation $(3.9)_2$ distinguishes three types of boundary conditions on the mesoscale: *uniform kinematic* (also called essential, or Dirichlet) boundary condition

$$\mathbf{v}(\mathbf{x}) = \dot{\mathbf{a}}^0 \cdot \mathbf{x} \quad \forall \mathbf{x} \in \partial \mathbf{B}_\delta \tag{3.10}$$

uniform traction (natural, or Neumann) boundary condition

$$\mathbf{t}(\mathbf{x}) = \mathbf{Y}^0 \cdot \mathbf{n} \quad \forall \mathbf{x} \in \partial \mathbf{B}_\delta \tag{3.11}$$

uniform kinematic-traction (also called orthogonal-mixed) boundary condition

$$(\mathbf{v}(\mathbf{x}) - \dot{\mathbf{a}}^0 \cdot \mathbf{x}) \cdot (\mathbf{t}(\mathbf{x}) - \mathbf{Y}^0 \cdot \mathbf{n}) = 0 \quad \forall \mathbf{x} \in \partial \mathbf{B}_\delta. \tag{3.12}$$

Each of these boundary conditions results in a different *apparent response*. Henceforth, we focus on the first two conditions because they provide bounds on the response under the third one. For any realization $B_\delta(\omega)$, a window's response on the mesoscale (δ finite) is, under these definitions, nonunique—because the response under (3.10) is not an inverse of the response under (3.11) almost surely (i.e., with probability one). Just as in elasticity studies, we use the term "apparent" to distinguish the mesoscale properties from the effective (macroscopic, global, or overall) ones. In the latter case, the fluctuations disappear in the limit $\delta \to \infty$ because of the ergodicity assumption of previous section.

We assume the composites to be made of just one size, d, of inclusions; we exclude slips and cracks. In the following it will be convenient to work with a nondimensional parameter

$$\delta = \frac{L}{d} \tag{3.13}$$

that characterizes any property associated with the windows such as those shown in Fig. 2. We shall refer to the case $\delta < \infty$ as a *mesoscale*, as opposed to $\delta \to \infty$ which is called a *macroscale*. $\delta \approx 1$ signifies *microscale*, or a micro-element ([21]).

4 Thermal Conductivity in Random Media

Let us first consider thermal conductivity, in a stationary state, in a two-dimensional random medium in the x_1, x_2-plane, governed locally by the Fourier's law

$$q_i = -K_{ij}(x, \omega) T_{,j}. \tag{4.1}$$

This paper's leitmotiv "microstructural randomness versus the RVE postulate" leads us, with reference to research papers mentioned in Section 3, to state two principal results in this area:

- order relation for any body $B(\omega) \in \mathbf{B}$

$$(\mathbf{R}^n_\delta(\omega))^{-1} \leqslant \mathbf{K}^e_\delta(\omega) \quad \forall \delta \tag{4.2}$$

- hierarchy of bounds for the ensemble $\mathbf{B}$

$$\langle \mathbf{R}^n_{\delta'} \rangle^{-1} \leqslant \langle \mathbf{R}^n_\delta \rangle^{-1} \leqslant \mathbf{K}^{\text{eff}} \leqslant \langle \mathbf{K}^e_\delta \rangle \leqslant \langle \mathbf{K}^e_{\delta'} \rangle \quad \forall \delta' < \delta. \tag{4.3}$$

The inequalities between any two second-rank tensors $\mathbf{A}$ and $\mathbf{B}$ are understood as $\mathbf{t}\cdot\mathbf{B}\cdot\mathbf{t} \leqslant \mathbf{t}\cdot\mathbf{A}\cdot\mathbf{t}$, $\forall t \neq 0$. $\mathbf{R}^n_\delta(\omega)$ and $\mathbf{K}^e_\delta(\omega)$ in the above are apparent resistivity and conductivity tensors obtained, respectively, under uniform natural ($q(\mathbf{x}) = q^0_i n_i$) and uniform essential ($T(\mathbf{x}) = T^0_{,i} x_i$) boundary conditions applied to the boundary ∂B_δ of $B_\delta(\omega)$.

Clearly, the hierarchy of bounds (4.3) may be expressed in terms of the apparent dissipation function $\Phi_\delta(\overline{\nabla T})$ and its dual $\Phi^*_\delta(\overline{\mathbf{q}})$

$$\begin{aligned} \langle \Phi_{\delta'}{}^*(q^0) \rangle \leqslant \langle \Phi_\delta{}^*(q^0) \rangle \leqslant \Phi^{\text{eff}}(\nabla T^0) \leqslant \langle \Phi_\delta(\nabla T^0) \rangle \\ \leqslant \Phi_{\delta'}(\nabla T^0) \quad \forall \delta' < \delta \end{aligned} \tag{4.4}$$

where, by virtue of ergodicity and stationarity assumptions, we have

$$\Phi_{\delta=\infty}{}^*(q^0) = \Phi^{*\text{eff}}(q^0) = \Phi^{\text{eff}}(\nabla T^0) = \Phi_{\delta=\infty}(\nabla T^0). \tag{4.5}$$

In other words, in (4.5) we have several equivalent statements: *effective*, *macroscopic*, *infinite size*, etc. This RVE situation ($\delta \to \infty$) is approached in practice only with some accuracy at a finite δ. The actual choice of accuracy—be it, say, two percent—is up to the researcher working on a given problem.

The joint dependence of material response on scale *and* on choice of independent variable (i.e., ∇T or $\mathbf{q}$) leads to a graphic representation of dissipation surfaces in the space of volume-averaged velocity $\overline{\dot{\mathbf{a}}}$ (i.e., thermal gradient $\overline{\nabla T}$) or force $\overline{\mathbf{Y}}$ (respectively, heat flux $\overline{\mathbf{q}}$) in Fig. 3. Note that $\overline{\nabla T} = \nabla T^0$ and $\overline{\mathbf{q}} = \mathbf{q}^0$ for a body with spatially continuous temperature and heat fields. Depending on how we take ensemble averages ([36]); we arrive at these Legendre transforms for finite-sized bodies:

(i) case of $\overline{\dot{\mathbf{a}}} = \dot{\mathbf{a}}^0$ being an independent variable

$$\Phi^*_\delta(\langle \overline{\mathbf{Y}} \rangle) + \langle \Phi_\delta(\overline{\dot{\mathbf{a}}}) \rangle = \langle \overline{\mathbf{Y}} \rangle \cdot \overline{\dot{\mathbf{a}}}. \tag{4.6}$$

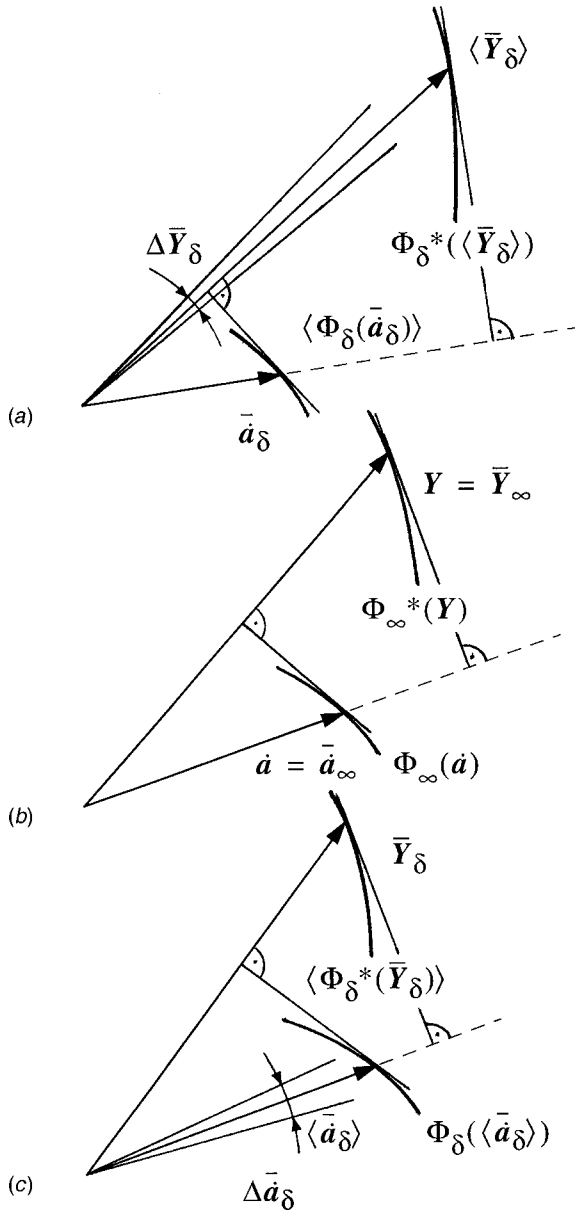


Fig. 3 Thermodynamic orthogonality in (*a*) the spaces of velocities $\bar{\dot{a}}_\delta$ and ensemble average forces $\langle \bar{Y}_\delta \rangle$ on mesoscale δ, with $\Delta \bar{Y}_\delta$ showing scatter in $\bar{Y}_\delta$; (*b*) the spaces of velocities $\dot{a} \equiv \bar{\dot{a}}_\infty$ and ensemble average forces $Y \equiv \bar{Y}_\infty$ on macroscale, where the scatter in $\dot{a}_\delta$ and $\bar{Y}_\delta$ is absent; (*c*) ensemble-average velocities $\langle \bar{\dot{a}}_\delta \rangle$ and forces $\bar{Y}_\delta$ on mesoscale, with $\Delta \bar{\dot{a}}_\delta$ showing scatter in $\dot{a}_\delta$. In all the cases, dissipation functions Φ and respective duals Φ^*, on mesoscale (parametrized by δ) or macroscale (parametrized by ∞) are shown.

(ii) case of $\overline{\mathbf{Y}} = \mathbf{Y}^0$ being an independent variable

$$\langle \Phi_\delta^*(\overline{\mathbf{Y}}) \rangle + \Phi_\delta(\overline{\dot{\mathbf{a}}}) = \overline{\mathbf{Y}} \cdot \langle \overline{\dot{\mathbf{a}}} \rangle. \tag{4.7}$$

In the $\delta \to \infty$ limit (4.6)–(4.7) become

$$\Phi^{*\text{eff}}(\overline{\mathbf{Y}}) + \Phi^{\text{eff}}(\dot{\mathbf{a}}) = \overline{\mathbf{Y}} \cdot \overline{\dot{\mathbf{a}}}. \tag{4.8}$$

Unfortunately, these Legendre transforms are not reversible because, with reference to heat conductivity, uniform boundary condition with $\overline{\nabla T} = \nabla T^0$ yields a different apparent response from that under nonuniform boundary condition of the same mean value $\overline{\nabla T}$. This, and the analogous observation on the natural

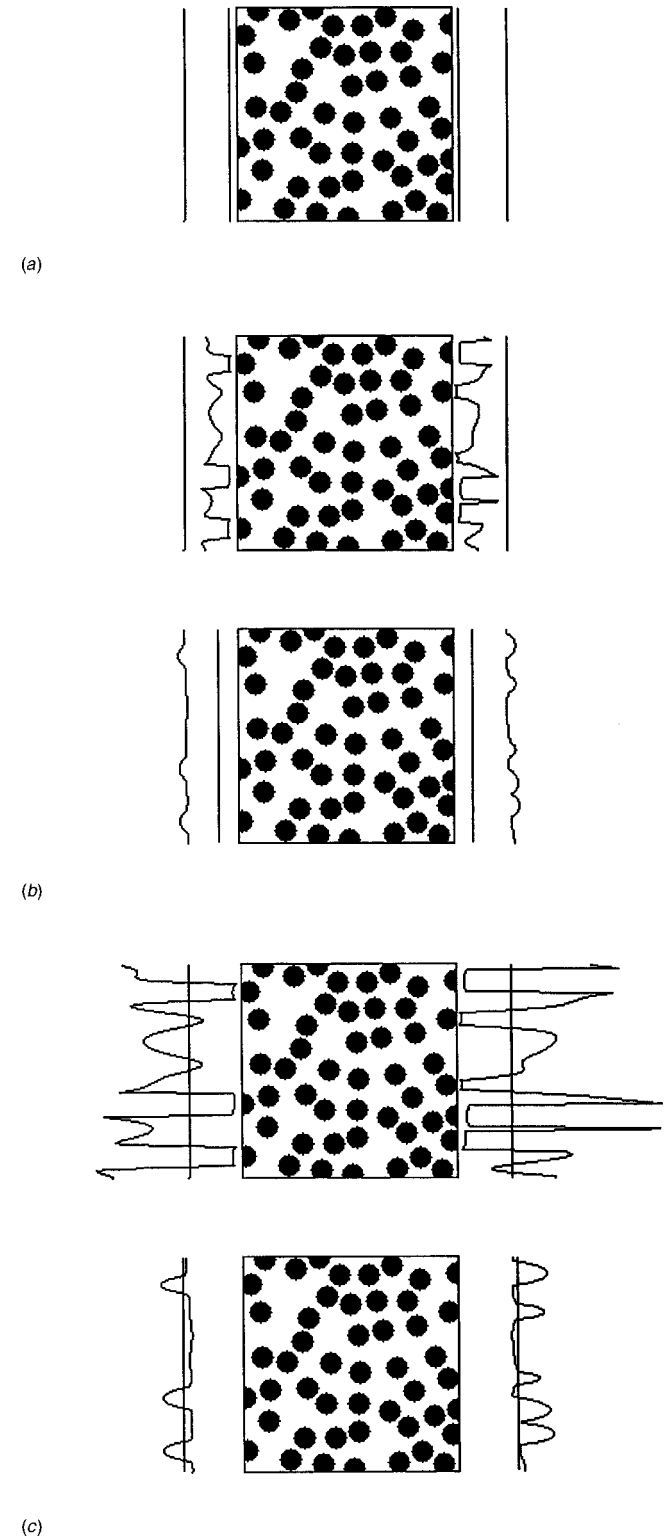


Fig. 4 Antiplane responses of a matrix-inclusion composite, with 35 percent volume fraction of inclusions, at decreasing contrasts: (*a*) $C^{(i)}/C^{(m)} = 1$, (*b*) $C^{(i)}/C^{(m)} = 0.2$, (*c*) $C^{(i)}/C^{(m)} = 0.05$, (*d*) $C^{(i)}/C^{(m)} = 0.02$. For (*b–d*), the first figure shows response under displacement b.c.'s ε_1^0, while the second one shows response under traction b.c.'s $\sigma_1^0 = \bar{\sigma}_1$ computed from the first problem.

Table 1

Thermal Conductivity	Antiplane Shear Elasticity
Temperature, T	Displacement, u
Temperature gradient, $g_i \equiv T_{,i}$	Strain, $\varepsilon_i = u_{i,3}$
Heat flux through a boundary, $q = q_i n_i$	Traction at a boundary, $t = \sigma_{i3} n_i$
Heat flux, q_i	Cauchy stress, σ_{i3}
Conductivity, $-K_{ij}$	Stiffness, C_{i3j3}
Resistivity, R_{ij}	Compliance, S_{i3j3}
Thermal dissipation,	Strain energy,
$\Phi/2T = q_i T_{,i}/2 = T_{,i} K_{ij} T_{,j}/2$	$\Psi = \sigma_i \varepsilon_i/2 = \varepsilon_i C_{i3j3} \varepsilon_j/2$
(Dual) thermal dissipation,	Complementary strain energy,
$\Phi^*/2T = q_i R_{ij} q_i/2$	$\Psi^* = \sigma_i S_{i3j3} \sigma_j/2$

boundary condition, is illustrated in terms of the boundary distributions for two basic types of boundary value problems on a matrix-inclusion composite in Fig. 4. In light of Section 3, we might also set up a reversible Legendre transformation in the case of uniform orthogonal-mixed boundary conditions on mesoscale, although there still remains a nonunique choice of the actual setup of the $\mathbf{Y}^0, \dot{\mathbf{a}}^0$-loading.

For the relations (4.6)–(4.7) one needs to assume that, for each specimen $B_\delta(\omega)$,
(i) the function $\Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega)$ depends on $\overline{\dot{\mathbf{a}}}$ alone, and is star-shaped, convex, and homogeneous of degree r

$$\overline{\dot{a}}_i \frac{\partial}{\partial \overline{\dot{a}}_i} \Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega) = r \Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega). \tag{4.9}$$

(ii) the function $\Phi^*_\delta(\overline{\mathbf{Y}}, \omega)$ is star-shaped, convex, and homogeneous of degree r

$$\overline{\mathbf{Y}}_i \frac{\partial}{\partial \overline{Y}_i} \Phi^{IL}_\delta(\overline{\mathbf{Y}}, \omega) = r \Phi^*_\delta(\overline{\mathbf{Y}}, \omega). \tag{4.10}$$

Note that $\Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega)$ and $\Phi^*_\delta(\overline{\mathbf{Y}}, \omega)$ are almost surely not inverses of one another because perfectly homogeneous domains of material carry probability zero in the Ω space.

It is of interest to note here that the conventional Onsager-Casimir reciprocity relations—that apply to Fig. 3(*b*)—need to be reconsidered depending on whether we work in the space of thermal gradients or the space of heat fluxes for finite-sized bodies in Figs. 3(*a*) and (*c*). Thus, in the first case we actually have two choices: when we are either on the surface $\langle \Phi_\delta(\overline{\dot{\mathbf{a}}}) \rangle$ of Fig. 3(*a*)

$$\frac{\partial \langle \overline{Y}_i \rangle}{\partial \overline{\dot{a}}_j} = \frac{\partial \langle \overline{Y}_j \rangle}{\partial \overline{\dot{a}}_i} \tag{4.11}$$

or on the surface $\Phi_\delta(\langle \overline{\dot{\mathbf{a}}} \rangle)$ of Fig. 3(*c*)

$$\frac{\partial \overline{Y}_i}{\partial \langle \overline{\dot{a}}_j \rangle} = \frac{\partial \overline{Y}_j}{\partial \langle \overline{\dot{a}}_i \rangle}. \tag{4.12}$$

When working in the space of heat fluxes we also have two choices: when we are on the surface $\Phi^*_\delta(\langle \overline{\mathbf{Y}} \rangle)$ of Fig. 3(*a*), we have

$$\frac{\partial \overline{\dot{a}}_i}{\partial \langle \overline{Y}_j \rangle} = \frac{\partial \overline{\dot{a}}_j}{\partial \langle \overline{Y}_i \rangle} \tag{4.13}$$

while on the surface $\langle \Phi^*_\delta(\overline{\mathbf{Y}}) \rangle$ of Fig. 3(*c*), we have

$$\frac{\partial \langle \overline{\dot{a}}_i \rangle}{\partial \overline{Y}_j} = \frac{\partial \langle \overline{\dot{a}}_j \rangle}{\partial \overline{Y}_i}. \tag{4.14}$$

In (4.12)–(4.13) averaging is to be conducted prior to differentiation. Noting the well-known analogy between the antiplane shear elasticity and the in-plane conductivity (Table 1), we see that all the results above apply to the antiplane elasticity of a two-dimensional random medium of the same microstructure, and governed locally by $\sigma_{i3} = C_{i3j3} \varepsilon_{j3}$.

This analogy confirms that $\overline{\nabla T}$ should be taken as a velocity-like variable and $\overline{\mathbf{q}}$ as a force-like variable in the thermomechanics of random media, which choice would reverse the roles of these variables conventionally assigned in TIV., but does agree with RT. Also, we note that whatever was said above for the irreversible thermodynamic process of heat conduction does also hold for the antiplane elasticity, and hence Fig. 3 may be interpreted in terms of the strain energies in the spaces of strains and stresses.

A very wide class of elastic/dissipative materials of nonlinear type may be obtained by postulating the local behavior to obey the thermodynamic orthogonality ([17,37]) as expressed by Fig. 3(*b*). The thermodynamic orthogonality, as well as the entire procedure of derivation of constitutive laws from the free energy and dissipation functions, are of primary interest with respect to materials with dissipative processes described by the intrinsic dissipation Φ_{intr} rather than the thermal dissipation Φ_{th} above. The next section, therefore, discusses thermodynamic orthogonality on mesoscale.

5 Thermodynamic Orthogonality on Mesoscale

5.1 Quasi-Homogeneous Dissipation Functions. A wide class of dissipative processes is described by dissipation functions $\Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega)$ of quasi-homogeneous type ([27]). Following the general framework given in ([36]), we now consider the apparent behavior to be described by dissipation functions of that type on mesoscale, so that $\Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega)$ pertains to a finite-sized body $B_\delta(\omega)$

$$\overline{\dot{a}}_i \frac{\partial}{\partial \overline{\dot{a}}_i} \Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega) = f(\Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega)) \tag{5.1}$$

where function f is arbitrary. This, of course, implies that the mesoscale dissipation functions in the space of dissipative forces, $\Phi^*_\delta(\overline{\mathbf{Y}}, \omega)$, are quasi-homogeneous too, that is

$$\overline{Y}_i \frac{\partial}{\partial \overline{Y}_i} \Phi^*_\delta(\overline{\mathbf{Y}}, \omega) = g(\Phi^*_\delta(\overline{\mathbf{Y}}, \omega)). \tag{5.2}$$

Given the nonuniqueness of the mesoscale response, these two functions are not perfectly dual of each other—just as was demonstrated by Fig. 4. Clearly, we have two alternatives:

(i) assume velocity $\overline{\dot{\mathbf{a}}}$ to be prescribed (controllable) for the body $B_\delta(\omega)$, the result being $\overline{\mathbf{Y}}$;
(ii) assume $\overline{\mathbf{Y}}$ to be prescribed (controllable) for the body $B_\delta(\omega)$, the result being $\overline{\dot{\mathbf{a}}}$.

In the first case, on account of (5.1), for any $B_\delta(\omega)$ we have

$$\overline{Y}_i(\omega) = \frac{\Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega)}{f(\Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega))} \frac{\partial}{\partial \overline{\nu}_i} \Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega). \tag{5.3}$$

If for every $B_\delta(\omega)$ we define a function $\phi(\overline{\dot{\mathbf{a}}}, \omega)$ from $\Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega)$ by

$$\phi_\delta(\overline{\dot{\mathbf{a}}}, \omega) = \int \frac{\Phi_\delta}{f(\Phi_\delta)} d\Phi_\delta \tag{5.4}$$

and let the additional constant in (5.4) be fixed by setting $\phi_\delta(\Phi_\delta(\overline{\dot{\mathbf{a}}}, \omega) = 0) = 0$, upon ensemble averaging, we obtain

$$\langle \overline{Y}_i \rangle = \left\langle \frac{\partial}{\partial \overline{\dot{a}}_i} \phi_\delta(\overline{\dot{\mathbf{a}}}) \right\rangle = \frac{\partial}{\partial \overline{\dot{a}}_i} \langle \phi_\delta(\overline{\dot{\mathbf{a}}}) \rangle. \tag{5.5}$$

Turning now to the space of dissipative forces, we may proceed in an analogous fashion. That is, we may either consider a random

dissipation function $\Phi_\delta^*(\overline{\mathbf{Y}},\omega)$ in the space of controllable forces resulting in $\overline{\dot{\mathbf{a}}}(\omega)$, or a deterministic $\Phi_\delta^*(\langle\overline{\mathbf{Y}}\rangle)$ in the space of average $\langle\overline{\mathbf{Y}}\rangle$ such that

$$\overline{\dot{a}}_i = \nu \frac{\partial}{\partial\langle\overline{Y}_i\rangle}\Phi_\delta^*(\langle\overline{\mathbf{Y}}\rangle). \tag{5.6}$$

Relevant to our analysis leading to (5.6) is the latter situation. On account of (5.2) the connection between $\overline{\dot{\mathbf{a}}}$ and $\langle\overline{\mathbf{Y}}\rangle$ reduces to

$$\overline{\dot{a}}_i = \frac{\Phi_\delta^*(\langle\overline{\mathbf{Y}}\rangle)}{g(\Phi_\delta^*(\langle\overline{\mathbf{Y}}\rangle))}\frac{\partial}{\partial\langle\overline{Y}_i\rangle}\Phi_\delta^*(\langle\overline{\mathbf{Y}}\rangle) = \mu\frac{\partial}{\partial\langle\overline{Y}_i\rangle}\Phi_\delta^*(\langle\overline{\mathbf{Y}}\rangle) \tag{5.7}$$

where

$$\mu = \Phi_\delta^*\left(\overline{Y}_i\frac{\partial}{\partial\overline{Y}_i}\Phi_\delta^*\right)^{-1}. \tag{5.8}$$

If we now define a function $\psi_\delta(\langle\overline{\mathbf{Y}}\rangle)$ from $\Phi_\delta^*(\langle\overline{\mathbf{Y}}\rangle)$ by

$$\psi_\delta(\langle\overline{\mathbf{Y}}\rangle) = \int\frac{\Phi_\delta^*}{g(\Phi_\delta^*)}d\Phi_\delta^* \tag{5.9}$$

and let $\psi_\delta(\Phi_\delta^*(\langle\overline{\mathbf{Y}}\rangle)=0)=0$, we can write, instead of (5.8),

$$\overline{\dot{a}}_i = \frac{\partial}{\partial\langle\overline{Y}_i\rangle}\psi_\delta(\langle\overline{Y}\rangle) \tag{5.10}$$

whereby

$$\phi_\delta(\overline{\dot{\mathbf{a}}}=0)=0 \quad \psi_\delta(\langle\overline{\mathbf{Y}}\rangle=0)=0. \tag{5.11}$$

We will now consider two curves: C in velocity space and its image C' in force space. Curve C connects the origin O with a point P with coordinates $\overline{\dot{\mathbf{a}}}$, while C' connects the origin O' with the image P' of P having coordinates $\langle\overline{\mathbf{Y}}\rangle$. Thus, we have

$$\int_C\langle\overline{Y}_i\rangle d\overline{\dot{a}}_i + \int_{C'}\langle\overline{\dot{a}}_i\rangle d\langle\overline{Y}_i\rangle = \int_C d(\langle\overline{Y}_i\rangle\overline{\dot{a}}_i) = \langle\overline{Y}_i\rangle\overline{\dot{a}}_i. \tag{5.12}$$

In light of (5.6), (5.11), and (5.12), this leads to a Legendre transformation corresponding to case (i),

$$\langle\phi_\delta(\overline{\dot{\mathbf{a}}})\rangle + \psi_\delta(\langle\overline{\mathbf{Y}}\rangle = \langle\overline{\mathbf{Y}}\rangle\cdot\overline{\dot{\mathbf{a}}} = \Phi_\delta^*(\langle\overline{\mathbf{Y}}\rangle). \tag{5.13}$$

An analogous analysis for case (ii) results in a very similar Legendre transformation (duality between the results in the velocity space and those in the force space)

$$\phi_\delta(\langle\overline{\dot{\mathbf{a}}}\rangle) + \langle\psi_\delta(\overline{\mathbf{Y}})\rangle = \overline{\mathbf{Y}}\cdot\langle\overline{\dot{\mathbf{a}}}\rangle = \Phi_\delta(\langle(\overline{\dot{\mathbf{a}}}\rangle) \tag{5.14}$$

where

$$\overline{Y}_i = \frac{\partial}{\partial\langle\overline{\dot{\mathbf{a}}}_i\rangle}\phi_\delta(\langle\overline{\dot{\mathbf{a}}}\rangle) \tag{5.15}$$

and

$$\langle\overline{\dot{\mathbf{a}}}_i\rangle = \frac{\partial}{\partial\overline{Y}_i}\langle\psi_\delta(\overline{\mathbf{Y}})\rangle. \tag{5.16}$$

The functions $\phi_\delta(\langle\overline{\dot{\mathbf{a}}}\rangle)$ and $\langle\psi_\delta(\overline{\mathbf{Y}})\rangle$ in the above are defined by

$$\phi_\delta(\langle\overline{\dot{\mathbf{a}}}\rangle) = \int\frac{\Phi_\delta}{f(\Phi_\delta)}d\Phi_\delta \quad \Phi_\delta \equiv \Phi_\delta(\langle\overline{\dot{\mathbf{a}}}\rangle) \tag{5.17}$$

and, for every $B_\delta(\omega)$,

$$\psi_\delta(\overline{\mathbf{Y}},\omega) = \int\frac{\Phi_\delta^*}{g(\Phi_\delta^*)}d\Phi_\delta^* \quad \Phi_\delta^* \equiv \Phi_\delta^*(\overline{\mathbf{Y}},\omega). \tag{5.18}$$

5.2 Extremum Principles. The foregoing generalization of the formulas relating the dissipative force with the velocity via functions Φ_δ and Φ_δ^* leads us now to a generalization of the extremum principles of deterministic thermomechanics ([17,18]) to the random medium $\mathbf{B}_\delta=\{B_\delta(\omega);\omega\in\Omega\}$. Let us discuss these principles for $\Phi_\delta(\overline{\dot{\mathbf{a}},\omega})$; the same results will then carry over automatically for $\Phi_\delta^*(\overline{\mathbf{Y}},\omega)$ by the argument of duality. Two different approaches—depending on whether velocities or forces are prescribed—were already considered, and their relation to the extremum principles for the case of apparent homogeneous dissipation functions of order r ($\dot{a}_i\partial\Phi_\delta(\dot{\mathbf{a}},\omega)/\partial\dot{a}_i = r\Phi_\delta(\dot{\mathbf{a}},\omega)$ and $\dot{Y}_i\partial\Phi_\delta^*(\dot{\mathbf{a}},\omega)/\partial\dot{Y}_i = r\Phi_\delta(\overline{\mathbf{Y}},\omega)$) is expounded in the following.

Approach 1. $\overline{\dot{\mathbf{a}}}$ is prescribed and $\overline{\mathbf{Y}}(\omega)$ follows from the ensemble of random dissipation surfaces $\Phi_\delta(\overline{\dot{\mathbf{a}},\omega})$ according to $\dot{Y}_i(\omega)=\lambda(\omega)\partial\Phi_\delta(\overline{\dot{\mathbf{a}},\omega})/\partial\dot{a}_i$; Fig. 3(*a*).

The *principle of maximal dissipation rate* for a random medium $\mathbf{B}_\delta$ reads: Provided the dissipative force $\langle\overline{\mathbf{Y}}\rangle$ is prescribed, the actual velocity $\overline{\dot{\mathbf{a}}}$ maximizes the dissipation rate $L_\delta^{(d)}=\langle\overline{\mathbf{Y}}\rangle\cdot\overline{\dot{\mathbf{a}}}$ subject to the side condition

$$\langle\Phi_\delta(\overline{\dot{\mathbf{a}}})\rangle = \langle\overline{\mathbf{Y}}\rangle\cdot\overline{\dot{\mathbf{a}}} = L_\delta^{(d)} > 0. \tag{5.19}$$

The *principle of least dissipative force* for a random medium $\mathbf{B}_\delta$ reads: Provided the value $\langle\Phi_\delta(\overline{\dot{\mathbf{a}}})\rangle$ of the dissipation function and the direction $\mathbf{n}$ of the dissipative force $\langle\overline{\mathbf{Y}}\rangle$ are prescribed, the actual velocity $\overline{\dot{\mathbf{a}}}$ minimizes the magnitude of $\langle\overline{\mathbf{Y}}\rangle$ subject to the side condition (5.19).

Approach 2. $\overline{\mathbf{Y}}$ is prescribed and $\overline{\dot{\mathbf{a}}}(\omega)$ follows from the ensemble of random dissipation surfaces $\Phi_\delta^*(\overline{\mathbf{Y}},\omega)$ according to $\overline{\dot{\mathbf{a}}}_i(\omega)=\mu(\omega)\partial\Phi_\delta^*(\overline{\mathbf{Y}},\omega)/\partial\overline{Y}_i$; Fig. 3(*c*).

The *principle of maximal dissipation rate* reads now: Provided the dissipative force $\overline{\mathbf{Y}}$ is prescribed, the actual velocity $\langle\overline{\dot{\mathbf{a}}}\rangle$ maximizes the mesoscale dissipation rate $L_\delta^{(d)}=\overline{\mathbf{Y}}\cdot\langle\overline{\dot{\mathbf{a}}}\rangle$ subject to the side condition

$$\Phi_\delta(\langle\overline{\dot{\mathbf{a}}}\rangle) = \overline{\mathbf{Y}}\cdot\langle\overline{\dot{\mathbf{a}}}\rangle = L_\delta^{(d)} > 0. \tag{5.20}$$

The *principle of least dissipative force* for a random medium $\mathbf{B}_\delta$ reads: Provided the value $\Phi_\delta(\langle\overline{\dot{\mathbf{a}}}\rangle)$ of the dissipation function and the direction $\mathbf{n}$ of the dissipative force $\overline{\mathbf{Y}}$ are prescribed, the actual velocity $\langle\overline{\dot{\mathbf{a}}}\rangle$ minimizes the magnitude of $\overline{\mathbf{Y}}$ subject to the side condition (5.20).

Clearly, the heterogeneity of material microstructure is the key cause of random constitutive behavior. The rate of dissipation per mesoscale volume of $\mathbf{B}_\delta$ varies from one specimen to another, unless the microstructure is perfectly deterministic (i.e., periodic) or contains an infinite number of elements (e.g., grains). In general, therefore, the spatial distribution is a cooperative stochastic process, a subject considered in the next section.

6 Material With Elasticity Coupled to Damage

6.1 Basic Considerations. Let us now consider a material whose elasticity law—as described in Section 7.5.1 in ([4])—is described by

$$\sigma_{ij} = (1-D)C_{ijkl}\varepsilon_{kl} \tag{6.1}$$

where C_{ijkl} is isotropic, and which must be coupled with a law of isotropic damage, that is

$$\dot{D} = \partial\Phi^*/\partial Y \tag{6.2}$$

with $Y=-\partial\Psi/\partial\varepsilon$, Ψ being the free energy. In particular, the scalar D evolves with elastic dilatation strain $\varepsilon=\varepsilon_{ii}$, which is taken as a time-like parameter, according to

$$\frac{dD}{d\varepsilon} = \begin{cases} (\varepsilon/\varepsilon_0)^{s^*} & \text{when} \quad \varepsilon=\varepsilon_D \quad \text{and} \quad d\varepsilon=d\varepsilon_D>0 \\ 0 & \text{when} \quad \varepsilon<\varepsilon_D \quad \text{and} \quad d\varepsilon<0. \end{cases} \tag{6.3}$$

Integration from the initial conditions $D=\varepsilon_D=0$ up to the total damage, $D=1$, gives

$$D=(\varepsilon/\varepsilon_R)^{s^*+1} \quad \varepsilon_R=[(s^*+1)\varepsilon_0^{s^*}]^{s^*+1}$$

$$\sigma=[1-(\varepsilon/\varepsilon_R)^{s^*+1}]E\varepsilon \tag{6.4}$$

where $\sigma=\sigma_{ii}$.

All of the above are to be understood as an effective law for the RVE, that is

$$C_{ijkl}^{\text{eff}}\equiv C_{ijkl_\infty} \quad D^{\text{eff}}\equiv D_\infty \quad \Psi^{\text{eff}}\equiv\Psi_\infty \quad \Phi^{\text{eff}}\equiv\Phi_\infty \ldots \tag{6.5}$$

as well as a guidance for adopting apparent responses on mesoscales. Thus, assuming that the same types of formulas hold for any finite δ, we have an apparent response for any specimen $B_\delta(\omega)$

$$\overline{\sigma}(\omega)=(1-D_\delta^d)C_\delta^d(\omega)\cdot\varepsilon^0 \tag{6.6}$$

under uniform displacement boundary condition: $\mathbf{u}(\mathbf{x})=\boldsymbol{\varepsilon}^0\cdot\mathbf{x}$. The notation D_δ^d expresses the fact that material damage is dependent on the mesoscale δ and the type of boundary conditions applied (i.e., d). In fact, while we could formally write another apparent response $\overline{\varepsilon}(\omega)=(1-D_\delta^t)^{-1}S_\delta^t(\omega)\cdot\boldsymbol{\sigma}^0$, we shall not do so because a damage process under traction boundary condition $(\mathbf{t}(\mathbf{x})=\boldsymbol{\sigma}^0\cdot\mathbf{n})$ would be unstable.

6.2 Scaling of Damage Parameter *D*. It is now possible to obtain scale-dependent bounds on D_δ^d through a procedure analogous to that in linear elasticity without damage ([26]). To this end, we partition a square-shaped window $B_\delta(\omega)$, of volume V_δ, into four smaller square-shaped windows $B_{\delta_s'}(\omega)$, $s=1,\ldots,4$, of scale $\delta'=\delta/2$ and volume $V_{\delta'}$ each. Next, we define two types of uniform displacement boundary conditions, in terms of a prescribed constant strain $\boldsymbol{\varepsilon}^0$, over the window $B_\delta(\omega)$:
unrestricted

$$\mathbf{u}(\mathbf{x})=\boldsymbol{\varepsilon}^0\cdot\mathbf{x} \quad \forall\mathbf{x}\in\partial B_\delta \tag{6.7}$$

and: restricted

$$\mathbf{u}^r(\mathbf{x})=\boldsymbol{\varepsilon}^0\cdot\mathbf{x} \quad \forall\mathbf{x}\in\partial B_{\delta_s'} \quad s=1,\ldots,4 \tag{6.8}$$

superscript r in (6.8) indicates a ''restriction.'' That is, (6.7) is given on the external boundary of the large window, whereas (6.8) is given on the boundaries of each of the four subwindows. Let us note, by the strain averaging theorem, that the volume average strain is the same in each subwindow and also equals that in the large window

$$\boldsymbol{\varepsilon}^0=\overline{\boldsymbol{\varepsilon}}=\overline{\boldsymbol{\varepsilon}}^s \quad s=1,\ldots,4. \tag{6.9}$$

Let $\tilde{\boldsymbol{\sigma}},\tilde{\boldsymbol{\varepsilon}}$ be any kinematically admissible fields: They satisfy everywhere the local stress-strain relations (6.1) and the displacement boundary condition (6.7), while $\tilde{u}$ is derivable from a continuous function such that $\tilde{\varepsilon}_{ij}=\tilde{u}_{(i,j)}$ but $\tilde{\boldsymbol{\sigma}}$ is not necessarily in equilibrium. Now, there is a minimum potential energy principle for the fields $\tilde{\sigma},\tilde{\varepsilon}$ in $B_\delta(\omega)$:

$$\int_{\partial B_\delta^t}\overline{\mathbf{u}}\cdot\tilde{t}\,dS-\frac{1}{2}\int_{B_\delta}\overline{\boldsymbol{\varepsilon}}\cdot\tilde{\boldsymbol{\sigma}}\,dV\leqslant\int_{\partial B_\delta^t}u\cdot t\,dS-\frac{1}{2}\int_{B_\delta}\boldsymbol{\varepsilon}\cdot\boldsymbol{\sigma}\,dV. \tag{6.10}$$

For the displacement boundary condition, $\partial B_\delta=\partial B_\delta^u$ and $\partial B_\delta^t=\varnothing$, so that

$$\frac{1}{2}\int_{B_\delta}\boldsymbol{\varepsilon}\cdot\boldsymbol{\sigma}\,dV=\Psi(\omega,\boldsymbol{\varepsilon})\leqslant\Psi(\omega,\tilde{\boldsymbol{\varepsilon}})=\frac{1}{2}\int_{B_\delta}\tilde{\boldsymbol{\varepsilon}}\cdot\tilde{\boldsymbol{\sigma}}\,dV \tag{6.11}$$

or

$$\overline{\boldsymbol{\varepsilon}\cdot\boldsymbol{\sigma}}\leqslant\overline{\tilde{\boldsymbol{\varepsilon}}\cdot\tilde{\boldsymbol{\sigma}}}. \tag{6.12}$$

However, the Hill condition (3.8), combined with the fact that $\boldsymbol{\varepsilon}^0=\overline{\boldsymbol{\varepsilon}}$, allows us to write

$$\overline{\tilde{\boldsymbol{\varepsilon}}}\cdot\overline{\boldsymbol{\sigma}}\leqslant\overline{\tilde{\boldsymbol{\varepsilon}}}\cdot\overline{\tilde{\boldsymbol{\sigma}}}. \tag{6.13}$$

Because the solution $\tilde{\boldsymbol{\sigma}}^r,\tilde{\boldsymbol{\varepsilon}}^r$ under the restricted condition (6.8) is an admissible distribution under unrestricted condition (6.7) (but not vice versa), from the above we have

$$\overline{\boldsymbol{\varepsilon}}\cdot\overline{\boldsymbol{\sigma}}\leqslant\overline{\boldsymbol{\varepsilon}^r}\cdot\overline{\boldsymbol{\sigma}^r}. \tag{6.14}$$

In view of (6.9), we find

$$\overline{\sigma}(\omega)=(1-D_\delta^d(\omega))\mathbf{C}_\delta^d(\omega)\cdot\boldsymbol{\varepsilon}^0$$

$$\equiv\frac{1}{4}\sum_{s=1}^{4}(1-D_{\delta_s'}^d(\omega))\mathbf{C}_{\delta_s'}^d(\omega)\cdot\boldsymbol{\varepsilon}^0 \tag{6.15}$$

so that upon substitution into (6.13) we obtain

$$(1-D_\delta^d(\omega))\mathbf{C}_\delta^d\leqslant(1-D_{\delta'}^d(\omega))\mathbf{C}_{\delta'}^d \quad \forall\,\delta'=\delta/2. \tag{6.16}$$

We may now recall that in the virgin (no damage) state

$$\mathbf{C}_\delta^d(\omega)\leqslant\mathbf{C}_{\delta'}^d(\omega)\equiv\frac{1}{4}\sum_{s=1}^{4}\mathbf{C}_{\delta_s'}^d(\omega). \tag{6.17}$$

Thus, we conclude that the damage parameter satisfies a scale-effect relation opposite to that seen for effective moduli

$$D_{\delta'}^d(\omega)\leqslant D_\delta^d(\omega) \quad \forall\,\delta'=\delta/2. \tag{6.18}$$

In view of the assumed WSS and ergodicity properties of the material, this results in ensemble averages

$$\langle D_{\delta'}^d\rangle\leqslant\langle D_\delta^d\rangle \quad \forall\,\delta'=\delta/2. \tag{6.19}$$

By applying this inequality to ever larger windows ad infinitum we get a hierarchy of bounds on $\langle D_\infty^d\rangle=D^{\text{eff}}\equiv D_\infty$ from above

$$\langle D_{\delta'}^d\rangle\leqslant\langle D_\delta^d\rangle\leqslant\ldots\leqslant\langle D_\infty^d\rangle \quad \forall\,\delta'=\delta/2. \tag{6.20}$$

The above inequalities are consistent with the much more phenomenological Weibull-type modeling of brittle solids: The larger the specimen the more likely it is to fail (e.g., [11,38]). We have thus provided a derivation of the scaling law in such materials via mechanics of random media.

6.3 Stochastic Evolution of *D*. Of interest is formulation of a stochastic model of evolution of D_δ in function of ε to replace $(6.3)_1$; in other words, we need a stochastic process $D_\delta(\omega,\varepsilon)$; $\omega\in\Omega,\varepsilon\in[0,\varepsilon_R]$; recall that ε is a time-like parameter. Assuming, for simplicity of discussion, just as in ([4]), that $s^*=2$, we may consider this setup

$$dD_\delta(\omega,\varepsilon)=D_\delta(\omega,\varepsilon)+3\varepsilon^2[1+r_\delta(\omega)]dt \quad D_\delta(\omega,0)=0 \tag{6.21}$$

where $r_\delta(\omega)$ is a zero-mean random variable taking values from $[-a_\delta,a_\delta]$, $1/\delta=a_\delta<1$. This stochastic process has the following properties:

(i) its sample realizations display scatter ω-by-ω for $\delta<\infty$, i.e., for finite body sizes;
(ii) it becomes deterministic as the body size goes to infinity in the RVE limit $(\delta\to\infty)$;
(iii) its sample realizations are weakly increasing functions of ε;
(iv) its sample realizations are continuous;
(v) the scale effect inequality (6.20) is satisfied, providing we take ε_R a function of δ with a property

$$\varepsilon_R(\delta)<\varepsilon_R(\delta') \quad \forall\,\delta'=\delta/2. \tag{6.22}$$

Let us observe, however, that, given the presence of microstructure, mesoscale damage should be considered as a sequence of micro/mesoscopic events, thus rendering the apparent damage process $D_\delta(\omega,\varepsilon);\omega\in\Omega,\varepsilon\in[0,\varepsilon_R]$ one with discontinuous paths having increments dD_δ occurring at discrete time instants, Fig. 5(c). To satisfy this requirement one should, in place of the above,

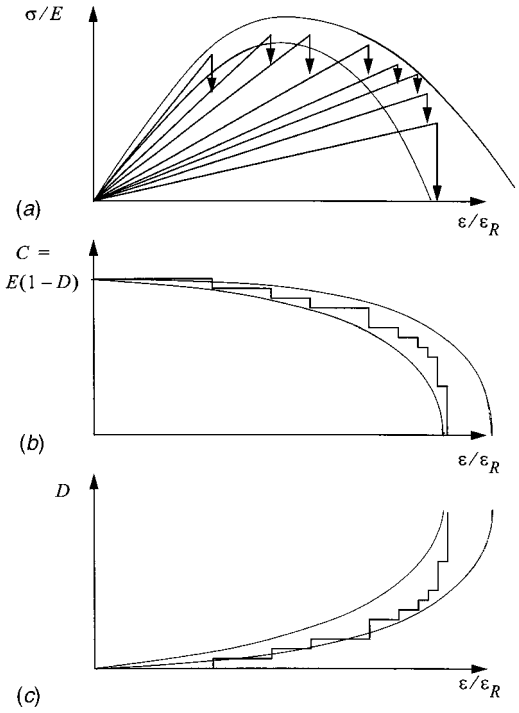


Fig. 5 Constitutive behavior of a material with elasticity coupled with damage, where $\varepsilon/\varepsilon_R$ plays the role of a controllable, time-like parameter of the stochastic process. (*a*) Stress-strain response of a single specimen $B_\delta(\omega)$ from B_δ, having a zigzag realization, (*b*) deterioration of stiffness, (*c*) evolution of the damage variable *D*. Curves shown in (*a*–*c*) indicate the scatter in stress, stiffness and damage at finite scale δ. Assuming spatial ergodicity, this scatter would vanish in the limit ($\delta \equiv L/d \to \infty$), whereby unique response curves of continuum damage mechanics would be recovered.

take a *Markov jump process* whose range is a subset [0, 1] of real line (i.e., where D_δ takes values). This process would be specified by an evolution propagator, or, more precisely, by a *next-jump probability density function* defined as follows:

$p(\varepsilon', D'_\delta | D_\delta, \varepsilon) d\varepsilon' dD'_\delta \equiv$ probability that, given the process is in state D_δ at time ε, its next jump will occur between times $\varepsilon+\varepsilon'$ and $\varepsilon+\varepsilon'+d\varepsilon'$, and will carry the process to some state between $D_\delta+D'_\delta$ and $D_\delta+D'_\delta+dD'_\delta$.

Figure 5(*b*) shows one realization $C_\delta(\omega,\varepsilon); \omega \in \Omega; \varepsilon \in [0,\varepsilon_R]$ of the apparent, mesoscale stiffness, corresponding to $D_\delta(\omega,\varepsilon); \omega \in \Omega; \varepsilon \in [0,\varepsilon_R]$ of Fig. 5(*c*). In Fig. 5(*a*) we see the resulting constitutive response $\sigma_\delta(\omega,\varepsilon); \omega \in \Omega; \varepsilon \in [0,\varepsilon_R]$.

Calibration of this model (just as the simpler one above)—that is, a specification of $p(\varepsilon', D'_\delta | D_\delta, \varepsilon) d\varepsilon' dD'_\delta$—may be conducted by either laboratory or computer experiments such as those in ([39,40]). As pointed out in the first of these references, in the macroscopic picture ($\delta \to \infty$) the zigzag character and randomness of an effective stress-strain response vanish. However, these studies—as well as many other works in mechanics/physics of fracture of random media (e.g., [41]), indicate that the homogenization with $\delta \to \infty$ is generally very slow, and hence that the assumption of WSS and ergodic random fields may be too strong for may applications.

Extension of the model from isotropic to (much more realistic) anisotropic damage will require vector, rather than scalar, Markov processes. This will lead to a somewhat greater mathematical complexity which may be balanced by choosing the first model of this subsection rather than the latter. These issues are secondary. The essence of this section is to outline a *stochastic continuum damage mechanics* that (i) is based on, and consistent with, micromechanics of random media as well as the classical thermomechanics formalism, and (ii) reduces to the classical continuum damage mechanics in the infinite volume limit.

7 Conclusions

Continuum thermomechanics of homogeneous media hinges on the concept of RVE, which is well defined in two situations only: (i) unit cell in a periodic microstructure, and (ii) volume containing a very large (mathematically infinite) number of microscale elements (e.g., molecules, grains, crystals), possessing ergodic properties. Modern materials, however, increasingly require one to work with small domains where neither of these two cases is met. As a result, response of finite volumes of material displays statistical scatter and is dependent on the scale and boundary conditions (typically kinematic or traction controlled). The need for stochastic homogenization of material response, which has been met by scale-dependent hierarchies of bounds for elastic materials, is extended here to dissipative/irreversible phenomena within the framework of thermomechanics with internal variables. In particular, the free-energy function and the dissipation function become stochastic functionals whose scatter tends to decrease to zero as the material volume is increased. These functionals are linked to their duals via Legendre transforms either in the spaces of ensemble-average kinematic variables or ensemble average force variables. It is in the limit of infinite volumes (RVE limit (ii) above) that all the functionals become deterministic, and that the classical Legendre transforms of deterministic thermomechanics hold. The procedure is illustrated by two constitutive behaviors—thermal expansion coefficients and elastic-brittle damage—for a wide class of materials with random microstructures.

In the case of heterogeneous media, thermomechanical response laws are not unique—they depend on the scale and the choice of boundary conditions applied to the given material domain. Such response laws are called apparent (or mesoscopic); they become effective (or macroscopic) in the limit of infinitely large volumes. For the elastic part of response we can prove that the apparent laws bound the effective law. For the inelastic part of response, when there is a field variational principle (plastic behavior) or some link (e.g., elastic-brittle damage) or analogy (e.g., thermal conductivity) to elastic behavior, we can also prove that the apparent dissipative responses bound the effective dissipative response. When such a link is not present, the latter is only a conjecture.

To be statistically representative, the RVE needs to possess some type of ergodicity and statistical homogeneity. A random field that lacks either of these properties cannot lead to a homogeneous continuum model of a random medium. Thus, a random field that is stationary and nonergodic does not offer a chance of reaching a statement in the sense of (2.8). However, a random field that is ergodic and nonstationary—such as encountered in the interphase zone of functionally graded materials—could be smeared out by an inhomogeneous continuum ([42]). In any case, the approximating mesoscale random field is nonunique (due to three types of boundary conditions admissible by the Hill condition) and almost surely anisotropic pointwise.

Scale dependence discussed in this paper indicates that many issues require further research. For example, assuming we have a material governed on the microscale by some homogeneous dissipation function of order $n \neq 2$, are the apparent dissipation functions on mesoscales, as well as the effective dissipation functions on macroscale ($\delta \to \infty$), of the same order? Or, considering that inclusion of higher gradients of displacement field provides a better description of the spatially inhomogeneous plasticity and damage processes ([43,44]) what is the formulation of such models for random media?

The proposed generalization of thermomechanics to grasp random nature of materials offers a basis on which to set up stochas-

tic finite element methods. Such an extension has been given so far in the case of elastic materials ([31]), where a window with apparent properties—of Dirichlet and Neumann type—played the role of a mesoscale finite element. It may now be conjectured that, in the case of inelastic materials, both dissipation functions—based on ensemble-average velocities, and ensemble average forces, respectively—will allow bounds on global response.

Finally, another example where the SVE rather than RVE—i.e., a statistical rather than a deterministic continuum—needs to be adopted is wavefront dynamics in random media, whether analyzed via TIV., RT, or ET. Thus, the wavefront (of thickness L) is a traveling SVE that becomes a deterministic RVE in the limit of microheterogeneity being infinitesimal relative to the thickness ($d/L \rightarrow 0$); as a result, the evolution is stochastic and its average may be different from the solution of an idealized homogeneous medium problem ([45]).

Acknowledgment

This research was made possible by the National Science Foundation through a grant CMS-9713764 and the USDA through a grant 99-35504-8672.

References

[1] Maugin, G. A., 1999, *The Thermomechanics of Nonlinear Irreversible Behaviors—An Introduction*, World Scientific, Singapura.
[2] Ziegler, H., 1963, ''Some Extremum Principles in Irreversible Thermodynamics With Applications to Continuum Mechanics,'' *Progress in Solid Mechanics*, Vol. 42, North-Holland, Amsterdam, pp. 91–191.
[3] Muschik, W., and Fang, J., 1989, ''Statistical Foundations of Non-equilibrium Contact Quantities. Bridging Phenomenological and Statistical Non-equilibrium Thermodynamics,'' Acta Phys. Hung., **66**, pp. 39–57.
[4] Lemaitre, J., and Chaboche, J.-L., 1990, *Mechanics of Solid Materials*, Cambridge University Press, Cambridge, UK.
[5] Maugin, G. A., 1992, *The Thermomechanics of Plasticity and Fracture*, Cambridge University Press, Cambridge, UK.
[6] Mei, C. C., Auriault, J.-L., and Ng, C.-O., 1996, ''Some Applications of the Homogenization Theory,'' Adv. Appl. Mech., **32**, pp. 277–348.
[7] Dvorak, G. J., 1997, ''Thermomechanics of Heterogeneous Media,'' J. Therm. Stresses, **20**, pp. 799–817.
[8] Drugan, W. J., and Willis, J. R., 1996, ''A Micromechanics-Based Nonlocal Constitutive Equation and Estimates of Representative Volume Element Size for Elastic Composites,'' J. Mech. Phys. Solids, **44**, pp. 497–524.
[9] Gusev, A. A., 1997, ''Representative Volume Element Size for Elastic Composites: A Numerical Study,'' J. Mech. Phys. Solids, **45**, No. 9, pp. 1449–1459.
[10] Jiang, M., Ostoja-Starzewski, M., and Jasiuk, I., 2001, ''Scale-Dependent Bounds on Effective Elastoplastic Response of Random Composites,'' J. Mech. Phys. Solids, **49**, No. 3, pp. 655–673.
[11] Bažant, Z. P., and Planas, J., 1998, *Fracture and Size Effect in Concrete and Other Quasibrittle Materials*, CRC Press, Boca Raton, FL.
[12] Chudnovsky, A., 1977, *The Principles of Statistical Theory of the Long Time Strength*, Novosibirsk (in Russian).
[13] Krajcinovic, D., 1996, *Damage Mechanics*, North-Holland, Amsterdam.
[14] Krajcinovic, D., 1997, ''Essential Structure of the Damage Mechanics Theories,'' *Theor. Appl. Mech.*, T. Tatsumi, E. Watanabe, and T. Kambe, eds., Elsevier Science, Amsterdam, pp. 411–426.
[15] Hill, R., 1963, ''Elastic Properties of Reinforced Solids: Some Theoretical Principles,'' J. Mech. Phys. Solids, **11**, pp. 357–372.
[16] Hazanov, S., and Huet, C., 1994, ''Order Relationships for Boundary Conditions Effect in Heterogeneous Bodies Smaller Than the Representative Volume,'' J. Mech. Phys. Solids, **41**, pp. 1995–2011.
[17] Ziegler, H., 1983, *An Introduction to Thermomechanics*, North-Holland, Amsterdam.
[18] Ziegler, H., and Wehrli, C., 1987, ''The Derivation of Constitutive Relations From Free Energy and the Dissipation Functions,'' *Adv. Appl. Mech.*, **25**, Academic Press, New York, pp. 183–238.
[19] Germain, P., Nguyen, Quoc Son, and Suquet, P., 1983, ''Continuum Thermodynamics,'' ASME J. Appl. Mech., **50**, pp. 1010–1020.
[20] Beran, M. J., 1974, ''Application of Statistical Theories for the Determination of Thermal, Electrical, and Magnetic Properties of Heterogeneous Materials,'' *Mechanics of Composite Materials*, Vol. 2, G. P. Sendeckyj, ed., Academic Press, San Diego, CA, pp. 209–249.
[21] Nemat-Nasser, S., and Hori, M., 1993, *Micromechanics: Overall Properties of Heterogeneous Materials*, North-Holland, Amsterdam.
[22] Jeulin, D., and Ostoja-Starzewski, M., eds., 2001, *Mechanics of Random and Multiscale Microstructures* (CISM Courses and Lectures), Springer-Verlag, New York, in press.
[23] Stoyan, D., Kendall, W. S., and Mecke, J., 1987, *Stochastic Geometry and its Applications*, J. Wiley and Sons, New York.
[24] Dempsey, J. P., 2000, ''Research Trends in Ice Mechanics,'' Int. J. Solids Struct., **37**, No. 1–2.
[25] Rytov, S. M., Kravtsov, Yu. A., and Tatarskii, V. I., 1989, *Principles of Statistical Radiophysics*, Vol. 3, Springer-Verlag, Berlin.
[26] Huet, C., 1990, ''Application of Variational Concepts to Size Effects in Elastic Heterogeneous Bodies,'' J. Mech. Phys. Solids, **38**, pp. 813–841.
[27] Sab, K., 1992, ''On the Homogenization and the Simulation of Random Materials,'' Eur. J. Mech. A/Solids, **11**, pp. 585–607.
[28] Ostoja-Starzewski, M., 1994, ''Micromechanics as a Basis of Continuum Random Fields,'' Appl. Mech. Rev., **47**, (No. 1, Part 2), pp. S221–S230.
[29] Ostoja-Starzewski, M., 1998, ''Random Field Models of Heterogeneous Materials,'' Int. J. Solids Struct., **35**, No. 19, pp. 2429–2455.
[30] Ostoja-Starzewski, M., 1999, ''Scale Effects in Materials With Random Distributions of Needles and Cracks,'' Mech. Mater., **31**, No. 12, pp. 883–893.
[31] Ostoja-Starzewski, M., 1999, ''Microstructural Disorder, Mesoscale Finite Elements, and Macroscopic Response,'' Proc. R. Soc. London, Ser. A, **A455**, pp. 3189–3199.
[32] Hazanov, S., 1998, ''Hill Condition and Overall Properties of Composites,'' Arch. Appl. Mech., **66**, pp. 385–394.
[33] Hazanov, S., 1999, ''On Apparent Properties of Nonlinear Heterogeneous Bodies Smaller Than the Representative Volume,'' Acta Mech., **134**, pp. 123–134.
[34] Huet, C., 1999, ''Coupled Size and Boundary-Condition Effects in Viscoelastic Heterogeneous Composite Bodies,'' Mech. Mater., **31**, No. 12, pp. 787–829.
[35] Hazanov, S., and Amieur, M., 1995, ''On Overall Properties of Elastic Heterogeneous Bodies Smaller Than the Representative Volume,'' Int. J. Eng. Sci., **33**, No. 9, pp. 1289–1301.
[36] Ostoja-Starzewski, M., 1990, ''A Generalization of Thermodynamic Orthogonality to Random Media,'' J. Appl. Math. Phys., **41**, pp. 701–712.
[37] Ziegler, H., 1968, ''A Possible Generalization of Onsager's Theory,'' *IUTAM Symp. Irreversible Aspects of Continuum Mechanics*, H. Parkus and L. I. Sedov, eds., Springer-Verlag, New York, pp. 411–424.
[38] Ashby, M. F., and Jones, D. R. H., 1986, *Engineering Materials*, Vol. 2, Pergamon Press, Oxford.
[39] Ostoja-Starzewski, M., Sheng, P. Y., and Jasiuk, I., 1997, ''Damage Patterns and Constitutive Response of Random Matrix-Inclusion Composites,'' Eng. Fract. Mech., **58**, Nos. 5 and 6, pp. 581–606.
[40] Alzebdeh, K., Al-Ostaz, A., Jasiuk, I., and Ostoja-Starzewski, M., 1998, ''Fracture of Random Matrix-Inclusion Composites: Scale Effects and Statistics,'' Int. J. Solids Struct., **35**, No. 19, pp. 2537–2566.
[41] Hermann, H. J., and Roux, S., eds., 1990, *Statistical Models for Fracture of Disordered Media*, Elsevier, Amsterdam.
[42] Ostoja-Starzewski, M., Jasiuk, I., Wang, W., and Alzebdeh, K. 1996, ''Composites With Functionally Graded Interfaces: Meso-continuum Concept and Effective Properties,'' Acta Mater., **44**, No. 5, pp. 2057–2066.
[43] Zbib, H., and Aifantis, E. C., 1989, ''A Gradient Dependent Flow Theory of Plasticity: Application to Metal and Soil Instabilities,'' Appl. Mech. Rev., (Mechanics Pan-America 1989, C. R. Steele, A. W. Leissa, and M. R. M. Crespo da Silva, eds.), **42**, No. 11, Part 2, pp. S295–S304.
[44] Lacy, T. E., McDowell, D. L., and Talreja, R., 1999, ''Gradient Concepts for Evolution of Damage,'' Mech. Mater., **31**, No. 12, pp. 831–860.
[45] Ostoja-Starzewski, M., and Trębicki, J., 1999, ''On the Growth and Decay of Acceleration Waves in Random Media,'' Proc. R. Soc. London, Ser. A, **A455**, pp. 2577–2614.

On Perturbation Solutions for Nearly Circular Inclusion Problems in Plane Thermoelasticity

C.-H. Wang
Graduate Student

C.-K. Chao[1]
Professor

Department of Mechanical Engineering,
National Taiwan University of Science and Technology,
43 Keelung Road, Section 4,
Taipei, Taiwan 106, R.O.C.

An approximate analytical solution to the nearly circular inclusion problems of arbitrary shape in plane thermoelasticity is provided. The shape of the inclusion boundary considered in the present study is assumed to have the form $r = a_0[1 + A(\theta)]$, *where* a_0 *is the radius of the unperturbed circle and* $A(\theta)$ *is the radius perturbation magnitude that is represented by a Fourier series expansion. The proposed method in this study is based on the complex variable theory, analytical continuation theorem, and the boundary perturbation technique. Originating from the principle of superposition, the solution of the present problem is composed of the reference and the perturbation terms that the reference term is the known exact solution pertaining to the case with circular inclusion. First-order perturbation solutions of both temperature and stress fields are obtained explicitly for elastic inclusions of arbitrary shape. To demonstrate the derived general solutions, two typical examples including elliptical and smooth polygonal inclusions are discussed in detail. Compared to other existing approaches for elastic inclusion problems, our methodology presented here is remarked by its efficiency and applicability to inclusions of arbitrary shape in a plane under thermal load.* [DOI: 10.1115/1.1410367]

1 Introduction

In many engineering applications, determination of thermal stresses is taking a more and more important role in analyzing the service life of advanced composite materials exposed to high temperature environments. In the cases that there is an inhomogeneity existing in an infinite matrix subjected to a remote heat flux, the insertion would induce nonuniform temperature disturbance that can create nearby stress concentration and result in material degradation. The example of such problems was studied earlier by Florence and Goodier [1,2] who solved the stress field for insulated ovaloid holes or spherical cavities embedded in an infinite matrix. Since then a number of the hole or inclusion problems have been studied, such as Chen [3] for orthotropic materials with an elliptic hole based on the complex variable technique developed by Green and Zerna [4], Hwu [5], for an anisotropic medium with an elliptic hole based on Stroh formalism (Stroh, [6]), Kattis and Meguid [7], for isotropic media with an elastic inclusion based on the complex variable representations, and Chao and Shen [8], for an elliptic inclusion problem in a generally anisotropic body based on the method of Lekhnitskii formulation and the technique of conformal mapping. Among the aforementioned studies, the shape geometry of elastic inclusion problems can be successfully treated up to an ellipse by using the conformal mapping technique with additional restriction in the inclusion domain to remedy the discontinuity to obtain one-to-one transformation (Hwu and Yen, [9]). It is known that the technique of conformal mapping is one of the most powerful methods for the solution of boundary value problems for awkwardly shaped regions. However, due to the lack of a conformal mapping, which maps simultaneously the exterior and interior of the inclusion onto the one with simpler shape, this technique is unable to directly apply elastic inclusions of any shape (see, e.g., Jaswon and Bhargava [10], Sendeckyj [11], and Ru [12]). The idea of neutrality in the context of inhomogeneities of various shapes imbedded in a two-dimensional elastic body is another important topic (see, e.g., Ru [13], and Benveniste and Miloh, [14]). For a nonideal interface between the matrix and the inhomogeneity, Ru [13] showed that if the interface parameter which relates the discontinuity of the displacement to the tractions at the interface is properly chosen, then a two-dimensional stress field in the original body will remain undisturbed after the introduction of the inhomogeneity. The existence of neutral inhomogeneities of various shapes in a conduction problem was studied by Benveniste and Miloh [14] who derived the conditions to be satisfied by the field variables at a nonideal interface with a variable interface parameter. To the best of our knowledge, no analytical exact solution is available for inclusions of any shape except the one with a family of ellipses. Hence, from a practical viewpoint, an approximate general methodology that gives an analytical solution for elastic inclusions of any shape is of great interest. The present study is then triggered by such a desire.

The magnification of stresses around the inhomogeneity embedded in an infinite matrix under a remote uniform heat flux is of great importance in engineering design. The problem becomes more and more attractive and applicable when the inclusion is of noncircular shape since the local stress could be significantly affected by the shape of inclusions. In the present work, the boundary value problem of an inclusion of arbitrary shape in plane thermoelasticity is considered. The boundary of the inclusion can be arbitrary, that is characterized by a Fourier series expansion. As mentioned in the last paragraph, the technique of conformal mapping is unable to solve the inclusion problem with nonelliptically shaped regions. An approximate methodology based on the perturbation technique (Gao [15]) by introducing a small real quantity that denotes the deviation of the inclusion from that of a circle is adopted in this study. Based on the principle of superposition, the solution of the present problem is composed of the reference

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, October 2, 2000; final revision, June 5, 2001. Associate Editor: J. R. Barber. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

and the perturbation terms that the reference term is the analytical exact solution corresponding to the problem associated with a circular inclusion. First-order solutions are derived in an explicit form for nearly circular inclusions of arbitrary shape that are viewed as being perturbed from a reference circular inclusion. Two typical examples associated with elliptical and smooth polygonal inclusions are solved explicitly and discussed in detail. In order to verify the validness and efficiency of the present methodology, the solutions associated with the hole problems, which are reduced from the present results, are compared with the existing exact solutions (Kattis [16]).

2 Plane Thermoelasticity for Circular Inclusion Problem

Consider a circular inclusion of radius a_0 existing in and being perfectly bonded to an infinite matrix (Fig. 1). We denote the regions occupied by the inclusion and the matrix as S_1 and S_2, respectively, throughout this paper; the quantities for the regions of their own are denoted by the corresponding subscripts. In the regular procedure on solving the thermoelasticity problem, the temperature function must be determined first and then for the stress functions. For the two-dimensional heat conduction problem, the resultant heat flow Q_j and the temperature T_j are related to the temperature functions $g_j'(z)$ as

$$Q_j = -k_j \,\mathrm{Im}[g_j'(z)], \tag{1}$$

$$T_j = \mathrm{Re}[g_j'(z)], \tag{2}$$

where the notations Re and Im stand for the real part and imaginary part of a complex value function, respectively, and k_j for the heat conductivity. Regarding the formulations for thermal stresses, the two components of the displacements and traction force in the Cartesian coordinates can be expressed in the following equations ([17]):

$$2\mu_j(u_j+iv_j) = \kappa_j\phi_j(z) - z\overline{\phi_j'(z)} - \overline{\psi_j(z)} + 2\mu_j\beta_j \times \int g_j'(z)dz, \tag{3}$$

$$-Y_j+iX_j = \phi_j(z) + z\overline{\phi_j'(z)} + \overline{\psi_j(z)}, \tag{4}$$

where μ_j is the shear modulus, and $\kappa_j=(3-\nu_j)/(1+\nu_j)$, $\beta_j=\alpha_j$ for the plane stress and $\kappa_j=(3-4\nu_j)$, $\beta_j=(1+\nu_j)\alpha_j$ for the plane strain with ν_j being the Poisson's ratios and α_j being the linear thermal expansion coefficients. In the following derivations, the stress-state notations are adopted in the Cartesian coordinates as

$$\Theta_j = (\tau_{xx}+\tau_{yy})_j = 2[\phi_j'(z)+\overline{\phi_j'(z)}], \tag{5}$$

$$\Sigma_j = (\tau_{yy}-\tau_{xx}+2i\tau_{xy})_j = 2[\overline{z}\phi_j''(z)+\psi_j'(z)], \tag{6}$$

where the Θ_j and Σ_j are known, respectively, as the hydrostatic and deviatoric of a stress state.

Assume that there is no heat source (or singularity) being situated in the inclusion, the temperature functions in the circular inclusion and in the matrix, respectively, can be written as

$$g_{0(1)}'(z) = \tilde{g}_{(1)}'(z), \quad \text{for } |z|\leqslant a_0, \tag{7}$$

$$g_{0(2)}'(z) = g_{00}'(z) + \tilde{g}_{(2)}'(z), \quad \text{for } |z|\geqslant a_0, \tag{8}$$

where $g_{00}'(z)$ is the temperature function for a homogeneous infinite solid and $\tilde{g}_{(j)}'(z)$ are the perturbed functions caused by field perturbation due to the existence of the inclusion. From Eqs. (1) and (2) with the thermal interface continuity conditions, $Q_1=Q_2$ and $T_1=T_2$, along $z=\zeta=a_0e^{i\theta}$, the temperature functions are found as ([18])

$$g_{0(1)}'(z) = \frac{2k_2}{k_2+k_1}g_{00}'(z), \quad \text{for } |z|\leqslant a_0, \tag{9}$$

and

$$g_{0(2)}'(z) = g_{00}'(z) + \frac{k_2-k_1}{k_2+k_1}\overline{g_{00}}\left(\frac{a_0^2}{z}\right), \quad \text{for } |z|\geqslant a_0. \tag{10}$$

For solving the stress field of elastic inclusion problems, the determination of the forms of the stress functions must be properly chosen such that both the displacements and resultant force are single-valued for any enclosed loop either in the matrix or in the inclusion field. For this consideration, the two stress functions, $\phi_{0(j)}$ and $\psi_{0(j)}$, are now expressed as

$$\phi_{0(j)}(z) = f_{\phi(j)}(z) + \phi_{0(j)}^*(z), \tag{11}$$

$$\psi_{0(j)}(z) = f_{\psi(j)}(z) + \psi_{0(j)}^*(z), \tag{12}$$

where the functions $f_{\phi(j)}$ and $f_{\psi(j)}$ must satisfy the following equations ([18]):

$$[\kappa_j f_{\phi(j)} - z\overline{f_{\phi(j)}'} - \overline{f_{\psi(j)}}]_{c_j} = -2\mu_j\beta_j \oint_{c_j} g_j'(t)dt, \tag{13}$$

and

$$[f_{\phi(j)} + z\overline{f_{\phi(j)}'} + \overline{f_{\psi(j)}}]_{c_j} = 0, \tag{14}$$

with the notation $[\,]_{c_j}$ being the increment of a function in the bracket when enclosing any contour c_j in S_j. The parts $\phi_{0(j)}^*$ and $\psi_{0(j)}^*$ are two holomorphic functions in their corresponding field S_j that can be expressed in series form as

$$\phi_{0(1)}^*(z) = \sum_{m=0}^{\infty} N_m z^m; \quad \psi_{0(1)}^*(z) = \sum_{m=0}^{\infty} P_m z^m; \tag{15}$$

$$\phi_{0(2)}^*(z) = \sum_{m=0}^{\infty} L_m z^{-m}; \quad \psi_{0(2)}^*(z) = \sum_{m=0}^{\infty} M_m z^{-m} \tag{16}$$

where the coefficients in the above equations can be determined from the interface continuity conditions.

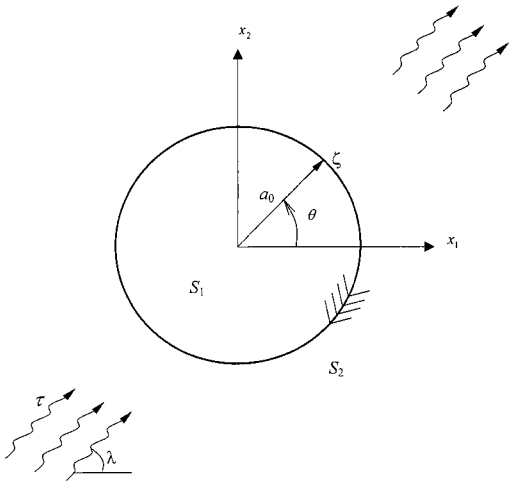


Fig. 1 Circular inclusion in an infinite plane under remote heat flow

3 Plane Thermoelasticity for Nearly Circular Inclusion Problems

3.1 Temperature Field Induced by a Nearly Circular Inclusion. Consider a nearly circular inclusion existing in an infinite solid that the boundary of the inclusion is slightly different

from a reference circle of radius a_0 and the points along the interface between the matrix and inclusion can be expressed as

$$z=r(\theta)e^{i\theta}=[1+A(\theta)]a_0e^{i\theta}, \tag{17}$$

where the small real quantity $A(\theta)$ is obviously the perturbation magnitude at the given position θ (Fig. 2). Denote the point ζ on the circumference $|z|=a_0$ and $\tilde{\zeta}$ for the point on the actual interface, Eq. (17) can be equivalently rewritten as

$$\tilde{\zeta}=[1+A(\zeta)]\zeta, \quad \zeta=a_0e^{i\theta}, \tag{18}$$

where ζ and $\tilde{\zeta}$ are of the same polar coordinate argument θ. The temperature functions $g'_j(z)$ for the nearly circular inclusion problems can be expressed in the perturbation form as follows:

$$g'_j(z)=g'_{0(j)}(z)+g'_{\varepsilon(j)}(z), \quad j=1,2 \text{ for } z\in S_j, \tag{19}$$

where $g'_{0(j)}(z)$ are the temperature functions of the circular inclusion problem, say the reference temperature functions, and $g'_{\varepsilon(j)}(z)$ are the perturbation terms due to the reference circumference at $|z|=a_0$ being perturbed by $A(\zeta)$ as Eq. (17) indicated. Based on a first-order approximation, Eq. (19) at the interface can then be expressed as

$$\begin{aligned} g'_j(\tilde{\zeta})&=g'_{0(j)}(\tilde{\zeta})+g'_{\varepsilon(j)}(\tilde{\zeta})\\ &\cong g'_{0(j)}(\zeta)+A(\zeta)\zeta g''_{0(j)}(\zeta)+g'_{\varepsilon(j)}(\zeta), \quad j=1,2. \end{aligned} \tag{20}$$

Since the resultant heat flux and the temperature are assumed to be continuous across the interface and according to the expressions in Eq. (1) and (2), the thermal continuity conditions lead to

$$[kg'(\tilde{\zeta})-k\overline{g'(\tilde{\zeta})}]_1^2=0, \tag{21}$$

and

$$[g'(\tilde{\zeta})+\overline{g'(\tilde{\zeta})}]_1^2=0, \tag{22}$$

where the symbol $[f]_1^2$ denotes for the function quantity jump across the interface from the inclusion boundary to the matrix boundary, i.e., f_2-f_1. Similarly, the thermal continuity conditions for the reference temperature function can also be expressed as

$$[kg'_0(\zeta)-k\overline{g'_0(\zeta)}]_1^2=0, \tag{23}$$

and

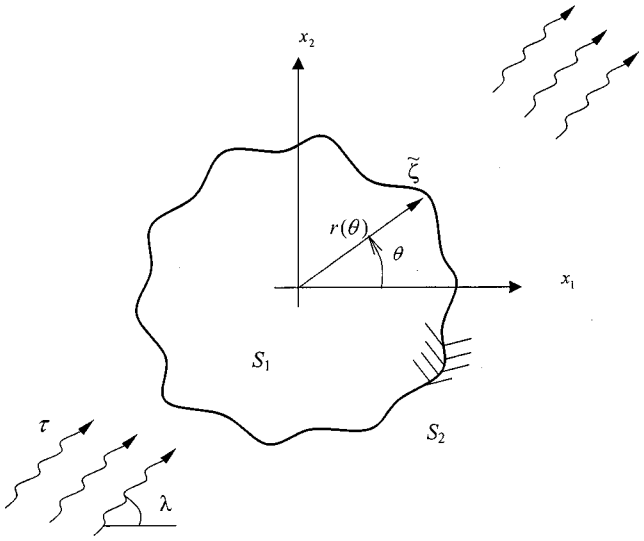


Fig. 2 Inclusion of arbitrary shape in an infinite plane under remote heat flow

$$[g'_0(\zeta)+\overline{g'_0(\zeta)}]_1^2=0. \tag{24}$$

Substituting Eq. (20) into Eq. (21) and (22) and using Eq. (23) and (24), we obtain a set of equations as

$$\begin{aligned} &k_1A(\zeta)\zeta g''_{0(1)}(\zeta)-k_2A(\zeta)\zeta g''_{0(2)}(\zeta)+k_1g'_{\varepsilon(1)}(\zeta)-k_2g'_{\varepsilon(2)}(\zeta)\\ &\quad-k_1\overline{A(\zeta)\zeta g''_{0(1)}(\zeta)}+k_2\overline{A(\zeta)\zeta g''_{0(2)}(\zeta)}-k_1\overline{g'_{\varepsilon(1)}(\zeta)}\\ &\quad+k_2\overline{g'_{\varepsilon(2)}(\zeta)}=0, \end{aligned} \tag{25}$$

and

$$\begin{aligned} &A(\zeta)\zeta g''_{0(1)}(\zeta)-A(\zeta)\zeta g''_{0(2)}(\zeta)+g'_{\varepsilon(1)}(\zeta)-g'_{\varepsilon(2)}(\zeta)\\ &\quad+\overline{A(\zeta)\zeta g''_{0(1)}(\zeta)}-\overline{A(\zeta)\zeta g''_{0(2)}(\zeta)}+\overline{g'_{\varepsilon(1)}(\zeta)}-\overline{g'_{\varepsilon(2)}(\zeta)}=0. \end{aligned} \tag{26}$$

Once the reference temperature functions $g'_{0(j)}(z)$ are obtained, the perturbation terms of the temperature functions $g'_{\varepsilon(j)}(z)$ can be determined from (25) and (26) by some basic methods of solution, such as analytical continuation, Cauchy integral, and series expansion, etc. In the present study, the method of analytical continuation is used which is much more powerful and allows boundary value problems to be solved with comparative ease.

3.2 Stress Field Induced by a Nearly Circular Inclusion

The two stress functions for the nearly circular inclusion problem can be expressed as

$$\phi_j(z)=\phi_{0(j)}(z)+\phi_{\varepsilon(j)}(z), \tag{27}$$

$$\psi_j(z)=\psi_{0(j)}+\psi_{\varepsilon(j)}(z), \tag{28}$$

where $\phi_{0(j)}(z)$ and $\psi_{0(j)}(z)$ are the stress functions for the reference circular inclusion problem while $\phi_{\varepsilon(j)}(z)$ and $\psi_{\varepsilon(j)}(z)$ are the perturbation terms due to the reference circle at $|z|=a_0$ being perturbed by $A(\zeta)$. These stress functions at the interface can be approximately expressed as

$$\phi_j(\tilde{\zeta})\cong\phi_{0(j)}(\zeta)+A(\zeta)\zeta\phi'_{0(j)}(\zeta)+\phi_{\varepsilon(j)}(\zeta), \tag{29}$$

$$\psi_j(\tilde{\zeta})\cong\psi_{0(j)}(\zeta)+A(\zeta)\zeta\psi'_{0(j)}(\zeta)+\psi_{\varepsilon(j)}(\zeta). \tag{30}$$

Since both the traction and the displacements at the actual interface must be continuous and according to the expressions in (3) and (4), the interface continuity conditions lead to

$$[\phi(\tilde{\zeta})+\tilde{\zeta}\overline{\phi'(\tilde{\zeta})}+\overline{\psi(\tilde{\zeta})}]_1^2=0, \tag{31}$$

and

$$\left[\frac{-\kappa\phi(\tilde{\zeta})+\tilde{\zeta}\overline{\phi'(\tilde{\zeta})}+\overline{\psi(\tilde{\zeta})}}{\mu}-2\beta g(\tilde{\zeta})\right]_1^2=0. \tag{32}$$

For the reference circular inclusion problem, the interface continuity conditions can also be expressed as

$$[\phi_0(\zeta)+\zeta\overline{\phi'_0(\zeta)}+\overline{\psi_0(\zeta)}]_1^2=0, \tag{33}$$

and

$$\left[\frac{-\kappa\phi_0(\zeta)+\zeta\overline{\phi'_0(\zeta)}+\overline{\psi_{0(\zeta)}}}{\mu}-2\beta g_0(\zeta)\right]_1^2=0. \tag{34}$$

Differentiating (34) with respect to ζ while noting that $\overline{\zeta}=a_0^2/\zeta$ yields

$$\begin{aligned} \left[\frac{-\kappa\phi'_0(\zeta)+\overline{\phi'_0(\zeta)}}{\mu}\right]_1^2&=\frac{a_0^2}{\zeta^2}\left[\frac{\zeta\overline{\phi''_0(\zeta)}+\overline{\psi'_0(\zeta)}}{\mu}\right]_1^2+[2\beta g'_0(\zeta)]_1^2\\ &=\frac{\overline{\zeta}}{2\zeta}\left[\frac{\overline{\Sigma_0(\zeta)}}{\mu}\right]_1^2+[2\beta g'_0(\zeta)]_1^2 \end{aligned} \tag{35}$$

where

$$\Sigma_0(\zeta)=(\tau_{yy}-\tau_{xx}+2i\tau_{xy})_0=2[\overline{\zeta}\phi_0''(\zeta)+\psi_0'(\zeta)]. \quad (36)$$

By substituting Eqs. (29) and (30) into (31) and (32) and making use of Eqs. (33), (34), and (35), we obtain a system of equations as

$$\phi_{\varepsilon(1)}(\zeta)-\phi_{\varepsilon(2)}(\zeta)+\zeta\overline{\phi'_{\varepsilon(1)}(\zeta)}-\zeta\overline{\phi'_{\varepsilon(2)}(\zeta)}+\overline{\psi_{\varepsilon(1)}(\zeta)}-\overline{\psi_{\varepsilon(2)}(\zeta)}$$
$$=-\overline{\zeta}A(\zeta)[\overline{\Sigma_{0(1)}(\zeta)}-\overline{\Sigma_{0(2)}(\zeta)}], \quad (37)$$

and

$$\frac{-\kappa_1\phi_{\varepsilon(1)}(\zeta)+\zeta\overline{\phi'_{\varepsilon(1)}(\zeta)}+\overline{\psi_{\varepsilon(1)}(\zeta)}}{\mu_1}$$
$$+\frac{\kappa_2\phi_{\varepsilon(2)}(\zeta)-\zeta\overline{\phi'_{\varepsilon(2)}(\zeta)}-\overline{\psi_{\varepsilon(2)}(\zeta)}}{\mu_2}$$
$$=-\overline{\zeta}A(\zeta)\left[\frac{\overline{\Sigma_{0(1)}(\zeta)}}{\mu_1}-\frac{\overline{\Sigma_{0(2)}(\zeta)}}{\mu_2}\right]$$
$$+2\beta_1 g_{\varepsilon(1)}(\zeta)-2\beta_2 g_{\varepsilon(2)}(\zeta). \quad (38)$$

The stress perturbation functions $\phi_{\varepsilon(j)}(z)$ and $\psi_{\varepsilon(j)}(z)$ can be immediately determined from (37) and (38) once the reference stress functions and the temperature perturbation functions are obtained. Before doing this, the function forms of the stress perturbation functions must be defined first that will make the derivation more obvious and easier. The typical solving procedure will be illustrated in the following section.

4 Elastic Inclusion of Arbitrary Shape Under Remote Uniform Heat Flux

4.1 Temperature Field. Consider an elastic inclusion of arbitrary shape embedded in an infinite matrix under remote uniform heat flux (Fig. 2). The boundary of the inclusion considered in this work is represented by a Fourier series expansion such that the radius perturbation magnitude $A(\zeta)$ in (18) along the inclusion boundary can be expressed as

$$A(\zeta)=\sum_{m=-\infty}^{\infty}C_m\left(\frac{\zeta}{a_0}\right)^m. \quad (39)$$

The temperature function for an infinite solid under a uniform heat flux is given by

$$g'_{00}(z)=\tau e^{-i\lambda}z, \quad (40)$$

where τ is the temperature gradient and λ the angle of the heat flux with respect to the positive real axis. By substituting Eq. (40) into (9) and (10), the reference temperature functions are found as

$$g'_{0(1)}(z)=\omega z, \quad \text{for } z\in S_1, \quad (41)$$

$$g'_{0(2)}(z)=\alpha z+\beta a_0^2\frac{1}{z}, \quad \text{for } z\in S_2 \quad (42)$$

where

$$\alpha=\tau e^{-i\lambda}; \quad \beta=\frac{k_2-k_1}{k_1+k_2}\tau e^{i\lambda}; \quad \omega=\frac{2k_2}{k_1+k_2}\tau e^{-i\lambda}. \quad (43)$$

Applying (39), (41), and (42) at the interface gives

$$A(\zeta)\zeta g''_{0(1)}(\zeta)=\sum_{m=-\infty}^{\infty}\omega a_0^{1-m}C_{m-1}\zeta^m, \quad (44)$$

$$A(\zeta)\zeta g''_{0(2)}(\zeta)=\sum_{m=-\infty}^{\infty}(\alpha C_{m-1}-\beta C_{m+1})a_0^{1-m}\zeta^m. \quad (45)$$

By substituting the above Eqs. (44) and (45) into Eqs. (25) and (26) and putting all the terms of positive or zero order of ζ on the one side and the terms of negative order of ζ on the other side, one can obtain a system of equations through the use of analytical continuation method as

$$\sum_{m=0}^{\infty}[(k_1\omega-k_2\alpha)C_{m-1}+k_2\beta C_{m+1}+(k_2\overline{\alpha}-k_1\overline{\omega})\overline{C_{-m-1}}$$
$$-k_2\overline{\beta C_{1-m}}]a_0^{1-m}\zeta^m+k_1 g'_{\varepsilon(1)}(\zeta)+k_2\overline{g'_{\varepsilon(2)}(\zeta)}=0, \quad (46)$$

$$\sum_{m=0}^{\infty}[(\omega-\alpha)C_{m-1}+\beta C_{m+1}-(\overline{\alpha}-\overline{\omega})\overline{C_{-m-1}}+\overline{\beta C_{1-m}}]a_0^{1-m}\zeta^m$$
$$+g'_{\varepsilon(1)}(\zeta)-\overline{g'_{\varepsilon(2)}(\zeta)}=0. \quad (47)$$

Solving (46) and (47) for the temperature perturbation functions $g'_{\varepsilon(j)}(z)$ yields

$$g'_{\varepsilon(1)}=\sum_{m=0}^{\infty}\left[\frac{-2k_2}{k_1+k_2}\beta(C_{1+m}+\overline{C_{-1-m}})\right]a_0^{1-m}z^m, \quad z\in S_1 \quad (48)$$

$$g'_{\varepsilon(2)}(z)=\sum_{m=1}^{\infty}\left[-\frac{k_2-k_1}{k_2+k_1}\overline{\beta}(C_{-1-m}+\overline{C_{1+m}})+\beta(C_{1-m}\right.$$
$$\left.+\overline{C_{-1+m}})\right]a_0^{1+m}z^{-m}, \quad z\in S_2. \quad (49)$$

Integration of (48) and (49) with respect to z gives

$$g_{\varepsilon(1)}(z)=\sum_{m=1}^{\infty}\left[\frac{-2k_2}{m(k_1+k_2)}\beta(C_m+\overline{C_{-m}})\right]a_0^{2-m}z^m, \quad z\in S_1, \quad (50)$$

$$g_{\varepsilon(2)}(z)=\gamma_\varepsilon\log z+\sum_{m=1}^{\infty}\left[\frac{k_2-k_1}{m(k_2+k_1)}\overline{\beta}(C_{-2-m}+\overline{C_{2+m}})\right.$$
$$\left.+\beta(C_{-m}+\overline{C_m})\right]a_0^{2+m}z^{-m}, \quad z\in S_2 \quad (51)$$

where

$$\gamma_\varepsilon=\left[-\frac{k_2-k_1}{k_2+k_1}\overline{\beta}(C_{-2}+\overline{C_2})+\beta(C_0+\overline{C_0})\right]a_0^2. \quad (52)$$

4.2 Thermal Stress Field. According to the obtained reference temperature functions given in Eqs. (41) and (42), the reference stress functions are assumed to be

$$\phi_{0(j)}(z)=B_{0(j)}\ln z+\phi^*_{0(j)}(z), \quad (53)$$

$$\psi_{0(j)}(z)=C_{0(j)}\ln z+\psi^*_{0(j)}(z) \quad (54)$$

where the four constants $B_{0(j)}$ and $C_{0(j)}$ can be found according to Eqs. (13) and (14) and the coefficients of the four holomorphic functions $\phi^*_{0(j)}$ and $\psi^*_{0(j)}(z)$ can be determined from the interface continuity conditions. The final results of the reference stress functions are obtained as [16]

$$\phi_{0(1)}(z)=-\frac{\mu_1\mu_2}{\mu_2\kappa_1+\mu_1}(\beta_1\omega-\beta_2\alpha)z^2, \quad (55)$$

$$\psi_{0(1)}(z)=0, \quad (56)$$

$$\phi_{0(2)}(z)=-\frac{2\mu_2\beta_2}{1+\kappa_2}\beta a_0^2\ln z, \quad (57)$$

$$\psi_{0(2)}(z)=-\frac{2\mu_2\beta_2}{1+\kappa_2}\overline{\beta}a_0^2\ln z-\left[\frac{\mu_1}{\mu_2\kappa_1+\mu_1}(\beta_1\overline{\omega}-\beta_2\overline{\alpha})-\frac{2\beta_2\beta}{1+\kappa_2}\right]\mu_2 a_0^4 z^{-2}. \quad (58)$$

Similar to the expressions of the reference stress functions as indicated in Eqs. (53) and (54), the stress perturbation functions can also be assumed as

$$\phi_{\varepsilon(j)}(z)=B_{\varepsilon(j)}\ln z+\phi^*_{\varepsilon(j)}(z), \quad (59)$$

$$\psi_{\varepsilon(j)}(z)=C_{\varepsilon(j)}\ln z+\psi^*_{\varepsilon(j)}(z), \quad (60)$$

where the holomorphic functions $\phi^*_{\varepsilon(j)}$ and $\psi^*_{\varepsilon(j)}$ are expressed in Laurent's series form as

$$\phi^*_{\varepsilon(1)}(z)=\sum_{m=0}^{\infty}N_{\varepsilon(m)}z^m;\quad \psi^*_{\varepsilon(1)}(z)=\sum_{m=0}^{\infty}P_{\varepsilon(m)}z^m; \quad (61)$$

$$\phi^*_{\varepsilon(2)}(z)=\sum_{m=0}^{\infty}L_{\varepsilon(m)}z^{-m};\quad \psi^*_{\varepsilon(2)}(z)=\sum_{m=0}^{\infty}M_{\varepsilon(m)}z^{-m}. \quad (62)$$

Making use of Eqs. (50), (51), (59), and (60), the constants appearing in the logarithmic terms in Eqs. (59) and (60) are obtained according to Eqs. (13) and (14) as

$$B_{\varepsilon(1)}=0, \quad (63)$$

$$B_{\varepsilon(2)}=-\frac{2\mu_2\beta_2}{\kappa_2+1}\gamma_\varepsilon \quad (64)$$

$$C_{\varepsilon(1)}=0, \quad (65)$$

$$C_{\varepsilon(2)}=\frac{2\mu_2\beta_2}{\kappa_2+1}\overline{\gamma_\varepsilon}. \quad (66)$$

With the help of Eqs. (36), (39) and (55)–(58), the terms $\overline{\zeta}A(\zeta)\overline{\Sigma_{0(j)}(\zeta)}$ in Eqs. (37) and (38) can be expanded in series form as

$$\overline{\zeta}A(\zeta)\overline{\Sigma_{0(1)}(\zeta)}=\sum_{m=-\infty}^{\infty}2A_1a_0^{2-m}C_m\zeta^m, \quad (67)$$

$$\overline{\zeta}A(\zeta)\overline{\Sigma_{0(2)}(\zeta)}=\sum_{m=-\infty}^{\infty}2[(A_2-\overline{A_3})C_{m-2}+A_3C_m]a_0^{2-m}\zeta^m, \quad (68)$$

in which

$$A_1=-\frac{2\mu_1}{\mu_2\kappa_1+\mu_1}\left(\frac{2k_2\beta_1}{k_1+k_2}-\beta_2\right)\mu_2\tau e^{i\lambda}; \quad (69)$$

$$A_2=\left[\frac{2\mu_1}{\mu_2\kappa_1+\mu_1}\left(\frac{2k_2\beta_1}{k_1+k_2}-\beta_2\right)-\frac{4\beta_2}{1+\kappa_2}\left(\frac{k_2-k_1}{k_2+k_1}\right)\right]\mu_2\tau e^{-i\lambda}; \quad (70)$$

$$A_3=-\frac{2\beta_2}{1+\kappa_2}\frac{(k_2-k_1)}{(k_2+k_1)}\mu_2\tau e^{i\lambda}. \quad (71)$$

Similar to the previous approach, substitution of (50), (51), (59), (60), (67), and (68) into (37) and (38) gives the following two sets of equations as

$$\sum_{m=1}^{\infty}[N_{\varepsilon(m)}-\overline{M_{\varepsilon(m)}}a_0^{-2m}-2(A_2-\overline{A_3})a_0^{2-m}C_{m-2}-2(A_3-A_1)a_0^{2-m}C_m]\zeta^m+\sum_{m=3}^{\infty}(m-2)\overline{L_{\varepsilon(m-2)}}\times a_0^{-2m+2}\zeta^m+\overline{N_{\varepsilon(1)}}\zeta-\overline{B_{\varepsilon(2)}}a_0^{-2}\zeta^2=0 \quad (72)$$

$$\sum_{m=1}^{\infty}\left[-\frac{\kappa_1}{\mu_1}N_{\varepsilon(m)}-\frac{1}{\mu_2}\overline{M_{\varepsilon(m)}}a_0^{-2m}-\frac{2}{\mu_2}(A_2-\overline{A_3})a_0^{2-m}C_{m-2}-2\left(\frac{1}{\mu_2}A_3-\frac{1}{\mu_1}A_1\right)a_0^{2-m}C_m\right]\zeta^m+\sum_{m=3}^{\infty}\frac{(m-2)}{\mu_2}\overline{L_{\varepsilon(m-2)}}a_0^{-2m+2}\zeta^m+\sum_{m=1}^{\infty}\frac{4\beta_1k_2\beta}{m(k_1+k_2)}(C_m+\overline{C_{-m}})a_0^{2-m}\zeta^m+\frac{\overline{N_{\varepsilon(1)}}}{\mu_1}\zeta-\frac{\overline{B_{\varepsilon(2)}}}{\mu_2}a_0^{-2}\zeta^2=0 \quad (73)$$

and

$$\sum_{m=1}^{\infty}[L_{\varepsilon(m)}-(m+2)\overline{N_{\varepsilon(m+2)}}a_0^{2+2m}-\overline{P_{\varepsilon(m)}}a_0^{2m}]\zeta^{-m}+\sum_{m=1}^{\infty}[2(A_2-\overline{A_3})C_{-2-m}+2(A_3-A_1)C_{-m}]a_0^{2+m}\zeta^{-m}=0 \quad (74)$$

$$\sum_{m=1}^{\infty}\left[-\frac{\kappa_2}{\mu_2}L_{\varepsilon(m)}-\frac{(m+2)}{\mu_1}\overline{N_{\varepsilon(m+2)}}a_0^{2+2m}-\frac{1}{\mu_1}\overline{P_{\varepsilon(m)}}a_0^{2m}\right]\zeta^{-m}-\sum_{m=1}^{\infty}\frac{2\beta_2}{m}\left[\frac{(k_2-k_1)}{(k_2+k_1)}\overline{\beta}(C_{-2-m}+\overline{C_{2+m}})+\beta(C_{-m}+\overline{C_m})\right]a_0^{2+m}\zeta^{-m}+\sum_{m=1}^{\infty}\left[\frac{2}{\mu_2}(A_2-\overline{A_3})C_{-2-m}+2\left(\frac{A_3}{\mu_2}-\frac{A_1}{\mu_1}\right)C_{-m}\right]a_0^{2+m}\zeta^{-m}=0. \quad (75)$$

Solving the above four equations for the unknowns $N_{\varepsilon(m)}$, $P_{\varepsilon(m)}$, $L_{\varepsilon(m)}$, and $M_{\varepsilon(m)}$ yields

$$N_{\varepsilon(1)}=\frac{1}{(\mu_1-\mu_2)+\mu_2\kappa_1}\left[2(\mu_2-\mu_1)A_1a_0C_1+\frac{4\beta_1\mu_1\mu_2k_2\beta a_0}{(k_1+k_2)}(C_1+\overline{C_{-1}})\right], \quad (76)$$

$$N_{\varepsilon(m)}=\frac{2(\mu_2-\mu_1)}{\mu_1+\mu_2\kappa_1}A_1a_0^{2-m}C_m+\frac{4\beta_1\mu_1\mu_2k_2}{(\mu_1+\mu_2\kappa_1)(k_1+k_2)}\frac{\beta}{m}a_0^{2-m}(C_m+\overline{C_{-m}}),\quad m\geqslant 2 \quad (77)$$

$$L_{\varepsilon(m)}=\frac{2(\mu_2-\mu_1)}{\mu_2+\mu_1\kappa_2}[C_{-2-m}-A_3C_{-m}]a_0^{2+m}-\frac{2\beta_2\mu_1\mu_2}{m(\mu_2+\mu_1\kappa_2)}\left[\frac{(k_2-k_1)}{(k_2+k_1)}\overline{\beta}(C_{-2-m}+\overline{C_{2+m}})+\beta(C_{-m}+\overline{C_m})\right]a_0^{2+m},\quad m\geqslant 1 \quad (78)$$

$$P_{\varepsilon(m)}=\overline{L_{\varepsilon(m)}}a_0^{-2m}+(m+2)N_{\varepsilon(m+2)}a_0^2+2[(\overline{A_2}-A_3)\overline{C_{-2-m}}+(\overline{A_3}-\overline{A_1})\overline{C_{-m}}]a_0^{2-m},\quad m\geqslant 1 \quad (79)$$

$$M_{\varepsilon(1)}=(N_{\varepsilon(1)}+\overline{N_{\varepsilon(1)}})a_0^2+2[(\overline{A_1}-\overline{A_3})\overline{C_1}-(\overline{A_2}-A_3)\overline{C_{-1}}]a_0^3 \quad (80)$$

$$M_{\varepsilon(2)}=\overline{N_{\varepsilon(2)}}a_0^4-B_{\varepsilon(2)}a_0^2+2[(\overline{A_1}-\overline{A_3})\overline{C_2}-(\overline{A_2}-A_3)\overline{C_0}]a_0^4 \tag{81}$$

$$M_{\varepsilon(m)}=\overline{N_{\varepsilon(m)}}a_0^{2m}+(m-2)L_{\varepsilon(m-2)}a_0^2+2[(\overline{A_1}-\overline{A_3})\overline{C_m}-(\overline{A_2}-A_3)\overline{C_{m-2}}]a_0^{2+m},\quad m\geqslant 3. \tag{82}$$

The expressions (59)–(66) and (76)–(82) provide the first-order analytical solution for the inclusion problem of arbitrary shape in an entire plane under a remote uniform heat flux.

5 Examples

5.1 Elliptical Inclusion Problem. We first consider the elliptical inclusion problem of which the radius perturbation magnitude $A(\zeta)$ can be expressed as

$$A(\zeta)=\frac{\varepsilon}{2}\left(\frac{\zeta^2}{a_0^2}+\frac{a_0^2}{\zeta^2}\right) \tag{83}$$

where ε is a relatively small quantity as compared to unity. The inclusion boundary represented by (83) is an ellipse with semiaxes being $a_0(1+\varepsilon)$ and $a_0(1-\varepsilon)$. Comparing Eq. (83) to Eq. (39), the coefficients C_m are then given by

$$C_m=\begin{cases}\varepsilon/2, & m=-2,2,\\ 0, & m\neq -2,2.\end{cases} \tag{84}$$

Applying (84) into (48) and (49), we can have the temperature functions as

$$g'_{\varepsilon(1)}(z)=-\frac{2\varepsilon k_2\beta}{k_1+k_2}z,\quad z\in S_1, \tag{85}$$

and

$$g'_{\varepsilon(2)}(z)=-\varepsilon\left[\frac{k_2-k_1}{k_2+k_1}\overline{\beta}a_0^2z^{-1}-\beta a_0^4z^{-3}\right],\quad z\in S_2. \tag{86}$$

After having the constants $B_{\varepsilon(2)}$, $N_{\varepsilon(m)}$, $L_{\varepsilon(m)}$, $P_{\varepsilon(m)}$, and $M_{\varepsilon(m)}$ from the general expressions of Eqs. (64), (76)–(82), the four stress functions $\phi_j(z)$ and $\psi_j(z)$ can be determined as

$$\phi_1(z)=\left[-\frac{\mu_1\mu_2}{\mu_2\kappa_1+\mu_1}(\beta_1\omega-\beta_2\alpha)+N_{\varepsilon(2)}\right]z^2, \tag{87}$$

$$\psi_1(z)=P_{\varepsilon(2)}z^2, \tag{88}$$

for $z\in S_1$, and

$$\phi_2(z)=\left[-\frac{2\mu_2\beta_2}{(1+\kappa_2)}\beta a_0^2+B_{\varepsilon(2)}\right]\ln z+L_{\varepsilon(2)}z^{-2}, \tag{89}$$

$$\psi_2(z)=\left[\frac{-2\mu_2\beta_2}{(1+\kappa_1)}\overline{\beta}a_0^2+\overline{B}_{\varepsilon(2)}\right]\ln z-\left[\frac{\mu_1\mu_2}{\mu_2\kappa_1+\mu_1}(\beta_1\overline{\omega}-\beta_2\alpha)a_0^4+\frac{2\mu_2\beta_2}{1+\kappa_2}\beta a_0^4+M_{\varepsilon(2)}\right]z^{-2}+M_{\varepsilon(4)}z^{-4}, \tag{90}$$

for $z\in S_2$, where

$$N_{\varepsilon(2)}=\varepsilon\left[\frac{(\mu_2-\mu_1)A_1+2\mu_1\mu_2\beta_1\beta k_2(k_1+k_2)^{-1}}{\mu_2\kappa_1+\mu_1}\right], \tag{91}$$

$$L_{\varepsilon(2)}=-\varepsilon\left[\frac{(\mu_2-\mu_1)A_3+\mu_1\mu_2\beta_2\beta}{\mu_1\kappa_2+\mu_2}\right]a_0^4, \tag{92}$$

$$P_{\varepsilon(2)}=\overline{L_{\varepsilon(2)}}a_0^{-4}+\varepsilon(\overline{A_3}-\overline{A_1}), \tag{93}$$

$$M_{\varepsilon(2)}=\overline{N_{\varepsilon(2)}}-B_{\varepsilon(2)}a_0^{-2}+\varepsilon(\overline{A_1}-\overline{A_3})a_0^4, \tag{94}$$

$$M_{\varepsilon(4)}=2L_{\varepsilon(2)}a_0^2-\varepsilon(\overline{A_2}-A_3)a_0^6, \tag{95}$$

$$B_{\varepsilon(2)}=\varepsilon\frac{2\mu_2\beta_2}{\kappa_2+1}\left(\frac{k_2-k_1}{k_2+k_1}\right)\overline{\beta}a_0^2. \tag{96}$$

It is seen that from (87) and (88) the stresses inside the elliptic inclusion are always linear functions of the coordinate z which is different from the result of uniform stresses existing in the inclusion of the corresponding isothermal elastic problem. The interfacial stresses along the inclusion boundary can be performed by using field solutions of the matrix ($j=2$) or inclusion ($j=1$) as

$$\tau_{\xi\xi(j)}=\mathrm{Re}\left\{\phi_j'(\tilde{\zeta})+\overline{\phi_j'(\tilde{\zeta})}-[\tilde{\zeta}\overline{\phi_j''(\tilde{\zeta})}+\overline{\psi_j'(\tilde{\zeta})}]\left(\frac{d\tilde{x}_2}{ds}+i\frac{d\tilde{x}_1}{ds}\right)^2\right\}, \tag{97}$$

$$\tau_{\eta\eta(j)}=\mathrm{Re}\left\{\phi_j'(\tilde{\zeta})+\overline{\phi_j'(\tilde{\zeta})}+[\tilde{\zeta}\overline{\phi_j''(\tilde{\zeta})}+\overline{\psi_j'(\tilde{\zeta})}]\left(\frac{d\tilde{x}_2}{ds}+i\frac{d\tilde{x}_1}{ds}\right)^2\right\}, \tag{98}$$

$$\tau_{\eta\xi(j)}=\mathrm{Im}\left\{\phi_j'(\tilde{\zeta})+\overline{\phi_j'(\tilde{\zeta})}-[\tilde{\zeta}\overline{\phi_j''(\tilde{\zeta})}+\overline{\psi_j'(\tilde{\zeta})}]\left(\frac{d\tilde{x}_2}{ds}+i\frac{d\tilde{x}_1}{ds}\right)^2\right\}, \tag{99}$$

where

$$\frac{d\tilde{x}_1}{ds}=\frac{1}{|J|}\frac{d\tilde{x}_1}{d\theta}=\frac{1}{|J|}(-\sin\theta-\varepsilon\cos 2\theta\sin\theta-2\varepsilon\sin 2\theta\cos\theta), \tag{100}$$

$$\frac{d\tilde{x}_2}{ds}=\frac{1}{|J|}\frac{d\tilde{x}_2}{d\theta}=\frac{1}{|J|}(\cos\theta+\varepsilon\cos 2\theta\cos\theta-2\varepsilon\sin 2\theta\sin\theta), \tag{101}$$

with

$$J=\sqrt{[r'(\theta)]^2+r^2}=\sqrt{(-2\varepsilon\sin 2\theta)^2+(1+\varepsilon\cos 2\theta)^2}. \tag{102}$$

When the inclusion is assumed to be an insulated and traction-free hole, the hoop stress along the hole boundary can be obtained from (98) by letting $k_1=0$ and $\mu_1=0$ as

$$\tau_{\eta\eta}=\mathrm{Re}\{-4\delta a_0^2\zeta^{-1}+\varepsilon(2\delta\zeta+4\overline{\delta}a_0^2\zeta^{-1}-6\delta a_0^4\zeta^{-3})\},\qquad \zeta=a_0e^{i\theta}, \tag{103}$$

where

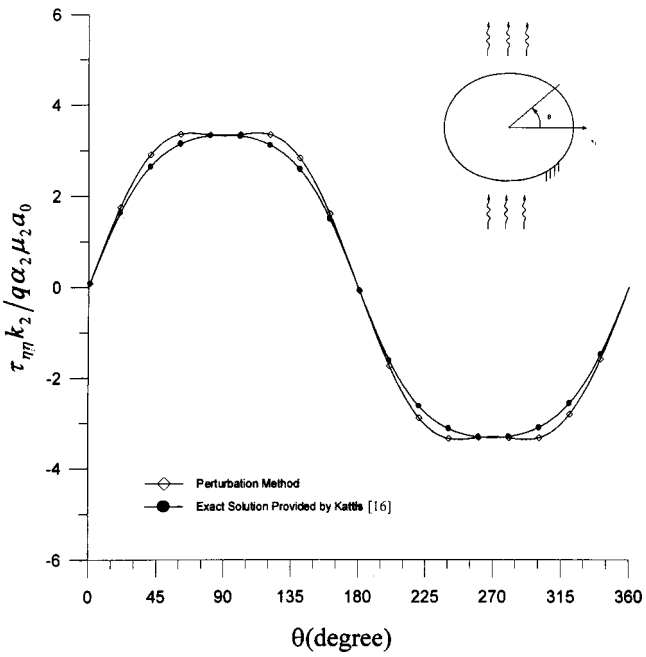


Fig. 3 Dimensionless hoop stress for the elliptical hole problem with λ=90 deg

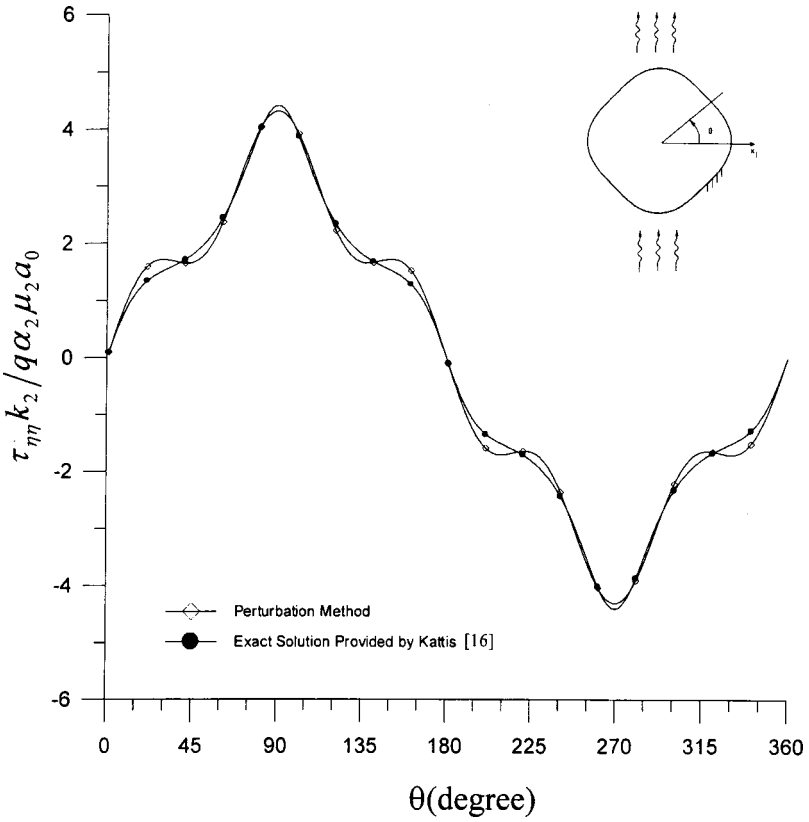


Fig. 4 Dimensionless hoop stress for the quadrilateral hole problem with λ = 90 deg

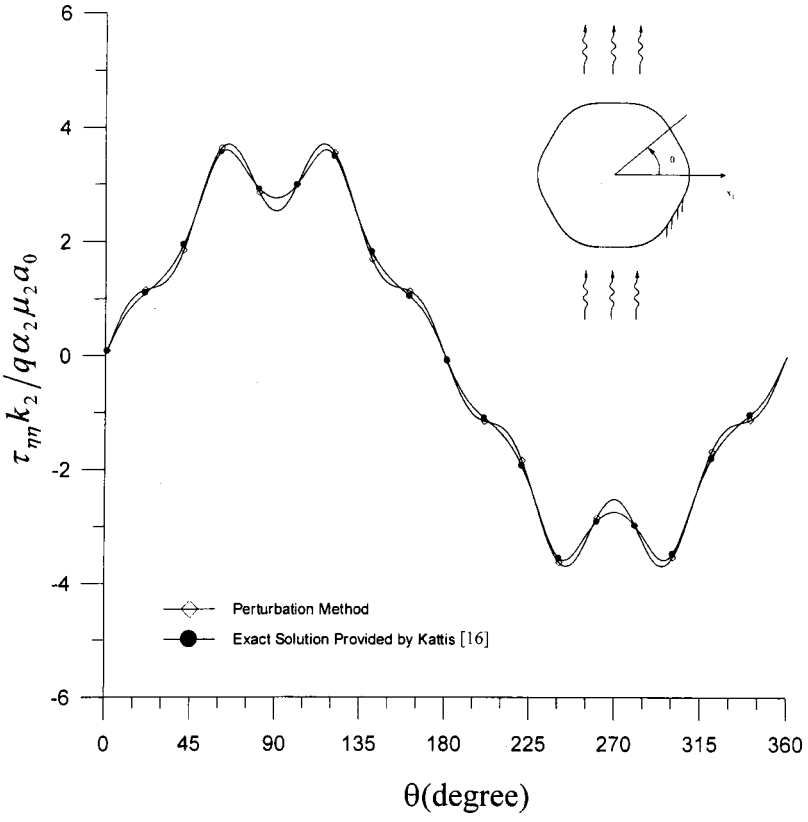


Fig. 5 Dimensionless hoop stress for the hexagonal hole problem with λ = 90 deg

$$\delta=\frac{2\mu_2\beta_2}{1+\kappa_2}\tau e^{i\lambda}. \tag{104}$$

Comparison of the dimensionless hoop stress between the present result (103) and that derived by Kattis [16] displayed in Fig. 3 shows that a good accuracy of the present first-order perturbation solution is achieved.

5.2 Smooth Polygonal Inclusion Problems. Next, we consider polygonal inclusion problems of which the radius perturbation magnitude $A(\zeta)$ can be given by

$$A(\zeta)=\frac{\varepsilon}{2}\left(\frac{\zeta^n}{a_0^n}+\frac{a_0^n}{\zeta^n}\right), \quad n\geqslant 3, \tag{105}$$

where

$$\varepsilon=\frac{1}{1+n^2}. \tag{106}$$

These inclusions are of practical interest since they provide good approximations to regular polygonal inclusions. For example, the inclusion shape with $n=3$ resembles a triangle, $n=4$ for quadrangle, $n=5$ for pentagon, and so forth. Although the polygonal elastic inclusion problems have been studied by some authors (see, e.g., Ru [12] and Gao [15]), the corresponding thermoplasticity problems are not available in the literature. According to Eq. (105), the coefficients C_m appearing in (39) are given by

$$C_m=\begin{cases}\varepsilon/2, & m=-n,n,\\ 0, & m\neq -n,n.\end{cases} \tag{107}$$

Applying (107) to the general expressions derived in Section 4, the temperature and stress functions for inclusion and matrix fields, respectively, can be immediately determined as follows:

$$g_1'(z)=\omega z-\varepsilon\frac{2k_2\beta}{k_1+k_2}a_0^{2-n}z^{n-1}, \tag{108}$$

$$\phi_1(z)=-\frac{\mu_1\mu_2}{\mu_2\kappa_1+\mu_1}(\beta_1\omega-\beta_2\alpha)z^2+N_{\varepsilon(n)}z^n, \tag{109}$$

$$\psi_1(z)=P_{\varepsilon(n-2)}z^{n-2}+P_{\varepsilon(n)}z^n, \tag{110}$$

for $z\in S_1$, and

$$g_2'(z)=\alpha z+\beta a_0^2 z^{-1}-\varepsilon\left(\frac{k_1-k_2}{k_1+k_2}\overline{\beta}a_0^n z^{1-n}-\beta a_0^{n+2}z^{-n-1}\right), \tag{111}$$

$$\phi_2(z)=-\frac{2\mu_2\beta_2}{1+\kappa_2}\beta a_0^2\ln z+L_{\varepsilon(n-2)}z^{-(n-2)}+L_{\varepsilon(n)}z^{-n}, \tag{112}$$

$$\psi_2(z)=-\frac{2\mu_2\beta_2}{1+\kappa_2}\overline{\beta}a_0^2\ln z-\left[\frac{\mu_1\mu_2}{\mu_2\kappa_1+\mu_1}(\beta_1\overline{\omega}-\beta_2\overline{\alpha})-\frac{2\mu_2\beta_2}{1+\kappa_2}\beta\right]a_0^4z^{-2}+M_{\varepsilon(n)}z^{-n}+M_{\varepsilon(n+2)}z^{-(n+2)}, \tag{113}$$

for $z\in S_2$, where

$$N_{\varepsilon(n)}=\varepsilon\frac{(\mu_2-\mu_1)A_1+4k_2\mu_1\mu_2\beta_1\beta(k_1+k_2)^{-1}n^{-1}}{\mu_1+\mu_2\kappa_1}a_0^{2-n}, \tag{114}$$

$$L_{\varepsilon(n)}=-\varepsilon\frac{(\mu_2-\mu_1)A_3+2n^{-1}\mu_1\mu_2\beta\beta_2}{\kappa_2\mu_1+\mu_2}a_0^{2+n}, \tag{115}$$

$$L_{\varepsilon(n-2)}=\varepsilon\frac{(\mu_2-\mu_1)(A_2-\overline{A_3})-2(2-n)^{-1}\mu_1\mu_2\beta_2\overline{\beta}(k_1-k_2)(k_1+k_2)^{-1}}{\kappa_2\mu_1+\mu_2}a_0^n, \tag{116}$$

$$M_{\varepsilon(n)}=\overline{N_{\varepsilon(n)}}a_0^{2n}+(n-2)L_{\varepsilon(n-2)}a_0^2+\varepsilon(\overline{A_1}-\overline{A_3})a_0^{2+n}, \tag{117}$$

$$M_{\varepsilon(n+2)}=nL_{\varepsilon(n)}a_0^2-\varepsilon(\overline{A_2}-A_3)a_0^{n+4}, \tag{118}$$

$$P_{\varepsilon(n)}=\overline{L_{\varepsilon(n)}}a_0^{-2n}+\varepsilon(\overline{A_3}-\overline{A_1})]a_0^{2-n}, \tag{119}$$

$$P_{\varepsilon(n-2)}=\overline{L_{\varepsilon(n-2)}}a_0^{-2n+4}+nN_{\varepsilon(n)}a_0^2+\varepsilon(\overline{A_2}-A_3)a_0^{4-n}. \tag{120}$$

Similar to the previous approach for the elliptical inclusion problem, the interfacial stresses can also be expressed from (97)–(99) by replacing (100)–(102) with

$$\frac{d\tilde{x}_1}{ds}=\frac{1}{|J|}\frac{d\tilde{x}_1}{d\theta}=\frac{1}{|J|}(-\sin\theta-\varepsilon\cos n\theta\sin\theta-n\varepsilon\sin n\theta\cos\theta), \tag{121}$$

$$\frac{d\tilde{x}_2}{ds}=\frac{1}{|J|}\frac{d\tilde{x}_2}{d\theta}=\frac{1}{|J|}(\cos\theta+\varepsilon\cos n\theta\cos\theta-n\varepsilon\sin n\theta\sin\theta), \tag{122}$$

$$J=\sqrt{[r'(\theta)]^2+r^2}=\sqrt{(-\varepsilon n\sin n\theta)^2+(1+\varepsilon\cos n\theta)^2}. \tag{123}$$

For the special case of an insulated and traction-free hole problem, the hoop stress along the polygonal hole boundary can be obtained from (98) by letting $k_1=0$ and $\mu_1=0$ as

$$\tau_{\eta\eta}=\mathrm{Re}\{-4\delta a_0^2\zeta^{-1}+\varepsilon[2\delta a_0^{2-n}\zeta^{n-1}+(8-4n)\overline{\delta}a_0^n\zeta^{1-n}+(2-4n)\delta a_0^{n+2}\zeta^{-1-n}]\}. \tag{124}$$

Comparisons of the dimensionless hoop stress between the present results (124) and the exact solutions derived by Kattis [16] for the case of a quadrilateral hole, $n=4$, and a hexagonal hole, $n=6$, are shown in Figs. 4 and 5, respectively. The agreement between these two results shows that our proposed methodology is satisfactory.

6 Conclusion

The general perturbation solutions for the two-dimensional thermoelastic problem with a nearly circular elastic inclusion of arbitrary shape existing in an infinite matrix are provided. The boundary of the inclusion is characterized by a Fourier series expansion that allows us to obtain an analytical solution for boundary value problems of an inclusion of arbitrary shape. Based on the method of analytical continuation and the boundary perturbation technique, first-order perturbation solutions are given explicitly for an inclusion of arbitrary shape. It is should be noted that higher-order perturbation solutions can also found by the present approach without difficulty. Besides, the corresponding inclusion problem under a point heat source with the strength q located at the point $z=z_0$ in the matrix can also be treated if one replaces the homogeneous solution in Eq. (40) with $-q/2\pi k_2\ln(z-z_0)$. As an application, two typical examples associated with elliptical and smooth polygonal elastic inclusions are solved completely and discussed in detail. Due to the intrinsic deficiency of the technique

of conformal mapping, the present method has its advantage of both efficiency and applicability to inclusions of arbitrary shape.

References

[1] Florence, A. L., and Goodier, J. N., 1959, ''Thermal Stress at Spherical Cavities and Circular Holes in Uniform Heat Flow,'' ASME J. Appl. Mech., **26**, pp. 293–294.

[2] Florence, A. L., and Goodier, J. N., 1960, ''Thermal Stress due to Disturbance of Uniform Heat Flow by an Insulated Ovaloid Hole,'' ASME J. Appl. Mech., **27**, pp. 635–639.

[3] Chen, W. T., 1967, ''Plane Thermal Stress at an Insulated Hole Under Uniform Heat Flow in an Orthotropic Medium,'' ASME J. Appl. Mech., **34**, pp. 133–136.

[4] Green, A. E., and Zerna, W., 1954, *Theoretical Elasticity*, Oxford University Press, London.

[5] Hwu, C., 1990, ''Thermal Stresses in an Anisotropic Plate Disturbed by an Insulated Elliptic Hole or Crack,'' ASME J. Appl. Mech., **57**, pp. 916–922.

[6] Stroh, A. N., 1958, ''Dislocations and Cracks in Anisotropic Elasticity,'' Philos. Mag., **7**, pp. 625–646.

[7] Kattis, M. A., and Mequid, S. A., 1995, ''Two-Phase Potentials for the Treatment of an Elastic Inclusion in Plane Thermoelasticity,'' ASME J. Appl. Mech., **62**, pp. 7–12.

[8] Chao, C. K., and Shen, M. H., 1998, ''Thermal Stresses in a Generally Anisotropic Body With an Elliptic Inclusion Subject to Uniform Heat Flow,'' ASME J. Appl. Mech., **65**, pp. 51–58.

[9] Hwu, C., and Yen, W. J., 1993, ''On the Anisotropic Elastic Inclusions in Plane Elastostatics,'' ASME J. Appl. Mech., **60**, pp. 626–632.

[10] Jaswon, M. A., and Bhargave, R. D., 1961, ''Two-Dimensional Elastic Inclusion Problem,'' Proc. Cambridge Philos. Soc., **57**, pp. 669–680.

[11] Sendeckyi, G. P., 1970, ''Elastic Inclusion Problems in Plane Elastostatics,'' Int. J. Solids Struct., **6**, No. 6, pp. 1535–1543.

[12] Ru, C. Q., 1999, ''Analytic Solution for Eshelby's Problem of an Inclusion of Arbitrary Shape in a Plane or Half-Plane,'' ASME J. Appl. Mech., **66**, pp. 315–322.

[13] Ru, C. Q., 1998, ''Interface Design of Neutral Elastic Inclusions,'' Int. J. Solids Struct., **35**, pp. 557–572.

[14] Benveniste, Y., and Miloh, T., 1999, ''Neutral Inhomogeneities in Conduction Phenomena,'' J. Mech. Phys. Solids, **47**, pp. 1873–1892.

[15] Gao, H., 1991, ''A Boundary Perturbation Analysis for Elastic Inclusions and Interfaces,'' Int. J. Solids Struct., **28**, No. 6, pp. 703–725.

[16] Kattis, M. A., 1991, ''Thermoelastic Plane Problems With Curvilinear Boundaries,'' Acta Mech., **87**, pp. 93–103.

[17] Bogdanoff, J. L., 1954, ''Note on Thermal Stress,'' ASME J. Appl. Mech., **21**, p. 88.

[18] Chao, C. K., and Shen, M. H., 1997, ''On Bonded Circular Inclusions in Plane Thermoelasticity,'' ASME J. Appl. Mech., **64**, pp. 1000–1004.

A. Yavari[1]
Graduate Research Assistant
Assoc. Mem. ASME

S. Sarkani
Professor

E. T. Moyer, Jr.
Adjunct Associate Professor

School of Engineering and Applied Science,
The George Washington University,
Washington, DC 20052

On Fractal Cracks in Micropolar Elastic Solids

In this paper we review the fracture mechanics of smooth cracks in micropolar (Cosserat) elastic solids. Griffith's fracture theory is generalized for cracks in micropolar solids and shown to have two possible forms. The effect of fractality of fracture surfaces on the powers of stress and couple-stress singularity is studied. We obtain the orders of stress and couple-stress singularities at the tip of a fractal crack in a micropolar solid using dimensional analysis and an asymptotic method that we call "method of crack-effect zone." It is shown that orders of stress and couple-stress singularities are equal to the order of stress singularity at the tip of the same fractal crack in a classical solid.
[DOI: 10.1115/1.1409258]

1 Introduction

Fractal geometry, which has been argued to be a better geometry for modeling natural objects than Euclidean geometry, was introduced by Mandelbrot [1,2]. The term ''fractal'' was coined by Mandelbrot [2] from the Latin verb *frangere*, ''to break,'' and the corresponding adjective *fractus*, ''fragmented and irregular.'' Fractal geometry has found applications in many fields of science and engineering in recent years. So far fractal geometry's major applications to solid mechanics problems are in contact mechanics and fracture mechanics. Fractal fracture mechanics is a nonclassical fracture mechanics in which cracks are assumed to be fractal curves (surfaces) (Cherepanov et al. [3], Balankin [4]). In classical fracture mechanics it is assumed that cracks are rectifiable curves (surfaces), i.e., curves (surfaces) with finite lengths (areas). Cracks are modeled by smooth curves (surfaces) with probably a finite number of kinks. These simplifying assumptions make fracture mechanics problems mathematically tractable.

Mandelbort et al. [5] experimentally showed that fracture surfaces of steel are fractals. Since that pioneering work many other experimental studies have been done (for example, Brown and Scholz [6]; Power and Tullis [7]; Saouma et al. [8]; Saouma and Barton [9]; Wong et al. [10]). Now we know that cracks can be modeled by fractals in a wide (but finite) range of length scales. A number of theoretical studies have been conducted to date. Mosolov [11] and Gol'dshteĭn and Mosolov [12,13] studied the singularity of stresses at the tip of a mode I self-similar fractal crack showing that the power of stress singularity is a linear function of fractal dimension of the crack. Yavari et al. [14] calculated the orders of stress singularity for mode I, II, and III fractal cracks. Yavari [15], Yavari et al. [16], and Balankin [4] studied HRR singularity for self-similar and self-affine fractal cracks.

Mosolov [17] and Balankin [4] investigated the path independence of *J*-integral for fractal cracks and modified the *J*-integral for fractal cracks. They argued that the modified *J*-integrals are path-independent. This problem was later discussed in Yavari et al. [16]. They mentioned that a fractal *J*-integral should be equal to the potential energy release per unit of a fractal measure. They explained that the modified *J*-integrals defined by Mosolov and by Balankin are only locally path-independent and have no physical meaning. Crack growth in compression was explained by Mosolov and Borodich [18], Mosolov [19], and Balankin [4].

Yavari [15] and Yavari et al. [14] introduced a new mode of fracture in fractal fracture mechanics and called it ''the fourth mode'' or ''the axial mode.'' They pointed out that the existence of this new mode of fracture could make some single-mode problems of classical fracture mechanics, mixed-mode problems in fractal fracture mechanics. Later, Yavari et al. [16] showed that there are actually three new fractal modes. Xie [20] studied crack branching using a fractal model. Xie and Sanderson [21] explained a paradox in dynamic fracture mechanics using their fractal model. Borodich [22,23] realized that Griffith's criterion must be modified for fractal cracks. He showed that in the modified criterion, the specific surface energy must be defined per unit of a fractal measure (not length or area) of fractal crack growth. Yavari [24] generalized Barenblatt's cohesive fracture theory and developed a fractal cohesive fracture theory.

To our best knowledge, there is no investigation into fractal cracks in micropolar (Cosserat) solids. This paper aims to explore some interesting problems of micropolar fractal fracture mechanics. In Section 2, micropolar elasticity is reviewed and its basic concepts and definitions are explained. Section 3 discusses fracture mechanics of rectilinear cracks in micropolar solids. The effects of couple-stresses in fracture mechanics are reviewed and Griffith's criterion is generalized for both smooth and fractal cracks in micropolar solids in Section 4. Section 5 studies self-similar and self-affine fractal cracks in micropolar solids. Using dimensional analysis and the method of crack-effect zone, it is shown that stresses and couple-stresses at the tip of a fractal crack in a micropolar solid have equal orders of singularity. The Appendix presents some basic definitions and techniques of fractal geometry that are directly relevant to our investigation.

2 Micropolar Elasticity

This section presents a brief introduction to generalized continuum theories and their history. Here we discuss only those aspects of micropolar elasticity theory that are necessary for our investigation of fractal cracks in a micropolar solid. A literature review for fracture mechanics of rectilinear cracks in micropolar solids will be given in the next section.

In classical continuum mechanics, at each point only translational degrees-of-freedom $u_i\,(i=1,2,3)$ are considered and it is assumed that the interaction between two material points along an arbitrary surface S is completely described by a stress vector $\boldsymbol{\sigma}$ defined on S. These assumptions lead to a mathematically consistent theory of continuum mechanics. Experience has shown that most analytical solutions obtained in the framework of classical continuum mechanics agree very well with the experimental re-

[1]Currently Graduate Research Assistant, Graduate Aeronautical Laboratories, California Institute of Technology, Pasadena, CA 91125.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, May 5, 2000; final revision, January 16, 2001. Associate Editor: W. J. Drugan. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

sults. To data all engineering designs are based on the use of classical continuum mechanics and sometimes even with some more simplifying assumptions.

The curiosity of some distinguished researchers led them to question the above-mentioned hypotheses and to develop generalized continuum theories. It was clear for them that considering only translational degrees-of-freedom might not be enough for continua with microstructure (see [25–37]).

In the original Cosserat brothers' formulation ([26]), rotations $\phi_i\,(i=1,2,3)$ were considered to be independent of displacement components $u_i\,(i=1,2,3)$. However, later most analytical solutions were reported for a special case that is now known as couple-stress theory or constrained Cosserat theory. In couple-stress theory, microrotations are assumed to be equal to macrorotations, i.e., $\phi_i=1/2\varepsilon_{ijk}u_{k,j}$. This is the theory that was developed independently by Grioli [28], by Aero and Kuvshinskii [31], and by Mindlin and Tiersten [32]. Eringen and his co-workers elaborately studied the theory of Cosserat continua and again assumed that microrotations are independent of displacement components. Eringen [36,37] renamed the Cosserat continuum theory and called it micropolar continuum theory. Cowin [38–40] discovered a continuous transition from couple-stress theory to micropolar theory by introducing a coupling number $N\,(0\leqslant N\leqslant 1)$, where $N=1$ corresponds to the couple-stress theory, $N=0$ corresponds to the classical theory, and N between zero and one $(0<N<1)$ corresponds to the micropolar theory. It is known that in couple-stress theory of elasticity two new constants appear and one of them, l has the dimension of length and is called the characteristic length. On the other hand, in micropolar elasticity there are four new material constants and two of them, l_t and l_b have dimensions of length and are called characteristic lengths in torsion and bending, respectively. This means that in generalized continuum theories there is at least one internal length scale and therefore these theories should be able to analytically predict size effects.

Several authors investigated the effects of couple-stresses in different problems of solid mechanics such as stress concentration in the presence of holes and inclusions and the change of size effect in rigidity of different structural members (see [41–62]). Recently, there have been some investigations into strain gradient plasticity (see [63] and references therein). These theories seem to be promising in design of very small structures.

As was mentioned at the beginning of this section, generalized continuum theories attracted theoreticians because of their beauty. To date these theories have not been applied to practical problems. Here we have an example of a field in which experimental studies are far behind the theory. There are several experimental investigations into the mechanical properties of micropolar elastic materials. What we have at this time are just some ranges of these material constants (Schijve [64] and Lakes [60]). So far we have only some qualitative sense of the influences of couple stresses. We are hopeful that future advances in experimental mechanics will make these elegant theories applicable to real engineering problems.

It is worth mentioning that there is a recent interest in generalized continuum theories because of the superiority they have in localization analyses. These studies are beyond the scope of this section and will not be mentioned here. Now we present the basic concepts, definitions, and balance equations of the theory of micropolar elasticity. Here we mainly follow Eringen [37].

In a continuous medium with microstructure each material element contains several micromaterial elements. In micropolar continuum mechanics only microrotations are considered for micro-elements. Therefore, for each material point, in addition to the three displacements, three microrotations are considered. Microrotations are assumed to be different from macrorotations. Displacement components are denoted by u_i, microrotations by ϕ_i, and macrorotations by r_i. Macrorotations have the following relations with displacements:

$$r_i=\frac{1}{2}\varepsilon_{ijk}u_{k,j} \tag{1}$$

where ε_{ijk} is the permutation symbol. Macrostrains e_{ij} and microstrains ε_{ij} are defined as

$$e_{ij}=\frac{1}{2}(u_{i,j}+u_{j,i}) \tag{2a}$$

$$\varepsilon_{ij}=e_{ij}+e_{ijk}(r_k-\phi_k). \tag{2b}$$

Curvature tensor is defined by

$$\chi_{ij}=\phi_{j,i}. \tag{3}$$

As a consequence of the assumption that each point has six degrees-of-freedom, in a micropolar continuum both stresses and couple-stresses exist and Cauchy's theorem holds for them, i.e.,

$$\sigma_i=\sigma_{ji}n_j \tag{4a}$$

$$m_i=m_{ji}n_j \tag{4b}$$

where σ_i and m_i are components of stress and couple-stress vectors, respectively, n_i is the unit normal vector to an arbitrary surface S, and σ_{ij} and m_{ij} are respective stress and couple-stress tensors. Stress and couple-stress tensors are in general asymmetric. The equilibrium equations are

$$\sigma_{ji,j}=0 \tag{5a}$$

$$m_{ji,j}+\varepsilon_{ijk}\sigma_{j,k}=0. \tag{5b}$$

For a centrosymmetric isotropic micropolar material the stress-strain relations are

$$\sigma_{ij}=\lambda e_{kk}\delta_{ij}+(2\mu+\kappa)e_{ij}+\kappa\varepsilon_{ijk}(r_k-\phi_k) \tag{6a}$$

$$m_{ij}=\alpha\phi_{k,k}\delta_{ij}+\beta\phi_{i,j}+\gamma\phi_{j,i} \tag{6b}$$

where λ and μ are the classical Lamé constants and α, β, γ, and κ are new micropolar constants with the following dimensions:

$$[\alpha]=[\beta]=[\gamma]=F=\frac{M}{L}\ \text{ and }\ [\kappa]=\frac{F}{L^2}=\frac{M}{L^3} \tag{7}$$

where F, M, and L are dimensions of force, moment, and length, respectively. The strain energy density has the following form:

$$W=\frac{1}{2}(\sigma_{ij}\varepsilon_{ij}+m_{ij}\chi_{ij})=\frac{1}{2}[\lambda e_{kk}e_{mm}+(2\mu+\kappa)e_{ij}e_{ij}]+\kappa(r_k-\phi_k)(r_k-\phi_k)+\frac{1}{2}(\alpha\phi_{k,k}\phi_{m,m}+\beta\phi_{i,j}\phi_{j,i}+\gamma\phi_{i,j}\phi_{i,j}). \tag{8}$$

The following technical elastic constants have clearer physical meanings ([59]):

$$E=\frac{(2\mu+\kappa)(3\lambda+2\mu+\kappa)}{2\lambda+2\mu+\kappa},\quad G=\frac{2\mu+\kappa}{2},\quad \nu=\frac{\lambda}{2\lambda+2\mu+\kappa} \tag{9a}$$

$$l_t=\sqrt{\frac{\beta+\gamma}{2\mu+\kappa}},\quad l_b=\sqrt{\frac{\gamma}{2(2\mu+\kappa)}} \tag{9b}$$

$$N=\sqrt{\frac{\kappa}{2(\mu+\kappa)}},\quad \psi=\frac{\beta+\gamma}{\alpha+\beta+\gamma} \tag{9c}$$

where E, G, ν, l_t, l_b, N, and ψ are Young's modulus, Poisson's ratio, the characteristic length in torsion, the characteristic length in bending, coupling number, and polar ratio, respectively. These constants have the following dimensions:

$$[E]=[G]=FL^{-2},\quad [\nu]=1$$
$$[l_t]=[l_b]=L \tag{10}$$

$$[N]=[\psi]=1.$$

It is seen that two internal characteristic lengths exist. Therefore, this theory is capable of analytically predicting size effects. It is worth mentioning that these characteristic lengths appear in the stress field solutions even for force control loading conditions. Therefore, in dimensional analysis formulations these characteristic lengths must be taken into account.

3 Fracture Mechanics of Smooth Cracks in Micropolar Solids

In this section, fracture mechanics of smooth cracks in a micropolar solid is reviewed. Here, the effects of couple-stresses on the stress distribution around the tip of a smooth crack are discussed. In 1960s and 1970s when generalized continuum theories were rediscovered and elaborately developed, several researchers became interested in examining the influence of couple-stresses in problems in which classical theory predicts infinite stresses. One such problem with great practical importance was the stress distribution near the tip of a smooth crack. It was known that stresses and strains around the tip of a crack are unbounded and have an $r^{-1/2}$ singularity. Researchers were hopeful not to see this pathological problem in higher-order continuum theories. Unfortunately, higher-order continuum theories could not solve this pathological problem; both stresses and couple-stresses were observed to be unbounded at the crack tip. There is a limited number of investigations in micropolar fracture mechanics, which will be reviewed in this section.

The first investigation into fracture mechanics of smooth cracks in Cosserat continua was performed by Sternberg and Muki [65]. They solved the problem of an infinite two-dimensional plane-strain linear couple-stress medium with a finite crack under a uniform tensile stress state perpendicular to the crack axis at infinity. They showed that both stresses and couple-stresses have an $r^{-1/2}$ singularity at the crack tip. They observed that couple-stresses only change the angular variation of stresses around the crack tip; the form of the radial variation of stresses remains unchanged. They found the following asymptotic expressions for stresses and couple-stresses:

$$\sigma_{xx}(r,\theta)=-(1-2\nu)\frac{K_I}{\sqrt{2r}}\left[\cos\frac{\theta}{2}-\frac{1}{2}\sin\theta\sin\frac{3\theta}{2}\right]+O(r^0) \tag{11a}$$

$$\sigma_{yy}(r,\theta)=\frac{K_I}{\sqrt{2r}}\left[(3-2\nu)\cos\frac{\theta}{2}-\frac{1}{2}(1-2\nu)\sin\theta\sin\frac{3\theta}{2}\right]+O(r^0) \tag{11b}$$

$$\sigma_{xy}(r,\theta)=-\frac{K_I}{\sqrt{2r}}\left[4(1-\nu)\sin\frac{\theta}{2}+\frac{1}{2}(1-2\nu)\times\sin\theta\cos\frac{3\theta}{2}\right]+O(r^0) \tag{11c}$$

$$\sigma_{yx}(r,\theta)=-(1-2\nu)\frac{K_I}{\sqrt{2r}}\left[\frac{1}{2}\sin\theta\cos\frac{3\theta}{2}\right]+O(r^0) \tag{11d}$$

and

$$m_{xz}(r,\theta)=-\frac{\hat{K}_l}{\sqrt{2r}}\left[\frac{a}{2}\sin\frac{\theta}{2}\right]+O(r^0) \tag{12a}$$

$$m_{xz}(r,\theta)=\frac{\hat{K}_l}{\sqrt{2r}}\left[\frac{a}{2}\cos\frac{\theta}{2}\right]+O(r^0) \tag{12b}$$

where σ_{xx}, σ_{yy}, σ_{xy}, σ_{yx} are (force-) stresses and m_{xz} and m_{yz} are couple-stresses and

$$K_I=f\left(\frac{l}{a},\nu\right)\sigma^\infty\sqrt{a}\ \text{ and }\ \hat{K}_l=\hat{f}\left(\frac{l}{a},\nu\right)\sigma^\infty\sqrt{a} \tag{13}$$

where K_I and $\hat{K}_l$ are stress- and couple-stress intensity factors, ν is Poisson's ratio, l is the characteristic length of the couple-stress material, and $2a$ is the crack length. Later Sih and Liebowitz [66] found the asymptotic expressions of the displacement and rotation components as shown below:

$$u_r(r,\theta)=\frac{K_I\sqrt{2r}}{8\mu}\left[3(1-2\nu)\cos\frac{\theta}{2}-(7-6\nu)\cos\frac{3\theta}{2}\right]+O(r)$$

$$u_\theta(r,\theta)=\frac{K_I\sqrt{2r}}{8\mu}\left[-(1-2\nu)\sin\frac{\theta}{2}+(7-6\nu)\sin\frac{3\theta}{2}\right]+O(r) \tag{14}$$

$$\omega_z(r,\theta)=\frac{\hat{K}_I\sqrt{2r}}{8\mu l^2}\left[a\sin\frac{\theta}{2}\right]+O(r).$$

For a crack in a Cosserat continuum, strain energy release may be calculated as

$$G=\lim_{\delta a\to 0}\frac{1}{\delta a}\int_0^{\delta a}[\sigma_{yy}(\delta a-\xi,0)u_y(\xi,\pi)+m_{yz}(\delta a-\xi,0)\omega_z(\xi,\pi)]d\xi. \tag{15}$$

Using the above formula, Sih and Liebowitz [66] found the strain energy release rate.

$$G=\frac{\pi}{2\mu}\left[(1-\nu)(3-2\nu)K_I^2+\frac{1}{16}\left(\frac{a}{l}\right)^2\hat{K}_I^2\right] \tag{16}$$

There are some other interesting investigations into fracture mechanics of cracks in micropolar solids (see [67–73]). Now it is known that the classical theory underestimates the value of K_I and overestimates the energy release rate G.

Another interesting investigation into micropolar fracture mechanics was conducted by Atkinson and Leppington [74]. They analyzed two problems: (1) a semi-infinite crack under an internal stress acting on the crack faces and (2) a finite crack in an infinite solid under a uniform stress at infinity. They solved the second problem only for cases in which l/a is very small ($l/a\ll 1$). They showed that both stresses and couple-stresses at the tip of a crack in a couple-stress or micropolar medium have an $r^{-1/2}$ singularity. They also demonstrated that the angular variations of stresses and couple-stresses in couple-stress and micropolar continua are a little different but have a similar form. Atkinson and Leppington defined the J-integral for both couple-stress and micropolar theories and showed that J-integral is path-independent. Recently, Lubarda and Markenscoff [75] studied some conservation integrals for linear couple-stress elasticity.

4 Micropolar Griffith's Criterion

For finding the orders of stress and couple-stress singularity at the tip of a fractal crack, we utilize an energy approach. The fractal crack is in equilibrium and hence the virtual work of all forces in a virtual displacement, which is an infinitesimal crack growth, is zero. For a cracked body, the principle of virtual work must be modified to take into account the work done in a crack propagation and strain energy release due to a crack growth. Griffith's [76,77] criterion is actually a modified energy balance for cracked bodies. In this section we generalized Griffith's theory for smooth and fractal cracks in micropolar solids.

4.1 Griffith's Theory for a Smooth Crack in a Micropolar Solid. When a crack propagates, new free surfaces are created. For creating these new free surfaces some amount of surface energy is needed to overcome the cohesive forces. This amount of energy is provided by an equal amount of strain energy release. This is Griffith's criterion [76,77], which was originally stated for

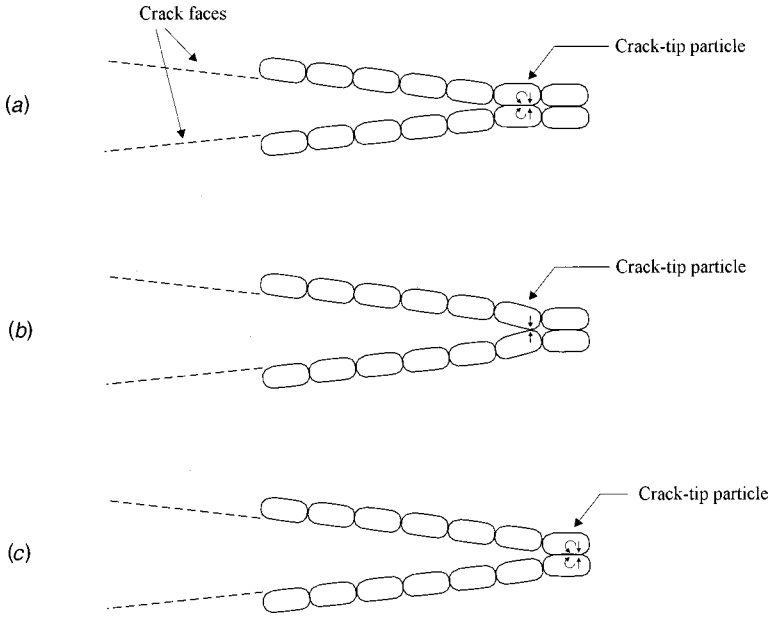


Fig. 1 Mechanism of crack propagation in a micropolar continuum: (*a*) crack-tip particles withstand rotation and separation, (*b*) the first step in crack propagation—crack-tip particles rotate with respect to each other, and (*c*) the second step in crack propagation—crack-tip particles move apart and neighboring particles become the next crack-tip particles

a rectilinear crack in a classical continuum. Mosolov [11] used this criterion for fractal cracks assuming that the specific surface energy per unit length remains unchanged and only the length of the crack increases in the case of a fractal crack. Later, Borodich [22,23] noticed that Griffith's criterion must be modified and in the modified criterion the specific surface energy must be defined per unit of a fractal measure. To our best knowledge, there is no discussion on Griffith's criterion for cracks in micropolar solids. This theory can be easily generalized for smooth and fractal cracks in a micropolar solid, as we show below.

In a micropolar continuum each material point can rotate and translate independently. Now suppose that there is a finite crack of length 2a in a micropolar solid. Figure 1(*a*) shows a crack and some particles (material points) on the crack surfaces. When the crack propagates, crack-tip particles separate from each other. Unlike a crack in a classical solid, this separation of crack-tip particles is a two-step process as shown in Figs. 1(*b*) and 1(*c*). In the first step crack-tip particles rotate with respect to each other but do not move, i.e.,

$$\Delta\phi = \phi_2 - \phi_1 \neq 0 \quad \text{and} \quad \Delta u = u_2 - u_1 = 0. \tag{17}$$

In the next step, crack-tip particles move apart but do not rotate, i.e.,

$$\Delta u = u_2 - u_1 \neq 0 \quad \text{and} \quad \Delta\phi = \phi_2 - \phi_1 = 0. \tag{18}$$

After this step, these particles are no longer crack-tip particles; they belong to two free surfaces (see Fig. 1(*c*)). Obviously, the surface energy δU_s needed for creating the new free surfaces has two parts, δU_s^{ϕ} and δU_s^{u}, where δU_s^{ϕ} is the surface energy spent on rotating particles in the path of crack growth and δU_s^{u} is the surface energy spent on separating these particles from each other. Figure 2 shows a crack and the dashed line is the crack propagation path. Crack-tip particles on the path of crack propagation are shown in this figure. Similar to the surface energy release rate, strain energy release rate is composed of two parts: stress part δU_e^{σ}, and couple-stress part δU_e^{m}.

Griffith's criterion for a crack in a micropolar solid may be stated in two different forms, depending on whether the effects of stresses and couple-stresses are considered uncoupled or coupled.

(I) Uncoupled Micropolar Griffith's Criterion. This form of Griffith's criterion states that a crack propagates by an amount δa if the following conditions are satisfied simultaneously:

$$\delta U_e^{\sigma} = \delta U_s^{u} 2t\gamma_u \delta a \tag{19a}$$

$$\delta U_e^{m} = \delta U_s^{\phi} = 2t\gamma_{\phi}\delta a \tag{19b}$$

where t, γ_u, and γ_{ϕ} are thickness, displacement, and rotation specific surface energies, respectively. Dimensions of these two surface energies are $[\gamma_u] = [\gamma_{\phi}] = FL^{-1}$.

(II) Coupled Micropolar Griffith's Criterion. In this form of the Griffith's criterion effects of stresses and couple-stresses are

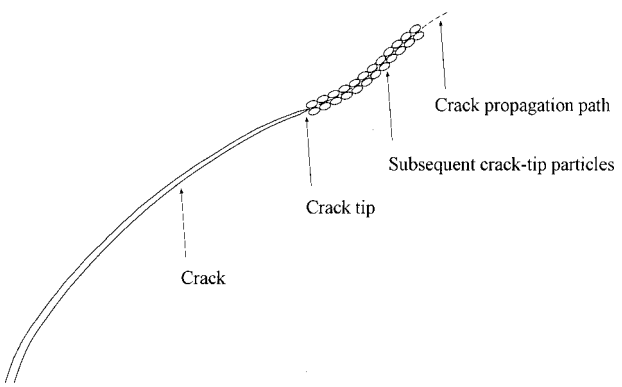


Fig. 2 A crack in a micropolar solid and its propagation path. The particles shown are the particles on the subsequent free surfaces.

assumed to be coupled. Coupled micropolar Griffith's criterion states that the crack propagates by an amount δa if

$$\delta U_e = \delta(U_e^\sigma + U_e^m) = \delta U_s = 2t\gamma_m \delta a = 2t(\gamma_u + \gamma_\phi)\delta a. \quad (20)$$

It should be noted that the micropolar specific surface energy γ_m is generally different from the classical specific energy γ. Obviously, if (19*a*) and (19*b*) are satisfied, (20) is automatically satisfied. In other words, the uncoupled criterion is stronger than the coupled criterion.

4.2 Griffith's Criterion for a Fractal Crack in a Micropolar Solid. For a smooth crack, surface energy required for crack propagation is proportional to the length (area) of the newly created free surfaces. In the case of a fractal crack the true length (area) of new free surfaces should be considered. Because the true length (area) of a fractal curve (surface) is infinity, a fractal measure should be utilized. The surface energy required to create a fractal crack in a classical solid is

$$U_s = 2t\gamma_f(D)m_D \quad (21)$$

where t is the plate thickness, $\gamma_f = \gamma_f(D)$ is the specific surface energy per unit of a fractal measure, and m_D is the corresponding fractal measure and is proportional to a^D (see the Appendix). Specific surface energy per unit of a fractal measure was introduced by Borodich [22,23] and has the dimension $[\gamma_f] = FL^{-D}$, where F and L are dimensions of force and length, respectively. There are two important issues arising from Borodich's generalization of Griffith's criterion that should be explained: (1) It should be noted that γ_f is not a material property. In general, it is possible to have cracks with different fractal dimensions in the same material. Therefore, in Eq. (21) γ_f cannot be a material property; it depends on both the material and the fractal dimensions of the fractal crack. (2) ''Fractal measure'' is an ambiguous term; there are different definitions of dimension and consequently these different dimensions have different corresponding measures. For self-similar fractals all different dimensions have the same value and hence the corresponding measures they define are identical. Therefore, for self-similar fractals ''fractal measure'' is not an ambiguous term. However, this is not the case for self-affine fractals; different definitions of dimension give completely different dimensions for the same self-affine fractal set. Obviously, the relevant fractal dimension for calculating the surface energy of a fractal crack is the divider (latent) fractal dimension. Therefore, the specific surface energy should be defined per divider fractal measure, although it can be defined for other fractal measures as well.

For a fractal crack in a micropolar solid, Griffith's criterion again has uncoupled and coupled forms and only the surface energies should be modified as

$$\delta U_s^u = 2t\gamma_u^f(D_{D_D})\delta m_{D_D} \quad \text{and} \quad \delta U_s^\phi = 2t\gamma_\phi^f(D_D)\delta m_{D_D} \quad (22a)$$

$$\delta U_s = \delta U_s^u + \delta U_s^\phi = 2t[\gamma_u^f(D_D) + \gamma_\phi^f(D_D)]\delta m_{D_D} \quad (22b)$$

where γ_u^f and γ_ϕ^f are fractal specific surface energies per unit of latent fractal measure and m_{D_D} is the latent fractal measure. Thus we have the following two forms of fractal micropolar Griffith's criterion.

(I) Uncoupled Fractal Micropolar Griffith's Criterion. A fractal crack with divider fractal dimension D_D propagates by an amount δm_{D_D} if the following two conditions are satisfied simultaneously:

$$\delta U_e^\sigma = \delta U_s^u = 2t\gamma_u^f \delta m_{D_D} \quad (23a)$$

$$\delta U_e^m = \delta U_s^\phi = 2t\gamma_\phi^f \delta m_{D_D}. \quad (23b)$$

(II) Coupled Fractal Micropolar Griffith's Criterion. A fractal crack with divider fractal dimension D_D propagates by an amount δm_{D_D} if the following condition is satisfied:

$$\delta U_e = \delta(U_e^\sigma + U_e^m) = \delta U_s = 2t\gamma_m^f \delta m_{D_D} = 2t(\gamma_u^f + \gamma_\phi^f)\delta m_{D_D}. \quad (24)$$

In the next section we use both forms of micropolar Griffith's criterion for calculating the orders of stress and couple-stress singularity at the tip of a fractal crack. We will show that both criteria give equal orders of stress and couple-stress singularity.

5 Fractal Cracks in Micropolar Elastic Solids

In this section radial variations of stresses and couple-stresses around the tip of a fractal crack in a micropolar solid are investigated. To the best of our knowledge, there is no investigation into this problem in the literature. Without loss of generality, a mode I problem is solved. Consider an infinite medium made of a micropolar material with a finite crack of nominal length $2a$. It is assumed that the cracked solid is under a uniform tensile stress σ^∞ perpendicular to the crack axis applied at infinity (see Fig. 3(*a*)).

One major difference between this problem and the similar problem of a fractal crack in a classical solid is that a micropolar material has two internal characteristic length scales, l_b and l_t. Here l_b and l_t are characteristic lengths in bending and torsion, respectively. On the other hand, a couple-stress material has only one characteristic length l. For a micropolar material in a two-dimensional problem only one of the characteristic lengths appears in the equilibrium equations. Therefore, it is assumed that the medium has a characteristic length and it is denoted by l. It is known that even for a force control loading this characteristic length appears in the stress solutions in the form of l/a, where ''a'' is a geometric characteristic length of the problem, for example hole radius or crack length.

The method of crack-effect zone, which was introduced by Yavari et al. [16], is utilized. When the system shown in Fig. 3(*a*) is uncracked, only one of stress components is nonzero and has a uniform distribution; all other stresses and all couple-stresses are identically zero. When the crack is formed, stresses and couple-stresses are perturbed. This stress perturbation is significant only in a finite zone around the crack. For the cracked system all stresses and couple-stresses are nonzero in the crack-effect zone. The crack-effect zone may be covered by a disk $\Re_c$, as shown in Fig. 3(*a*). We assume that the micropolar material is centrosymmetric isotropic and homogeneous. The strain energy of the system may be written as

$$U_e = U_e^\sigma + U_e^m = \int_\Re \left(\frac{1}{2}\sigma_{ij}\varepsilon_{ij} + \frac{1}{2}m_{ij}\chi_{ij}\right)dV \quad (25)$$

where σ_{ij}, m_{ij}, ε_{ij}, and χ_{ij} are stresses, couple-stresses, strains, and curvatures, respectively. The strain energy can be decomposed into two parts as follows:

$$U_e = \int_{\Re - \Re_c} \left(\frac{1}{2}\sigma_{ij}\varepsilon_{ij} + \frac{1}{2}m_{ij}\chi_{ij}\right)dV + \int_{\Re_c} \left(\frac{1}{2}\sigma_{ij}\varepsilon_{ij} + \frac{1}{2}m_{ij}\chi_{ij}\right)dV. \quad (26)$$

When the crack propagates by an infinitesimal amount δa, the change of the strain energy in $\Re_c$ is dominant, hence

$$\delta U_e \cong \delta \int_{\Re_c} \left(\frac{1}{2}\sigma_{ij}\varepsilon_{ij} + \frac{1}{2}m_{ij}\chi_{ij}\right)dV. \quad (27)$$

For a centrosymmetric material stress-strain and couple-stress-curvature relations are uncoupled, i.e., stresses are not functions of curvatures and couple-stresses are not functions of strains. Therefore, the constitutive equations may be written as

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl} \quad \text{and} \quad m_{ij} = \hat{C}_{ijkl}\chi_{kl} \quad (28)$$

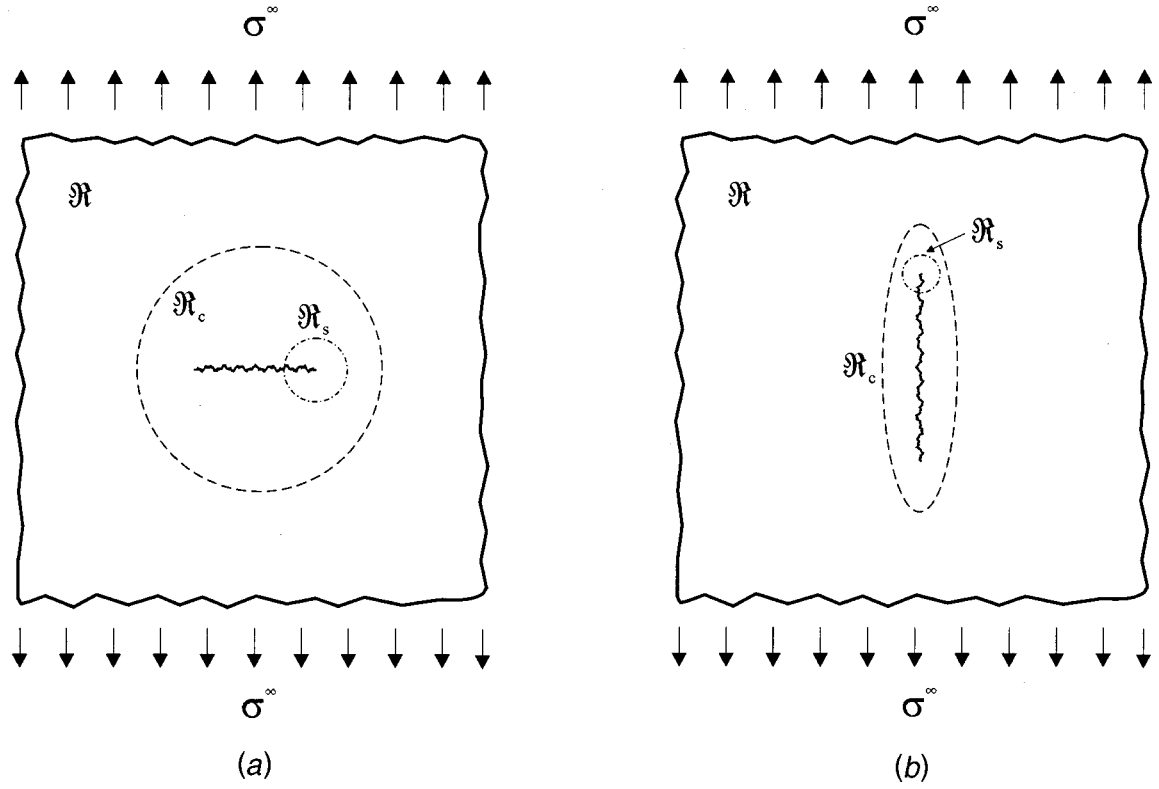


Fig. 3 (*a*) A two-dimensional micropolar solid with a finite fractal crack perpendicular to the applied stresses, (*b*) an infinite micropolar solid with a finite fractal crack parallel to the applied stresses

where C_{ijkl} and $\hat{C}_{ijkl}$ are fourth-order tensors and are mechanical properties of the material. The following asymptotic stresses and couple-stresses are assumed at the crack tip

$$\sigma_{ij}(r,\theta)=K_I r^{-\alpha_1} f_{ij}\left(\theta,\nu,\frac{l}{a},H\right) \tag{29a}$$

$$m_{ij}(r,\theta)=\hat{K}_I r^{-\alpha_2} \hat{f}_{ij}\left(\theta,\nu,\frac{l}{a},H\right) \tag{29b}$$

where K_I^f and $\hat{K}_I^f$ are fractal stress and couple-stress intensity factors, respectively, and H is the Hurst exponent (see the Appendix). We will calculate α_1 and α_2 using the method of crack-effect zone. The above asymptotic stresses and couple-stresses are dominant for $r \leqslant r_{s_1}$ and $r \leqslant r_{s_2}$, respectively. Therefore, Eqs. (29*a*) and (29*b*) are valid in a disk $\Re_s$ with radius $r_s=\min(r_{s_1},r_{s_2})$. Here, the method of crack-effect zone should be used very carefully. Because the change of U_e in $\Re_s$ is dominant, the change of strain energy may be expressed as

$$\delta U_e=\delta\int_{\Re_c-\Re_s}\left(\frac{1}{2}\sigma_{ij}\varepsilon_{ij}+\frac{1}{2}m_{ij}\chi_{ij}\right)dV+\delta\int_{\Re_s}\left(\frac{1}{2}\sigma_{ij}\varepsilon_{ij}+\frac{1}{2}m_{ij}\chi_{ij}\right)dV\cong\delta\int_{\Re_s}\left(\frac{1}{2}\sigma_{ij}\varepsilon_{ij}+\frac{1}{2}m_{ij}\chi_{ij}\right)dV. \tag{30}$$

From (29) and (30) we have

$$\delta U_e=\delta U_e^{\sigma}+\delta U_e^{m} \tag{31a}$$

$$\delta U_e^{\sigma}\propto\delta(r_s^{-\alpha_1}r_s^{-\alpha_1}r_s^2) \quad \text{and} \quad \delta U_e^{m}\propto\delta(r_s^{-\alpha_2}r_s^{-\alpha_2}r_s^2) \tag{31b}$$

where "$\propto$" means "proportional to." For a rectilinear or fractal crack in a micropolar continuum r_s is not necessarily proportional to "*a*." Here r_s is a function of *a*, *l*, ν, and *H*, i.e., $r_s=f(a,l,\nu,H)$. Using Buckingham's ([78,79]) Π theorem we must have

$$\frac{r_s}{a}=\hat{\Phi}\left(\frac{l}{a},\nu,H\right) \quad \text{or} \quad r_s=a\hat{\Phi}\left(\frac{l}{a},\nu,H\right). \tag{32}$$

As it is seen from (32), r_s is not necessarily proportional to "*a*." The functional form of $\hat{\Phi}$ cannot be found using dimensional analysis and this makes the use of crack-effect zone method very difficult. But we know that for most engineering materials *l* is very small ($l/a \ll 1$). We also know that the following limit exists:

$$\lim_{l/a\to 0}\hat{\Phi}\left(\frac{l}{a},\nu,H\right)=\hat{\Phi}(0,\nu,H)=\Phi(\nu,H) \tag{33}$$

because when l/a tends to zero for a constant "*a*" we approach the classical theory and obviously $\hat{\Phi}$ is defined for the classical theory and is finite. Thus we have a complete similarity or a similarity of the first kind (see, for example, the excellent book of Barenblatt [80]). Therefore according to dimensional analysis for very small $l/a(l/a \ll 1)\hat{\Phi}$ can be considered independent of l/a and replaced by its limit Φ. Therefore, $r_s \sim a$ for $l/a \ll 1$. As a matter of fact, we do not need to limit ourselves to the case $l/a \ll 1$. We can show that $\hat{\Phi}$ is not a function of l/a as we see in the following. We know that for a smooth crack both stresses and couple-stresses have an $r^{-1/2}$ singularity regardless of the size of the characteristic length(s) of the cracked micropolar material. Suppose that the radius of the dominant zone of stress and couple-stress singularity is r_s. Thus, $\delta U_s \propto \delta(r_s^{2-2\alpha})=\delta(r_s)$ and $\delta U_s \propto \delta(a)$. Therefore, according to Griffith's criterion we must have $r_s \propto a$. Thus

$$\hat{\Phi}\left(\frac{l}{a},\nu,H=1\right)=f(\nu). \tag{34}$$

For a fractal crack H is a local parameter while l/a is a global parameter. The reason is that H is defined only for the fractal crack, which has 2-measure (area) zero but l/a is defined for all points of the domain other than the crack. Therefore, we expect $\hat{\Phi}$ to be separable, i.e.,

$$\hat{\Phi}\left(\frac{l}{a},\nu,H\right)=\hat{\Phi}_1\left(\frac{l}{a},\nu\right)\hat{\Phi}_2(\nu,H). \tag{35}$$

From (34) and (35) we obtain

$$\hat{\Phi}\left(\frac{l}{a},\nu,H\right)=\tilde{\Phi}(H,\nu). \tag{36}$$

Therefore, r_s is always proportional to "a" regardless of the value of l/a.

From (31b) we have

$$\delta U_e^{\sigma}\propto\delta(a^{2-2\alpha_1}) \quad \text{and} \quad \delta U_e^{m}\propto\delta(a^{2-2\alpha_2}). \tag{37}$$

The next thing we need is the asymptotic form of surface energies. From (22) we have

$$\delta U_s^{u}\propto\delta U_s^{\phi}\propto\delta U_s\propto\begin{cases}\delta(a^{1/H}) & \frac{1}{2}\leqslant H<1\\ \delta(a^2) & 0<H\leqslant\frac{1}{2}\end{cases}. \tag{38}$$

We use both forms of Griffith's criterion and show that they give us the same result.

Using uncoupled micropolar Griffith's criterion is easier and yields

$$\delta(a^{2-2\alpha_1})\propto\begin{cases}\delta(a^{1/H}) & \frac{1}{2}\leqslant H<1\\ \delta(a^2) & 0<H\leqslant\frac{1}{2}\end{cases} \tag{39a}$$

$$\delta(a^{2-2\alpha_2})\propto\begin{cases}\delta(a^{1/H}) & \frac{1}{2}\leqslant H<1\\ \delta(a^2) & 0<H\leqslant\frac{1}{2}\end{cases}. \tag{39b}$$

Thus

$$\alpha_1=\alpha_2=\begin{cases}\frac{2H-1}{2H} & \frac{1}{2}\leqslant H<1\\ 0 & 0<H\leqslant\frac{1}{2}\end{cases}. \tag{40}$$

Using coupled Griffith's criterion is tricky. From (20), (38), and (39) we obtain

$$C_1\delta(a^{2-2\alpha_1})+C_2\delta(a^{2-2\alpha_2})\propto\begin{cases}\delta(a^{1/H}) & \frac{1}{2}\leqslant H<1\\ \delta(a^2) & 0<H\leqslant\frac{1}{2}\end{cases} \tag{41}$$

where C_1 and C_2 are not functions of a. We prove by contradiction that α_1 and α_2 must be equal. Suppose that $\alpha_1\neq\alpha_2$ and for example $\alpha_1>\alpha_2$. Notice that "$\propto$" means "proportional to" and that Eq. (41) holds for an arbitrary crack length "a." For a very large "a" ($a\gg1$) we can write

$$a^{2-2\alpha_2}\gg a^{2-2\alpha_1}. \tag{42}$$

Hence

$$C_1\delta(a^{2-2\alpha_1})+C_2\delta(a^{2-2\alpha_2})\cong C_2\delta(a^{2-2\alpha_2})\propto\begin{cases}\delta(a^{1/H}) & \frac{1}{2}\leqslant H<1\\ \delta(a^2) & 0<H\leqslant\frac{1}{2}\end{cases} \tag{43}$$

Thus

$$\alpha_2=\begin{cases}\frac{2H-1}{2H} & \frac{1}{2}\leqslant H<1\\ 0 & 0\leqslant H\leqslant\frac{1}{2}\end{cases} \tag{44}$$

On the other hand, for a very small "a" ($a\ll1$) we can write

$$a^{2-2\alpha_1}\gg a^{2-2\alpha_2}. \tag{45}$$

Thus

$$C_1\delta(a^{2-2\alpha_1})+C_2\delta(a^{2-2\alpha_2})\cong C_1\delta(a^{2-2\alpha_1})\propto\begin{cases}\delta(a^{1/H}) & \frac{1}{2}\leqslant H<1\\ \delta(a^2) & 0<H\leqslant\frac{1}{2}\end{cases}. \tag{46}$$

Hence

$$\alpha_2=\begin{cases}\frac{2H-1}{2H} & \frac{1}{2}\leqslant H<1\\ 0 & 0<H\leqslant\frac{1}{2}\end{cases}. \tag{47}$$

From (44) and (47) we see that $\alpha_1=\alpha_2$, which is a contradiction. Therefore our assumption was false and α_1 and α_2 must be equal, i.e.,

$$\alpha_1=\alpha_2=\begin{cases}\frac{2H-1}{2H} & \frac{1}{2}\leqslant H<1\\ 0 & 0<H\leqslant\frac{1}{2}\end{cases}. \tag{48}$$

Therefore

$$\sigma_{ij}\sim r\,\frac{-2H-1}{2H},\quad m_{ij}\sim r\,\frac{-2H-1}{2H}\ \text{ as }\ r\to0\quad\frac{1}{2}\leqslant H<1 \tag{49a}$$

$$\sigma_{ij}\sim r^0,\quad m_{ij}\sim r^0\quad\text{as}\quad r\to0\quad 0<H\leqslant\frac{1}{2}. \tag{49b}$$

It is seen that both forms of Griffith's criterion yield the same result: stresses and couple-stresses have equal orders of singularity and this order of singularity is the same as that of stresses at the tip of a fractal crack in a classical continuum. This result is similar to that reached by Sternberg and Muki [65]: that in a couple-stress medium at the tip of a smooth crack both stresses and couple-stresses have $r^{-1/2}$ singularities. This is also true for self-similar cracks; orders of stress and couple-stress singularity are equal.

A similar result can be reached for mode IV self-affine cracks. A mode IV fractal crack is shown in Fig. 3(b). This new mode of fractal fracture was introduced by Yavari [15] and Yavari et al. [14]. As was done for a mode I fractal crack, the orders of stress and couple-stress singularity can be calculated. The only modification in the analysis is to change Eq. (37) to read ([14,16])

$$\delta U_e^{\sigma}\propto\delta(a^{1+H-2\alpha_1H}) \quad \text{and} \quad \delta U_e^{m}\propto\delta(a^{1+H-2\alpha_2H}) \tag{50}$$

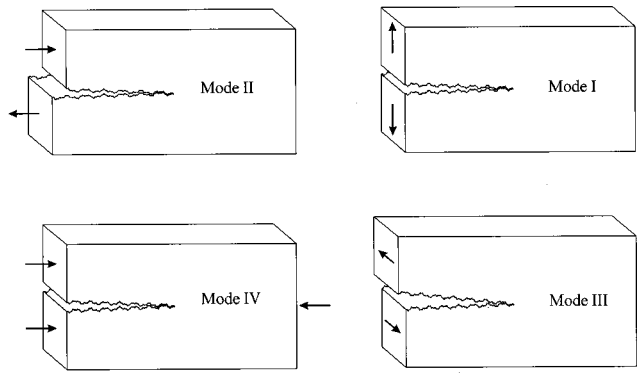


Fig. 4 The four modes of fractal fracture: mode I (opening mode), mode II (shearing mode), mode III (tearing mode), and mode IV (axial mode)

The stresses and couple-stresses have the following asymptotic forms:

$$\sigma_{ij}\sim r\,\frac{-H^2-H+1}{2H^2},\quad m_{ij}\sim r\,\frac{-H^2-H+1}{2H}\quad \text{as}\quad r\to 0\quad \frac{1}{g}\leqslant H<1 \tag{51a}$$

$$\sigma_{ij}\sim r^0,\quad m_{ij}\sim r^0\quad \text{as}\quad r\to 0\quad 0<H\leqslant\frac{1}{g} \tag{51b}$$

where $g=(\sqrt{5}+1)/2$ is the Golden ratio. All four modes of fracture are shown in Fig. 4. (Actually, there are six modes. We found the fifth and sixth modes very recently ([16])).

For three-dimensional cracked bodies made of a couple-stress material or a micropolar material a similar conclusion can be reached.

6 Conclusions

Fracture mechanics of smooth cracks in micropolar continua is reviewed. Griffith's fracture theory is generalized for rectilinear and fractal cracks in micropolar continua. It is seen that Griffith's criterion can have two forms: uncoupled micropolar Griffith's criterion and coupled micropolar Griffith's criterion. Using dimensional analysis and the method of crack-effect zone it is shown that both forms of Griffith's criterion predict that stresses and couple-stresses have the same order of singularity. This order of stress and couple-stress singularity is shown to be equal to that of stresses at the tip of a fractal crack in a classical continuum.

Acknowledgments

The first author is grateful to Prof. J. W. Hutchinson (of Harvard University) for helpful discussions on generalization of Griffith's theory for smooth cracks in micropolar solids and to Prof. K. G. Hockett (of The George Washington University) for helpful discussions on fractal geometry.

Appendix

Fractal Geometry. This Appendix presents some concepts and definitions of fractal geometry. Here we discuss only those aspects of fractal geometry that are directly relevant to our investigation. For more details the reader may refer to Mandelbrot [2,81–83], Feder [84], Vicsek [85], and Falconer [86,87].

Suppose that $U\neq\varnothing$ is a subset of $\mathbb{R}^n$. The diameter of U is defined as

$$\text{diam}(U)=\sup\{|x-y|:x,y\in U\}. \tag{A1}$$

An ε-cover of S is a countable or finite collection of sets $\{U_i\}$ such that

1. $0<\text{diam}(U_i)\leqslant\varepsilon$,
2. $S\subset\cup_{i=1}^{\infty}U_i$.

Consider a set $S\subset\mathbb{R}^n$. An affine transformation of real scaling ratios $r_1,r_2,\ldots,r_n$ $(0<r_i<1)$ transforms each $x=(x_1,x_2,\ldots,x_n)\in S$ into $r(x)=(r_1x_1,r_2r_2,\ldots,r_nx_n)\in r(S)$. The set S is self-affine if it is composed of N nonoverlapping subsets congruent to $r(S)$. If the above property holds for S when $r_1=r_2=\ldots=r_n=r$, it is called a self-similar set. A self-similar fractal is invariant under an isotropic length-scale transformation while a self-affine fractal is invariant under a transformation with different length scales in different directions.

Roughly speaking, measure of a set $S\subset\mathbb{R}^n$ tells us about the size of the set and is denoted by $\mu(S)$. In other words, measure is a generalized size. μ is a measure on $\mathbb{R}^n$ if it assigns a nonnegative real number (possibly $+\infty$) to each subset of $\mathbb{R}^n$ and satisfies the following requirements:

1. $\mu(\varnothing)=0$
2. $\mu(A)\leqslant\mu(B)$ if $A\subset B$
3. If $A_1,A_2,\ldots$ is a finite or countable sequence of subsets of $\mathbb{R}^n$ then

$$\mu\left(\overset{\infty}{\underset{i=1}{\cup}}A_i\right)\leqslant\sum_{i=1}^{\infty}\mu(A_i) \tag{A2}$$

with equality if A_i's are disjoint subsets of $\mathbb{R}^n$.

Now suppose that $S\subset\mathbb{R}^n$ and $D\in\mathbb{R}^+\cup\{0\}$. The D-dimensional Hausdorff measure of S is denoted by $\mathcal{H}^D(S)$ and is defined as

$$H^D(S)=\lim_{\varepsilon\to 0}H^D_\varepsilon(S) \tag{A3}$$

where

$$H^D_\varepsilon(S)=\inf\left\{\sum_{i=1}^{\infty}\text{diam}(U_i)^D:\{U_i\}\ \text{is an}\ \varepsilon-\text{cover of}\ S\right\} \tag{A4}$$

It can be shown that $\mathcal{H}^D$ has all the properties of a measure. It can be proved that for any set S, $\mathcal{H}^D(S)$ has a jump from $+\infty$ to 0 for one and only one value of D, which is called the Hausdorff dimension of S, i.e.,

$$D_H=\inf\{D:H^D(S)=0\}=\sup\{D:H^D(S)=+\infty\}. \tag{A5}$$

Therefore

$$H^D(S)=\begin{cases}+\infty & D<\text{Dim}_H S\\ 0 & D>\text{Dim}_H S\end{cases}. \tag{A6}$$

There are many other definitions of dimension. One disadvantage of Hausdorff dimension is the difficulty of calculating it, which makes it impractical. Here we discuss two other important dimensions, namely box dimension and divider dimension. All different dimensions somehow measure the complexity of irregularity of a set. It should be emphasized that dimension provides only limited information about a fractal set. In most definitions there is a measurement at scale ε. For each ε irregularities below this scale are ignored and the behavior of measurements as $\varepsilon\to 0$ is studied.

Box dimension: Let $S\neq\varnothing$ be a subset of $\mathbb{R}^n$ and let $N^B_\varepsilon(S)$ be the smallest number of sets of diameter at most ε which can cover S. Box dimension of S is D_B if

$$N^B_\varepsilon(S)=0(\varepsilon^{-D_B})\ \text{ as }\ \varepsilon\to 0\ \text{ or }\ D_B=\lim_{\varepsilon\to 0}\frac{\log N^B_\varepsilon(S)}{-\log\varepsilon} \tag{A7}$$

where O is Landau's order symbol. It can be shown that always $D_H\leqslant D_B$. For self-similar fractals the equality holds. Box measure m_{D_B} is defined as

$$m^{\varepsilon}_{D_B}=N^{B}_{\varepsilon}(S)\varepsilon^{D_B}$$

$$=\inf\left\{\sum_i \varepsilon^{D_B}:\{U_i\} \text{ in a finite } \varepsilon-\text{cover of } S\right\},$$

$$m_{D_B}=\lim_{\varepsilon\to 0} m^{\varepsilon}_{D_B}. \tag{A8}$$

In calculating Hausdorff measure different weights $|U_i|^s$ are assigned to covering sets U_i while in box measure the same weight ε^{D_B} is used for all covering sets. It should be noted that m_{D_B} is not a mathematical measure on subsets of $\mathbb{R}^n$ because it is not σ-additive. (It is actually a "content.")

Divider dimension: This is the most important dimension in applications to fractal fracture mechanics problems. Consider a Jordan curve C (a curve that does not intersect itself) $f:[a,b]\to\mathbb{R}^n$, where f is a bijection (a one-to-one and onto function). For $\varepsilon>0$ define $N^{D}_{\varepsilon}(C)$ to be the maximum number of points $x_0,x_1,x_2,\ldots,x_m$ on C such that $|x_k-x_{k-1}|=\varepsilon$ for $k=1,2,\ldots,m$. Therefore, the approximate length of the curve $L_\varepsilon(C)$ is $L_\varepsilon(C)=O[(N^{D}_{\varepsilon}(C)-1)\varepsilon]$. The divider dimension of C is D_D if

$$L_\varepsilon(C)=O(\varepsilon^{1-D_D}) \text{ as } \varepsilon\to 0 \text{ or } N^{D}_{\varepsilon}(C)=O(\varepsilon^{-D_D}) \text{ as } \varepsilon\to 0. \tag{A9}$$

We know that $N^{D}_{\varepsilon}(C)$ is dimensionless while ε has dimension of length. Therefore from (A9) we conclude that

$$N^{D}_{\varepsilon}(C)\sim\left(\frac{\varepsilon}{L_0}\right)^{-D_D} \text{ or } N^{D}_{\varepsilon}(C)\sim\varepsilon^{-D_D}L_0^{D_D} \tag{A10}$$

where L_0 is the nominal length of C. It can be shown that for any Jordan curve C, $D_D\geqslant D_B$. For self-similar curves the equality holds. Divider measure m_{D_D} is defined as

$$m^{\varepsilon}_{D_D}=N^{D}_{\varepsilon}(C)\varepsilon^{D_D}, \quad m_{D_D}=\lim_{\varepsilon\to 0} m^{\varepsilon}_{D_D}. \tag{A11}$$

From (A10) and (A11) we can write

$$m_{D_D}\sim L_0^{D_D} \quad \text{or} \quad m_{D_D}=\eta L_0^{D_D}. \tag{A12}$$

Like box measure, divider measure is not a mathematical measure because it is not σ-additive.

Consider a topologically one-dimensional set. Suppose that this set can be expressed as the graph of a single-valued function $F(t)$ embedded in $\mathbb{R}^2$. Then $F(t)$ is a self-affine function if

$$F(t)=r^{-H}F(rt) \quad \forall r,t\in R \tag{A13}$$

where $H(0<H<1)$ is the Hurst (roughness) exponent. Weierstrass-Mandelbrot function is an example of a self-affine function and is defined as

$$WM(x)=\sum_{-\infty}^{+\infty} a^{-nH}[1-\cos(a^n x)] \quad a>1, \quad 0<H<1. \tag{A14}$$

This function satisfies the invariance relation (A13). It can be shown that for a self-affine curve locally ($\varepsilon\leqslant\varepsilon_x$) we have

$$D_B=2-H \text{ and } D_D=\begin{cases} \dfrac{1}{H} & \dfrac{1}{2}\leqslant H<1 \\ 2 & 0<H\leqslant\dfrac{1}{2} \end{cases} \tag{A15a}$$

and globally ($\varepsilon\geqslant\varepsilon_x$)

$$D_B=D_D=1. \tag{A15b}$$

The length scale ε_x is called the crossover length ($\varepsilon_x^H\cong\varepsilon_x$). In general, for a self-affine fractal (with Hurst exponent H) embedded in $\mathbb{R}^n$ the divider and box dimensions are locally related to roughness exponent by

$$D_D=\begin{cases} \dfrac{n-1}{H} & \dfrac{n-1}{n}\leqslant H<1 \\ n & 0<H\leqslant\dfrac{n-1}{n} \end{cases} \tag{A16a}$$

$$D_B=n-H. \tag{A16b}$$

And globally, $D_D=D_B=n-1$.

References

[1] Mandelbrot, B. B., 1975, *Les Objets Fractals: Forme, Hasard et Dimension*, Flammarion, Paris.

[2] Mandelbrot, B. B., 1983, *Fractal Geometry of Nature*, W. H. Freeman and Company, New York.

[3] Cherepanov, G. P., Balankin, A. S., and Ivanova, V. S., 1995, "Fractal Fracture Mechanics A Review," Eng. Fract. Mech., **51**, No. 6, pp. 997–1033.

[4] Balankin, A. S., 1997, "Physics of Fracture and Mechanics of Self-Affine Cracks," Eng. Fract. Mech., **57**, No. 2, pp. 135–203.

[5] Mandelbrot, B. B., Passoja, D. E., and Paullay, A. J., 1984, "Fractal Character of Fracture Surfaces in Metals," Nature (London), **308**, pp. 721–722.

[6] Brown, S. R., and Scholz, C. H., 1985, "Broad Bandwidth Study of the Topography of Natural Rock Surfaces," J. Geophys. Res., **90B**, pp. 12575–12582.

[7] Power, W. L., and Tullis, T. E., 1991, "Euclidean and Fractal Models for the Description of Rock Surface Roughness," J. Geophys. Res., **96B**, pp. 415–424.

[8] Saouma, V. E., Barton, C. C., and Gamaledin, N. A., 1990, "Fractal Characterization of Fracture Surface in Concrete," Eng. Fract. Mech., **35**, pp. 47–53.

[9] Saouma, V. E., and Barton, C. C., 1994, "Fractals, Fractures, and Size Effect in Concrete," J. Eng. Mech., **120**, No. 4, pp 835–854.

[10] Wong, P., Howard, J., and Lin, J. S., 1986, "Surface Roughening and The Fractal Nature of Rocks," Phys. Rev. Lett., **57**, pp. 637–640.

[11] Mosolov, A. B., 1991, "Cracks with Fractal Surfaces," Dokl. Akad. Nauk SSR, **319**, No. 4, pp. 840–844.

[12] Gol'dshteĭn, R. V., and Mosolov, A. B., 1991, "Cracks With a Fractal Surface," Sov. Phys. Dokl., **36**, No. 8, pp. 603–605.

[13] Gol'dshteĭn, R. V., and Mosolov, A. B., 1992, "Fractal Cracks," J. Appl. Math. Mech., **56**, No. 4, pp. 563–571.

[14] Yavari, A., Hockett, K. G., and Sarkani, S., 2000, "The Fourth Mode of Fracture in Fractal Fracture Mechanics," Int. J. Fract., **101**, No. 4, pp. 365–384.

[15] Yavari, A., 2000, "Fracture Mechanics of Fractal Cracks in Classical and Micropolar Solids," M.S. thesis, The George Washington University, Washington, DC.

[16] Yavari, A., Sarkani, S., and Moyer, E. T., "The Mechanics of Self-Similar and Self-Affine Fractal Cracks," submitted for publication.

[17] Mosolov, A. B., 1991, "Fractal J-integral in Fracture," Sov. Tech. Phys. Lett., **17**, pp. 698–700.

[18] Mosolov, A. B., and Borodich, F. M., 1992, "Fractal Fracture of Brittle Bodies During Compression," Sov. Phys. Dokl., **37**, No. 5, pp. 263–265.

[19] Mosolov, A. B., 1993, "Mechanics of Fractal Cracks in Brittle Solids," Europhys. Lett., **24**, No. 8, pp. 673–678.

[20] Xie, H., 1989, "The Fractal Effect of Irregularity of Crack Branching on the Fracture Toughness of Brittle Materials," Int. J. Fract., **41**, pp. 267–274.

[21] Xie, H., and Sanderson, D. J., 1995, "Fractal Effects of Crack Propagation on Dynamic Stress Intensity Factors and Crack Velocities," Int. J. Fract., **74**, pp. 29–42.

[22] Borodich, F. M., 1992, "Fracture Energy in a Fractal Crack Propagating in Concrete or Rock," Doklady Rossiyskoy Akademii Nauk, **325**, No. 6, pp. 1138–1141.

[23] Borodich, F. M., 1997, "Some Fractal Models of Fracture," J. Mech. Phys. Solids, **45**, No. 2, pp. 239–259.

[24] Yavari, A., "Generalization of Barenblatt's Fracture Theory for Fractal Cracks," Fractals, to appear.

[25] Voigt, W., 1887, "Theoretische Studien über die Elastizitätsverhältnisse der Kristalle," Abh. Ges. Wiss. Göttingen 34.

[26] Cosserat, E., and Cosserat, F., 1909, *Theorie des Corps Deformables*, Hermann et Fils, Paris.

[27] Truessdel, C., and Toupin, R. A., 1960, *The Classical Field Theories* (Encyclopedia of Physics, Vol. III/1, Secs. 200, 203, 205), Springer, Berlin.

[28] Grioli, G., 1960, "Elastica Asimmetrica," Ann. Mat. Pura appl., **50**, Ser. IV, p. 389.

[29] Toupin, R. A., 1962, "Elastic Material With Couple-Stresses," Arch. Ration. Mech. Anal., **11**, pp. 385–414.

[30] Toupin, R. A., 1964, "Theories of Elasticity With Couple-Stresses," Arch. Ration. Mech. Anal., **17**, pp. 85–112.

[31] Aero, E. L., and Kuvshinskii, E. V., 1961, "Fundamental Equations of the Theory of Elastic Media with Rotationally Interacting Particles," Sov. Phys. Solid State, **2**, pp. 1272–1281.

[32] Mindlin, R. D., and Tiersten, H. F., 1962, "Effects of Couple-Stresses in Linear Elasticity," Arch. Ration. Mech. Anal., **11**, pp. 415–448.

[33] Koiter, W. T., 1964, "Couple-Stresses in the Theory of Elasticity," Proc. K. Ned. Akad. Wet., **67**, p. 17.

[34] Eringen, A. C., and Suhubi, E. S., 1964, ''Nonlinear Theory of Simple Micro-Elastic Solids—I,'' Int. J. Eng. Sci., **2**, pp. 189–203.
[35] Suhubi, E. S., and Eringen, A. C., 1964, ''Nonlinear Theory of Simple Micro-Elastic Solids—II,'' Int. J. Eng. Sci., **2**, pp. 389–404.
[36] Eringen, A. C., 1966, ''Linear Theory of Micropolar Elasticity,'' J. Math. Mech., **15**, pp. 909–923.
[37] Eringen, A. C., 1968, *Theory of Micropolar Elasticity* (in Fracture), H. Liebowitz, ed., Academic Press, New York.
[38] Cowin, S. C., 1969, ''Singular Stress Concentrations in Plane Cosserat Elasticity,'' Z. Angew. Math. Phys., **20**, pp. 979–982.
[39] Cowin, S. C., 1970, ''Stress Functions for Cosserat Elasticity,'' Int. J. Solids Struct., **6**, pp. 389–398.
[40] Cowin, S. C., 1970, ''An Incorrect Inequality in Micropolar Elasticity Theory,'' Z. Angew. Math. Phys., **21**, pp. 494–497.
[41] Mindlin, R. D., 1962, ''Influence of Couple-Stresses on Stress Concentrations,'' Exp. Mech., **3**, pp. 1–7.
[42] Neuber, H., 1966, ''Über Probleme der Spannungskonzentration in Cosserat-Körper,'' Acta Mech., **2**, pp. 48–69.
[43] Ariman, T., 1967, ''On the Stresses Around a Circular Hole in Micropolar Elasticity,'' Acta Mech., **3**, pp. 216–229.
[44] Kaloni, P. N., and Ariman, T., 1967, ''Stress Concentration Effects in Micropolar Elasticity,'' Z. Angew. Math. Phys., **18**, pp. 136–141.
[45] Itou, S., 1973, ''The Effect of Couple-Stresses on The Stress Concentration Around an Elliptical Hole,'' Acta Mech., **16**, pp. 289–296.
[46] Krishnaswamy, S., Jin, Z.-H., and Batra, R. C., 1998, ''Stress Concentration in an Elastic Cosserat Plate Undergoing Extensional Deformations,'' ASME J. Appl. Mech., **65**, pp. 66–70.
[47] Kishida, M., Sasaki, K., and Ishii, S., 1990, ''Torsion of a Circular Bar With Annular Groove in Couple-Stress Theory,'' Int. J. Eng. Sci., **28**, pp. 773–781.
[48] Hartranft, R. J., and Sih, G. C., 1965, ''The Effect of Couple-Stresses on Stress Concentration of a Circular Inclusion,'' ASME J. Appl. Mech., **32**, pp. 429–431.
[49] Weitsman, Y., 1965, ''Couple-Stress Effects on Stress Concentration Around a Cylindrical Inclusion in a Field of Uniaxial Tension,'' ASME J. Appl. Mech., **32**, pp. 424–428.
[50] Banks, Jr., C. B., and Sokolowski, M., 1968, ''On Certain Two-Dimensional Applications of the Couple Stress Theory,'' Int. J. Solids Struct., **3**, pp. 757–770.
[51] Itou, S., 1977, ''The Effect of Couple-Stresses on the Stress Concentration Around a Rigid Circular Inclusion in a Strip Under Tension,'' Acta Mech., **27**, pp. 261–268.
[52] Wang, T. T., 1970, ''The Effect of Couple-Stress on Maximum Stress and Its Location Around Spherical Inclusions,'' ASME J. Appl. Mech., **37**, pp. 865–868.
[53] Ellis, R. W., and Smith, C. W., 1967, ''A Thin-Plate Analysis and Experimental Evaluation of Couple-Stress Effects,'' Exp. Mech., **7**, pp. 372–380.
[54] Perkins, R. W., Jr., and Thompson, D., 1973, ''Experimental Evidence of Couple-Stress Effect,'' AIAA J., **11**, pp. 1053–1055.
[55] Gauthier, R. D., and Jahsman, W. E., 1975, ''A Quest for Micropolar Elastic Constants,'' ASME J. Appl. Mech., **42**, pp. 369–374.
[56] Gauthier, R. D., and Jahsman, W. E., 1976, ''Bending of a Curved Bar of Micropolar Elastic Material,'' ASME J. Appl. Mech., **43**, pp. 502–503.
[57] Reddy, G. V., and Venkatasubramanian, N. K., 1978, ''On the Flexural Rigidity of a Micropolar Elastic Circular Cylinder,'' ASME J. Appl. Mech., **45**, pp. 429–431.
[58] Park, H. C., and Lakes, R. S., 1987, ''Torsion of a Micropolar Elastic Prism of Square Cross-Section,'' Int. J. Solids Struct., **23**, pp. 485–503.
[59] Lakes, R., 1986, ''Experimental Microelasticity of Two Porous Solids,'' Int. J. Solids Struct., **22**, pp. 55–63.
[60] Lakes, R., 1991, ''Experimental Micro Mechanics Methods for Conventional and Negative Poisson's Ratio Cellular Solids as Cosserat Continua,'' ASME J. Appl. Mech., **113**, pp. 148–155.
[61] Lakes, R., 1995, ''Experimental Methods for Study of Cosserat Elastic Solids and Other Generalized Elastic Continua,'' *Continuum Methods for Materials With Microstructure*, H.-B. Mühlhaus, ed., John Wiley and Sons, New York.
[62] Anthoine, A., 2000, ''Effect of Couple-Stresses on the Elastic Bending of Beams,'' Int. J. Solids Struct., **37**, pp. 1003–1018.
[63] Hutchinson, J. W., 2000, ''Plasticity in the Micron Scale,'' Int. J. Solids Struct., **37**, pp. 225–238.
[64] Schijve, J., 1966, ''Note on Couple Stresses,'' J. Mech. Phys. Solids, **14**, pp. 113–120.
[65] Sternberg, E., and Muki, R., 1967, ''The Effect of Couple-Stresses on the Stress Concentration Around a Crack,'' Int. J. Solids Struct., **3**, pp. 69–95.
[66] Sih, G. C., and Liebowitz, H., 1968, *Mathematical Theories of Brittle Fracture* (In Fracture, Vol. II), H. Liebowitz, ed., Academic Press, New York, pp. 67–190.
[67] Ejike, U. B. C. O., 1969, ''The Plane Circular Crack Problem in the Linearized Couple-Stress Theory,'' Int. J. Eng. Sci., **7**, pp. 947–961.
[68] Kim, B. S., and Eringen, A. C., 1973, ''Stress Distribution Around an Elliptic Hole in an Infinite Micropolar Elastic Plate,'' Lett. Appl. Eng. Sci., **1**, pp. 381–390.
[69] Sláadek, J., and Sláadek, V., 1984, ''The Effect of Couple Stresses on the Stress Field Around a Penny-Shaped Crack,'' Int. J. Fract., **25**, pp. 109–119.
[70] Nakamura, S., and Lakes, R. S., 1988, ''Finite Element Analysis of Stress Concentration Around a Blunt Crack in a Cosserat Elastic Solid,'' Comput. Methods Appl. Mech. Eng., **66**, pp. 257–266.
[71] Han, S. Y., Narasimhan, M. N. L., and Kennedy, T. C., 1990, ''Dynamic Propagation of a Finite Crack in a Micropolar Elastic Solid,'' Acta Mech., **85**, pp. 179–191.
[72] Kennedy, T. C., and Kim, J. B., 1993, ''Dynamic Analysis of Cracks in Micropolar Elastic Materials,'' Eng. Fract. Mech., **44**, pp. 207–216.
[73] Yadava, R. N., Roy, A., and Katiyar, R. K., 1994, ''The Effect of Internal Pressure on a Penny-Shaped Crack at the Interface of Two Bonded Dissimilar Micropolar Elastic Half-Spaces,'' Int. J. Fract., **65**, pp. 19–30.
[74] Atkinson, C., and Leppington, F. G., 1977, ''The Effect of Couple Stresses on the Tip of a Crack,'' Int. J. Solids Struct., **13**, pp. 1103–1122.
[75] Lubarda, V. A., and Markenscoff, X., 2000, ''Conservation Integrals in Couple Stress Elasticity,'' J. Mech. Phys. Solids, **48**, pp. 553–564.
[76] Griffith, A. A., 1920, ''The Phenomenon of Rupture and Flow in Solids,'' Philos. Trans. R. Soc. London, Ser. A, **A221**, pp. 163–198.
[77] Griffith, A. A., 1924, *Proceedings of the 1st International Congress for Applied Mechanics*, Delft, p. 55.
[78] Buckingham, E., 1914, ''On Physically Similar Systems: Illustrations of the Use of Empirical Equations,'' Phys. Rev., **IV**, No. 4, pp. 345–376.
[79] Buckingham, E., 1915, ''Model Experiments and the Form of Empirical Equations,'' Trans. ASME, **37**, pp. 263–296.
[80] Barenblatt, G. I., 1996, *Scaling, Self-Similarity, and Intermediate Asymptotics*, Cambridge University Press, New York.
[81] Mandelbrot, B. B., 1985, ''Self-Affine Fractals and Fractal Dimension,'' Phys. Scr., **32**, pp. 257–260.
[82] Mandelbrot, B. B., 1986, ''Self-Affine Fractal Sets, I: The Basic Fractal Dimensions,'' *Fractals in Physics*, L. Pietronero and E. Tosatti, eds., Elsevier, New York, pp. 3–16.
[83] Mandelbrot, B. B., 1986, ''Self-Affine Fractal Sets, II: Length and Surface Dimensions,'' *Fractals in Physics*, L. Pietronero and E. Tosatti, eds., Elsevier, New York, pp. 17–20.
[84] Feder, J., 1988, *Fractals*, Plenum Press, New York.
[85] Vicsek, T., 1989, *Fractal Growth Phenomena*, World Scientific, Singapore.
[86] Falconer, K., 1990, *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley and Sons, Chichester.
[87] Falconer, K., 1997, *Techniques in Fractal Geometry*, John Wiley and Sons, Chichester.

Electromechanical Effects of a Screw Dislocation Around a Finite Crack in a Piezoelectric Material

J. H. Kwon

K. Y. Lee[1]
Professor,
e-mail: fracture@yonsei.ac.kr

Department of Mechanical Engineering,
Yonsei University,
Seoul 120-749, Korea

The interaction between a screw dislocation and a finite crack in an unbounded piezoelectric medium is studied in the framework of linear piezoelectric theory. A straight screw dislocation with the Burgers vector, which is normal to the isotropic basal plane, positioned around the tip of a finite crack is considered. In addition to having a discontinuous electric potential across the slip plane, the dislocation is assumed to be subjected to a line force and a line charge at the core. The explicit solution is derived by means of complex variable and conformal mapping methods. All field variables such as stress, strain, electric field, electric displacement near the crack tip, and the forces on a screw dislocation, the field intensity factors, and the energy release rate are determined under the combined out-of-plane mechanical and in-plane electrical loads. Also, the effects of screw dislocation and electrical loads are numerically analyzed.
[DOI: 10.1115/1.1427692]

Introduction

Studies on piezoelectric materials were reported actively during the last decade. It is because the piezoelectric materials have been used in various applications of the electromechanical devices such as actuators, sensors, and transducers, since the brothers Curie discovered the electromechanical coupling phenomenon of piezoelectric material in 1880. It is well known that defects, such as dislocations, cracks, cavities, and inclusions, can adversely influence the performance of such piezoelectric devices. For example, these defects carrying charges in piezoelectric semiconductors can be sources of internal electro-elastic fields ([1]). Therefore, in order to predict the performance and integrity of piezoelectric devices, it is important that the behaviors of various defects are analyzed and studied under the electrical and mechanical fields.

Deeg [2] studied the general defect mechanics of a piezoelectric material using Green's function method and modeling the defect with a collinear dislocation and charge dipole line. Later, Pak [3] considered the fracture mechanics problem of a finite crack, and he ([4]) derived the generalized Peach-Koehler forces acting on a screw dislocation in an unbounded piezoelectric material subjected to far-field electromechanical loads. Recently, efforts have been made to develop the inclusion models in a piezoelectric material (Wang [5], Chen [6], Liang et al. [7], Chung and Ting [8], and Zhong and Meguid [9]). However, the interaction between defects in piezoelectric media have been little considered. More recently, Meguid and Deng [10] discussed the electro-elastic interaction between a screw dislocation and an elliptical inhomogeneity in piezoelectric media, but they obtained only the distributions of the fields. Lee et al. [11] presented the forces on the screw dislocation and the energy release rate under the interaction between a semi-infinite crack and a screw dislocation in a piezoelectric medium.

In this work, to extend our previous work ([11]), we provide the interaction between a screw dislocation and a finite crack in an unbounded piezoelectric medium. A simple continuum model of a single screw dislocation around a finite crack in a hexagonal piezoelectric crystal subjected to antiplane mechanical and in-plane electrical loads is considered. The screw dislocation has a line force and a line charge along its core. The analyses are carried out in the framework of linear piezoelectric theory using conformal mapping and complex variable methods. The field variables and the forces acting on a piezoelectric screw dislocation around a finite crack are determined. Also, the effects of a screw dislocation and the loading conditions on the field intensity factors and energy release rate are discussed. The solutions are given for a screw dislocation located in arbitrary position around a crack.

Mathematical Model

Let's consider an unbounded piezoelectric medium containing a charged screw dislocation located at a point (x_0, y_0) around a finite crack of length $2a$, as shown in Fig. 1. The Cartesian coordinates (x,y,z) are set at the center of the crack for reference. A screw dislocation core is assumed to be straight and infinitely long in the z-direction, and subjected to a line force and a line charge. It is known that in ionic crystals, dislocations can have charges, which can be transported along with the dislocation ([12]). The piezoelectric medium is considered to be transversely isotropic with hexagonal symmetry, which has an isotropic basal plane of xy-plane and a poling direction of z-axis.

The piezoelectric boundary value problem is simplified in the case of out-of-plane mechanical displacement and in-plane electric fields such that

$$u_x = u_y = 0, \quad u_z = w(x,y), \tag{1}$$

$$E_x = E_x(x,y), \quad E_y = E_y(x,y), \quad E_z = 0. \tag{2}$$

In this case, the constitutive relations ([13]) become

$$\sigma_{zi}(x,y) = c_{44}\gamma_{zi}(x,y) - e_{15}E_i(x,y), \tag{3}$$

$$D_i(x,y) = e_{15}\gamma_{zi}(x,y) + \in_{11} E_i(x,y), \tag{4}$$

where $\sigma_{zi}(x,y)$, $\gamma_{zi}(x,y)$, $E_i(x,y)$, and $D_i(x,y)$ $(i = x,y)$ are the components of the stress, strain, electric field, and electric dis-

[1]To whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 9, 1999; final revision, August 9, 2001. Associate Editor: B. M. Moran. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

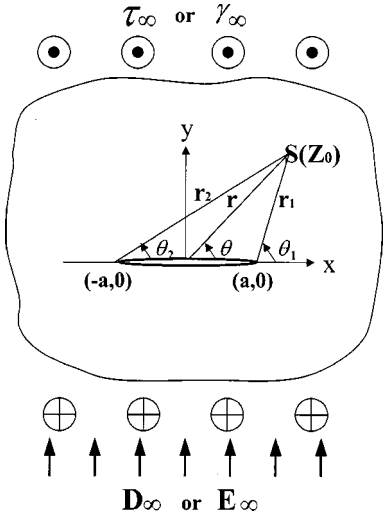


Fig. 1 A piezoelectric medium containing a screw dislocation around a finite crack of length 2*a* subjected to far-field mechanical and electrical loads

placement vectors, respectively. Also, c_{44} is the elastic modulus of the piezoelectric material measured in a constant electric field, e_{15} is the piezoelectric constant, $\in_{11}$ is the dielectric permittivity measured at a constant strain.

The electric field components are related to the electric potential ϕ by

$$E_i(x,y) = -\phi(x,y)_{,i} \quad (i=x,y). \tag{5}$$

The antiplane governing equations are simplified to

$$c_{44}\nabla^2 w(x,y) + e_{15}\nabla^2\phi(x,y) = 0, \tag{6}$$

$$e_{15}\nabla^2 w(x,y) - \in_{11}\nabla^2\phi(x,y) = 0, \tag{7}$$

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the two-dimensional Laplacian operator.

The boundary conditions on the upper and the lower surfaces of the crack are to be free of surface traction and surface charge, i.e.,

$$\sigma_{zy}(x,0^{\pm}) = 0, \quad D_y(x,0^{\pm}) = 0 \quad \text{at } |x| < a. \tag{8}$$

The following far-field boundary conditions are considered:

$$\text{Case 1:} \quad \sigma_{zy}(x,\pm\infty) = \tau_\infty \quad \text{and} \quad D_y(x,\pm\infty) = D_\infty, \tag{9}$$

$$\text{Case 2:} \quad \gamma_{zy}(x,\pm\infty) = \gamma_\infty \quad \text{and} \quad E_y(x,\pm\infty) = E_\infty, \tag{10}$$

$$\text{Case 3:} \quad \sigma_{zy}(x,\pm\infty) = \tau_\infty \quad \text{and} \quad E_y(x,\pm\infty) = E_\infty, \tag{11}$$

$$\text{Case 4:} \quad \gamma_{zy}(x,\pm\infty) = \gamma_\infty \quad \text{and} \quad D_y(x,\pm\infty) = D_\infty, \tag{12}$$

where τ_∞, γ_∞, E_∞, and D_∞ are uniform shear stress, uniform shear strain, uniform electric field, and uniform electric displacement, respectively.

Solution Procedure

By solving the simultaneous equations, the governing Eqs. (6) and (7) give

$$\nabla^2 w(x,y) = 0, \quad \nabla^2\phi(x,y) = 0. \tag{13}$$

The solution to Eq. (13) can be found by letting w and ϕ be the certain complex analytic functions such that

$$w(x,y) = W(Z), \quad \phi(x,y) = \Phi(Z), \tag{14}$$

where $Z = x + iy$ is a complex variable.

A crack on the x-axis is constructed using the following conformal mapping function, which transforms the circle $|\zeta| = 1$ in the ζ-plane onto a finite crack of length $2a$ along the real axis in the z-plane,

$$\zeta = \frac{1}{a}[Z + \sqrt{Z^2 - a^2}]. \tag{15}$$

By referring to the solution for an elastic material ([14–16]), in case of a screw dislocation subjected to a line force and a line charge, the complex potentials for displacement and electric potential are obtained, respectively, as follows:

$$\begin{aligned} W(\zeta) = {} & A_1\left[\log(\zeta-\zeta_0) + \log\left(\zeta - \frac{1}{\overline{\zeta_0}}\right)\right] \\ & + iA_2\left[\log(\zeta-\zeta_0) - \log\left(\frac{1}{\zeta} - \overline{\zeta_0}\right)\right. \\ & \left. + m\log\zeta - \log\zeta + \log(-\overline{\zeta_0})\right] + iA_3\left[\zeta - \frac{1}{\zeta}\right], \end{aligned} \tag{16}$$

$$\begin{aligned} \Phi(\zeta) = {} & B_1\left[\log(\zeta-\zeta_0) + \log\left(\zeta - \frac{1}{\overline{\zeta_0}}\right)\right] \\ & + iB_2\left[\log(\zeta-\zeta_0) - \log\left(\frac{1}{\zeta} - \overline{\zeta_0}\right)\right. \\ & \left. + m\log\zeta - \log\zeta + \log(-\overline{\zeta_0})\right] + iB_3\left[\zeta - \frac{1}{\zeta}\right], \end{aligned} \tag{17}$$

where A_j and B_j $(j=1,2,3)$ are real constants, and m implies the effect of an extra image dislocation considered to prevent the residual stress. If the real screw dislocation is emitted from the crack which is originally stress free, m should be zero. In contrast, m should be unity if the real screw dislocation originates elsewhere and comes near the crack which is originally stress free ([16]). Therefore, the potential functions (16) and (17) have three terms, respectively: the first corresponds to the line force or charge and its image, the second to the screw dislocation and its image, the third to the uniform external loads.

The strain, electric field, stress, and electric displacements can then be expressed in the forms

$$\gamma(x,y)=\gamma_{zx}(x,y)-i\gamma_{zy}(x,y)$$
$$=\frac{dW(\zeta)}{dZ}=\frac{dW(\zeta)}{d\zeta}\frac{d\zeta}{dZ}$$
$$=\frac{1}{\sqrt{Z^2-a^2}}\left[A_1\left(\frac{\zeta}{\zeta-\zeta_0}+\frac{\zeta\overline{\zeta_0}}{\zeta\overline{\zeta_0}-1}\right)\right.$$
$$\left.+iA_2\left(\frac{\zeta}{\zeta-\zeta_0}+\frac{1}{1-\zeta\overline{\zeta_0}}+m-1\right)+iA_3\left(\zeta+\frac{1}{\zeta}\right)\right], \tag{18}$$

$$E(x,y)=E_x(x,y)-iE_y(x,y)$$
$$=-\frac{d\Phi(\zeta)}{dZ}=-\frac{d\Phi(\zeta)}{d\zeta}\frac{d\zeta}{dZ}$$
$$=-\frac{1}{\sqrt{Z^2-a^2}}\left[B_1\left(\frac{\zeta}{\zeta-\zeta_0}+\frac{\zeta\overline{\zeta_0}}{\zeta\overline{\zeta_0}-1}\right)\right.$$
$$\left.+iB_2\left(\frac{\zeta}{\zeta-\zeta_0}+\frac{1}{1-\zeta\overline{\zeta_0}}+m-1\right)+iB_3\left(\zeta+\frac{1}{\zeta}\right)\right], \tag{19}$$

$$\sigma(x,y)=\sigma_{zx}(x,y)-i\sigma_{zy}(x,y)$$
$$=c_{44}\frac{dW(\zeta)}{dZ}+e_{15}\frac{d\Phi(\zeta)}{dZ}$$
$$=\frac{1}{\sqrt{Z^2-a^2}}\left[(c_{44}A_1+e_{15}B_1)\left(\frac{\zeta}{\zeta-\zeta_0}+\frac{\zeta\overline{\zeta_0}}{\zeta\overline{\zeta_0}-1}\right)\right.$$
$$+i(c_{44}A_2+e_{15}B_2)\left(\frac{\zeta}{\zeta-\zeta_0}+\frac{1}{1-\zeta\overline{\zeta_0}}+m-1\right)$$
$$\left.+i(c_{44}A_3+e_{15}B_3)\left(\zeta+\frac{1}{\zeta}\right)\right], \tag{20}$$

$$D(x,y)=D_x(x,y)-iD_y(x,y)$$
$$=e_{15}\frac{dW(\zeta)}{dZ}-\epsilon_{11}\frac{d\Phi(\zeta)}{dZ}$$
$$=\frac{1}{\sqrt{Z^2-a^2}}\left[(e_{15}A_1-\epsilon_{11}B_1)\left(\frac{\zeta}{\zeta-\zeta_0}+\frac{\zeta\overline{\zeta_0}}{\zeta\overline{\zeta_0}-1}\right)\right.$$
$$+i(e_{15}A_2-\epsilon_{11}B_2)\left(\frac{\zeta}{\zeta-\zeta_0}+\frac{1}{1-\zeta\overline{\zeta_0}}+m-1\right)$$
$$\left.+i(e_{15}A_3-\epsilon_{11}B_3)\left(\zeta+\frac{1}{\zeta}\right)\right]. \tag{21}$$

The six unknown constants, A_j and B_j $(j=1,2,3)$ can be determined by the theories of the force and charge balance conditions at the core, the displacement and electric potential jump conditions across the slip plane, and the far-field loading conditions at infinity ([11]) in the forms

$$A_1=\frac{-\epsilon_{11}p_z+e_{15}q_z}{2\pi(c_{44}\epsilon_{11}+e_{15}^2)},\quad B_1=-\frac{e_{15}p_z+c_{44}q_z}{2\pi(c_{44}\epsilon_{11}+e_{15}^2)}, \tag{22}$$

$$A_2=-\frac{b_z}{2\pi},\quad B_2=-\frac{\Delta\phi}{2\pi}, \tag{23}$$

Case 1:

$$A_3=-\frac{a}{2}\frac{\epsilon_{11}\tau_\infty+e_{15}D_\infty}{c_{44}\epsilon_{11}+e_{15}^2},\quad B_3=\frac{a}{2}\frac{c_{44}D_\infty-e_{15}\tau_\infty}{c_{44}\epsilon_{11}+e_{15}^2}, \tag{24}$$

Case 2:

$$A_3=-\frac{a}{2}\gamma_\infty,\quad B_3=\frac{a}{2}E_\infty \tag{25}$$

Case 3:

$$A_3=-\frac{a}{2}\frac{\tau_\infty+e_{15}E_\infty}{c_{44}},\quad B_3=\frac{a}{2}E_\infty \tag{26}$$

Case 4:

$$A_3=-\frac{a}{2}\gamma_\infty,\quad B_3=\frac{a}{2}\frac{D_\infty-e_{15}\gamma_\infty}{\epsilon_{11}} \tag{27}$$

where b_z, $\Delta\phi$, p_z, and q_z are the Burgers vector, electric potential jump, line force, and line charge, respectively.

Forces on a Screw Dislocation

The forces acting on a screw dislocation with a line force and a line charge can be obtained, using the relations introduced by Eshelby [17] and Pak [4], as follows:

$$F_x=b_z\sigma_{zy}^T+\Delta\phi D_y^T+p_z^S\gamma_{zx}^T+q_z^SE_x^T,$$
$$F_y=-b_z\sigma_{zx}^T-\Delta\phi D_x^T+p_z^S\gamma_{zy}^T+q_z^SE_y^T, \tag{28}$$

where the superscripts S and T represent two conceptional domains, i.e., S is the internal domain in which a screw dislocation exists, and T is the external domain in which a crack subjected to the mechanical and electrical loads exists. Therefore, Eq. (28) can determine the forces acting on a screw dislocation in the domain S due to the stress, strain, electric field, and electric displacement generated by the crack and the external loads in the domain T. The expressions for the field variables γ_{zj}^T, σ_{zj}^T, E_i^T, and D_i^T by the crack and the external loads can be obtained by subtracting the fields generated by the dislocation subjected to a line force and a line charge in an unbounded piezoelectric medium from the foregoing field Eqs. (18)–(21), and then by taking a limit at $Z=Z_0$. Substituting the results into Eq. (28), we obtain the forces acting on a screw dislocation subjected to a line force and a line charge in the presence of a finite crack as follows:

$$F_x=-[b_z(c_{44}A_1+e_{15}B_1)+\Delta\phi(e_{15}A_1-\in_{11}B_1)]\left[-\frac{1}{\sqrt{r_1r_2}}\sin\left(\frac{\theta_1+\theta_2}{2}\right)-\frac{r}{2r_1r_2}\left\{\sin(\theta-\theta_1-\theta_2)+\cos\theta\tan\left(\frac{\theta_1+\theta_2}{2}\right)\right\}\right]$$

$$-[b_z(c_{44}A_2+e_{15}B_2)+\Delta\phi(e_{15}A_2-\in_{11}B_2)]\left[\frac{m}{\sqrt{r_1r_2}}\cos\left(\frac{\theta_1+\theta_2}{2}\right)-\frac{r}{2r_1r_2}\{\cos(\theta-\theta_1-\theta_2)+\cos\theta\}\right]$$

$$-\frac{2}{a}[b_z(c_{44}A_3+e_{15}B_3)+\Delta\phi(e_{15}A_3-\in_{11}B_3)]\left[\frac{r}{\sqrt{r_1r_2}}\cos\left(\theta-\frac{\theta_1+\theta_2}{2}\right)\right]$$

$$+(p_zA_1-q_zB_1)\left[\frac{1}{\sqrt{r_1r_2}}\cos\left(\frac{\theta_1+\theta_2}{2}\right)-\frac{r}{2r_1r_2}\{\cos(\theta-\theta_1-\theta_2)-\cos\theta\}\right]$$

$$-(p_zA_2-q_zB_2)\left[-\frac{m}{\sqrt{r_1r_2}}\sin\left(\frac{\theta_1+\theta_2}{2}\right)-\frac{r}{2r_1r_2}\left\{\sin(\theta-\theta_1-\theta_2)-\cos\theta\tan\left(\frac{\theta_1+\theta_2}{2}\right)\right\}\right]$$

$$-\frac{2}{a}(p_zA_3-q_zB_3)\left[\frac{r}{\sqrt{r_1r_2}}\sin\left(\theta-\frac{\theta_1+\theta_2}{2}\right)\right], \tag{29}$$

$$F_y=-[b_z(c_{44}A_1+e_{15}B_1)+\Delta\phi(e_{15}A_1-\in_{11}B_1)]\left[\frac{1}{\sqrt{r_1r_2}}\cos\left(\frac{\theta_1+\theta_2}{2}\right)-\frac{r}{2r_1r_2}\{\cos(\theta-\theta_1-\theta_2)-\cos\theta\}\right]$$

$$+[b_z(c_{44}A_2+e_{15}B_2)+\Delta\phi(e_{15}A_2-\in_{11}B_2)]\left[-\frac{m}{\sqrt{r_1r_2}}\sin\left(\frac{\theta_1+\theta_2}{2}\right)-\frac{r}{2r_1r_2}\left\{\sin(\theta-\theta_1-\theta_2)-\cos\theta\tan\left(\frac{\theta_1+\theta_2}{2}\right)\right\}\right]$$

$$+\frac{2}{a}[b_z(c_{44}A_3+e_{15}B_3)+\Delta\phi(e_{15}A_3-\in_{11}B_3)]\left[\frac{r}{\sqrt{r_1r_2}}\sin\left(\theta-\frac{\theta_1+\theta_2}{2}\right)\right]$$

$$-(p_zA_1-q_zB_1)\left[-\frac{1}{\sqrt{r_1r_2}}\sin\left(\frac{\theta_1+\theta_2}{2}\right)-\frac{r}{2r_1r_2}\left\{\sin(\theta-\theta_1-\theta_2)+\cos\theta\tan\left(\frac{\theta_1+\theta_2}{2}\right)\right\}\right]$$

$$-(p_zA_2-q_zB_2)\left[\frac{m}{\sqrt{r_1r_2}}\cos\left(\frac{\theta_1+\theta_2}{2}\right)-\frac{r}{2r_1r_2}\{\cos(\theta-\theta_1-\theta_2)+\cos\theta\}\right]-\frac{2}{a}(p_zA_3-q_zB_3)\left[\frac{r}{\sqrt{r_1r_2}}\cos\left(\theta-\frac{\theta_1+\theta_2}{2}\right)\right]. \tag{30}$$

Equations (29) and (30) give the solution for a screw dislocation located in arbitrary position around a crack.

Zhang and Li [16] presented the x-direction force on a screw dislocation on the x-axis for an elastic material without any external load. If all the electrical quantities and the line and external loads vanish from Eq. (29), it agrees with their solution.

Field Intensity Factors and Energy Release Rate

Extending the traditional concept of stress intensity factor K^σ to other field variables, we can introduce the strain intensity factor K^ε, the electric field intensity factor K^E, and the electric displacement intensity factor K^D. These field intensity factors can be obtained by Eqs. (18)–(21), in the case of a screw dislocation located in arbitrary position around a crack, in the forms

$$K^\varepsilon(\pm a)=\lim_{Z\to\pm a}\sqrt{\pm 2\pi(Z\mp a)}\,\gamma_{zy}(x,y)$$

$$=\sqrt{\frac{\pi}{a}}\left[A_1\left(\mathrm{Im}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\right)-A_2\left(m-\mathrm{Re}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\right)\mp 2A_3\right], \tag{31}$$

$$K^E(\pm a)=\lim_{Z\to\pm a}\sqrt{\pm 2\pi(Z\mp a)}E_y(x,y)$$

$$=\sqrt{\frac{\pi}{a}}\left[-B_1\left(\mathrm{Im}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\right)+B_2\left(m-\mathrm{Re}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\right)\mp 2B_3\right], \tag{32}$$

$$K^\sigma(\pm a)=\lim_{Z\to\pm a}\sqrt{\pm 2\pi(Z\mp a)}\,\sigma_{zy}(x,y)$$

$$=\sqrt{\frac{\pi}{a}}\left[(c_{44}A_1+e_{15}B_1)\left(\mathrm{Im}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\right)-(c_{44}A_2+e_{15}B_2)\left(m-\mathrm{Re}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\right)\mp 2(c_{44}A_3+e_{15}B_3)\right], \tag{33}$$

$$K^D(\pm a)=\lim_{Z\to\pm a}\sqrt{\pm 2\pi(Z\mp a)}D_y(x,y)$$
$$=\sqrt{\frac{\pi}{a}}\Bigg[(e_{15}A_1-\in_{11}B_1)\Bigg(\mathrm{Im}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\Bigg)$$
$$-(e_{15}A_2-\in_{11}B_2)\Bigg(m-\mathrm{Re}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\Bigg)$$
$$\mp 2(e_{15}A_3-\in_{11}B_3)\Bigg],\qquad(34)$$

where

$$\mathrm{Re}\sqrt{\frac{Z_0+a}{Z_0-a}}=\sqrt{\frac{r_2}{r_1}}\cos\Bigg(\frac{\theta_2-\theta_1}{2}\Bigg),$$
$$\mathrm{Im}\sqrt{\frac{Z_0+a}{Z_0-a}}=\sqrt{\frac{r_2}{r_1}}\sin\Bigg(\frac{\theta_2-\theta_1}{2}\Bigg),\qquad(35)$$
$$\mathrm{Re}\sqrt{\frac{Z_0-a}{Z_0+a}}=\sqrt{\frac{r_1}{r_2}}\cos\Bigg(\frac{\theta_1-\theta_2}{2}\Bigg),$$
$$\mathrm{Im}\sqrt{\frac{Z_0-a}{Z_0+a}}=\sqrt{\frac{r_1}{r_2}}\sin\Bigg(\frac{\theta_1-\theta_2}{2}\Bigg).\qquad(36)$$

As shown in Eqs. (31)–(34), the intensity factors are divided into three terms: the first is by the line loads, the second by the dislocation, and the last by the external loads.

If only far-field external loading conditions are applied, Eqs. (31)–(34) reduce to Pak's result ([3]). Also, if all the electrical quantities, the line loads and the far-field loading conditions are eliminated from Eq. (33), it is reduced to the solution of the purely elastic case proposed by Zhang and Li [16].

The energy release rate G for piezoelectric materials under the antiplane shear case can be evaluated by the generalized path-independent J-integral ([3]) in the form

$$J=G=\frac{K^{\varepsilon}K^{\sigma}-K^{E}K^{D}}{2}.\qquad(37)$$

For a screw dislocation located in an arbitrary position around a crack, the energy release rates at both crack tips can be derived by substituting the field intensity factors into Eq. (37), as follows:

$$G(\pm a)=\frac{\pi}{2a}\Bigg[(c_{44}A_1^2+2e_{15}A_1B_1-\in_{11}B_1^2)\Bigg(\mathrm{Im}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\Bigg)^2$$
$$-2(c_{44}A_1A_2+e_{15}A_1B_2+e_{15}A_2B_1$$
$$-\in_{11}B_1B_2)\mathrm{Im}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\Bigg(m-\mathrm{Re}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\Bigg)$$
$$\mp 4(c_{44}A_1A_3+e_{15}A_1B_3+e_{15}B_1A_3$$
$$-\in_{11}B_1B_3)\mathrm{Im}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}+(c_{44}A_2^2+2e_{15}A_2B_2$$
$$-\in_{11}B_2^2)\Bigg(m-\mathrm{Re}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\Bigg)^2\pm 4(c_{44}A_2A_3+e_{15}A_2B_3$$
$$+e_{15}B_2A_3-\in_{11}B_2B_3)\Bigg(m-\mathrm{Re}\sqrt{\frac{Z_0\pm a}{Z_0\mp a}}\Bigg)$$
$$+4(c_{44}A_3^2+2e_{15}A_3B_3-\in_{11}B_3^2)\Bigg].\qquad(38)$$

The energy release rates for four possible boundary conditions can be, respectively, determined by inserting Eqs. (22)–(27) into Eq. (38).

Table 1 Material properties of PZT-5H ceramic

Elastic modulus	c_{44}	2.3×10^{10} (N/m^2)
Piezoelectric constant	e_{15}	17.0 (C/m^2)
Dielectric permittivity	$\in_{11}$	150.4×10^{-10} (C/Vm)
Critical energy release rate	G_{cr}	5.0

Table 2 Basic values for numerical example

Burgers vector	b_z	1.0×10^{-9} (m)
Electric potential jump	$\Delta\phi$	1.0 (V)
Line force	p_z	10 (N/m)
Line charge	q_z	1.0×10^{-8} (C/m)

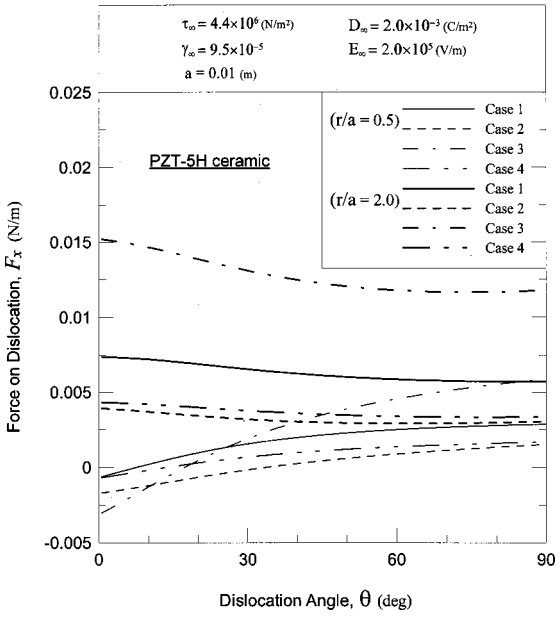


Fig. 2 Force on a screw dislocation, F_x versus dislocation angle, θ

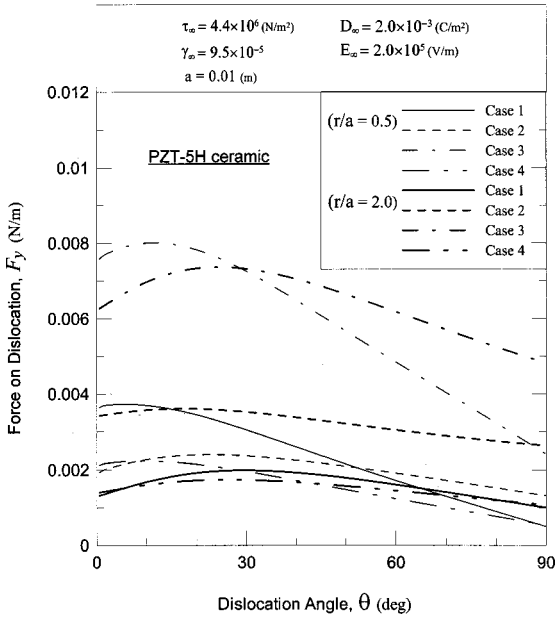


Fig. 3 Force on a screw dislocation, F_y versus dislocation angle, θ

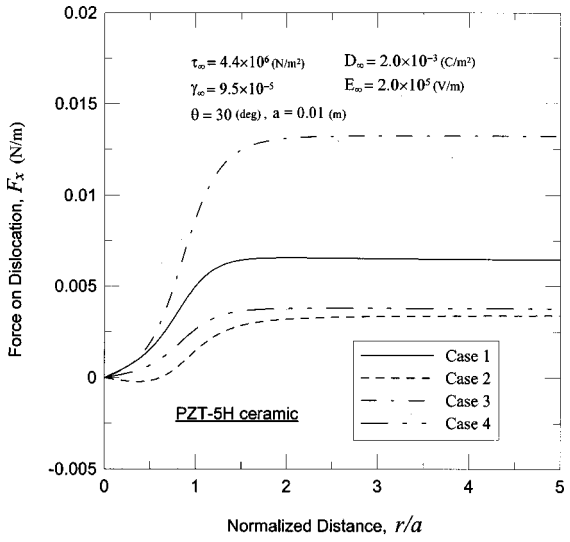


Fig. 4 Force on a screw dislocation, F_x versus normalized distance, r/a

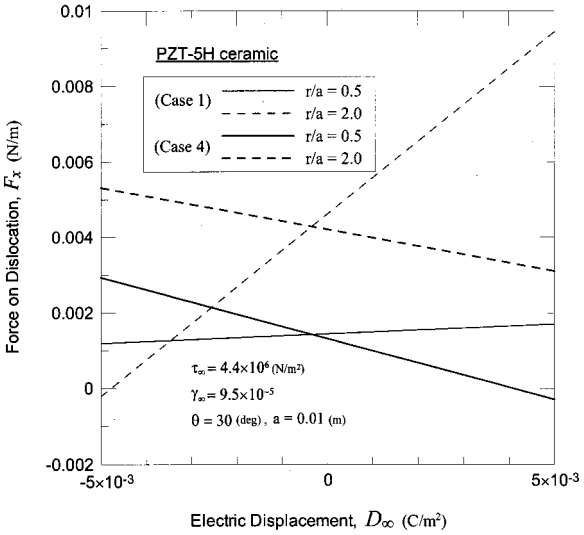


Fig. 6 Force on a screw dislocation, F_x versus electric displacement, D_∞

If the line force, line charge, and screw dislocation are eliminated from Eq. (38), it agrees with the solution for the purely external loading case of Pak [3].

Discussion

A lead zirconate titanate piezoceramic ceramic of PZT-5H is considered for the numerical analyses, and its material properties are given in Table 1. Table 2 presents the values of the Burgers vector, electric potential jump, line force, and line charge applied for numerical examples. The effect of extra image dislocations is observed in the case $m=0$ in which a real screw dislocation is emitted from the crack. Because of the symmetry in geometry and loading, it is sufficient to consider only the case that a screw dislocation exists near the right tip of a crack in the first quarter of the rectangular coordinate system of Fig. 1.

Figures 2 and 3 show the forces acting on a screw dislocation as a function of the dislocation angle θ. For all cases of loading conditions, the force in the x-direction F_x increases with the increase of angle for $r/a=0.5$, but decreases for $r/a=2.0$. In contrast, the force in the y-direction F_y increases first and then decreases with the increase of angle regardless of the normalized distance r/a.

Figures 4 and 5 show the forces acting on a screw dislocation as a function of the normalized distance r/a. F_x increases with the increase of normalized distance in $0<r/a<1.5$ approximately in spite of some exceptional trend of Case 2, and is little affected by the change of normalized distance for $r/a>1.5$. F_y increases up to $r/a=1.0$ and then decreases with the increase of r/a.

Figures 6 and 7 show the forces acting on a screw dislocation as a function of the electric displacement D_∞. In Case 1 subjected to shear stress as the mechanical load, F_x and F_y are proportional to the electric displacement. However, in Case 4 subjected to shear strain, F_x is inversely proportional to the electric displacement. Similar trends are obtained for Cases 2 and 3 under the electric field E_∞.

Figure 8 shows the normalized energy release rates $G(\pm a)/G_{cr}$ as a function of the dislocation angle θ under Case 1. The normalized energy release rate at the right crack tip $G(a)/G_{cr}$ shows the different results for $r/a=0.5$ and $r/a=2.0$, because of

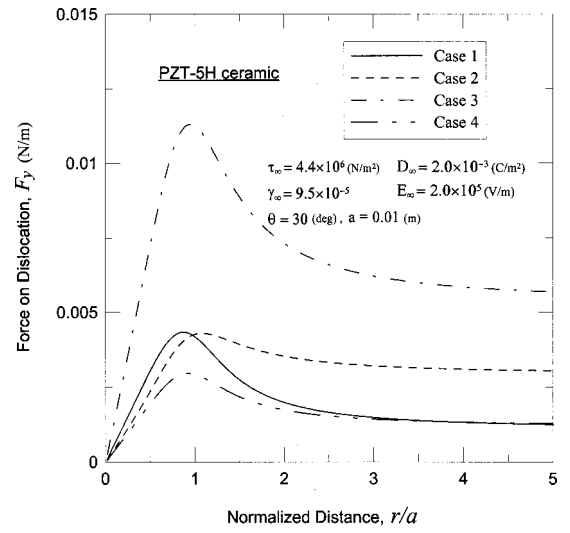


Fig. 5 Force on a screw dislocation, F_y versus normalized distance, r/a

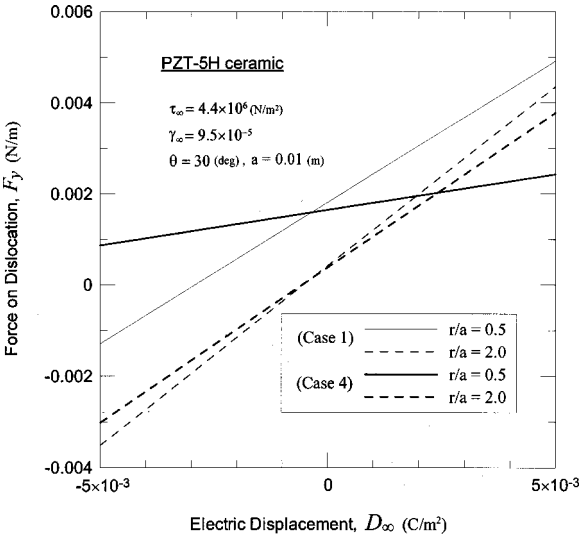


Fig. 7 Force on a screw dislocation, F_y versus electric displacement, D_∞

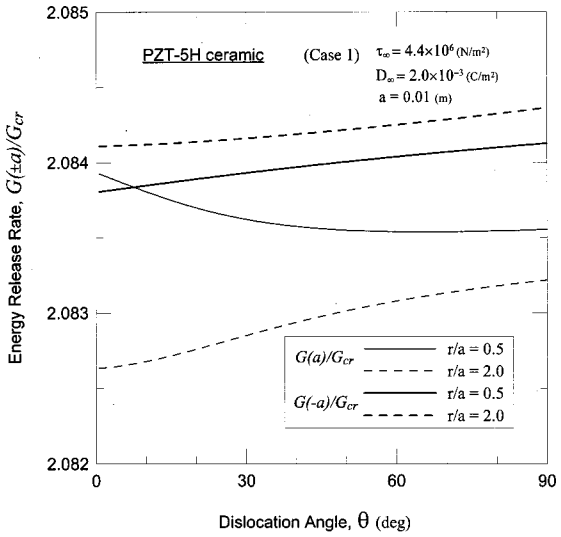


Fig. 8 Normalized energy release rate, $G(\pm a)/G_{cr}$ versus dislocation angle, θ

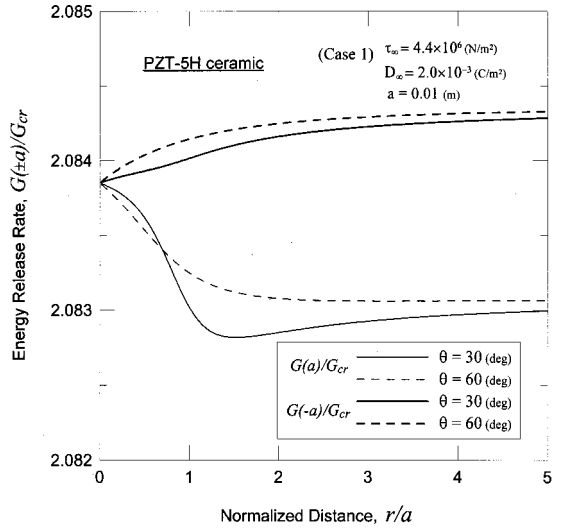


Fig. 9 Normalized energy release rate, $G(\pm a)/G_{cr}$ versus normalized distance, r/a

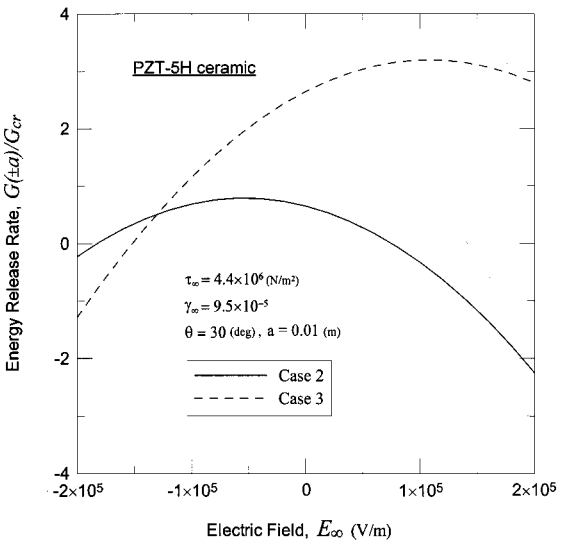


Fig. 10 Normalized energy release rate, $G(\pm a)/G_{cr}$ versus electric field, E_∞

the effect of a screw dislocation. On the other hand, $G(-a)/G_{cr}$ at the left crack tip increases with the increase of angle for both $r/a=0.5$ and 2.0.

Figure 9 shows the normalized energy release rates $G(\pm a)/G_{cr}$ as a function of the normalized distance r/a under Case 1. Generally, $G(a)/G_{cr}$ decreases with the increase of r/a, and $G(-a)/G_{cr}$ increases.

Figure 10 shows the normalized energy release rates $G(\pm a)/G_{cr}$ as a function of the electric field E_∞ under Cases 2 and 3. In this figure, the normalized energy release rates at both crack tips have almost similar values in the overall viewpoint. Here, it is noted that $G(\pm a)/G_{cr}$ may have negative values under the special electric field. This means that crack growth in a piezoelectric material can be arrested by changing the direction and magnitude of the electric field, because of the electromechanical coupling phenomenon.

The similar trends with Figs. 8–10 are obtained for other loading cases.

In the case of $m=1$, where the real screw dislocation originates elsewhere and comes near the crack, the results have the exactly same trend with the above figures in the case of $m=0$. The numerical values of $m=1$ are larger than those of $m=0$ a little, but the differences of values between two cases are too small to distinguish in the overall viewpoint.

Conclusion

A theoretical analysis was performed for a screw dislocation with a line force and a line charge around a finite crack in a hexagonal piezoelectric crystal under the far-field antiplane mechanical and in-plane electrical loads. The field variables near the crack tip and the forces acting on the dislocation were obtained by a complex variable and a conformal mapping technique. The field intensity factors and energy release rates at both crack tips were also determined as a function of the location of a screw dislocation subjected to line and external loads. The results were verified by comparing with previous works. The numerical analyses presented the interaction between the defects by showing that the forces acting on the dislocation and the energy release rates are affected by the position of a screw dislocation and loading conditions.

Acknowledgments

This work was supported by the Brain Korea 21 Project.

References

[1] Faivre, G., and Saada, G., 1972, ''Dislocations in Piezoelectric Semiconductors,'' Phys. Status Solidi B, **52**, p. 127.
[2] Deeg, W. F., 1980, ''The Analysis of Dislocation, Crack, and Inclusion Problems in Piezoelectric Solids,'' Ph.D. thesis, Stanford University, Stanford, CA.
[3] Pak, Y. E., 1990, ''Crack Extension Force in a Piezoelectric Material,'' ASME J. Appl. Mech., **57**, pp. 647–653.
[4] Pak, Y. E., 1990, ''Force on a Piezoelectric Screw Dislocation,'' ASME J. Appl. Mech., **57**, pp. 863–869.
[5] Wang, B., 1992, ''Three-Dimensional Analysis of an Ellipsoidal Inclusion in Piezoelectric Material,'' Int. J. Solids Struct., **29**, pp. 293–308.
[6] Chen, T., 1993, ''The Rotation of a Rigid Inclusion Embedded in an Anisotropic Piezoelectric Medium,'' Int. J. Solids Struct., **30**, pp. 1983–1995.
[7] Liang, J., Han, J. -C., Wang, B., and Du, S., 1995, ''Electroelastic Modelling of Anisotropic Piezoelectric Materials With an Elliptic Inclusion,'' Int. J. Solids Struct., **32**, pp. 2989–3000.
[8] Chung, M. Y., and Ting, T. C. T., 1996, ''Piezoelectric Solid With an Elliptical Inclusion or Hole,'' Int. J. Solids Struct., **33**, pp. 3343–3361.
[9] Zhong, Z., and Meguid, S. A., 1997, ''Interfacial Debonding of a Circular Inhomogeneity in Piezoelectric Materials,'' Int. J. Solids Struct., **34**, pp. 1965–1984.
[10] Meguid, S. A., and Deng, Wei, 1998, ''Electro-electic Interaction Between a Screw Dislocation and an Elliptical Inhomogeneity in Piezoelectric Materials,'' Int. J. Solids Struct., **35**, pp. 1467–1482.
[11] Lee, K. Y., Lee, W. G., and Pak, Y. E., 2000, ''Interaction Between a Semi-infinite Crack and a Screw Dislocation in a Piezoelectric Material,'' ASME J. Appl. Mech., **67**, pp. 165–170.
[12] Hirth, J. P., and Lothe, J., 1982, *Theory of Dislocations*, John Wiley and Sons, New York, p. 436.
[13] Berlincourt, D. A., Curran, D. R., and Jaffe, H., 1964, ''Piezoelectric and

Piezomagnetic Materials and Their Function in Transducers,'' W. P. Mason, ed., Academic Press, New York, Phys. Acoust., **1A**, p. 177.
[14] Majumdar, B. S., and Burns, S. J., 1981, ''Crack Tip Shielding—An Elastic Theory of Dislocations and Dislocation Arrays Near a Sharp Crack,'' Acta Metall., **29**, pp. 579–588.
[15] Lee, S., 1987, ''The Image Force on the Screw Dislocation Around a Crack of Finite Size,'' Eng. Fract. Mech., **27**, No. 5, pp. 539–545.
[16] Zhang, T.-Y., and Li, J. C. M., 1991, ''Image Forces and Shielding Effects of a Screw Dislocation Near a Finite-Length Crack,'' Mater. Sci. Eng., A, **142**, pp. 35–39.
[17] Eshelby, J. D., 1956, ''The Continuum Theory of Lattice Defects,'' Solid State Phys., **3**, p. 79.

A Critical Reexamination of Classical Metal Plasticity

C. D. Wilson
Department of Mechanical Engineering,
Tennessee Technological University,
Cookeville, TN 38505-0001
e-mail: chriswilson@tntech.edu

P. W. Bridgman's early work on flow and fracture in the presence of hydrostatic pressure showed no systematic effect on strain hardening. This experimental observation led to the conclusions that yielding does not depend on hydrostatic stress and that the yielded material is incompressible. Classical plasticity theory was largely built on these observations. New experiments and nonlinear finite element analyses of 2024-T351 aluminum notched round bars has quantified the effect of hydrostatic tensile stresses on yielding. Nonlinear finite element analyses using the von Mises (yielding is independent of hydrostatic stress) and the Drucker-Prager (yielding is linearly dependent on hydrostatic stress) yield functions was performed. The von Mises results overestimated experimental load-displacement curves by 10–65 percent. The Drucker-Prager results essentially matched the experimental results. The only additional data requirement for the Drucker-Prager yield function is the compressive yield strength. [DOI: 10.1115/1.1412239]

Introduction

Bridgman and Classical Metal Plasticity. In his book, *Studies in Large Plastic Flow and Fracture*, P. W. Bridgman [1] summarized his work on plastic flow and fracture in the presence of hydrostatic pressure. Bridgman's research was motivated by the discovery that the ductility of mild steels increases greatly when exposed to hydrostatic pressures greater than 300,000 psi. During World War II, the National Defense Research Committee and later, the Watertown Arsenal, funded an extensive investigation of the effect of hydrostatic pressure on the ductility of ballistic steels. Bridgman's earlier work ([2]) showed no systematic effect of pressure on strain hardening. However, his later work was characterized by more precise measurements which established a definite effect of hydrostatic pressure on the strain hardening curves of mild steel. His experiments were essentially tensile tests on smooth specimens in a pressure chamber.

Under the conditions of hydrostatic pressures up to 3100 MPa (450,000 psi), Bridgman found that a major effect of hydrostatic pressure was increased ductility. In other words, much larger strains before fracture were obtained when hydrostatic pressure was applied as a boundary condition to a tensile test. Additionally, Bridgman found that the material volume in the gage section did not change for very large plastic strain changes. Therefore, a metal was assumed to have incompressible plastic strains. These two experimental observations about metals—no influence of hydrostatic pressure on yielding and incompressibility for plastic strain changes—are two of the basic tenets of classical metal plasticity.

Bridgman's tests were conducted on unnotched tensile bars with pressures exceeding 100,000 psi. Such externally applied pressure levels are rarely seen in typical applications, leading many researchers to assume that the effect of all hydrostatic stress—both tensile and compressive—is negligible. Generations of material scientists and engineers have studied classical metal plasticity based on the postulate that yield is not affected by hydrostatic stress. However, notched components loaded in tension or bending can internally develop large hydrostatic stresses.

Generations of material scientists and engineers have studied classical metal plasticity based on Bridgman's two observations. Plasticity textbooks by Hill [3] in 1950 and Mendelson [4] in 1968 are examples of classical plasticity. Even modern treatments on plasticity, such as Lubliner [5] in 1990 and Stouffer and Dame [6] in 1996 assume that there is no hydrostatic pressure dependence for yield and incompressibility as fundamental assumptions for metal plasticity. These assumptions are also the basis for metal plasticity in commercial finite element programs. For example, the theory manual of ABAQUS [7] directs potential users to use a yield criterion for metals that has no dependence on hydrostatic stress.

It is important to examine hydrostatic stress and incompressibility in mathematical terms. To simplify the discussion, assume that the principal stresses are given by σ_1, σ_2, and σ_3. The principal stresses are the roots of the stress cubic

$$\sigma^3 - I_1\sigma^2 - I_2\sigma - I_3 = 0, \tag{1}$$

where σ is a principal stress and I_1, I_2, and I_3 are functions of the stress state called the stress invariants. In terms of principal stresses, the stress invariants are

$$I_1 = \sigma_1 + \sigma_2 + \sigma_3$$
$$I_2 = -(\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_1)$$
$$I_3 = \sigma_1\sigma_2\sigma_3. \tag{2}$$

The hydrostatic or mean stress σ_m is defined as $\sigma_m = I_1/3$ and the hydrostatic pressure p is defined as $p = -\sigma_m$.

Bridgman's first observation that the hydrostatic pressure had no effect on the yield behavior of metals until very large hydrostatic pressures led engineers to develop a plasticity theory that subtracts the mean stress from the principal stresses. The resulting stresses are called the deviatoric stresses S_1, S_2, and S_3, and are written as

$$S_1 = \sigma_1 - \sigma_m$$
$$S_2 = \sigma_2 - \sigma_m$$
$$S_3 = \sigma_3 - \sigma_m. \tag{3}$$

The deviatoric stress invariants are the coefficients J_1, J_2, and J_3 of the cubic

$$S^3 - J_1S^2 - J_2S - J_3 = 0 \tag{4}$$

where the invariants are functions of the principal stresses and the mean stress

$$J_1 = 0,$$

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, March 1, 2001; final revision, June 12, 2001. Associate Editor: M. Ortiz. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

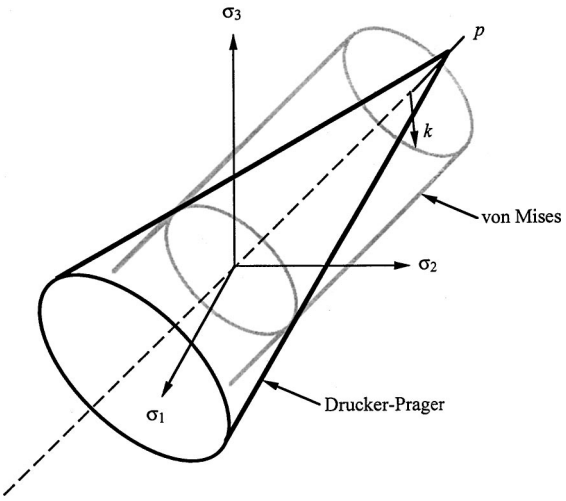


Fig. 1 von Mises and Drucker-Prager yield surfaces

$$J_2=\frac{1}{6}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2],$$

$$J_3=(\sigma_1-\sigma_m)(\sigma_2-\sigma_m)(\sigma_3-\sigma_m). \tag{5}$$

The deviatoric stress invariants can also be written in terms of the deviatoric stresses

$$J_2=\frac{1}{2}(S_1^2+S_2^2+S_3^2),$$

$$J_3=S_1S_2S_3. \tag{6}$$

For classical metal plasticity, a yield function is a function $f=f(\sigma_1,\sigma_2,\sigma_3)$ such that when $f<0$ the material behavior is elastic. For $f=0$, the material behavior is plastic and yielding occurs. Assuming that yield is independent of the hydrostatic stress leads to a yield function $f=f(J_2,J_3)$. The most commonly used yield function, the von Mises yield function, assumes further that yielding is not a function of J_3

$$f=\sqrt{J_2}-k, \tag{7}$$

where k is the yield strength in pure shear. For hardening materials, k is a function of plastic strain. The square root of $3J_2$ is the von Mises or effective stress and can be written in terms of the principal stresses

$$\sigma_{\text{eff}}=\sqrt{3J_2}=\sqrt{\frac{1}{2}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2]}. \tag{8}$$

For $f=0$, Eq. (7) can be written as $\sigma_{\text{eff}}=\sigma_0$, where σ_0 is the yield strength in pure tension $\sigma_0=\sqrt{3}k$.

Rewritting Eq. (7) for $f=0$ in the form $J_2=k^2$ leads to a graphical interpretation of the yield function in the principal stress space. For the von Mises yield function, the yield surface is a circular cylinder of radius k whose axis is defined in the direction of the hydrostatic stress. Any cross section taken perpendicular to the cylinder axis will be a circle of radius k. A von Mises yield surface is shown in Fig. 1.

The discussion thus far has focused on the initial yield surface, where a material first starts to yield. Many materials exhibit strain hardening, a process where the yield surface changes shape or location or both as the material is plastically deformed. For many metals, isotropic hardening is an appropriate approximation. Essentially, isotropic hardening means that the yield surface expands equally in all directions. For the von Mises yield function, the radius k of the yield surface grows larger.

Isotropic hardening implies that the tensile and compressive yield behaviors are the same. While this is approximately true for some materials, it is not true in general. Some materials exhibit different magnitudes of yield strength for tension than for compression. This effect is called the Bauschinger effect and becomes important whenever stresses are to be predicted after a reversal of loading. Kinematic hardening is often used to describe materials with a pronounced Bauschinger effect. For kinematic hardening, the new yield surface is assumed to have the same radius k as the initial yield surface. However, the axis of the yield surface shifts in principal stress space.

Different metals have differing amounts of Bauschinger effect. Purely isotropic hardening represents an extreme of no Bauschinger effect. Kinematic hardening represents the other extreme. However, a linear combination of the two models is useful in describing real materials. In practice, isotropic hardening is easier to implement and is more often used.

The second basic tenet of classical metal plasticity is incompressibility. It was observed that volume change during plastic deformation is nearly elastic. In terms of principal strain increments $d\varepsilon_1$, $d\varepsilon_2$, and $d\varepsilon_3$, the sum of the plastic strain increments $d\varepsilon_{ii}^{pl}$ (or plastic dilation rate) must be zero. Mathematically, this condition is

$$d\varepsilon_{ii}^{pl}=d\varepsilon_1^{pl}+d\varepsilon_2^{pl}+d\varepsilon_3^{pl}=0. \tag{9}$$

The relationship between plastic strain increments and the yield condition is given by an associated flow rule. Thus, a flow rule governs the postyield behavior of a material. The general form of an associated flow rule is

$$d\varepsilon_i^{pl}=d\lambda\frac{\partial f}{\partial\sigma_i}, \tag{10}$$

where $d\lambda$ is a positive scalar. This type of flow rule is also called a normality flow rule because it assumes that the strain increment is normal to the yield surface. Drucker and Prager [8] showed that the plastic dilation rate $d\varepsilon_{ii}^{pl}$ can be summed from Eq. (10) to obtain

$$d\varepsilon_{ii}^{pl}=3d\lambda\frac{\partial f}{\partial I_1}. \tag{11}$$

Noting that the von Mises yield function is not a function of I_1 leads to the result in Eq. (9). In other words, the von Mises yield function f does not depend on hydrostatic stress and the plastic dilation rate $d\varepsilon_{ii}^{pl}$ must be zero. Thus, the two observations of Bridgman and the two tenets of classical plasticity are closely related.

Richmond and Yield Dependence on Hydrostatic Pressure. Although classical metal plasticity has a great deal of inertia in engineering practice, the basic tenets were challenged in the 1970s by the experiments of Richmond, Spitzig, and Sober ([9,10]). They studied the effects of hydrostatic pressure up to 1,100 MPa (160 kpsi) on the yield strength of four steels (4330, 4310, maraging steel, and HY80). These pressure levels were significantly less than the maximum pressures used by Bridgman years earlier.

Richmond found that the yield strength was a linear function of hydrostatic pressure. For high-strength steels, Richmond found that a yield function identical to one proposed by Drucker and Prager [8] for soils described the yielding process. The Drucker-Prager yield function is

$$\sigma_{\text{eff}}=\sigma_0(\varepsilon_{pl})-aI_1, \tag{12}$$

where a is a material constant (referred to in this paper as the Drucker-Prager constant) related to the theoretical cohesive strength of the material σ_c. The theoretical cohesive strength is the stress required to overcome cohesive forces between neighboring atoms. The cohesive strength can be expressed as a fraction of Young's modulus E. Dieter [11] gives the range of theoretical

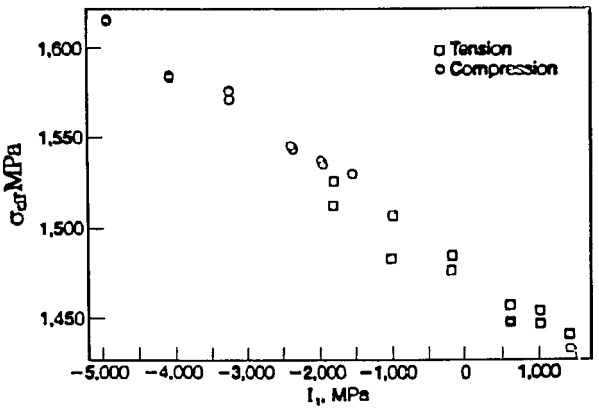


Fig. 2 Effective Stress σ_{eff} versus I_1 for 4330 steel ([9])

cohesive strength for metals as $E/15$ to $E/4$ with a typical value of $E/5.5$. As with the von Mises yield function, σ_0 is the yield strength for pure tension and is a function of plastic strain ε_{pl}.

Graphically, a can be interpreted as the slope of the graph of σ_{eff} versus I_1. The value of σ_{eff} for $I_1=S_y$ corresponds to the typically reported yield strength for a tensile test. The value of I_1 for $\sigma_{\text{eff}}=0$ corresponds to the theoretical cohesive strength of the material. For the 4330 steel shown in Fig. 2, the yield strength for pure tension was 1475 MPa and the cohesive strength was 59,000 MPa. For E of 200,000 MPa, the theoretical cohesive strength was approximately $0.3E$.

Determining the theoretical cohesive strength is a difficult task; therefore, another method is needed for determining the Drucker-Prager constant a. If the tensile yield strength S_y and the compressive yield strength S_{yc} are known, a can be calculated by

$$a=\frac{S_{yc}-S_y}{S_{yc}+S_y}. \tag{13}$$

In terms of a yield surface, the Drucker-Prager yield function is a cone whose apex is at a hydrostatic stress equal to the cohesive strength. This is shown schematically in Fig. 1. For small to moderate amounts of hydrostatic stress, this cone would locally be approximated by a cylinder. For this situation, the von Mises yield function would give comparable results to the Drucker-Prager yield function. However, for high levels of hydrostatic stress, the Drucker-Prager yield function is preferable to von Mises. However, the Drucker-Prager yield function is rarely used for metal plasticity.

Table 1 Room temperature material properties for 2024-T351 aluminum alloy

E (psi)	10.4×10^6
ν	0.33
S_y (kpsi)	55.7
S_{ut} (kpsi)	69
n	15
S_{yc} (kpsi)	59.1
a	0.03

The Drucker-Prager yield function in Eq. (12) can be written as

$$f=\sqrt{J_2}+\alpha I_1-k, \tag{14}$$

where α is a constant related to the Drucker-Prager constant a. The plastic dilation rate for the Drucker-Prager yield function in Eq. (14) is no longer zero since f depends on I_1. Thus, the assumption of incompressibility in classical metal plasticity is no longer true and $d\varepsilon_{ii}^{pl}=3d\lambda\alpha$.

Research Program

A combined experimental and analytical research program was designed to study the hydrostatic tensile stress effects on the room temperature yield behavior of 2024-T351 aluminum alloy ([12]). Two geometries were studied: a smooth tensile bar with a round cross section and a notched round bar (NRB).

Experimental Program. Tensile tests on smooth, unnotched round bars were conducted for each material with the resulting tensile strength and hardening properties (0.2 percent offset yield strength S_y, ultimate strength S_{ul}, and Ramberg-Osgood hardening exponent n) given in Table 1. The values for Young's modulus E and Poisson's ratio ν in Table 1 are typical handbook values. Compression tests were also conducted. The resulting compressive yield strength S_{yc} and the Drucker-Prager constant a are also given in Table 1.

The NRB specimen details are shown in Fig. 3. The aluminum alloy specimens had a nominal radius R of 0.25 in. and a neck radius r of 0.125 in., thus $r/R=0.5$. All specimens had an notch

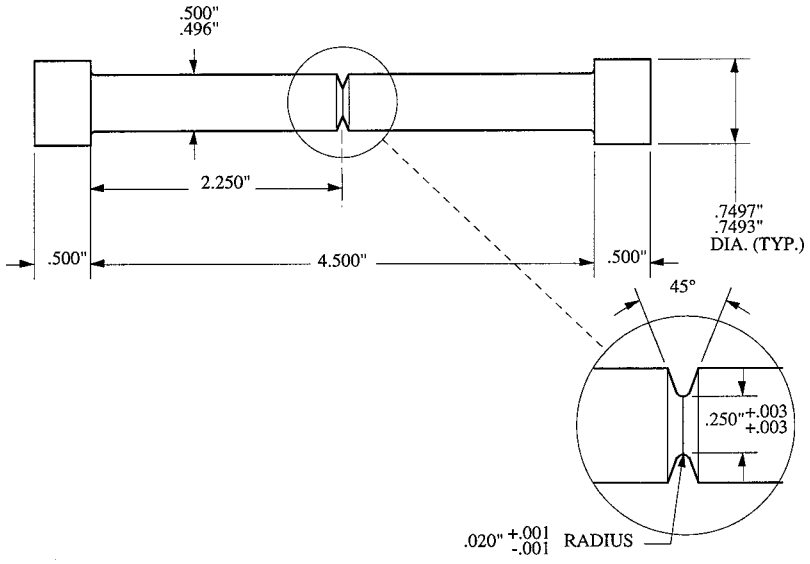


Fig. 3 Engineering drawing of the notched round bar (NRB) specimen

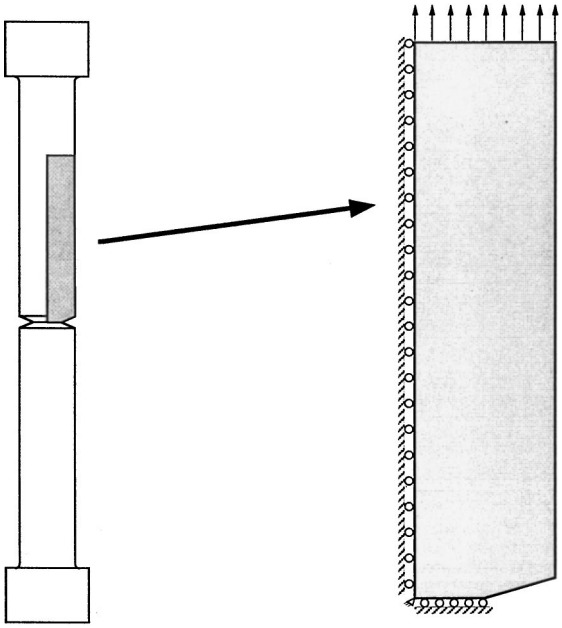


Fig. 4 Schematic of axisymmetric model of a notched round bar (NRB)

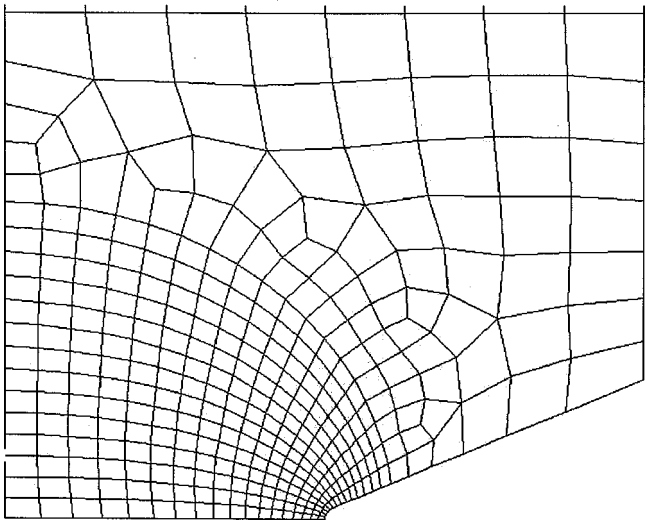


Fig. 5 Coarse mesh in the notch region

flank angle of 45 deg. The notch root radii ρ chosen were 0.005, 0.010, and 0.020 in. All notch root radii ρ had a tolerance of ± 0.001 in.

All tests were conducted on screw-driven universal testing machines in stroke control. An extensometer with a gage length of 0.4 in. and a ten percent extension range was used in all NRB tests. Load and gage displacement data were recorded with a *x-y* plotter and digitized afterward. No attempts were made to experimentally determine volume changes in the test specimens.

Analytical Program. Nonlinear finite element analyses were used to model the load-displacement response of the 2024-T351 aluminum specimens tested. The Sandia National Laboratory computer program FASTQ [13] was used for preprocessing of meshes and boundary conditions. The Q4 element type was used for all meshes. Axisymmetric models of notched round bars were developed with FASTQ in three levels of mesh refinement in the notch region. A schematic showing the region of the notched round bar actually modeled is given in Fig. 4. A coarse mesh typically had 250 elements in the notch region. Medium and fine meshes typically had 500 and 1000 elements, respectively, in the notch region. An example of a coarse mesh is given in Fig. 5.

A commercial finite element program, ABAQUS [7], was used for the finite element analyses and postprocessing of results. Large strain analysis and reduced integration options were used. The nonlinear material response was modeled using isotropic hardening with a von Mises yield criterion, and again with a Drucker-Prager yield criterion. Associated flow rules were assumed for both yield criteria. The true stress versus plastic strain data from the tensile test was used as input for the hardening curve.

Results and Conclusions

The experimental NRB load-displacement P-ν curves for 2024-T351 are shown in Fig. 6 for each notch root radius ([12]). Multiple specimens were used for each notch root radius, but a single representative load-displacement curve for each radius is plotted in Fig. 6. The P-ν records for all three values of ρ seem to follow a common curve with the smallest ρ failing at a lower load. It was

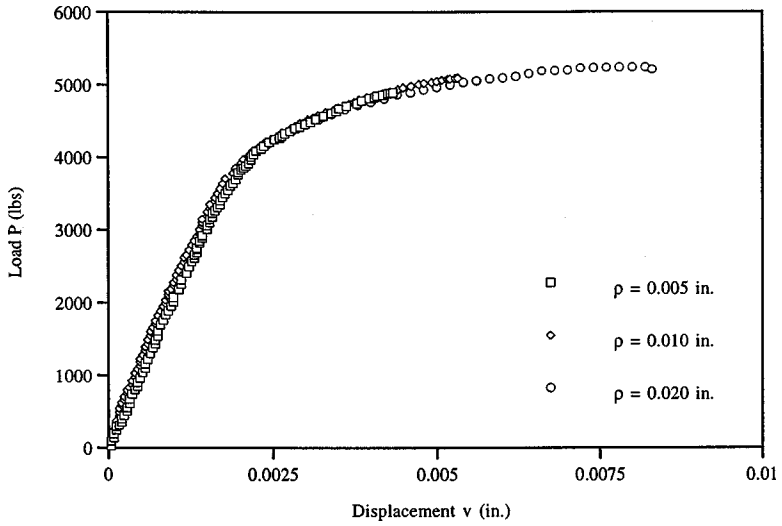


Fig. 6 Notched round bar (NRB) load-displacement results for all ρ for 2024-T351

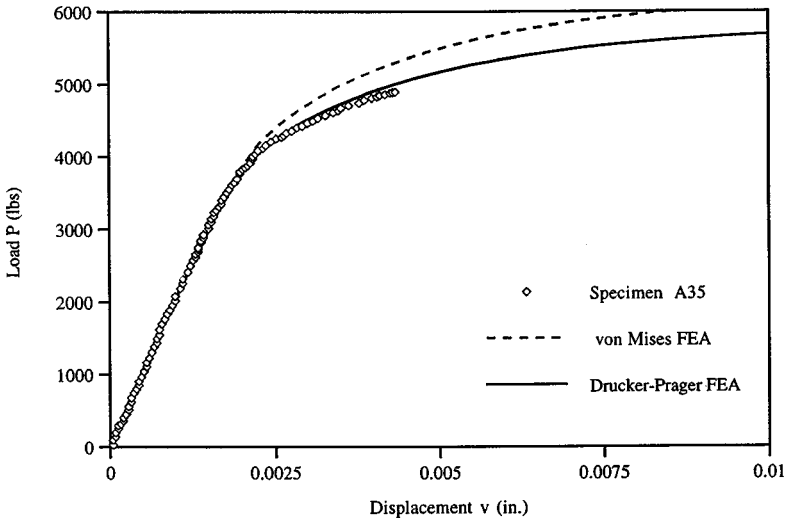


Fig. 7 Notched round bar (NRB) load-displacement results (ρ=0.005 in.) for 2024-T351

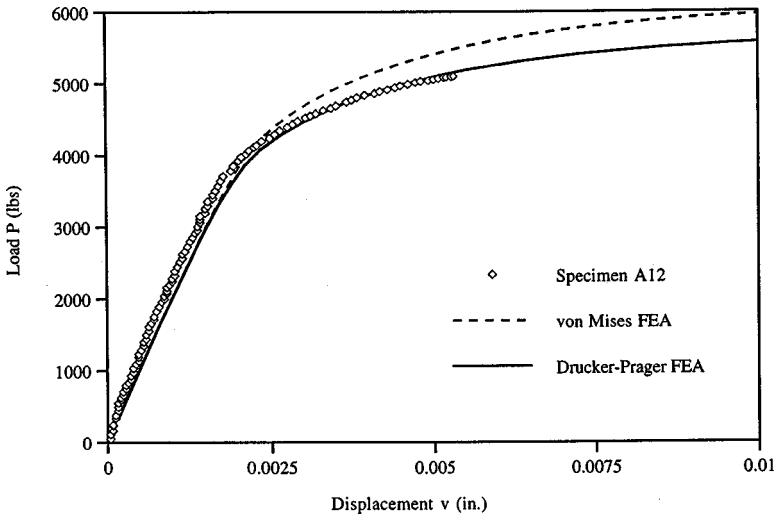


Fig. 8 Notched round bar (NRB) load-displacement results (ρ=0.010 in.) for 2024-T351

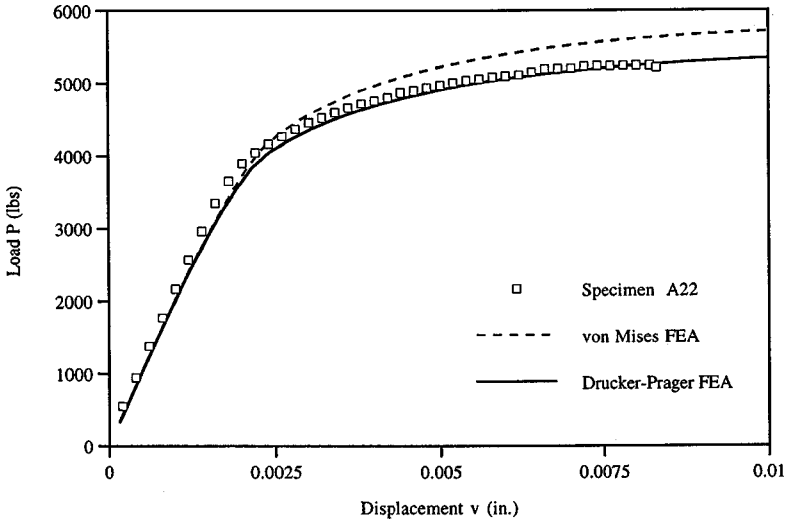


Fig. 9 Notched round bar (NRB) load-displacement results (ρ=0.020 in.) for 2024-T351

observed that all specimens failed before the load reached an instability point predicted by $dP=0$. No visible surface cracking was observed during the tests.

The finite element results using a von Mises yield function and a Drucker-Prager yield function are compared to the experimental results in Figs. 7–9. The von Mises yield function consistently overpredicts the actual load-displacement response of the NRB specimens. For the gage displacement at failure, the failure loads predicted using the Drucker-Prager yield function essentially matched the test data. The failure loads predicted using the von Mises yield function overestimated the experimental failure loads by approximately ten percent. For the load at failure, the failure displacements predicted using the Drucker-Prager yield function essentially matched the test data. The failure displacements predicted using the von Mises yield function overestimated the experimental failure displacements by 20 to 65 percent. Therefore, the Drucker-Prager yield function was consistently more accurate than the von Mises yield function.

The notched round bar is an axisymmetric geometry with a state of stress that is similar to plane strain. It develops a larger hydrostatic stress than a thin, flat double edge notched bar in a state of plane stress. Thus, the errors reported for notched round bars using the von Mises criterion for 2024-T351 are probably upperbound errors. The differences between the von Mises and Drucker-Prager yield criteria may be much smaller for plane stress conditions.

In terms of stress concentration factors (k_t), the notched round bars had k_t ranging from 2.70 (ρ of 0.020 in.) to 5.17 (ρ of 0.005 in.) using equations developed by Neuber [14]. For milder notches in notched round bars, the hydrostatic effect will be less pronounced. However, notches in nominally plane stress geometries may still have hydrostatic effects because the local notch conditions are more like plane strain than the far-field plane stress condition.

There is a small Bauschinger effect or strength-differential (SD) phenomenon observed in 2024-T351 (S_y of 55.7 kpsi versus S_{yc} of 59.1 kpsi). This effect appears to be adequately handled using the Drucker-Prager yield function. The good agreement between experimental and numerical results indicates that isotropic hardening, when coupled with the Drucker-Prager yield function, adequately describes the strain hardening of 2024-T351.

In summary, the yield behavior of 2024-T351 aluminum alloy is more accurately modeled using a yield function that includes a hydrostatic stress term. The Drucker-Prager yield criterion was shown to capture the hydrostatic tensile stress effects on yielding without introducing complications and additional expense. The only additional data requirement for the Drucker-Prager yield function is the yield strength in compression.

References

[1] Bridgman, P. W., 1952, *Studies in Large Plastic Flow and Fracture With Special Emphasis on the Effects of Hydrostatic Pressure*, McGraw-Hill, New York.

[2] Bridgman, P. W., 1947, ''The Effect of Hydrostatic Pressure on the Fracture of Brittle Substances,'' J. Appl. Phys., **18**, p. 246.

[3] Hill, R., 1950, *The Mathematical Theory of Plasticity*, Clarendon Press, Oxford, UK.

[4] Mendelson, A., 1968, *Plasticity: Theory and Application*, Macmillan, New York.

[5] Lubliner, J., 1990, *Plasticity Theory*, Macmillan, New York.

[6] Stouffer, D. C., and Dame, L. T., 1996, *Inelastic Deformation of Metals*, John Wiley and Sons, New York.

[7] ABAQUS Theory Manual, Version 5.5, 1995, Hibbitt, Karlsson, and Sorensen, Inc., 1995.

[8] Drucker, D. C., and Prager, W., 1952, ''Soil Mechanics and Plastic Analysis for Limit Design,'' Q. Appl. Math., **10**, pp. 157–165.

[9] Spitzig, W. A., Sober, R. J., and Richmond, O., 1976, ''The Effect of Hydrostatic Pressure on the Deformation Behavior of Maraging and HY-80 Steels and Its Implications for Plasticity Theory,'' Metall. Trans. A, **7A,** Nov., pp. 1703–1710.

[10] Richmond, O., and Spitzig, W. A., 1980, ''Pressure Dependence and Dilatancy of Plastic Flow,'' *International Union of Theoretical and Applied Mechanics Conference Proceedings*, pp. 377–386.

[11] Dieter, G. E., 1976, *Mechanical Metallurgy*, McGraw-Hill, New York.

[12] Wilson, C. D., August 1997, ''Fracture Toughness Testing With Notched Round Bars,'' Ph.D. dissertation, The University of Tennessee, Knoxville, TN.

[13] Blacker, T. D., June 1988, ''FASTQ Users Manual Version 1.2,'' Sandia Report SAND88-1326, UC-705, Sandia National Laboratories.

[14] Neuber, H., 1937, ''Theory of Notch Stresses,'' (translated from The German for The David Taylor Model Basin, U.S. Navy, by F. A. Raven, 1946).

L. Andersen
e-mail: la@civil.auc.dk

S. R. K. Nielsen
e-mail: soren.nielsen@civil.auc.dk

Department of Civil Engineering,
Aalborg University,
Sohngaardsholmsvej 57,
9000 Aalborg, Denmark

R. Iwankiewicz
School of Mechanical Engineering,
University of the Witwatersrand,
Johannesburg, Private Bag 3,
Wits, 2050, South Africa
e-mail: radiwank@hertz.mech.wits.ac.za

Vehicle Moving Along an Infinite Beam With Random Surface Irregularities on a Kelvin Foundation

The paper deals with the stochastic analysis of a single-degree-of-freedom vehicle moving at a constant velocity along an infinite Bernoulli-Euler beam with surface irregularities supported by a Kelvin foundation. Both the Bernoulli-Euler beam and the Kelvin foundation are assumed to be constant and deterministic. This also applies to the mass, spring stiffness, and damping coefficient of the vehicle. At first the equations of motion for the vehicle and beam are formulated in a coordinate system following the vehicle. The frequency response functions for the displacement of the vehicle and beam are determined for harmonically varying surface irregularities. Next, the surface irregularities are modeled as a random process. The variance response of the mass of the vehicle as well as the displacement variance of the beam under the oscillator are determined in terms of the autospectrum of the surface irregularities. [DOI: 10.1115/1.1427339]

1 Introduction

In structural dynamics it is common practice to model a road, runway, or railway track as a beam structure. Typically, a Kelvin-type foundation is used to describe an underlying material, and traffic is modeled as moving loads. Many papers deal with a deterministic analysis of the problem. However, in reality the parameters of the track structure as well as the loads are usually random variables. Thus a stochastic analysis of the problem has a significant interest, not only from an academic but also from a practical point of view.

Different aspects of uncertainties for the beam system and loads or, possibly, a vehicle moving along the structure have been examined in the literature. An analysis of a beam on a random Kelvin foundation due to a random excitation was performed by Sobczyk [1]. Frýba [2] gave an analytical solution to a random force moving at a constant velocity along a simply supported beam, and Zibdeh [3] analyzed the response of an axially loaded, simply supported beam for a random, moving load with time-varying velocity.

Iwankiewicz and Śniady [4] studied the behavior of a simply supported beam subjected to a stream of deterministic point forces moving at constant and equal velocity but with random interarrival time along the beam. Ricciardy [5] expanded the analysis to loads with random amplitudes, whereas Zibdeh and Rackwitz [6] studied the influence of loads with random velocities. Recently Śniady et al. [7] performed a study of the beam response for a loading process where the randomness of both the interarrival times and loading amplitudes as well as the velocities was taken into account.

Finite elements have been used for numerical stochastic analysis of the beam response in a number of papers. Yoshimura et al. [8] performed a finite element analysis of a vehicle moving along a simply supported beam with random surface irregularities and varying cross section. Frýba et al. [9] examined the behavior of an infinitely long Euler beam on a Kelvin foundation with randomly varying parameters along the beam. The response due to a constant load moving uniformly along the beam was analyzed using stochastic finite elements in a moving frame of reference. Chang and Liu [10] performed an analysis of a single-degree-of-freedom vehicle moving along a finite beam with random surface on a nonlinear deterministic Kelvin foundation using finite elements and Monte Carlo simulation.

Lombaert et al. [11] studied the behavior of a beam on an elastic half-space due to the load from a moving vehicle. The vehicle-track interaction was disregarded based on the previous findings by Cebon [12], who showed that for road traffic the interaction is insignificant. Mertrine and Vrouwenvelder [13] used a beam situated within a two-dimensional soillayer to represent a railway tunnel. The tunnel and surface vibrations due to random loading on the beam was analyzed. Again vehicle-track interaction was, however, not taken into consideration. Furthermore the analysis was performed for a fixed reference system. Hence the response became nonstationary with time even though a stationary random loading process was considered.

For the purpose of analysing a railway track or a road on a subsoil, a Kelvin foundation with frequency-independent parameters is unrealistic. A better model may be formulated by the use of a Kelvin foundation which is equivalent to a viscoelastic half-space or a layer over a bedrock as proposed by Dieterman and Metrikine [14] and Metrikine and Popp [15], respectively.

In the present paper the parameters of the foundation are nevertheless assumed to be frequency independent and deterministic. The surface of the beam is on the other hand irregular, described by a weakly homogeneous random process. An analytical method will be presented for the analysis of a single-degree-of-freedom vehicle moving uniformly along the beam. The problem is formulated in a local coordinate frame, which follows the vehicle, and the interaction between the vehicle and the beam is taken into account. No Monte Carlo simulation is necessary since the system is assumed to be linear. A numerical example is given for surface irregularities with an autospectrum which is typically used to describe the irregularities of a road surface. The analysis shows that for certain configurations of the vehicle, the beam and the support (which may correspond to a real railway structure) the interaction

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 17, 2000; final revision, June 12, 2001. Editor: N. C. Perkins. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

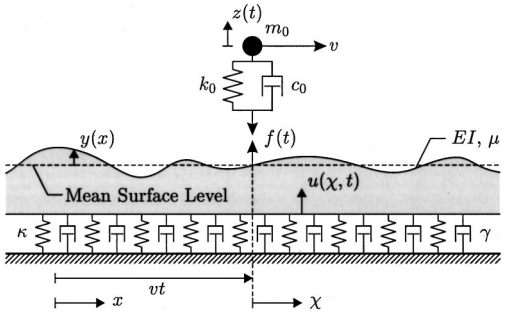


Fig. 1 Single-degree-of-freedom vehicle moving along a Bernoulli-Euler beam with irregular surface on a Kelvin foundation

has a significant influence on the dynamic amplification of the displacement response of both the vehicle mass and the beam.

2 Theory

A vehicle modeled as a single-degree-of-freedom system with mass m_0, spring stiffness k_0 and viscous damping c_0 is moving uniformly in permanent contact along the surface of a Bernoulli-Euler beam at the velocity v thus having the along beam position $x=vt$ at time t. The beam has the bending stiffness EI and the mass μ per unit length and rests on a Kelvin foundation with stiffness κ and viscous damping γ per unit length of the beam. The material properties of the vehicle and beam are homogeneous and constant in time, and the beam axis is supposed to form a straight line in the state of static equilibrium. However, the surface of the beam has the irregularities $y=y(x)$ measured from the mean level of the surface perpendicular to the beam axis. Alternatively, $y=y(x)$ may be used to describe wheel irregularities or a combination of wheel and surface irregularities (see Fig. 1).

A local coordinate system following the vehicle is introduced by the transformation $\chi=x-vt$. Due to convection the partial derivative with respect to time in the moving frame of reference includes a spatial derivative,

$$\left.\frac{\partial}{\partial t}\right|_{\chi}=\left.\frac{\partial}{\partial t}\right|_{x}+v\,\frac{\partial}{\partial x}. \tag{1}$$

Let $z=z(t)$ be the vertical displacement of the point mass and let $u=u(\chi,t)$ denote the vertical displacement of the beam in the local coordinate system, both relative to the respective positions in the state of static equilibrium for no surface irregularity. Putting $\dot{u}=\left.\partial u/\partial t\right|_{\chi}$ and making use of the fact that $y(x)=y(\chi+vt)$, the equation of motion may in turn be written

$$m_0\frac{d^2z}{dt^2}+c_0\left(\frac{dz}{dt}-\dot{u}(0,t)-vy'(vt)\right)+k_0(z-u(0,t)-y(vt))=0, \tag{2}$$

where the prime denotes differentiation with respect to the argument. It is assumed that the mean level $y_m(x)$ of the beam surface does not change along the beam. However, if $y_m(x)$ has a deterministic variation with x, the linearity of the problem implies that the total response may be expressed as a sum of the response due to the irregularities $y_m(x)$ and $y(x)$.

For the beam the equation of motion in the moving coordinate frame may be written

$$EI\frac{\partial^4u}{\partial\chi^4}+\mu\left(\ddot{u}-2v\,\frac{\partial\dot{u}}{\partial\chi}+v^2\frac{\partial^2u}{\partial\chi^2}\right)+\gamma\left(\dot{u}-v\,\frac{\partial u}{\partial\chi}\right)+\kappa u=f(t)\,\delta(\chi). \tag{3}$$

Here, $\ddot{u}=\left.\partial^2u/\partial t^2\right|_{\chi}$ is the local acceleration, whereas $f(t)$ is the moving load and $\delta(\chi)$ is the delta function. In the present case the force on the beam originates from the single-degree-of-freedom vehicle. By d'Alembert's principle $f(t)=-m_0d^2z/dt^2$. It should be noticed that besides this dynamic load the vehicle will also provide the deterministic static load $f_s=-m_0g$, where g is the gravitational acceleration. In what follows, the static load is disregarded and only the stochastic dynamic response is considered. A determination of the total response could easily be carried out by application of the principle of linear superposition.

2.1 Harmonic Surface Irregularities. Harmonically varying surface irregularities with wavelength L are given by

$$y(x)=Ye^{ikx}, \tag{4}$$

where Y is the amplitude, $i=\sqrt{-1}$ is the imaginary unit, and $k=2\pi/L$ is the wave number. Inserting Eq. (4) with $x=vt$ into Eq. (2), the equation of motion for the vehicle reads

$$m_0\frac{d^2z}{dt^2}+c_0\left(\frac{dz}{dt}-\dot{u}(0,t)-i\omega Ye^{i\omega t}\right)+k_0(z-u(0,t)-Ye^{i\omega t})=0, \tag{5}$$

in which $\omega=kv$ denotes the apparent circular frequency of the surface irregularities as seen from the vehicle in the local coordinate system.

Solutions to Eqs. (5) and (3) with $f(t)=-m_0d^2z/dt^2$ are on the form

$$z(t)=Z(\omega)e^{i\omega t},\quad u_j(\chi,t)=U_j(\chi,\omega)e^{i\omega t},\quad j=1,2,3,4, \tag{6}$$

where $Z(\omega)$ and $U_j(\chi,\omega)=m_0\omega^2Z(\omega)\tilde{U}_j(\omega)e^{iK_j\chi}$ are the vehicle amplitude and the amplitude functions for the four bending wave components in the beam, respectively. Here $\tilde{U}_j(\omega)$ are the amplitudes at $\chi=0$ for harmonic excitation with unit amplitude, i.e., $f(t)=e^{i\omega t}$, and the wave numbers $K_j=K_j(\omega)$ correspond to the roots of the characteristic polynomial,

$$K_j^4-\frac{\mu v^2}{EI}K_j^2-\frac{i\gamma v-2\mu\omega v}{EI}K_j+\frac{\kappa-\mu\omega^2+i\gamma\omega}{EI}=0. \tag{7}$$

For $v=\gamma=0$ a *cutoff frequency*, $\omega_c=\sqrt{\kappa/\mu}$, exists. When $\omega<\omega_c$, no travelling waves without attenuation will propagate.

Physically, waves with increasing amplitudes in the far-field are invalid. Furthermore only travelling waves with no attenuation and a group velocity away from the force can exist. The requirements may be formulated as

$$\mathcal{I}(K_j)\chi\geqslant 0,\quad \mathcal{R}(K_j)\chi\leqslant 0\quad\text{for}\quad \mathcal{I}(K_j)=0, \tag{8}$$

respectively. Here $\mathcal{R}(K_j)$ and $\mathcal{I}(K_j)$ denote the real and imaginary part of the wave number, respectively. It may be shown that only two of the solutions to Eq. (7) fulfill the requirements given in Eq. (8) on either side of the load; see, e.g., [16]. In what follows the subscripts $j=1,2$ will be used for the components existing at $\chi\leqslant 0$ (i.e., behind and under the load), whereas the components existing in front of the load have subscripts $j=3,4$. Thus the amplitude $U(\chi,\omega)$ of the beam displacement field becomes

$$U(\chi,\omega)=m_0\omega^2Z(\omega)\sum_{j=j_1}^{j_2}\tilde{U}_j(\omega)e^{iK_j\chi},\qquad \begin{cases}\{j_1,j_2\}=\{1,2\} & \text{for}\quad \chi\leqslant 0\\ \{j_1,j_2\}=\{3,4\} & \text{for}\quad \chi>0\end{cases}. \tag{9}$$

Alternatively the wave components $j=3$ and $j=4$ could be used for $\chi=0$. A summation of all four components is inappropriate, since this would imply a discontinuity of the displacement field under the load.

At $\chi=0$ the displacement, rotation, and bending moment of the beam must be continuous functions of χ. A unit amplitude load inquires a jump of 1 in the shear force at the point where the load is applied. Hence, with the positive directions defined in Fig. 1 the amplitudes $\tilde{U}_j(\omega)$ are obtained by the following system of equations:

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ iK_1 & iK_2 & -iK_3 & -iK_4 \\ (iK_1)^2 & (iK_2)^2 & -(iK_3)^2 & -(iK_4)^2 \\ (iK_1)^3 & (iK_2)^3 & -(iK_3)^3 & -(iK_4)^3 \end{bmatrix} \begin{bmatrix} \tilde{U}_1 \\ \tilde{U}_2 \\ \tilde{U}_3 \\ \tilde{U}_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1/EI \end{bmatrix}. \tag{10}$$

Insertion of the wave field given in Eq. (9) into Eq. (5) leads to the following solution for the amplitude of the single-degree-of-freedom mass, $Z(\omega)$:

$$Z(\omega)=H_{ZY}(\omega)Y, \quad H_{ZY}(\omega)=(i\omega c_0+k_0)/D(\omega), \tag{11}$$

where the denominator, $D(\omega)$, is given by

$$D(\omega)=(-\omega^2 m_0+i\omega c_0+k_0)-(i\omega^3 m_0 c_0+\omega^2 m_0 k_0)\sum_{j=1}^{2}\tilde{U}_j(\omega). \tag{12}$$

Analogously Eq. (9) for the beam displacement may be rewritten as

$$U(\chi,\omega)=H_{UY}(\chi,\omega)Y, \quad H_{UY}(\chi,\omega)=H_{UZ}(\chi,\omega)H_{ZY}(\omega). \tag{13}$$

Here the frequency response function $H_{UZ}(\chi,\omega)$ is given by

$$H_{UZ}(\chi,\omega)=m_0\omega^2\sum_{j=j_1}^{j_2}\tilde{U}_j(\omega)e^{iK_j\chi}, \quad \begin{cases} \{j_1,j_2\}=\{1,2\} & \text{for } \chi\leqslant 0 \\ \{j_1,j_2\}=\{3,4\} & \text{for } \chi>0 \end{cases}. \tag{14}$$

2.2 Random Surface Irregularities. In practice the surface irregularities are not harmonically varying, but will instead be described by a weakly homogeneous stochastic process. Given the one-sided autospectral density $S_Y(\omega)$ for the surface irregularities, the one-sided autospectral density $S_Z(\omega)$ for the displacement of the single-degree-of-freedom vehicle can be found. Also, the one-sided cross-spectral density $S_{UU}(\chi_1,\chi_2,\omega)$ for the beam displacement at two points χ_1 and χ_2 on the beam axis can be calculated. Thus, see, e.g., Lin [17],

$$S_Z(\omega)=|H_{ZY}(\omega)|^2 S_Y(\omega), \tag{15}$$

$$S_{UU}(\chi_1,\chi_2,\omega)=H^*_{UY}(\chi_1,\omega)H_{UY}(\chi_2,\omega)S_Y(\omega), \tag{16}$$

where $H_{ZY}(\omega)$ and $H_{UY}(\chi,\omega)$ are defined previously and $H^*_{UY}(\chi_1,\omega)$ is the complex conjugate of $H_{UY}(\chi_1,\omega)$. The principle of superposition is valid, because the governing equations are all linear.

From the Wiener-Khintchine relation the auto-covariance function $\kappa_{ZZ}(\tau)$ for the vehicle displacement and the cross-covariance function $\kappa_{UU}(\chi_1,\chi_2,\tau)$ for the beam displacement may be expressed as

$$\kappa_{ZZ}(\tau)=\int_0^\infty \cos(\omega\tau)S_Z(\omega)d\omega, \tag{17}$$

$$\kappa_{UU}(\chi_1,\chi_2,\tau)=2\int_0^\infty(\cos(\omega\tau)S^{\mathcal{R}}_{UU}-\sin(\omega\tau)S^{\mathcal{J}}_{UU})d\omega, \tag{18}$$

respectively, where $S^{\Re}_{UU}$ and $S^{\Im}_{UU}$ are the real and imaginary parts of $S_{UU}(\chi_1,\chi_2,\omega)$, respectively.

2.3 Nondimensional Parameter Description. The dynamic response of the vehicle and beam is self-induced, i.e., no external load acts on the system. Therefore, only the relative size of the system parameters is of importance. Dimension analysis and a further study of the system equations would imply that the following nondimensional identities govern the problem:

$$\Omega_0=\frac{\omega_0}{\omega_r}, \quad \Omega_c=\frac{\omega_c}{\omega_0}, \quad M_0=\frac{m_0}{\mu L_c}, \quad \nu=\frac{v}{v_c}, \tag{19}$$

$$\zeta_0=\frac{c_0}{2\sqrt{m_0 k_0}}, \quad \zeta_c=\frac{\gamma}{2\sqrt{\mu\kappa}}. \tag{20}$$

Here $\omega_0=\sqrt{k_0/m_0}$ is the circular eigenfrequency of the single-degree-of-freedom vehicle, and $\omega_r=v k_r$ is a characteristic circular frequency of the surface roughness, k_r being the corresponding characteristic wave number. L_c is a characteristic wavelength for bending waves in the beam, and v_c is the corresponding characteristic phase velocity. They are defined, respectively, as

$$L_c=\frac{2\pi v_c}{\omega_c}, \quad v_c=\sqrt[4]{\frac{4EI\kappa}{\mu^2}}. \tag{21}$$

It should be noticed that ν serves as a kind of Mach number putting the vehicle velocity relative to the velocity of the bending waves.

3 Numerical Examples

Irregularities of a road surface are often modeled as a weakly homogeneous process $\{Y(x),x\in R\}$ having the one-sided autospectral density

$$S_Y(k)=\begin{cases} 0, & k\notin[k_1,k_2] \\ \sigma_Y^2\dfrac{k_1k_2}{k_2-k_1}\dfrac{1}{k^2}, & k\in[k_1,k_2] \end{cases}, \tag{22}$$

where σ_Y is the standard deviation of the surface irregularities. The spectrum has a slope of 1:2 in double-logarithmic mapping and is valid for wavelengths between $L_2=0.1$ m and $L_1=10$ m corresponding to $k_1=2\pi/10\ \text{m}^{-1}$ and $k_2=2\pi/0.1\ \text{m}^{-1}$. The characteristic wave number for the surface roughness is chosen as $k_r=\sqrt{k_1k_2}$, which lies in the middle of the $[k_1,k_2]$ interval on a logarithmic scale.

A transformation from wavenumber domain to frequency domain is obtained by substituting $k=\omega/v$ into Eq. (22),

$$S_Y(\omega)=\begin{cases} 0, & \omega\notin[\omega_1,\omega_2] \\ \sigma_Y^2\dfrac{\omega_1\omega_2}{\omega_2-\omega_1}\dfrac{v}{\omega^2}, & \omega\in[\omega_1,\omega_2] \end{cases}, \tag{23}$$

where $\omega_1=k_1v$ and $\omega_2=k_2v$. Subsequently $S_Z(\omega)$ is found by insertion into Eq. (15).

In the following an analysis will be carried out for the variance of the single-degree-of-freedom mass displacement response, $\sigma_Z^2=\kappa_{ZZ}(0)$, and the variance of the beam displacement response directly under the vehicle, $\sigma_U^2=\kappa_{UU}(0,0,0)$. Due to the linearity of the problem σ_Z^2 and σ_U^2 are proportional to σ_Y^2. Hence, it is convenient to describe the respective variances by the *dynamic variance amplification factors*,

$$s_Z=\frac{\sigma_Z^2}{\sigma_Y^2}, \quad s_U=\frac{\sigma_U^2}{\sigma_Y^2}. \tag{24}$$

A beam with the mass per unit length $\mu=100$ kg/m is considered. The vehicle has the circular eigenfrequency $\omega_0=2\pi\ \text{s}^{-1}$ and the damping ratio $\zeta_0=1$, which are assumed to be typical values. The analysis is performed for different sizes of the point mass, $m_0=100$, 1000, 10,000 kg, and the characteristic wavelength of bending waves and the circular cutoff frequency are both varied. The values $L_c=1$, 10, 100 m and $\Omega_c=0.1$, 1, 10 are used. It should be noticed that with the definitions in Eqs. (19) to (21) an increase of Ω_c will imply an increase in both κ and *EI* for constant L_c, whereas an increase in L_c only leads to an increase in *EI* for constant Ω_c.

Figures 2–4 show the dynamic amplification factors s_Z and s_U as functions of the vehicle velocity in the $v\in[1\ \text{m/s},100\ \text{m/s}]$ interval. Further to the results obtained using the indicated theory (the *continuous* curves), reference results (the *dashed* curves) are shown for a situation where the interaction between the beam and

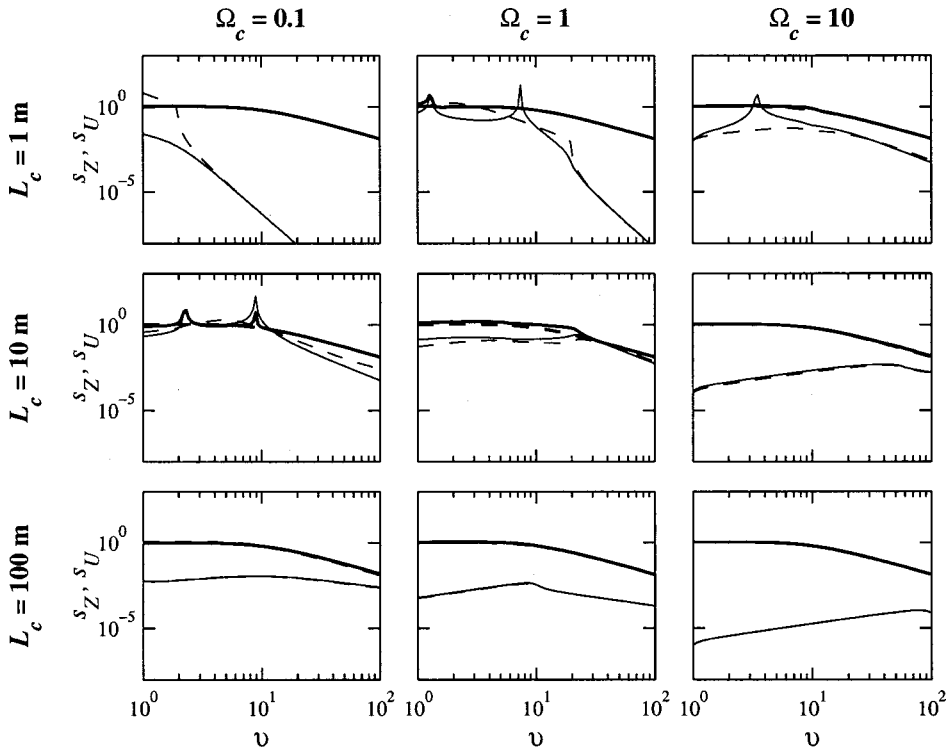


Fig. 2 Dynamic amplification of vehicle mass response (▬) and beam response (—) at $\chi=0$. The dashed lines indicate the response when interaction between vehicle and beam is neglected. $m_0=100$ kg, $\omega_0=2\pi\ \mathrm{s}^{-1}$, $\mu=100$ kg/m, $\zeta_0=1$ and $\zeta_c=0.1$.

the vehicle is not taken into consideration, i.e., the vehicle does not feel the displacements of the beam, but only the surface roughness. In this case the denominator $D(\omega)=(-\omega^2 m_0+i\omega c_0+k_0)$ is used instead of the original denominator, Eq. (12).

In the reference case with no interaction (the dashed curves) the response of the point mass is seen to decrease for increasing values of v beyond 10 m/s. This could be expected since $\Omega_0=1$ when $v=1$ m/s and ω_0 moves to the bottom of the frequency range for the autospectra when $v=10$ m/s. For velocities of more than 10 m/s, ω_0 lies outside the frequency range.

When interaction of the vehicle and beam is taken into account the response of both the single-degree-of-freedom mass and the beam is almost unchanged, if the spring stiffness of the Kelvin foundation and the bending stiffness of the beam is sufficiently high. This also applies to the single-degree-of-freedom mass response when the stiffness of the beam and support is very low. In this case the beam response is, however, not estimated correctly in the reference calculation at low velocities. As seen from the figures the deviation grows with an increase of m_0.

For some configurations of the beam and support other than the aforementioned there is a general amplification of the response of both the single-degree-of-freedom mass and the beam, as long as the velocity of the vehicle lies beneath a certain value. The amplification tends to become stronger when the single-degree-of-freedom mass is increasing and the velocity, where the amplification drops off, at the same time becomes smaller. However, an increase of the bending stiffness of the beam leads to an increase of the velocity where the dropoff appears. Hence, the dropoff velocity for $m_0=1000$ kg, $\Omega_c=1$ and $L_c=10$ m is almost the same as the dropoff velocity for $m_0=10{,}000$ kg, $\Omega_c=1$ and $L_c=100$ m.

Another interesting feature of the response is that for certain combinations of beam configurations, point mass and velocities, the beam response, and in some cases also the single-degree-of-freedom mass response, is amplified drastically. Hence, values of both amplification factors as high as 1000 are found. The most dramatic reinforcement of the response takes place for $m_0=1000$ kg, $\Omega_c=1$ and $L_c=10$ m.

A real road structure is typically very stiff in proportion to the suspension of the vehicle, i.e., Ω_c and L_c are very large. Further, analyses show that in such cases the local peaks in the amplification curves will move to higher velocities and the interaction between the vehicle and the beam structure becomes less important, as indicated by Cebon [12]. However, for heavy freight trains or high-speed passenger trains running on soft soil layers or rubber damping devices in metro tunnels, a low stiffness of the beam and support relative to the stiffness of the vehicle may be expected. Hence, the interaction may be of significant importance in the analysis of such problems.

For $\zeta_0=1$, i.e., critical damping of the vehicle, Figs. 2–4 show that in most cases there is no amplification of the single-degree-of-freedom mass displacement response beyond the displacements directly corresponding to the surface irregularities. Also it may be noticed that an amplification of the single-degree-of-freedom mass response will only take place in conjunction with an amplification of the beam displacement, though an amplification of the beam response may not necessarily lead to an amplification of the single-degree-of-freedom mass response.

To test whether the same feature applies when the vehicle is not critically damped, the system response has been analyzed for various combinations of the damping ratios ζ_0 and ζ_c. The remaining parameters of the system are kept constant with the following parameters being used:

$$m_0=1000\ \mathrm{kg},\ \omega_0=2\pi\ \mathrm{s}^{-1},\ \mu=100\ \mathrm{kg/m},\ \Omega_c=10,\ L_c=10\ \mathrm{m}. \tag{25}$$

The results of the analysis are illustrated in Fig. 5. As it could be expected, the dynamic amplification of the single-degree-of-freedom mass displacement increases significantly when the damping ratio ζ_0 is decreased. There is almost a one-to-one cor-

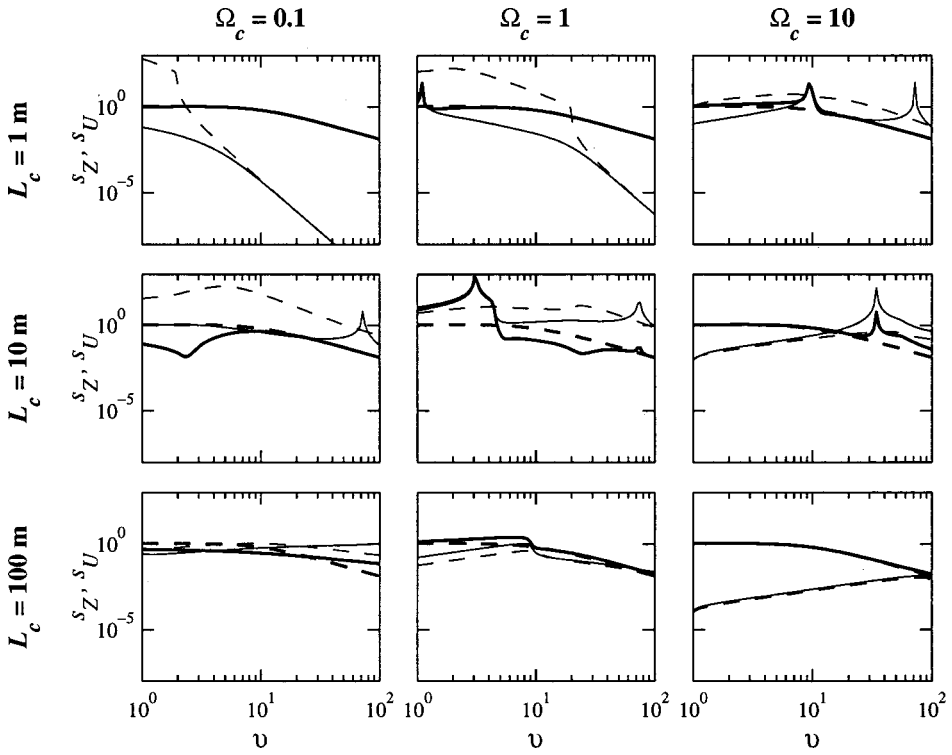


Fig. 3 Dynamic amplification of vehicle mass response (——) and beam response (——) at $\chi = 0$. The dashed lines indicate the response when interaction between vehicle and beam is neglected. $m_0 = 1000$ kg, $\omega_0 = 2\pi$ s^{-1}, $\mu = 100$ kg/m, $\zeta_0 = 1$ and $\zeta_c = 0.1$.

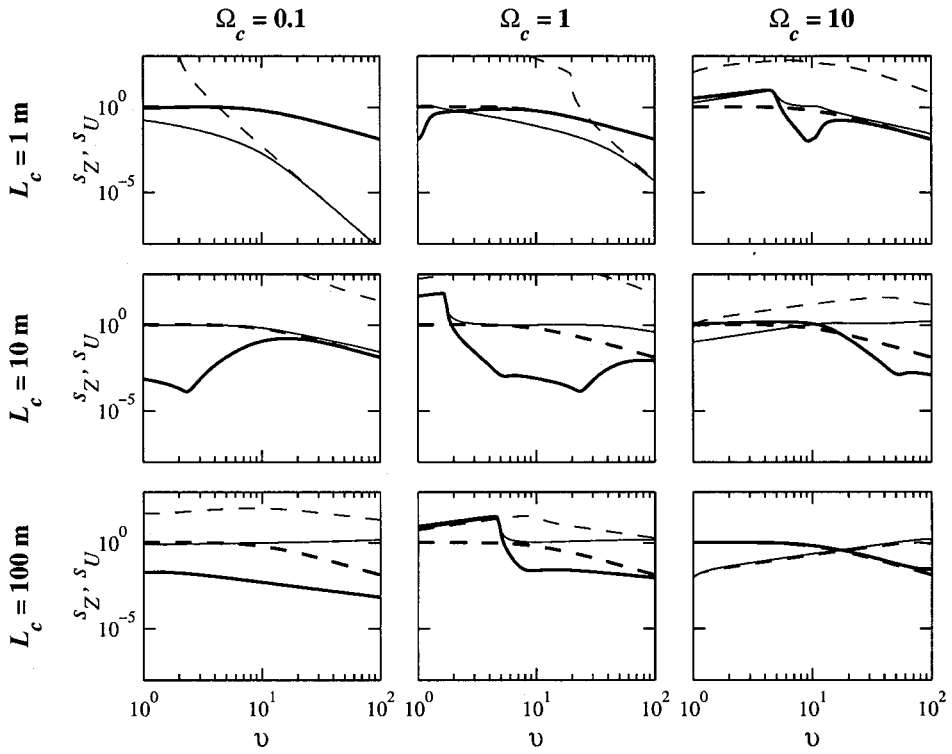


Fig. 4 Dynamic amplification of vehicle mass response (——) and beam response (——) at $\chi = 0$. The dashed lines indicate the response when interaction between vehicle and beam is neglected. $m_0 = 10{,}000$ kg, $\omega_0 = 2\pi$ s^{-1}, $\mu = 100$ kg/m, $\zeta_0 = 1$ and $\zeta_c = 0.1$.

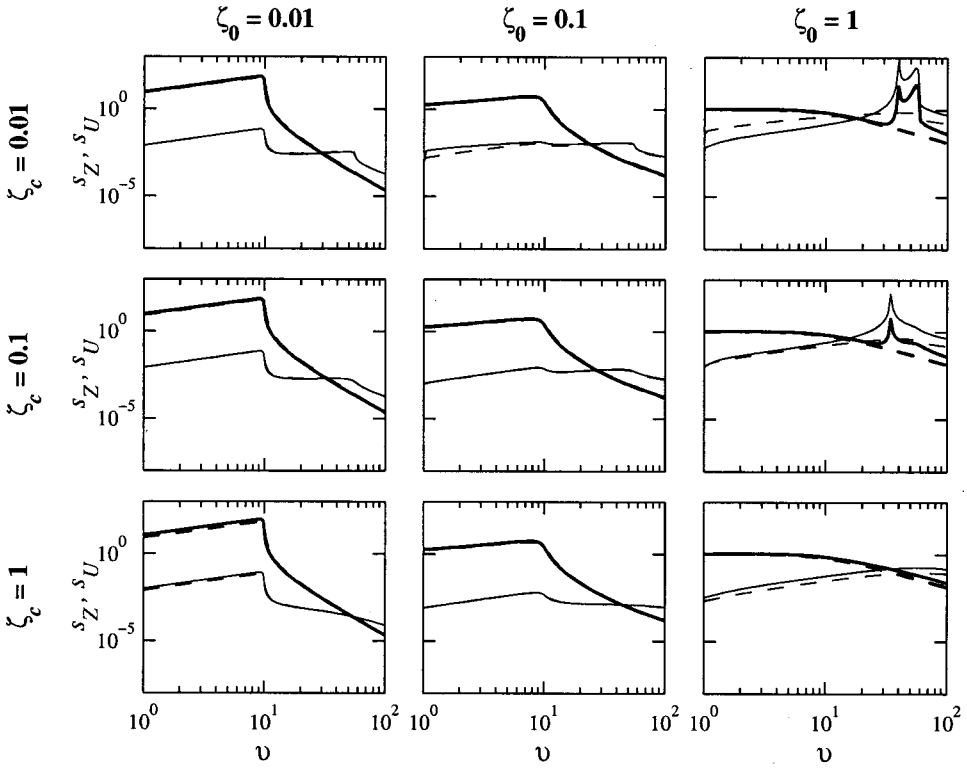


Fig. 5 Dynamic amplification of vehicle mass response (——) and beam response (——) at $\chi=0$. The dashed lines indicate the response when interaction between vehicle and beam is neglected. $m_0=1000$ kg, $\omega_0=2\pi$ s^{-1}, $\mu=100$ kg/m, $\Omega_c=10$ and $L_c=10$ m.

respondence between $1/\zeta_0$ and s_Z for velocities close to 10 m/s. Here the eigenfrequency ω_c lies near the lower bound of the frequency interval for the autospectra where the major part of the variation in the surface irregularities is present.

However, the curves in Fig. 5 indicate that for velocities beyond 10 m/s an increase of the damping ratio of the vehicle is a disadvantage, especially when the damping in the support is relatively low. Thus, some quite significant peaks arise for the combination $\zeta_0=1$, $\zeta_c=0.01$ in the interval $v\approx 40$ to 60 m/s. At these velocities the eigenfrequency of the vehicle lies way below the bottom frequency of the autospectrum for the surface roughness. Hence, the origin of the peaks must be resonance in the beam. For $\zeta_c=0.01$ a peak is still present, though it is less pronounced. However, when both the vehicle and the beam are critically damped, the peaks have vanished. Here the interaction has no influence on the displacement response of neither the single-degree-of-freedom mass nor the beam for the entire velocity range considered (the dashed and continuous curves coincide).

In practice vehicles are (close to) being critically damped. However, depending on the material of the underlying structure it seems likely that ζ_c will be of the order 0.01 to 0.1. This means that a strong displacement amplification will actually occur at a certain critical velocity of the vehicle.

4 Conclusions

The response of a single-degree-of-freedom vehicle moving uniformly along an Euler beam on a viscoelastic foundation of Kelvin type has been investigated. Only the stochastic part of the response due to random surface irregularities has been considered.

The analysis shows that when the beam and support are relatively stiff compared to the suspension of the vehicle the influence of interaction between the beam and vehicle is insignificant. This is in accordance with the assumptions made in [11]. Also, when the beam and support are relatively soft, the vehicle response goes unchanged from the situation where interaction is neglected. The beam response is, however, estimated incorrectly in this situation when interaction is not taken into account.

For the given autospectrum of the surface roughness, some configurations of the beam and support along with the velocity of the vehicle may prove critical. Thus, even for a vehicle, which is critically damped, a significant amplification of the order 10^3 to the surface roughness may take place. Resonance of the vehicle as an isolated system is not the main problem in this case, which indicates that for a vehicle moving on a structure, which may interact with the vehicle, the mechanical design of both vehicle and structure is of great importance. Actually, it has been found in the analysis that a reduction of the damping in the vehicle may prove beneficial at high velocities. Nevertheless, when the damping of the vehicle is reduced, the dynamic amplification of the response is increased drastically at velocities where the eigenfrequency of the single-degree-of-freedom system lies within the frequency range of the autospectrum for the surface roughness.

The analysis indicates that the interaction between the vehicle and beam structure should not be neglected in the analysis of, e.g., heavy freight trains running on a railway track structure with a low bending stiffness and a low cutoff frequency. Also it should be noticed that neglecting the interaction in the derived analytical method does not make the calculation of the displacement amplifications any simpler. Anyway, the terms of the beam displacement that influence the vehicle mass displacement must be found, if the beam displacement under the vehicle should be determined.

An analysis of a realistic vehicle requires a model with more degrees-of-freedom and contact points along the beam. For the single-degree-of-freedom vehicle it is relatively simple to put forward the frequency response matrix for both the vehicle itself and for the beam. Already when a two-degree-of-freedom vehicle in contact with the beam at two separate points along the structure is considered, however, the interaction of the vehicle and the under-

lying structure becomes far more complicated. Hence, an analytical approach is inconvenient. For a vehicle with even more degrees-of-freedom only a numerical solution may be found. Discrete frequency response functions may be determined using, e.g., a finite element scheme. A finite element formulation of the problem with a moving load on an infinite Euler beam-Kelvin foundation system in the moving frame of reference has previously been suggested by Andersen et al. [16].

Acknowledgments

The participation of L. Andersen and S. R. K. Nielsen in the present research has been supported by The Danish Technical Research Council within the project ''Damping Mechanisms in Dynamics of Structures and Materials.''

References

[1] Sobczyk, K., 1970, ''Random Vibrations of Statistically Inhomogeneous Elastic Systems,'' Proc. Vib. Probl., **11**, No. 4, pp. 369–381.
[2] Frýba, L., 1976, ''Non-stationary Response of a Beam to a Moving Random Force,'' J. Sound Vib., **46**, No. 3, pp. 323–338.
[3] Zibdeh, H. S., 1995, ''Stochastic Vibration of an Elastic Beam due to Random Moving Loads and Deterministic Axial Forces,'' Eng. Struct., **17**, No. 7, pp. 530–535.
[4] Iwankiewicz, R., and Śniady, P., 1984, ''Vibration of a Beam Under a Random Stream of Moving Forces,'' J. Struct. Mech., **12**, No. 1, pp. 13–26.
[5] Ricciardi, G., 1994, ''Random Vibration of Beam Under Moving Loads,'' J. Eng. Mech., **120**, No. 11, pp. 2361–2380.
[6] Zibdeh, H. S., and Rackwitz, R., 1995, ''Response Moments of an Elastic Beam Subjected to Poissonian Moving Loads,'' J. Sound Vib., **188**, No. 4, pp. 479–495.
[7] Śniady, P., Biernat, S., Sieniawska, R., and Żukowski, S., 2001, ''Vibrations of the Beam due to a Load Moving With Stochastic Velocity,'' Prob. Eng. Mech., **16**, pp. 53–59.
[8] Yoshimura, T., Hino, J., Kamata, T., and Ananthanarayana, N., 1988, ''Random Vibration of a Non-linear Beam Subjected to a Moving Load: A Finite Element Analysis,'' J. Sound Vib., **122**, No. 2, pp. 317–329.
[9] Frýba, L., Nakagiri, S., and Yoshikawa, N., 1993, ''Stochastic Finite Elements for a Beam on a Random Foundation With Uncertain Damping Under a Moving Force,'' J. Sound Vib., **163**, No. 1, pp. 31–45.
[10] Chang, T.-P., and Liu, Y.-N., 1996, ''Dynamic Finite Element Analysis of a Nonlinear Beam Subjected to a Moving Load,'' Int. J. Solids Struct., **33**, No. 12, pp. 1673–1680.
[11] Lombaert, G., Degrande, G., and Clouteau, D., 2000, ''Numerical Modelling of Free Field Traffic-Induced Vibrations,'' Soil Dyn. Earthquake Eng., **19**, No. 7, pp. 473–488.
[12] Cebon, D., (1993), ''Interaction Between Heavy Vehicles and Roads,'' Technical report, Cambridge University Engineering Department, Cambridge, UK.
[13] Metrikine, A., and Vrouwenvelder, A. C. W. M., 2000, ''Surface Ground Vibrations due to a Moving Train in a Tunnel: Two-Dimensional Model,'' J. Sound Vib., **234**, No. 1, pp. 43–66.
[14] Dieterman, H. A., and Metrikine, A., 1996, ''The Equivalent Stiffness of a Half-Space Interacting With a Beam: Critical Velocities of a Moving Load Along the Beam,'' Eur. J. Mech., A/Solids, **15**, No. 1, pp. 67–90.
[15] Metrikine, A., and Popp, K., 2000, ''Steady-State Vibrations of an Elastic Beam on a Visco-elastic Layer Under a Moving Load,'' Arch. Appl. Mech., **70**, pp. 399–408.
[16] Andersen, L., Nielsen, S. R. K., and Kirkegaard, P. H., 2001, ''Finite Element Modelling of Infinite Euler Beams on Kelvin Foundations Exposed to Moving Loads in Convected Co-ordinates,'' J. Sound Vib., **241**, No. 4, pp. 587–604.
[17] Lin, Y. K., (1967), *Probabilistic Theory of Structural Dynamics*, McGraw-Hill, New York.

Intersonic Crack Propagation—Part II: Suddenly Stopping Crack

Y. Huang
Department of Mechanical and Industrial Engineering,
University of Illinois,
Urbana, IL 61801

H. Gao
Division of Mechanics and Computation,
Stanford University,
Stanford, CA 94305

In Part I of this series, we have obtained the fundamental solution for a mode II intersonic crack which involves a crack moving uniformly at a velocity between the shear and longitudinal wave speeds while subjected to a pair of concentrated forces suddenly appearing at the crack tip and subsequently acting on the crack faces. The fundamental solution can be used to generate solutions for intersonic crack propagation under arbitrary initial equilibrium fields. In this paper, Part II of this series, we study a mode II crack suddenly stopping after propagating intersonically for a short time. The solution is obtained by superposing the fundamental solution and the auxiliary problem of a static crack emitting dynamic dislocations such that the relative crack face displacement in the fundamental solution is negated ahead of where the crack tip has stopped. We find that, after the crack stops moving, the stress intensity factor rapidly rises to a finite value and then starts to change gradually toward the equilibrium value for a static crack. A most interesting feature is that the static value of stress intensity is reached neither instantaneously like a suddenly stopping subsonic crack nor asymptotically like a suddenly stopping edge dislocation. Rather, the dynamic stress intensity factor changes continuously as the shear and Rayleigh waves catch up with the stopped crack tip from behind, approaches negative infinity when the Rayleigh wave arrives, and then suddenly assumes the equilibrium static value when all the waves have passed by. This study is an important step toward the study of intersonic crack propagation with arbitrary, nonuniform velocities. [DOI: 10.1115/1.1410936]

1 Introduction

Freund [1] used the Wiener-Hopf method to obtain the fundamental solution for mode I dynamic crack propagation. A semi-infinite crack in an infinite solid was subjected to a pair of concentrate normal forces suddenly applied at the crack tip at time $t=0$, and at the same instant the crack tip started to propagate at a constant velocity v. This fundamental solution can be used to construct the general solution for dynamic crack propagation subject to arbitrary initial equilibrium field ([2]). More importantly, it can also be used to obtain the general solution for crack growth at nonuniform crack-tip speed ([3]). This is because that, once the crack tip suddenly stops propagating, the dynamic stress intensity factor instantaneously reaches its static counterpart for the same geometry and loading ([3]). A more general conclusion has been established (e.g., [2,4]) that the stress intensity factor around a dynamically propagating crack tip is simply its static counterpart (for the same crack length) multiplied by a universal function k of the crack tip velocity,

$$K(a,\dot{a})=k(\dot{a})K(a,0), \tag{1}$$

where a is the crack length, $\dot{a}$ is the crack-tip velocity, the universal function k is approximately given by $k(v)=(1-v/c_R)/\sqrt{1-v/c_l}$, and c_R and c_l are the Rayleigh and longitudinal wave speeds, respectively.

The corresponding mode II analysis was reported by Fossum and Freund [5], and the approach was also generalized to other cases of loading ([6–9]). It should be pointed out that these analyses hold strictly for sub-Rayleigh crack growth, i.e., the crack-tip velocity v is less than the Rayleigh wave speed c_R. Using a different analytic method, Willis [10,11] obtained the solution for nonuniform crack growth in the entire subsonic range, i.e., the crack-tip velocity can exceed the Rayleigh wave speed but less than the shear wave speed c_s.

The present study is motivated by recent experiments on intersonic crack propagation by Rosakis et al. [12,13] who investigated shear dominated crack growth along weak planes in a brittle polyester resin under far-field asymmetrical loading. They observed that the crack-tip velocity not only exceeded the Rayleigh and shear wave speeds, c_R and c_s, but also approached the longitudinal wave speed, c_l. Evidence of shear crack propagation in excess of the shear wave speed has also been provided from observations of shallow crustal earthquakes ([14–17]). There are also analytical and numerical studies on various aspects of shear intersonic crack propagation in homogeneous solids with weak planes, such as the stress singularity and stability regime in isotropic solids ([18–21]) and orthotropic solids ([22,23]); crack propagation at a constant velocity subjected to uniform shear on crack faces ([24]); radiation-free crack-tip velocity and relations with intersonic dislocations ([25]); slip-weakening or cohesive models ([26–29]) and atomic simulations ([30,31]) of intersonic fracture.

The aim of this paper is to extend Freund's analysis ([1,3]) for subsonic crack growth ($v<c_s$) to intersonic crack propagation ($c_s<v<c_l$). The fundamental solution for intersonic shear crack propagation was obtained in Part I of this paper ([32]). A pair of concentrate shear forces was suddenly applied at the crack tip at time $t=0$, and at the same instant the crack tip started to propagate at a velocity v between the shear and longitudinal wave speeds, c_s and c_l. It was established that only at a single crack-tip velocity of $v=\sqrt{2}c_s$ the crack tip has the square-root singularity, and the crack-tip energy release rate at this velocity is slightly less than one quarter of its counterpart for a stationary crack tip (for the same crack length). The present study focuses on an intersonically propagating shear crack tip that suddenly stops moving. This study provides answers to two important but interrelated questions on intersonic crack propagation. First, once the crack tip stops moving, will the crack-tip stress intensity factor instantaneously

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, May 30, 2001. Associate Editor: L. T. Wheeler. Discussion on the paper should be addressed to the Editor, Professor Lewis T. Wheeler, Department of Mechanical Engineering, University of Houston, Houston, TX 77204-4792, and will be accepted until four months after final publication of the paper itself in the ASME JOURNAL OF APPLIED MECHANICS.

reach its static counterpart as in subsonic crack growth ([1,3])? Second, the solution for nonuniform subsonic crack growth can be obtained from that for uniform crack growth ([1,3]). Does this conclusion hold for intersonic crack propagation? Unfortunately, the answer to both questions are negative, as shown in this paper.

2 Suddenly Stopping Crack

The fundamental solution is given in Part I of this paper ([32]) for an infinite linear elastic solid containing a semi-infinite crack on the negative x axis. For time $t \leqslant 0$, the solid is stress free and at rest everywhere, and the crack tip is at the origin (0,0). At time $t=0$, the crack tip begins to move intersonically at a constant velocity v ($c_s<v<c_l$) in the positive x-direction. As the tip moves away, a pair of concentrated shear forces (in the x-direction) of constant magnitude τ^* is left at the origin. In the present study, the loading, geometry and crack-tip motion are the same as those in the fundamental solution ([32]), except that the crack tip suddenly stops moving at time $t=t^*$ after propagating a distance of vt^*. This suddenly stopping crack problem can be decomposed into the following two problems; (i) the fundamental solution that persists for $t>t^*$, i.e., crack tip continues to propagate after time t^*; and (ii) the negation of the sliding displacement across the crack face in the fundamental solution for $vt>x>vt^*$ by emitting dislocations from the crack tip along $x>0, y=0$ in just the appropriate sequence. Freund [33] also used the moving dislocation solution to determine the dynamic stress intensity factor due to normal impact loading on the crack faces.

2.1 Solution for a Moving Dislocation. An infinite linear elastic solid contains a semi-infinite crack on the negative x-axis. For time $t \leqslant 0$, the solid is stress free and at rest everywhere, and the crack tip is at the origin (0,0). An edge dislocation is emitted from the crack tip at time $t=0$, and propagates along the positive x-axis at a constant velocity w less than the longitudinal wave speed c_l (therefore including both subsonic and intersonic regimes). Freund [2,33] used the Wiener-Hopf method to solve the above problem with an edge dislocation climbing in the x-direction, leading to a mode I crack tip. The solution for an edge dislocation gliding in the x-direction, corresponding to a mode II crack tip, is provided in this section to pave the way for analyzing a suddenly stopping crack in the following sections.

The boundary conditions can be written as

$$\sigma_{yy}(x,y=0,t)=0,$$

$$\sigma_{xy}(x<0,y=0,t)=0, \tag{2}$$

$$u_x(x>0,y=0,t)=bH(wt-x),$$

where b is the Burgers vector, and H is the unit step function. Since the analysis is nearly identical to that for mode I ([2,33]), we only provide the final solution of the crack-tip stress intensity factor k_{II},

$$k_{II}=bk_0(t,w), \tag{3}$$

where the linear dependence on the Burgers vector b is made explicit, and k_0 is a function of time t and dislocation velocity w given by

$$k_0(t,w)=-2\mu\left(1-\frac{c_s^2}{c_l^2}\right)\frac{1+\dfrac{w}{c_R}}{\sqrt{1+\dfrac{w}{c_s}}}s_{0-}\left(-\frac{1}{w}\right)\sqrt{\frac{2}{\pi wt}}, \tag{4}$$

μ is the shear modulus, and

$$s_{0-}\left(-\frac{1}{w}\right)=\exp\left[-\frac{1}{\pi}\int_{1/c_l}^{1/c_s}\tan^{-1}\times\left[\frac{4r^2\sqrt{\left(r^2-\dfrac{1}{c_l^2}\right)\left(\dfrac{1}{c_s^2}-r^2\right)}}{\left(2r^2-\dfrac{1}{c_s^2}\right)^2}\right]\frac{dr}{r+\dfrac{1}{w}}\right]. \tag{5}$$

2.2 Crack-Face Sliding Displacement in the Fundamental Solution. It is established in Part I of this paper ([32]) that the sliding displacement across the crack face in the fundamental solution, $\delta=u_1(x<vt,y=0+)-u_1(x<vt,y=0-)$, depends on time t and coordinate x only through their ratio x/t,

$$\delta(x<vt,t)=\delta\left(\frac{x}{t}\right), \tag{6}$$

where v is the crack-tip velocity, and δ is given by

$$\delta(w)=\frac{2}{\pi}PV\int_{1/(c_l+v)}^{1/(v-w)}\mathrm{Im}[U_-(\eta)]d\eta H(w+c_l). \tag{7}$$

Here H is the unit step function; PV stands for the Cauchy principal value integral because $\mathrm{Im}[U_-(\eta)]$ has simple poles at $1/(v-c_R)$ and $1/(v+c_R)$ and a double pole at $1/v$, and

$$\mathrm{Im}[U_-(\eta)]=\frac{\sqrt{c_l v}}{c_s^2}\frac{\tau^*}{\mu}\frac{s_+(\eta)}{s_+\left(\dfrac{1}{v}\right)}(1-v\eta)\sqrt{(c_l-v)\eta+1}\frac{F_n(\eta)}{F_d(\eta)}, \tag{8}$$

τ^* is the concentrated shear force on the crack face in the fundamental solution, μ is the shear modulus,

$$\begin{aligned}F_n(\eta)=&4\eta^2\sqrt{(c_l-v)\eta+1}\sqrt{(c_l+v)\eta-1}[(v-c_s)\eta-1]\\&\times[(v+c_s)\eta-1]-c_lc_s\left[2\eta^2-\frac{1}{c_s^2}(v\eta-1)^2\right]^2\\&\times\sqrt{1-(v-c_s)\eta}\sqrt{(v+c_s)\eta-1}H[1-(v-c_s)\eta]\\&\times H[(v+c_s)\eta-1],\end{aligned} \tag{9}$$

$$\begin{aligned}F_d(\eta)=&16\eta^4[(c_l-v)\eta+1][(c_l+v)\eta-1][(v-c_s)\eta-1]\\&\times[(v+c_s)\eta-1]+c_l^2c_s^2\left[2\eta^2-\frac{1}{c_s^2}(v\eta-1)^2\right]^4,\end{aligned} \tag{10}$$

$$\frac{s_+(\eta)}{s_+\left(\dfrac{1}{v}\right)}=\exp\left[-\frac{v\eta-1}{\pi}\int_{1/(c_l-v)}^{\infty}\frac{\tan^{-1}[V(r)]}{(vr+1)(r+\eta)}dr\right], \tag{11}$$

and

$$V(r)=\frac{c_l c_s\left[2r^2-\frac{1}{c_s^2}(vr+1)^2\right]^2}{4r^2\sqrt{(c_l-v)r-1}\sqrt{(c_l+v)r+1}\sqrt{(v-c_s)r+1}\sqrt{(v+c_s)r+1}}. \tag{12}$$

The crack-face displacement depending only on the ratio x/t implies that any given displacement level $\delta(w)=\delta(x/t)$ radiates out along the x-axis at a constant speed $w=x/t$ ([2,33]). As discussed in the next section, this has important implications on the negation of the sliding displacement in the fundamental solution by the moving dislocation solution.

2.3 Stress Intensity Factor Around a Suddenly Stopping Crack. The method used to determine the stress intensity factor in this section is same as that developed by Freund [2,33]. At any instant $t>t^*$, the sliding displacement needs to be negated from the current crack-tip position $x=vt$ in the fundamental solution to $x=vt^*$, where the crack tip is supposed to stop moving. The corresponding speed $w=x/t$ ranges from v to vt^*/t. For a given velocity w in this range, the time at which the corresponding displacement level $\delta(w)=\delta(x/t)$ arrives at the stopped crack-tip location $x=vt^*$ is

$$t_w=\frac{vt^*}{w}. \tag{13}$$

The stress intensity factor due to a moving dislocation starting to propagate at $t=0$ with velocity w and Burgers vector b is $bk_0(t,w)$, as given in (3). If a dislocation with Burgers vector $d\delta$ begins moving at time $t=t_w$, instead of at $t=0$, then the stress intensity factor is $k_0(t-t_w,w)d\delta$. Since both δ in (7) and t_w in (13) are prescribed functions of w, the stress intensity factor can be summed from v to vt^*/t with respect to w. Stresses in the fundamental solution are not singular around $x=vt^*$ because the current crack tip already propagates to $x=vt$. In other words, only the moving dislocation solution contributes to the stress intensity factor K_{II} at the stopped crack tip $x=vt^*$. This gives K_{II} as

$$K_{II}(t)=-\int_v^{vt^*/t} k_0(t-t_w,w)\frac{d\delta}{dw}dw. \tag{14}$$

The substitution of (7) and (13) into (14) and the change of integration variable to $\eta=1/(v-w)$ give the analytic expression of the stress intensity factor at the stopped crack tip as

$$K_{II}(t)=-\sqrt{\frac{2c_s}{\pi}}\frac{2\mu}{\pi c_R}\left(1-\frac{c_s^2}{c_l^2}\right)PV \times\int_{1/[v(1-t^*/t)]}^{+\infty}\frac{(v+c_R)\eta-1}{\sqrt{v(t-t^*)\eta-t}\sqrt{(v+c_s)\eta-1}} \times s_{0-}\left(-\frac{\eta}{v\eta-1}\right)\mathrm{Im}[U_-(\eta)]d\eta, \tag{15}$$

where PV stands for the Cauchy principal value integral because of the possible Cauchy-type singularity around $1/(v-c_R)$, and s_{0-} and $\mathrm{Im}[U_-]$ are given in (5) and (8), respectively.

It can be shown, after some lengthy calculations, that the stress intensity approaches its equilibrium limit as time $t\to+\infty$,

$$K_{II}(\infty)=\tau^*\sqrt{\frac{2}{\pi vt^*}}, \tag{16}$$

where the right-hand side is the stress intensity factor for a static crack tip subjected to a pair of concentrated shear forces τ^* at a distance vt^* on the crack faces behind the tip. It is also interesting to study the other limit, $t\to t^*+0$, i.e., immediately after the crack tip stops moving. For crack-tip velocity $v\neq\sqrt{2}c_s$, stresses around an intersonically propagating crack tip changes from a weak singularity ($r^{-q}, q<1/2$) ([18,21–23,25,32]) to the conventional square-root singularity ($r^{-1/2}$) as the crack tip suddenly stops moving, but the stress intensity factor in (15) remains zero at this instant $t=t^*$. For crack-tip velocity $v=\sqrt{2}c_s$, the propagating crack tip also has the square-root singularity, and the stress intensity factor in (15) does not vanish at $t=t^*$, but becomes 0.68 times the equilibrium value $K(\infty)$ for Poisson's ratio $\nu=1/3$. These observations imply that, contrary to subsonic crack growth, the dynamic stress intensity factor does not reach its static counterpart instantaneously once an intersonically propagating crack tip suddenly stops moving. In fact, this is quite reasonable because both shear and Rayleigh waves are trailing behind the intersonically propagating crack tip. As the crack tip suddenly stops propagating at time $t=t^*$, the time for the shear and Rayleigh waves to reach the stopping crack tip are $t=vt^*/c_s$ and $t=vt^*/c_R$, respectively. It will be interesting to examine the stress intensity factor at the stopping crack tip after all waves have passed, i.e., $t>vt^*/c_R$.

Figure 1 shows the stress intensity factor K_{II} in (15) for a suddenly stopping crack tip for three crack-tip velocities, $v/c_s=1.1$, $\sqrt{2}$, and 1.7, where K_{II} is normalized by its equilibrium value $K_{II}(\infty)$, and time t is normalized by the crack propagation time t^*. Once the crack tip stops moving, the stress intensity factor rapidly increases to a finite value, and then starts to decrease. A sharp vertex in the figure corresponds to the arrival of the shear wave ($t=vt^*/c_s$). The stress intensity factor approaches to negative infinity once the Rayleigh wave arrives ($t=vt^*/c_R-0$). However, immediately after the Rayleigh wave arrives ($t=vt^*/c_R+0$), the stress intensity factor K_{II} reaches its equilibrium value $K_{II}(\infty)$. This indicates that, even though the dynamic stress intensity factor does not reach its equilibrium value instantaneously once an intersonically propagating crack tip suddenly stops moving, the equilibrium value is reached after all waves have passed. In other words, the stress intensity factor around a suddenly stopped intersonic crack tip displays a finite delay in reaching its static counterpart, contrary to the case of subsonic crack growth. It also behaves different from a suddenly stopping dislocation, for which the corresponding static dislocation field is approached only asymptotically (i.e., as time approaching infinity).

3 Concluding Remarks

In Part I of this paper ([32]) we obtained the fundamental solution for an intersonically propagating crack subjected to a pair of suddenly applied concentrate shear forces on the crack faces. In the present study we have studied the crack-tip behavior and stress intensity factor when the intersonically propagating crack tip is suddenly arrested. It is established that, unlike subsonic crack growth ([1–3]), the dynamic crack-tip stress intensity does not instantaneously reach its equilibrium value when an intersonically propagating crack tip suddenly stops moving. The equilibrium stress intensity factor is reached after a finite delay, i.e., after all (shear and Rayleigh) waves have passed the stopped crack tip. Because of these observations, one cannot obtain the solution for

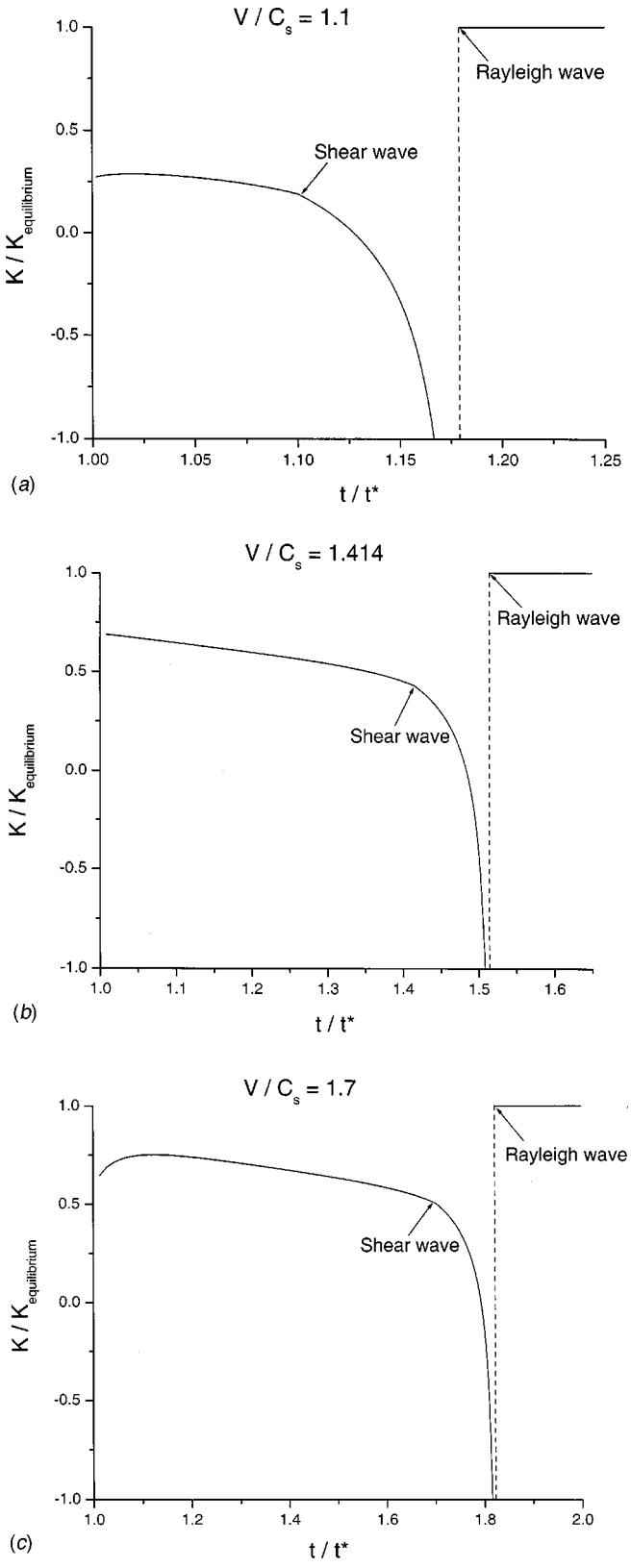


Fig. 1 The stress intensity factor, K_{II}, around a suddenly stopped crack tip that was propagating intersonically; $K_{II}(\infty)$ is the equilibrium value of the stress intensity factor; time t is normalized by the crack propagation time t^*; Poisson's ratio $\nu=1/3$; the crack tip velocities are (*a*) $v=1.1c_s$, (*b*) $v=\sqrt{2}c_s$, and (*c*) $v=1.7c_s$, respectively, and c_s is the shear wave speed

nonuniform intersonic crack propagation from its counterpart for uniform crack growth, as suggested by Freund for subsonic cracks ([2,3]).

Acknowledgments

This investigation was supported by the Office of Naval Research grant #N00014-01-1-0205 (Dr. Y. D. S. Rajapakse, Scientific Officer). Y. H. gratefully acknowledges this support and helpful discussions with Prof. L. B. Freund of Brown University and Prof. J. R. Willis of Cambridge University. The work of H. G. was supported by the NSF through grant CMS9820988.

References

[1] Freund, L. B., 1972, "Crack Propagation in an Elastic Solid Subjected to General Loading. I. Constant Rate of Extension," J. Mech. Phys. Solids, **20**, pp. 129–140.
[2] Freund, L. B., 1990, *Dynamic Fracture Mechanics*, Cambridge University Press, Cambridge, UK.
[3] Freund, L. B., 1972, "Crack Propagation in an Elastic Solid Subjected to General Loading. II. Nonuniform Rate of Extension," J. Mech. Phys. Solids, **20**, pp. 141–152.
[4] Broberg, K. B., 1999, *Cracks and Fracture*, Academic Press, San Diego, CA.
[5] Fossum, A. F., and Freund, L. B., 1975, "Nonuniformly Moving Shear Crack Model of a Shallow Focus Earthquake Mechanism," J. Geophys. Res., **80**, pp. 3343–3347.
[6] Freund, L. B., 1973, "Crack Propagation in an Elastic Solid Subjected to General Loading. III. Stress Wave Loading," J. Mech. Phys. Solids, **21**, pp. 47–61.
[7] Kostrov, B. V., 1975, "On the Crack Propagation With Variable Velocity," Int. J. Fract., **11**, pp. 47–56.
[8] Kostrov, B. V., and Osaulenko, V. I., 1976, "Crack Propagation at Arbitrary Variable Rate Under the Action of Static Solids," Mech. Solids, **11**, pp. 76–92.
[9] Burridge, R., 1976, "An Influence Function for the Intensity Factor in Tensile Fracture," Int. J. Eng. Sci., **14**, pp. 725–734.
[10] Willis, J. R., 1989, "Accelerating Cracks and Related Problems," *Elasticity, Mathematical Methods and Applications, the Ian Sneddon 70th Birthday Volume* R. W. Ogden, ed., Ellis Horwood, Chichester, pp. 397–409.
[11] Willis, J. R., 1992, "The Stress Field Near the Tip of an Accelerating Crack," J. Mech. Phys. Solids, **40**, pp. 1671–1681.
[12] Rosakis, A. J., Samudrala, O., and Coker, D., 1999, "Cracks Faster Than the Shear Wave Speed," Science, **284**, pp. 1337–1340.
[13] Rosakis, A. J., Samudrala, O., and Coker, D., 2000, "Intersonic Shear Crack Growth Along Weak Planes," Mater. Res. Innovations, **3**, pp. 236–243.
[14] Archuleta, R. J., 1982, "Analysis of Near Source Static and Dynamic Measurements From the 1979 Imperial Valley Earthquake," Bull. Seismol. Soc. Am., **72**, pp. 1927–1956.
[15] Beroza, G. C., and Spudich, P., 1988, "Linearized Inversion for Fault Rupture Behavior—Application to the 1984 Morgan-Hill, California, Earthquake," J. Geophys. Res., **93**, pp. 6275–6296.
[16] Wald, D. J., and Heaton, T. H., 1994, "Spatial and Temporal Distribution of Slip for the 1992 Landers, California, Earthquake," Bull. Seismol. Soc. Am., **84**, pp. 668–691.
[17] Ellsworth, W. L., and Celebi, M., 1999, "Near Field Displacement Time Histories of the M 7.4 Kocaeli (Izimit), Turkey, Earthquake of August 17, 1999," EOS Trans. Am. Geophys. Union, **80**, No. 46, p. F648.
[18] Freund, L. B., 1979, "The Mechanics of Dynamic Shear crack propagation," J. Geophys. Res., **84**, pp. 2199–2209.
[19] Burridge, R., Conn, G., and Freund, L. B., 1979, "The Stability of a Plane Strain Shear Crack With Finite Cohesive Force Running at Intersonic Speeds," J. Geophys. Res., **84**, pp. 2210–2222.
[20] Simonov, I. V., 1983, "Behavior of Solutions of Dynamic Problems in the Neighborhood of the Edge of a Cut Moving at Transonic Speed in an Elastic Medium," Mech. Solids, **18**, pp. 100–106.
[21] Broberg, K. B., 1994, "Intersonic Bilateral Slip," Geophys. J. Int., **119**, pp. 706–714.
[22] Broberg, K. B., 1999, "Intersonic Crack Propagation in an Orthotropic Material," Int. J. Fract., **99**, pp. 1–11.
[23] Huang, Y., Wang, W., Liu, C., and Rosakis, A. J., 1999, "Analysis of Intersonic Crack Growth in Unidirectional Fiber-Reinforced Composites," J. Mech. Phys. Solids, **47**, pp. 1893–1916.
[24] Brock, L. M., 1977, "Two Basic Problems of Plane Crack Extension: A Unified Treatment," Int. J. Eng. Sci., **15**, pp. 527–536.
[25] Gao, H., Huang, Y., Gumbsch, P., and Rosakis, A. J., 1999, "On Radiation-Free Transonic Motion of Cracks and Dislocations," J. Mech. Phys. Solids, **47**, pp. 1941–1961.
[26] Andrews, D. J., 1976, "Rupture Velocity of Plane Strain Shear Cracks," J. Geophys. Res., **81**, pp. 5679–5687.
[27] Needleman, A., 1999, "An Analysis of Intersonic Crack Growth Under Shear Loading," ASME J. Appl. Mech., **66**, pp. 847–857.

[28] Yu, H. H., and Suo, Z., 2000, ''Intersonic Crack Growth on an Interface,'' Proc. R. Soc. London, Ser. A, **A456**, pp. 223–246.
[29] Geubelle, P. H., and Kubair, D., 2001, ''Intersonic Crack Propagation in Homogeneous Media Under Shear Dominated Loading: Numerical Analysis,'' J. Mech. Phys. Solids, **49**, pp. 571–587.
[30] Abraham, F. F., and Gao, H., 2000, ''How Fast Can Cracks Propagate?'' Phys. Rev. Lett., **84**, pp. 3113–3116.
[31] Gao, H., Huang, Y., and Abraham, F. F., 2001, ''Continuum and Atomistic Studies of Intersonic Crack Propagation,'' J. Mech. Phys. Solids, **49**, pp. 2113–2132.
[32] Huang, Y., and Gao, H., 2001, ''Intersonic Crack Propagation. I. The Fundamental Solution,'' ASME J. Appl. Mech., **68**, pp. 169–175.
[33] Freund, L. B., 1974, ''The Stress Intensity Factor due to Normal Impact Loading of the Faces of a Crack,'' Int. J. Eng. Sci., **12**, pp. 179–189.

Journal of Applied Mechanics

Brief Notes

A Brief Note is a short paper that presents a specific solution of technical interest in mechanics but which does not necessarily contain new general methods or results. A Brief Note should not exceed 1500 words *or equivalent* (a typical one-column figure or table is equivalent to 250 words; a one line equation to 30 words). Brief Notes will be subject to the usual review procedures prior to publication. After approval such Notes will be published as soon as possible. The Notes should be submitted to the Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on the Brief Notes should be addressed to the Editorial Department, ASME International, Three Park Avenue, New York, NY 10016-5990, or to the Editor of the JOURNAL OF APPLIED MECHANICS. Discussions on Brief Notes appearing in this issue will be accepted until two months after publication. Readers who need more time to prepare a Discussion should request an extension of the deadline from the Editorial Department.

Elastic Solutions for a Solid Rotating Disk With Cubic Anisotropy

F. Zhou[1] **and A. Ogawa**
Aeroengine Division, National Aerospace Laboratory,
7-44-1 Jindaijihigashi, Chofu, Tokyo 182-8522, Japan

Elastic solutions of a rotating solid disk made of cubic anisotropic material are obtained using direct displacement method. Displacement, strain, and stress distributions within the disk are expressed as simple functions of polar coordinates.
[DOI: 10.1115/1.1406958]

1 Introduction

Rotating disks are important structural components in turbomachinery or flywheel systems. Finding elastic solutions of a rotating disk is always a critical issue. The problem of isotropic disk is quite simple ([1]). If the disk is anisotropic, the problem becomes complicated. Two typical anisotropy forms are cylindrical (polar) orthotropic and Cartesian (XY) orthotropic. In the former case, closed-form solutions are obtainable because the deformation of the disk is cylindrically symmetric ([2–8]). In the latter case, the seemingly simple problem turns out to be difficult to solve. There were some works dealing with the Cartesian orthotropic disk problem ([9,10]), where the stress functions of Lekhnitskii [11] were used as the start point of the analyses. Closed-form solutions were obtained only for some special cases.

Among the general Cartesian orthotropic materials, *cubic anisotropic* materials are a special category. These materials have the same elastic property in the material principal directions. Some examples of cubic anisotropic materials are cubic single crystals and balanced crossply ($[0/90]_s$) laminates. Even for this simple type of anisotropy, solutions are still difficult to obtain. Recently, Vigdergauz, and Givoli presented a perturbation method to analyze thermal stress state in a *solid* disk ([12]). Their method is limited to *weak* anisotropic material and is also an approximate method.

In our group, centrifugal spin experiments have been performed to test the bursting strength of plain-woven and [0/90] laminated C/C composites. The specimens are thin annular ring disks. The finite element method was used to evaluate the experimental results ([13]). We also used the Ritz method to analyze the problem. Though no exact solutions were obtained for *rings*, we found that for a *solid, cubic anisotropic* rotating disk, simple closed-form elastic solutions exist. The results are reported below.

2 Formulation

The problem is illustrated in Fig. 1. The disk, with radius R and density ρ, is rotating at angular speed ω. Cartesian and polar coordinate systems are established, with the X-axis as one material principal direction. The disk material is cubic anisotropic with Young's modulus: $E_x=E_y=E$; shear modulus: $G_{xy}=G$; and Poisson's ratio $\nu_{xy}=\nu_{yx}=\nu$. More general, we assume that the disk is loaded by a uniform external tensile stress σ_0, which may come from the tension of blades attached to the disk.

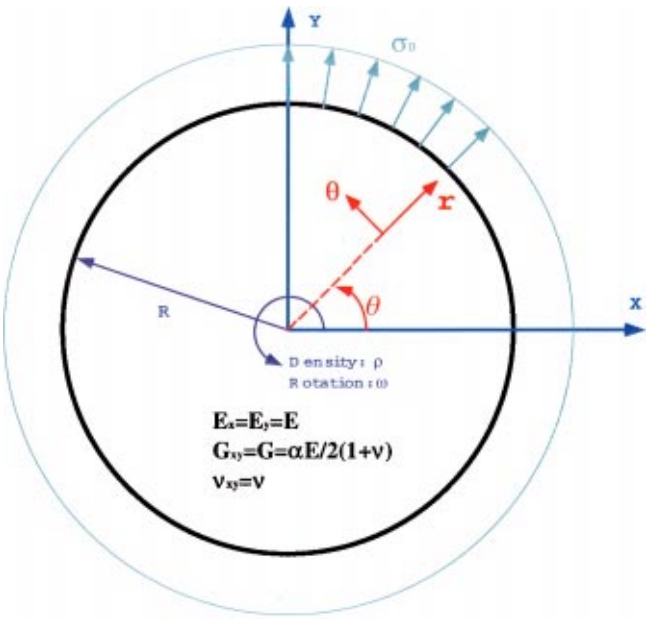


Fig. 1 Geometry

[1]Present address: Department of Mechanical Engineering, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218.

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS. Manuscript received by the ASME Applied Mechanics Division, Nov. 12, 2000; final revision, May 29, 2001. Associate Editor: A. K. Mal.

We limit the problem to the plane stress case. The case of plane strain can be treated similarly when certain transformations of elastic parameters are made. Because of material anisotropy, the displacement components in radial and hoop directions u_r and u_θ are functions of polar coordinates (r,θ):

$$u_r = u_r(r,\theta), \quad u_\theta = u_\theta(r,\theta).$$

The strain components in polar coordinates are

$$\begin{Bmatrix} \varepsilon_r \\ \varepsilon_\theta \\ \gamma_{r\theta} \end{Bmatrix} = \begin{Bmatrix} \dfrac{\partial u_r}{\partial r} \\ \dfrac{u_r}{r} + \dfrac{\partial u_\theta}{r\partial\theta} \\ \dfrac{\partial u_r}{r\partial\theta} + \dfrac{\partial u_\theta}{\partial r} - \dfrac{u_\theta}{r} \end{Bmatrix}. \tag{1}$$

Cubic anisotropic materials have three basic elastic parameters. The constitutive relationship in Cartesian coordinates is

$$\begin{Bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{Bmatrix} = \tilde{E}_0 \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{Bmatrix} = \frac{E}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \dfrac{(1-\nu^2)G}{E} \end{bmatrix} \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{Bmatrix}.$$

In polar coordinates, the constitutive relationship is

$$\begin{Bmatrix} \sigma_r \\ \sigma_\theta \\ \tau_{r\theta} \end{Bmatrix} = \tilde{E}(\theta) \begin{Bmatrix} \varepsilon_r \\ \varepsilon_\theta \\ \gamma_{r\theta} \end{Bmatrix} = \frac{E}{4(1-\nu^2)} \begin{bmatrix} c_{rr} & c_{r\theta} & c_{rs} \\ & c_{\theta\theta} & c_{\theta s} \\ \text{Sym.} & & c_{ss} \end{bmatrix} \begin{Bmatrix} \varepsilon_r \\ \varepsilon_\theta \\ \gamma_{r\theta} \end{Bmatrix}. \tag{2}$$

The elastic matrix $\tilde{E}(\theta)$ is symmetric. Its elements are expressed as follows:

$$\begin{aligned} c_{rr} &= c_{\theta\theta} = 3+\alpha+(1-\alpha)\nu+(1-\alpha)(1-\nu)\cos 4\theta \\ c_{r\theta} &= -1+\alpha-(3+\alpha)\nu+(1-\alpha)(1-\nu)\cos 4\theta \\ c_{rs} &= -(1-\alpha)(1-\nu)\sin 4\theta \\ c_{\theta s} &= (1-\alpha)(1-\nu)\sin 4\theta \\ c_{ss} &= (1-\nu)[1+\alpha-(1-\alpha)\cos 4\theta] \end{aligned}$$

where the parameter $\alpha = 2(1+\nu)G/E$ represents the anisotropy of the material. For an isotropic material: $\alpha=1$, $\tilde{E}(\theta)=\tilde{E}_0$, meaning that the material properties are identical in all directions. For a crossply laminated (or woven) composite, α is smaller than 1; for a single crystal superalloy, α is larger than 1.

The equilibrium equations are

$$\frac{\partial\sigma_r}{\partial r} + \frac{\partial\tau_{r\theta}}{r\partial\theta} + \frac{\sigma_r-\sigma_\theta}{r} + \rho\omega^2 r = 0 \tag{3.1}$$

$$\frac{\partial\sigma_\theta}{r\partial\theta} + \frac{\partial\tau_{r\theta}}{\partial r} + \frac{2\tau_{r\theta}}{r} = 0. \tag{3.2}$$

We wish to find the functions $\{u_r, u_\theta\}$, $\{\sigma_r, \sigma_\theta, \tau_{r\theta}\}$ and $\{\varepsilon_r, \varepsilon_\theta, \gamma_{r\theta}\}$ that simultaneously satisfy Eqs. (1)–(3) and other boundary conditions.

3 Solutions for a Solid Rotating Disk

If the disk is solid, i.e., no internal holes exist, there is only one boundary $(r=R)$, on which the stress conditions are

$$\sigma_r(R,\theta) = \sigma_0, \quad \tau_{r\theta}(R,\theta) = 0. \tag{4}$$

Based on experience, we propose that the displacement components take the following form:

$$\begin{aligned} u_r(r,\theta) &= ar + br^3 + cr^3\cos 4\theta \\ u_\theta(r,\theta) &= -cr^3\sin 4\theta \end{aligned} \tag{5}$$

where a, b, c are unknown constants. At the disk center, u_r and u_θ are zero, omitting possible rigid translation. By direct differentiation, the strain and stress expressions are deduced. It is seen then that the second equilibrium Eq. (3.2) is automatically satisfied. The other equilibrium Eq. (3.1) and the boundary conditions (4) lead to four algebraic equations for three unknowns a, b, c. However, two of the four equations are identical, so that we get determinate solutions for the unknown constants

$$\begin{aligned} a &= \frac{(1-\nu)(1+2\alpha+\nu)\rho\omega^2R^2}{2E(1+3\alpha+\nu-\alpha\nu)} + \frac{1-\nu}{E}\sigma_0 \\ b &= -\frac{(1-\nu^2)(1+\alpha)\rho\omega^2}{4E(1+3\alpha+\nu-\alpha\nu)} \\ c &= \frac{(1-\nu^2)(1-\alpha)\rho\omega^2}{12E(1+3\alpha+\nu-\alpha\nu)}. \end{aligned} \tag{6}$$

Combining Eqs. (5) and (6), we get explicit expressions for *displacement* components, and hence, *strain* and *stress* components. These quantities are expressed as

$$\begin{aligned} u_r &= \frac{(1-\nu)\rho\omega^2 r}{4E(1+3\alpha+\nu-\alpha\nu)}\left[2(1+2\alpha+\nu)R^2-(1+\alpha)(1+\nu)r^2 + \frac{(1-\alpha)(1+\nu)}{3}r^2\cos 4\theta\right] + \frac{1-\nu}{E}\sigma_0 r \\ u_\theta &= -\frac{(1-\nu^2)(1-\alpha)\rho\omega^2}{12E(1+3\alpha+\nu-\alpha\nu)}r^3\sin 4\theta \end{aligned} \tag{7}$$

$$\varepsilon_r = \frac{(1-\nu)\rho\omega^2}{4E(1+3\alpha+\nu-\alpha\nu)}[2(1+2\alpha+\nu)R^2-3(1+\alpha)(1+\nu)r^2 + (1-\alpha)(1+\nu)r^2\cos 4\theta] + \frac{(1-\nu)\sigma_0}{E}$$

$$\varepsilon_\theta = \frac{(1-\nu)\rho\omega^2}{4E(1+3\alpha+\nu-\alpha\nu)}[2(1+2\alpha+\nu)R^2-(1+\alpha)(1+\nu)r^2 - (1-\alpha)(1+\nu)r^2\cos 4\theta] + \frac{(1-\nu)\sigma_0}{E} \tag{8}$$

$$\gamma_{r\theta} = -\frac{(1-\nu^2)(1-\alpha)\rho\omega^2 r}{2E(1+3\alpha+\nu-\alpha\nu)}\sin 4\theta$$

$$\begin{aligned} \sigma_r &= \frac{(1+2\alpha+\nu)(R^2-r^2)}{2(1+3\alpha+\nu-\alpha\nu)}\rho\omega^2 + \sigma_0 \\ \sigma_\theta &= \frac{(1+2\alpha+\nu)R^2-(1+\nu+2\alpha\nu)r^2}{2(1+3\alpha+\nu-\alpha\nu)}\rho\omega^2 + \sigma_0, \quad \tau_{r\theta} = 0. \end{aligned} \tag{9}$$

These functions simultaneously satisfy governing Eqs. (1)–(3) and boundary condition (4). They are therefore the required elastic solutions for the present problem. For isotropic materials, $\alpha=1$, Eqs. (7)–(9) render the same results as those given in [1].

From Eq. (9), it is seen that shearing stress components $\tau_{r\theta}$ is zero everywhere within the disk; the normal stress components are independent of the coordinate θ. This means that, although the deformation of the disk is not cylindrically symmetric when $\alpha\neq 1$, the stress distribution within the disk is always cylindrically symmetric.

The maximum stress values appear at the center of the disk. When the external loading σ_0 is omitted, these values are

$$(\sigma_r)_{\max} = (\sigma_\theta)_{\max} = \frac{(1+2\alpha+\nu)\rho\omega^2R^2}{2(1+3\alpha+\nu-\alpha\nu)}. \tag{10}$$

When $\alpha=1$, we get $\sigma_{max}(\text{iso})=(3+\nu)\rho\omega^2R^2/8$, a formula frequently given in design handbooks.

We define the maximum stress ratio β as

$$\beta=\frac{\sigma_{max}(\text{aniso})}{\sigma_{max}(\text{iso})}=\frac{4(1+2\alpha+\nu)}{(1+3\alpha+\nu-\alpha\nu)(3+\nu)}. \qquad (11)$$

For ordinary cubic anisotropic materials, $\nu>0$, $\alpha>0$. The variation of the β value in this region is not significant. The maximum stress ratio varies between values 8/9 ($\alpha\to\infty$, $\nu\to0$) and 4/3 ($\alpha\to0$, $\nu\to0$).

4 Concluding Remarks

Although the solutions (7)–(9) are identical to the results of Chang [14], we obtained these results by an alternative, more direct approach. In our approach, we do the analysis by following steps:

1. Assume the type of displacements functions, which contain unknown coefficients.
2. Deduce the strain and the stress functions.
3. Substitute the stress function into equilibrium and BC equations, to determine the unknown coefficients.
4. Confirm that all governing equations are satisfied.

If the disk is not solid, i.e., there is a hole in the center, then the inner boundary and the outer boundary influence on each other. As a result, the stress field within the ring is distorted, and it is impossible to find simple displacement functions that satisfy all governing equations. For the ring problem, the approximate analyses based on variation theory (finite element method or Ritz method) are effective. The analyzing steps outlined above can be easily adapted for Ritz analysis, only that the unknown coefficients are determined by the minimum potential theory instead. Moreover, the displacement functions of the solid disk problem, namely (5), provide good hints for assuming the deformation shape of the ring ([15]).

References

[1] Timoshenko, S., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd Ed., McGraw-Hill, New York, pp. 80–83.
[2] Tang, S., 1969, "Elastic Stresses in Rotating Anisotropic Disks," Int. J. Mech. Sci., **11**, pp. 509–517.
[3] Murthy, D. N. S., and Sherbourne, A. N., 1970, "Elastic Stresses in Anisotropic Disks of Variable Thickness," Int. J. Mech. Sci., **12**, pp. 627–640.
[4] Yella Reddy, T., and Srinath, H., 1974, "Elastic Stresses in a Rotating Anisotropic Annular Disk of Variable Thickness and Variable Density," Int. J. Mech. Sci., **16**, pp. 85–89.
[5] Leissa, A. W., and Vagins, M., 1978, "The Design of Orthotropic Materials for Stress Optimization," Int. J. Solids Struct., **14**, pp. 517–526.
[6] Tutuncu, N., 1995, "Effect of Anisotropy on Stresses in Rotating Discs," Int. J. Mech. Sci., **37**, pp. 873–881.
[7] Jain, R., Ramachandra, K., and Simha, K. R. Y., 1999, "Rotating Anisotropic Disc of Uniform Strength," Int. J. Mech. Sci., **41**, pp. 639–648.
[8] Jain, R., Ramachandra, K., and Simha, K. R. Y., 2000, "Singularity in Rotating Orthotropic Disc and Shells," Int. J. Solids Struct., **37**, pp. 2035–2058.
[9] Chang, C. I., 1975, "The Anisotropic Rotating Disks," Int. J. Mech. Sci., **17**, pp. 397–402.
[10] Genta, G., and Gola, M., 1981, "The Stress Distribution on Orthotropic Rotating Disks," ASME J. Appl. Mech., **48**, pp. 559–562.
[11] Lekhnitskii, S. G., 1963, *Theory of Elasticity of an Anisotropic Body*, Holden-Day, San Francisco, pp. 109–110.
[12] Vigdergauz, S., and Givoli, D., 1999, "Thermoelastic Stresses in a Cylinder or Disk With Cubic Anisotropy," Int. J. Solids Struct., **36**, pp. 2109–2125.
[13] Ogawa, A., Zhou, F., Hashimoto, R., and Yonaiyama, M., 1997, "Spin Strength of 3D Woven C/C Turbine Disk Model," 12th Autumn JSGT Symposium, Nara, Oct.
[14] Chang, C. I., 1974, "A Closed-Form Solution for an Orthotropic Rotating Disk," ASME J. Appl. Mech., **41**, pp. 1122–1123.
[15] Zhou, F., Ogawa, A., and Hashimoto, R., 2001, "Strain and Stress Distribution in a Rotating Disk Made by 2D C/C Laminated Composites," Presented at the 13th International Conference on Composite Materials, Beijing, June 25–29.

A New Wave Technique for Free Vibration of a String With Time-Varying Length

S.-Y. Lee
Associate Professor, Department of Mechanical Engineering, Sogang University, Sinsudong, Mapoku, Seoul 121-742, Korea
e-mail: sylee@sogang.ac.kr. Mem. ASME

M. Lee
Associate Professor, Department of Mechanical Engineering, Sejong University, Kunjadong, Kwanjinku, Seoul 143-747, Korea. Assoc. Mem. ASME

We introduce a new technique to analyze free vibration of a string with time-varying length by dealing with traveling waves. When the string length is varied, the natural frequencies and vibration energy are not constant. Thus, free response is not represented by discrete standing waves but by traveling waves, and a given phase of oscillation travels along the string. String tension and nonzero instantaneous transverse velocity at the moving boundary results in energy variation. When the string undergoes retraction, free vibration energy increases exponentially with time, causing dynamic instability. The new wave technique gives the time-varying natural frequency and the exact amount of energy transferred into the vibrating string at the moving boundary.

[DOI: 10.1115/1.1427337]

1 Introduction

Strings and beams, with time-varying length, are the common models of many mechanical systems such as ropes of elevators, mining lifts, and type-supporting wires of typewriters. Time-dependent continuous models are often used for precise mechanical systems such as flexible appendages under retraction motion on spacecrafts, and robotic manipulators with prismatic joints ([1,2]). The *Spaghetti problem* is the typical example for unstable motion of a time-varying system ([3,4]).

A similar continuous model is an axially moving system, including belts, chains, tapes, paper sheets, and pipes conveying fluids, treated by many authors ([5,6]). In this problem, the continuous system translates at a velocity between two fixed supports with constant separation. In recent years, Lee and Mote [7,8] studied the vibration characteristics and energy transfer mechanisms between a moving string and various boundaries or constraints using the traveling wave method.

In general, the free vibration of continuous systems with a constant length is analyzed by applying the method of separation of variables. However, the classical method for treating constant length systems is not applicable to continuous systems with time-varying length, because boundary conditions are dependent on time.

Yamamoto et al. [9] studied the free and forced vibrations of a string with variable length using a perturbation method. Kotera [10] solved the same problem by defining new parameters to apply the separation of variable method to the time-varying string. Ram and Caldwell [11] obtained the D'Almebert solution of a

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, Aug. 29, 2000; final revision, Aug. 8, 2001. Associate Editor: A. K. Mal.

similar system using the image method. Terumichi et al. [12] investigated the free response of an axially moving string with time-varying length and a mass-spring boundary; that is the common model for an elevator system.

In this paper, the time-varying natural frequency and energy of free vibration of a string with variable length are investigated using traveling waves. By calculating the frequency variation and energy transferred while traveling waves reflect from the moving boundary, we derive the exact solutions for the time-dependent frequency and energy.

2 Free Vibration Analysis

2.1 The Equation of Motion. A string with time-varying length is shown in Fig. 1. The linear equation of motion for the transverse vibration, $w(x,t)$ of the string is

$$\rho \frac{\partial^2 w}{\partial t^2} - P \frac{\partial^2 w}{\partial x^2} = f(x,t). \tag{1}$$

Here the string tension P and linear density ρ are constant, and $f(x,t)$ is external distributed force. Both boundaries are vertically fixed, but the right boundary moves at a constant velocity v. The time-dependent boundary conditions are

$$w(0,t) = w(l(t),t) = 0$$

where the string length is $l(t) = l_0 \pm vt$.

2.2 Standing and Traveling Waves. In the infinite string problem, all frequencies are permissible in free vibration. However, for a finite string with boundary conditions, free oscillation is described by standing waves (natural modes) with discrete frequencies. In general, a standing wave can be decomposed into two equal but opposite traveling waves,

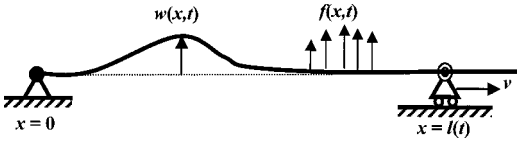


Fig. 1 String with length varying at a constant velocity *v*

$$w(x,t) = A \sin kx \cos \omega t = \frac{1}{2} A \sin(kx - \omega t) + \frac{1}{2} A \sin(kx + \omega t). \tag{3}$$

Here the first and second terms of the right-hand side indicate the right-going and left-going traveling waves, respectively. Wave number is $k = \omega/c$ and the wave propagation speed is $c = \sqrt{P/\rho}$. Consider the string motion with two opposite traveling waves with different frequencies, ω_r and ω_l. Using trigonometric identities, the wave motion becomes

$$w(x,t) = \frac{1}{2} A \sin(kx - \omega_r t) + \frac{1}{2} A \sin(kx + \omega_l t)$$

$$= A \sin\left(kx - \frac{\omega_r - \omega_l}{2} t \right) \cos\left(\frac{\omega_r + \omega_l}{2} t \right). \tag{4}$$

Equation (4) describes that free vibration by the two waves with different frequencies is no longer represented by standing waves. It is seen that the phase is not constant along the string, and there is a phase shift between any two points. A given phase of oscillation travels downstream at a speed $(\omega_r - \omega_l)/2k$ if $\omega_r > \omega_l$ or upstream if $\omega_r < \omega_l$. The speed is generally described as the *phase propagation speed.*

2.3 Time-Varying Natural Frequency. Figure 2(*a*) shows that an initial standing wave of the fundamental frequency at $t = 0$ is decomposed into two equal but opposite traveling waves. When the right-going wave is incident on the moving boundary, the reflected wave has a different frequency. As shown in Fig. 2(*b*), the changed wave frequency is less than the incident one, when the string is being lengthened. The string motion at the instant is a superposition of the two traveling waves with different frequencies. The phase of the superposed motion of the original and reflected traveling waves lags behind that of the original two opposite traveling waves. Thus, it is similar to the case of Eq. (4) that the superposed wave motion is described as a phase-traveling wave where a given phase of oscillation travels downstream. However, in this case, both wave number and frequency change during reflection at the moving boundary. When the string length is decreased, a phase of oscillation travels upstream.

Figure 2(*c*) shows how the wave motion recovers its original phase after a cycle. As the initial left-going wave is reflected from $x = 0$, the reflected wave keeps its original frequency. When this wave propagates and reflects from the moving boundary $x = l(t)$, the wave frequency changes. At this time, all traveling waves have the same frequency, and all the points of the string recover their original phases.

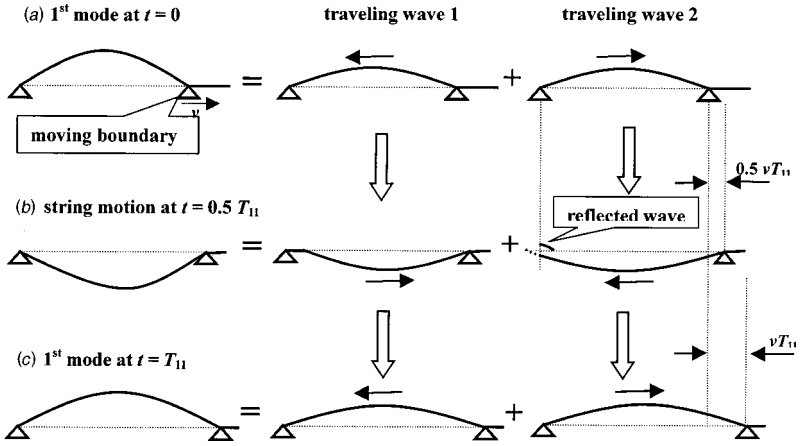


Fig. 2 Traveling wave pattern of the fundamental vibration mode over a period

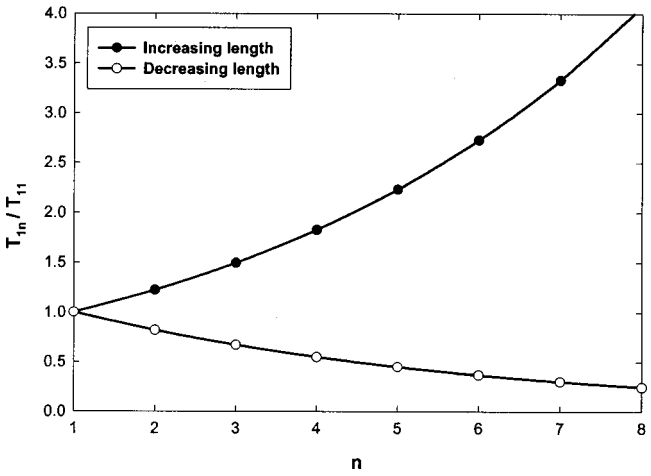


Fig. 3 Time-varying vibration periods of the fundamental mode when the string length changes at $v=0.1c/l_0$

The first period T_{11} of the fundamental mode is defined as the time for the string to recover its original phase. It is actually equal to the time for a traveling wave to pass through the string downstream and upstream. The oscillation period is calculated as follows: For the increasing length problem, the distance for a left-going wave starting at $x=l_0$ to reach $x=l_0+vT_{11}$, after reflection at $x=0$, gives $cT_{11}=2l_0+vT_{11}$. Solving this expression for the first period of the fundamental mode yields

$$T_{11}=\frac{2l_0}{c-v}. \tag{5}$$

The corresponding natural frequency is $\omega_{11}=\pi(c-v)/l_0$. When $v=0$, Eq. (5) becomes the fundamental period of the fixed length string.

The vibration period and natural frequency are also given by the phase-closure principle ([13,14]). As a wave propagates downstream and upstream, the total phase changes become $k(l_0+vT)$ and kl_0, respectively. The phase difference, induced during reflection from both vertically fixed boundaries, is π. If the total phase change is an integer multiple of 2π, the condition describes the natural frequency of the system. The total phase change over the first cycle of the nth vibration mode, satisfying the principle

$$\omega_{n1}\left(\frac{l_0+vT_{n1}}{c}+\frac{l_0}{c}\right)+\pi+\pi=2\pi n, \tag{6}$$

gives the first natural frequency of the nth vibration mode,

$$\omega_{n1}=\frac{\pi(cn-v)}{l_0}, \tag{7}$$

where $n=1,2,3,\cdots$. The natural frequency of the fundamental vibration mode after m cycles is also given by

$$\omega_{1m}=\frac{\pi(c-v)}{l_{m-1}}=\frac{\pi(c-v)^m}{l_0(c+v)^{m-1}} \tag{8}$$

where $l_m=l_{m-1}+vT_{1m}$ is the string length at $t=\Sigma_{k=1}^{m-1}T_{1k}$.

When the length of the string decreases, the corresponding period and natural frequency are obtained by replacing v with $-v$ in Eqs. (5)–(8). Figure 3 shows time-dependent periods $T_{1m}=2\pi/\omega_{1m}$ of the fundamental mode for both increasing and decreasing lengths when $v=0.1c/l_0$.

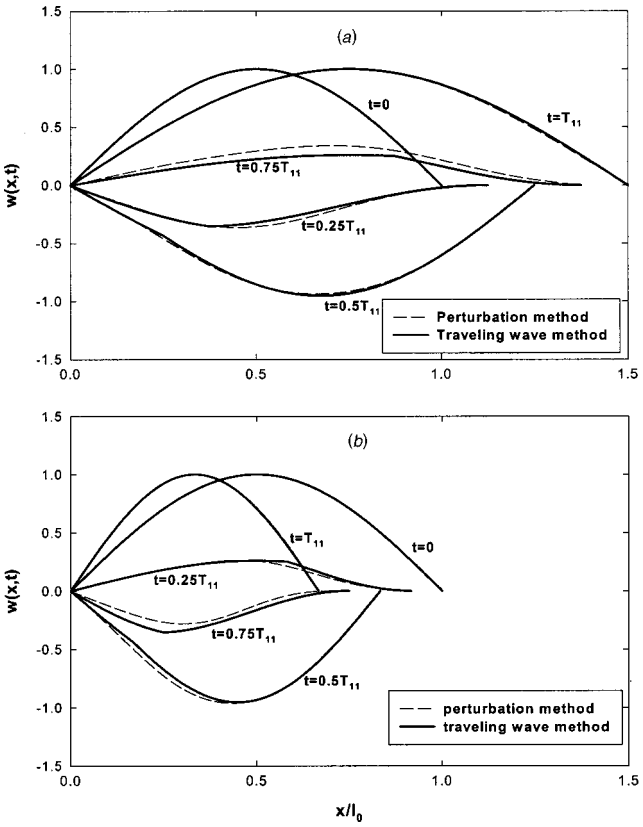


Fig. 4 String motion history during a period when $v=0.2c/l_0$ for the initial displacement of the fundamental mode; (*a*) increasing length, (*b*) decreasing length

2.4 Free Response by the Superposition of Traveling Waves. Time-invariant natural modes do not exist for this time-varying string. However, we can represent free response by decomposing initial standing waves into the corresponding traveling waves, and by superposing the traveling waves at a certain time. Figures 4(*a*) and (*b*) show the motion history of the fundamental mode over a period by superposing traveling waves for both cases of the increasing and decreasing lengths, respectively. The results are plotted with the perturbation solution by Yamamoto et al. [9]. It is noted that the traveling wave method gives the exact solution for free vibration, compared to the approximate solution.

In the figure, all the points of the string start to vibrate with the same phase at $t=0$. As time goes, the vibration at the point downstream lags behind that at the point downstream until $t=T_{11}$. While the string displacements for the constant length problem are zero at $t=0.25T_{11}$ and $0.75T_{11}$, the nonconstant phase make the time-varying string having a different displacement pattern at those times. All the string points recover their original phases after the fundamental period.

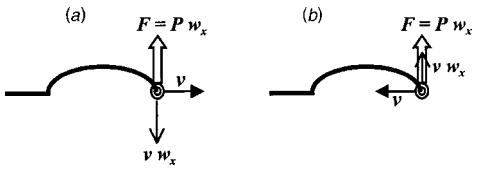


Fig. 5 Power flow by string tension at the moving boundary; (*a*) increasing length, (*b*) decreasing length

3 Free Vibration Energy

3.1 Energy Reflection Coefficient. Lee and Mote [7] defined the energy reflection coefficient to quantify energy transferred when a wave is reflected at various boundary supports. The energy reflection coefficient R is defined as the ratio between the incident wave energy and reflected one,

$$R=\frac{E_r}{E_i}=\frac{Z_r\omega_r}{Z_i\omega_i}\gamma^2. \tag{9}$$

Here $Z=P/c$ is the mechanical impedance of the string, and the subscripts i and r denote incident and reflected waves. $r=A_r/A_i$ is the amplitude ratio of the incident and reflected waves, that is usually called as the *reflection coefficient*. For the fixed length string, the energy reflection coefficient, $R=r^2$, is determined by only the reflection coefficient, because $Z_i=Z_r$ and $\omega_i=\omega_r$. However, when the string length changes with time ($\omega_i\neq\omega_r$) or the string translates axially between the two fixed boundaries ($Z_i\neq Z_r$), the reflected energy is different from the incident one ($R\neq 1$), even though the amplitude of the reflected wave is equal to the incident one ($r=1$).

3.2 Energy Variation. When a traveling wave reflects from the fixed boundary ($x=0$), the amplitude and impedance of the reflected wave are equal to the incident ones. However, when the wave reflects from the moving boundary at $x=l(t)$, the reflected frequency ω_r differs from the incident one ω_i. By the use of Eq. (5), the frequency ratio

$$\frac{\omega_r}{\omega_i}=\frac{l_0}{l_0+vT_{11}}=\frac{c-v}{c+v} \tag{10}$$

is expressed in terms of by the wave speed and the moving velocity. Finally, the energy reflection coefficient is given by

$$R=\frac{E_r}{E_i}=\frac{\omega_r}{\omega_i}=\frac{c-v}{c+v}. \tag{11}$$

When the string length increases, the energy coefficient become less than 1, and free vibration energy decreases. The wave energy E_n after n periods becomes

$$E_n=E_0R^n \tag{12}$$

where E_0 is the initial energy of the string. For the case of decreasing length, the energy coefficient becomes $R=(c+v)/(c-v)>1$. The free vibration energy E_n increases exponentially with time. The source of the energy increase is the external energy required for delivering the boundary support at v. When a wave reflects from the moving boundary, a part of the external energy is transferred from the boundary into the vibration energy.

3.3 Energy Transfer Mechanism. The transverse displacement of the vertically fixed but moving boundary is always zero. However, when the string length increases at v, the instantaneous velocity is $-vw_x(l,t)$, as shown in Fig. 5(a). This nonzero velocity causes power flow together with the vertical component of string tension, $Pw_x(l,t)$. The magnitude of the power flow is given by

$$P_l=-Pvw_x^2(l,t)<0. \tag{13}$$

This value is always negative, and the energy is transferred from the string to the moving boundary. For the decreasing length problem, the energy transfer mechanism is contrary and the associated power flow is always positive (Fig. 5(b)). In this case, the energy is transferred into the string.

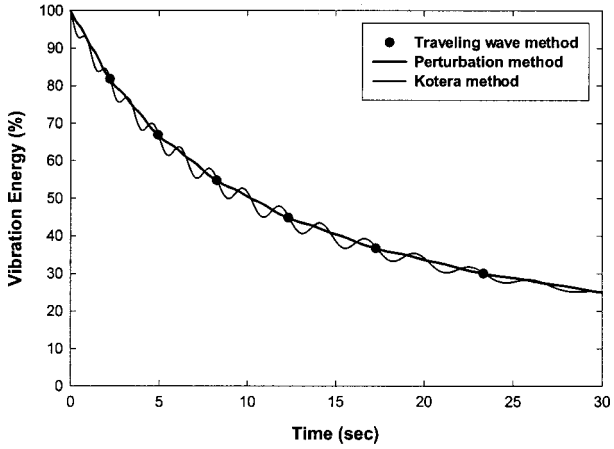


Fig. 6 Free vibration energy when the length increases at $v=0.1c/l_0$

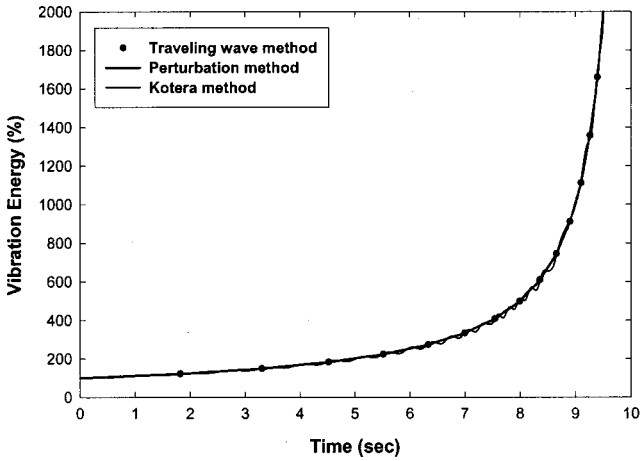


Fig. 7 Free vibration energy when the length decreases at $v=0.1c/l_0$

3.4 Comparisons. The analytical solution (12) for the free vibration energy of the time-varying string is compared with two earlier results solved by the approximated methods; a perturbation method ([9]) and Kotera's method ([10]). Figure 6 shows free vibration energy when the length increases at $v=0.1c/l_0$. The vibration energy by the traveling wave method at each cycle calculated from Eq. (12), and it is marked as "$\cdot$". Vibration energies by the two approximated methods are calculated by implementing the numerical integration of free responses derived in the papers. It is seen that the vibration energy decreases with at a rate $R=0.9/1.1=81.8$ percent over each period. The traveling wave solution, calculated from the simple Eq. (12), gives the exact solution of the free vibration energy. The approximated solution by the perturbation method is in a good agreement with the exact one. However, Kotera's solution shows a sinusoidal disturbance around the exponentially decreasing value.

Figure 7 shows the vibration energy when the length decreases at $v=0.1c/l_0$. As the string length gets closer to zero ($t\rightarrow 10$ sec), the corresponding energy shows a dramatic increase, and it becomes infinite at the zero length. This explains qualitatively the unstable motion of various continuous systems with the decreasing length.

When the moving velocity is increased to $v=0.2c/l_0$, free vibration energies for increasing and decreasing cases are plotted in Figs. 8(a)and (b). The rate of change of free vibration energy is larger than the case of $v=0.1c/l_0$. For the decreasing case, a half

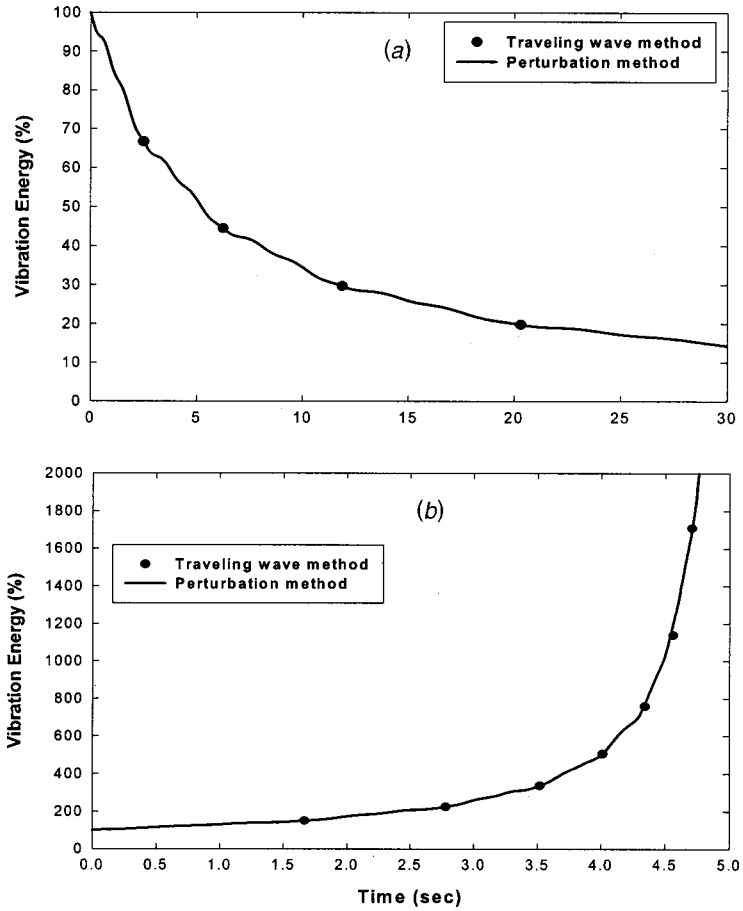


Fig. 8 Free vibration energy when $v=0.2c/l_0$; (*a*) increasing length, (*b*) decreasing length

of the initial energy increases only after a period ($E=1.5E_0$). There exists a little difference between the traveling wave and perturbation results after six periods. This is because the increase of the moving velocity amplifies the error of the approximated solution, while the linear solution by the traveling wave is exact.

4 Conclusions

In this paper, the time-varying frequency and energy of free vibration of a string with variable length are analyzed by a new wave technique. When the string length increases with a constant velocity v, the first natural frequency the nth vibration mode is given by $\omega_{n1}=\pi(cn-v)/l_0$, and vibration energy decreases at a constant ratio $R=(c-v)/(c+v)$ over a cycle. A given phase of oscillation travels from the fixed boundary toward the moving boundary at a phase propagation speed. When the string length is being shortened, free vibration energy increases exponentially with time, causing dynamic instability. In this case, string tension and nonzero instantaneous transverse velocity at the moving boundary results in positive power flow, and energy is transferred from the moving boundary into the vibrating string.

Acknowledgments

This work was supported by grant No. 1999-1-30400-004-2 from the interdisciplinary research program of the KOSEF. The authors also thank Mr. Sangkyu Park for his efforts in drawing figures.

References

[1] Downer, J. D., and Park, K. C., 1993, "Formulation and Solution of Inverse Spaghetti Problem: Application to Beam Deployment Dynamics," AIAA J., **31**, pp. 339–347.
[2] Tadikonda, S. K., and Baruh, H., 1992, "Dynamics and Control of a Translating Flexible Beam With a Prismatic Joint," ASME J. Dyn. Syst., Meas., Control, **114**, pp. 422–427.
[3] Carrier, G. F., 1949, "The Spagetti Problem," Am. Math. Monthly, **56**, pp. 669–672.
[4] Renshaw, A. A., 1997, "Energetics of Winched Strings," ASME J. Vibr. Acoust., **119**, No. 4, pp. 643–644.
[5] Wickert, J. A., and Mote, C. D., 1988, "Current Research on the Vibration and Stability of Axially Moving Materials," Shock Vib. Dig., **20**, pp. 3–13.
[6] Li, G. X., and Paidoussis, M. P., 1993, "Pipes Conveying Fluid: A Model of Dynamical Problem," J. Fluids Struct., **7**, pp. 137–204.
[7] Lee, S.-Y., and Mote, C. D., 1997, "A Generalized Treatment of the Energetics of Translating Continua, Part I: Strings and Tensioned Pipes," J. Sound Vib., **204**, pp. 735–753.
[8] Lee, S.-Y., and Mote, C. D., 1998, "Traveling Wave Dynamics in a Translating String Coupled to Stationary Constraints: Energy Transfer and Mode Localization," J. Sound Vib., **212**, pp. 1–22.
[9] Yamamoto, T., Yasuda, K., and Kato, M., 1978, "Vibrations of a String With Time-Variable Length," Bull. JSME, **21**, pp. 1677–1684.
[10] Kotera, T., 1978, "Vibrations of String With Time-Varying Length," Bull. JSME, **21**, pp. 1469–1474.
[11] Ram, Y. M., and Caldwell, J., 1996, "Free Vibration of a String With Moving Boundary Conditions by the Method of Distorted Images," J. Sound Vib., **194**, No. 1, pp. 35–47.
[12] Terumichi, Y., and Ohtsuka, M., et al., 1993, "Nonstationary Vibrations of a String With Time-Varying Length and a Mass-Spring System Attached at the Lower End," *Dynamics and Vibration of Time-Varying Systems and Structures*, DE-Vol. 56, ASME, New York, pp. 63–69.
[13] Cremer, L., Heckel, M., and Ungar, E. E., 1988, *Structure-Borne Sound*, Springer–Verlag, Berlin.
[14] Mead, D. J., 1994, "Waves and Modes in Finite Beams: Application of the Phase-Closure Principle," J. Sound Vib., **171**, pp. 695–702.

On the Accuracy of Benchmark Tables and Graphical Results in the Applied Mechanics Literature[1]

J. Helsing
Department of Solid Mechanics and NADA, Royal Institute of Technology, SE-100 44 Stockholm, Sweden
e-mail: helsing@nada.kth.se

A. Jonsson
Department of Solid Mechanics, Royal Institute of Technology, SE-100 44 Stockholm, Sweden

Converged normalized stress intensity factors for a matrix crack interacting with an elastic cylinder are presented. The new results differ from previously published results in several examples. The need for better error analysis in computational fracture mechanics is emphasized. [DOI: 10.1115/1.1427691]

Introduction

The purpose of this note is to initiate a discussion of the accuracy of benchmark tables and graphical results presented in the applied mechanics literature. Accurate benchmark results are essential in the development of new software. Programming errors easily occur. If one cannot find at least three-digit accurate results for standard nontrivial setups to verify against, many errors will go unnoticed.

Stress intensity factors are frequently tabulated and presented in graphs. These factors are considered difficult to compute, even though the underlying physical problem often is well conditioned. The chief difficulties are to find and to implement efficient numerical algorithms and to assess the accuracy of the final result. There are many pitfalls. Finding the correct branch of the square root of complex numbers in the context of computing weight functions is just one example of a nonstandard task which may occur and where even the properties of the compiler must be taken into account. Also, the orientation of coordinate systems and the various normalization factors and symbols used by different authors may cause confusion. Not surprisingly, many of the numerical results presented in the literature are of questionable quality. Convergence studies are seldom, if ever, presented. We believe that there is a particular need to reexamine previously published results in this area.

In order to illustrate the points made above we consider an example involving two papers presenting results for normalized stress intensity factors of a matrix crack in the presence of an elastic cylinder: one classic paper by Erdogan, Gupta, and Ratwani [1], and a recent paper by Cheeseman and Santare [2]. In the latter paper the authors validate their algorithm by comparing with results from the former paper. ''Good agreement'' is noted, but the statement is not supported by numerical results.

[1]This work was supported by NFR, TFR, and The Knut and Alice Wallenberg Foundation under TFR contracts 98-568 and 99-380.

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, 2000; final revision, Aug. 21, 2001. Associate Editor: A. Needleman.

Results

We simply recompute some results of Erdogan, Gupta, and Ratwani [1] and Cheeseman and Santare [2] using an algorithm based on a pair of integral equations for the crack and inclusion problem developed by Helsing and Peters [3]. The integral equations, number (48) and number (49) in Helsing and Peters [3], are of Fredholm's second kind with compact operators. This allows for stable convergence. The integral equations are solved using a Nyström scheme with composite quadrature on a uniform mesh. We use 16-point Gauss-Legendre quadrature on all quadrature panels except for the two panels containing the crack tips. There we use Gauss-Jacobi quadrature. Great care is devoted to avoiding round-off error throughout the code. The setups under investigation are depicted in Fig. 1. The shear moduli of the matrix and of the cylinder are $\mu_1=1$ and $\mu_2=23$. The Poisson's ratios of matrix and of the cylinder are $\nu_1=0.35$ and $\nu_2=0.30$. The two-dimensional bulk modulus κ, used in Helsing and Peters, is $\kappa=\mu/(1-2\nu)$. This bulk modulus should not be mixed up with the ''kappa'' used by many authors including Erdogan, Gupta, and Ratwani. The latter ''kappa'' corresponds to the quantity $\varkappa=3-4\nu$, in Muskhelishvili's notation.

Our new converged results do not always agree with the previously published results. This can be seen in Table 1 and in Fig. 2. In many cases the results differ considerably, in digits and also in signs (for the secondary factors k_{21} and k_{22}). It is hazardous to speculate in the reasons for this discrepancy. One thing is certain, however. Our results have converged stably. See Fig. 3 for an example where the relative error for a stress intensity factor settles on a level of 10^{-15} as the mesh is refined.

Discussion

This note stresses the need for more error analysis in computational fracture mechanics. An algorithm may be correct in a mathematical sense. The results it produces on a computer may still be wrong if the problem is not properly resolved, if the algorithm is unstable, or if there is a bug in the code or in the compiler.

It is difficult to prove, rigorously, that a numerical solution to a nontrivial problem is accurate to a certain number of digits. Accurate benchmarks can, in our opinion, best be established by the agreement of several calculations performed by independent investigators. To this end, the presentation of numerical results in terms of numbers is essential. Graphs alone are not sufficient. The presentation of convergence studies, further, helps remove doubts about underresolution and instability. We encourage the inclusion of this type of information in forthcoming papers. We challenge other scientists in computational mechanics to confirm or disprove our new numerical results.

Table 1 Comparison between our new results and those of Table 3 in Erdogan, Gupta, and Ratwani [1] for the geometry of the left image in our Fig. 1. The relations between the shear moduli of the inclusion and the matrix is $\mu_2=23\mu_1$. The normalized mode *i* stress intensity factor at crack-tip *j* is denoted as k_{ij}.

c/a	k_{11}^{EGR}	k_{11}^{new}	k_{12}^{EGR}	k_{12}^{new}	k_{21}^{EGR}	k_{21}^{new}	k_{22}^{EGR}	k_{22}^{new}
0.3	0.784	0.790	0.225	0.235	−0.004	−0.023	0.072	0.073
0.5	0.792	0.797	0.341	0.347	−0.006	−0.037	0.101	0.102
1.0	0.817	0.817	0.613	0.613	−0.005	−0.067	0.057	0.061
1.5	0.839	0.833	0.763	0.755	0.008	−0.074	−0.007	0.012
2.0	0.860	0.850	0.845	0.830	0.034	−0.058	−0.021	0.018
3.0	0.905	0.897	0.953	0.936	0.089	−0.035	−0.001	0.067
4.0	0.951	0.947	1.014	1.003	0.117	0.032	0.002	0.079
8.0	1.020	1.022	1.043	1.043	0.088	0.032	−0.026	0.032

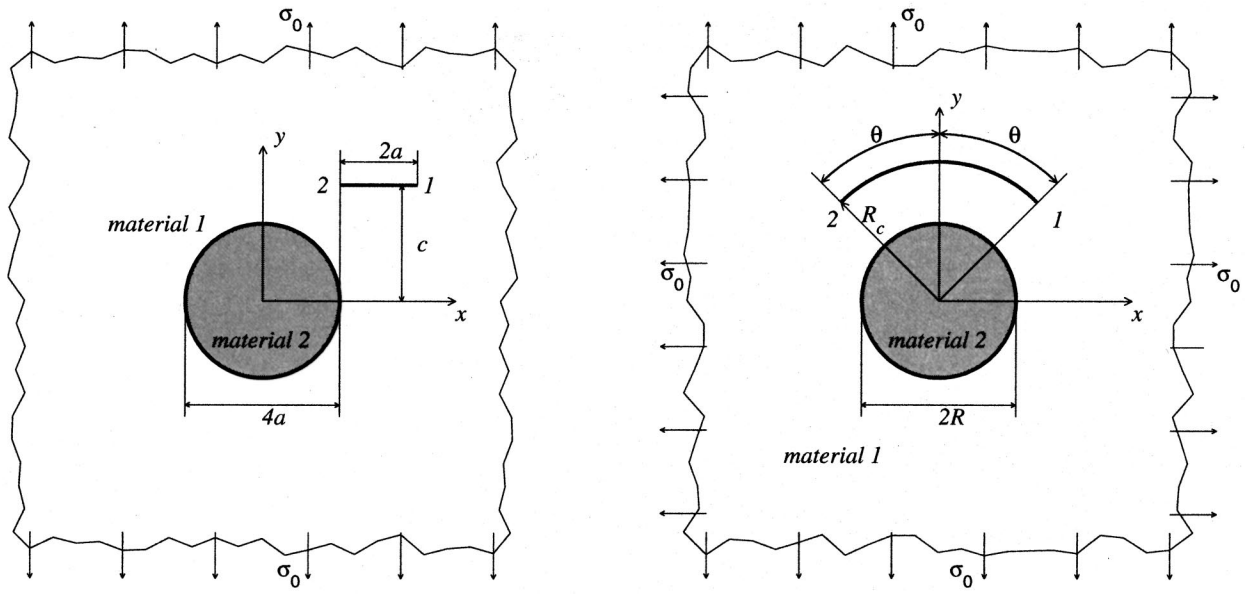


Fig. 1 Left, a straight crack outside an inclusion under uniaxial tension. This is the setup of Erdogan, Gupta, and Ratwani [1] corresponding to their Table 3. Right, an arc-shaped crack outside a circular inclusion under biaxial tension. This is the setup of Cheeseman and Santare [2] corresponding to their Fig. 8.

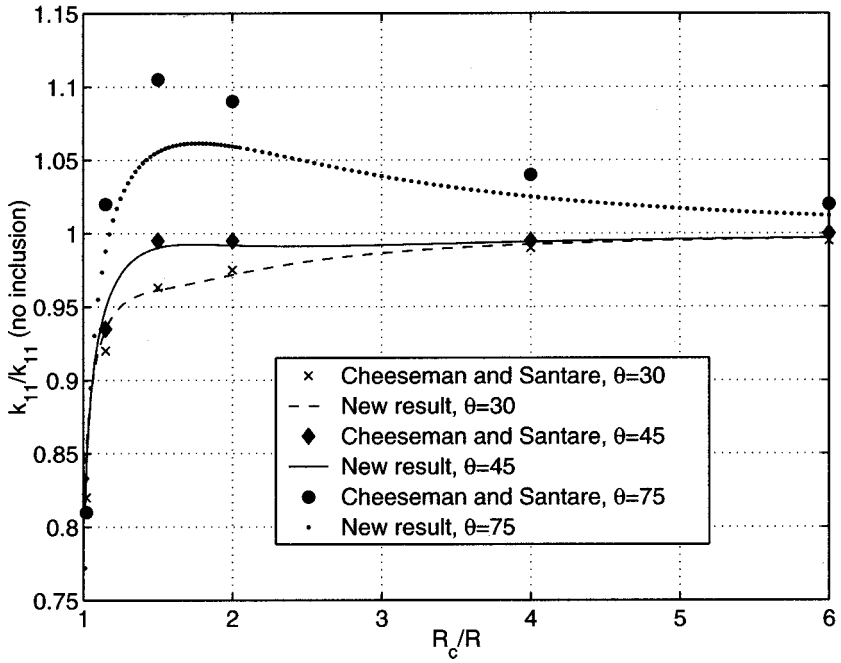


Fig. 2 Normalized mode I stress intensity factors of the setup in Fig. 8 in Cheeseman and Santare [2] (the right image of our Fig. 1) versus dimensionless distance for a circular arc-shaped crack interacting with an inclusion

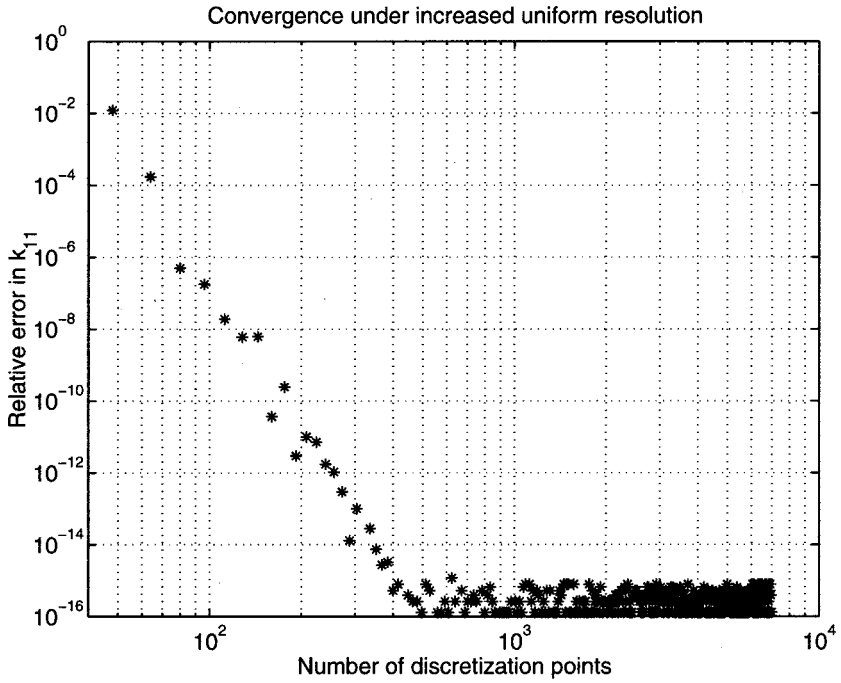


Fig. 3 Convergence of the stress intensity factor k_{11} of Erdogan, Gupta, and Ratwani [1] for $c=2a$ in the left image of our Fig. 1. The mesh is uniformly refined. The number of distgretization points is N. Double precision arithmetic is used. The reference value $k_{11}=0.8497339474770513$ is computed with 592, or more, discretization points in quadruple precision arithmetic. Relative errors smaller than machine epsilon as displayed as $1.11\cdot10^{-16}$.

References

[1] Erdogan, F., Gupta, G. D., and Ratwani, M., 1974, ''Interaction Between a Circular Inclusion and an Arbitrarily Oriented Crack,'' ASME J. Appl. Mech., **41**, pp. 1007–1013.

[2] Cheeseman, B. A., and Santare, M. H., 2000, ''The Interaction of a Curved Crack With a Circular Elastic Inclusion,'' Int. J. Fract., **103**, pp. 259–277.

[3] Helsing, J., and Peters, G., 1999, ''Integral Equation Methods and Numerical Solutions of Crack and Inclusion Problems in Planar Elastostatics,'' SIAM (Soc. Ind. Appl. Math.) J. Appl. Math., **59**, pp. 965–982.

Journal of Applied Mechanics

Book Review

Error Analysis With Applications in Engineering, by W. Szczepinski and Z. Kotulski. Lastran Corp., Rochester, NY, 2000.

REVIEWED BY M. OSTOJA–STARZEWSKI[1]

This is a well-laid-out and well-written introductory book for engineers (especially mechanical engineers) interested in analysis of random effects in mechanics. The key word here is introductory—as it would be suitable for an undergraduate (but not really a graduate) course—as opposed to an advanced level which would involve stochastic differential equations plus possibly random processes and fields. As such, the book therefore offers an introduction to statistical error analysis methodology for anyone in solid/structural and/or rigid-body mechanics. With this book, a mechanical (but also civil, aerospace, materials...) engineer can learn concepts of applied probability theory—especially calculus of random variables—through very clear expositions, numerous mechanics problems and examples (e.g., positioning accuracy of robot manipulators). Indeed, many people do need such physical motivation to go through the basics of probability—and this may define the potentially wide market for the book.

The book contains seven chapters plus an appendix. The chapters are 1 Basic characteristics of error distribution; histograms; 2 Sample points, random variables, and probability; 3 Functions of independent random variables; 4 Two-dimensional distributions; 5 Two-dimensional functions; 6 Three-dimensional distributions; 7 Three-dimensional functions of independent random variables; and the Appendix—Some useful definitions and facts of probability theory. In all these chapters much attention has been paid to the practical significance of error analysis, although some background on pertinent mathematical foundations has also been included.

Both authors are well-established mechanicians—the first one being known for his work in plasticity theory, and the second one for his studies in stochastic mechanics (specifically, stochastic wave propagation). Perhaps the only criticism that might be raised is that there are other books on applied probability methods in engineering in the English speaking world. But the presentation and scope of ''*Error...*'' are not in overlap with any one of these, and some aspects are truly original. The book is therefore recommended.

[1]Department of Mechanical Engineering, McGill University, 817 Sherbrooke Street West, Montreal, Quebec H3A 2K6, Canada.

