Journal of Applied Mechanics

Published Bimonthly by ASME

VOLUME 74 • NUMBER 4 • JULY 2007

Transactions of the ASME, Journal of Applied Mechanics (ISSN 0021-8936) is published bimonthly (Jan., Mar., May, July, Sept., Nov.) by The American Society of Mechanical Engineers, Three Park Avenue, New York, NY 10016. Periodicals postage paid at New York, NY and additional mailing offices. 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).

Canadian Goods & Services Tax Registration #126148048.

MEMORIAM

TECHNICAL PAPERS

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

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

(**Contents continued**)

Journal of Applied Mechanics **Volume 74, Number 4** **JULY 2007**

TECHNICAL BRIEFS

ERRATA

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

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

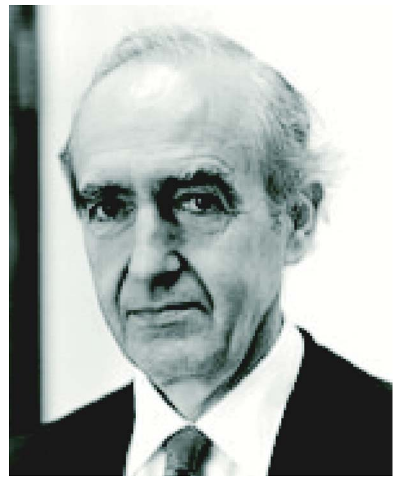
Journal of Applied Mechanics

Memoriam

In Memoriam: Erastus H. Lee

Erastus H. Lee, professor emeritus of Stanford University and RPI, and a prominent researcher, with fundamental contributions to plasticity, viscoelasticity and wave propagation, died at the age of 90 on May 17, 2006 in Lee, New Hampshire.

Ras Lee was born on February 2, 1916, in Southport, England. He graduated from Cambridge University in 1937 with a First Class Honours degree in mechanical sciences and mathematics. After a year of postgraduate study at Cambridge with Professor C. E. Inglis, Ras was awarded a fellowship from the Commonwealth Fund of New York to study with Professor Stephen P. Timoshenko at Stanford University. There he met Shirley, and later completed his Ph.D. and married her, both in 1940. Immediately thereafter, he became involved in the British war effort. He worked first as a progress officer in the British Purchasing Commission in New York and later in the British Air Commission in Washington. Ras was responsible for planning aircraft deliveries from U.S. companies and for specifying modifications required to meet British needs. He and Shirley returned to England during the war, where Ras first worked at the Ordnance Board and then at the Ministry of Supply Armament Research Department. He was elected a Fellow of his College, Gonville and Caius, Cambridge, in 1944, and became Assistant Director in charge of the Technical Engineering Section of the Production Department of the newly established British Atomic Energy Authority in 1946.



Erastus H. Lee
February 2, 1916–May 17, 2006

After an offer from Professor William Prager, Lee and his family returned to the United States in 1948, where he was a Professor of Applied Mathematics at Brown University for 14 years. He served as Chairman of the Applied Mathematics Division for five years. During these years, faculty members in the Divisions of Applied Mathematics and Engineering, which included Dan Drucker, Harry Kolsky, Allen Pipkin, Paul Symonds, Ronald Rivlin, Dick Shield, and Eli Sternberg in addition to Prager and Lee, made Brown the worldwide center for research in solid mechanics. In 1962, Ras was appointed as a Professor in the Division of Applied Mechanics and the Department of Aeronautics and Astronautics at Stanford University, joining Norman Goodier, Wilhelm Flügge, Nick Hoff, and Miklos Hetenyi in the widely acclaimed Stanford applied mechanics group. Almost every graduate student in solid mechanics during that time took Lee's sequence of three courses (each two quarters long) in nonlinear continuum mechanics, viscoelasticity, and plasticity. He remained at Stanford for 20 years (1962–1982), taking mandatory retirement at the age of 65. For the last 10 years of his professional career, Ras was the Rosalind and John J. Redfern, Jr. Chair Professor of Engineering at Rensselaer Polytechnic Institute.

In his early work, Lee made fundamental contributions to the development of solutions for elastic-plastic problems and slip-line methods for metal forming processes. This includes a series of papers, written in England with R. Hill and S. J. Tupper, on the theory of the autofrettage process, wedge indentation in ductile metals, and compression of a block between rough plates. This was followed by research with his students at Brown on the stress discontinuities in plane plastic flow, the analysis of plastic flow in deeply notched bars, and discontinuous machining and chip formation. At Brown, Lee also made significant contributions to the analysis of boundary value problems in the theory of plastic wave propagation, including the determination of moving plastic-elastic boundaries, known as loading and unloading waves, with particular application to normal impact between a cylinder and rigid target at rest. Extending his research interests to polymers, he contributed significantly to the development of solution methods for viscoelastic stress analysis, by reducing them to more tractable elasticity problems, which is now known as the correspondence principle. He studied the effects of residual stresses and temperature variations on viscoelastic response (the well-known time-temperature shift), viscoelastic contact problems, and viscoelastic wave propagation. His research papers in this field are regularly referenced in contemporary publications, monographs, and books devoted to viscoelasticity.

Lee continued his research on inelastic wave propagation at Stanford, by developing a finite-strain elastic-plastic theory with application to plane-wave analysis, as arises in dynamic plate impact problems, which culminated in his 1969 paper "Elastic-Plastic Deformation at Finite Strain," published in the *Journal of Applied Mechanics*. Through this research, he developed a framework for the constitutive analysis of large elastic-plastic deformations based on the multiplicative decomposition of the deformation gradient ($F=FeFp$), now commonly referred to as Lee's decomposition. This decomposition had a great impact on subsequent developments of elastoplastic constitutive theories for polycrystalline materials and single crystals. With his students at Stanford and RPI, Lee applied his decomposition to develop rate-type theories of elastoplastic deformation at finite strain for both isotropic and anisotropic materials. His other contributions to mechanics include studies of shock waves in elastic-plastic solids, wave propagation in composite materials with periodic structure, elastic-plastic stress and deformation analysis of metal-forming processes, providing the first calculations of residual stresses, and the modeling of anisotropic strain hardening.

Lee was elected to the National Academy of Engineering in 1975 and was awarded the Timoshenko Medal in 1976, in recognition of his distinguished contributions to the field of applied mechanics. He was a Fellow of the American Academy of Mechanics, and a Life Fellow of the American Society of Mechanical Engineers, with frequent publications in the *Journal of Applied Mechanics* throughout his career. With contributions from his colleagues and former students, an anniversary volume, entitled *Topics in Plasticity*, was published in 1991 by AM Press on the occasion of his 75th birthday. He delivered invited lectures throughout the world; he was a Guggenheim Fellow in 1975 and an Alexander von Humboldt Fellow in 1986.

Ras Lee is survived by his four children and four grandchildren. He was predeceased by his wife, Shirley. Ras had a delightful personality and was well liked and admired by his many colleagues. He also inspired admiration and gratitude among his many post-doctoral and graduate students.

Ras Lee will be sorely missed, but his mechanics legacy will live on.

Vlado Lubarda

Jeng Luen Liou
Assistant Professor
Department of Military Meteorology Engineering,
Air Force Institute of Technology,
Kaohsiung 820, Taiwan,
R.O.C.

Jen Fin Lin[1]
Professor
Department of Mechanical Engineering,
National Cheng Kung University,
Tainan 701, Taiwan,
R.O.C.
e-mail: jflin@mail.ncku.edu.tw

A New Microcontact Model Developed for Variable Fractal Dimension, Topothesy, Density of Asperity, and Probability Density Function of Asperity Heights

In the present study, the fractal theory is applied to modify the conventional model (the Greenwood and Williamson model) established in the statistical form for the microcontacts of two contact surfaces. The mean radius of curvature (R) *and the density of asperities* (η) *are no longer taken as constants, but taken as variables as functions of the related parameters including the fractal dimension* (D)*, the topothesy* (G)*, and the mean separation of two contact surfaces. The fractal dimension and the topothesy varied by differing the mean separation of two contact surfaces are completely obtained from the theoretical model. Then the mean radius of curvature and the density of asperities are also varied by differing the mean separation. A numerical scheme is thus developed to determine the convergent values of the fractal dimension and topothesy corresponding to a given mean separation. The topographies of a surface obtained from the theoretical prediction of different separations show the probability density function of asperity heights to be no longer the Gaussian distribution. Both the fractal dimension and the topothesy are elevated by increasing the mean separation. The density of asperities is reduced by decreasing the mean separation. The contact load and the total contact area results predicted by variable D,* G^**, and* η *as well as non-Gaussian distribution are always higher than those forecast with constant D,* G^**,* η*, and Gaussian distribution.*
[DOI: 10.1115/1.2338059]

1 Introduction

In order to analyze tribological problems such as sealing [1], thermal and electrical contact resistance [2,3], and friction and wear [4] between two rough surfaces, a deep understanding of the deformation behavior of contacting asperities and an accurate characterization of the contact is fundamentally important.

The field of microcontacts was pioneered by Greenwood and Williamson (the GW model) [5], with their elastically basic contact model, or "asperity-based model." The plastic deformation of asperities has also been studied by Abbott and Firestone [6], who constructed their "surface microgeometry model" for the plastic deformation; they assumed that the real contact area of a rough surface with a rigid and flat surface is the geometrical intersection of the flat with the undeformed profile of the asperity, and the contact pressure is equal to the flow pressure.

Chang et al. [7] proposed an elastoplastic asperity model (the Chang, Etsion, and Bogy model) for the analysis of contact surfaces. Based on the concept of volume conservation, they connected the limiting cases of the purely elastic and fully plastic deformations without considering the elastoplastic deformation regime. Zhao et al. [8] presented an elastic-plastic asperity microcontact model (the Zhao model), which was modeled as logarithmic and fourth-order polynomial functions for contact between two nominally flat surfaces.

The finite element method was used to solve the elastoplastic contact of a single asperity [9–12]. An elastic-plastic finite element model for the frictionless contact of a deformable sphere pressed by a rigid flat was presented by Kogut and Etsion [13]. The evolution of the elastic-plastic contact with increasing interference was analyzed, revealing three distinct deformation regimes that range from fully elastic through elastoplastic to a fully plastic contact interface.

Nevertheless, all of these studies mentioned above were based on the assumptions made by Greenwood and Williamson [5]. They assumed that the model surface was composed of hemispherical asperities with the same radius. The heights of summits of the asperities varied randomly in Gaussian distribution, and the interactions between neighboring asperities on the same surface are neglected. Thus, the mean radius of curvature of asperities, R, and the density of asperities, η, are assumed invariant, and the probability density function of asperity heights is also assumed to be always in the Gaussian distribution. However, this is unrealistic when two rough surfaces have contact deformations. The motivation of this study tries to improve these inappropriate assumptions made in the GW model.

Another branch for the study of microcontacts has developed by applying the fractal theory. A rough surface has fractal-like features; it has wiggly features over a large range of length scales and sometimes does follow the self-similar hierarchy. Mathematical fractals follow self-repetition over all these length scales, so rough surfaces have higher and lower length scale limits between which the fractal behavior is observed. Majumdar and Bhushan [14] and Bhushan and Majumdar [15] used scale-independent pa-

[1]Author to whom correspondence should be addressed.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received April 26, 2005; final manuscript received April 19, 2006. Review conducted by A. Maniatty. Discussion on the paper should be addressed to the Editor, Prof. Robert M. McMeeking, Journal of Applied Mechanics, Department of Mechanical and Environmental Engineering, University of California–Santa Barbara, Santa Barbara, CA 93106-5070, and will be accepted until four months after final publication in the paper itself in the ASME Journal of Applied Mechanics.

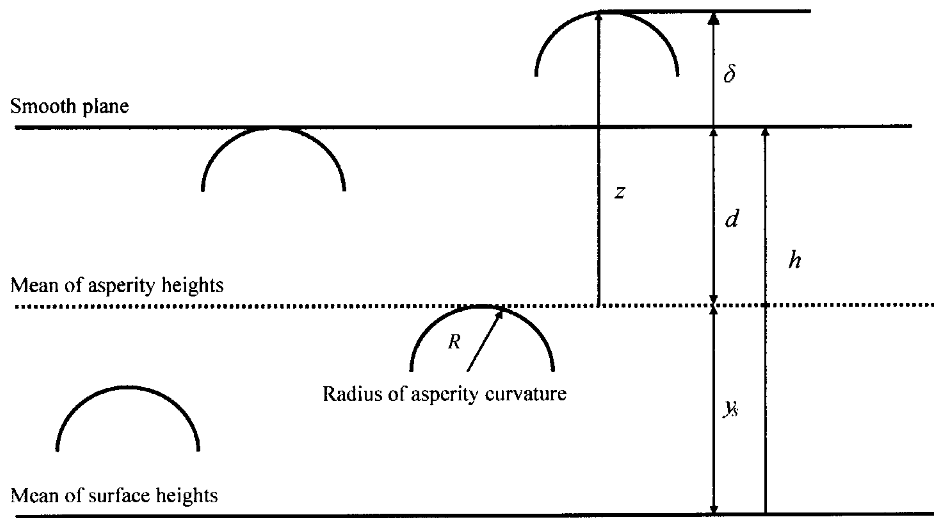


Fig. 1 The schematic diagram of two contact surfaces with deformation

rameters (fractal dimension D), instead of using the conventional statistical parameter, to describe the load contact of rough surfaces. A new fractal-based functional model for anisoptropic rough surface developed by Blackmore and Zhou [16] was used to devise and test two methods for the approximate computation of the fractal dimension $(2<D<3)$ of surfaces. A two-variable fractal surface description was incorporated in a three-dimensional elastic-plastic contact mechanics analysis by Yan and Komvopoulos [17]. There are some other researchers who have studied microcontact with the self-affine fractal distribution of surface roughness. Persson [18] developed a theory of rubber friction when a rubber block slid over a hard rough surface with roughness on many different length scales. The explicit results are presented for self-affine fractal surface with elastic and elastoplastic considerations. Persson [19] further developed a theory of adhesion between an elastic solid and a hard randomly rough substrate. Zhang and Zhao [20] developed a theoretical model to describe the adhesion between plastically deformable fractal surfaces whose asperity heights conformed to a general distribution.

Nevertheless, all of the above studies related to the fractal theory were conducted by assuming the fractal dimension and the topothesy to be invariant with the mean separation, and the probability density function is assumed to be always in the Gaussian distribution. However, in real circumstances the topography of each surface will be continuously changed when two rough surfaces occur contact deformations. According to the experimental results shown in the study of Othmani and Kaminsky [21], surface asperities after experiencing contacts of different separations were found to be satisfied by a non-Gaussian probability density function. Their experimental results also provided the motivation for Chung and Lin [22] to investigate the behavior of these contact parameters varying with the mean separation. In their study, the fractal dimensions of surface asperities at different interferences were obtained on the basis of the experimental data of the number of contact spots $(N(a))$ with their contact area larger than a, which were reported in the study of Othmani and Kaminsky [21]. The topothesy corresponding to its fractal dimension was thus determined from the relationship among the scaling coefficient C_p, the fractal dimension D, and the topothesy G. This relationship was established by the equivalence of the structure functions developed by two different ways. Instead of the fractal analyses on the basis of the experimental results, fractal dimension D and topothesy G varying with the separation of two contact surfaces are now predicted purely by the theoretical model developed in the present study. This model of variable fractal parameters is applied to modify the conventional models established in the statistical form for the microcontacts of two contact surfaces. The mean radius of curvature (R) and the density of asperities (η) in the present microcontact model are thus no longer taken as constants, but taken as variables as a function of the related parameters including the fractal dimension (D), the topothesy (G), and the mean separation of two contact surfaces.

In the present study, the relationship between the fractal dimension and the mean separation is analyzed first. For a fractal surface, the number of contact spots $(N(a))$ with their contact area larger than a satisfies the power-law relation [23]. The slope of a $N(a)-a$ curve is thus equal to $(1-D)/2$. The relationship between $N(a)$ and a is determined purely by theoretical analyses. Through the equality of the real contact area formulas expressed by two different forms, the topothesies evaluated at different deformation regimes can be expressed as a function of the fractal dimension and the mean separation. A numerical scheme is then developed in this study to incorporate these relationships to determine the convergent values of fractal dimension and topothesy corresponding to a given mean separation. The topographies of a surface obtained from the theoretical prediction of different separations show the probability density function (g) to be no longer in the Gaussian distribution. In this model, the elastic-plastic microcontact behavior of two rough surfaces is developed to investigate the effect of variable radius of curvature (R), density of asperity (η), and non-Gaussian probability density function of asperity heights (g) on the total contact area and the contact load.

2 Theoretical Analysis for Contact Surfaces

The contact of two rough surfaces (see Fig. 1) can be modeled by a flat and smooth surface in contact with a rough surface. z is the height of an asperity measured from the mean surface of asperity heights. The asperity interference δ is given as

$$\delta=z-d \tag{1}$$

where d denotes the separation between two contact surfaces. If the mean radii of curvature of the asperities on surface 1 and surface 2 are R_1 and R_2, respectively, the equivalent rough surface can be expressed as having the radius of curvature, R, satisfying $1/R=1/R_1+1/R_2$. $(\sigma_1)_0$ and $(\sigma_2)_0$ denote the standard deviations of the asperity heights of surface 1 and surface 2 before contact deformations occur, respectively. The standard deviation, σ_0, for this equivalent rough surface before any contact deformation satisfies $\sigma_0=\sqrt{(\sigma_1)_0^2+(\sigma_2)_0^2}$.

2.1 Contacts Parameters at Elastic and Fully Plastic Deformations. According to the Hertz theory, the elastic contact area, a_e, the elastic contact load, F_e, and the average contact pressure, P_e, produced by a sphere with a radius of R in contact with a smooth flat with an elastic interference, δ, are given as [24]

$$a_e = \pi R \delta \tag{2}$$

$$F_e = \tfrac{4}{3} E R^{1/2} \delta^{3/2} \tag{3}$$

$$P_e = \frac{4}{3}\frac{E}{\pi}\left(\frac{\delta}{R}\right)^{1/2} \tag{4}$$

where E denotes the effective Young's modulus of two solid contact surfaces (surface 1 and surface 2) with the Young's moduli, E_1 and E_2, and the Poisson ratios, ν_1 and ν_2, respectively. It is stated as

$$\frac{1}{E} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2}$$

In the fully plastic deformation regime, the asperity's contact area, a_p, the contact load, F_p, and the average contact pressure, P_p, can be expressed as [24]

$$a_p = 2\pi R \delta \tag{5}$$

$$F_p = H a_p \tag{6}$$

$$P_p = H \tag{7}$$

where H is the hardness of the softer material of two contact solids.

2.2 The Critical Interference and Contact Parameters in Elastoplastic Deformation Regime. The critical interference, δ_c, which marks the transition from the elastic deformation to elastoplastic deformation, is given by [25]

$$\delta_c = \left(\frac{\pi K H}{2E}\right)^2 R \tag{8}$$

where the maximum contact pressure factor K is related to the Poisson ratio (ν) of the softer material. It is expressed as [8]

$$K = 0.454 + 0.41\nu$$

Kogut and Etsion [13] used a finite element method to solve the elastoplastic contact problem of a single asperity and found that the entire elastoplastic regime extends over the dimensionless interference values in the range of $1 \leqslant \delta/\delta_c \leqslant 110$. The asperity's contact area, a_{ep}, the contact load, F_{ep}, and the average contact pressure, P_{ep}, in the elastoplastic deformation regime are presented in a dimensionless form as [13]

$$\frac{a_{ep}}{\pi R \delta_c} = a_1\left(\frac{\delta}{\delta_c}\right)^{b_1} \tag{9}$$

$$\frac{F_{ep}}{2/3 K H \pi R \delta_c} = a_2\left(\frac{\delta}{\delta_c}\right)^{b_2} \tag{10}$$

$$\frac{P_{ep}}{H/2.8} = a_3\left(\frac{\delta}{\delta_c}\right)^{b_3} \tag{11}$$

where $a_1, b_1, a_2, b_2, a_3, b_3$ are constants, they are summarized in distinct elastoplastic regimes given in this study [13].

2.3 Interference and Radius of Curvature of an Asperity in the Fractal Model. The asperity interference, δ, and the effective radius of curvature, R, of an asperity obtained by the fractal analyses are given as [17]

$$\delta = 2^{(4-D)} G^{(D-2)} (\ln \alpha)^{1/2} \pi^{(D-3)/2} a^{(3-D)/2} \tag{12a}$$

$$R = \frac{2^{(D-4)} G^{(2-D)} a^{(D-1)/2}}{\pi^{(D-1)/2} (\ln \alpha)^{1/2}} \tag{12b}$$

where $D (2 < D < 3)$ is the fractal dimension of the surface, G is the topothesy, which is a height scaling parameter independent of spatial frequency, a is the contact area of an asperity, and α represents a parameter that determines the density of frequency shown in the surface asperities, chosen here to be 1.5 [17]. The dimensionless topothesy, G^*, and the dimensionless contact area, a^*, are normalized by the standard deviation, σ_0, as $G^* \equiv G/\sigma_0$ and $a^* \equiv a/\sigma_0^2$.

The dimensionless interference $\delta^* (\delta^* \equiv \delta/\sigma_0)$ and the dimensionless effective radius of curvature $R^* (R^* \equiv R/\sigma_0)$ can thus be expressed in the fractal form as [23]

$$\delta^* = 2^{(4-D)} G^{*(D-2)} (\ln \alpha)^{1/2} \pi^{(D-3)/2} a^{*(3-D)/2} \tag{13a}$$

$$R^* = \frac{2^{(D-4)} G^{*(2-D)} a^{*(D-1)/2}}{\pi^{(D-1)/2} (\ln \alpha)^{1/2}} \tag{13b}$$

δ^* becomes δ_c^* if a^* in Eq. (13a) is replaced by a_c^*. Substitutions of this δ_c^* expression and Eq. (13b) into Eq. (8) give a_c^* as

$$a_c^* = \left[2^{(10-2D)} \pi^{(D-4)} G^{*(2D-4)} (\ln \alpha) \left(\frac{E}{KH}\right)^2\right]^{1/(D-2)} \tag{14}$$

Asperities with $a < a_c$ are elastically deformed since they satisfy the condition of $\delta < \delta_c$; whereas asperities with $a \geqslant a_c$ are thus considered to be operating in the elastoplastic deformation, even in the fully plastic deformation regime.

In the elastic deformation regime ($\delta/\delta_c \leqslant 1$), the substitutions of Eqs. (1) and (2) into Eq. (13b) give the expression of the dimensionless radius curvature R_e^*; in the first elastoplastic deformation regime ($1 < \delta/\delta_c \leqslant 6$), the substitutions of Eqs. (1) and (9) into Eq. (13b) give the expression of R_{ep1}^*; Similarly, in the second elastoplastic deformation regime ($6 < \delta/\delta_c \leqslant 110$), the dimensionless radius of curvature R_{ep2}^* can be expressed. In the fully plastic deformation regime ($110 < \delta/\delta_c$), the substitutions of Eqs. (1) and (5) into Eq. (13b) give the expression of R_p^*. The dimensionless radii of curvature in the elastic, elastoplastic and fully plastic deformation regimes can be expressed in a general fractal form as

$$R^* = \left(\frac{1}{8}\right)^{C_1} \left[\frac{2^{2C_2(D-4)} G^{*2(2-D)} (z^* - d^*)^{C_3(D-1)} C_4^{(D-1)} \delta_c^{*C_5(1-D)}}{(\ln \alpha)}\right]^{1/(3-D)} \tag{15}$$

where $z^* \equiv z/\sigma_0$; $d^* \equiv /\sigma_0$; $R_e^* \equiv R_e/\sigma_0$; $R_{ep1}^* \equiv R_{ep1}/\sigma_0$; $R_{ep2}^* \equiv R_{ep2}/\sigma_0$; and $R_p^* \equiv R_p/\sigma_0$. The parameters of C_1, C_2, C_3, C_4, C_5 in Eq. (15) are constants, they are summarized in distinct deformation regimes given in Table. 1. Equation (15) can be applied to calculate the radii of curvature in different deformation regimes if the fractal dimension D and the topothesy G^* corresponding to an interference are available.

Table 1 The parameters of C_1, C_2, C_3, C_4, C_5 in Eq. (15)

Deformation regime	C_1	C_2	C_3	C_4	C_5
$R_e^* (\delta/\delta_c < 1)$	0	1	1	1	0
$R_{ep1}^* (1 < \delta/\delta_c \leqslant 6)$	0	1	1.136	0.93	0.136
$R_{ep2}^* (6 < \delta/\delta_c \leqslant 110)$	0	1	1.146	0.94	0.146
$R_p^* (\delta/\delta_c \geqslant 110)$	1	0	1	1	0

2.4 Area Density of Asperities in a Fractal Surface. Define η to be the area density of asperities. According to the study of Nayak [26], the η parameter can be expressed as

$$\eta = \frac{1}{6\pi\sqrt{3}}\frac{m_4}{m_2} \tag{16}$$

where m_2, m_4 are the second, and fourth moments of the power spectral density, respectively. The relationship between m_2, m_4 and the fractal parameters G and D can also be obtained from the power spectrum of the surface profile, and they are given as follows [27]:

$$m_2 = \frac{1}{2}\frac{(3-D)}{(D-2)\sin[\pi(2D-5)/2]\Gamma(2D-5)}\left(\frac{G}{l}\right)^{2(D-2)} \tag{17}$$

$$m_4 = \frac{1}{2G^2}\frac{(3-D)}{(D-1)\sin[\pi(2D-5)/2]\Gamma(2D-5)}\left(\frac{G}{l}\right)^{2(D-1)} \tag{18}$$

where l is the resolution of the surface measuring instrument. Substitutions of Eqs. (17) and (18) into Eq. (16) give

$$\eta = \frac{1}{6\pi\sqrt{3}}\frac{(D-2)}{(D-1)}l^{-2} \tag{19}$$

According to Eq. (19), it is obvious that the asperity density η has its magnitude dependent upon the instrument resolution l and the fractal dimension D. The l parameter is independent of the contact procedure, it can thus be obtained from Eq. (19) by substituting the initial values of the fractal dimension (D_0) and the area density of asperities (η_0) before surface contacts.

2.5 Non-Gaussian Probability Density Function Varying With Mean Separation. The topographies obtained from the experimental results [21] of surface contacts at different separations (or different interferences) are generally varied by the non-Gaussian distribution. The probability density function of surface asperities actually varies with the mean separation of two surfaces and thus is expressed as a function of z^*. The equation for the non-Gaussian probability density function, $g(z^*)$, can be expressed as [26]

$$g(z^*) = \frac{\sqrt{3}}{2\pi}\left(\begin{aligned} & e^{-C_1 z^{*2}}\left[\frac{3(2\alpha'-3)}{\alpha'^2}\right]^{1/2} z^* + \frac{3\sqrt{2\pi}}{2\alpha'}e^{-0.5z^{*2}}(1+\operatorname{erf}\beta)(z^{*2}-1) \\ & + \sqrt{2\pi}\left[\frac{\alpha'}{3(\alpha'-1)}\right]^{1/2}\exp\left\{-\left[\frac{(\alpha' z^{*2})}{2(\alpha'-1)}\right]\right\}(1+\operatorname{erf}\gamma)\end{aligned}\right) \tag{20}$$

where erf(x) is the error function and α' denotes the bandwidth parameter and it is given as

$$\alpha' = \frac{m_0 m_4}{m_2^2} \tag{21}$$

where m_0, m_2, m_4 are the zeroth, second, and fourth moments of the power spectral density, respectively. The expressions for m_2 and m_4 have been shown in Eqs. (17) and (18). C_1, β, and γ in Eq. (20) are written as

$$C_1 = \frac{\alpha'}{(2\alpha'-3)}; \quad \beta = \left[\frac{3}{2(2\alpha'-3)}\right]^{1/2} z^*;$$

$$\gamma = \left[\frac{3}{2(\alpha'-1)(2\alpha'-3)}\right]^{1/2} z^*$$

When $\alpha' \to \infty$, the asperity heights show a Gaussian distribution. The expression for $g(z^*)$ is noticed to be dependent upon the parameter α'. However, the zeroth moment, m_0, which results from the analysis of the power spectrum variations, is expressed as [27]

$$m_0 = \frac{G^{2(D-2)}L^{2(3-D)}}{2\sin(\pi(2D-5)/2)\Gamma(2D-5)} \tag{22}$$

Therefore, the substitutions of Eqs. (22), (17), and (18) into Eq. (21) give

$$\alpha' = \frac{(D-2)^2}{(3-D)(D-1)}\left(\frac{L}{l}\right)^{2(3-D)} \tag{23}$$

where L denotes the sample length ($L=10^{-5}$ m in the present study). According to Eq. (23), the bandwidth parameter, α', can be expressed as a function of the fractal dimension D. The fractal dimension D is varied with the mean separation of two surfaces will be developed in Sec. 2.6. Therefore, its value at different mean separation can thus be obtained. The distribution form of probability density function, $g(z^*)$, varying with the mean separation of two surfaces can be predicted in the theoretical method.

2.6 Relationship Between Fractal Dimension (D) and Mean Separation (d^*). Define the profile dimension of a fractal surface as D_s and the surface dimension as D. In his study of the geomorphology of the Earth, Mandelbrot [28] found that the cumulative size distribution of islands on Earth's surface follows the power law, $N(a) \approx a^{-D_s/2}$, where N is the total number of islands with area larger than a, and D_s is related to the surface dimension D by $D=(D_s+1)$. For a fractal surface, the number of contact spots ($N(a^*)$) with their contact area larger than a^* satisfies the following power law relation [25]

$$N(a^*) \approx B(a^*)^{(1-D)/2} \tag{24}$$

If $N(a^*)-a^*$ plot in the log-log form can be obtained from the experimental results of $N(a^*)$ and a^*, the slope of a $N(a^*)-a^*$ curve is thus equal to $(1-D)/2$. Then, the fractal dimension D of this curve can be determined. In the present study, the relationship between $N(a^*)$ and a^* is determined by theoretical analyses. If the size distribution parameter, $n(a^*)$, of the asperities at each of the elastic, elastoplastic, and fully plastic deformation regimes is available, the real contact area of a surface in the dimensionless form ($A_r^* \equiv A_r/\sigma_0^2$) can be expressed as

$$\begin{aligned} A_r^* = A_e^* + A_{ep1}^* + A_{ep2}^* + A_p^* = & \int_{a_c^*}^{a_L^*} n_e(a^*)a_e^* da^* \\ & + \int_{\left(\frac{1}{6}\right)^{1/(D-2)}a_c^*}^{a_c^*} n_{ep1}(a^*)a_{ep1}^* da^* + \int_{\left(\frac{1}{100}\right)^{1/(D-2)}a_c^*}^{\left(\frac{1}{6}\right)^{1/(D-2)}a_c^*} n_{ep2}(a^*)a_{ep2}^* da^* \\ & + \int_0^{\left(\frac{1}{110}\right)^{1/(D-2)}a_c^*} n_p(a^*)a_p^* da^* \end{aligned} \tag{25}$$

where a_L^* represents the dimensionless largest contact area. The size distribution functions in the three deformation regimes, $n_e(a^*)$, $n_{ep1}(a^*)$, $n_{ep2}(a^*)$, and $n_p(a^*)$, can thus be determined if the real contact area A_r^* in Eq. (25) can be obtained by another way. If the probability density function of asperity heights, $g(z^*)$, is known, the real contact area, which was developed by Kogut and Etsion [13], can be modified by considering the area density of surface asperities to be a variable here. Then, the asymptotic expression of the real contact area is expressed as

$$A_r^* = A_e^* + A_{ep1}^* + A_{ep2}^* + A_p^* = A_n^*\left\{\int_{d^*}^{d^*+\delta_c^*} \eta a_e^* g(z^*)dz^* + \int_{d^*+\delta_c^*}^{d^*+6\delta_c^*} \eta a_{ep1}^* g(z^*)dz^* + \int_{d^*+6\delta_c^*}^{d^*+110\delta_c^*} \eta a_{ep2}^* g(z^*)dz^* + \int_{d^*+110\delta_c^*}^{\infty} \eta a_p^* g(z^*)dz^*\right\} \quad (26)$$

where $\delta_c^* = \delta_c/\sigma_0$ and $z^* = z/\sigma_0$. The equality between Eqs. (25) and (26) can be established first, then the substitution of the $g(d^*)$ into the resulting equivalence allows us to obtain the largest contact area (a_L^*) of an asperity among all real contact spot areas. The first right term of Eq. (26) is applied to determine the size distribution function, $n_e(a^*)$, for the elastic deformation regime. Applying the theorem of calculus to Eqs. (25) and (26), one can obtain the size distribution function for the elastic, elastoplastic, and fully plastic regimes, respectively. The detailed procedures of derivations have been shown in the study of Chung and Lin [22], the expressions of $n_e(a^*)$, $n_{ep1}(a^*)$, and $n_{ep2}(a^*)$, $n_p(a^*)$ can thus be given as follows:

$$n_e(a^*) = \eta A_n^* g(d^*)\left(\frac{3-D}{2}\right)2^{4-D}G^{*(D-2)}(\ln\alpha)^{1/2}\pi^{(D-3)/2}a^{*(1-D)/2} \quad (27)$$

$$n_{ep1}(a^*) = \eta A_n^* g(d^*+\delta_c^*)\left\{(1.1236-0.3386D) \times\left(\frac{KH}{E}\right)^{0.239}G^{*(0.76D-1.522)}2^{(3.044-0.76D)} \times(\ln\alpha)^{0.38}\pi^{(0.38D-1.021)}a^{*(0.261-0.38D)}\right\} \quad (28a)$$

$$n_{ep2}(a^*) = \eta A_n^* g(d^*+6\delta_c^*)\left\{(1.1-0.33D) \times\left(\frac{KH}{E}\right)^{0.255}G^{*(0.746D-1.49)}2^{(2.981-0.746D)} \times(\ln\alpha)^{0.373}\pi^{(0.373D-0.99)}a^{*(0.245-0.373D)}\right\} \quad (28b)$$

$$n_p(a^*) = \eta A_n^* g(d^*+110\delta_c^*)\left(\frac{3-D}{4}\right)G^{*(D-2)}2^{(4-D)} \times(\ln\alpha)^{1/2}\pi^{(D-3)/2}a^{*(1-D)/2} \quad (29)$$

By applying Eq. (27) to Eq. (29), the relationships of $N(a^*)$ developed for the three deformation regimes can be found theoretically. Since the slope of an $N(a^*)-a^*$ curve is equal to $(1-D)/2$, the fractal dimension D at each deformation regime can thus be determined.

2.7 Relationship Between Topothesy (G^*), Fractal Dimension (D), and Mean Separation (d^*). According to the results obtained above, the relationship between fractal dimension (D) and mean separation (d^*) has been established. The relationship among topothesy (G^*), fractal dimension (D), and mean separation (d^*) can also be derived. According to Eq. (26), the dimensionless real contact areas, A_e^*, A_{ep1}^*, A_{ep2}^*, and A_p^* corresponding to the contact behavior operating in the elastic, elastoplastic and fully plastic deformation regimes, respectively. The normalization of Eq. (2) gives $a_e^* = \pi R_e^* \delta^* = \pi R_e^*(z^*-d^*)$. Therefore, the dimensionless contact area in the elastic deformation regime, A_e^*, can be expressed as

$$A_e^* = \pi A_n^* \int_{d^*}^{d^*+\delta_c^*} \eta R_e^*(z^*-d^*)g(z^*)dz^* \quad (30)$$

Substituting R_e^* shown in Eq. (15) into Eq. (31) gives

$$A_e^* = \pi A_n^* \int_{d^*}^{d^*+\delta_c^*} \eta\left[\frac{2^{2(D-4)}G^{*2(2-D)}}{\ln\alpha}\right]^{1/(3-D)}(z^*-d^*)^{2/(3-D)}g(z^*)dz^* \quad (31)$$

The equality of Eqs. (30) and (31) obtains

$$G_e^* = \left[\frac{\int_{d^*}^{d^*+\delta_c^*} \eta\pi R_e^*(z^*-d^*)g(z^*)dz^*}{\int_{d^*}^{d^*+\delta_c^*} \eta\pi 2^{2(D-4)/(3-D)}(\ln\alpha)^{1/(D-3)}(z^*-d^*)^{2/(3-D)}g(z^*)dz^*}\right]^{(3-D)/2(2-D)} \quad (32)$$

Following the method mentioned above, the dimensionless topothesies for the first elastoplastic regime (G_{ep1}^*, $1<\delta/\delta_c\leqslant 6$), the second elastoplastic regime (G_{ep2}^*, $6<\delta/\delta_c\leqslant 110$) and the fully plastic regime (G_p^*, $\delta/\delta_c>110$) are also can be found.

By incorporating Eqs. (27), (24), and (32), variable G^* and D in the elastic regime as a function of d^* can thus be found in the following way. In Eq. (32), the numerator value is determined first by substituting (D_0) and (G_0^*) values in the flow chart of numerical iterations (see Fig. 2). However, the D value given in the denominator is now determined by using the $D-d^*$ relationship developed in Sec. 2.6. Then the D values corresponding to different d^* values can be substituted into Eq. (32) to obtain the G^* values. The values of these (D, G^*) sets corresponding to different d^* values are then substituted into Eq. (27) to calculate $n_e(a^*)$. The $N(a^*)$ values can thus be obtained from the formula $N(a^*) = \int_{a^*}^{a_L^*} n_e(a^*)da^*$. These $N(a^*)$ values are then applied to determine the new D values by Eq. (24). These new D values and the previously obtained G^* values are then replaced with the (D_0) and (G_0^*) values, respectively, in order to start the next numerical iteration. This procedure is repeated in the successive iterations until the D and G^* are convergent. Then, these convergent values of the (D, G^*) sets show the variations of these two roughness parameters in the elastic regime with d^*. The methods similar to

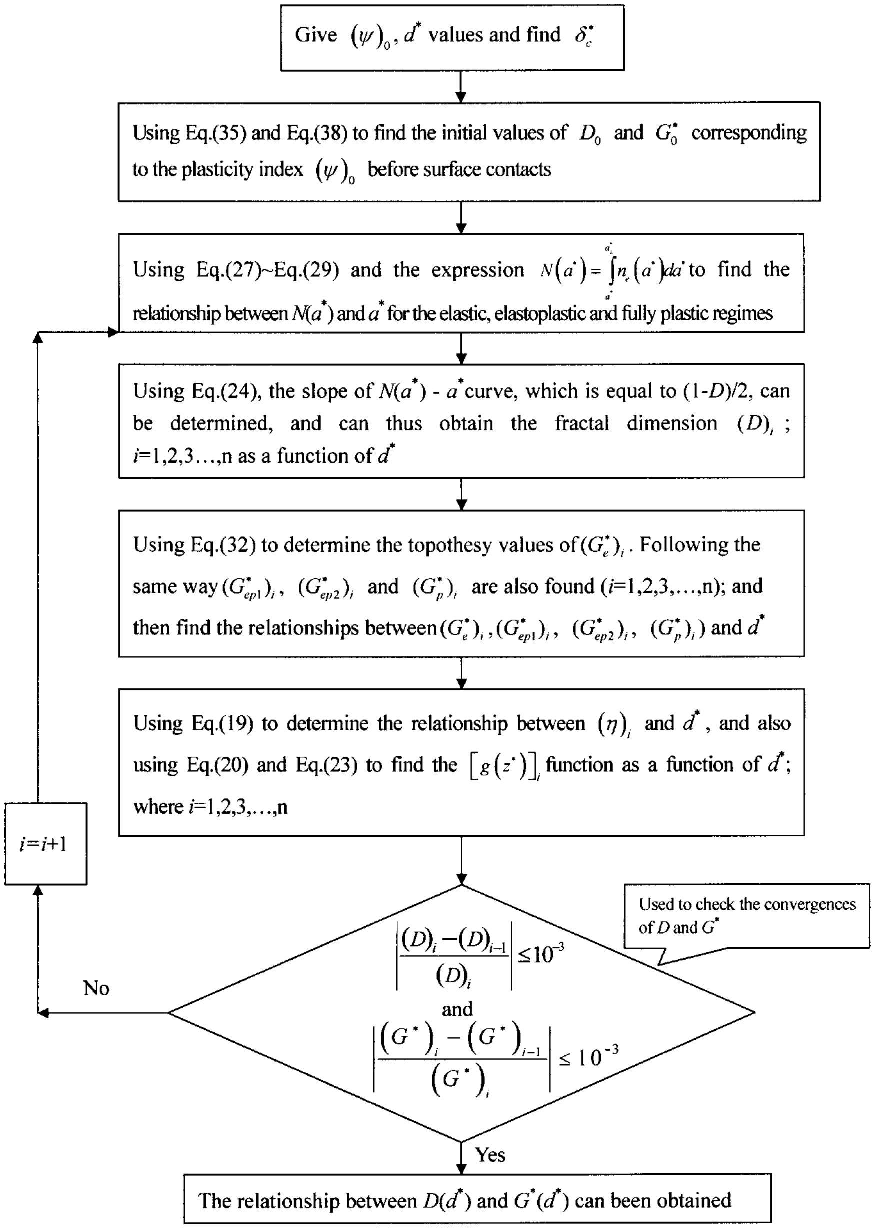


Fig. 2 Flow chart for the numerical analyses of $D(d^*)$ and $G^*(d^*)$

the procedure mentioned above can be applied to determine the (D,G^*) sets for the elastoplastic and fully plastic regimes. The flow chart of finding the (D,G^*) values arising at different d^* values is shown in Fig. 2.

2.8 Dimensionless Topothesy (G_0^*), Fractal Dimension (D_0) Corresponding to a Plasticity Index ($(\psi)_0$) Before Asperity Contact Deformations. In the present study, the fractal theory is applied to modify the conventional GW model by incorporating the fractal dimension (D) and topothesy (G) with the mean radius of curvature (R) and the density of asperities (η). Thus, it is necessary to find the relationship among the topothesy, the fractal dimension and a given value of the plasticity index before contact deformations occur. G_0^*, D_0, and $(\psi)_0$ represent the initial values of the dimensionless topothesy, the fractal dimension and the plasticity index before contact deformations, respectively. The plasticity index introduced by Greenwood and Williamson [5] can be expressed as

$$(\psi)_0=\left(\frac{\delta_c}{\sigma_{s0}}\right)^{-1/2}=\left(\frac{\delta_c}{\sigma_0}\right)^{-1/2}\left(\frac{\sigma_0}{\sigma_{s0}}\right)^{-1/2} \tag{33}$$

where the δ_c expression is shown in Eq. (8). The δ_c/σ_0 parameter shown in Eq. (33) is thus expressed as

$$\left(\frac{\delta_c}{\sigma_0}\right)^{-1/2}=\frac{2E}{\pi KH}[R_0^*(a_c^*)]^{-1/2} \tag{34}$$

where $R_0^*(a_c^*)=R_0(a_c)/\sigma_0$, R_0: the radius of curvature before contact deformations) is obtained if the a^* parameter shown in Eq. (13b) is replaced by a_c^*. Equation (34) can thus be rewritten as

$$\left(\frac{\delta_c}{\sigma_0}\right)^{-1/2}=2^{(2D_0-7)/2(D_0-2)}(\ln\alpha)^{1/4(2-D_0)}\left(\frac{E}{\pi KH}\right)^{(D_0-3)/2(D_0-2)}G_0^{*-1/2} \tag{35}$$

In the study of McCool [29], the ratio of σ_{s0} and σ_0 in Eq. (33) can be expressed as $(\sigma_{s0}/\sigma_0)^2=1-3.717\times10^{-4}/(\eta_0\sigma_0R_0)^2$, where

Table 2 Contrasts for plasticity index, surface topography, topothesy, and fractal dimension

$(\psi)_0$	$\eta_0\sigma_0R_0$	σ_0/R_0	σ_0 (m)	D_0	G_0^*
0.75	5.292	6.586×10^{-3}	2.94×10^{-6}	2.19	1.86×10^{-6}
2.0	46.898	4.834×10^{-2}	21.58×10^{-6}	2.3	1.15×10^{-4}

the $(\eta_0\sigma_0R_0)$ values corresponding to the plasticity indices of $(\psi)_0=0.75$ and 2 were obtained from the experimental study of Bhushan and Dugger [30]. By Eq. (33), the $(\delta_c/\sigma_0)^{-1/2}$ values corresponding to $(\psi)_0=0.75$ and 2 are 0.7516 and 2.041, respectively.

According to Eq. (35), a given value of $(\delta_c/\sigma_0)^{-1/2}$ may not be enough to determine the values of D_0 and G_0^* corresponding to a $(\psi)_0$ value. Therefore, it is necessary to find another relationship which incorporates with Eq. (35) to determine the exact values of D_0 and G_0^*. The scale dependence of the standard deviation of surface heights, σ_0, which comes from the fractal power-law analysis for the power spectrum, can be written as [27]

$$\sigma_0=\frac{G_0^{(D_0-2)}L^{(3-D_0)}}{\sqrt{2}\left[\sin\left(\pi(2D_0-5)/2\right)\Gamma(2D_0-5)\right]^{1/2}} \tag{36}$$

where L denotes the sample length. By Eq. (36) G_0^* can thus be written as

$$G_0^*\equiv\frac{G_0}{\sigma_0}=\sqrt{2}G_0^{(3-D_0)}L^{(D_0-3)}\left[\sin\frac{\pi(2D_0-5)}{2}\Gamma(2D_0-5)\right]^{1/2} \tag{37}$$

the rearrangement of Eq. (37) gives G_0 as

$$G_0=L\left(\frac{1}{\sqrt{2}}G_0^*\right)^{1/(3-D_0)}\left[\sin\frac{\pi(2D_0-5)}{2}\Gamma(2D_0-5)\right]^{1/2(D_0-3)} \tag{38}$$

Equations (35) and (38) are combined to solve the D_0 and G_0^* values of a rough surface before asperity contacts and deformations occur. Because the fractal dimension of the surface is ranged between 2 and 3 $(2<D_0<3)$. Thus, numerous sets of D_0 and G_0^* values can be obtained and satisfied by Eq. (35). By Eq. (38), numerous G_0 values can thus be obtained by substituting these sets of D_0 and G_0^* values corresponding to this $(\psi)_0$ value. Thus, numerous σ_0 values can be determined if the G_0 and G_0^* values are available. The σ_0 value corresponding to a $(\psi)_0$ value can be found in the experiments [30] and is now shown in Table 2. This σ_0 value is then applied to determine the genuine D_0 and G_0^* values.

3 Results and Discussion

In the present study, all the contact parameters related to the topography including the fractal dimension (D), the topothesy (G), the area density (η), and the probability density function (g) of surface asperities are assumed to be variables as a function of the mean separation distance of two contact surfaces. In order to investigate the fractal dimensions varying with the dimensionless separation, the $N(a)$-a curves should be obtained first by designating the material properties (E,H) and the plasticity index before contact deformations $(\psi)_0$. Figure 3(a) shows the $N(a)$-a curves evaluated at different separations and a plasticity index of 0.75; whereas Fig. 3(b) shows the curves evaluated at a plasticity index of 2.0. In Fig. 3(a), three of these five curves are marked in somewhere by $(\delta/\delta_c=1,a=a_c)$. This sign is given to indicate the border of the elastic and elastoplastic regimes. In the region that $a<a_c$, elastic deformation is shown; whereas in the region of $a>a_c$, elastoplastic deformation behavior is exhibited. No mark of $(\delta/\delta_c=1,a=a_c)$ is given in the curves with $d^*=3$ and $d^*=4$. This implies that the N/A_n values corresponding to $d^*=3$ and $d^*=4$ are obtained from the asperities operating in the elastic regime only. The zigzag shown at the point marked by $(\delta/\delta_c=1,a=a_c)$ is caused due to the application of the power-form expressions developed in the study of Kogut and Etsion [13] for the contact parameters including the size distribution functions. In the Kosut and Etsion model, there exists a problem that discontinuities in the asperity's contact parameters are found to be present at the inception and the end of the elastoplastic deformation regime. These $N(a)$-a curves show the behavior that the absolute values of these negative slopes are increased by increasing the dimensionless separation; i.e., the fractal dimension is increased by increasing the separation between two contact surfaces.

Figure 3(b) shows the (N/A_n)-a curves evaluated by the elevation of $(\psi)_0$ to 2.0. All these five curves show the mark of $(\delta/\delta_c=1,a=a_c)$ to note the border of the elastic and elastoplastic regimes. Four of these five curves show the mark of $(\delta/\delta_c=6)$ to note the border beyond it, where the second elastoplastic regime of an asperity is stretched. The results shown in Figs. 3(a) and 3(b) reveal that these contacts at different separations are mostly operating in the elastic and elastoplastic regime, irrespective of the $(\psi)_0$ value. The elastic regime is prevailing in a wide range of

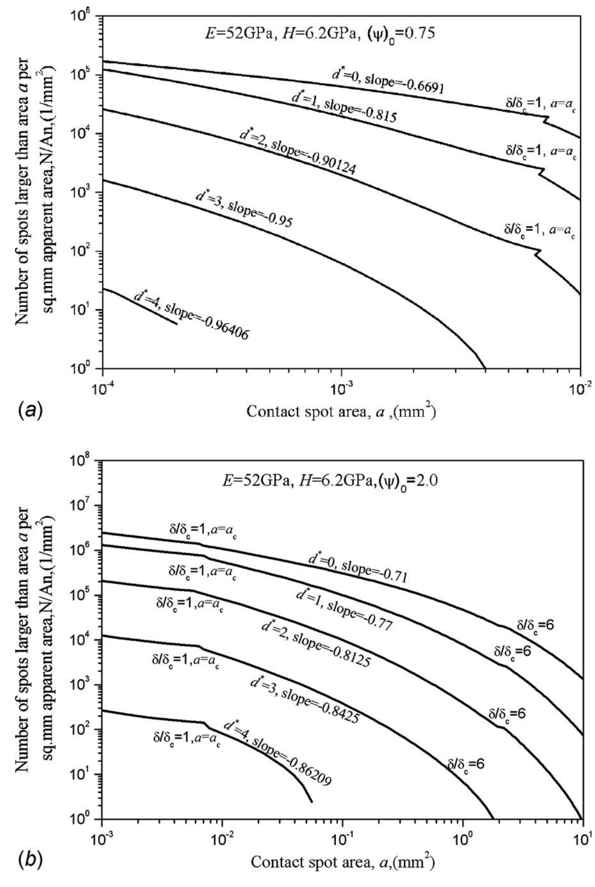


Fig. 3 The theoretical results of N/A_n expressed as a function of the contact spot area a: (a) $(\psi)_0=0.75$ and (b) $(\psi)_0=2.0$

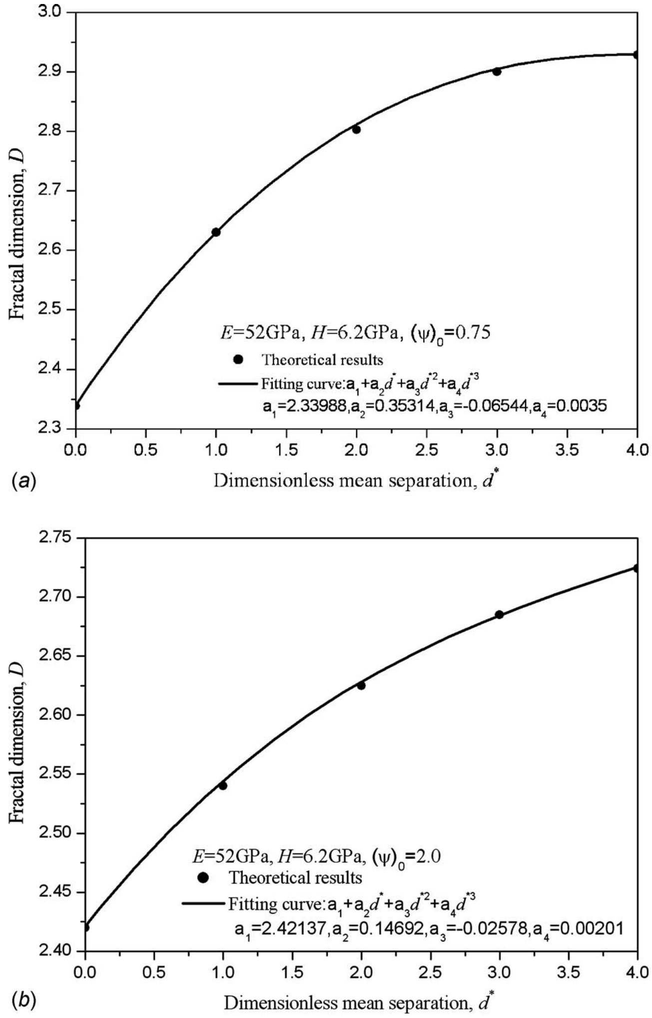


Fig. 4 The fractal dimensions varying with the dimensionless mean separation. These data of *D* are obtained from the slope values of those five curves shown in Fig. 3. (*a*) $(\psi)_0=0.75$ and (*b*) $(\psi)_0=2.0$.

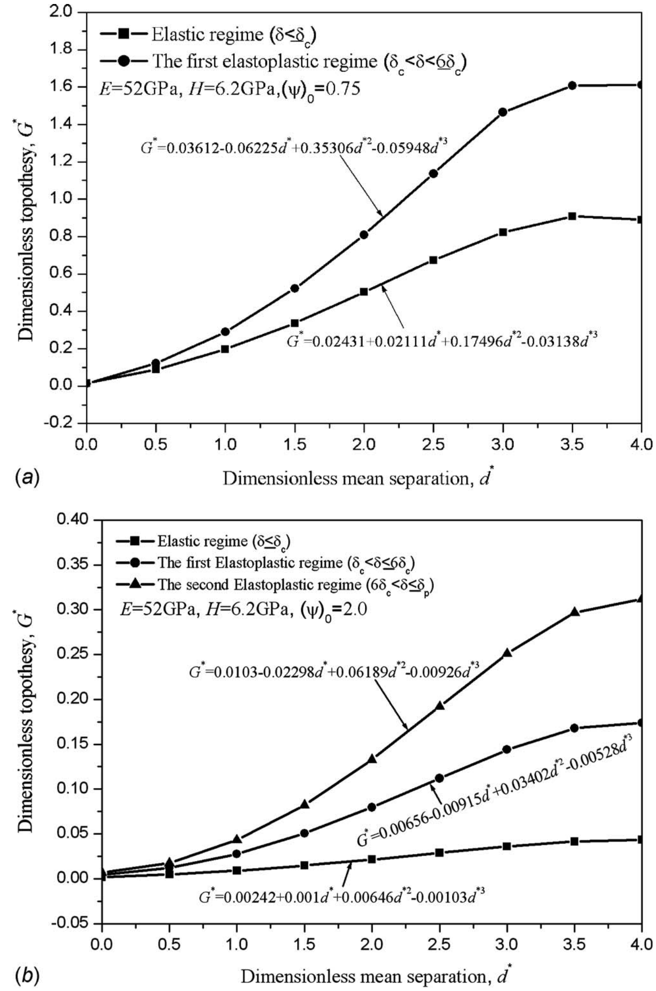


Fig. 5 The dimensionless topothesy varying with the dimensionless mean separation. The values of G^* are obtained for the elastic, elastoplastic and fully plastic regimes. (*a*) $(\psi)_0=0.75$ and (*b*) $(\psi)_0=2.0$.

contact spot areas (a) if a small value of plasticity index is assumed. This elastic regime is then narrowed significantly as the plasticity index is elevated to 2.0.

The results of the fractal dimension corresponding to the $(N(a)/A_n)$-a curves shown in Figs. 3(a) and 3(b) are shown in Figs. 4(a) and 4(b), respectively. The curve in either Fig. 4(a) or Fig. 4(b) shows the fractal dimension to be increased by increasing the dimensionless mean separation (d^*). The polynomials used in fitting these results are shown in the respective figures. In the same range of mean separations, the rate of increase in the fractal dimension due to the rise in the mean separation is elevated by reducing the plasticity index if the material properties are fixed.

The variations of the dimensionless topothesy with the dimensionless mean separation, according to the G^* expressions developed in Sec. 2.7, are obtained independently for these different deformation regimes. They are shown in Fig. 5(a) for $(\psi)_0=0.75$ and in Fig. 5(b) for $(\psi)_0=2.0$. In both figures, the topothesies evaluated at different deformation regimes are asymptotic to almost the same value as the mean separation is lowered to zero. However, the differences between (or among) them are significantly enlarged by increasing the mean separation. The magnitudes of G^* shown in different deformation regimes always satisfy the sequence that $(G^*)_{\text{elastic}} < (G^*)_{\substack{\text{first}\\\text{elastoplastic}}} < (G^*)_{\substack{\text{second}\\\text{elastoplastic}}}$.

The area density of asperities (η), which is widely regarded as a constant value, is actually varied with the mean separation of two contact surfaces. In Eq. (19), the area density of asperities η can be expressed as a function of the fractal dimension D. According to the results shown in Fig. 4, the fractal dimension is lowered

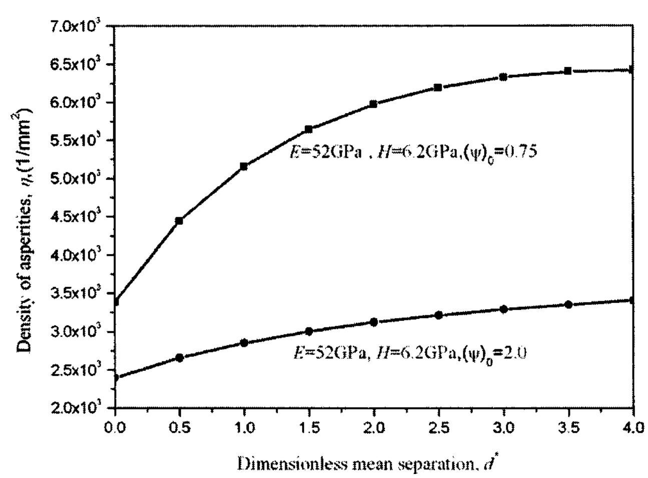
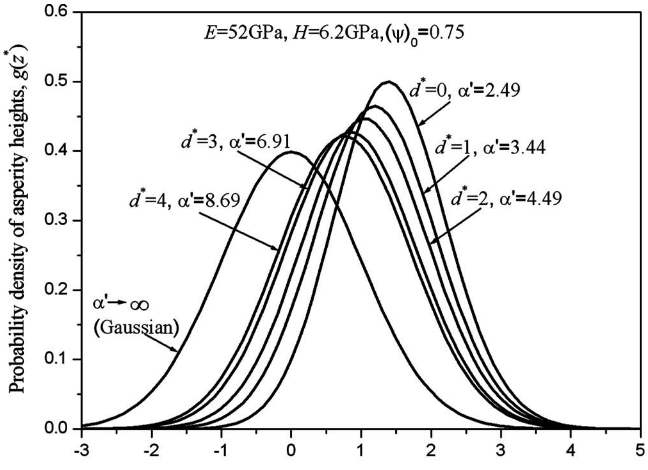


Fig. 6 Density of asperities varying with the dimensionless mean separation



(*a*)

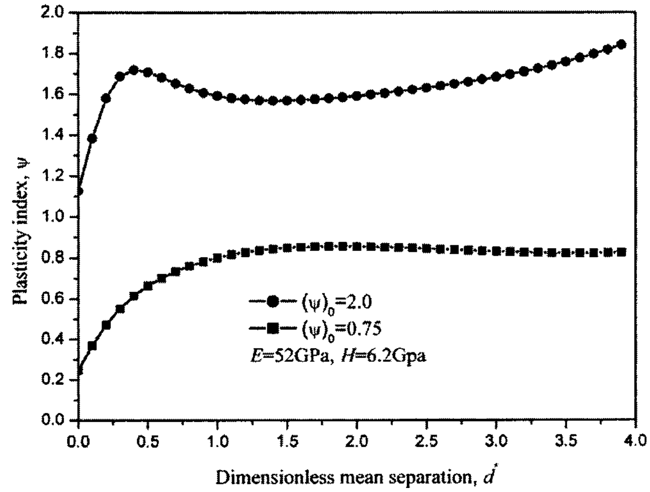
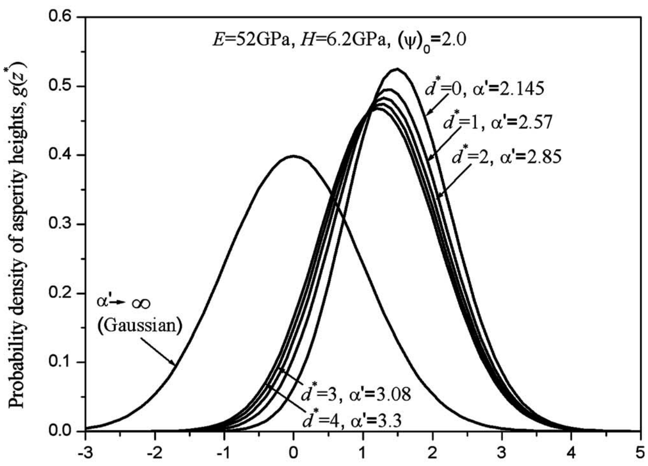


Fig. 8 Plasticity index of rough surfaces varying with the dimensionless mean separation



(*b*)

Fig. 7 Probability density functions of asperity heights varying with the dimensionless asperity height: (*a*) $(\psi)_0$=0.75 and (*b*) $(\psi)_0$=2.0

by decreasing the mean separation of two contact surfaces. The area density of asperities varying with the dimensionless mean separation is shown in Fig. 6. This figure shows that the area density of asperities is always decreased by lowering the mean separation of two contact surfaces, regardless of the original plasticity index $(\psi)_0$ before any surface contact. In general, the contact spots cannot operate independent of each other because they are connected by the solid bodies that can sustain some elastic or plastic deformation. When the loads are low, only the small asperities are deformed. As the load is further increased, they can merge to form a larger spot, thus resulting in the lowering of the area density of asperities (η).

The distribution form of probability density function $g(z^*)$ is also taken as a variable when evaluated at different mean separations (d^*). The variations of $g(z^*)$ for $(\psi)_0=0.75$ are shown in Fig. 7(*a*); whereas the variations of $g(z^*)$ for $(\psi)_0=2.0$ are shown in Fig. 7(*b*). According to the study of Nayak [26], the probability density function was developed for the asperity heights as a function of the bandwidth parameter α'. Since this bandwidth parameter is considered to be a variable as a function of the mean separation (d^*), the value of α' corresponding to a mean separation is also marked behind the mean separation. As the two surfaces are separated by a large distance, $\alpha' \to \infty$ and the Gaussian distribution is assumed for the asperity heights. If the separation is reduced, the α' value is quickly lowered too. Then, the probability density function becomes non-Gaussian and the profile is no longer symmetric with respect to the axis of $z^*=0$. The peak value is shifted rightwards ($z^*>0$) and is elevated by reducing the separation between the two contact surfaces. Detailed investigation of these two figures finds the characteristic that the shift distance between $z^*=0$ and the z^* value corresponding to the peak value of a profile is related to the bandwidth parameter α'.

According to the definition of the plasticity index from Eq. (34), it is expressed as a function of the critical interference (δ_c) and the standard deviation of asperity height (σ_{s0}). According to Eq. (36), the critical interference can be expressed as a function of the following parameters: fractal dimension D, dimensionless topothesy, G^*, and the material properties (hardness H, Young's modulus E). In the present study, the fractal dimension and dimensionless topothesy, is varied with the different mean separation. Therefore, the plasticity index is expected to vary with the mean separation, rather than remain a constant value. Figure 8 shows the variations of ψ with d^* for the cases of $(\psi)_0=0.75$ and $(\psi)_0=2.0$. These two curves are shown for d^* in the range of 0 to 4.0 only. As the d^* value is further increased to be sufficiently large, these two curves will be asymptotic to $(\psi)_0=0.75$ and $(\psi)_0=2.0$, respectively.

The results of dimensionless contact load F_t^* obtained by the assumptions of constant D, G^*, and Gaussian $g(z^*)$ are compared with variable D, G^*, and non-Gaussian $g(z^*)$, and this is shown in Fig. 9. These results are evaluated at a value of $(\psi)_0=2.0$. The contact load results for $(\psi)_0=0.75$ show behavior similar to that exhibited for $(\psi)_0=2.0$, but they are not presented here. The contact load results predicted by the condition of variable D and G^* and non-Gaussian g are always higher than those evaluated by the assumption of constant D and G^* and Gaussian g if they are obtained at the same value of d^*. The difference between these two kinds of contact loads is significantly enlarged by decreasing the mean separation between the two contact surfaces. Under the same conditions given in Fig. 9, the contact area predicted by the condition of variable D and G^* and non-Gaussian is also always higher than that predicted by the assumption of constant D and G^* and Gaussian $g(z^*)$ (see Fig. 10). Similarly, the difference between these two kinds of contact areas is enlarged by reducing the mean separation (d^*).

Figure 10 shows the variations of the dimensionless real contact area with the dimensionless total load applied to the contact surfaces for $(\psi)_0=0.75$ and $(\psi)_0=2.0$. In the log-log plot, the real contact area and the total contact load are presented as having a

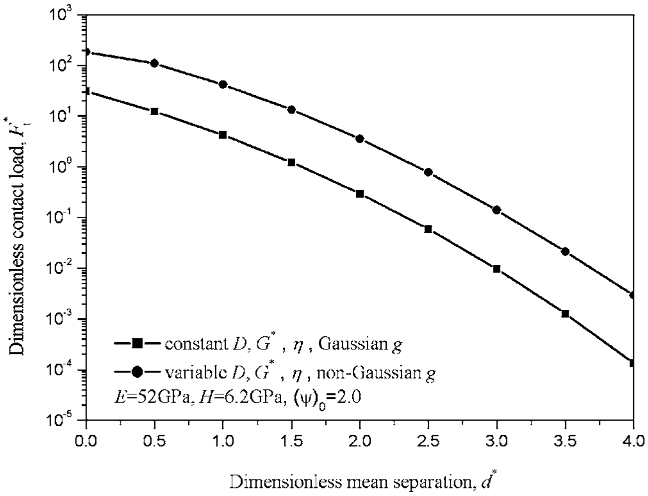


Fig. 9 Variations of the dimensionless contact loads with the dimensionless mean separation. They are presented to compare the evaluations based on variable D, G^*, η, and non-Gaussian g with the evaluations based on constant D, G^*, η, and Gaussian g.

linear relationship. The slope of a straight line is dependent upon the value of $(\psi)_0$ and the conditions set for D, G^*, η, and $g(z^*)$. In each of these two figures, the slope of the straight line obtained by the assumption of constant D, G^*, and η as well as Gaussian $g(z^*)$ is larger than that exhibited in the line obtained by variable D, G^*, and η as well as non-Gaussian $g(z^*)$, regardless of the $(\psi)_0$ value.

4 Conclusions

(1) Instead of a general consideration of the microcontacts based on the GW model, the mean radius of curvature and the density of asperities are varied with the mean separation between two contact surfaces (thus the interference), these two parameters can be expressed as a function of variable fractal dimension and topothesy. The probability density function of asperity heights is also found to be a function of variable fractal dimension. Moreover, the fractal dimension and the topothesy are related to the mean separation. Thus, these parameters are found to vary with the interference.

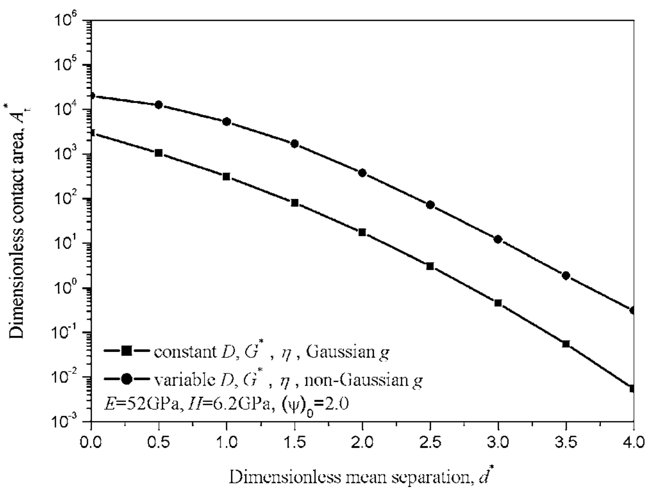


Fig. 10 Variations of the dimensionless total contact area with the dimensionless mean separation. They are presented to compare the evaluations based on variable D, G^*, η, and non-Gaussian g with the evaluations based on constant D, G^*, η, and Gaussian g.

(2) The fractal dimension is always reduced by decreasing the mean separation of the two contact surfaces, regardless of the initial plasticity index before surface contacts. The topothesies arising at different deformation regimes also show different behavior in the variations. Nevertheless, the topothesy is also reduced by decreasing the mean separation, regardless of the operating deformation regime of an asperity and the initial plasticity index.

(3) The density of surface asperities is lowered by reducing the mean separation between the two contact surfaces, regardless of the initial plasticity index before contact deformations. As two rough surfaces experience contact deformations, the topographies of each surface will be changed. The plasticity index is no longer a constant value, but varies with the mean separation.

(4) The contact load and the total contact area predicted by the assumption of constant D, G^*, and η as well as Gaussian g is lower than the load predicted by variable D, G^*, and η as well as non-Gaussian g.

Nomenclature

a = area of a contact spot
A_r = real contact area
A_n = apparent area
d = separation based on asperity heights
D = 3D fractal dimension $(2<D<3)$
D_s = 2D fractal dimension $(1<D_s<2)$
E = effective Young's modulus
F = contact load
$g(z^*)$ = probability density function of summit heights
G = topothesy
h = separation based on surface heights
H = hardness of the softer material in contact
K = maximum contact pressure factor
N = total number of contact spots
P = mean contact pressure
R = the equivalent radius
Y = yield stress of the material in simple tension or compression
y_s = separation between the mean of asperity heights and that of surface heights
z = height of asperity measured from the mean of asperity heights
α = 1.5, a parameter that determined the density of frequencies in the surface
α' = the bandwidth parameter
δ = interference of asperity
σ = standard deviation of surface heights
σ_s = standard deviation of asperity heights
η = the area density of asperities
ν = Poisson's ratio
ψ = plasticity index

Subscripts or Superscripts

0 = initial value of the surface before contact
c = critical value
e = elastic deformation
ep = elastoplastic deformation
p = plastic deformation
t = total summation
$*$ = dimensionless

References

[1] Polycarpou, A. A., and Etsion, I., 2000, "A Model for the Static Sealing Performance of Compliant Metallic Gas Seal Including Surface Roughness and Rarefraction Effects," STLE Tribol. Trans., **43**(2), pp. 237–244.

[2] Kogut, L., and Etsion, I., 2000, “Electrical Conductivity and Friction Force Estimation in Compliant Electrical Connectors,” Tribol. Trans., **43**, pp. 816–822.
[3] Jeng, Y. R., and Horng, J. H., 2001, “A Microcontact Approach for Ultrasonic Wire Bonding in Microelectronics,” ASME J. Tribol., **123**, pp. 725–731.
[4] Halling, J., and Nuri, K. A., 1988, “The Elastic-Plastic Contact of Rough Surfaces and its Relevance in the Study of Wear,” Proc. Inst. Mech. Eng., Part C: Mech. Eng. Sci., **202**(c4), pp. 269–274.
[5] Greenwood, J. A., and Williamson, J. B. P., 1966, “Contact of Nominally Flat Surfaces,” Proc. R. Soc. London, Ser. A, **295**, pp. 300–319.
[6] Abbott, E. J., and Firestone, F. A., 1933, “Specifying Surfaces Quality-A Method Based on Accurate Measurement and Comparison,” Mech. Eng. (Am. Soc. Mech. Eng.), **55**, pp. 569–572.
[7] Chang, W. R., Etsion, I., and Bogy, D. B., 1987, “An Elastic-Plastic Model for the Contact of Rough Surfaces,” ASME J. Tribol., **110**, pp. 50–56.
[8] Zhao, Y., Maietta, D. M., and Chang, L., 2000, “An Asperity Microcontact Model Incorporating the Transition From Elastic Deformation to Fully Plastic Flow,” ASME J. Tribol., **122**(1), pp. 86–93.
[9] Kral, E. R., Komvopolous, K., and Bogy, D. B., 1933, “Elastic-Plastic Finite Element Analysis of Repeated Indentation of a Half-Space by a Rigid Sphere,” ASME J. Appl. Mech., **60**, pp. 829–841.
[10] Kucharski, S., Klimczak, T., Polijaniuk, A., and Kaczmarek, J., 1994, “Finite-Elements Model for the Contact of Rough Surfaces,” Wear, **177**, pp. 1–13.
[11] Lin, G., Zhu, J., and Wang, Q. J., 2001, “Elasto-Plastic Contact of Rough Surfaces,” Tribol. Trans., **44**, pp. 437–443.
[12] Vu-Quoc, L., Zhang, X., and Lesburg, L., 2000, “A Normal Force-Displacement Model for Contacting Spheres Accounting for Plastic Deformation: Force-Drive Formulation,” ASME J. Appl. Mech., **67**, pp. 363–371.
[13] Kogout, L., and Etsion, I., 2002, “Elastic-Plastic Contact Analysis of a Sphere and a Rigid Flat,” ASME J. Appl. Mech., **69**, pp. 657–662.
[14] Majumdar, A., and Bhushan, B., 1991, “Fractal Model of Elastic-Plastic Contact Between Rough Surfaces,” ASME J. Tribol., **113**, pp. 1–11.
[15] Bhushan, B., and Majumdar, A., 1992, “Elastic-Plastic Contact for Bifractal Surfaces,” Wear, **153**, pp. 53–64.
[16] Blackmore, D., and Zhou, G., 1998, “A New Fractal Model for Anisotropic Surfaces,” Int. J. Mach. Tools Manuf., **38**, pp. 551–557.
[17] Yan, W., and Komvopoulos, K., 1998, “Contact Analysis of Elastic-Plastic Fractal Surfaces,” J. Appl. Phys., **84**(7), pp. 3617–3624.
[18] Persson, B. N. J., 2001, “Theory or Rubber Friction and Contact Mechanics,” J. Chem. Phys., **115**(8), pp. 3840–3861.
[19] Persson, B. N. J., 2002, “Adhesion Between an Elastic Body and a Randomly Rough Hard Surface,” Eur. Phys. J. E, **8**(4), pp. 385–401.
[20] Zhang, L. X., and Zhao, Y. P., 2004, “Adhesion of Rough Surfaces With Plastic Deformation,” J. Adhes. Sci. Technol., **18**(6), pp. 715–729.
[21] Othmani, A., and Kaminsky, C., 1998, “Three Dimensional Fractal Analysis of Sheet Metal Surfaces,” Wear, **214**, pp. 147–150.
[22] Chung, J. C., and Lin, J. F., 2006, “Variation in Fractal Properties and Non-Gaussian Distributions of Microcontact Between Elastic-Plastic Rough Surfaces with Mean Surface Separation,” ASME J. Appl. Mech., **73**, pp. 143–152.
[23] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge.
[24] Chang, W. R., Etsion, I., and Bogy, D. B., 1988, “Static Friction Coefficient Model for Metallic Rough Surfaces,” ASME J. Tribol., **110**, pp. 57–63.
[25] Chung, J. C., and Lin, J. F., 2004, “Fractal Model Developed for Elliptic Elastic-Plastic Asperity Microcontacts of Rough Surfaces,” ASME J. Tribol., **126**, pp. 646–654.
[26] Nayak, P. R., 1971, “Random Process Model of Rough Surfaces, J. Lubr. Technol., **93**, pp. 398–407.
[27] Bhushan, B., 1999, *Handbook of Micro/Nanotribology*, 2nd ed., CRC Press, Boca Raton.
[28] Mandelbrot, B. B., 1975, “Stochastic Models for the Earth’s Relief, the Sharp and the Fractal Dimension of the Coastlines, and the Number-Area Rule for Islands,” Proc. Natl. Acad. Sci. U.S.A., **72**, pp. 3825–3828.
[29] McCool, J. I., 1986, “Comparison of Model for Contact of Rough Surfaces,” Wear, **107**, pp. 37–60.
[30] Bhushan, B., and Dugger, M. T., 1990, “Real Contact Area Measurements on Magnetic Rigid Disks,” Wear, **137**, pp. 41–50.

The Mixed Mode I and II Interface Crack in Piezoelectromagneto–Elastic Anisotropic Bimaterials

R. Li
Post Doctoral Fellow

G. A. Kardomateas
Professor
e-mail: george.kardomateas@aerospace.gatech.edu

School of Aerospace Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332-0150

Taking the electric–magnetic field inside the interface crack into account, the interface crack problem of dissimilar piezoelectromagneto (PEMO)–elastic anisotropic bimaterials under in-plane deformation is investigated. The conditions to decouple the in-plane and anti-plane deformation is presented for PEMO–elastic biaterials with a symmetry plane. Using the extended Stroh's dislocation theory of two-dimensional space and the analytic continuition principle of complex analysis, the interface crack problem is turned into a nonhomogeneous Hilbert equation in matrix notation. Four possible eigenvalues as well as four eigenvectors for the fundamental solution to the corresponding homogeneous Hilbert equation are found, so are four modes of singularities for the fields around the interface crack tip. These singularities are shown to have forms of $r^{-(1/2)\pm i\epsilon_1}$ and $r^{-(1/2)\pm i\epsilon_2}$, in which the bimaterial constants ϵ_1 and ϵ_2 are proven to be real numbers for practical dissimilar PEMO–elastic bimaterials. Compared with the solution for the interface crack of dissimilar elastic bimaterials without electro–magnetic properties, two new additional singularities are discovered for the interface crack in the PEMO–elastic bimaterial media. The electric–magnetic field inside the crack is solved by employing the "energy method," which is based on finding the stationary point of the saddle surface of the energy release rate with respect to the electro–magnetic field inside the crack. Closed form expressions for the extended crack tip stress fields and crack open displacements are formulated, so are some other fracture characteristic parameters, such as the extended stress intensity factors and energy release rate (G) for dissimilar PEMO–elastic bimaterial solids. Finally, fundamental results and some conclusions are presented, which could have applications in the failure of piezoelectro/magneto–elastic devices.
[DOI: 10.1115/1.2424468]

Keywords: piezoelectromagneto-elastic solids, dissimilar, anisotropic, interface crack, bimaterials

1 Introduction

The simultaneous presence of piezo-electric and piezomagnetic material properties [1,2] usually lends a device some exceptional features such as converting energy from one form to the other form [3,4] and flat frequency responses [5]. These types of media find such applications in smart structure sensors, actuators, magnetoelectric memory apparatus, and broadband magnetic probes. In these applications, dissimilar bimaterials or layered composites are often incorporated. Having been considered as one of the common failure modes, an interface crack/delamination could be developed in structures made of piezoelectro magneto (PEMO)–elastic bimaterials, and then deteriorate the performance of the devices.

The interface crack phenomenon has been investigated for decades by many authors [6–13]. Morevover, although the piezomagnetic material properties were not included, there are many studies on piezoelectric media or smart materials such as those by McMeeking [14], Kuo and Barnett [15] and Suo et al. [16]. In their studies, the singularities around the interface crack tip were found to be of the form $r^{-(1/2)\pm i\epsilon}$ and $r^{-(1/2)\pm\kappa}$, where ϵ and κ are real numbers. In particular, the paper by Suo et al. [16] has investigated this type of interface crack in detail.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received November 16, 2005; final manuscript received June 2, 2006. Review conducted by Zhigang Suo.

As addressed in the literature (e.g., Refs. [2,4,5]), the simultaneous presence of the piezoelectric and piezomagnetic material properties usually have a big influence on the behavior of PEMO–elastic solids or layered structures. Thus, these piezoelectromagnetic material properties would also affect the interface fracture behavior of PEMO–elastic bimedia. Several papers on the study of cracks in monolithic PEMO–elastic solids are available such as Sih and Song [17], Song and Sih [18], and Gao et al. [19] etc. But few papers can be found for the problem of the interface crack in PEMO bimaterial solids. Gao et al. [20] presented a solution for a permeable interface crack and presented the singularities of the interface crack of the form as $r^{-(1/2)\pm i\epsilon_\alpha}$, but did not show whether ϵ_α are real or complex numbers. Furthermore, another important fracture parameter, the energy release rate G, has not been addressed in the literature for the in-plane interface crack of dissimilar anisotropic PEMO-elastic bimaterial solids.

In this research, the impermeable and permeable interface cracks in dissimilar PEMO bimaterial solids are investigated by employing the Stroh's dislocation theory [21], extended to PEMO–elastic media (e.g., Refs. [1,22]). The electric–magnetic field inside the crack is also considered. The Mode III interface crack solution has been analyzed in the authors' earlier work (Ref. [23]), and the current paper deals with the mixed mode I and II in-plane problems.

The paper is organized as follows: In Sec. 2, the conditions to decouple the in-plane and anti-plane deformations are derived and basic equations for the in-plane deformation are presented in the form of the extended Stroh's dislocation theory. In Sec. 3, a nonhomogenous Hilbert equation is obtained in matrix notation by using the analytic continuation principle of complex analysis. Four roots (i.e., four eigenvalues) to the corresponding homogenous Hilbert equation are found and so are four eigenvectors. Four possible singularities are then found in the form of $r^{-(1/2)\pm i\epsilon_1}$ and $r^{-(1/2)\pm i\epsilon_2}$. The bimaterial property constants ϵ_1 and ϵ_2 are proved to be real numbers for practical dissimilar bimaterial media. Compared with the solutions for conventional dissimilar bimaterials and piezoelectric bimaterials, two new types of singularities can be observed in this solution due to the simultaneous presence of piezoelectric and piezomagnetic material properties. Fracture parameters such as the extended stress intensity factor and the extended crack open displacement are presented in closed form for uniform applied remote loading.

The "energy method," which is based on finding the stationary point of the saddle surface of the energy release rate with respect to the electromagnetic field inside the crack [23] is employed to find the solution for the electric–magnetic fields inside the interface crack. Compact formulas for the energy release rate are derived for impermeable and permeable interface cracks. As a special solution, a crack in a monolithic anisotropic PEMO–elastic medium is also discussed by setting the upper and lower media identical. The conventional singularity of $r^{-(1/2)}$ is found for the crack tip fields in monolithic materials. This result is in good agreement with the results in the literature [18]. In Sec. 4, numerical results are presented to verify the characteristics of some bimaterial parameters and demonstrate the influence of the piezoelectromagnetic material properties on the energy release rate. The behavior of the energy release rate, G, is also studied under various loading conditions. In Sec. 5 we provide some useful conclusions.

2 Basic Equations

The basic equations, in extended Stroh's formalism, for PEMO–elastic material under generalized deformation are summarized in this section. The conditions to decouple the in-plane and anti-plane deformation are also discussed and some formulas are developed for the in-plane deformation.

In a fixed Cartesian coordinate system (x_1,x_2,x_3), the generalized Hooke's law for an elastic material considering both piezoelectric and piezomagnetic material properties may be written in the following form

$$\sigma_{ij} = c_{ijkl}u_{k,l} + e_{lij}\varphi^E_{,l} + \varrho_{lij}\varphi^H_{,l}$$

$$D_i = e_{ikl}u_{k,l} - \varepsilon_{il}\varphi^E_{,l} - \alpha_{il}\varphi^H_{,l}$$

$$B_i = \varrho_{ikl}u_{k,l} - \alpha_{li}\varphi^E_{,l} - \mu_{il}\varphi^H_{,l} \qquad (1)$$

where i,j,k,l range in $\{1,2,3\}$ and the repeated indices imply summation; the comma stands for differentiation with respect to corresponding coordinate variables; σ_{ij} is the elastic stress, u_k the elastic displacement; c_{ijkl} the elastic moduli tensor; D_i the electric displacements; φ^E the electrostatic potential; ε_{il} the electric permittivity; B_i the magnetic induction (magnetic fluxes); φ^H the magnetic scalar potential; μ_{il} the magnetic permeability; and e_{ikl}, ϱ_{ikl}, and α_{li} the piezoelectric, piezomagnetic, and magnetoelectric coefficients, respectively. For the material constants, the following relationships hold

$$c_{ijkl} = c_{jikl} = c_{ijlk} = c_{klij}; \quad e_{ikl} = e_{ilk}; \quad \varrho_{ikl} = \varrho_{ilk}$$

$$\alpha_{il} = \alpha_{li}; \quad \varepsilon_{il} = \varepsilon_{li}; \quad \mu_{il} = \mu_{li} \qquad (2)$$

The equilibrium equations read

$$\sigma_{ij,i} + f_j = 0, \quad D_{i,i} - f_e = 0, \quad B_{i,i} - f_m = 0 \qquad (3)$$

If one defines the extended displacements as

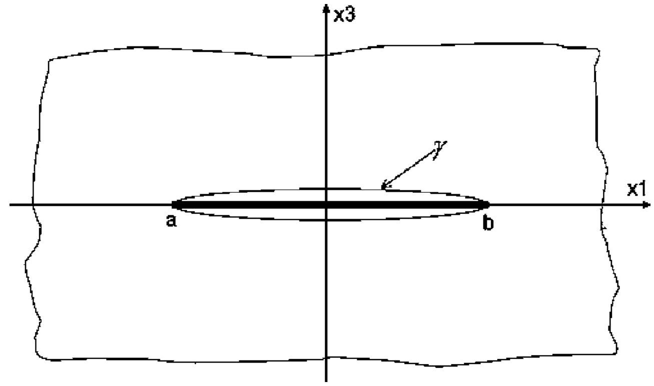


Fig. 1 An interface delamination between dissimilar piezoelectromagneto–elastic anisotropic bimedia and the associated contour integral path

$$U = [u_1, u_2, u_3, \varphi^E, \varphi^H]^T$$

or

$$U_K = u_k, \quad \text{for } K = 1,2,3; \quad U_4 = \varphi^E; \quad U_5 = \varphi^H \qquad (4)$$

and, correspondingly, extends the conventional 3×3 stress tensor to a 3×5 stress tensor

$$\sigma_{i_J} = \sigma_{ij}, \quad \text{for } J = 1,2,3; \quad \sigma_{i4} = D_i; \quad \sigma_{i5} = B_i \qquad (5)$$

then the equilibrium equations could be rewritten as

$$C_{iJKl}U_{K,li} + \mathbf{f}_J = 0 \qquad (6)$$

where C_{iJKl} are the extended material constants

$$C_{iJKl} = \begin{cases} C_{ijkl} & J,\ K = 1,2,3 \\ e_{l_J i} & J = 1,2,3;\ K = 4 \\ e_{i_K l} & J = 4;\ K = 1,2,3 \\ \varrho_{l_J i} & J = 1,2,3;\ K = 5 \\ \varrho_{i_K l} & J = 5;\ K = 1,2,3 \\ -\alpha_{il} & J = 4,\ K = 5 \\ -\alpha_{li} & J = 5,\ K = 4 \\ -\varepsilon_{il} & J = K = 4 \\ -\mu_{il} & J = K = 5 \end{cases} \qquad (7)$$

and $\mathbf{f}_J$ is the extended body force

$$\mathbf{f}_J = f_j, \quad \text{for } J = 1,2,3; \quad \mathbf{f}_4 = -f_e; \quad \mathbf{f}_5 = -f_m \qquad (8)$$

in which, f_i, f_e, f_m are the body force, electric charge, and magnetic charge, respectively.

2.1 Decoupling the In-Plane and Anti-Plane Deformation. For a plane system, the extended displacement field depends on two variables, namely x_1 and x_3 (Fig. 1). Then, expanding the equilibrium Eq. (3) leads to the expressions

$$C_{1JK1}U_{K,11} + (C_{1JK3} + C_{3JK1})U_{K,13} + C_{3JK3}U_{K,33} = \mathbf{f_J}, \quad J,K = 1, \ldots ,5 \qquad (9)$$

Rewriting Eq. (9) gives

$$C_{1JK1}U_{K,11} + (C_{1JK3} + C_{3JK1})U_{K,13} + C_{3JK3}U_{K,33} + C_{1J21}U_{2,11} + (C_{1J23} + C_{3J21})U_{2,13} + C_{3J23}U_{2,33} = \mathbf{f_J}, \quad J,K = 1,3,4,5$$

$$C_{12K1}U_{K,11} + (C_{12K3} + C_{32K1})U_{K,13} + C_{32K3}U_{K,33} + C_{1221}U_{2,11} + (C_{1223} + C_{3221})U_{2,13} + C_{3223}U_{2,33} = \mathbf{f}_2, \quad K = 1,3,4,5 \qquad (10)$$

To decouple the anti-plane and in-plane deformation, the coefficients for the terms involving U_2 in Eq. $(10)_1$ and not involving U_2 in Eq. $(10)_2$ should vanish, leading to the following conditions

$$C_{1J21} = C_{1J23} = C_{3J21} = C_{3J23} = 0, \quad J = 1,3,4,5$$

$$C_{12K1} = C_{12K3} = C_{32K1} = C_{32K3} = 0, \quad K = 1,3,4,5 \tag{11}$$

or, in contracted form

$$C_{14} = C_{16} = C_{34} = C_{36} = C_{54} = C_{56} = 0$$

$$e_{16} = e_{14} = e_{36} = e_{34} = 0, \quad \varrho_{16} = \varrho_{14} = \varrho_{36} = \varrho_{34} = 0 \tag{12}$$

Equation $(12)_1$ is the condition which decouples the anti-plane and in-plane deformation for an anisotropic material with no piezo-electromagnetic properties. One may call it the mechanical decoupling condition. Unlike the conventional anisotropic media, one may see that if a PEMO–elastic material only satisfies this mechanical decoupling condition, the in-plane loading may still produce an anti-plane deformation, or vice versa.

2.2 Basic Equations for In-Plane Deformation. Since the anti-plane interface crack problem was studied by Li and Kardomateas [23], the current work focuses on the interface crack problem under in-plane deformation. The extended displacements Eq. (4) may be redefined as

$$U = [u_1, u_3, \varphi^E, \varphi^H]^T$$

or

$$U_1 = u_1, \quad U_2 = u_3, \quad U_3 = \varphi^E; \quad U_4 = \varphi^H \tag{13}$$

A nontrivial displacement solution to Eq. (6) with the corresponding stress function $\psi_k (k=1,2,3,4)$, in the absence of the extended body force, takes the form

$$U = \sum_{J=1}^{4} [\mathbf{a_J g_J(z_J)} + \mathbf{\bar{a}_J \bar{g}_J(\bar{z}_J)}], \quad \psi = \sum_{\mathbf{J=1}}^{\mathbf{4}} [\mathbf{b_J g_J(z_J)} + \mathbf{\bar{b}_J \bar{g}_J(\bar{z}_J)}],$$

$$\mathbf{z_J = x_1 + p_J x_3} \tag{14}$$

where $\bar{z}$ denotes the conjugate of a complex z; p_J is a complex number; $\mathbf{a}_j$ is a column vector; and $g(z_J)$ is a function vector to be determined from the boundary conditions.

The stresses can be written in term of a stress function, ψ, as

$$\sigma_{i1} = -\frac{\partial \psi_i}{\partial x_3}, \quad \sigma_{i3} = \frac{\partial \psi_i}{\partial x_1} \tag{15}$$

Substitution of Eq. (14) into Eq. (3) leads to the following eigenequation

$$[Q + p_J(R + R^T) + p_J^2 T]\mathbf{a}_J = 0 \tag{16}$$

where

$$Q_{JK} = C_{1JK1}, \quad R_{JK} = C_{1JK3}, \quad T_{JK} = C_{3JK3}, \quad J,K = 1,3,4,5 \tag{17}$$

Specifically, when contracted notation is employed, one has

$$[Q_{JK}] = \begin{bmatrix} c_{11} & c_{15} & e_{11} & \varrho_{11} \\ c_{15} & c_{55} & e_{15} & \varrho_{15} \\ e_{11} & e_{15} & -\varepsilon_{11} & -\alpha_{11} \\ \varrho_{11} & \varrho_{15} & -\alpha_{11} & -\mu_{11} \end{bmatrix}$$

$$[R_{JK}] = \begin{bmatrix} c_{15} & c_{13} & e_{31} & \varrho_{31} \\ c_{55} & c_{53} & e_{35} & \varrho_{35} \\ e_{15} & e_{13} & -\varepsilon_{13} & -\alpha_{13} \\ \varrho_{15} & \varrho_{13} & -\alpha_{31} & -\mu_{13} \end{bmatrix} \tag{18}$$

$$[T_{JK}] = \begin{bmatrix} c_{55} & c_{35} & e_{35} & \varrho_{35} \\ c_{35} & c_{33} & e_{33} & \varrho_{33} \\ e_{35} & e_{33} & -\varepsilon_{33} & -\alpha_{33} \\ \varrho_{35} & \varrho_{33} & -\alpha_{33} & -\mu_{33} \end{bmatrix} \tag{19}$$

As for the elastic and the piezoelectric cases (Suo et al. [16] and Lothe and Barnett [24]), it can be shown that the p_J are complex, and that if p_J is an eigenvalue of Eq. (16), then $\bar{p}_J$ is also an eigenvalue of Eq. (16) [1]. The roots p_J will be assumed to be all distinct, and in this paper equal roots are viewed as the limiting case of the distinct roots. From the relationship $\sigma_{13}=\sigma_{31}$, one may obtain

$$\mathbf{b}_J = (R^T + p_J T)\mathbf{a}_J = -\frac{1}{p_J}(Q + p_J R)\mathbf{a}_J \tag{20}$$

The combination of Eqs. (16) and (20) readily leads to

$$N\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} N_1 & N_2 \\ N_3 & N_1^T \end{bmatrix}\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = p\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \tag{21}$$

where N is an 8×8 matrix with $N_1=-T^{-1}R^T$, $N_2=T^{-1}$, $N_3=RT^{-1}R^T-Q$; and the superscript T stands for the transpose of a matrix.

For the convenience of writing, we denote the extended traction vector on a surface $x_3=$constant, as

$$\mathbf{t} = [\sigma_{31}, \sigma_{33}, D_3, B_3]^T \tag{22}$$

Expression (14) may also be rewritten in vector form

$$\mathbf{u} = Ag(z_J) + \bar{A}\bar{g}(\bar{z}_J), \quad \psi = Bg(z_J) + \bar{B}\bar{g}(\bar{z}_J); \quad z_J = x_1 + p_J x_3 \tag{23}$$

where $A=[\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4]$, $B=[\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3,\mathbf{b}_4]$, and $p_J(J=1,2,3,4)$; these satisfy the orthogonality relations [25] after being properly normalized

$$\begin{bmatrix} B^T & A^T \\ \bar{B}^T & \bar{A}^T \end{bmatrix} \times \begin{bmatrix} A & \bar{A} \\ B & \bar{B} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{24}$$

Here, three useful matrices may be defined as

$$M = iAB^{-1}, \quad L = -2iBB^T, \quad S = i(2AB^T - I) \tag{25}$$

where $I=\mathrm{diag}[1,1,1,1]$ is the unit matrix. One can see from Eq. (24) that H and L are real and symmetric, whereas S and SL^{-1} are real and anti-symmetric. Moreover, the following relations can be verified.

$$M = L^{-1} + iL^{-1}S^T = L^{-1} - iSL^{-1} \tag{26}$$

which tells that M is Hermitian. The M matrix may be partitioned as

$$M = \begin{pmatrix} M_{11} & M_{13} & M_{14} \\ M_{31} & M_{33} & M_{34} \\ M_{41} & M_{43} & M_{44} \end{pmatrix} \tag{27}$$

where

$$M_{11} \sim [\text{elasticity}]^{-1}, \quad M_{33} \sim -[\text{permittivity}]^{-1}$$

$$M_{44} \sim -[\text{permeability}]^{-1}$$

$$M_{13} = \bar{M}_{31}^T \sim [\text{piezoelectricity}]^{-1} \tag{28}$$

$$M_{14} = \bar{M}_{41}^T \sim [\text{piezomagneticity}]^{-1}$$

$$M_{34} = \bar{M}_{43}^T \sim [\text{magnetoelectricity}]^{-1}$$

and M_{11} is positive definite [24], but $M_{33}<0$ and $M_{44}<0$.

If the coordinate system (x_1,x_2,x_3) transfers to a new coordinate system (x_1^*,x_2^*,x_3^*) by the in-plane rotation

$$\left[\frac{\partial x_i^*}{\partial x_j}\right]=\begin{bmatrix}\cos(\theta) & \sin(\theta) & 0\\ -\sin(\theta) & \cos(\theta) & 0\\ 0 & 0 & 1\end{bmatrix} \tag{29}$$

then one can easily show that

$$S(\theta)=\Omega^T(\theta)S\Omega(\theta), \quad L(\theta)=\Omega^T(\theta)L\Omega(\theta) \tag{30}$$

where

$$\Omega(\theta)=\begin{bmatrix}\cos(\theta) & \sin(\theta) & 0 & 0\\ -\sin(\theta) & \cos(\theta) & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{31}$$

A transformation similar to Eq. (30) was also addressed in Suo et al. [16] in their fracture mechanics study of piezoelectric material.

3 Interface Cracks in PEMO-Elastic Bimaterial Media

3.1 Statement of the Problem. Let the medium I occupy the upper half space (denoted by L) and medium II be in the lower-half space (denoted by R); the interface crack is assumed to be located in the region $a<x_1<b$, $-\infty<x_2<\infty$ of the plane $x_3=0$ (Fig. 1). The $\sigma_{i3}^\infty=p^\infty=[\sigma_{13}^\infty,\sigma_{33}^\infty,D_3^\infty,B_3^\infty]^T$ is applied at infinity. Under applied external loading, the crack may open and be filled with vacuum or air, in which an electric–magnetic field, denoted by D_3^0 and B_3^0, would be built up. This field is uniform if the applied loading σ_{i3}^∞ is uniform [23]. By the superposition principle, this interface crack problem is equivalent to the one under the applied loading

$$p=[\sigma_{13}^\infty,\ \sigma_{33}^\infty,\ \Delta D_3^0,\ \Delta B_3^0]^T; \quad \Delta D_3^0=D_3^\infty-D_3^0, \quad \Delta B_3^0=B_3^\infty-B_3^0 \tag{32}$$

acting on the interface crack surfaces while the loading vanishes at infinity.

3.2 Formulation of the Solution to the Interface Crack. The procedure to derive the solution is similar to the one employed in Li and Kardomateas [26]. From Eq. (23), one can have the following expressions for this bimedia

$$U^{\mathrm{I}}=A_{\mathrm{I}}\phi_{\mathrm{I}}(z_J)+\bar{A}_{\mathrm{I}}\bar{\phi}_{\mathrm{I}}(\bar{z}_J)$$

$$\psi^{\mathrm{I}}=B_{\mathrm{I}}\phi_{\mathrm{I}}(z_J)+\bar{B}_{\mathrm{I}}\bar{\phi}_{\mathrm{I}}(\bar{z}_J) \tag{33}$$

where $U^{\mathrm{I}},\psi^{\mathrm{I}}$ are displacement and stress functions for $z_J\in L$; and for medium II

$$U^{\mathrm{II}}=A_{\mathrm{II}}\phi_{\mathrm{II}}(z_J)+\bar{A}_{\mathrm{II}}\bar{\phi}_{\mathrm{II}}(\bar{z}_J)$$

$$\psi^{\mathrm{II}}=B_{\mathrm{II}}\phi_{\mathrm{II}}(z_J)+\bar{B}_{\mathrm{II}}\bar{\phi}_{\mathrm{II}}(\bar{z}_J) \tag{34}$$

where $U^{\mathrm{II}},\psi^{\mathrm{II}}$ are displacement and stress functions for $z_J\in R$. For the convenience of writing, the symbols I and II, denoting the quantities in medium L and R, respectively, may be put as superscripts or subscripts.

Making use of Eq. $(15)_2$, the boundary condition for this problem can be written for the interface rack region $(a<x_1<b, x_3=0)$ as

$$\psi_+'^{\mathrm{I}}(x_1)=\psi_-'^{\mathrm{II}}(x_1)=-p(x_1) \tag{35}$$

and along the interface outside the crack $(x_1<a$ and $b<x_1$, $x_3=0)$

$$U_+^{\mathrm{I}}(x_1)=U_-^{\mathrm{II}}(x_1), \quad \psi_+'^{\mathrm{I}}(x_1)=\psi_-'^{\mathrm{II}}(x_1) \tag{36}$$

and at infinity

$$\sigma_{ij}^{\mathrm{I}}=\sigma_{ij}^{\mathrm{II}}=0, \quad \text{at infinity} \tag{37}$$

where the convention $\psi(x_1,x_3)=\psi_\pm(x_1)$ as $x_3\to 0^\pm$ for any function $\psi(x_1,x_3)$ was used and will be employed in the following sections.

The displacement continuity along the bonded interface gives

$$A_{\mathrm{I}}\phi_{\mathrm{I}+}(x_1)-\bar{A}_{\mathrm{II}}\bar{\phi}_{\mathrm{II}+}(x_1)=A_{\mathrm{II}}\phi_{\mathrm{II}-}(x_1)-\bar{A}_{\mathrm{I}}\bar{\phi}_{\mathrm{I}-}(x_1) \tag{38}$$

One may define a function

$$\Phi(z)=\begin{cases}A_{\mathrm{I}}\phi_{\mathrm{I}}(z)-\bar{A}_{\mathrm{II}}\bar{\phi}_{\mathrm{II}}(z), & z\in L\\ A_{\mathrm{II}}\phi_{\mathrm{II}}(z)-\bar{A}_{\mathrm{I}}\bar{\phi}_{\mathrm{I}}(z), & z\in R\end{cases} \tag{39}$$

which automatically satisfies the condition (38) and is analytic on the whole plane except the cut along the interface crack.

Differentiation of Eq. (39) yields

$$\Phi'(z)=\begin{cases}A_{\mathrm{I}}\phi_{\mathrm{I}}'(z)-\bar{A}_{\mathrm{II}}\bar{\phi}_{\mathrm{II}}'(z), & z\in L\\ A_{\mathrm{II}}\phi_{\mathrm{II}}'(z)-\bar{A}_{\mathrm{I}}\bar{\phi}_{\mathrm{I}}'(z), & z\in R\end{cases} \tag{40}$$

The traction continuity on the bonded interface leads to

$$B_{\mathrm{I}}\phi_{\mathrm{I}+}'(x_1)-\bar{B}_{\mathrm{II}}\bar{\phi}_{\mathrm{II}+}'(x_1)=B_{\mathrm{II}}\phi_{\mathrm{II}-}'(x_1)-\bar{B}_{\mathrm{I}}\bar{\phi}_{\mathrm{I}-}'(x_1) \tag{41}$$

Similarly to the displacement continuity, a function which automatically satisfies the condition Eq. (41) may be defined as

$$\omega(z)=\begin{cases}B_{\mathrm{I}}\phi_{\mathrm{I}}'(z)-\bar{B}_{\mathrm{II}}\bar{\phi}_{\mathrm{II}}'(z) & z\in L\\ B_{\mathrm{II}}\phi_{\mathrm{II}}'(z)-\bar{B}_{\mathrm{I}}\bar{\phi}_{\mathrm{I}}'(z), & z\in R\end{cases} \tag{42}$$

which is analytical on the whole plane except the cut along the interface crack.

Solving Eqs. (40) and (42) gives for $z\in L$

$$B_{\mathrm{I}}\phi_{\mathrm{I}}'(z)=N[i\Phi'(z)+\bar{M}_{\mathrm{II}}\omega(z)]$$

$$\bar{B}_{\mathrm{II}}\bar{\phi}_{\mathrm{II}}'(z)=B_{\mathrm{I}}\phi_{\mathrm{I}}'(z)-\omega(z) \tag{43}$$

and for $z\in R$

$$B_{\mathrm{II}}\phi_{\mathrm{II}}'(z)=\bar{N}[i\Phi'(z)+\bar{M}_{\mathrm{I}}\omega(z)]$$

$$\bar{B}_{\mathrm{I}}\bar{\phi}_{\mathrm{I}}'(z)=B_{\mathrm{II}}\phi_{\mathrm{II}}'(z)-\omega(z) \tag{44}$$

In the above equations, the following matrix was used

$$N^{-1}=M_{\mathrm{I}}+\bar{M}_{\mathrm{II}}=D+iW, \quad D=L_1^{-1}+L_2^{-1}, \quad W=S_2L_2^{-1}-S_1L_1^{-1} \tag{45}$$

The matrix N is Hermitian since M_{I} and M_{II} are Hermitian; matrix D can be easily shown to be real symmetric and W to be real anti-symmetric.

Substituting Eqs. (43) and (44) into the boundary conditions Eqs. $(35)_{1,2}$, respectively, gives

$$B_{\mathrm{I}}\phi_{\mathrm{I}+}'(x_1)+B_{\mathrm{II}}\phi_{\mathrm{II}-}'(x_1)-\omega_-(x_1)=-p(x_1)$$

$$B_{\mathrm{II}}\phi_{\mathrm{II}-}'(x_1)+B_{\mathrm{I}}\phi_{\mathrm{I}+}'(x_1)-\omega_+(x_1)=-p(x_1) \tag{46}$$

Subtraction of Eq. $(46)_2$ from Eq. $(46)_1$ yields.

$$\omega_+(x_1)-\omega_-(x_1)=0 \tag{47}$$

which implies that $\omega(z)$ is continuous on the whole interface.

By the analytic continuation principle [27], the function $\omega(z)$ is analytical on the whole plane. But according to Liouville's theorem [27], this $\omega(z)$ must be a constant function in the whole domain. However, the condition in Eq. (37) imposes that this function vanish at infinity. Therefore, this constant function must be identically zero in the whole plane, i.e.

$$\omega(z)=0, \quad \text{for all } z \tag{48}$$

Then, either Eq. $(46)_1$ or Eq. $(46)_2$ leads to a general Hilbert equation in matrix notation

$$N\Phi_+'(x_1)+\bar{N}\Phi_-'(x_1)=ip(x_1), \quad a<x_1<b \tag{49}$$

The homogenous equation corresponding to the above general Hilbert Eq. (49) can be written as

$$NX_+(x_1)+\bar{N}X_-(x_1)=0, \quad a<x_1<b \tag{50}$$

The following function vector may be considered as possible solution to Eq. (50)

$$\chi(z)=v(z-a)^{-\delta}(z-b)^{\delta-1} \tag{51}$$

which is analytic in the whole plane except the cut along (a,b) and has the property

$$z\chi(z)\to 1, \quad \text{as} \quad |z|\to\infty \tag{52}$$

Substitution of Eq. (51) into Eq. (50) leads to a 4×4 eigenvalue system

$$(N+e^{2\pi i\delta}\bar{N})v=0, \quad \delta=1/2+i\epsilon \tag{53}$$

Since the procedure to obtain the solution to this eigenvalue problem (53) is significant for one to understand the singularities of the fields at the interface crack tip, a detailed study of Eq. (53) is presented in Appendix A, in which four possible eigenvalues of δ are found as

$$\delta_{1,2}=1/2\pm i\epsilon_1, \quad \delta_{3,4}=1/2\pm i\epsilon_2 \tag{54}$$

It is also shown in Appendix B that the bimaterial parameters ϵ_1 and ϵ_2 are real numbers and the corresponding four eigenvectors, $v_i(i=1\ldots 4)$, are complex and satisfy the following conditions

$$v_2=\bar{v}_1, \quad v_3=\bar{v}_4 \tag{55}$$

The matrix defined as $\mathbf{v}=[v_1,v_2,v_3,v_4]$ can also be written as $\mathbf{v}=[v_1,\bar{v}_1,v_3\bar{v}_3]$.

Since N is a Hermitian matrix, the following identity can be readily verified

$$\bar{\mathbf{v}}^TN\mathbf{v}=\begin{bmatrix}\bar{v}_1^TNv_1 & 0 & 0 & 0\\ 0 & v_1^TN\bar{v}_1 & 0 & 0\\ 0 & 0 & \bar{v}_3^TNv_3 & 0\\ 0 & 0 & 0 & v_3^TN\bar{v}_3\end{bmatrix} \tag{56}$$

Denoting

$$\gamma_1=\bar{v}_1^TNv_1, \quad \gamma_2=v_1^TN\bar{v}_1, \quad \gamma_3=\bar{v}_3^TNv_3, \quad \gamma_4=v_3^TN\bar{v}_3 \tag{57}$$

then all the γs are real numbers, and $\gamma_1\neq\gamma_2$, $\gamma_3\neq\gamma_4$, unless N is real symmetric. One can further show

$$\gamma_1=e^{2\pi\epsilon_1}\gamma_2 \quad \gamma_3=e^{2\pi\epsilon_2}\gamma_4 \tag{58}$$

and normalize $\mathbf{v}$ as

$$\bar{\mathbf{v}}^TN\mathbf{v}=\mathrm{diag}[\tilde{\gamma}_1e^{\pi\epsilon_1},\tilde{\gamma}_1e^{-\pi\epsilon_1},\tilde{\gamma}_2e^{\pi\epsilon_2},\tilde{\gamma}_2e^{-\pi\epsilon_2}] \tag{59}$$

where $\tilde{\gamma}_1=\gamma_1e^{-\pi\epsilon_1}$ and $\tilde{\gamma}_2=\gamma_3e^{-\pi\epsilon_2}$ are real numbers.

Therefore, the fundamental solution to the homogeneous Hilbert Eq. (50) would take the form

$$X(z)=\frac{1}{\sqrt{(z-a)(z-b)}}\mathbf{v}\Delta(z;\epsilon_1,\epsilon_2)$$
$$\Delta(z;\epsilon_1,\epsilon_2)=\mathrm{diag}\left[\left(\frac{z-b}{z-a}\right)^{i\epsilon_1},\left(\frac{z-b}{z-a}\right)^{-i\epsilon_1},\left(\frac{z-b}{z-a}\right)^{i\epsilon_2},\left(\frac{z-b}{z-a}\right)^{-i\epsilon_2}\right] \tag{60}$$

One may see that there are four modes of singularities for the crack tip fields and these singularities have the following form

$$(x_1-a)^{-(1/2)\mp i\epsilon_1}(x_1-b)^{-(1/2)\pm i\epsilon_1}, \quad (x_1-a)^{-(1/2)\mp i\epsilon_2}(x_1-b)^{-(1/2)\pm i\epsilon_2} \tag{61}$$

Hence, a solution to the nonhomogeneous Hilbert Eq. (49), which vanishes at infinity, can be formulated as

$$\Phi'(z)=\frac{X(z)}{2\pi i}\int_{\overline{ab}}\frac{[X_+(x_1)]^{-1}N^{-1}[ip(x_1)]dx_1}{x_1-z} \tag{62}$$

It can be seen that once the applied loading is given, a specific expression to Eq. (62) would be obtained, as would the displacement and stress functions.

For the applied constant loading $p(x_1)=p$, a closed form solution can be found by the contour integral method (Appendix B) as

$$\Phi_p'(z)=\mathbf{v}\left[I-\frac{\Delta(z;\epsilon_1,\epsilon_2)}{\sqrt{(z-a)(z-b)}}\Xi(z;\epsilon_1,\epsilon_2)\right]\mathbf{v}^{-1}[N+\bar{N}]^{-1}(ip) \tag{63}$$

where Ξ is defined as

$$\Xi(z;\epsilon_1,\epsilon_2)=\mathrm{diag}\left[z_1-\frac{(b+a)}{2}+(b-a)i\epsilon_1,z_2-\frac{(b+a)}{2}-(b-a)i\epsilon_1,z_3-\frac{(b+a)}{2}+(b-a)i\epsilon_2,z_4-\frac{(b+a)}{2}-(b-a)i\epsilon_2\right] \tag{64}$$

Further integration of Eq. (63) leads to

$$\Phi_p(z)=\mathbf{v}[\Pi(z)-\sqrt{(z-a)(z-b)}\mathbf{\Delta}(z;\epsilon_1,\epsilon_2)]\mathbf{v}^{-1}[N+\bar{N}]^{-1}(ip) \tag{65}$$

where

$$\Pi(z)=\mathrm{diag}[z_1,z_2,z_3,z_4] \tag{66}$$

and the terms contributing to rigid body motion have been omitted.

If we let r be the distance ahead of the crack tip, then, from expressions (63) and $(33)_2$ (or $(34)_2$), one can find that the crack tip fields, such as the extended stress field, can be expressed as the combination of

$$\sigma_{i_J}\sim r^{-(1/2)\pm i\epsilon_1}, \quad r^{-(1/2)\pm i\epsilon_2} \tag{67}$$

i.e., a combination of four different singularities in piezomagnetoelectro–elastic dissimilar bimaterials. It should be mentioned that for conventional dissimilar bimedia, only two singularities of the form $r^{-(1/2)\pm i\epsilon}$ exist (ϵ is real [6]) and for the piezoelectric dissimilar bimaterials, four possible singularities of the form $r^{-(1/2)\pm i\epsilon}$ and $r^{-(1/2)\pm\kappa}$ were found (ϵ and κ are real [15,16]). In Eq. (67), two new singularities of the form $r^{-(1/2)\pm i\epsilon_2}$ (ϵ_2 is real) can be observed. These new types of singularities reflect the effects from the simultaneous presence of the piezoelectric and the piezomagnetic material properties.

3.3 Interface Crack Characteristic Parameters. With the solution to the stress functions in the foregoing section, some interesting fracture characteristic parameters such as the crack tip stress intensity factors and the extended displacements jump near the crack tip; furthermore, the energy release rate can be readily derived.

From Eqs. $(43)_1$ and $(43)_2$, the extended traction along the interface can be expressed as

$$\mathbf{t}(x_1) = Ni\Phi'_+(x_1) + \bar{N}i\Phi'_-(x_1) \tag{68}$$

We will show that the right hand side of Eq. (68) is real as required.

Substituting the stress function Eq. (63) into Eq. (68) leads to

$$\mathbf{t}(x_1) = -p + [NX_+(x_1) + \bar{N}X_-(x_1)]\Xi(x_1,\epsilon_1,\epsilon_2)(Nv + \bar{N}v)^{-1}p \tag{69}$$

or, when Eqs. (50) and (53) are employed, it reads

$$\mathbf{t}(x_1) = \begin{cases} -p + (N+\bar{N})v\dfrac{\Delta(x_1;\epsilon_1,\epsilon_2)\Xi(x_1;\epsilon_1,\epsilon_2)}{\sqrt{(x_1-a)(x_1-b)}}v^{-1}(N+\bar{N})^{-1}p & x_1 < a \text{ and } b < x_1 \\ -p & a < x_1 < b \end{cases} \tag{70}$$

Making use of Eqs. (56) and (59), the extended traction at a distance r ahead of the crack tip such as b (Fig. 1) can be expressed in the form

$$\begin{aligned}\mathbf{t}(r) = \frac{1}{\sqrt{2\pi r}}\sqrt{\pi(b-a)/2}(N+\bar{N})\Bigg[&\frac{(1/2+i\epsilon_1)r^{i\epsilon_1}}{(b-a)^{i\epsilon_1}\tilde{\gamma}_1\cosh(\pi\epsilon_1)}v_1\bar{v}^T{}_1p \\ &+ \frac{(1/2-i\epsilon_1)r^{-i\epsilon_1}}{(b-a)^{-i\epsilon_1}\tilde{\gamma}_1\cosh(\pi\epsilon_1)}\bar{v}_1v_1^Tp \\ &+ \frac{(1/2+i\epsilon_2)r^{i\epsilon_2}}{(b-a)^{i\epsilon_2}\tilde{\gamma}_2\cosh(\pi\epsilon_2)}v_3\bar{v}_3^Tp \\ &+ \frac{(1/2-i\epsilon_2)r^{-i\epsilon_2}}{(b-a)^{-i\epsilon_2}\tilde{\gamma}_2\cosh(\pi\epsilon_2)}\bar{v}_3v_3^Tp\Bigg]\end{aligned} \tag{71}$$

where

$$p = [\sigma_{31}^\infty, \sigma_{33}^\infty, \Delta D_3^0, \Delta B_3^0]^T \tag{72}$$

One can easily see that the right side of Eq. (71) is a real vector, an expected result. The $v_i^Tp(i=1,3)$ are scalar (complex or real). Therefore, the interface traction ahead of the crack tip may be expressed in the space spanned by two eigenvectors (v_1,v_3) as

$$\begin{aligned}\mathbf{t}(r) = (N+\bar{N})\Bigg[&\frac{r^{i\epsilon_1}K_\sigma v_1}{\sqrt{2\pi r}\tilde{\gamma}_1\cosh^2(\epsilon_1\pi)} + \frac{r^{-i\epsilon_1}\bar{K}_\sigma\bar{v}_1}{\sqrt{2\pi r}\tilde{\gamma}_1\cosh^2(\epsilon_1\pi)} \\ &+ \frac{r^{i\epsilon_2}K_{DB}v_3}{\sqrt{2\pi r}\tilde{\gamma}_2\cosh^2(\epsilon_2\pi)} + \frac{r^{-i\epsilon_2}\bar{K}_{DB}\bar{v}_3}{\sqrt{2\pi r}\tilde{\gamma}_2\cosh^2(\epsilon_2\pi)}\Bigg]\end{aligned} \tag{73}$$

where K_σ and K_{DB} are complex numbers, defined as

$$K_\sigma = K_\mathrm{I} + iK_\mathrm{II} = \sqrt{\pi(b-a)/2}(1/2+i\epsilon_1)(b-a)^{-i\epsilon_1}\cosh(\epsilon_1\pi)\bar{v}_1^Tp$$

$$K_{DB} = K_D + iK_B = \sqrt{\pi(b-a)/2}(1/2+i\epsilon_2)(b-a)^{-i\epsilon_2}\cosh(\epsilon_2\pi)\bar{v}_3^Tp \tag{74}$$

These Ks can be called the extended stress intensity factors (ESIFs); similar notations have also been defined in the literature.

One can also extend the conventional crack open displacement (COD) to PEMO–electric materials. From Eqs. (33), (34), and (39), this extended crack open displacements (ECOD) can readily be evaluated by

$$\Delta\mathbf{u}(x_1) = \mathbf{u}_+^\mathrm{I}(x_1) - \mathbf{u}_-^\mathrm{II}(x_1) = \Phi_+(x_1) - \Phi_-(x_1) = \begin{cases}[(x_1-a)(b-x_1)]^{1/2}v[\Delta_+(x_1;\epsilon_1,\epsilon_2) + \Delta_-(x_1;\epsilon_1,\epsilon_2)]v^{-1}(N+\bar{N})^{-1}p, & a < x_1 < b \\ 0, & x_1 < a \text{ or } b < x_1\end{cases} \tag{75}$$

Then the ECOD at a small distance r behind the tip of the interface crack reads

$$\begin{aligned}\Delta\mathbf{u}(r) = 2\sqrt{\frac{r}{2\pi}}\sqrt{\pi(b-a)/2}v\,\mathrm{diag}\Bigg[&\frac{r^{i\epsilon_1}}{(b-a)^{i\epsilon_1}\tilde{\gamma}_1}, \\ &\frac{r^{-i\epsilon_1}}{(b-a)^{-i\epsilon_1}\tilde{\gamma}_1}, \frac{r^{i\epsilon_2}}{(b-a)^{i\epsilon_2}\tilde{\gamma}_2}, \frac{r^{-i\epsilon_2}}{(b-a)^{-i\epsilon_2}\tilde{\gamma}_2}\Bigg]\bar{v}^Tp\end{aligned} \tag{76}$$

The ECOD can be further expressed in terms of the ESIF

$$\begin{aligned}\Delta\mathbf{u}(r) = 2\sqrt{\frac{r}{2\pi}}\Bigg[&\frac{r^{i\epsilon_1}K_\sigma v_1}{(1/2+i\epsilon_1)\tilde{\gamma}_1\cosh(\pi\epsilon_1)} \\ &+ \frac{r^{-i\epsilon_1}\bar{K}_\sigma\bar{v}_1}{(1/2-i\epsilon_1)\tilde{\gamma}_1\cosh(\pi\epsilon_1)} + \frac{r^{i\epsilon_2}K_{DB}v_3}{(1/2+i\epsilon_2)\tilde{\gamma}_2\cosh(\pi\epsilon_2)} \\ &+ \frac{r^{-i\epsilon_2}\bar{K}_{DB}\bar{v}_3}{(1/2-i\epsilon_2)\tilde{\gamma}_2\cosh(\pi\epsilon_2)}\Bigg]\end{aligned} \tag{77}$$

a real vector, as expected.

Next, the energy release rate G can be computed and it reads

$$\begin{aligned}G = \frac{1}{2}\lim_{\delta\to0^+}\frac{1}{\delta}\int_0^\delta \mathbf{t}(r)^T\Delta\mathbf{u}(\delta-r)dr = &\frac{\bar{v}_1^T(N+\bar{N})v_1}{2\tilde{\gamma}_1\cosh^4(\epsilon_1\pi)}|K_\sigma|^2 \\ &+ \frac{v_1^T(N+\bar{N})\bar{v}_1}{2\tilde{\gamma}_1\cosh^4(\epsilon_1\pi)}|K_\sigma|^2 + \frac{v_3^T(N+\bar{N})\bar{v}_3}{2\tilde{\gamma}_2\cosh^4(\epsilon_2\pi)}|K_{DB}|^2 \\ &+ \frac{\bar{v}_3^T(N+\bar{N})v_3}{2\tilde{\gamma}_2\cosh^4(\epsilon_2\pi)}|K_{DB}|^2\end{aligned} \tag{78}$$

In deriving Eq. (78), the following identity was employed

$$\begin{aligned}\lim_{\delta\to0^+}\frac{1}{\delta}\int_0^\delta r^{-(1/2)\pm i\epsilon}(\delta-r)^{(1/2)\mp i\epsilon}dr &= \int_0^1 s^{-(1/2)\pm i\epsilon}(1-s)^{(1/2)\mp i\epsilon}ds \\ &= (1/2 \mp i\epsilon)\frac{\pi}{\cosh(\epsilon\pi)}\end{aligned} \tag{79}$$

Substituting Eq. (74) into Eq. (78), one can obtain

$$G = \frac{\pi(b-a)}{8} p^T H p \tag{80}$$

where

$$H = (1/2 + 2\epsilon_1^2)[\bar{v}_1\bar{v}_1^T(N+\bar{N})v_1v_1^T + v_1v_1^T(N+\bar{N})\bar{v}_1\bar{v}_1^T]/[\tilde{\gamma}_1{}^2\cosh^2(\epsilon_1\pi)] + (1/2+2\epsilon_2^2)[\bar{v}_3\bar{v}_3^T(N+\bar{N})v_3v_3^T + v_3v_3^T(N+\bar{N})\bar{v}_3\bar{v}_3^T]/[\tilde{\gamma}_2{}^2\cosh^2(\epsilon_2\pi)] \tag{81}$$

a symmetric real matrix.

All the formulas developed so far are functions of the unknown D_3^0 and B_3^0, the electric–magnetic field inside the interface crack. By finding the stationary point of the saddle surface of energy release rate with respect to the electromagnetic field inside the crack ("energy method" [23]), one can have the following equations in terms of D_3^0 and B_3^0

$$\frac{\partial G}{\partial D_3^0} = H_{13}\sigma_{31}^\infty + H_{23}\sigma_{33}^\infty + H_{33}\Delta D_3^0 + H_{34}\Delta B_3^0 = 0$$

$$\frac{\partial G}{\partial B_3^0} = H_{14}\sigma_{31}^\infty + H_{24}\sigma_{33}^\infty + H_{34}\Delta D_3^0 + H_{44}\Delta B_3^0 = 0 \tag{82}$$

which lead to

$$\Delta D_3^0 = D_3^\infty - D_3^0 = -\frac{H_{13}H_{44}-H_{14}H_{34}}{H_{33}H_{44}-H_{34}^2}\sigma_{13}^\infty - \frac{H_{23}H_{44}-H_{24}H_{34}}{H_{33}H_{44}-H_{34}^2}\sigma_{33}^\infty$$

$$\Delta B_3^0 = B_3^\infty - B_3^0 = -\frac{H_{14}H_{33}-H_{13}H_{34}}{H_{33}H_{44}-H_{34}^2}\sigma_{13}^\infty - \frac{H_{24}H_{33}-H_{23}H_{34}}{H_{33}H_{44}-H_{34}^2}\sigma_{33}^\infty \tag{83}$$

Then, the D_3^0 and B_3^0 can be calculated as

$$D_3^0 = D_3^\infty - \Delta D_3^0 = D_3^\infty - \frac{H_{14}H_{34}-H_{13}H_{44}}{H_{33}H_{44}-H_{34}^2}\sigma_{13}^\infty - \frac{H_{24}H_{34}-H_{23}H_{44}}{H_{33}H_{44}-H_{34}^2}\sigma_{33}^\infty$$

$$B_3^0 = B_3^\infty - \Delta B_3^0 = B_3^\infty - \frac{H_{13}H_{34}-H_{14}H_{33}}{H_{33}H_{44}-H_{34}^2}\sigma_{13}^\infty - \frac{H_{23}H_{34}-H_{24}H_{33}}{H_{33}H_{44}-H_{34}^2}\sigma_{33}^\infty \tag{84}$$

Now, one may further express the energy release rate in more explicit forms for two types of interface cracks: the impermeable and permeable interface cracks [14,16].

1. Impermeable interface crack, for which $D_3^0=0$ and $B_3^0=0$. The energy release rate reads

$$G_{\text{imp}} = \frac{\pi(b-a)}{8}[\sigma_{13}^\infty, \sigma_{33}^\infty, D_3^\infty, B_3^\infty]H[\sigma_{13}^\infty, \sigma_{33}^\infty, D_3^\infty, B_3^\infty]^T \tag{85}$$

2. Permeable interface crack, for which the electric–magnetic field, D_3^0 and B_3^0, inside the crack, is considered and given by Eq. (83). Substituting Eq. (83) into Eq. (80), one can obtain the energy release rate in a more explicit form as

$$G_{\text{perm}} = \frac{\pi(b-a)}{8}\left[\frac{\det(\tilde{H}_{22})}{\det(\hat{H})}(\sigma_{13}^\infty)^2 + \frac{\det(\tilde{H}_{12}+\tilde{H}_{21})}{\det(\hat{H})}(\sigma_{13}^\infty\sigma_{33}^\infty) + \frac{\det(\tilde{H}_{11})}{\det(\hat{H})}(\sigma_{33}^\infty)^2\right] \tag{86}$$

where, det( ) is the determinant of a square matrix; matrices $\tilde{H}_{\alpha\beta}$ $(\alpha,\beta=1,2)$ are the submatrices of H obtained by striking out the αth column and the βth row, and

$$\hat{H} = \begin{pmatrix} H_{33} & H_{34} \\ H_{43} & H_{44} \end{pmatrix} \tag{87}$$

In Eq. (86), one may clearly see that the applied mechanical loading σ_{13}^∞ and σ_{33}^∞ has a coupling effect on the energy release rate.

3.4 Special Case: A Crack in a Monolithic Piezoelectromagnetic Medium. As an illustration, this section will show that the solution to the Griffith type crack in monolithic piezomagnetoelectro–elastic solids can be conveniently obtained by setting the two media identical in the foregoing formulas of the interface crack problem. Specifically, the bimaterial matrix $D = L_1^{-1}+L_2^{-1} = 2L_1^{-1} = 2L_2^{-1} = 2L^{-1}$. Also, the N reduces to a 4×4 positive definite matrix, i.e.

$$N = \bar{N} = (2L^{-1})^{-1} = \tfrac{1}{2}L \tag{88}$$

The nonhomogenous Hilbert Eq. (49) then turns to

$$\Phi'_+(x_1) + \Phi'_-(x_1) = 2L^{-1}ip(x_1), \quad a < x_1 < b \tag{89}$$

A solution which vanishes at infinity is [26]

$$\Phi'(z) = \frac{1}{2\pi i}\operatorname{diag}\left[\frac{1}{\sqrt{(z-a)(z-b)}}\right] \times \int_{\overline{ab}} \frac{\left\{\operatorname{diag}\left[\frac{1}{\sqrt{(z-a)(z-b)}}\right]_+\right\}^{-1} L^{-1}[2ip(x_1)]dx_1}{x_1 - z} \tag{90}$$

If the applied loading is constant, then

$$\Phi'(z) = \operatorname{diag}\left[1 - \frac{z-(a+b)/2}{\sqrt{(z-a)(z-b)}}\right]L^{-1}(ip) \tag{91}$$

Integrating Eq. (91), results in

$$\Phi(z) = \operatorname{diag}[z - \sqrt{(z-a)(z-b)}]L^{-1}(ip) \tag{92}$$

where the constant contributing rigid body motion has been omitted.

Let

$$K = [K_{\text{II}}, K_{\text{I}}, K_D, K_B]^T \tag{93}$$

one then may easily show that the expression Eq. (93) becomes

$$K = \sqrt{\frac{\pi(b-a)}{2}}p \tag{94}$$

an interesting result that is similar in form to the conventional isotropic SIF. This result is also valid for some bimaterials with null bimaterial matrix W. Expressions (73) and (76) reduce, respectively, to.

$$\mathbf{t}(r) = \sqrt{\frac{1}{2\pi r}}K$$

and

Table 1 Properties of piezoelectromagneto–elastic bimaterial systems

Properties	Medium I	Medium II (1)	(2)	(3)	(4)
c_{11} (GPa)	86.74	166.0	202.0	226.0	250.0
c_{13} (GPa)	27.15	78.0	105.0	124.0	142.7
c_{33} (GPa)	102.83	162.0	194.2	216.0	237.3
c_{35} (GPa)	68.81	43.0	43.7	44.0	44.6
e_{11} (c/m^2)	0.171	0.0	0.0	0.0	0.0
e_{13} (c/m^2)	−0.0187	0.0	0.0	0.0	0.0
e_{35} (c/m^2)	−0.0761	0.0	0.0	0.0	0.0
e_{31} (c/m^2)	0.0	−4.4	3.08	−2.2	−1.32
e_{33} (c/m^2)	0.0	18.6	13.02	9.3	5.58
e_{15} (c/m^2)	0.0	0.0	8.12	5.8	3.48
ρ_{15} (N/Am)	550	550.0	174.1	275.0	385.0
ρ_{31} (N/Am)	580.3	550.0	165.0	290.2	406.2
ρ_{33} (N/Am)	669.7	699.7	209.9	350.0	489.8
ε_{11} (c/Nm^2)	39.21×10^{-12}	11.2×10^{-10}	78.6×10^{-10}	56.4×10^{-10}	34.2×10^{-10}
ε_{13} (c/Nm^2)	0.86×10^{-12}	0.0	0.0	0.0	0.0
ε_{31} (c/Nm^2)	0.86×10^{-12}	0.0	0.0	0.0	0.0
ε_{33} (c/Nm^2)	40.42×10^{-12}	12.6×10^{-10}	88.5×10^{-10}	63.5×10^{-10}	38.5×10^{-10}
μ_{11} (Ns^2/C^2)	5.50×10^{-6}	5.0×10^{-6}	180.5×10^{-6}	297.0×10^{-6}	414.5×10^{-6}
μ_{33} (Ns^2/C^2)	10.0×10^{-6}	10.0×10^{-6}	541.0×10^{-6}	83.5×10^{-6}	112.9×10^{-6}

$$\Delta\mathbf{u}(r) = 4\sqrt{\frac{r}{2\pi}}L^{-1}K \tag{95}$$

Equations (95) can also be directly obtained from the functions (91) and (92).

Equation (95) shows that the crack tip field for the monolithic material is proportional to the inverse of the square root of r, i.e.

$$\sigma_{i_J} \sim \frac{1}{\sqrt{2\pi r}} \tag{96}$$

a result that is in agreement with the one obtained in Song and Sih [18].

The energy release rate reads as

$$G_0 = \frac{1}{2}K^T L^{-1}K = \frac{\pi(b-a)}{4}p^T L^{-1}p \tag{97}$$

which can also be obtained by substituting Eq. (95) into Eq. $(78)_1$.

4 Numerical Results and Discussion

The data for the piezoelectric and piezomagnetic properties of the upper and lower media of the dissimilar bimaterial systems are selected from the literature [4,18] and recorded in Table 1. The constituents of each of the bimaterial systems are PEMO—elastic materials.

Table 2 gives the results of some bimaterial parameters such as c_2 and c_4 [defined by Eq. (102)], $\beta_{1,2}$, and $\beta_{3,4}$ [defined by Eq. (103)]. One can see from these numerical results that c_2 and c_4 are larger than zero. $\beta_{1,2}$, $\beta_{3,4}$ are real numbers, and so are ϵ_1 and ϵ_2. These observations are in agreement with the results proved in Appendix A and show that four possible singularities of the form $r^{-(1/2)\pm i\epsilon_1}$ and $r^{-(1/2)\pm i\epsilon_2}$ with real ϵ_1 and ϵ_2 exist around the interface crack tip in PEMO–elastic bimaterials.

The results in Figs. 2 and 3 are used to demonstrate the influence of the bimaterial parameter c_2 on the energy release rate, G. These results show that the energy release rate increases as c_2 increases both for a permeable and an impermeable interface crack; the energy release rate of a permeable interface crack is larger than that of an impermeable interface crack if only the loading σ_{33} (far field tension normal to the crack) is applied (Fig. 2), while the energy release rate for a permeable interface crack is the same as that for an impermeable interface when the loading is only σ_{13} (far field in-plane shear, Fig. 3).

The bimaterial system Medium I—Medium II (1) ($\epsilon_1 = 0.00950057$, $\epsilon_2 = 0.00337206$) will be used as an example in a further study to illustrate the energy release behavior of interface cracks in PEMO—elastic bimaterial solids.

Figure 4 plots the results of the energy release rate G of a permeable and an impermeable interface crack under any combination of loading σ_{33} and σ_{13}. It can be seen that the value of the energy release rate, G, for an interface crack, if considered as permeable, is larger than that of an interface crack if considered as impermeable, for any given pair of $(\sigma_{33}, \sigma_{13})$. Some details of this observation are shown in Figs. 5 and 6, where Fig. 5 is the variation of energy release rate versus the change of applied loading

Table 2 Bimaterial parameters

Bimaterial parameters	Bimaterial systems: Medium I–Medium II(1)	Medium I–Medium II(2)	Medium I–Medium II(3)	Medium I–Medium II(4)
c_2	2.0104×10^{-3}	2.4613×10^{-3}	5.1277×10^{-3}	7.5233×10^{-3}
c_4	1.6038×10^{-6}	5.5027×10^{-6}	7.6662×10^{-6}	1.01275×10^{-6}
$\beta_{1,2}$	±0.05976423	±0.05662615	±0.0971787	±0.11975132
$\beta_{3,4}$	±0.02119044	±0.04142584	±0.0284918	±0.02657485
ε_1	0.00950057	0.00900272	0.01541805	0.01896868
ε_2	0.00337206	0.00658936	0.00453338	0.00422852

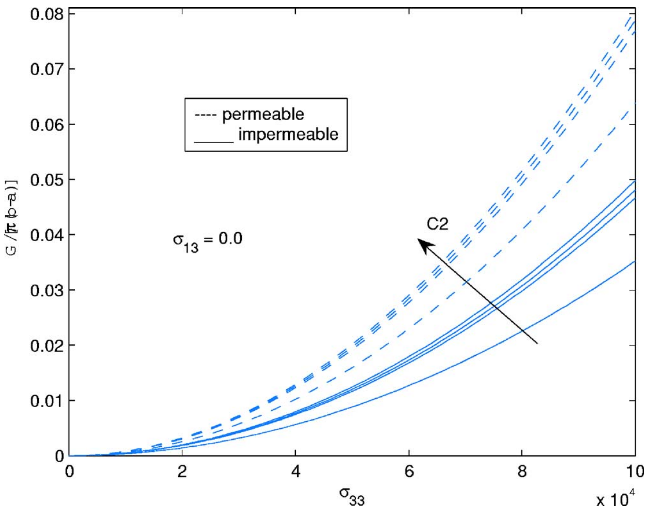


Fig. 2 Energy release rate versus bimaterial constant c_2 under pure mechanical loading σ_{33}

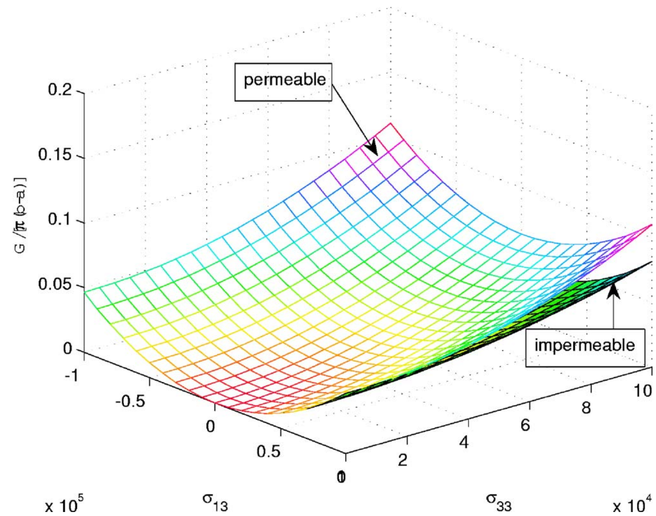


Fig. 4 Energy release rate for the combined mechanical loading σ_{13} and σ_{33}

σ_{33} for two given values of σ_{13}, and Fig. 6 is the variation of the energy release rate versus the change of applied loading σ_{13} for three given values of σ_{33}.

Figure 7 shows the results of energy release rate values for an impermeable interface crack under any combined electric–magnetic loading (D_3, B_3). An interesting phenomenon that can be seen here is that all the values of energy release rate G are less than or equal to zero. Negative energy release rates are physically impossible. This observation implies that a pure electric–magnetic loading (D_3, B_3) would be expected to retard the propagation of an interface crack in PEMO–elastic bimaterial systems. This retardation mechanism will be more clearly seen in the following discussions. Moreover, this retardation phenomenon has also been reported in the literature for cracks in monolithic electromagnetic materials [17].

Figures 8–11, respectively, demonstrate the influence of the applied electric or magnetic field on the energy release rate G under applied mechanical tensile loading σ_{33}. Figure 8 is the variation of G versus the applied loading σ_{33} for two given values of D_3 applied in different directions, namely, positive direction ($D_3 > 0$) and negative direction ($D_3 < 0$) and Fig. 9 is the variation of G versus any combination of loading (σ_{33}, D_3). Figure 10 is the G versus the applied loading σ_{33} for two given values of B_3 and Fig. 11 is G versus any combination of loading (σ_{33}, B_3). The results in Figs. 8 and 10 show that for a given D_3 or B_3, the applied mechanical loading σ_{33} has to exceed a certain value in order to reach a positive G. Here, we may call this value the thrust value, denoted as σ_{33}^{thr}.

Figures 8 and 10 also show that the values of the σ_{33}^{thr} are different for the applied D_3 or B_3 with the same amplitudes but different directions. One can see that the direction of applied D_3 or B_3 has an influence on the energy release rate, G. The influence of the direction of the electric or magnetic field can be viewed more clearly in Fig. 12, which shows the variation of G versus σ_{33} under the combined influence from a given (D_3, B_3). Here, a more subtle observation needs to be pointed out. The results in Fig. 7 show that the bigger the value of pure applied D_3 or/and B_3 is, the bigger a negative value G reaches. But this does not mean that the bigger value of applied D_3 or/and B_3 always makes the energy release rate G smaller (than the corresponding G with a smaller value of applied D_3 or/and B_3) when a mechanical loading is present. This can be easily verified from the results in Figs. 8 and

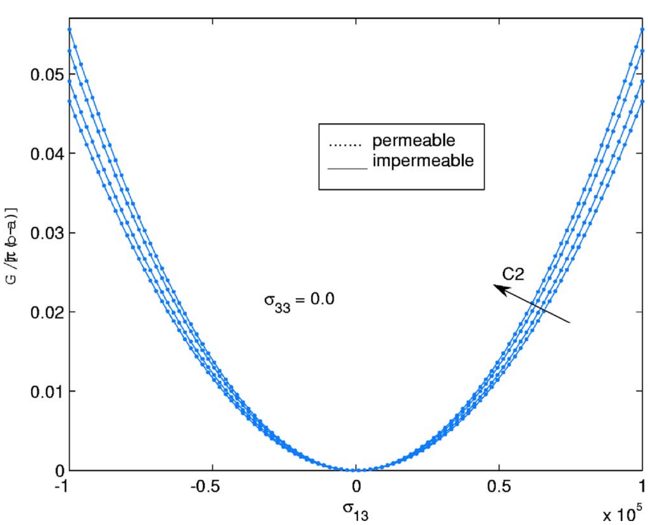


Fig. 3 Energy release rate versus bimaterial constant c_2 under pure mechanical loading σ_{13}

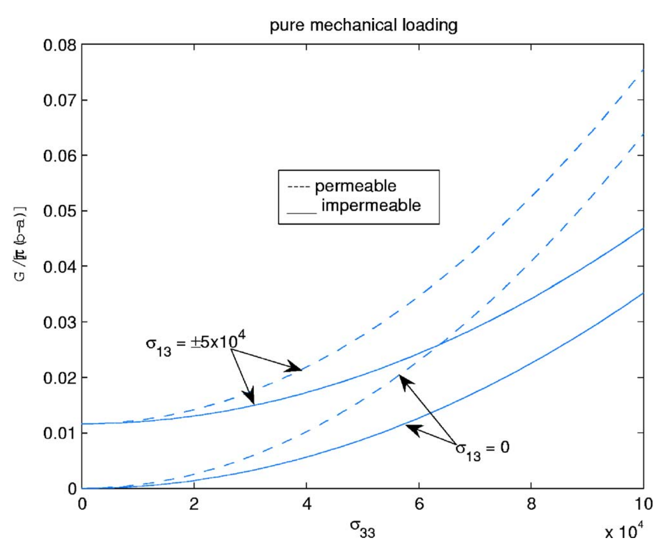


Fig. 5 Energy release rate versus σ_{33} for a given σ_{13}

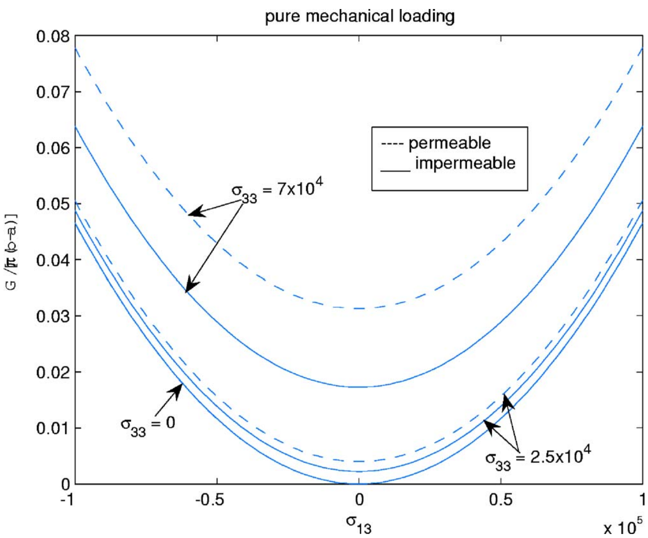


Fig. 6 Energy release rate versus σ_{13} for a given σ_{33}

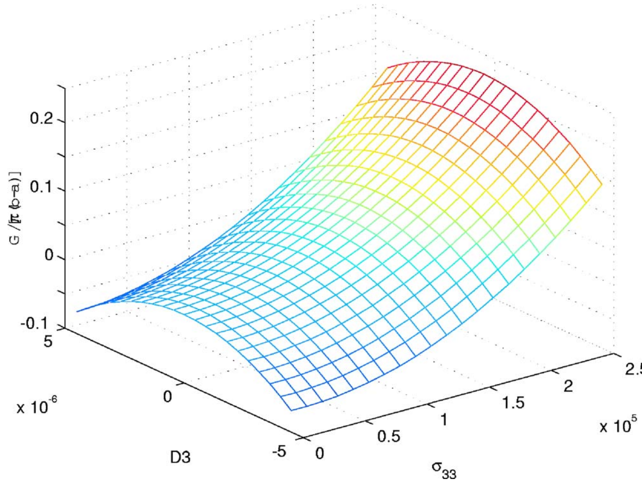


Fig. 9 Energy release rate under combined mechanical σ_{33} and electrical D_3 loading

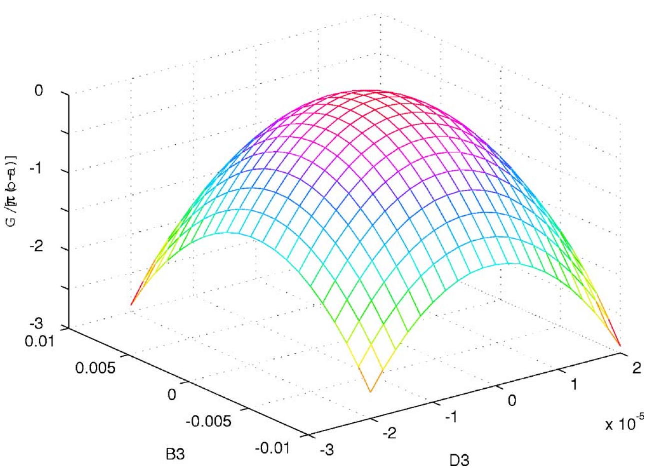


Fig. 7 Energy release rate under pure electric and magnetic applied loading

10. Particularly in Fig. 10, when $\sigma_{33} < 2.8 \times 10^5$, the G is bigger for $B_3 = 0.25 \times 10^{-2}$ than for $B_3 = 0.5 \times 10^{-2}$. But the trend reverses when $\sigma_{33} > 2.8 \times 10^5$.

One can further see that the surface of G is not symmetric with respect to the plane $D_3 = 0$ or $B_3 = 0$, as shown in Figs. 9 and 11. It is possible that the value of G with $D_3 = 0$ or $B_3 = 0$ is smaller than the value of G with $D_3 > 0$ or $B_3 > 0$ when σ_{33} reaches a certain value. This result can be more explicitly observed from the results in Fig. 11. Therefore, an important conclusion which can be drawn here is that the applied D_3 and B_3 do not always contribute a negative value to the energy release rate G when an applied mechanical loading σ_{33} is also present. This conclusion further suggests that the applied electric and magnetic loading does not always retard the propagation of an interface crack. Instead, under certain conditions of the applied mechanical loading, σ_{33}, they may actually speed the propagation of an interface crack.

Figures 13–15 study the influence of D_3 or B_3 on the energy release rate, G, under mechanical applied shear loading σ_{13}. Figure 13 is the variation of G versus any combination of (σ_{13}, D_3). One can see that the surface of G is symmetric both with respect to the axis $\sigma_{13} = 0$ and the $D_3 = 0$. This observation indicates that the direction of D_3 has no effect on the energy release rate G if the applied loading is only σ_{13} and $B_3 = 0$. This result can be easily

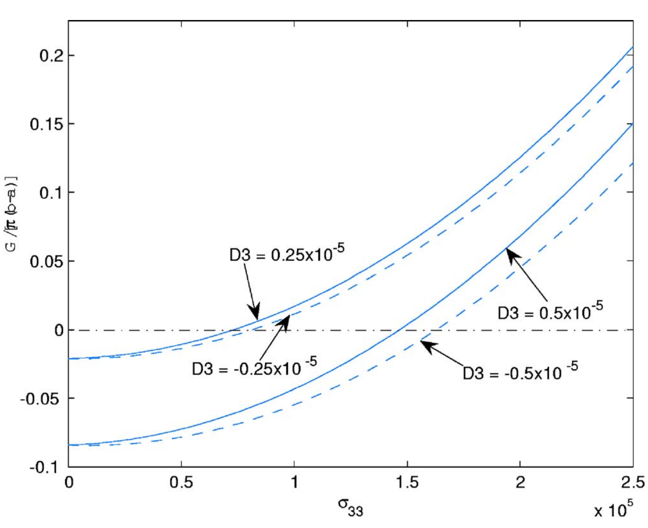


Fig. 8 Energy release rate versus σ_{33} for a given D_3

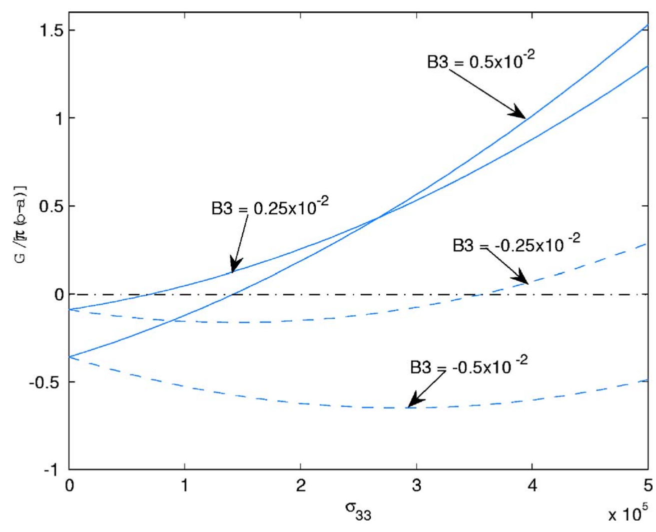


Fig. 10 Energy release rate versus σ_{33} for a given B_3

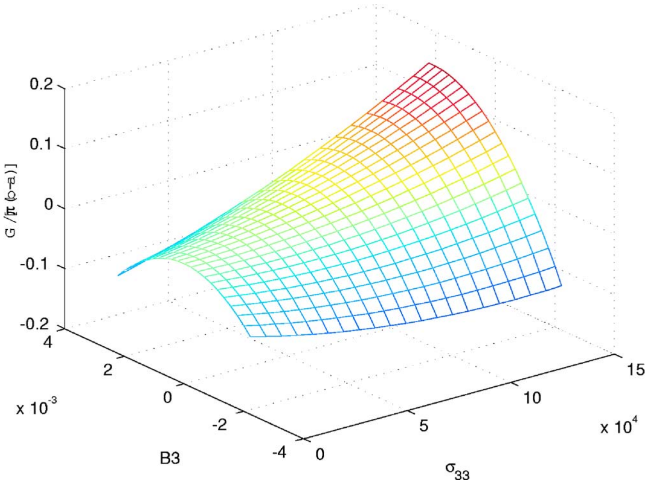


Fig. 11 Energy release rate under combined mechanical σ_{33} and magnetic B_3 loading

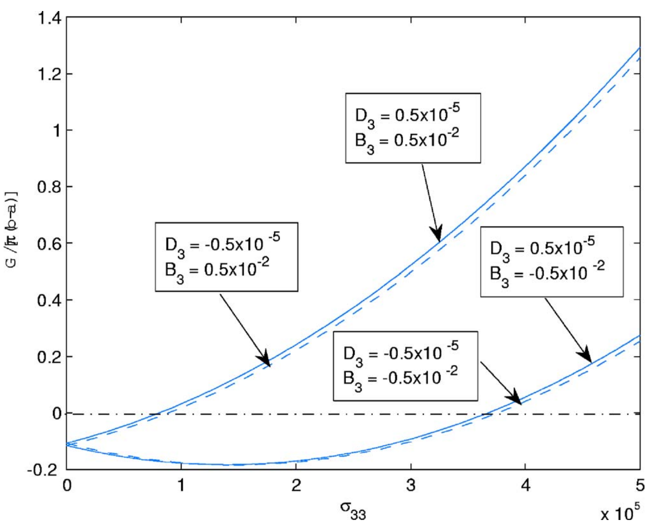


Fig. 12 Energy release rate under loading σ_{33} for a given ($\pm B_3, \pm D_3$)

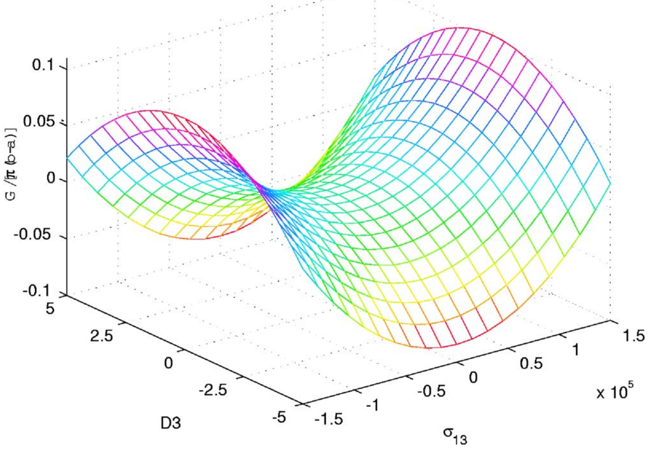


Fig. 13 Energy release rate under loading σ_{13} and D_3

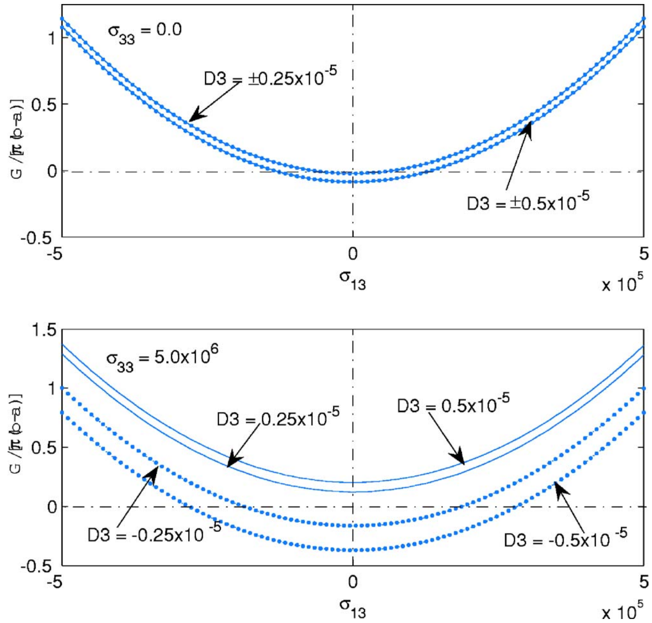


Fig. 14 Variation of energy release rate versus σ_{13}: top for a given D_3 only; bottom for a given pair (D_3, σ_{33})

seen from the top graphic in Fig. 14. A similar tendency can also be found for the case when only σ_{13} and B_3 are applied, as shown in the top graph of Fig. 15. As a comparison, the results with $\sigma_{33} \neq 0$ are also plotted at the bottom of Figs. 14 and 15. The results in Figs. 13–15 also show that the applied electric (D_3) or/and magnetic (B_3) field(s) usually retard(s) the propagation of an interface crack when the applied mechanical loading is only the shear loading σ_{13}.

Figures 16 and 17 plot the results of the energy release rate G under both tensile and shear in-plane applied loading. Figure 16 is the G versus $(\sigma_{33}, \sigma_{13})$ under simultaneously given D_3 and B_3, and Fig. 17 is a particular case for $\sigma_{13} = \pm 3.25 \times 10^5$. These results demonstrate that different combinations of the directions of D_3 and B_3 produce different results of the energy release rate G. For a given σ_{33} and σ_{13}, there exist a direction for D_3 and B_3 that makes the G maximum. One can also see that for any given D_3 and B_3, the G is symmetric with respect to σ_{13}. This observation, together with similar observations in Figs. 13 and 15, show that the direction of in-plane shear loading σ_{13} has no effect on G. Although when individually applied with $\sigma_{33}=0$, the direction of D_3 or B_3 does not affect the G as shown in Fig. 13, the top of Fig. 14, and the top of Fig. 15, the directions of D_3 or B_3 do have effects on the G when they are applied together as clearly shown in Fig. 16 and 17, say, values are different when directions of D_3 and B_3 are different with $\sigma_{33}=0$ and $\sigma_{13} = \pm 3.25 \times 10^5$, as depicted in Fig. 17.

5 Conclusions

Four possible singularities of the form $r^{-(1/2)\pm\epsilon_1}$ and $r^{-(1/2)\pm i\epsilon_2}$ exist for the fields around an interface crack tip in dissimilar PEMO—elastic bimaterial media. The bimaterial parameters ϵ_1 and ϵ_2 are proven to be real numbers for practical materials. The electric—magnetic field inside the crack is solved by finding the stationary point of the saddle surface of the energy release rate with respect to the electromagnetic field inside the crack. The energy release rate, G, can be expressed in compact form both for impermeable and permeable interface cracks. The value of G increases as the bimaterial parameter c_2 (defined by Eq. (102) in Appendix B) increases. When the only applied mechanical loading is σ_{13} (in-plane shear), the directions of separately applied D_3

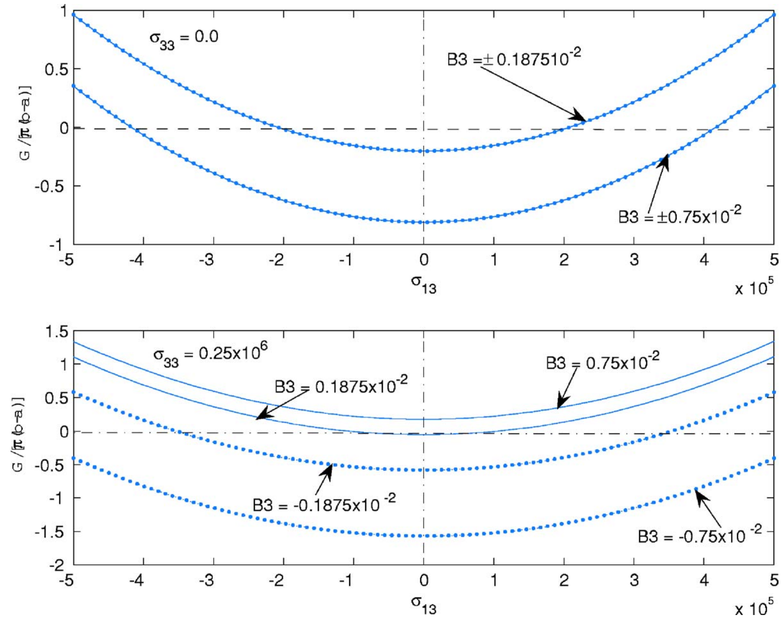


Fig. 15 Variation of energy release rate versus σ_{13}: top for a given B_3 only; bottom for a given pair (B_3, σ_{33})

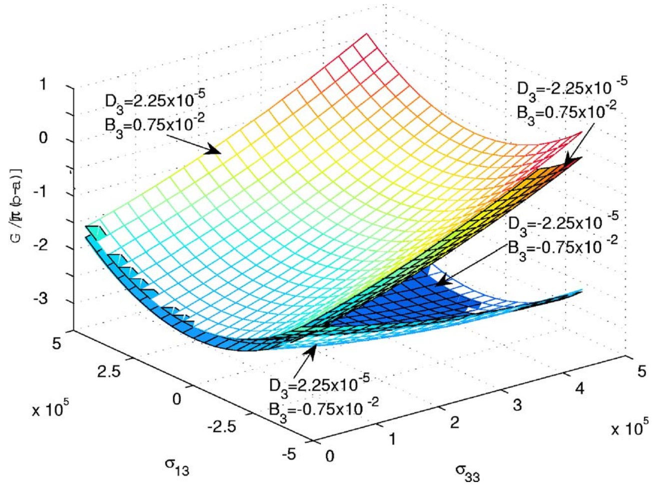


Fig. 16 Variation of energy release rate versus $(\sigma_{33}, \sigma_{13})$ for a given (D_3, B_3)

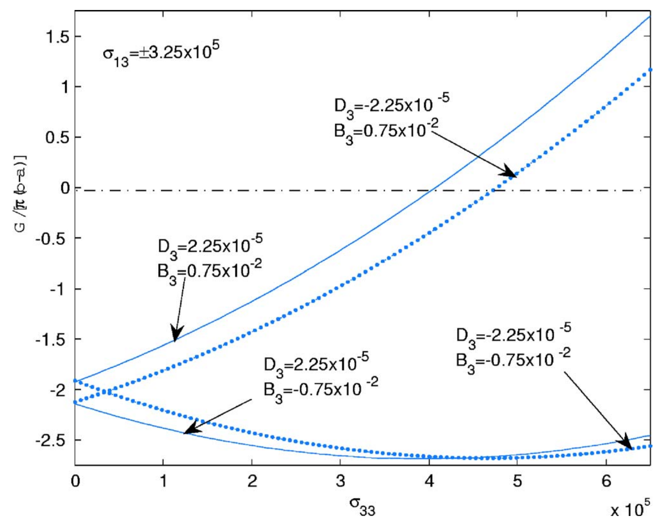


Fig. 17 Variation of energy release rate versus σ_{33} for a given (σ_{13}, D_3, B_3)

and B_3 do not affect the value of G while the directions of simultaneously applied D_3 and B_3 do. But the directions of applied D_3 and B_3 always have influences on the value of energy release rate G if the applied mechanical loading involves σ_{33} (tension). There exist a pair of directions of D_3 and B_3 which makes the G the maximum for each given mechanical loading. Pure applied electric–magnetic loading lowers G and therefore is expected to retard the propagation of an interface crack. However, the electric–magnetic loading does not always retard the propagation of an interface crack when the mechanically applied loading includes σ_{33} (tension). They can also foster the propagation of an interface crack if the applied mechanical loading reaches a certain value. The results or observations in this paper are still fundamental for the investigation of dissimilar piezoelectromagneto–elastic bimaterial solids. There are still more studies needed on this subject, such as in finding the criteria for the propagation of an interface crack, in understanding how the electric and magnetic fields inside an interface crack would interfere with the measured signals in broad band probes, etc. Nevertheless, the results in the current study may serve as a basis for more complex investigations.

Acknowledgment

The financial support of the Office of Naval Research, Grant No. N00014-90-J-1995, and the interest and encouragement of the Grant Monitor, Dr. Y. D. S. Rajapakse, are both gratefully acknowledged.

Appendix A. Bimaterial Constants: Real ϵ_1, ϵ_2

To ensure a nontrivial solution for the homogeneous Hilbert Eq. (50), the following condition should be satisfied

$$\|N + e^{2\pi i\delta}\bar{N}\| = \|N - e^{2\pi\epsilon}\bar{N}\| = 0 \tag{98}$$

where $\delta = 1/2 - i\epsilon$. Substituting Eq. $(45)_2$ into Eq. (98) leads to

$$\|W + i\beta D\| = 0 \tag{99}$$

where

$$\beta = \frac{e^{2\pi\epsilon} - 1}{e^{2\pi\epsilon} + 1}, \quad \text{or} \quad \beta = \tanh(\pi\epsilon) \tag{100}$$

Since N is Hermitian, from the definition of Eq. (100) one may easily see that if β and $\bar{\beta}$ are roots of Eq. (99), so are $-\beta$ and $-\bar{\beta}$. Therefore, Eq. (99) should have the form

$$p(i\beta) = (i\beta)^4 + 2c_2(i\beta)^2 + c_4 = 0 \tag{101}$$

where

$$2c_2 = -\tfrac{1}{2}\mathrm{tr}(D^{-1}W)^2, \quad c_4 = \|(D^{-1}W)\| \tag{102}$$

Then, the roots of Eq. (101) can be expressed as follows

$$\beta_{1,2} = \pm\sqrt{c_2 + \sqrt{(c_2)^2 - c_4}}$$

$$\beta_{3,4} = \pm\sqrt{c_2 - \sqrt{(c_2)^2 - c_4}} \tag{103}$$

These βs can be verified as real numbers. Actually, for the square matrices D and W, one has $\|(D^{-1}W)\| = \|D^{-1}\| \times \|W\|$, but $\|W\| \geqslant 0$, a property of an anti-symmetric matrix of even order, and $\|D^{-1}\| \geqslant 0$, a result from Eq. (27) and (28). So $c_4 = \|(D^{-1}W)\| \geqslant 0$. Mathematically, $(c_2)^2$ could be less than c_4. But for practical PEMO—elastic bimaterials, if $(c_2)^2 < c_4$, then $c_2 \pm \sqrt{(c_2)^2 - c_4}$ would be complex numbers with nonzero real and imaginary parts. Consequently, all the βs would be complex numbers, as would all the ϵs. This would contradict the fact in the literature that at least two singularities should have the form of $r^{-1/2 \pm i\tilde{\epsilon}}$ with real $\tilde{\epsilon}$, in the case of bimedia with no piezoelectric/piezomagnetic material properties. Hence $(c_2)^2 \geqslant c_4$, as shown in Table 2. Further, the c_2 also should not be less than zero for practical materials. If c_2 is less than zero, one could find that all the βs in Eq. (103) would be pure imaginary numbers. Then from Eq. (100), all the singularities would have the form $r^{-1/2 \pm \hat{\epsilon}}$ with real $\hat{\epsilon}$. This assertion would also contradict the fact that for bimedia with no piezoelectric/piezomagnetic material properties, at least two singularities have the form of $r^{-1/2 \pm i\tilde{\epsilon}}$ with real $\tilde{\epsilon}$. Hence, $c_2 \geqslant 0$ (also shown in Table 2 for practical PEMO—elastic bimaterials). Therefore, one can conclude that the $\beta_{1,2}$, $\beta_{3,4}$ are real numbers, and so are the ϵ_1 and ϵ_2. One may realize that c_2 and c_4 are simultaneously equal to zero if W is null, a special case similar to the one discussed by Qu and Li [13] for conventional dissimilar anisotropic bimaterials.

Appendix B. Contour Integral for $\Phi(z)'$ and $\Phi(z)$

The method used here can be viewed as the generalization of the technique in Muskhelishvili (Ref. [28] Secs. 110, 70), which is for a single equation. Let γ be a contour which includes the arc $\overline{ab}$, and let this contour shrink into the arc $\overline{ab}$ (Fig. 1), then for the $q(x_1)$ constant

$$\int_\gamma \frac{[X(\xi)]^{-1}N^{-1}}{\xi - z}d\xi = \int_{\overline{ab}} \frac{[X_+(x_1)]^{-1}N^{-1}}{x_1 - z}dt + \int_{\overline{ba}} \frac{[X_-(x_1)]^{-1}\bar{N}}{x_1 - z}dx_1$$
$$= \int_{\overline{ab}} \frac{[X_+(x_1)]^{-1}N^{-1}}{x_1 - z}dx_1 - \int_{\overline{ab}} \frac{[X_-(x_1)]^{-1}N^{-1}}{x_1 - z}dx_1 \tag{104}$$

From Eq. (50), one can have

$$X_-(x_1) = -\bar{N}^{-1}NX_+(x_1), \quad a < x_1 < b \tag{105}$$

Substituting Eq. (105) into Eq. (104) leads to

$$\int_\gamma \frac{[X(\xi)]^{-1}N^{-1}}{\xi - z}d\xi = \int_{\overline{ab}} \frac{[X_+(x_1)]^{-1}N^{-1}[I + \bar{N}N^{-1}]}{x_1 - z}dx_1 \tag{106}$$

Then,

$$\int_{\overline{ab}} \frac{[X_+(x_1)]^{-1}N^{-1}}{x_1 - z}dx_1 = \int_\gamma \frac{[X(\xi)]^{-1}N^{-1}[I + \bar{N}N^{-1}]^{-1}}{\xi - z}d\xi$$
$$= \int_\gamma \frac{[X(\xi)]^{-1}[N + \bar{N}]^{-1}}{\xi - z}d\xi \tag{107}$$

but

$$J = \frac{1}{2\pi i}\int_\gamma \frac{[X(\xi)]^{-1}[N + \bar{N}]^{-1}}{\xi - z}d\xi = \{X(z)^{-1}v - [\Xi + \Pi]\}v^{-1}[N + \bar{N}]^{-1} \tag{108}$$

where Ξ and Π are defined in Eq. (64). Therefore

$$\Phi'(z) = \frac{X(z)}{2\pi} \times 2\pi i J = v\left\{I - \frac{\Delta(z;\epsilon_1,\epsilon_2)}{\sqrt{(z-a)(z-b)}}[\Xi + \Pi]\right\}v^{-1}[N + \bar{N}]^{-1}ip \tag{109}$$

References

[1] Alshits, V. I., Darinskii, A. N., and Lothe, J., 1992, "On the Existence of Surface Waves in Half-Infinite Anisotropic Elastic Media With Piezoelectric and Piezomagnetic Properties," Wave Motion, **16**, pp. 265–283.
[2] Benveniste, Y., 1995, "Magnetoelectric Effect in Fibrous Composites With Piezoelectric and Piezomagnetic Phases," Phys. Rev. B, **51**(22), pp. 16424–16427.
[3] Chung, M. Y., and Ting, T. C. T., 1995, "The Green Function for a Piezoelectric Piezomagnetic Anisotropic Elastic Medium With an Elliptic Hole or Rigid Inclusion," Philos. Mag. Lett., **72**(6), pp. 405–410.
[4] Pan, E., 2002, "Three-Dimensional Green's Functions in Anisotropic Magneto-Electro-Elastic Bimaterials," J. Math. Phys., **53**, pp. 815–838.
[5] Bichurin, M. I., Filippov, D. A., Petrov, V. M., Laletsin, V. M., and Paddubnaya, N., 2003, "Resonance Magnetoelectric Effect in Layered Magnetostrictive-Piezoelectric Composites," Phys. Rev. B, **68**, pp. 132408-1–132408-4.
[6] Williams, M. L., 1959, "The Stress Around a Fault or Crack in Dissimilar Media," Bull. Seismol. Soc. Am., **49**(2), pp. 199–204.
[7] Erdogan, F., 1963, "The Stress Distribution in a Nonhomogeneous Elastic Plane With Cracks," J. Appl. Mech., **85**, pp. 232–236.
[8] England, A. H., 1965, "A Crack Between Dissimilar Media," J. Appl. Mech., **32**, pp. 400–402.
[9] Rice, J. R., and Sih, G. C., 1965, "Plane Problems of Cracks in Dissimilar Media," J. Appl. Mech., **32**, pp. 418–423.
[10] Clements, D. L., 1971, "A Crack Between Dissimilar Anisotropic Media," Int. J. Eng. Sci., **9**, pp. 257–265.
[11] Willis, R. J., 1971, "Fracture Mechanics of Interfacial Cracks," J. Mech. Phys. Solids, **19**, pp. 353–368.
[12] Suo, Z., and Hutchinson, J. W., 1990, "Interface Crack Between Two Elastic Layers," Int. J. Fract., **43**, pp. 1–18.
[13] Qu, J., and Li, Q. Q., 1991, "Interfacial Dislocation and Its Applications to Interface Cracks in Anisotropic Bimaterials," J. Elast., **26**, pp. 169–195.
[14] McMeeking, R. M., 1989, "Electrostrictive Forces Near Crack Like Flaws," J. Appl. Math. Phys., **40**, pp. 615–627.
[15] Kuo, C.-M., and Barnett, D. M., 1991, *Modern Theory of Anisotropic Elasticity and Applications* (SIAM Proceeding Series), J. J. Wu, T. C. T. Ting, and D. M. Barnett eds., SIAM, Philadelphia, PA, pp. 33–50.
[16] Suo, Z., Kuo, C.-M., and Willis, R. J., 1992, "Fracture Mechanics for Piezoelectric Ceramics," J. Mech. Phys. Solids, **40**, pp. 739–765.
[17] Sih, G. C., and Song, Z. F., 2003, "Magnetic and Electric Poling Effects Associated With Crack Growth in $BaTiO_3-CoFe_2O_4$ Composites," Theor. Appl. Fract. Mech., **39**, pp. 209–227.
[18] Song, Z. F., and Sih, G. C., 2003, "Crack Initiation Behavior in Magneto-Electro-Elastic Composite Under In-plane Deformation," Theor. Appl. Fract. Mech., **39**, pp. 189–207.
[19] Gao, C. F., Kessler, H., and Balke, H., 2003, "Crack Problems in Magneto-electroelastic Solids—Part I: Exact Solution of a Crack," Int. J. Eng. Sci., **41**, pp. 969–981.
[20] Gao, C. F., Tong, P., and Zhang, T.-Y., 2003, "Interfacial Crack Problems in Magneto-Electric Solids," Int. J. Eng. Sci., **41**(18), pp. 2105–2121.
[21] Stroh, A. N., 1958, "Dislocations and Cracks in Anisotropic Elasticity," Philos. Mag., **8**(3), pp. 625–646.

[22] Barnett, D. M., and Lothe, J., 1975, “Dislocations and Line Charges in Anisotropic Piezoelectric Insulators,” Phys. Status Solidi B, **67**, p. 105.
[23] Li, R., and Kardomateas, G. A., 2006, “Mode III Interface Crack of Piezo-Electro-Magneto-Elastic Dissimilar Bi-material Composites,” J. Appl. Mech., **73**(2), pp. 220–227.
[24] Lothe, J., and Barnett, D. M., 1976, “Integral Formalism for Surface Waves in Piezoelectric Crystals. Existence Considerations,” J. Appl. Phys., **47**(5), pp. 1799–1807.
[25] Barnett, D. M., and Lothe, J., 1973, “Synthesis of Sextic and the Integral Formalism for Dislocations, Greens Functions, and Surface Waves in Anisotropic Elastic Solids,” Phys. Norv., **7**(1), pp. 13–19.
[26] Li, R., and Kardomateas, G. A., 2005, “A Solution to the Thermoelastic Interface Delamination Branching Behavior for Dissimilar Anisotropic Bi-material Media,” Int. J. Solids Struct., in Press.
[27] Rudin, W., 1987, *Real and Complex Analysis*, McGraw-Hill, New York.
[28] Muskhelishvili, N. I., 1992, *Singular Integral Equations*, Dover, New York.

Quasi-Static Biaxial Plastic Buckling of Tubular Structures Used as an Energy Absorber

R. Baleh

A. Abdul-Latif[1]
e-mail: aabdul@iu2t.univ-paris8.fr

L3M,
Université Paris 8,
IUT de Tremblay,
93290 Tremblay-en-France,
France

The aim of this experimental study is to improve the energy absorption capacity of tubular metallic structures during their plastic buckling by increasing the strength properties of materials. Based on a novel idea, a change in the plastic strength of materials could be predictable through the loading path complexity concept. An original experimental device, which represents a patent issue, is developed. From a uniaxial loading, a biaxial (combined compression–torsion) loading path is generated by means of this device. Tests are carried out to investigate the biaxial plastic buckling behavior of several tubular structures made from copper, aluminum, and mild steel. The effects of the loading path complexity, the geometrical parameters of the structures, and loading rates (notably the tangential one) on the plastic flow mechanism, the mean collapse load, and the energy absorbed are carefully analyzed. The results related to the copper and aluminum metals show that the plastic strength properties of the tubes crushed biaxially change with the torsional component rate. This emphasizes that the energy absorption improves with increasing the applied loading complexity. However, the energy absorbed data for the mild steel tubular structures do not demonstrate the same sensitivity to the quasi-static loading path complexity. [DOI: 10.1115/1.2424470]

Keywords: plastic buckling, combined compression-torsion quasi-static loading, energy absorber

1 Introduction

It is well known that the most efficient energy absorber device should maintain the maximum and almost stable allowable force throughout the greatest stroke length. The problem of safe vehicle design with maximum impact energy absorption is an intensive research subject: e.g., vehicles crashworthiness analysis problem [e.g., Refs. [1–4]]. As a matter of fact, the energy absorbed during plastic deformation is one of the most significant factors dealing with the energy dissipating systems. Large plastic behavior of mechanical elements (plates, shells, tubes, stiffeners,..) when subjected to various types of load has been the subject of several research programs. They aim to understand the deformation modes, then the energy absorption patterns, and the resulting failure during collapse. For a given structure, this energy depends exclusively on main parameters: the magnitude, type, and method of application of loads, strain rates, deformation or displacement patterns, and material properties. Thus, the plastic flow has to be properly determined through experimental procedures to thoroughly understand the structure response during collapse.

In the recent decade, renewed interest has been also observed in the application of structures (as materials for energy absorber devices, like honeycombs, foams, etc.) [for example, Refs. [5,6]]. Beside the mass efficiency of such materials, their mechanical characteristics in compression demonstrate that they can be considered as an excellent energy absorber, offering an obvious plateau of almost constant force in the uniaxial compressive force–displacement curve [7–9]. In the light of this fact, the adopted solutions focus on the design of passive energy absorbing systems, which are frequently based on engineering materials always looking for a high-specific energy absorption ratio. Moreover, passive energy absorption devices are designed to successfully work in predefined collision scenarios.

According to such demands, the plastic buckling of tubular structures represents an appropriate compromise for a classical problem in solid mechanics. The literature shows that it provides an inexpensive and adaptable system [3,10,11]. In general, it offers one of the best devices in absorbing energy due to the stability of the average collapse load throughout the entire collapse process and due to the available stroke per unit mass. This is due to the fact that all of the tube material participates in the absorption of energy by plastic bending and stretching combination.

In a recent work [12], an experimental methodology has been developed where the plastic flow mechanism of the axial collapse of metallic hollow cylinders is controlled. Several tubular structures, made from copper and aluminum, are axially crushed under quasi-static compressive loading. As it is well known, the deformation mode of hollow cylinders is principally controlled by means of two geometrical parameters $\eta=R_m/t$ and $\lambda=R_m/L$ (R_m=mean radius, L=initial length, and t=thickness). From the energy absorption viewpoint, the axisymmetric mode provides enhancement with respect to the diamond fold mechanism for a given cylinder. Hence, the effects of both parameters on the mean collapse load and the energy absorbed have been studied. Thereby, two different structural solutions (fixed ends and subdivided) have been developed for encouraging the axisymmetric mode. With the subdivided solution, it has been found that the energy increases up to 21% compared to a classical uniaxial case for copper tubes.

Now, an important question becomes justifiable according to the following issue: How one can further increase the energy absorption for a tubular structure crushed axially having appropriate geometrical parameters, i.e., encouraging the axisymmetric mode?

Contrary to the standard passive systems, the proposed solution consists of creating a particular loading condition, which can be started during the deformation process. Actually, the philosophy

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received November 27, 2005; final manuscript received May 24, 2006. Review conducted by Yonggang Huang.

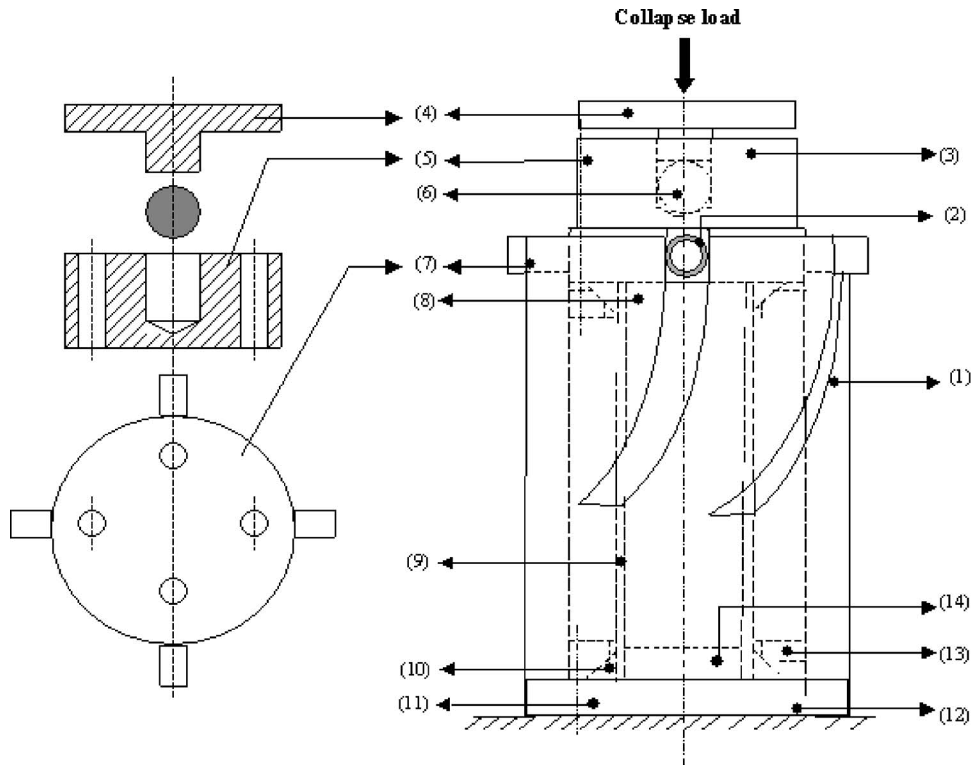


Fig. 1 Brief view of the ACTP device

of our step depends upon a novel idea, which aims to provoke a extra absorption of energy within a loaded structure via the loading path complexity notion.

Therefore, the intention of the present work is to offer further improvement in the energy absorption capacity of a tubular structure during deformation. Hence, a patented device, with which a combined compression–torsion biaxial loading can be generated over the tubes, is exploited. In fact, the strength properties of the loaded material change because of this loading type. Moreover, an investigation of interaction effects between the loading path complexity and the principal geometrical parameters is performed to interpret the obtained results. Thus, an extensive experimental study of plastic buckling is conducted using several structures made from copper, aluminum, and mild steel having different dimensions (i.e., different η and λ values). An integrity measure of the mean collapse load and consequently the corresponding energy absorbed show particularly the efficiency of the developed idea in improving the energy absorbed. As a result, the higher biaxial loading complexity, the greater the energy absorbed in copper and aluminum cases for a given structure. However, under the quasi-static loading condition, the mild steel behaves with a certain passivity vis-à-vis the loading path complexity.

It is important to emphasize that the logical continuation of this work is oriented toward a dynamic loading study. Considerable results are obtained showing some new remarkable results notably in the case of aluminum specimens [13,14].

2 General Scope

The main goal of this work is originated following the panoply of uniaxial solutions adopted, which made it possible to treat carefully the effect of the geometrical parameters on the plastic flow and the corresponding energy absorbed, therefore giving a maximum increase of 21% vis-à-vis the uniaxial buckling of copper structures [12].

In order to greater intensify the energy absorption capacity, it is now obvious that it is necessary to develop a new approach. Accordingly, an original device, referred to as "absorption par compression-torsion plastique" Patent No. WO 2005090822) (ACTP), is designed and tested. It is a simple mechanical assembly offering the advantage to function in quasi-static and dynamic modes. It transforms an external uniaxial compression load into a biaxial combined compression–torsion one (Fig. 1) within a loaded structure. It is particularly intended for industrial applications. Based on the same reference of crushing displacement for a given tube compressed uniaxially, the ACTP, as its fundamental role, allows a substantial enhancement in the energy absorption. In reality, a change in the strength properties of collapsed materials can be observed. In other words, this change is based undoubtedly on its work hardening evolution. The latter is mainly governed by the loading path complexity, which seems to be related to several (or one) mechanisms responsible for strength enhancement. Consequently, they play an important role in the local behavior, precisely, at the dislocation level probably generated by the slip systems multiplication phenomenon. In order to appropriately understand the effect of these key local mechanisms on the overall strength improvement of the structure, a research program is planned to be performed in the near future. This is due exclusively to the torsional component (generated by the ACTP), together with the compression controlled naturally by the external loading. In a precise manner, the plastic buckling becomes more and more complicated since three different strains (compressive, bending, and shear) take place concurrently, leading to a more complex load/unload condition.

Created by the proposed apparatus, two loading situations are considered. The first one involves a loading path, denoted as *integral loading*, where the compression and torsion components are applied simultaneously. Nevertheless, the second loading situation, referred to as *partial loading*, is characterized by the application of a purely uniaxial compression (from the beginning up to a chosen distance) followed by biaxial combined compression–torsion loading. To investigate the rate of change of the torsional component on the collapse operation, the device provides an opportunity to test several rates. The generation of such rates will be given in the next paragraph.

3 Description of the New Device

The ACTP device, illustrated schematically in Fig. 1, is constituted from a tempered steel hollow cylindrical body (1), on which four parallel helicoid grooves are machined. These grooves are characterized by a well defined inclination angle. The importance of the rate of change of torsional component in the collapse process within this device can be investigated. Therefore, three interchangeable cylindrical bodies: (1) are designed giving consequently three distinct propeller inclination angles of 30 deg, 37 deg, and 45 deg. The helicoid grooves are intended to receive a crosspiece (7) provided with four pivots and to guide it in its movement of descent by inculcating a rotational movement. Hence, the two principal parts of this apparatus form a slide–helicoid connection. This mechanism permits transformation of an initial external load of uniaxial nature into a biaxial combined compression–torsion one. In order to minimize the friction in the contact zone between the grooves and the crosspiece (7), the crosspiece pivots are equipped with bronze rollers (2).

Obviously, the crushed tube (9) is mutually dependent on the crosspiece and cylindrical body by the intermediary of a mechanical tube extremities fixation system. The system is made principally from two hard steel disks (11). Two half-conical shells (10) and a clip (13) over which these conical surfaces are machined and assembled in opposition attached to the disk (12), maintain the necessary tightening pressure in locking of both crushed tube extremities. Therefore, during its biaxial deformation, the specimen (9) is totally conditioned by the crosspiece in its movements of descent and rotation. Moreover, the assembling and the disassembling of the specimens within the ACTP are relatively simple.

With the ACTP, it is considerably difficult to evaluate the friction effect on the deformation operation during the biaxial tests. However, an approximate method is proposed to define its effect by means of the total absorbed energy. Actually, the following experimental methodology is adopted. As it is imposed by the ACTP design, a rolling friction type is used. Hence, six uniaxial crushing tests are realized using copper tubes of small initial lengths (giving $\lambda=0.6$). This value can be justified by the fact that it permits systematic generation of the axisymmetric deformation mode. These tubes are crashed under two different experimental conditions. In the first case, two uniaxial crushing tests of the free-ends situation are conducted. The average force–displacement of these tests is used to define the absorbed energy. The other four tests are carried out with the ACTP using the inclination angle of 45 deg (highest friction case), but with the existence of the crosspiece and without the tube extremities fixation system. Consequently, this leads to a classical free-ends uniaxial plastic buckling. Moreover, in order to limit the effect of friction on the contact surfaces between the crosspiece/specimen due to their relative rotation, an appropriate greasing is therefore used. Finally, the maximum recorded friction effect is about 6% as shown in Fig. 2.

4 Experimental Program

4.1 Tested Materials. In this study, three metallic materials are investigated, which are: commercial hardened copper (tensile yield stress: 310 MPa, Young's modulus: 117,000 MPa), annealed aluminum alloy (tensile yield stress: 150 MPa, Young's modulus: 70,000 MPa), and mild steel (tensile yield stress: 220 MPa, Young's modulus: 208,000 MPa), designated according to French standard as NFA51120 and AFNOR A506411, A50-451 (6060) and NFA 49 330 378 504523 NBK122, respectively. Note that each material has a good ductility.

The hollow cylindrical specimens employed have the following dimensions: two internal diameters (d) are chosen: 30 mm and 38 mm with 1 mm thickness (t), leading, respectively, to the following radial geometrical ratios ($\eta=R_m/t$): 15.5 and 19.5, where R_m is the mean tube radius. Characterized by three longitudinal ratios ($\lambda=R_m/L$): 0.1, 0.12, and 0.14, two initial tube lengths are defined for these metals.

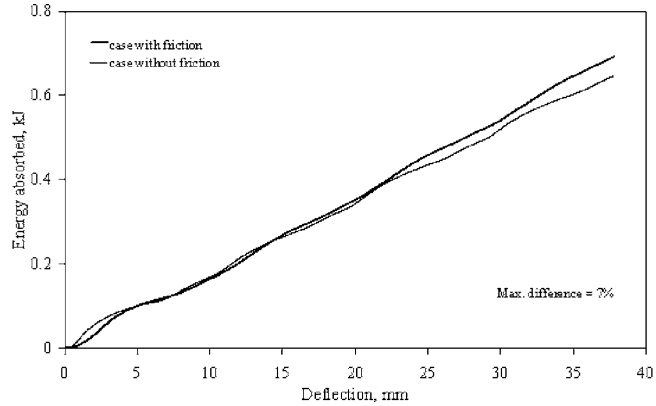


Fig. 2 Friction effect estimation within the ACTP device

All the specimens are crushed quasi-statically under compressive loading. They are not subjected either to heat treatment or to special machining operation. The defined lengths of the material cylinders used with their ratios are summed up in Table 1.

4.2 Experimental Procedure. All of the employed tubular structures are loaded between the two parallel platens of an Instron Universal Testing Machine (type 1186) under two compressive constant cross head speeds, namely, 5 mm/min and 500 mm/min at room temperature. Each employed compressive speed gives undoubtedly a quasi-static loading condition. The machine, where the ACTP is fixed, is connected to an acquisition chain to simultaneously record the force and the corresponding displacement during tube crushing operation.

In order to ensure the experimental results accuracy, each test is repeated twice under the same experimental conditions (applied speed and temperature). If the differences between the two responses exceed 3%, then another test should to be performed.

5 Results and Discussion

Three distinct deformation modes are regularly recorded, noted diamond mode (DM), axisymmetric mode (AM) and mixed mode (XM). Two deformation nuances are also obtained. Actually, the majority of the tests generates a mixed mode. The proportions of axisymmetric and diamond in the mixed mode differ from one material to another and/or from one rate of change of torsional component to another. Hence, one assigns the notation AXM to the mixed mode with axisymmetric predominance and the DXM to the mixed mode with diamond dominance. At the end of the crushing process for the three materials used, Figs. 3–5 collect typical deformed specimens loaded under two uniaxial loading situations (free ends and fixed ends: considered henceforth as biaxial-0 deg) and biaxial of the three inclination angles (30 deg, 37 deg, and 45 deg). Moreover, Tables 2–4 summarize the deformation modes and the corresponding mean collapse load (F_{av}), obtained under three different biaxial loading paths controlled essentially by the three inclination angles using the two cross head speeds (5 and 500 mm/min). In order to determine the mean collapse load with a maximum of objectivity, especially for these

Table 1 Chosen lengths of the material tubes with their geometrical ratios

Material used		Copper		Aluminum	Mild steel
Parameter η		15.5	19.5	15.5	19.5
Parameter λ	$L=136$ mm	0.11	0.14	0.11	0.14
	$L=160$ mm	0.1	0.12	0.1	0.12

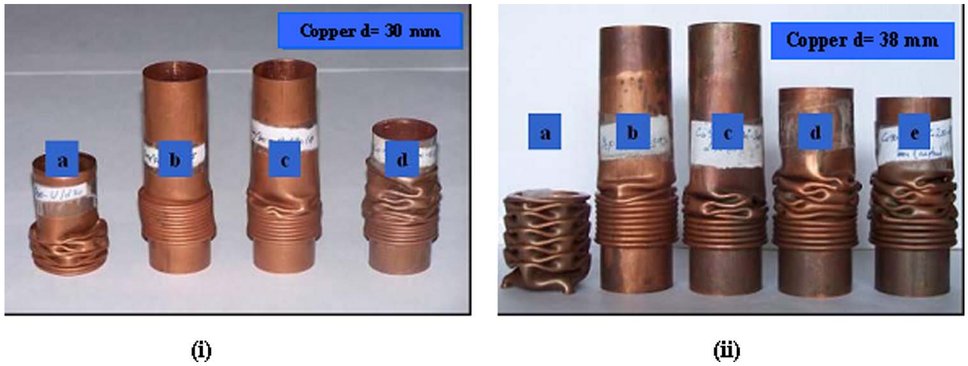


Fig. 3 Typical examples of the deformation modes related to the copper specimens having: (i) η=15.5 and (ii) η=19.5 with different crushing types: (*a*) free-ends (DM); (*b*) biaxial-0 deg (AM); (*c*) biaxial-30 deg (AXM); (*d*) biaxial-45 deg (XM); and (*e*) biaxial-37 deg (XM)

strain hardened metals, all recorded forces for the whole crushed distance are used to determine the mean collapse load.

The typical curves of the load and energy absorbed evolutions are presented in Figs. 8–11.

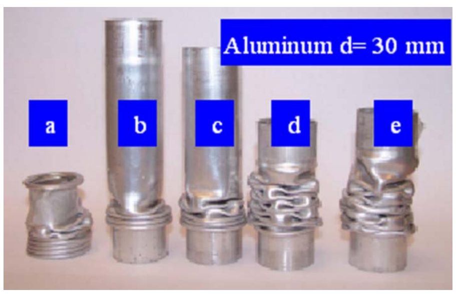


Fig. 4 Deformation modes concerning the aluminum specimens during their collapse under various loading paths: (*a*) free-ends (AXM); (*b*) biaxial-0 deg (AXM); (*c*) biaxial-30 deg (XM); (*d*) biaxial-37 deg (DM); and (*e*) biaxial-45 deg (DM)

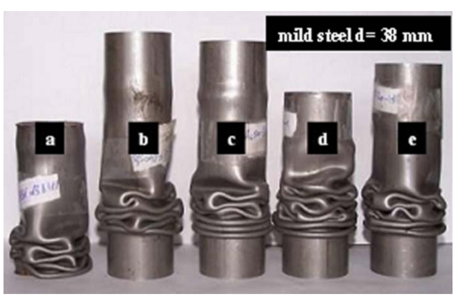


Fig. 5 Deformation modes of the collapsed mild steel specimens under various crushing types: (*a*) free-ends (DM); (*b*) biaxial-0 deg (DM); (*c*) biaxial-30 deg (XM); (*d*) biaxial-37 deg (DXM); and (*e*) biaxial-45 deg (DXM)

Tests are conducted under biaxial loading in order to validate the capacity of the ACTP in increasing the strength properties of the deformed material (i.e., improvement of the energy absorption) and to justify its efficiency. Indeed, examination of the recorded results exhibits a remarkable influence of the ACTP on the crushed structures behavior with respect to similar structures deformed uniaxially, i.e., enhancements in the F_{av}, F_{max} (peak collapse load).

5.1 Deformation Mechanisms. As far as the deformation mode is concerned, a rough similarity is pointed out between the deformation modes generated by the copper and aluminum specimens. In fact, with increasing in the loading path complexity, the AM and the AXM take place rather largely in the copper case (Fig. 3), whereas, for the aluminum, the XM and DM are obviously observed in Fig. 4.

Moreover, each test presents its own deformation mode for a given loading configuration. For the copper specimens (Table 2), the AM (or DM) is recorded in the uniaxial free-ends case and the AXM or XM under the biaxial loading (Fig. 3). The last types are characterized by a clear transition from the AXM toward XM, passing, respectively, from an inclination angle of 30 deg to 37 deg and finally to 45 deg. This observation remains also valid for the aluminum specimens (Table 3), i.e., the increase in the inclination angle supports the emergence of the diamond mode. The passage from 30 deg to 45 deg induces a deformation transition from the XM toward the DM (Fig. 4). Contrary to the uniaxial loading where the deformation modes are practically not affected by the loading rate, the deformation mode is strongly sensitive to the rates of change of the torsional component, especially for the inclination angle of 45 deg.

This observation is different in the case of mild steel (Fig. 5 and table 4), i.e., the increase in the inclination angle supports the emergence of the XM mode. The passage from 30 deg to 45 deg induces a deformation transition from the XM toward DXM, contrary to the uniaxial loading where the deformation modes are of

Table 2 Deformation modes and the mean collapse loads in kN for the copper tubes having different geometrical parameters under various loading situations using two cross head speeds

Copper (V, mm/min)		Free ends				Biaxial-0 deg				Biaxial-30 deg				Biaxial-37 deg				Biaxial-45 deg			
		5		500		5		500		5		500		5		500		5		500	
η=15.5	λ=0.11	DM	16	DM	16	AM	17.6	AM	17.6	AXM	19.2	AXM	19.2	XM	19.4	XM	19.4	XM	22	XM	22
	λ=0.1	DM	15.8	DM	15.8	AM	17	AM	17	AXM	18.6	AXM	18.6	AXM	18.9	XM	19	XM	21.6	XM	21.4
η=19.5	λ=0.14	DM	17	DM	17	AXM	17.8	AXM	17.8	AXM	18.8	AXM	18.8	AXM	19.4	XM	19.4	XM	22.4	XM	22.4

Table 3 Deformation modes and the mean collapse loads in kN for the (*a*) aluminum tubes and (*b*) mild steel tubes having different geometrical parameters under various loading situations using two cross head speeds

(*a*) Aluminum (*V*, mm/min)	Free ends 5	Free ends 500	Biaxial-0 deg 5	Biaxial-0 deg 500	Biaxial-30 deg 5	Biaxial-30 deg 500	Biaxial-37 deg 5	Biaxial-37 deg 500	Biaxial-45 deg 5	Biaxial-45 deg 500
$\eta=15.5$ $\lambda=0.11$	XM 10.5	XM 10.5	XM 10.8	XM 11.8	XM 11.6	XM 11.8	DM 12	DM 11.8	DM 12.2	DM 12.5
$\lambda=0.1$	DXM 10.4	DXM 10.5	XM 10.6	XM 11.7	XM 11.5	XM 11.5	DM 11.6	DM 11.6	DM 12	DM 12
(*b*) Mild steel (*V*, mm/min)	Free ends 5	Free ends 500	Biaxial-0 deg 5	Biaxial-0 deg 500	Biaxial-30 deg 5	Biaxial-30 deg 500	Biaxial-37 deg 5	Biaxial-37 deg 500	Biaxial-45 deg 5	Biaxial-45 deg 500
$\eta=19.5$ $\lambda=0.14$	XM 19	DXM 21.1	DX 20.4	DM 21.8	DXM 21.6	DXM 22	DM 22	DM 22	DM 22.1	DM 22.6
$\lambda=0.12$	XM 20.6	DXM 21.4	DM 20.3	DM 22	DXM 20.1	DXM 21.3	DM 21.7	DM 22.2	DM 21.5	DM 22.6

DM type. Hence, one can generally conclude, for a given material, that the deformation mode is significantly influenced by the rates of change of the torsional component.

The collapsed aluminum structures present three distinct deformation modes: AXM, XM, and DM (Fig. 4). It is important to keep in mind that the aluminum tubes are often a matter of macroscopic wall cracking, taking place at the plastic hinges. Such a cracking phenomenon does not take place either in copper or in mild steel specimens. This would be related to their intrinsic mechanical behaviors notably under biaxial loading.

In order to thoroughly study the deformation mode under biaxial loading condition, the biaxial-45 deg extreme case is considered as a typical example. In Fig. 6, some tube sections demonstrate the deformation progression of the copper tube (with $\eta=15.5$ and $\lambda=0.11$). In fact, the particular importance of the geometrical parameters (η and λ) is obvious in controlling the deformation mode at the beginning of the biaxial crushing process where the applied load axis coincides with that of the tube. This actually gives up to six folds of AM as shown in Fig. 6(*a*).

Thereafter, with the crushing progression, the torsional component becomes important and provokes a deviation with respect to this co-axiality, therefore leading to the appearance of the first fold of MD as clearly illustrated in Figs. 6(*a*) and 6(*b*). This is due to the fact that with the enhancement in the torsional component effect over the remaining nonbuckled length, the geometrical parameters, notably λ, cannot subsequently be the principal factors defining the deformation mode. Thus, a competition phenomenon occurs between these parameters and the torsional component effect. Obviously, the tangential disturbance becomes progressively more significant with collapse development. Undoubtedly, this leads to the DM up to the final collapse of the tubular structure (Fig. 6(*c*)).

Table 4 Energy absorbed under different biaxial loadings for two axial deflections of 30 and 60 mm for copper and aluminum structures.

Crashed distance / Loading path type	Aluminum Absorbed energy (kJ/kg) 30 mm	Aluminum Absorbed energy (kJ/kg) 60 mm	Copper Absorbed energy (kJ/kg) 30 mm	Copper Absorbed energy (kJ/kg) 60 mm
Biaxial-0 deg	20	19.5	41.7	40.7
Biaxial-30 deg	22.7	22	45.6	42.7
Biaxial-45 deg	26.2	25	47	45.8

5.2 Collapse Loads and Energy Absorbed. Let us now discuss the evolutions of the applied crushing load and the energy absorbed. First of all, it is important to study the loading rate influence on the material response. It is found that the responses of the tested materials are not affected by the employed loading rate range (Tables 2–4). As a typical example, load–deflection curves, in Fig. 7, reveal almost the same loading evolution during the copper tube collapse under the two speeds (5 and 500 mm/min) using the biaxial-37 deg (i.e., 37 deg inclination angle) case.

Figure 8 shows the crushing load evolution versus axial displacement in the case of the copper tubes having $\eta=15.5$, under a speed of 500 mm/min. Two loading situations are selected: the uniaxial free ends and the partial biaxial-30 deg. It is clear, in spite of the difference in these applied loading paths, that the

(a) (b) (c)

Fig. 6 Photos showing the deformation mode progression under the biaxial-45 deg case for the copper tubular structure having $\eta=15.5$ and $\lambda=0.11$

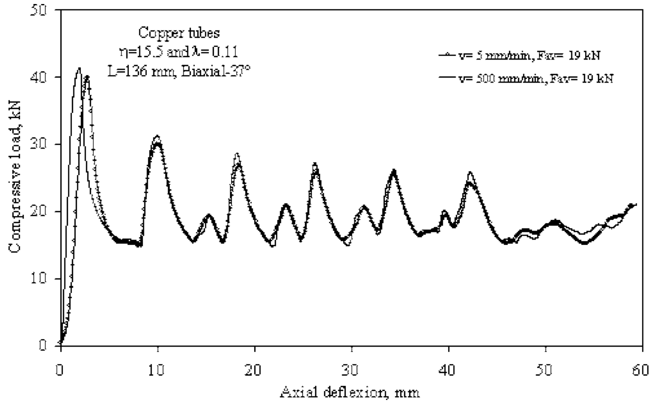
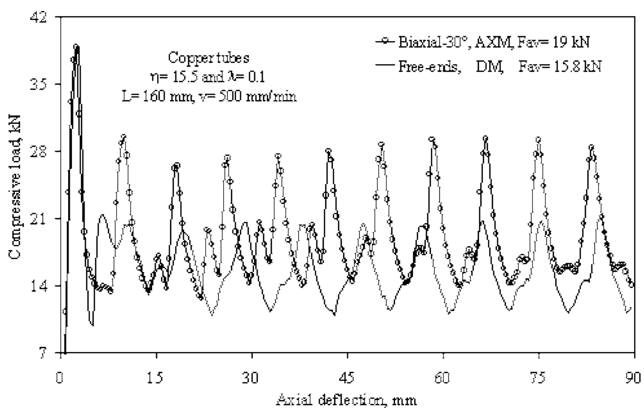


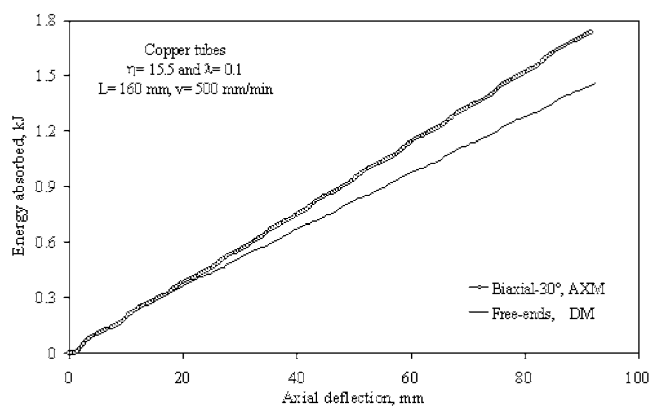
Fig. 7 Loading rate effect on the collapse loading evolution versus the axial deflection for the copper tubes in biaxial-37 deg case

nature of these curves is moreorless similar. An increase of 3.2 kN in the mean collapse load (F_{av}) in favor of the biaxial loading is recorded. Hence, this gives an enhancement of 20%.

However, the two peak collapse loads remain unchanged. This is due to the partial loading nature under which the torsional component value is equal to zero up to a certain axial deflection ($\delta = 15$ mm). Then, the biaxial combined compression–torsion loading starts to be applied. Under this loading complexity, an additional hardening can be provoked, notably when the material is severely loaded by three different strains: compressive, bending,

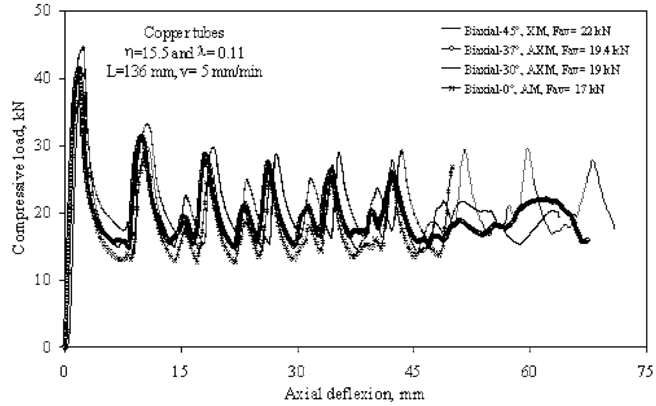


(*a*)

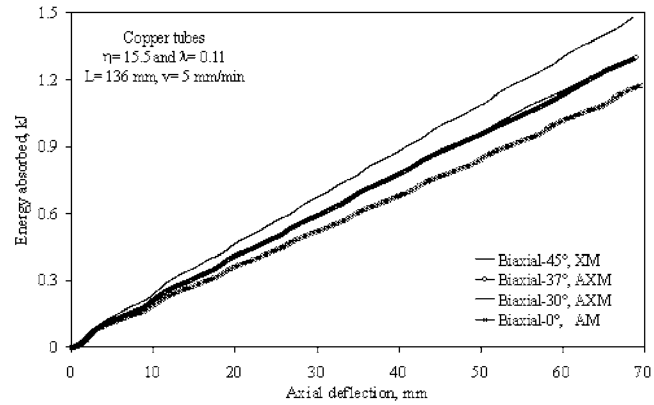


(*b*)

Fig. 8 Evolution of: (*a*) collapse load; and (*b*) energy absorbed versus the axial deflection for the copper tubes in uniaxial (free-ends) and biaxial-30 deg cases



(*a*)



(*b*)

Fig. 9 Evolution of: (*a*) collapse load; and (*b*) energy absorbed versus the axial deflection for the copper tubes (η=15.5) in biaxial (0 deg, 30 deg, 37 deg, and 45 deg) cases

and shear.

The deformation mode of the copper tubes evolves, consequently, passing from the DM under uniaxial loading to the AXM at biaxial loading (Fig. 3(*i*)). This observation, in addition to the increase in the F_{av}, can interpret the enhancement in the energy absorption for these tubes as shown in Fig. 8(*b*). As mentioned above, the absorbed energies under these loading conditions are identical during the first 15 mm; then these curves diverge when the combined compression–torsion loading is entering into action. Hence, the curve shows an additional absorption of energy in the biaxial-30 deg case. As a typical example, for δ=85 mm, the absorbed energies under uniaxial and biaxial-30 deg are, respectively, 1.35 kJ and 1.62 kJ, giving an enhancement of 20%.

For the copper tubes having η=15.5, Fig. 9(*a*) demonstrates the load–deflection curves for four different collapse situations: three integral biaxial loadings with inclination angles of 30 deg, 37 deg, and 45 deg, and the biaxial-0 deg employing a speed of 5 mm/min. It is recognized that the higher the inclination angle (i.e., the higher the loading complexity), the greater the rates of change of torsional component, the greater the mean collapse load and the corresponding energy consequently absorbed. As a result, the biaxial-45 deg gives the highest F_{av} and energy absorbed which has also been confirmed in Ref. [15].

This finding is systematically raised even for the mode DXM. In reality, the latter is characterized by the emergence of a strong proportion of DM, especially at the end of collapse. Note that the F_{av} value is 21.6 kN for the biaxial-45 deg, whereas it is equal to 17 kN for the biaxial-0 deg. Thus, a difference of 4.6 kN is recorded between these two cases giving an increase of 27%. As noted above, the curves in Fig. 9(*b*) reveal clearly that the biaxial-

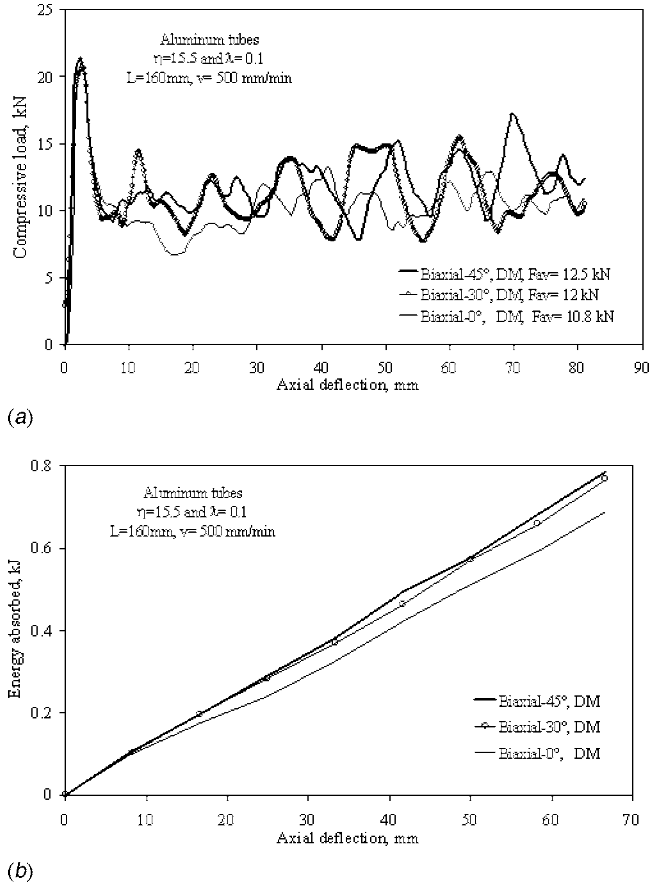


Fig. 10 Comparison of the (*a*) load and (*b*) energy absorbed evolutions during crushing process: under biaxial loadings (0 deg, 30 deg, and 45 deg) for the aluminum tubes

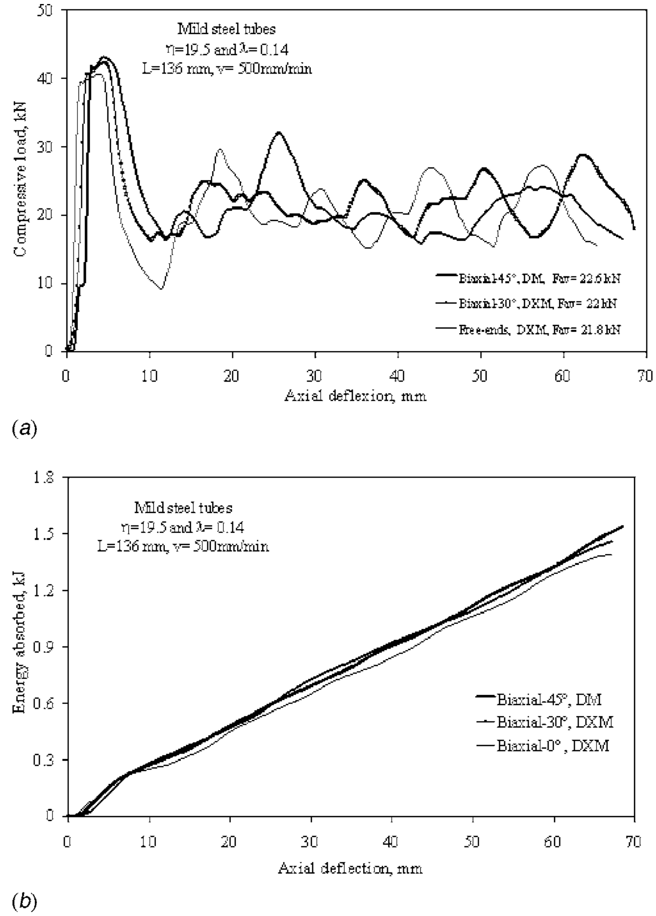


Fig. 11 Plots of the (*a*) load and (*b*) energy absorbed evolutions during crushing process: under biaxial loadings (0 deg, 30 deg, and 45 deg) for the mild steel tubes

45 deg case absorbs the greatest amount of energy. However, the intermediate curves for the biaxial-30 deg and 37 deg are relatively close, which is confirmed by the equivalence of their respective F_{av}, 19.2 kN and 19.4 kN. In fact, this result can be interpreted by a competition phenomenon between the change of the deformation mode and the change in the material strength via the loading path complexity. The traditional uniaxial crushing of reference always presents the weakest energy absorbed.

The biaxial buckling of the aluminum structures activates simultaneously two contradictory phenomena: a deformation mode change and a variation in the material strength properties. Indeed, the load–deflection curves of the aluminum specimens are recorded in Fig. 10(*a*) showing a comparison between integral biaxial-30 deg and 45 deg and 0 deg. The effect of the torsional component on the peak collapse loads (F_{max}) is not evident. One can note, however, a light advantage is observed for the biaxial-45 deg, i.e., F_{av} are 10.9 kN, 12 kN, and 12.5 kN for biaxial-0 deg, 30 deg, and 45 deg loadings, respectively. This means that the work hardening is moderately changed under biaxial quasi-static loading. For the energy absorbed, Fig. 10(*b*) illustrates the same form as in previous figures. In fact, for an axial deflection of 60 mm, the energy absorbed values are: 0.61, 0.68 kJ, and 0.7 kJ for biaxial-0 deg, 30 deg, and 45 deg, respectively, giving corresponding enhancements of 11% and 15% in favor of biaxial-30 deg and 45 deg. However, a radical change in the strength properties of this material is obviously recorded under dynamic loading conditions [14]. This demonstrates the real advantage functions of the ACTP (more than 150% increase in the energy absorbed in favor of biaxial-45 deg).

The behavior of the mild steel structures is also presented via the load–deflection curves (Fig. 11(*a*)). They have approximately atypical evolutions concerning the peak collapse loads under the different biaxial integral loading paths of 0 deg, 30 deg, and 45 deg. Indeed, the initial plasticity peak is rather replaced by a yield plateau (totally related to this well-known material behavior). This evolution translates a certain sensitivity related to "the elastic limit" of this material with respect to the torsion rate of change. Compared to copper, the F_{av} of mild steel evolves slightly with respect to the applied loading complexity, as pointed out in Fig. 11(*a*). Figure 11(*b*), illustrating the energy absorbed, reveals that the absorbed energies under the three loading paths are practically the same.

It is intriguing to note the parameters η and λ and their interaction play a primordial role in the deformation mode for a given structure. As mentioned above, the plastic flow is also conditioned by the coincidence of the applied load axis with that of the structure. However, under biaxial loading, the torsional component provokes after a certain axial deflection, a violation of this coaxiality. The experimental results confirmed that the two parameters, η and λ, control principally the deformation type. Progressively, a kind of competition phenomenon between the geometrical parameters (η and λ) and the torsional component effect takes place to determine the deformation mode. Obviously, the tangential disturbance becomes progressively more significant with collapse progression. This means that the higher the inclination angle, the higher the tangential disturbance. This demonstrates that the biaxial-45 deg is the most severe case. Consequently, the biaxial-45 deg loading often gives the DM (or DXM), as recorded in the aluminum structures (Fig. 4), whereas, with the biaxial-30 deg situation, it is rather the XM. In spite of this difference, the biaxial-45 deg case shows an energy absorption capacity more

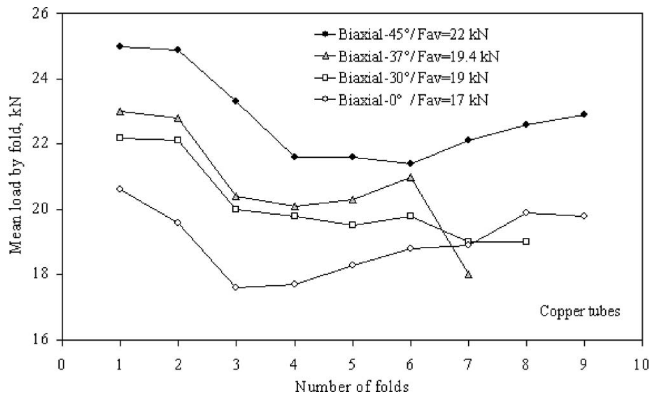


Fig. 12 Variation of the mean load for each fold occurring during copper tubes collapse under different biaxial (0 deg, 30 deg, 37 deg, and 45 deg) loadings

important than the biaxial-30 deg. This is certainly due to the strong competition between the two phenomena. It resides between the reduction in the energy absorption capacity, induced by the emergence and often the predominance of DM, and the enhancement in the material strength due to the biaxial-45 deg loading complexity, as it is clearly shown in Tables 2 and 3.

Let us now examine deeply the effect of the torsional component and its complexity on the crushing load evolution for each fold during collapse process. Extracted from the data in Fig. 9, Fig. 12 demonstrates clearly that the mean value of each fold is noticeably affected by the torsional component, i.e., the higher mean value is recorded for the higher level of loading complexity. For moderate loading complexity (biaxial-30 deg and 37 deg), the effect of the deformation mode becomes, in general more important than the torsional component, practically at the end of collapse; and a contrary result is found in the biaxial-45 deg case, where loading complexity is the most severe.

As far as the variation of the energy absorbed per unit weight is concerned, it is defined for the three loading paths (biaxial-0 deg, 30 deg, and 45 deg) for the copper and aluminum tubes using two selected axial deflections of 30 nm and 60 mm. As shown previously, the biaxial-45 deg loading has the best energy absorption capacity. It is significant to note that the choice of these two axial deflections is not arbitrary. In fact, since the AM absorbs more energy that the DM, it is easily demonstrated through Table 4 that for $\delta=30$ mm, this energy is always more important compared to $\delta=60$ mm for the two metals. This is due to the fact that after the first chosen axial deflection, a change in deformation mode (from AM or AXM to DM or DXM) generally takes place, therefore giving a decrease in the energy absorbed for the two metals (Table 4).

6 Closure

Based on the importance of the geometrical parameters of tubular structures (η and λ), recently studied [12] in controlling the energy absorbed via the deformation mode, the results show that the maximum gain of energy does not exceed 21% with respect to the classical uniaxial crushing case.

In order to further increase the energy absorption for tubular structures crushed axially, an original idea is thus explored based on the material strength modification concept during deformation. In fact, such an idea consists, a priori, via the loading path complexity, to induce local physical phenomena responsible for the change in strength properties (enhancement in the work hardening), consequently giving more energy absorbed. This step has led to the development of a new patented mechanical assembly, the ACTP. It generates, through a uniaxial external compression, a biaxial combined compression–torsion loading within the loaded structure. The effects of this induced torsional component and its rate on the absorbed energy are therefore carried out using three inclination angles (30 deg, 37 deg, and 45 deg). In the case of a biaxial situation, two loading situations are studied: the *partial and integral* biaxial loadings. The principal conclusions are as follows:

1. In the biaxial loading, the influence of the rate of change of the torsional component on the plastic flow and consequently on energy absorbed, can be translated by a maximum increase of 35% for the copper tubes compared to the uniaxial free-ends case;
2. In biaxial buckling, three principal parameters (η, and λ, and the torsional component) obviously control the plastic flow, for a given axial loading rate and material. This therefore leads to the fact that the loaded structure undergoes an extremely complex loading path of three different strains (compressive, bending, and shear) taking place simultaneously;
3. The change in mechanical behavior of the deformed material is provoked by local mechanism modifications;
4. Contrary to the mild steel tubular structures, the copper and aluminum specimens are obviously affected by the loading path complexity; and
5. The inexistence of a significant influence of the partial biaxial loading in comparison to the integral biaxial from the energy absorbed viewpoint.

Hence, it is generally recognized now that under biaxial loading, the higher the inclination angle, the greater the loading complexity, the higher the rate of change of torsional component, the greater consequently, are the mean collapse load and the corresponding absorbed energy, i.e., the biaxial-45 deg is considered as the most significant case giving the highest F_{av} as well as the energy absorbed.

References

[1] Ezra, A. A., and Fay, R. J., 1972, “An Assessment of Energy Absorbing Devices for Prospective Use in Aircraft Impact Situations,” *Dynamic Response of Structures*, G. Hermann and N. Perrone, eds., Pergamon, New York, p. 225.
[2] Abramowicz, W., and Jones, N., 1986, “Dynamic Progressive Buckling of Circular and Square Tubes,” Int. J. Impact Eng., **4**, p. 243.
[3] Johnson, W., and Reid, S. R., 1986, “Update to Metallic Energy Dissipating Systems,” Appl. Mech. Rev., **31**, pp. 277–288.
[4] Al-Ghamdi, A. A. A., 2001, “Collapsible Impact Energy Absorbers: An Overview,” Thin-Walled Struct., **39**, p. 189.
[5] Bastawros, A. F., Bart-Smith, H., and Evans, A. G., 2000, “Experimental Analysis of Deformation Mechanisms in a Closed-Cell Aluminum Alloy Foam,” J. Mech. Phys. Solids, **48**, p. 301.
[6] Tan, P. J., Reid, S. R., Harrigan, J. J., Zou, Z., and Li, S., 2005, “Dynamic Compressive Strength Properties of Aluminum Foams: Part I—Experimental Data and Observations,” J. Mech. Phys. Solids, **53**, p. 2174.
[7] Hanssen, A. G., Langseth, M., and Hopperstad, O. S., 2000, “Static and Dynamic Crushing of Circular Aluminum Extrusions With Aluminum Foam Filler,” Int. J. Impact Eng., **24**, p. 457.
[8] Harte, A. M., Norman, A. F., and Ashby, M. F., 2000, “Energy Absorption of Foam-Filled Circular Tubes With Braided Composites Walls,” Eur. J. Mech. A/Solids, **19**, p. 31.
[9] Lopatnikov, S. L., Gama, B. A., Haque, M. D. J., Krauthauser, C., and Gillespie, J. W., Jr., 2004, “High-Velocity Plate Impact of Metal Foams,” Int. J. Impact Eng., **30**, p. 421.
[10] Reid, S. R., 1993, “Plastic Deformation Mechanisms in Axially Compressed Metal Tubes Used as Impact Energy Absorbers,” Int. J. Mech. Sci., **35**, p. 1035.
[11] Jones, N., 1998, “Some Recent Developments and Future Trends in Thin-Walled Sections for Structural Crashworthiness,” Thin-Walled Struct., **32**, p. 231.
[12] Abdul-Latif, A., Baleh, R., and Aboura, Z., 2006, “Effect of Plastic Flow Mechanisms on the Large Deformation of Hallow Tubes Loaded Axially,” Int. J. Solids Struct., **43**, p. 1543.
[13] Baleh, R., 2004, “Flambage Plastique Quasi-Statique et Dynamique de Structures Tubulaires Métalliques Sous Sollicitations Simple et Complexe-Système d’Absorption d’Énergie-via un Nouveau Dispositif Expérimental,” thèse de doctorat, Université de Technologie de Troyes, France.
[14] Abdul-Latif, A., and Baleh, R., 2006, “Biaxial Plastic Buckling of Circular Shells Under Dynamic Loadings,” J. Appl. Mech., submitted.
[15] Baleh, R., and Abdul-Latif, 2006, “On the Biaxial Plastic Buckling with Large Torsional Rate of Change,” L3M, Université Paris 8, Rep. No. 012.

Z. Wei[1]
e-mail: zhensong@engineering.ucsb.edu

M. Y. He

A. G. Evans

Materials Department,
University of California,
Santa Barbara, CA 93106

Application of a Dynamic Constitutive Law to Multilayer Metallic Sandwich Panels Subject to Impulsive Loads

The present paper describes an investigation that implements and assesses a dynamic continuum constitutive law for all-metallic sandwich panels. It also demonstrates its application to multilayer panels subject to water blast. Finite element calculations of unit cells are used to calibrate the model, especially the hardening curves at different strain rates. Once calibrated, the law is assessed by comparison with two sets of experiments. The dynamic response of panels impacted by Al foam projectiles at impulses comparable to those expected in water blast. The response of a multilayer core to an impulse caused by an explosion occurring in a cylindrical water column. The comparisons reveal that the overall deformation, average core strain, peak transmitted pressure, and velocities of the front and back faces are adequately predicted, inclusive of fluid/structure interactions. The inherent limitations of the approach are the underprediction of the plastic strains in the faces and incomplete assessment of stress oscillations beyond the peak. The former deficiency would pertain for any continuum representation for the core and would lead to problems in the prediction of face tearing. The latter may adversely affect the predictions of the impulse. [DOI: 10.1115/1.2424471]

1 Introduction

Metallic sandwich structures subject to air and water blast exhibit several response mechanisms, characterized by differing trends in displacement and plastic strain [1–3]. For example, in water blast, because of fluid/structure interaction (FSI) effects, there are four possible responses: soft, strong, slap, and zero back face deflection [2,3]. When amenable to design in accordance with the "soft" mechanism, the panels exhibit the best performance, judged using metrics based on: center deflections, tearing susceptibility of the faces and the reaction forces at the supports [1–3]. However, a transition to slapping (Fig. 1) would not only eliminate the benefits, but actually degrade the response. Accordingly, the pursuit of preferred core topologies must be conducted with these transitions clearly understood. The core options include single and multilayer configurations. Multilayers have the potential advantage that successive regions of the core can have properties tailored to satisfy differing requirements imposed by the dynamics. An example of a multilayer pyramidal core panel is depicted in Fig. 2. To establish the requisite understanding, measurements conducted in a test system that includes FSI effects are needed, in combination with simulations. Relevant measurements include those conducted in the Dynocrusher facility [4]. To conduct simulations, it is not realistic to fully mesh all of the core members (Fig. 2). Instead, each layer of the core must be homogenized and an adequate constitutive law established. More generally, homogenization is also needed for large-scale (ship-level) simulations, even with unilayer cores.

With this background, the first objective of the present assessment is to choose a continuum constitutive law for the individual layers of the core, implement within ABAQUS Explicit [5] and demonstrate its application. The second objective is to examine the capability of the code and establish its limitations by comparing results calculated using this representation with two sets of existing experimental measurements [4,6] and with full three-dimensional (3D) simulations [1]. One set of experiments involves the dynamic response of panels impacted by Al foam projectiles at impulses comparable to those expected in water blast [6] (Fig. 3). These experiments incorporate the dynamic responses of the core, but do not address FSI effects. The second set of experiments examines the response of a multilayer core to an impulse imparted by an explosion occurring in a cylindrical water column [4] (Fig. 4). This measurement system includes the contribution from the fluid/structure interaction. The final part of the investigation involves the application of the code to explore the potential benefits of multilayer designs subject to water blast.

2 Constitutive Law

2.1 Background. Several continuum approaches have been developed to characterize core structures [7–12]. One has been developed for isotropic cores, applicable to metal foams [7], but not to truss or prismatic cores. Another is anisotropic [8], but limited to rate independent materials. It invokes a conical yield surface in shear/normal stress space, along with a nonassociated flow rule. Two approaches have been devised for rate-dependent, anisotropic cores [9,10]. Both incorporate strain and strain-rate hardening as well as strain softening. One uses a quasi-continuum approach [9], bypassing the introduction of a continuum constitutive law. It has the unfavorable feature that face/core interactions are ignored. Another invokes a constitutive law representative of the elastic–plastic response of orthotropic compressible materials [12] with strain-rate dependence [10,11]. The latter has been implemented in ABAQUS/Explicit through a user material subroutine applicable to all core topologies [10,11]. This implementation provides a level of convenience more suitable to the objectives of this investigation.

In the model, an ellipsoidal yield surface is invoked that generalizes Hill's surface for orthotropic plastically incompressible materials [13]. The ellipticity of the surface is allowed to change to account for differential hardening or softening upon stressing in different directions at different strain rates. Associated plastic flow assures that the plastic strain-rates are normal to the yield surface.

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received December 10, 2005; final manuscript received April 19, 2006. Review conducted by Horacio D. Espinosa.

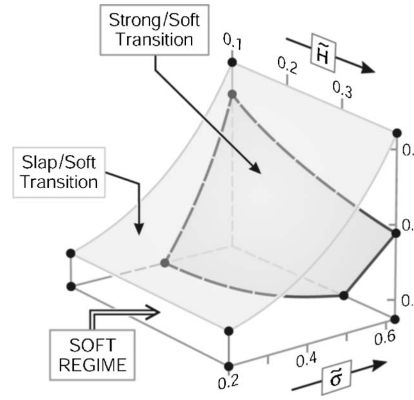


Fig. 1 A mechanism map indicating three possible responses of sandwich panels subject to impulsive loads. A fourth possibility, zero back face deflection, is small and not shown. The coordinates $\tilde{I}$, $\tilde{H}$, and $\tilde{\sigma}$ are the normalized impulse, core height, and strength, respectively [2,3]. They are defined as, $\tilde{I}=(I/M)\sqrt{\rho/\sigma_Y}$, $\tilde{H}$-H/L, and $\tilde{\sigma}=\sigma_{YD}^c/\bar{\rho}\sigma_Y$, where M is the mass per unit area of the sandwich panel, L its half span, ρ the density of the solid material, σ_Y its yield strength, and $\bar{\rho}$ the relative density of the core. The best response to water blast is found in the soft domain close to the transition to slapping.

The general form of the equations allows for strains in all orientations [11]. But, they can be simplified when it is reasonable to assume that the cores undergo relatively little transverse plastic strain upon uniaxial stressing. Then, the plastic strain-rate ratios can be equated to zero: whereupon, the ellipsoidal yield surface can be written in the form [10]

$$f \equiv \sigma_{\text{eff}} - \sigma_0 = 0 \tag{1a}$$

where the effective stress σ_{eff} is defined by

$$\left(\frac{\sigma_{\text{eff}}}{\sigma_0}\right)^2 = \left(\frac{\sigma_{11}}{\hat{\sigma}_{11}}\right)^2 + \left(\frac{\sigma_{22}}{\hat{\sigma}_{22}}\right)^2 + \left(\frac{\sigma_{33}}{\hat{\sigma}_{33}}\right)^2 + \left(\frac{\sigma_{12}}{\hat{\sigma}_{12}}\right)^2 + \left(\frac{\sigma_{13}}{\hat{\sigma}_{13}}\right)^2 + \left(\frac{\sigma_{23}}{\hat{\sigma}_{23}}\right)^2 = \sum_i \left(\frac{\sigma_i}{\hat{\sigma}_i}\right)^2 \tag{1b}$$

The quantity σ_0 is a fixed reference stress that can be chosen arbitrarily; it is simply a scaling factor. Six hardening functions $\hat{\sigma}_i$ are the basic inputs to the model. When σ_i is the only non-zero component, $\hat{\sigma}_i(\varepsilon_i^P, \dot{\varepsilon}_i^P)$ denotes the hardening (or softening) function, specifying the dependence of σ_i on the associated plastic strain component, ε_i^P. Independent hardening is used as the simplest option. That is, under multiaxial stressing, each of the six hardening functions, $\hat{\sigma}_i$, is affected only by the plastic strain component, ε_i^P, and its rate. Details can be found elsewhere [10,11].

The method to be used is based on that previously demonstrated for quasi-static situations [12,14,15]. In a first step, the input stress/strain curves for the core members are obtained using unit cell calculations, conducted in the important straining orientations. For implementation purposes, some simplifications are required, requiring judgments about the features to be retained and those to be discarded. The second step is to implement the constitutive law and reproduce the unit cell calculations to assure that the simplifications do not give rise to significant discrepancies. The third (and most important) step is to duplicate experiments on heterogeneously loaded sandwich panels containing these cores. Various truss, honeycomb and prismatic core topologies have been addressed [12,14,15]. The comparisons have revealed that most aspects of the load/deflection response are adequately duplicated. Discrepancies arise when multiple buckling modes with strong imperfection sensitivity dominate the behavior [14]. Note that, to establish consistency, realistic manufacturing imperfections needed to be incorporated into the unit cell calculations.

2.2 Dynamic Calibration. The objective is to find a procedure for determining the dynamic stress/strain curves for the core members in a manner that excludes the inertia [10]. Namely, since ABAQUS Explicit already includes the inertia of each element [5], the input stress/strain curves must include the material strain-rate sensitivity and buckling stabilization, but exclude the inertia. When the members yield before elastic buckling, as in the present circumstance, the following procedure is adopted. Unit cells (Fig. 5) are subject to a velocity profile, $\nu(H)$, on the top face corresponding to a constant effective strain rate in the core, $\dot{\varepsilon}_{\text{eff}} \equiv \nu/H$, where ν is the imposed velocity and H the current core height. The stress induced on the back face, which has no inertial component, is calculated. While this response does not reproduce the dynamic stress/strain behavior of the core member, the rapid drop in the transmitted stress [10] (see Fig. 6) does give the time at which the core starts to buckle plastically, and the corresponding strain, ε_{pb}. Once this strain has been determined, the input stress/strain response is ascertained in the following manner (Fig. 7). The members are considered to have the strain-rate dependent stress/strain characteristics of the constituent material at strains below ε_{pb}. At this strain, the material begins to soften with slope, $-4E_T$ (the choice is not critical), where E_T is the tangent modulus. Thereafter, the stress drops to a lower plateau level, σ_{pl}, dictated by the stress for dynamic crushing of the prebuckled core members.

Calibration curves have been obtained by introducing initial imperfections with the shape of the first buckling mode (Fig. 5). Prior assessments [14] have indicated that experimental measurements can be duplicated by allowing the imperfections to have amplitude, ξ=0.01 (based on the member length). Most calculations have been performed using ξ=0.01. But to explain some effects observed experimentally in the Dynocrusher tests [4], a few calculations are performed using ξ=0.02. Typical results are plotted in Fig. 6 using the true strain as the abscissa. A comparable set of calibration curves has been obtained for square honeycombs, similar to those presented elsewhere [10], but specialized to the core dimensions used in projectile impact experiments [6]. The calculations reveal the following features.

Initial yield (position a in Fig. 6) occurs at stress

$$\sigma_{33} = \sigma_Y' \bar{\rho} \quad \text{(square honeycomb)} \tag{2a}$$

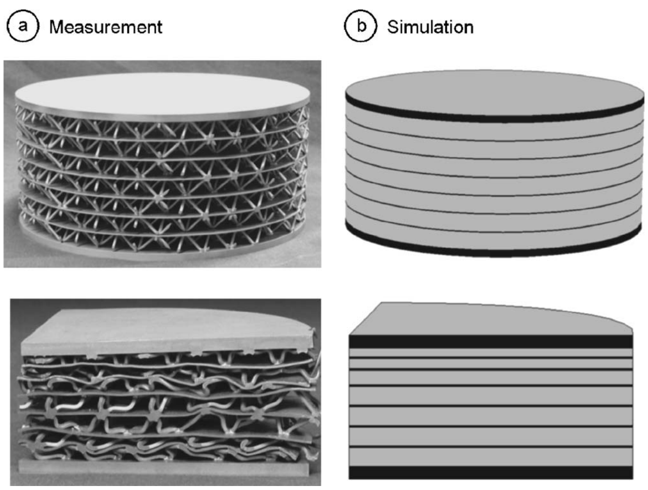


Fig. 2 Images of the seven-layer truss panel taken before and after testing. Also shown is the numerical model with homogenized cores before and after testing.

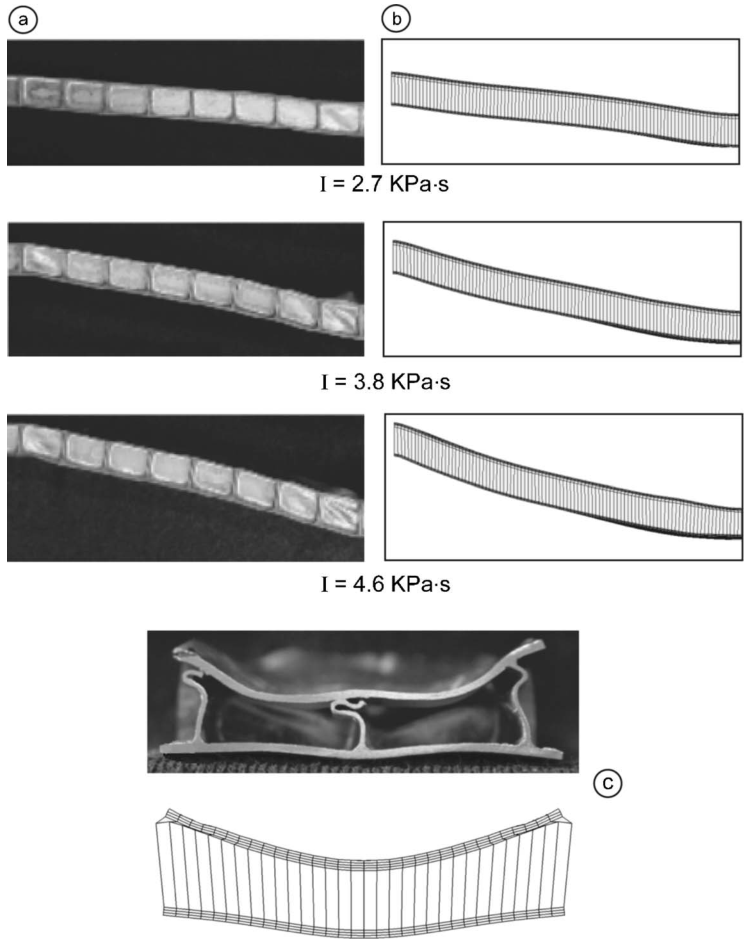


Fig. 3 (*a*) Images of the stainless steel sandwich panels with square honeycomb cores after impact by foam projectiles performed at various values of the nominal impulse [6]. (*b*) The corresponding simulations conducted using the dynamic constitutive law for the core. (*c*) The deformation of the core near the center of the panel shown in (*a*), impacted at the highest impulse. A comparison is made between the measured shape and that obtained using the dynamic continuum law using mesh scheme (iii).

$$\sigma_{33} = 0.5\sigma_Y' \bar{\rho} \quad \text{(pyramidal truss)} \tag{2b}$$

where σ_Y' is the yield strength of the constituent material at the imposed strain rate and $\bar{\rho}$ is the relative density. The members begin to buckle soon after yielding (position b in Fig. 6), but still strain harden. The hardening rate is

$$E_T^{\text{core}} \approx \bar{\rho} E_T \quad \text{(square honeycomb)} \tag{3a}$$

$$E_T^{\text{truss}} \approx 0.25\bar{\rho} E_T \quad \text{(pyramidal truss)} \tag{3b}$$

The strain hardening continues up to a strain ε_{pb} (location d in Fig. 6), which varies with strain rate as (Fig. 7(c))

$$\varepsilon_{\text{pb}} \approx 10\dot{\varepsilon}_{\text{eff}} H_c / c_{\text{el}} \quad \text{(square honeycomb and pyramidal truss)} \tag{4}$$

where H_c is the core thickness and c_{el} the elastic wave speed in the core member. For the pyramidal truss, softening beyond ε_{pb} is succeeded by a minimum and a second peak (position $e \rightarrow f$ in Fig. 6), which occurs when the top face contacts the buckled core members. At strains beyond ε_{pb} the stress oscillates about an *average*, σ_{pl}, given by

$$\sigma_{\text{pl}} / \sigma_Y \bar{\rho} \approx 0.035\sqrt{\dot{\varepsilon}_{\text{eff}} / \dot{\varepsilon}_0} \tag{5}$$

with $\dot{\varepsilon}_0 = 1/s$. For the square honeycomb, it has been assumed that $\sigma_{\text{pl}} / \sigma_Y \bar{\rho} \approx 1$. At strain ε_D, densification commences.

2.3 Input Curves. To perform meaningful calculations, dynamic stress/strain curves are needed that encapsulate these features ($a \rightarrow d$) in the most straightforward manner. The curves chosen for this purpose have the features plotted in Fig. 7. Before proceeding, the consequences of the simplification should be assessed. For this purpose, the unit cell calculations are duplicated by using the input curves. Some typical results are presented in Fig. 8. The results reveal that the homogenized model is stable and reproduces most of the important stress/strain features. However, for the pyramidal truss, it does not capture the oscillations associated with contact between the top face and the core. It remains to establish whether exclusion of this aspect of the core crushing is significant.

2.4 Multilayer Cores. An alternative procedure is preferred for multilayer structures with thin interlayer sheets in which interactions between layers strongly effects the response. To demonstrate this approach, a multilayer unit cell has been constructed

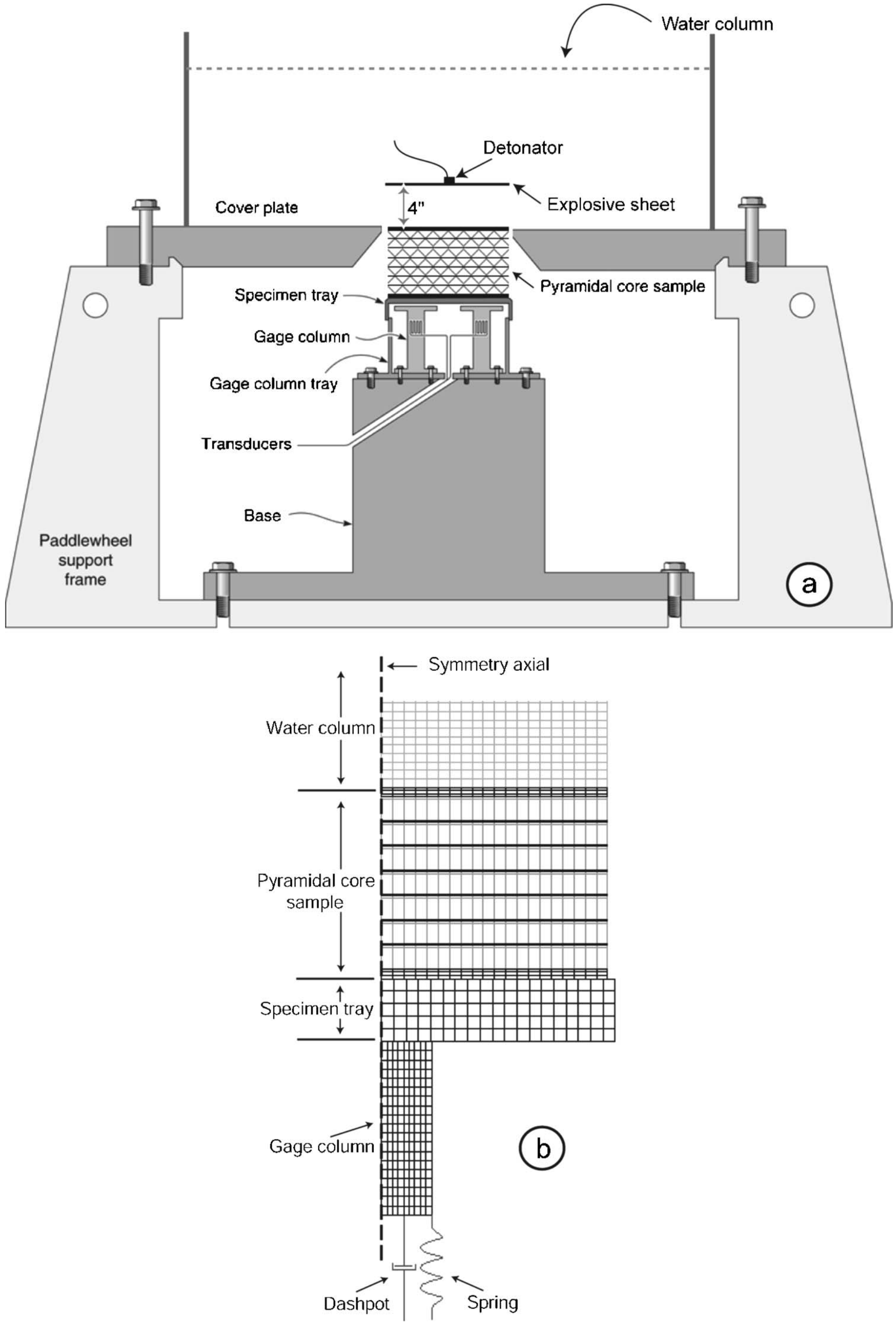


Fig. 4 (*a*) A schematic of the Dynocrusher test arrangement [4]. (*b*) The axisymmetric continuum finite element model built up for the test. Also shown are representative meshes.

(Fig. 9), comprising a one-unit column of the seven-layer truss core. Symmetry boundary conditions are imposed, and $\xi=0.01$ imperfections introduced into every truss member. Typical results plotted in Fig. 9 indicate that, in this case, the interactions between layers does not change either the deformation mechanism or the buckling mode, so that the unilayer calibration may still be used. By using the same imperfections in all layers, it will be shown that one of the experimental observations in the Dynocrusher test [4] (Fig. 3(*c*)) cannot be duplicated: namely, that the lowest layer exhibits more extensive crushing than the adjacent layers. To address this discrepancy, some calculations are performed by incorporating larger imperfections (magnitude, $\xi=0.02$) in those layers adjoined to each of the faces.

2.5 Meshing. Three meshing schemes have been explored [10,11,16], differentiated by the number of elements through the thickness of the core: (i) multiple elements; (ii) a single element; and (iii) a thin element next to the outer face plus an elongated element elsewhere. Scheme (i) has the disadvantage that it is not computationally efficient for large structures. Scheme (ii) is most efficient and computationally stable with complex hardening curves [11]. In this scheme, a half of the core mass is added to

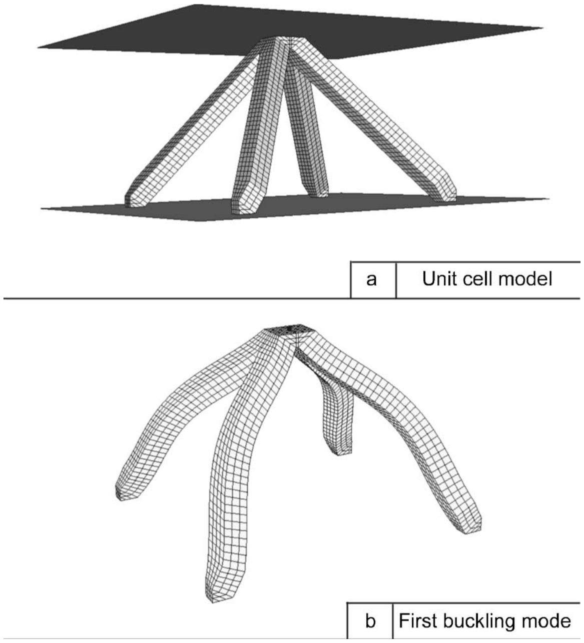


Fig. 5 (*a*) The unit cell for the pyramidal truss core with the relative density, $\bar{\rho}=0.04$, including the top and bottom faces to capture the truss/face contact. (*b*) The first buckling mode used to incorporate imperfections.

each of the two faces, causing problems when the top face mass plays a key role in the performance. In order to diminish the consequences of this problem, while maintaining computing efficiency, scheme (iii) was devised. Preliminary assessments have

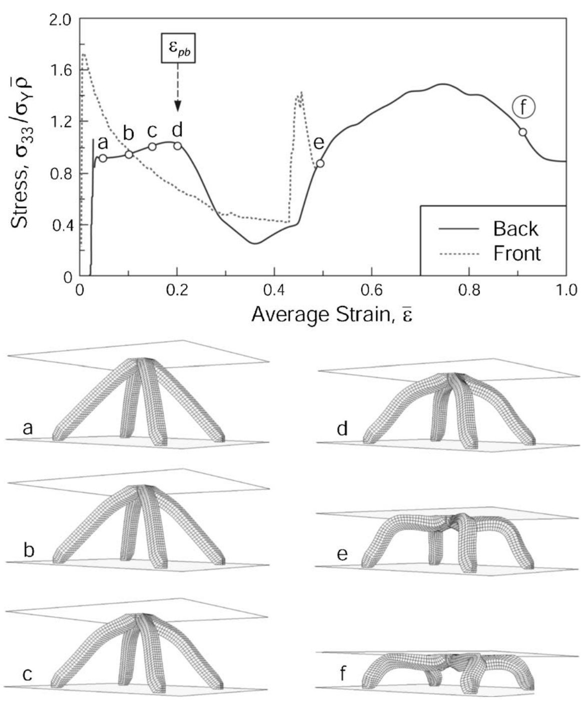


Fig. 6 A representative dynamic stress/strain plot for the truss core and the associated deformations at six different times after imposing the velocity on the front face (effective strain rate, $\dot{\varepsilon}_{eff}=1000/s$). The temporal variations of stress are shown for both the back and front faces. Note that, at location *e*, the front face contacts the core, causing a stress elevation.

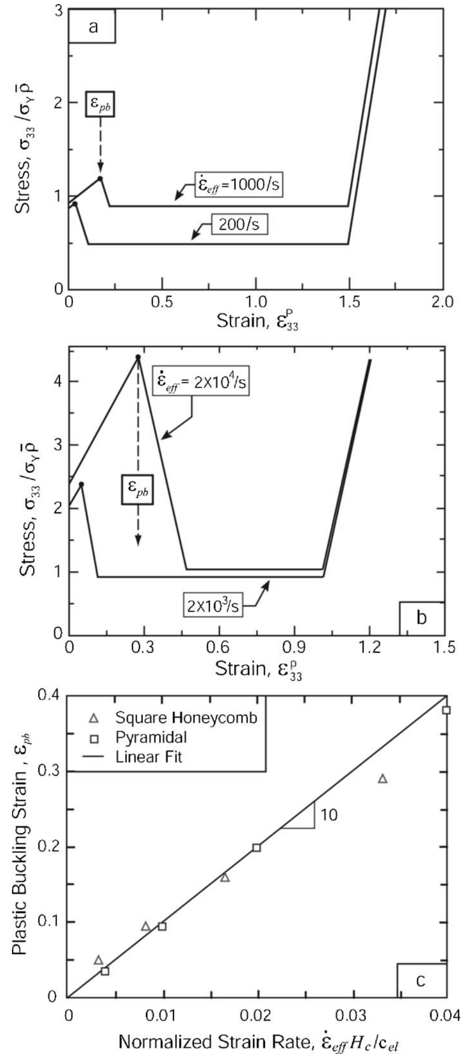


Fig. 7 (*a*) The dynamic stress/strain curves used to characterize truss cores. The curves are for core thickness, $H_c=10$ cm and relative density, $\bar{\rho}=0.04$. (*b*) The dynamic stress/strain curves used to characterize square honeycomb cores. The curves are for core thickness, $H_c=8.3$ mm and relative density, $\bar{\rho}=0.04$. (*c*) Relationship between the strain at which the core members begin to buckle and the effective strain rate, ascertained from (*a*),(*b*).

indicated satisfactory responses for impulsive loadings [16]. The present investigation provides additional insight into the utility of this scheme.

3 Comparisons With Experiments

3.1 Foam Projectile Impacts on Panels With Square Honeycomb Cores. Tests performed on panels with square honeycomb cores, impacted by foam projectiles, are used as one assessment of the dynamic model [6]. Images of the impact sequence (Figs. 3(*a*) and 3(*b*)) provide a synopsis of the test and its outcomes, expressed as the center displacement and the core crush strain, as functions of the incident impulse (Fig. 10). The results of the experiments, as well as simulations with fully meshed cores, are included. The simulations are repeated using the homogenized core model with the curves from Fig. 7 used to characterize the compressive behavior. The shear response is represented by the quasi-static shear stress/strain curve [15]. The dynamic properties of the foam projectile are described elsewhere [17]. It is apparent (Fig. 3 and 10) that the predictions of the

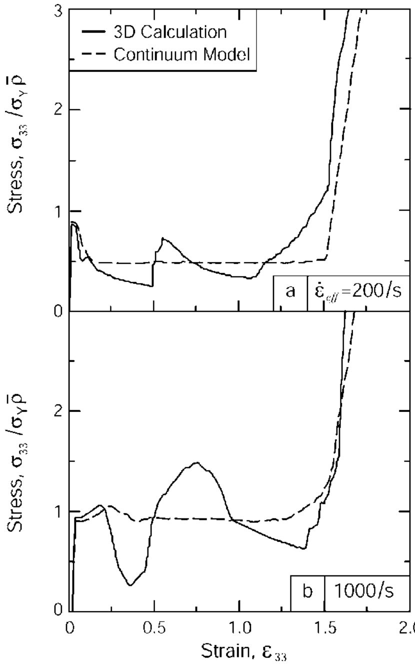


Fig. 8 Two examples of stress/strain variations on the back face comparing 3D results with the predictions obtained using the dynamic stress/strain curves from Fig. 7(*a*) in conjunction with the dynamic constitutive law.

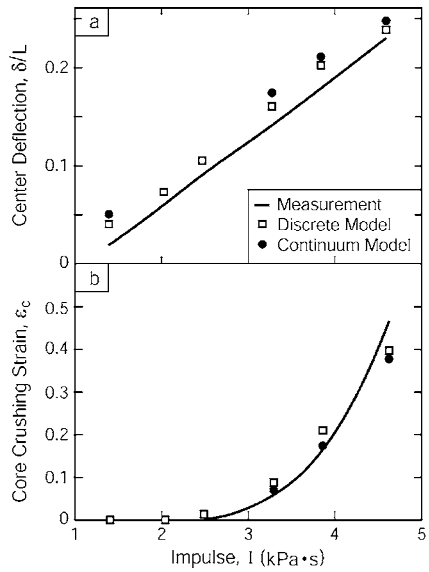


Fig. 10 Comparisons between measurements, full 3D simulations, and simulations conducted using the continuum law with two different mesh schemes, all corresponding to the tests summarized in Fig. 3: (*a*) center point displacements; and (*b*) core crushing strains.

center displacement, the overall panel deformations, and the core crushing strain reproduce the experimental measurements as well as the 3D simulations with reasonable fidelity. However, the deformations of a central cross section induced in a high impulse simulation (Fig. 3(*c*)) reveal two of the limitations of the homogenization approach. (i) The buckling experienced by the actual core members cannot be reproduced. Instead, the core deformation is replaced by a crush domain with crush front partially through the core, extending from the top face. (ii) The face bending is smooth and monotonic, excluding the local bulging where the core members connect to the face. *The largest plastic strains in the faces are thus underestimated*, with implications for the prediction of face tearing. In summary, the constitutive law approach has the attribute that the overall deformations and the average strains are predictable. But, the concentrated local strains where core and face members connect are underestimated, as with any continuum representation of the core. The likely consequence will be poor prediction of local face tearing at the connections with the core members.

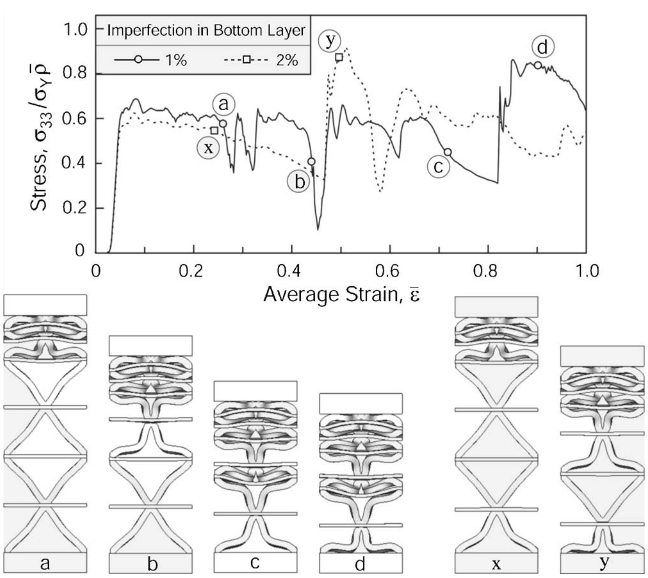


Fig. 9 Crushing sequences for the 3D one-unit column of the seven-layer core and the associated stresses induced on the back face. The solid curve is for a calculation with 1% imperfections in all layers. The corresponding deformation patterns are shown in the sequence *a*→*d*. The dotted curve is for the case where the two layers adjoined to the faces have 2% imperfections, with 1% imperfections elsewhere. Deformation patterns at two strain levels, *x* and *y*, are also shown. Note that, now, the bottom layer starts to crush before the adjacent layers and that there is a large stress pulse (at *y*) when contact occurs between the core and the face in this layer.

3.2 Dynocrusher Tests. An underwater explosive test method (Fig. 4) has been used to investigate the dynamic crushing of a seven-layer pyramidal truss system with average relative density, $\bar{\rho}=0.04$ [4,18]. The measurements reveal that the panel diminishes the pressure from an incident level, $p_0=260$ MPa to $p\approx 10$ MPa, accompanied by a time extension from $0.2\rightarrow 2$ ms. A sectioned side view (Fig. 2) indicates that the top three layers and the bottom layer have almost fully densified, while the three other layers have partially crushed (by plastic buckling of the truss members). The thickness has decreased from 82 mm to 40 mm corresponding to a compressive plastic strain, $\varepsilon_{crush}=0.51$. The test has been simulated using an axisymmetric continuum model (Fig. 4(*b*)). A spring and a dashpot have been introduced beneath the column to address the elasticity and energy absorption of the base, with coefficients calibrated by a reference test using a solid aluminum cylinder [4,18].

Initially, the attributes of the constitutive law have been assessed by placing the panel on a rigid base and computing the back-face stress and impulse as a function of time. For this situation, comparative results can be obtained by using the multilayer unit cell model. The comparison indicates that, despite some differences in detail, the stresses and impulse levels are captured quite faithfully by the continuum model (Figs. 11(*a*) and 11(*b*), respectively). The features missed are the oscillations in stress caused by the plastic buckling and contact events, evident in Fig.

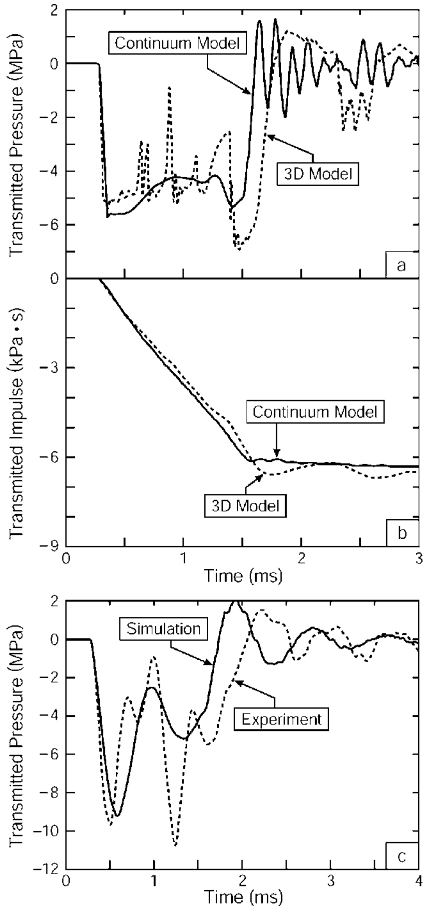


Fig. 11 Comparisons between fully meshed 3D calculations and continuum calculations for a Dynocrusher test with the bottom face rigidly fixed: (*a*) transmitted pressure and (*b*) transmitted impulse. Note the close correspondence for both metrics. (*c*) The pressures transmitted to the gauge columns: a comparison between measurements and values calculated using the constitutive law.

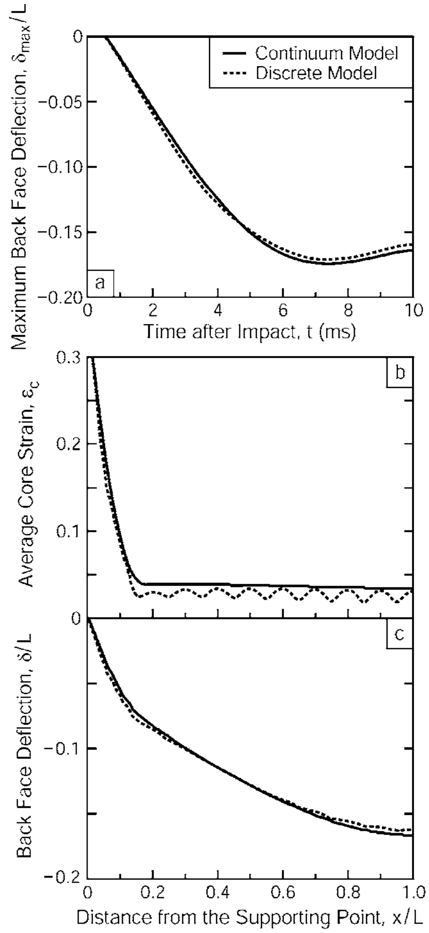


Fig. 12 Comparisons between full 3D calculations and simulations conducted using the continuum model conducted for panels with square honeycomb cores subjected to water blast [1]: (*a*) maximum back face deflection; (*b*) average core strain; and (*c*) back face deflection.

9. The most significant difference is in the amplitude and timing of the stress peak toward the conclusion of crushing, at time f in Fig. 9. This peak occurs because of contact between the core members in the lowest layer and the back face. This deficiency is believed to be the source of the difference between measurements and calculations discussed below.

Because of the discrepancy with the observations (Fig. 3(*c*)), the finding that the layer next to the base does not crush before the adjacent layers has motivated further calculations. Namely, larger imperfections (ξ=0.02) are used in the layers adjoined to the faces. The results are plotted in Fig. 9. Now, the lowest layer does, indeed, crush earlier than some of the adjacent layers. Moreover, this event results in additional stress pulse, which should influence the transmitted pressure. The conclusion is that imperfections can change the crushing sequence and hence, the transmitted pressure and final deformation patterns.

When the support column is inserted, the stresses at the strain gauges calculated using the continuum model are different (Fig. 11(*c*)) because of stress wave effects occurring in the columns. In particular, the stress calculated at the first peak is higher than that for a rigid base (10 MPa rather than 5 MPa) because the column acts as a waveguide. The continuum simulations reproduce this initial pressure pulse. The core crushing strain, $\varepsilon_{\text{crush}}$=0.5, is also closely reproduced (Fig. 2). However, the following discrepancies are apparent. (i) The time required to fully compact the core, characterized by a drop in pressure to a background level, is predicted to be, t_{crush}=1.7 ms, while the measured time is somewhat longer, t_{crush}=2.2 ms. (ii) There is a corresponding underprediction of the total transmitted impulse $I_T \approx 6.5$ kPa s instead of, $I_T \approx 8$ kPa s. (iii) The third pressure peak found in the experiments, occurring at time, t=1.2 ms, is not duplicated. This discrepancy may be present because the imperfection effect described above has not been included in the model.

In summary, the constitutive law approach has the attribute that the overall deformations and peak pressure levels are predicted quite closely. However, certain details are not reproduced, such as the pressure oscillations beyond the initial peak. It remains to ascertain whether the discrepancies present problems when using the continuum approach to predict responses to water blast.

4 Comparisons With Water Blast Simulations

The constitutive law has been used to predict the response to water blast of panels with square honeycomb cores for comparison with full 3D simulations described elsewhere [1]. The results provide another assessment of the constitutive law when fluid/structure interactions are involved. To be consistent with these studies, a longitudinal unit cell has been chosen and the panel supported only at the dry face [1]. The impulse is planar with peak pressure, p_0=100 MPa and duration, t_0=0.1 ms. The half span is, L=1 m, with ratio of core thickness to span, H_c/L=0.2. Results for a panel having a core with relative density, $\bar{\rho}$=0.04, are presented in Fig. 12 and 13, where they are compared with the dis-

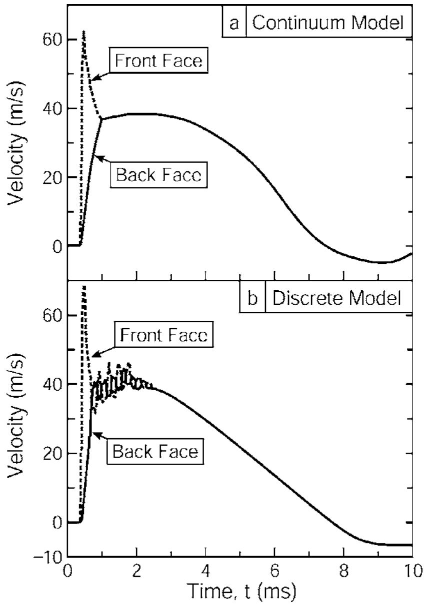


Fig. 13 Velocities of the wet and dry faces for the simulations presented in Fig. 12: (*a*) Simulations conducted using the continuum model; (*b*) full 3D simulations.

crete simulations. Note that the velocity characteristics are consistent with strong core behavior [1]: that is, the faces attain a common velocity and decelerate to rest along a common trajectory. The consistency in the deflections and crushing strains (Fig. 12), as well as the velocities of the faces (Fig. 13), imply that the continuum model captures fluid/structure interactions and provides quite accurate predictions of these metrics. However, deficiencies again arise in the prediction of the plastic strains in the faces (not shown).

5 Multilayer Panels Subject to Water Blast

The constitutive law has been used to ascertain the potential for using panels with multilayer cores to improve blast resistance. Results for a four-layer pyramidal core design are presented in Figs. 14 and 15, and Table 1. The four layers have the same relative density, with the ratio of the back to front face thickness, $\Delta=6$. Two panels (one with $\bar{\rho}=0.5\%$ at $H_c/L=0.3$ and the other with $\bar{\rho}=4\%$ at $H_c/L=0.2$) exhibit typical soft core and strong core mechanisms, respectively [1]. The distinctions are apparent in the

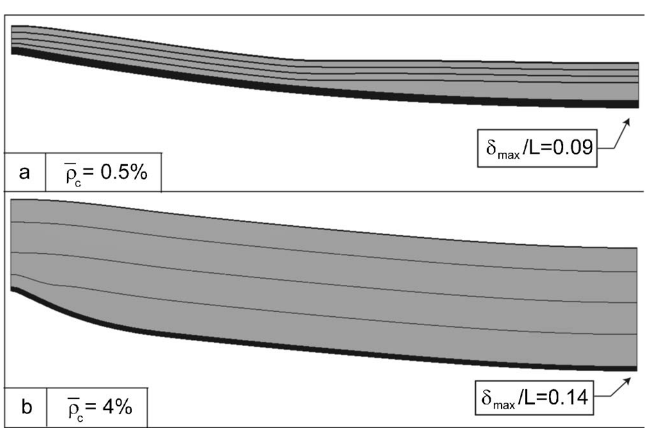


Fig. 14 The deformed shapes of two typical four-layer pyramidal core panels under water blast: (*a*) a soft core with $\bar{\rho}=0.5\%$; and (*b*) a strong core with $\bar{\rho}=4\%$.

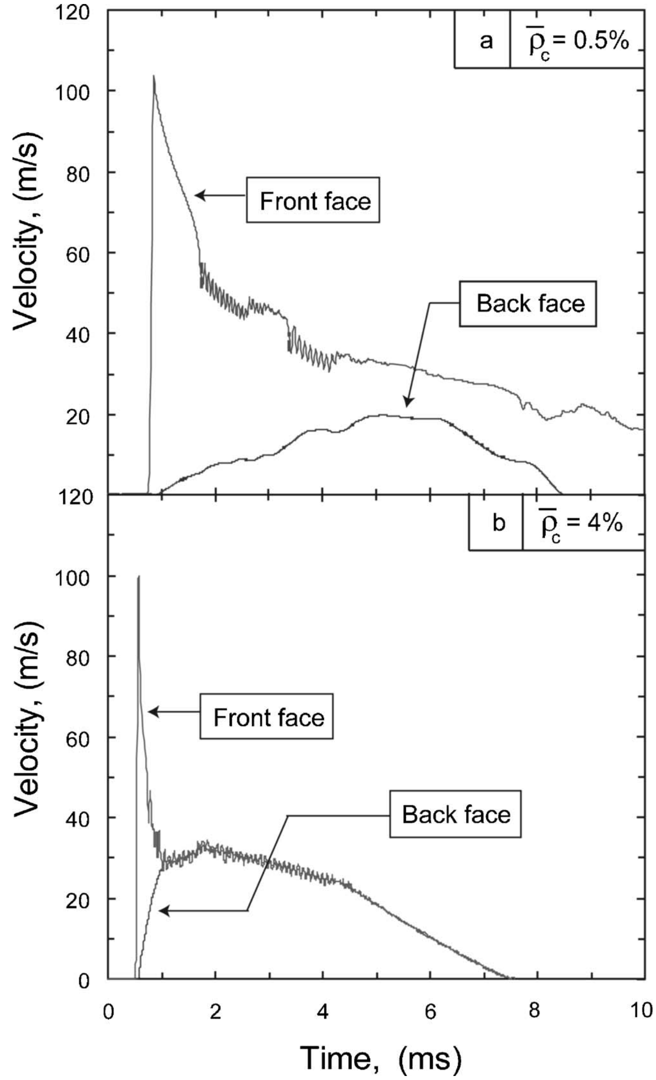


Fig. 15 The velocities of the front and back faces for the two panels from Fig. 14: (*a*) a soft core with $\bar{\rho}=0.5\%$ and (*b*) a strong core with $\bar{\rho}=4\%$.

deformed shapes (Fig. 14) and velocity characteristics (Fig. 15) [1]. The center point displacement and loads transmitted to the supports are compared with the best available results found for designs with *I*- and corrugated cores [1] (Table 1). Plastic strains in the faces have also been determined (Table 1), but because of the limitations of this approach, the results are de-emphasized. It is apparent that the multilayer with $\bar{\rho}=0.5\%$ enables smaller displacements than those found for the *I*- and corrugated cores but

Table 1 Comparisons of the best available results between pyramidal, corrugated, and *I* cores with $H_c/L=0.3$

	Pyramidal $\bar{\rho}=0.5\%$ $\Delta=6$	*I* $\bar{\rho}=1.3\%$ $\Delta=3$	Corrugated $\bar{\rho}=2\%$ $\Delta=5.5$
Back face deflection, δ_b/L	0.090	0.096	0.104
Maximum plastic strain in the front face, ε^p	(0.008)	0.035	0.126
Peak reaction force, P_{react} (kN/m)	1300	1100	850

the transmitted loads are larger. The preferred choice of core thus depends on the metric having greater relevance. Continued exploration of multilayer cores might reveal designs that benefit both metrics.

6 Concluding Comments

An investigation has been conducted that implements and assesses a dynamic continuum constitutive law for all-metallic sandwich panels. Its application to multilayer panels subject to water blast is demonstrated. It is shown that unit cell calculations can be used to calibrate the continuum model. Comparisons between simulations and experiments have been made for two cases. One involves the dynamic response of honeycomb core panels impacted by Al foam projectiles at impulses comparable to those expected in water blast. The second concerns the response of a multilayer truss core to an impulse caused by an explosion occurring in a cylindrical water column. The comparisons reveal that the overall deformation, average core strain, peak transmitted pressure, and velocities of the front and back faces are adequately predicted, inclusive of fluid/structure interactions. The inherent limitations of the approach are the underprediction of the plastic strains in the faces, as with any continuum representation of the core, and incomplete assessment of stress oscillations beyond the peak. The former leads to problems in the prediction of local face tearing.

New results have been presented for multilayer designs with pyramidal truss cores subject to water blast. When compared with results for other designs, the performance of a multilayer with a low-density core is shown to enable smaller displacements.

Acknowledgment

This work was supported by ONR, through both the MURI program on Blast Resistant Structures through a subcontract from Harvard University to the UCSB (Contract No. 123163-03), and the program on Multifunctional Implementations of Topologically Structured Materials through a subcontract from the University of Virginia to UCSB (Contract No. GG10376-114969).

References

[1] Liang, Y., Spuskanyuk, A. V., Flores, S. E., Hayhurst, D. R., Hutchinson, J. W., McMeeking, R. M., and Evans, A. G., 2006, "The Response of Metallic Sandwich Panels to Water Blast," J. Appl. Meteorol., in press.
[2] Tilbrook, M. T., Deshpande, V. S., and Fleck, N. A., 2006, "The Impulsive Response of Sandwich Beams: Analytical and Numerical Investigation of Regimes of Behaviour," J. Mech. Phys. Solids, **54**, pp. 2242–2280.
[3] McMeeking, R. M., Spuskanyuk, A. V., He, H. Y. , Deshande, V. S., Fleck, N. A., and Evans, A. G., 2006, "An Analytic Model for the Response to Water Blast of Unsupported Metallic Sandwich Panels," unpublished.
[4] Wadley, H. N. G., Dharmasena, K. P., Chen, Y., Dudt, P., Knight, D., Charette, R., and Kiddy, K., 2007, "Compressive Response of Multilayered Pyramidal Lattices During Underwater Shock Loading," Int. J. Impact Eng., in press.
[5] Hibbitt, Karlsson & Sorensen, Inc., 2005, *ABAQUS Analysis User's Manual*, Version 6.5, Pawtucket, RI.
[6] Rathbun, H. J., Radford, D. D., Xue, Z., He, M. Y., Yang, J., Deshpande, V., Fleck, N. A., Hutchinson, J. W., Zok, F. W., and Evans, A. G., 2006, "Performance of Metallic Honeycomb-Core Sandwich Beams Under Shock Loading," Int. J. Solids Struct., **43**, pp. 1746–1763.
[7] Desphande, V. S., Fleck, N. A., and Ashby, M. F., 2000, "Isotropic Constitutive Models for Metallic Foams," J. Mech. Phys. Solids, **48**, pp. 1253–1283.
[8] Mohr, D., and Doyoyo, M., 2004, "Large Plastic Deformation of Metallic Honeycomb: Orthotropic Rate-Independent Constitutive Model," Int. J. Solids Struct., **41**, pp. 4435–4456.
[9] Rabczuk, T., Kim, J. K., Samaniego, E., and Belytschko, T., 2004, "Homogenization of Sandwich Structures," Int. J. Numer. Methods Eng., **61**, pp. 1009–1027.
[10] Xue, Z., and Hutchinson, J. W., 2006, "Crush Dynamics of Square Honeycomb Sandwich Cores," Int. J. Numer. Methods Eng., **65**, pp. 2221–2245.
[11] Xue, Z., Vaziri, A., and Hutchinson, J. W., 2005, "Non-Uniform Hardening Constitutive Model for Compressible Orthotropic Materials With Application to Sandwich Plate Cores," Comput. Model. Eng. Sci., **10**, pp. 79–95.
[12] Xue, Z., and Hutchinson, J. W., 2004, "Constitutive Model for Quasi-Static Deformation of Metallic Sandwich Cores," Int. J. Numer. Methods Eng., **61**, pp. 2205–2238.
[13] Hill, R., 1947, "A Theory of the Yielding and Plastic Flow of Anisotropic Materials," Proc. R. Soc. London, Ser. A, **193**, pp. 281–297.
[14] Zok, F. W., Waltner, S. A., Wei, Z., Rathbun, H. J., McMeeking, R. M., and Evans, A. G., 2004, "A Protocol for Characterizing the Structural Performance of Metallic Sandwich Panels, Application to Pyramidal Truss Cores," Int. J. Solids Struct., **41**(22–23), pp. 6249–6271.
[15] Zok, F. W., Rathbun, H. J., He, M. H., Ferri, E., Mercer, C., McMeeking, R. M., and Evans, A. G., 2005, "Structural Performance of Metallic Sandwich Panels With Square Honeycomb Cores," Philos. Mag., **85**, pp. 3207–3234.
[16] Xue, Z., 2005, private communication.
[17] Radford, D. D., Deshpande, V. S., and Fleck, N. A., 2005, "The Use of Metal Foam Projectiles to Simulate Shock Loading on a Structure," Int. J. Impact Eng., **31**, pp. 1152–1171.
[18] Wei, Z., Dharmasena, K. P., Wadley, H. N. G., and Evans, A. G., 2006, "Analysis and Interpretation of a Test for Characterizing the Response of Metallic Sandwich Panels to Water Blast," Int. J. Impact Eng., submitted.

Effect of Boundary Conditions on Nonlinear Vibrations of Circular Cylindrical Panels

M. Amabili
Dipartimento di Ingegneria Industriale,
Università di Parma,
Parco Area delle Scienze 181/A,
Parma 43100,
Italy
e-mail: marco.amabili@unipr.it

Geometrically nonlinear vibrations of circular cylindrical panels with different boundary conditions and subjected to harmonic excitation are numerically investigated. The Donnell's nonlinear strain–displacement relationships are used to describe geometric nonlinearity; in-plane inertia is taken into account. Different boundary conditions are studied and the results are compared; for all of them zero normal displacements at the edges are assumed. In particular, three models are considered in order to investigate the effect of different boundary conditions: Model A for free in-plane displacement orthogonal to the edges, elastic distributed springs tangential to the edges and free rotation; Model B for classical simply supported edges; and Model C for fixed edges and distributed rotational springs at the edges. Clamped edges are obtained with Model C for the very high value of the stiffness of rotational springs. The nonlinear equations of motion are obtained by the Lagrange multimode approach, and are studied by using the code AUTO based on the pseudo-arclength continuation method. Convergence of the solution with the number of generalized coordinates is numerically verified. Complex nonlinear dynamics is also investigated by using bifurcation diagrams from direct time integration and calculation of the Lyapunov exponents and the Lyapunov dimension. Interesting phenomena such as (i) subharmonic response; (ii) period doubling bifurcations; (iii) chaotic behavior; and (iv) hyper-chaos with four positive Lyapunov exponents have been observed.
[DOI: 10.1115/1.2424474]

1 Introduction

The first studies on geometrically nonlinear vibrations of classical simply supported, circular cylindrical shallow shells are due to Reissner [1], Grigolyuk [2], and Cummings [3]. They used the Donnell's nonlinear shallow-shell theory, i.e., neglecting in-plane inertia, and single-mode approximation of the transverse deflection in order to reduce the solution to a single degree of freedom.

Leissa and Kadi [4] studied linear and nonlinear free vibrations of doubly curved shallow panels, classical simply supported at the four edges. Donnell's nonlinear shallow-shell theory was used in a slightly modified version to take into account the meridional curvature. A single mode expansion of the transverse displacement was used.

Extensive literature review on nonlinear vibration of circular cylindrical shells and panels is given by Amabili and Païdoussis [5] and therefore it is not reported here.

In particular, not many studies deal with nonlinear vibrations of circular cylindrical panels with boundary conditions different from classical simply supported edges. In particular, Chia [6] studied laminated shallow cylindrical panels with mixed boundary conditions resting on elastic foundation. Zero transverse displacement is assumed at the edges where rotational springs are introduced. In-plane boundary conditions are zero tangential force and given force orthogonal to the edges. A single mode analysis is carried out by using the Donnell's nonlinear shallow-shell theory. Fu and Chia [7] extended this study by considering a multimode solution.

Yamaguchi and Nagai [8] studied vibrations of a shallow cylindrical panel with a rectangular boundary, with zero transverse deflection and free rotation at the boundaries, and with in-plane elastic constraints orthogonal to the edges and free in-plane tangential to the edges. The panel was excited by an acceleration having a constant value plus a harmonic component. Donnell's nonlinear shallow-shell theory was utilized with the Galerkin projection along with a multimode expansion of the flexural displacement. In-plane boundary conditions were satisfied on the average. Initial imperfection was taken into account in the theoretical formulation but not in the calculations. The harmonic balance method and direct integration were used. The response of the panel was of softening type over the whole range of possible stiffness values of the in-plane springs (elastic support), becoming hardening for a vibration amplitude on the order of the shell thickness; in-plane constraints reduce the softening nonlinearity, which turns to hardening for smaller vibration amplitudes. The objective of the study was to investigate regions of chaotic motion; these regions were identified by means of Poincaré maps and Lyapunov exponents. It was found that, when approaching the static instability point (due to the constant acceleration load), chaotic shell behavior may be observed.

Free vibrations of doubly-curved, laminated, clamped shallow panels, including circular cylindrical panels, were studied by Abe et al. [9]. Both first-order shear deformation theory and a shell theory analogous to the Donnell's theory were used. Results obtained neglecting in-plane and rotary inertia are very close to those obtained retaining these effects. Two modes were considered to interact in the nonlinear analysis of the second mode (2,1) (two circumferential and one longitudinal half-waves), while a single mode was used for the first mode (1,1) (one circumferential and one longitudinal half-wave).

Amabili [10] studied the large-amplitude, forced vibrations of mode (1,1) of circular cylindrical panels with rectangular base, classical simply supported at the four edges, and subjected to radial harmonic excitation. Two different nonlinear strain–displacement relationships, from the Donnell's and the Novozhilov's shell theories, were used to calculate the elastic strain energy. In-plane inertia and geometric imperfections were taken into account. The solution was obtained by the Lagrangian approach. Convergence of the solution was shown and differences between the Donnell's and the Novozhilov's nonlinear shell theo-

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received August 31, 2005; final manuscript received February 6, 2006. Review conducted by N. Sri Namachchivaya.

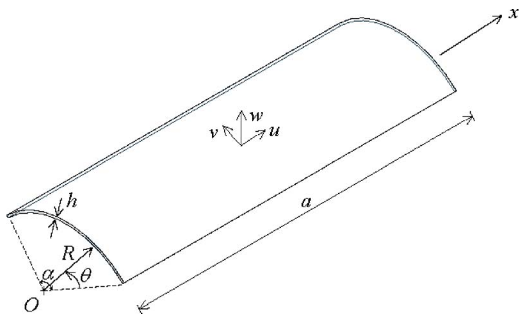


Fig. 1 Geometry of the panel, coordinate system and symbols

ries were fully negligible. A very satisfactory comparison with the results obtained by Kobayashi and Leissa [11] was given. Interaction of modes having integer ratio among their natural frequencies, giving rise to internal resonances, are also discussed.

In the present study, for the first time the effect of boundary conditions on the trend of nonlinearity of circular cylindrical panels is studied. This is of particular interest because results show that a panel with the same geometry presents a significant softening type nonlinearity if simply supported or with fixed edges, while it has a relatively strong, hardening type nonlinearity for free in-plane edges and for clamped edges. Here geometrically nonlinear vibrations of circular cylindrical panels with different boundary conditions and subjected to harmonic excitation are numerically investigated. The Donnell's nonlinear strain–displacement relationships are used to describe the geometric nonlinearity but in-plane inertia is taken into account. Different boundary conditions are studied and results are compared; for all of them, zero transverse displacements at the edges are assumed. In particular, three models are considered in order to investigate different boundary conditions: Model A, recently developed by Amabili [12], for free in-plane displacements orthogonal to the edges, elastic distributed springs tangential to the edges, and free rotation; Model B for classical simply supported edges, which has been previously developed and validated by the author [10]; and Model C for fixed edges and distributed rotational springs at the edges, especially developed for the present study. Clamped edges are obtained with Model C for a very high value of the stiffness of the rotational springs. The nonlinear equations of motion are obtained by the Lagrange multimode approach, and are studied by using the code AUTO based on the pseudo-arclength continuation method. The convergence of the solution with the number of generalized coordinates is numerically verified. Complex nonlinear dynamics is also investigated by using bifurcation diagrams from direct time integration and calculation of the Lyapunov exponents and the Lyapunov dimension. Interesting phenomena such as (i) subharmonic response; (ii) period doubling bifurcations; (iii) chaotic behavior; and (iv) hyper-chaos with four positive Lyapunov exponents have been observed.

2 Elastic Strain Energy of the Panel

A circular cylindrical panel with the cylindrical coordinate system $(O;x,r,\theta)$, having the origin O at the center of one end of the panel, is considered, as shown in Fig. 1. The displacements of an arbitrary point of coordinates (x,θ) on the middle surface of the panel are denoted by u, v, and w, in the axial, circumferential, and radial directions, respectively; w is taken positive outward. Initial imperfections of the circular cylindrical panel associated with zero initial tension are denoted by radial displacement w_0, and positive outward; only radial initial imperfections are considered.

The Donnell's strain–displacement relationships for thin shells are used; they are based on Love's first approximation assumptions. The strain components ε_x, ε_θ, and $\gamma_{x\theta}$ at an arbitrary point of the panel are related to the middle surface strains $\varepsilon_{x,0}$, $\varepsilon_{\theta,0}$, and $\gamma_{x\theta,0}$ and to the changes in the curvature and torsion of the middle surface k_x, k_θ, and $k_{x\theta}$ by the following three relationships

$$\varepsilon_x = \varepsilon_{x,0} + zk_x, \quad \varepsilon_\theta = \varepsilon_{\theta,0} + zk_\theta, \quad \gamma_{x\theta} = \gamma_{x\theta,0} + zk_{x\theta} \tag{1}$$

where z is the distance of the arbitrary point of the panel from the middle surface. The middle surface strain–displacement relationships and changes in the curvature and torsion for a circular cylindrical panel are [10] (see Ref. [5] for other classical references)

$$\varepsilon_{x,0} = \frac{\partial u}{\partial x} + \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2 + \frac{\partial w}{\partial x}\frac{\partial w_0}{\partial x} \tag{2a}$$

$$\varepsilon_{\theta,0} = \frac{\partial v}{R\partial\theta} + \frac{w}{R} + \frac{1}{2}\left(\frac{\partial w}{R\partial\theta}\right)^2 + \frac{\partial w}{R\partial\theta}\frac{\partial w_0}{R\partial\theta} \tag{2b}$$

$$\gamma_{x\theta,0} = \frac{\partial u}{R\partial\theta} + \frac{\partial v}{\partial x} + \frac{\partial w}{\partial x}\frac{\partial w}{R\partial\theta} + \frac{\partial w}{\partial x}\frac{\partial w_0}{R\partial\theta} + \frac{\partial w_0}{\partial x}\frac{\partial w}{R\partial\theta} \tag{2c}$$

$$k_x = -\frac{\partial^2 w}{\partial x^2} \tag{2d}$$

$$k_\theta = -\frac{\partial^2 w}{R^2\partial\theta^2} \tag{2e}$$

$$k_{x\theta} = -2\frac{\partial^2 w}{R\partial x\partial\theta} \tag{2f}$$

The elastic strain energy U_S of a circular cylindrical panel, neglecting σ_z as stated by Love's first approximation assumptions, is given by

$$U_S = \frac{1}{2}\int_0^\alpha \int_0^a \int_{-h/2}^{h/2} (\sigma_x\varepsilon_x + \sigma_\theta\varepsilon_\theta + \tau_{x\theta}\gamma_{x\theta})dz\, dx\, R(1+z/R)d\theta \tag{3}$$

where h is the panel thickness; R is the panel middle radius; a is the panel length; α is the angular dimension of the panel; and the stresses σ_x; σ_θ, and $\tau_{x\theta}$ are related to the strains for homogeneous and isotropic material ($\sigma_z=0$, case of plane stress) by

$$\sigma_x = \frac{E}{1-\nu^2}(\varepsilon_x + \nu\varepsilon_\theta), \quad \sigma_\theta = \frac{E}{1-\nu^2}(\varepsilon_\theta + \nu\varepsilon_x), \quad \tau_{x\theta} = \frac{E}{2(1+\nu)}\gamma_{x\theta} \tag{4}$$

where E is the Young's modulus and ν is Poisson's ratio. By using Eqs. (1), (3), and (4), the following expression is obtained

$$U_S = \frac{1}{2}\frac{Eh}{1-\nu^2}\int_0^\alpha \int_0^a \left(\varepsilon_{x,0}^2 + \varepsilon_{\theta,0}^2 + 2\nu\varepsilon_{x,0}\varepsilon_{\theta,0} + \frac{1-\nu}{2}\gamma_{x\theta,0}^2\right)dxR\,d\theta + \frac{1}{2}\frac{Eh^3}{12(1-\nu^2)}\int_0^\alpha \int_0^a \left(k_x^2 + k_\theta^2 + 2\nu k_x k_\theta + \frac{1-\nu}{2}k_{x\theta}^2\right)dxR\,d\theta + O(h^4) \tag{5}$$

where $O(h^4)$ is a higher-order term in h; the first term is the membrane (also referred to as stretching) energy; and the second one is the bending energy.

3 Boundary Conditions, Kinetic Energy, and External Loads

The kinetic energy T_S of a circular cylindrical panel, by neglecting rotary inertia, is given by

Table 1 Natural frequency of mode (1, 1) for different boundary conditions

Model	Boundary condition	Natural frequency (Hz)
A	$k=0$	549.3 (19 DOF) 539.4 (50 DOF)
A	$k=4\times10^9$ N/m^2	597.0 (19 DOF)
B	Simply supported	636.9 (9 DOF)
C	$c=0$	912.0 (39 DOF) 912.0 (79 DOF)
C	$c=5\times10^4$ N	1211.2 (39 DOF) 1168.3 (79 DOF)

$$T_S=\frac{1}{2}\rho_S h\int_0^{\alpha}\int_0^{a}(\dot{u}^2+\dot{v}^2+\dot{w}^2)dxR\,d\theta \tag{6}$$

where ρ_S is the mass density of the panel. In Eq. (6) the overdot denotes a time derivative.

The virtual work W done by the external forces is written as

$$W=\int_0^{\alpha}\int_0^{a}(q_x u+q_\theta v+q_r w)dxR\,d\theta \tag{7}$$

where q_x, q_θ, and q_r are the distributed forces per unit area acting in axial, circumferential, and radial directions, respectively. Initially, only a single harmonic radial force is considered; therefore $q_x=q_\theta=0$. The external radial distributed load q_r applied to the panel, due to the radial concentrated force $\tilde{f}$, is given by

$$q_r=\tilde{f}\delta(R\theta-R\tilde{\theta})\delta(x-\tilde{x})\cos(\omega t) \tag{8}$$

where ω is the excitation frequency; t is the time; δ is the Dirac delta function; $\tilde{f}$ gives the radial force amplitude positive in w direction; and $\tilde{x}$ and $\tilde{\theta}$ give the axial and angular positions of the point of application of the force, respectively. Equation (7) can be rewritten in the following form

$$W=\tilde{f}\cos(\omega t)(w)_{x=\tilde{x},\theta=\tilde{\theta}} \tag{9}$$

The following boundary conditions are introduced in the present study:

Model A

$$N_x=w=w_0=M_x=\partial^2 w_0/\partial x^2=0,\quad N_{x,y}=-kv,\quad \text{at } x=0,a \tag{10}$$

$$N_y=w=w_0=M_y=\partial^2 w_0/\partial y^2=0,\quad N_{y,x}=-ku,\quad \text{at } y=0,b \tag{11}$$

Model B

$$v=w=w_0=N_x=M_x=\partial^2 w_0/\partial x^2=0,\quad \text{at } x=0,a \tag{12}$$

$$u=w=w_0=N_y=M_y=\partial^2 w_0/\partial y^2=0,\quad \text{at } y=0,b \tag{13}$$

Model C

$$u=v=w=w_0=0,\quad M_x=\frac{Eh^3}{12(1-\nu^2)}(k_x+\nu k_y)=c\partial w/\partial x \quad \text{at } x=0,a \tag{14}$$

$$u=v=w=w_0=0,\quad M_y=\frac{Eh^3}{12(1-\nu^2)}(k_y+\nu k_x)=c\partial w/\partial y \quad \text{at } y=0,b \tag{15}$$

where $y=R\theta$; $b=R\alpha$; k is the distributed in-plane spring stiffness (N/m^2) where springs are distributed along the panel edges in the edge direction; M_x and M_y are the bending moments per unit length on the edges orthogonal to x and y, respectively; N_x and N_y are the normal forces per unit length; $N_{x,y}$ is the shear force per unit length; and c is the stiffness per unit length of the elastic, distributed rotational springs placed at the four edges, $x=0$, a and $y=0$, b. w is restrained at the four panel edges for all the three models. Model A: Eqs. (10) and (11) give fully free in-plane for $k=0$ and classical simply supported conditions in the limit case $k\to\infty$ this model has been developed in Ref. [12]. Model B: Eqs. (12) and (13) give the classical simply supported boundary conditions and the problem has been studied in Ref. [10]. Model C is developed in the present study and gives fixed edges in-plane with free rotation for $c=0$ and a perfectly clamped panel ($\partial w/\partial x=0$ and $\partial w/\partial y=0$) obtained as the limit for $c\to\infty$. In the case of c, unlike zero, an additional potential energy stored by the elastic rotational springs at the panel edges must be added. This potential energy U_R is given by

$$U_R=\frac{1}{2}\int_0^b c\left\{\left[\left(\frac{\partial w}{\partial x}\right)_{x=0}\right]^2+\left[\left(\frac{\partial w}{\partial x}\right)_{x=a}\right]^2\right\}dy+\frac{1}{2}\int_0^a c\left\{\left[\left(\frac{\partial w}{\partial y}\right)_{y=0}\right]^2+\left[\left(\frac{\partial w}{\partial y}\right)_{y=b}\right]^2\right\}dx \tag{16}$$

In Eq. (16) a nonuniform stiffness c (function of x or y, simulating a nonuniform constraint) can be assumed. In order to simulate clamped edges in numerical calculations, a very high value of the stiffness c must be assumed. This approach is usually referred to as the artificial spring method which can be regarded as a variant of the classical penalty method. The values of the spring stiffness simulating a clamped panel can be obtained by trials and errors or by evaluating the edge bending stiffness of the panel. In fact, it has been found that the natural frequencies of the system converge asymptotically to those of a clamped panel when c becomes very large. Only the development of Model C is discussed in the fol-

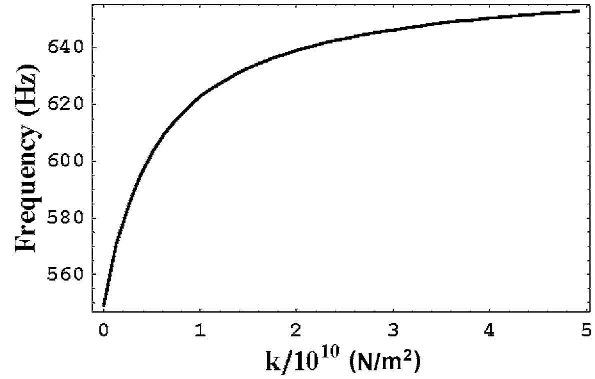


Fig. 2 Natural frequency of mode (1,1) of supported panel computed with model A versus *k*; 19 DOF

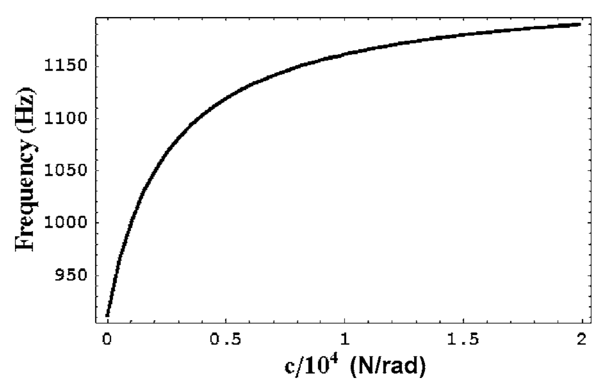


Fig. 3 Natural frequency of mode (1,1) of panel with fixed edges computed with model C versus *c*; 39 DOF

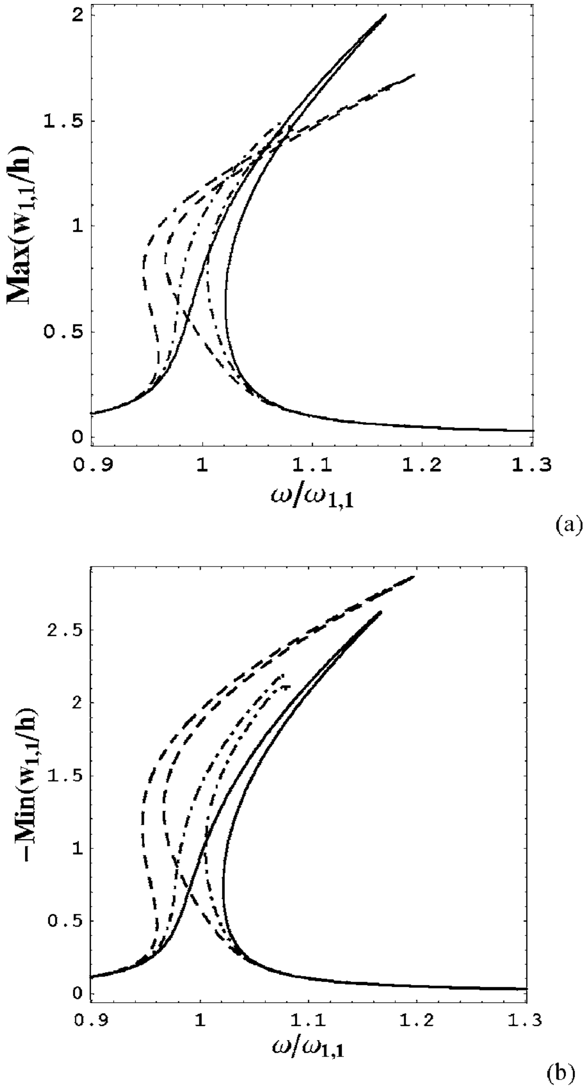


Fig. 4 Nondimensional response of the panel for different boundary conditions versus non-dimensional excitation frequency; mode (1,1), f=0.021, $\zeta_{1,1}$=0.004. (– –) classical simply supported panel (model B), 9 DOF; (— - —) model A with k=4 $\times 10^9$ N/m^2, 19 DOF; (—) model A with k=0 (in-plane free edges), 19 DOF. (*a*) Maximum of generalized coordinate $w_{1,1}$; and (*b*) minimum of $w_{1,1}$.

lowing section and the interested reader can find mathematical details on Models A and B in Refs. [12,10], respectively.

4 Mode Expansion and Lagrange Equations of Motion

In order to reduce the system to finite dimensions, the middle surface displacements u, v, and w are expanded by using the following approximate functions, which identically satisfy the geometric boundary conditions in Eqs. (14) and (15)

$$w(x,y,t) = \sum_{m=1}^{M_1} \sum_{n=1}^{N_1} w_{m,n}(t)\sin(m\pi x/a)\sin(n\pi y/b) \tag{17}$$

$$u(x,y,t) = \sum_{m=1}^{M_2} \sum_{n=1}^{N_2} u_{m,n}(t)\sin(m\pi x/a)\sin(n\pi y/b) \tag{18}$$

$$v(x,y,t) = \sum_{m=1}^{M_3} \sum_{n=1}^{N_3} v_{m,n}(t)\sin(m\pi x/a)\sin(n\pi y/b) \tag{19}$$

where m and n are the numbers of half-waves in x and y directions, respectively, and t is the time; $u_{m,n}(t)$, $v_{m,n}(t)$, and $w_{m,n}(t)$ are the generalized coordinates that are unknown functions of t. $M_\#$ and $N_\#$ indicate the terms necessary in the expansion of the displacements, where $\#=1,2,3$.

Only out-of-plane initial geometric imperfections of the panel are assumed; they are associated with zero initial stress. The imperfection w_0 is expanded in the same form of w, i.e., in a double Fourier sine series satisfying the boundary conditions (14) and (15) at the panel edges

$$w_0(x,y) = \sum_{m=1}^{\tilde{M}} \sum_{n=1}^{\tilde{N}} A_{m,n} \sin(m\pi x/a)\sin(n\pi y/b) \tag{20}$$

where $A_{m,n}$ are the modal amplitudes of imperfections; and $\tilde{N}$ and $\tilde{M}$ are integers indicating the number of terms in the expansion.

The nonconservative damping forces are assumed to be of viscous type and are taken into account by using Rayleigh's dissipation function

$$F = \frac{1}{2}c_d \int_0^a \int_0^b (\dot{u}^2 + \dot{v}^2 + \dot{w}^2)dx\, dy \tag{21}$$

where c_d has a different value for each term of the mode expansion. Simple calculations give

$$F = \frac{1}{2}(ab/4)\left[\sum_{n=1}^{N_1}\sum_{m=1}^{M_1} c_{m,n}\dot{w}_{m,n}^2 + \sum_{n=1}^{N_2}\sum_{m=1}^{M_2} c_{m,n}\dot{u}_{m,n}^2 + \sum_{n=1}^{N_3}\sum_{m=1}^{M_3} c_{m,n}\dot{v}_{m,n}^2\right] \tag{22}$$

The damping coefficient $c_{m,n}$ is related to the damping ratio by $\zeta_{m,n}=c_{m,n}/(2\mu_{m,n}\omega_{m,n})$, where $\omega_{m,n}$ is the natural circular frequency of mode (m,n) and $\mu_{m,n}$ is the modal mass of this generalized coordinate, given by $\mu_{m,n}=\rho_S h(ab/4)$.

The following notation is introduced for brevity

$$\mathbf{q} = \{u_{m,n}, v_{m,n}, w_{m,n}\}^T, \quad m = 1, \ldots M_{1 \text{ or } 2 \text{ or } 3} \quad \text{and } n = 1, \ldots N_{1 \text{ or } 2 \text{ or } 3} \tag{23}$$

The generic element of the time-dependent vector $\mathbf{q}$ is referred to as q_j, which is the generalized coordinate; the dimension of $\mathbf{q}$ is N, which is the number of degrees of freedom (DOF) used in the mode expansion.

The generalized forces Q_j are obtained by differentiation of Rayleigh's dissipation function and of the virtual work done by external concentrated harmonic force at the center of the panel

$$Q_j = -\frac{\partial F}{\partial \dot{q}_j} + \frac{\partial W}{\partial q_j} = -(ab/4)c_j\dot{q}_j + \begin{cases} 0 & \text{if } q_j = u_{m,n}, v_{m,n}; \text{or } w_{m,n} \quad \text{with } m \text{ or } n \text{ even} \\ \tilde{f}\cos(\omega t) & \text{if } q_j = w_{m,n} \quad \text{with both } m \text{ and } n \text{ odd} \end{cases} \tag{24}$$

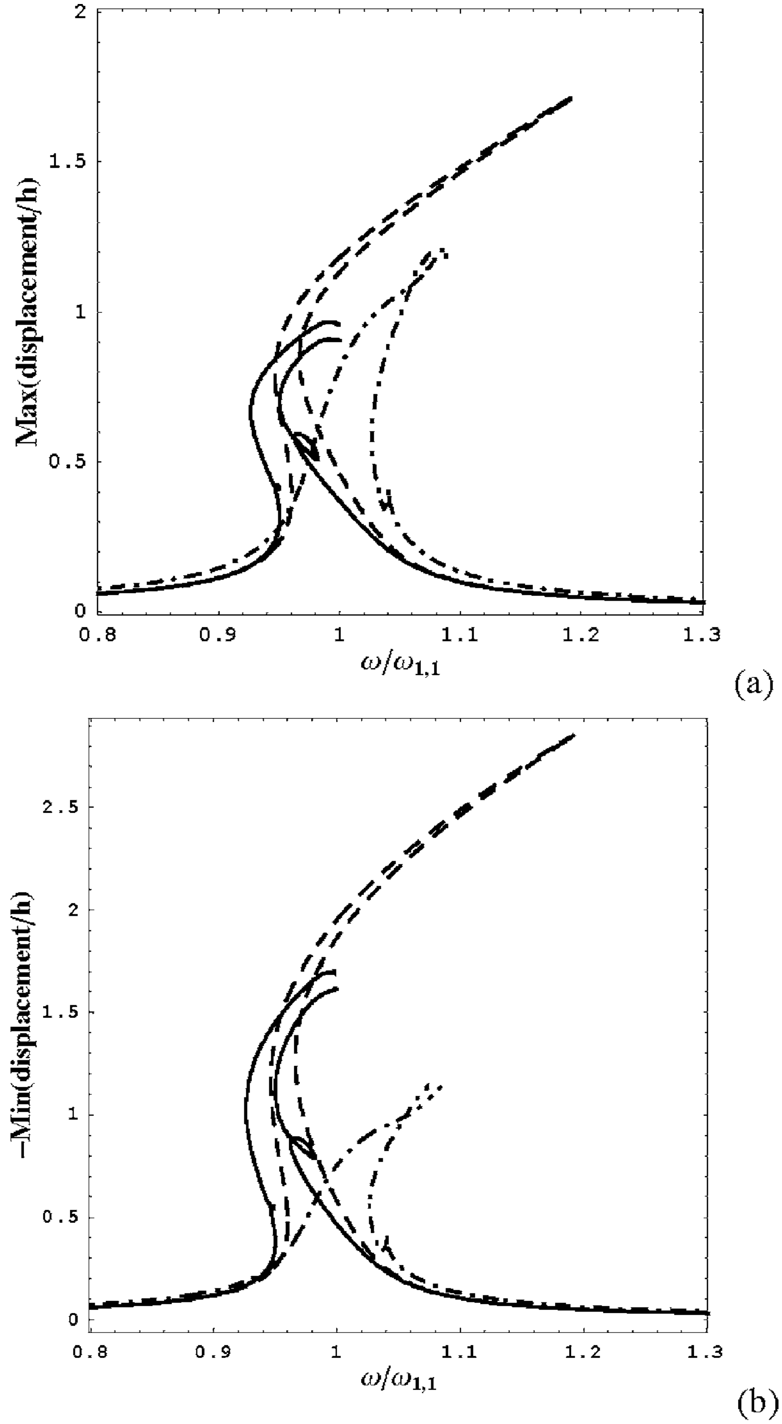


Fig. 5 Nondimensional response of the panel for different boundary conditions versus nondimensional excitation frequency; mode (1,1), $f=0.021$, $\zeta_{1,1}=0.004$. (– –) classical simply supported panel (model B), 9 DOF; (— - —) model C with $c=5\times10^4$ N/rad (practically clamped), 39 DOF; (—) model C with $c=0$ (fixed edges), 39 DOF. (*a*) Maximum transverse displacement at the center of the panel; and (*b*) minimum transverse displacement at the center of the panel.

The Lagrange equations of motion are

$$\frac{d}{dt}\left(\frac{\partial T_S}{\partial \dot{q}_j}\right)-\frac{\partial T_S}{\partial q_j}+\frac{\partial(U_S+U_R)}{\partial q_j}=Q_j, \quad j=1,\ldots N \tag{25}$$

where $\partial T_S/\partial q_j=0$. These second-order equations have very long expressions containing quadratic and cubic nonlinear terms. In particular

$$\frac{d}{dt}\left(\frac{\partial T_S}{\partial \dot{q}_j}\right)=\rho_S h(ab/4)\ddot{q}_j \tag{26}$$

which shows that no inertial coupling among the Lagrange equations exists for the panel with the mode expansion used.

The very complicated term giving quadratic and cubic nonlinearities can be written in the form

$$\frac{\partial(U_S+U_R)}{\partial q_j}=\sum_{k=1}^{\text{dofs}} q_k f_{k,j}+\sum_{i,k=1}^{\text{dofs}} q_i q_k f_{i,k,j}+\sum_{i,k,l=1}^{\text{dofs}} q_i q_k q_l f_{i,k,l,j} \tag{27}$$

where coefficients f have long expressions that also include geometric imperfections.

5 Numerical Techniques

The Lagrange equations have been obtained by using the Mathematica computer software [13] in order to perform analytical surface integrals of trigonometric functions (e.g., integrals in Eq. (5)). The generic jth Lagrange equation is divided by the mass of the jth generalized coordinate (associated with $\ddot{q}_j$) and then is transformed in two first-order equations. A nondimensionalization of variables is also performed for computational convenience: the frequencies are divided by the natural circular frequency $\omega_{m,n}$ of the mode (m,n) investigated, and the vibration amplitudes are divided by the panel thickness h. The resulting $2\times N$ equations are studied by using: (i) the software AUTO 97 [14] for continuation and bifurcation analysis of nonlinear ordinary differential equations, and (ii) direct integration of the equations of motion by using the DIVPAG routine of the Fortran library IMSL. The software AUTO 97 is capable of continuation of the solution, bifurcation analysis, and branch switching by using pseudo-arclength continuation and collocation methods; in the present study the program has been modified in order to handle more variables and it has been recompiled for the PC. In particular, the panel response under harmonic excitation has been studied by using an analysis in two steps: (i) first the excitation frequency has been fixed far enough from resonance and the magnitude of the excitation has been used as the bifurcation parameter; the solution has been started at zero force where the solution is the trivial undisturbed configuration of the panel and has been continued up to reach the desired force magnitude; and (ii) when the desired magnitude of excitation has been reached, the solution has been continued by using the excitation frequency as the bifurcation parameter.

Direct integration of the equations of motion by using Gear's BDF method (routine DIVPAG of the Fortran library IMSL) has also been performed to check the results and obtain the time behavior. Gear's algorithm has been used due to the relatively high dimension of the dynamical system. Indeed, when a high-dimensional phase space is analyzed, the problem can present stiff characteristics, due to the presence of different time scales in the response. In simulations with adaptive step size Runge–Kutta methods, spurious nonstationary and divergent motions can be obtained. Therefore, Gear's method, designed for stiff equations, was used.

The bifurcation diagram of the Poincaré maps was also used in the case of nonstationary response. It has been constructed by using the time integration scheme and by varying the force amplitude.

The maximum Lyapunov exponent has been computed with the algorithm described in Refs. [15,16] by simultaneous integration of the $4\times N$ first-order differential equations (the nonlinear equations of motion are integrated by using DIVPAG IMSL routine and the variational linear equations with time-varying coefficients [16] are integrated by using the adaptive step-size 4th–5th-order Runge–Kutta method). The excitation period has been divided into 10,000 integration steps in order to have accurate evaluation of the time-varying coefficients. To find a reference trajectory, 6×10^6 steps are skipped in order to eliminate the transient and 1×10^6 steps are skipped to eliminate the transitory on the variational equations. Then 1×10^6 steps are used for evaluation of the maximum Lyapunov exponent.

The Fortran computer program developed to calculate all the Lyapunov exponents has been described in Ref. [16]. In particular, 1×10^7 steps have been used to evaluate the Lyapunov exponents (ten times larger than for calculation of the maximum Lyapunov exponent, which has been evaluated for all the bifurcation diagram with high computational cost).

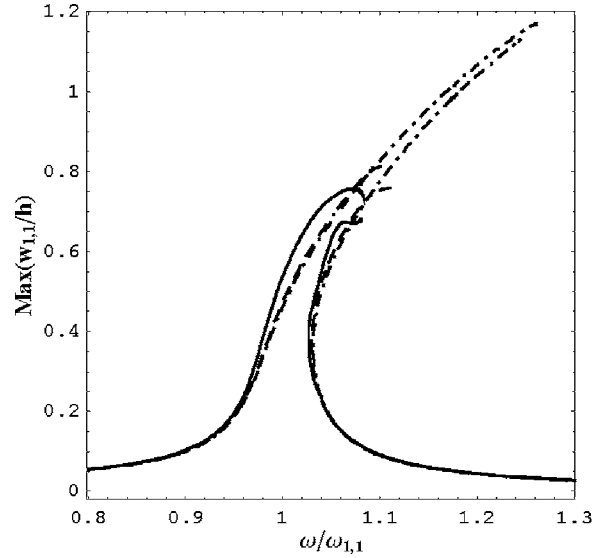


Fig. 6 Convergence of model C ($c=5\times10^4$ N/rad) for nonlinear forced vibrations of the clamped panel; nondimensional response of generalized coordinate $w_{1,1}$ versus non-dimensional excitation frequency; fundamental mode (1,1), $f=0.021$, $\zeta_{1,1}=0.004$. (— - —) 24 DOF; (– –) 27 DOF; (—) 39 DOF.

6 Numerical Results for Harmonic Response

Numerical calculations have been performed for the harmonic response of a circular cylindrical panel having the following dimension and material properties: length between supports $L=0.1$ m, radius of curvature $R=1$ m, thickness $h=0.001$ m, angular width between supports $\alpha=0.1$ rad (i.e., the panel has length equal to the circumferential width), Young's modulus $E=206\times10^9$ Pa, mass density $\rho=7800$ kg/m^3 and Poisson ratio $\nu=0.3$. A panel with the same dimension ratios ($R/L=10$, $h/L=0.01$, $R\alpha/L=1$, $\nu=0.3$) was previously studied by Kobayashi and Leissa [11] and Amabili [10]. The following generalized coordinates have been used for Model A, giving a 19 DOF model: $w_{1,1}$, $w_{1,3}$, $w_{3,1}$, $w_{3,3}$, $u_{1,0}$, $u_{1,2}$, $u_{1,4}$, $u_{3,0}$, $u_{3,2}$, $u_{3,4}$, $v_{0,1}$, $v_{2,1}$, $v_{4,1}$, $v_{0,3}$, $v_{2,3}$, $v_{4,3}$, $v_{0,5}$, $v_{2,5}$, $v_{4,5}$; convergence of this model has been shown in Ref. [12]. The following generalized coordinates have been used for Model B, giving a 9 DOF model: $w_{1,1}$, $u_{1,1}$, $u_{3,1}$, $u_{1,3}$, $u_{3,3}$, $v_{1,1}$, $v_{1,3}$, $v_{3,1}$, $v_{3,3}$; convergence of this model has been shown in Ref. [10]. The following generalized coordinates have been used for Model C, giving a 39 DOF model: $w_{1,1}$, $w_{3,1}$, $w_{5,1}$, $w_{7,1}$, $w_{1,3}$, $w_{3,3}$, $w_{5,3}$, $w_{7,3}$, $w_{1,5}$, $w_{3,5}$, $w_{5,5}$, $w_{1,7}$, $w_{3,7}$, $u_{2,1}$, $u_{4,1}$, $u_{6,1}$, $u_{8,1}$, $u_{2,3}$, $u_{4,3}$, $u_{6,3}$, $u_{8,3}$, $u_{2,5}$, $u_{4,5}$, $u_{6,5}$, $u_{2,7}$, $u_{4,7}$, $v_{1,2}$, $v_{3,2}$, $v_{5,2}$, $v_{7,2}$, $v_{1,4}$, $v_{3,4}$, $v_{5,4}$, $v_{7,4}$, $v_{1,6}$, $v_{3,6}$, $v_{5,6}$, $v_{1,8}$, $v_{3,8}$. Geometric imperfections are not considered in the following study.

The natural frequency of mode $(m=1,n=1)$ for different boundary conditions is given in Table 1. For Model A, which has a slow convergence of linear frequency, a model with 50 DOF has also been used for comparison. Model C has slow convergence in the case of high stiffness of rotational springs (almost clamped edges); in this case a model with 79 DOF has been used for comparison.

The effect of the stiffness k (N/m^2) of distributed in-plane springs parallel to the panel edges on natural frequencies of mode (1,1) is shown in Fig. 2, obtained with Model A with 19 DOF. Figure 2 shows that k on the order of 10^{11} N/m^2 is necessary to simulate the classical simply supported panel.

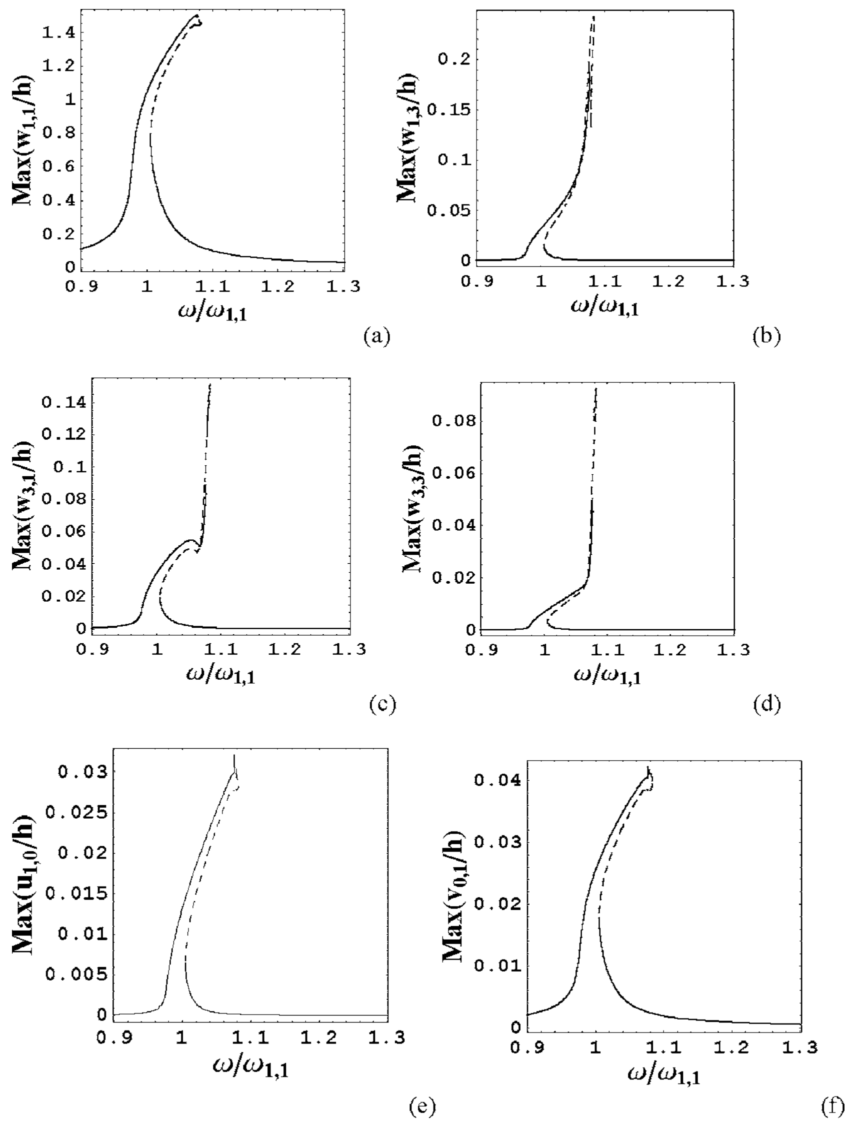


Fig. 7 Response of panel computed with model A for $k=4\times10^9$ N/m^2; fundamental mode (1,1), $f=0.021$, $\zeta_{1,1}=0.004$, 19 DOF. (—) stable periodic response; (– –) unstable periodic response. (*a*) Maximum of the generalized coordinate $w_{1,1}$; (*b*) maximum of the generalized coordinate $w_{1,3}$; (*c*) maximum of the generalized coordinate $w_{3,1}$; (*d*) maximum of the generalized coordinate $w_{3,3}$; (*e*) maximum of the generalized coordinate $u_{1,0}$; and (*f*) maximum of the generalized coordinate $v_{0,1}$.

The effect of the stiffness c (N/rad) of distributed rotational springs at panel edges on natural frequencies of mode (1,1) is shown in Fig. 3, obtained with Model C with 39 DOF. Figure 3 shows that c equal to or larger than 2×10^4 N/rad is necessary to simulate the clamped panel.

Forced vibrations of large amplitude are studied by using the software AUTO 97. The following nondimensional modal excitation on the generalized coordinate $w_{1,1}$ is introduced and its amplitude is immediately related to the harmonic point force excitation $\tilde{f}$ at $(x=\tilde{x},\ y=\tilde{y})$

$$f=\frac{\tilde{f}\sin(\pi\tilde{x}/a)\sin(\pi\tilde{y}/b)}{h^2\rho_S\omega_{1,1}^2(a/2)(b/2)}$$

Harmonic excitation of nondimensional amplitude $f=0.021$ (corresponding to $\tilde{f}=6.6$ N for the simply supported panel) has been imposed at the center of the panel in the frequency range around resonance of the fundamental mode (1, 1). In all the numerical simulations a modal damping $\zeta_{1,1}=0.004$ has been assumed.

In the present study, comparison with previously published results is not shown because Model B has already been validated in Ref. [10] by comparison with results by Kobayashi and Leissa [11] for the same panel studied here. Moreover, Model A has been satisfactorily compared to experimental results for two different panels in Ref. [12]. For clamped panels (Model C for $c\to\infty$) no results have been found for isotropic shells. A study showing a comparison with theoretical results obtained by using a completely different approach based on the R-function method [17] is under development.

Figure 4 shows the effect of the boundary conditions on the nonlinear response (generalized coordinate $w_{1,1}$ only, which is far from the most significant) of the panel considering Models A and B. In fact, three different boundary conditions are compared: classical simply supported (Model B, see Ref. [10]) versus Model A for $k=0$ and $k=4\times10^9$ N/m^2. The simply supported panel for mode (1,1) presents a significant softening nonlinearity while Model A with free in-plane edges ($k=0$) presents hardening nonlinearity; the case for $k=4\times10^9$ N/m^2 obviously lies inbetween.

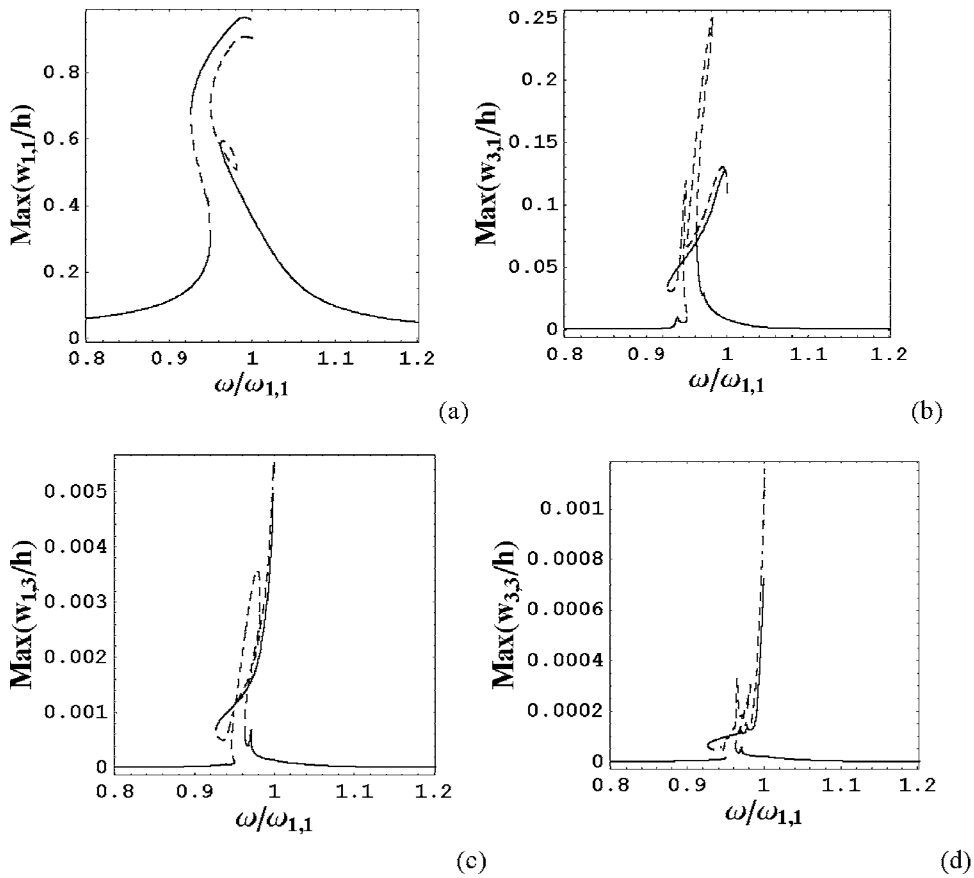


Fig. 8 Response of panel with fixed edges computed with model C for $c=0$; fundamental mode (1,1), $f=0.021$, $\zeta_{1,1}=0.004$, 39 DOF. (—) stable periodic response; (– –) unstable periodic response. (*a*) Maximum of the generalized coordinate $w_{1,1}$; (*b*) maximum of the generalized coordinate $w_{3,1}$; (*c*) maximum of the generalized coordinate $w_{1,3}$; and (*d*) maximum of the generalized coordinate $w_{3,3}$.

Comparison of Figs. 4(*a*) and 4(*b*) also show asymmetric behavior of the oscillation outward (maximum) and inwards (minimum) with respect to the center of the curvature of the panel. In particular, this asymmetric behavior is enhanced for the simply supported panel.

Figure 5 shows the effect of the boundary conditions on the nonlinear response (in this case, the transverse displacement w at the center of the panel is shown, due to the significant modal interaction for Model C) of the panel considering Models B and C. The panel with fixed in-plane edges and free rotation (Model C, $c=0$) initially presents an enhanced softening-type nonlinearity with respect to the simply supported panel, turning to hardening type for the larger value of the vibration amplitude. However, the behavior of the panel with clamped edges (Model C, $c=5\times10^4$ N/rad) is completely different and always shows a relatively strong hardening-type nonlinearity. Also for Model C asymmetric behavior of the oscillation outward and inward is observed. Response for Model C with $c=0$ presents a peculiar loop due to internal resonances and it will be discussed later in this section.

The convergence of the solution for clamped panel (Model C, $c=5\times10^4$ N/rad) versus different number of generalized coordinates retained in the expansion is shown in Fig. 6. In particular, three models are compared: 24, 27, and 39 DOF; the 27 DOF model has: $w_{1,1}$, $w_{3,1}$, $w_{5,1}$, $w_{1,3}$, $w_{3,3}$, $w_{5,3}$, $w_{1,5}$, $w_{3,5}$, $w_{5,5}$, $u_{2,1}$, $u_{4,1}$, $u_{6,1}$, $u_{2,3}$, $u_{4,3}$, $u_{6,3}$, $u_{2,5}$, $u_{4,5}$, $u_{6,5}$, $v_{1,2}$, $v_{3,2}$, $v_{5,2}$, $v_{1,4}$, $v_{3,4}$, $v_{5,4}$, $v_{1,6}$, $v_{3,6}$, $v_{5,6}$. The 24 DOF model has: $w_{1,1}$, $w_{3,1}$, $w_{5,1}$, $w_{1,3}$, $w_{3,3}$, $w_{1,5}$, $u_{2,1}$, $u_{4,1}$, $u_{6,1}$, $u_{2,3}$, $u_{4,3}$, $u_{6,3}$, $u_{2,5}$, $u_{4,5}$, $u_{6,5}$, $v_{1,2}$, $v_{3,2}$, $v_{5,2}$, $v_{1,4}$, $v_{3,4}$, $v_{5,4}$, $v_{1,6}$, $v_{3,6}$, $v_{5,6}$. The three models give a very close trend of hardening-type nonlinearity results but, especially for the 24 DOF model, the amplitude of the response of the generalized coordinate $w_{1,1}$ is different. In fact, for the model with 24 DOF only six out-of-plane coordinates are included in the model, instead of 13 (39 DOF) and 9 (27 DOF); this result indicates that, even if $w_{5,3}$, $w_{3,5}$, $w_{5,5}$ do not give significant contribution to the trend of nonlinear response, they absorb significant energy from the excitation. It can be observed that the 27 and 39 DOF models present a nonclassical response with a strange tip, due to internal resonances.

The six main generalized coordinates associated to the panel response given in Fig. 4 for Model A with $k=4\times10^9$ N/m^2 are reported in Fig. 7; only $w_{1,1}$, $w_{1,3}$, and $w_{3,1}$ have significant amplitude. An internal resonance 4:1 between $w_{1,1}$ and $w_{3,1}$ is detected very close to the peak of the response. This a typical phenomenon for panels: subharmonic resonances in parametrically excited circular cylindrical panels have been previously detected by Baumgarten and Kreuzer [18].

The four main generalized coordinates associated to the panel response given in Fig. 5 for Model C with $c=0$ (fixed edges) are reported in Fig. 8; only $w_{1,1}$ and $w_{3,1}$ have significant amplitude. An internal resonance 3:1 between $w_{1,1}$ and $w_{3,1}$ is detected for excitation frequency close to 0.94 $\omega_{1,1}$ giving a characteristic loop with unstable response in Fig. 8(*a*) (this is not an accuracy problem in the numerical calculation).

The five main generalized coordinates associated to the panel response given in Fig. 5 for Model C with $c=5\times10^4$ N/rad (practically clamped edges) are reported in Fig. 9; all the coordinates

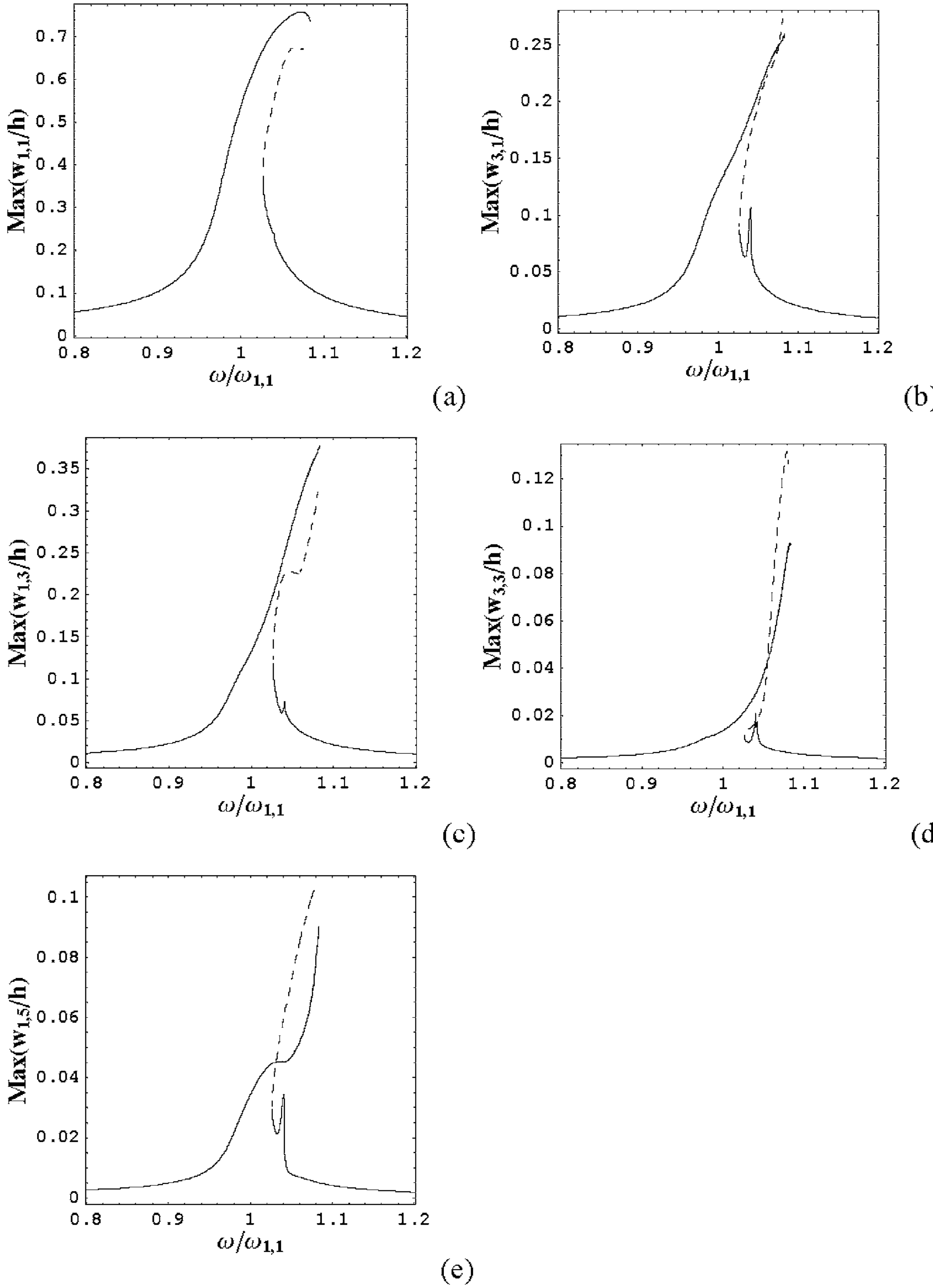


Fig. 9 Response of panel with fixed edges computed with model C for $c=5\times10^4$ N/rad (practically clamped); fundamental mode (1,1), $f=0.021$, $\zeta_{1,1}=0.004$, 39 DOF. (—) stable periodic response; (– –) unstable periodic response. (*a*) Maximum of the generalized coordinate $w_{1,1}$; (*b*) maximum of the generalized coordinate $w_{3,1}$; (*c*) maximum of the generalized coordinate $w_{1,3}$; (*d*) maximum of the generalized coordinate $w_{3,3}$; and (*e*) maximum of the generalized coordinate $w_{1,5}$.

have significant amplitude in this case. An internal resonance 3:1 between $w_{1,1}$ and $w_{3,1}$ is detected close to 1.04 $\omega_{1,1}$ giving a secondary peak in Fig. 9(*b*).

The time response computed for Model C with $c=5\times10^4$ N/rad (practically clamped edges) has been plotted in Fig. 10 for excitation frequency $\omega=1.06\ \omega_{1,1}$, i.e., close to the peak of the response; these results have been obtained by direct integration of the equations of motion by using the DIVPAG routine of the Fortran library IMSL, while all the previous ones have been obtained by using AUTO 97. Results show that the generalized coordinate $w_{1,1}$ has a harmonic response almost without superharmonics, but with significant translation inward. Other generalized

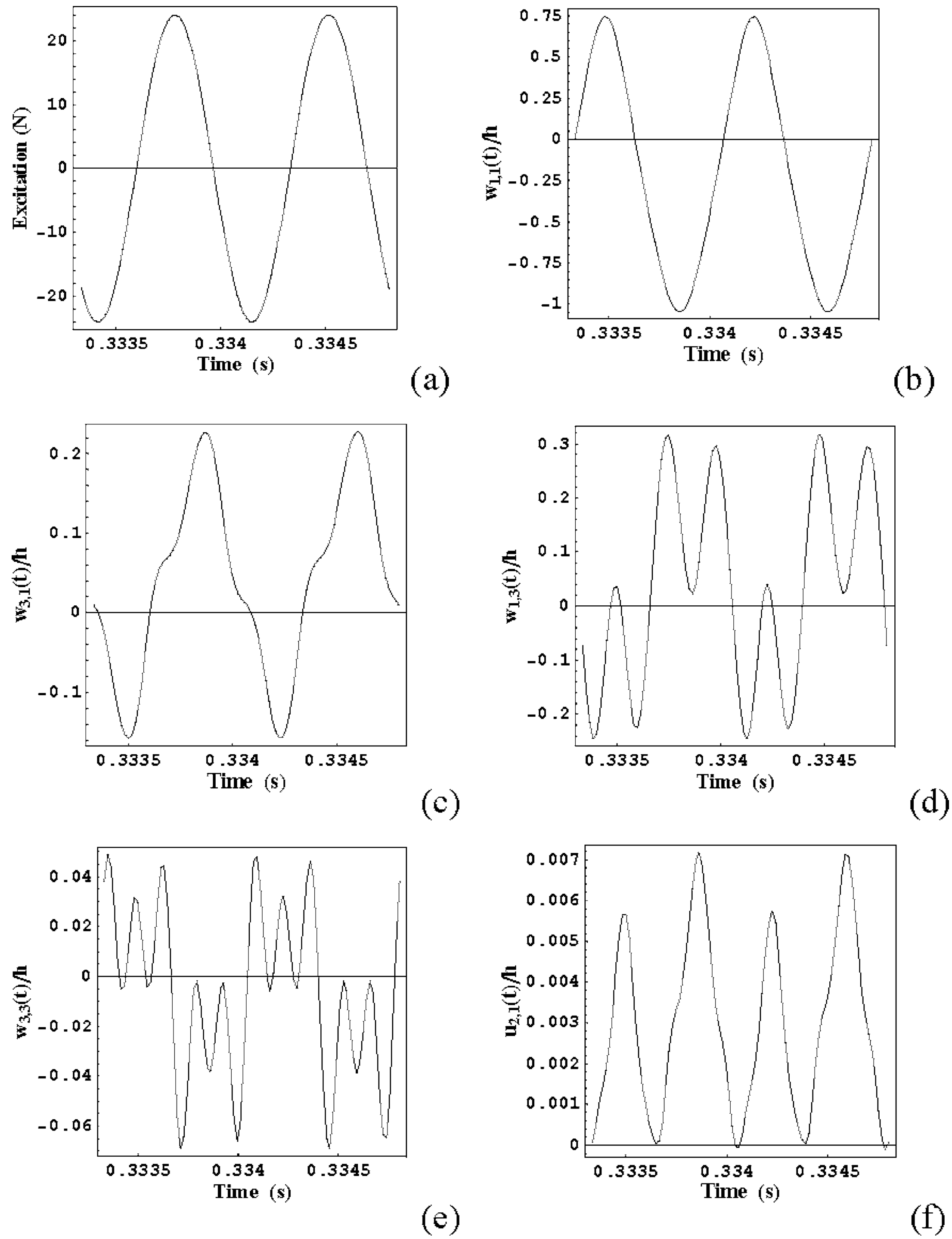


Fig. 10 Computed time response of the panel with fixed edges computed with model C for $c=5\times10^{4}$ N/rad (practically clamped) for excitation frequency $\omega=1.06\omega_{1,1}$; fundamental mode (1,1), $f=0.021$, $\zeta_{1,1}=0.004$, 39 DOF. (*a*) Force excitation; (*b*) Generalized coordinate $w_{1,1}$; (*c*) generalized coordinate $w_{3,1}$; (*d*) generalized coordinate $w_{1,3}$; (*e*) generalized coordinate $w_{3,3}$; and (*f*) generalized coordinate $u_{2,1}$.

coordinates present large superharmonics. Figure 10 also indicates the phase relationship with respect to the excitation. The presence of superharmonics and zero-frequency (mean value) component is clarified in Fig. 11 with the frequency spectra.

7 Nonperiodic Response

The same shell studied in Sec. 6 with free in-plane boundary conditions (Model A with $k=0$, $\zeta_{1,1}=0.004$) is considered here and the 19 DOF model is used. Poincaré maps have been computed by direct integration of the equations of motion, as described in Sec. 5. The excitation frequency has been kept constant, $\omega=\omega_{1,1}$ (linear resonance condition), and the excitation amplitude has been varied between 0 and about 5952 N. The force range has been divided into 800 steps; 600 periods have been discharged each time the force is changed in a step in order to eliminate the transient motion. The initial condition at the first step is zero displacement and zero velocity for all the variables; at the following steps the solution at the previous step, with addition of a small

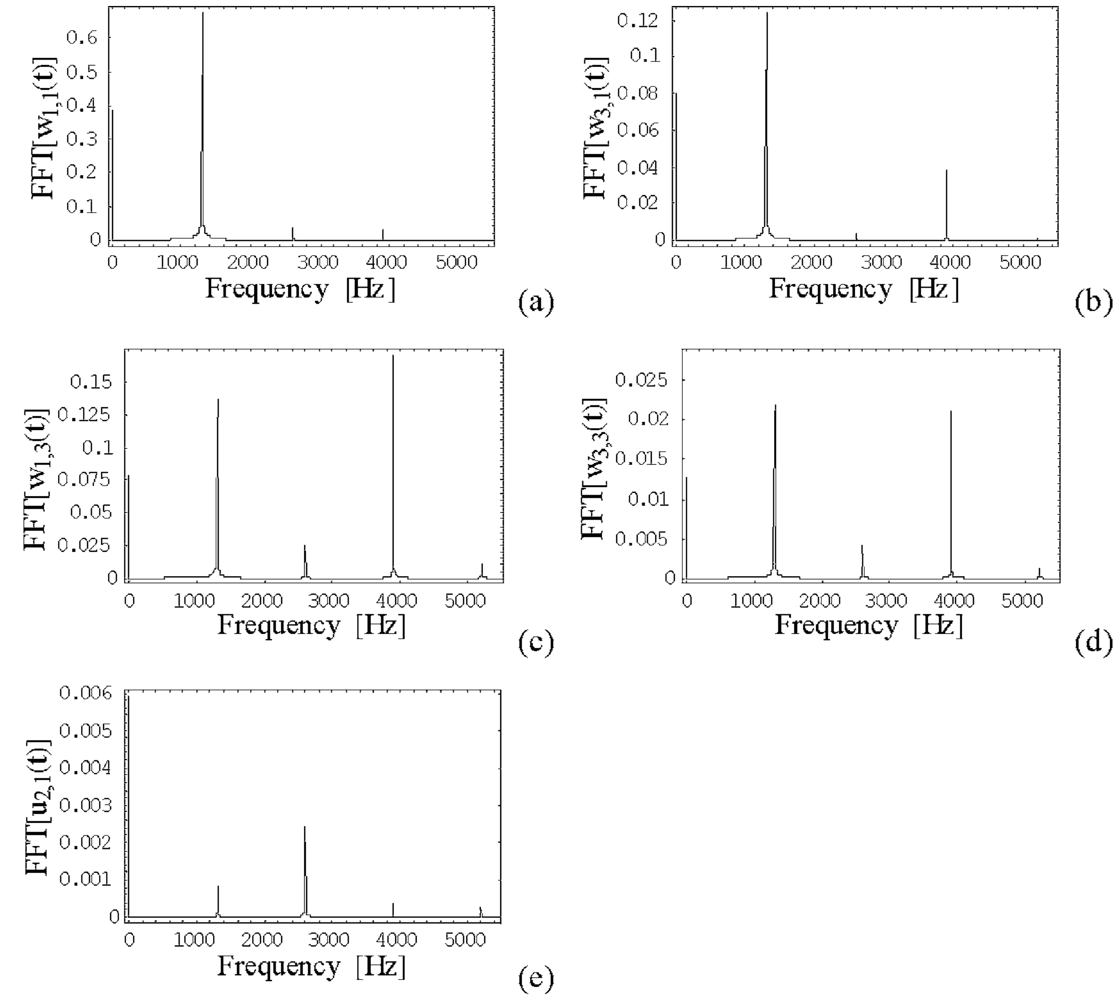


Fig. 11 Frequency spectrum of the response of the panel with fixed edges computed with model C for $c = 5 \times 10^4$ N/rad (practically clamped) for excitation frequency $\omega = 1.06\omega_{1,1}$; fundamental mode (1,1), $f = 0.021$, $\zeta_{1,1} = 0.004$, 39 DOF. (*a*) Generalized coordinate $w_{1,1}$; (*b*) generalized coordinate $w_{3,1}$; (*c*) generalized coordinate $w_{1,3}$; (*d*) generalized coordinate $w_{3,3}$; and (*e*) generalized coordinate $u_{2,1}$.

perturbation in order to find a stable solution, is used as the initial condition. The bifurcation diagram obtained by all these Poincaré maps is shown in Fig. 12 where the load is decreased from 5952.4 N to 0. Simple periodic motion, subharmonic response, amplitude modulations and chaotic response have been detected, as indicated in Figs. 12(*a*) and 12(*c*)–12(*e*). This shows very rich and complex nonlinear dynamics of the circular cylindrical shallow shell subject to large harmonic excitation. In particular, for excitation of 5952.4 N a chaotic response is obtained, which is transformed into a subharmonic response with a period nine times the excitation period around 5200 N. Then the response becomes quasi-periodic (amplitude modulation), returning to a simply periodic response around 4850 N. Around 4500 N there is a period-doubling bifurcation, clearly visible in Fig. 12(*c*), after which the response again becomes simply periodic. In the range between 2500 and 800 N several regions of quasi-periodic response appear. A period-doubling bifurcation is detected around 670 N, which ends in a chaotic region at 580 N. After that, a simply periodic oscillation is detected, with a final jump to the undisturbed configuration at zero excitation.

Figure 12(*b*) gives the maximum Lyapunov exponent σ_1 associated to the bifurcation diagram. It can be easily observed that: (i) for periodic forced vibrations $\sigma_1 < 0$; (ii) for amplitude modulated response (quasi-periodic) $\sigma_1 = 0$; and (iii) for chaotic response $\sigma_1 > 0$. Therefore σ_1 can be conveniently used for identification of the system dynamics. A three-dimensional representation of the bifurcation diagram for the generalized coordinate $w_{1,1}$ in the displacement, velocity, and load space is shown in Fig. 12(*e*), while in Figs. 12(*a*), 12(*c*), and 12(*d*) the bifurcation diagrams have been projected on a plane orthogonal to the velocity axis.

The study of the complete spectrum of the Lyapunov exponents is often reported only for simpler systems [19]. In the present case, all 38 Lyapunov exponents have been evaluated for the case with excitation $\tilde{f} = 5952.4$ N, corresponding to chaotic response, and are given in Fig. 13. In this case, four positive Lyapunov exponents have been identified, allowing us to classify this response as hyperchaos. The Lyapunov dimension [15] in this case is $d_L = 27.56$. The shape of the curve joining the exponents in Fig. 13 is nearly antisymmetric. This is not surprising because the system has small damping and for conservative systems the curve can be proved to be exactly antisymmetric. Moreover, most of the exponents have a very similar value (slightly negative), which can be related to damping; they form a characteristic nearly horizontal segment.

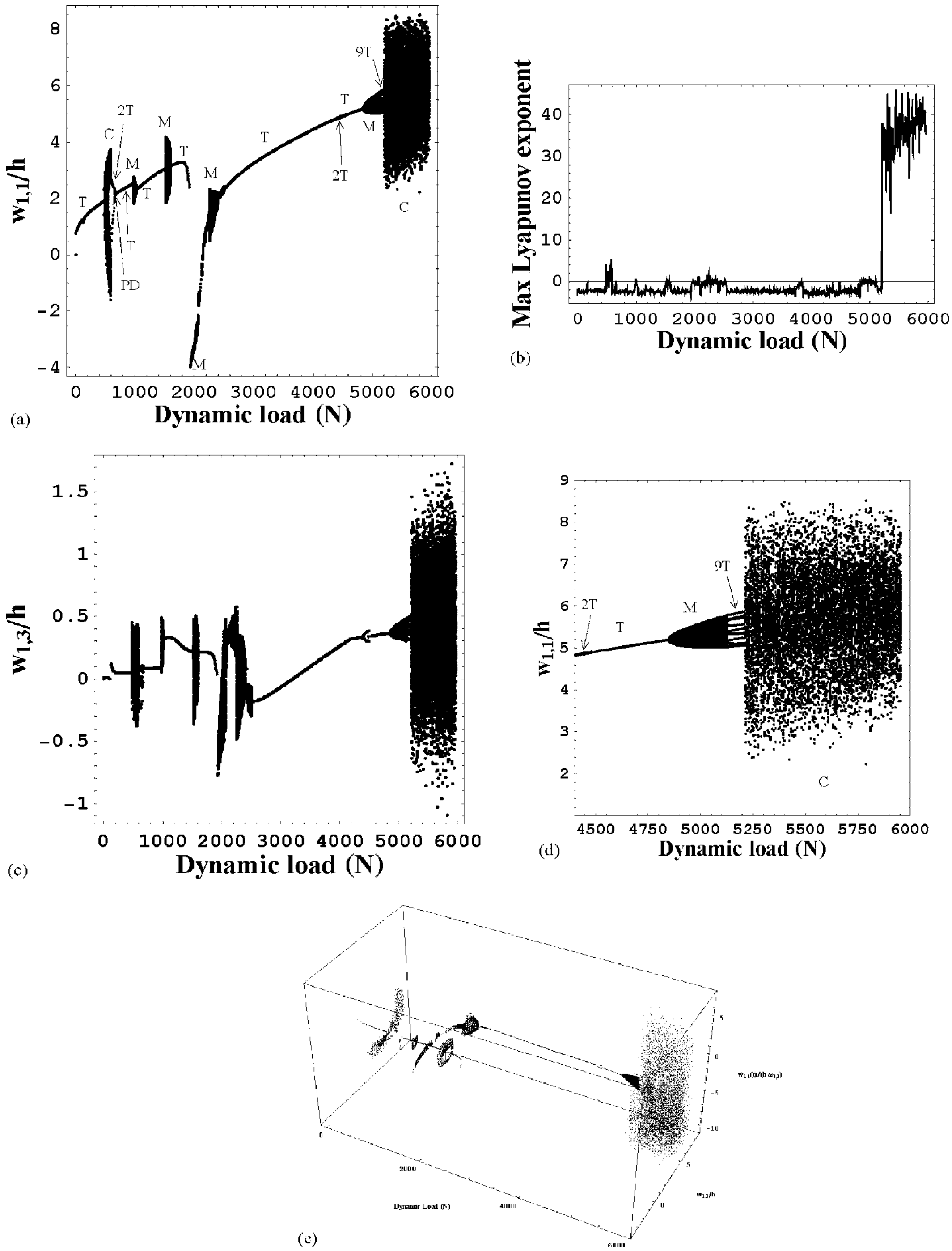


Fig. 12 Bifurcation diagram of Poincaré maps and maximum Lyapunov exponent for the panel with in-plane free edges (model A with $k=0$) under decreasing harmonic load $\tilde{f}$ with frequency $\omega=\omega_{1,1}$ (linear resonance condition); $\zeta_{1,1}=0.004$; 19 DOF model. (*a*) Bifurcation diagram: generalized coordinate $w_{1,1}$; T=response period equal to excitation period; $2T$=periodic response with two times the excitation period; $9T$=periodic response with nine times the excitation period; (PD) period-doubling bifurcation; (*M*) amplitude modulations; (*C*) chaos; (*b*) maximum Lyapunov exponent; (*c*) bifurcation diagram: generalized coordinate $w_{1,3}$; (*d*) bifurcation diagram: generalized coordinate $w_{1,1}$, enlarged scale; and (*e*) 3D representation of the bifurcation diagram: generalized coordinate $w_{1,1}$.

8 Conclusions

In the present study, for the first time the effect of boundary conditions on the trend of nonlinearity of circular cylindrical panels is studied. This is of particular interest because results show that a panel with the same geometry presents a significant softening type nonlinearity if simply supported or with fixed edges (intermediate cases), while it has a relatively strong, hardening type nonlinearity for free in-plane edges and for clamped edges (extreme cases).

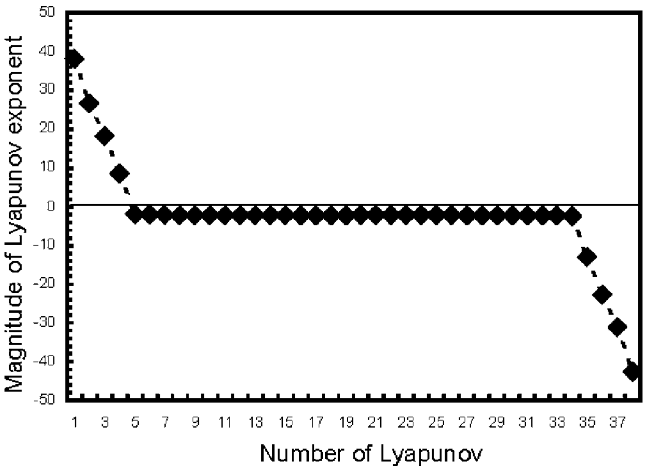


Fig. 13 All the 38 Lyapunov exponents for the panel with in-plane free edges (model A with $k=0$); excitation frequency $\omega = \omega_{1,1}$ (linear resonance condition); $\tilde{f}=5952.4$ N; $\zeta_{1,1}=0.004$, 19 DOF model.

For the specific boundary condition of free in-plane edges, complex nonlinear dynamics is also investigated by using bifurcation diagrams from direct time integration and calculation of the Lyapunov exponents and the Lyapunov dimension. Hyperchaos has been detected, confirming the result of Yamaguchi and Nagai [8] for a circular cylindrical shell with different dimensions (two times thicker) and similar boundary conditions. In Ref. [8] a smaller dimension of the model (8 DOF) was used and different excitation (acceleration excitation with very different frequency, in the subharmonic region); only two positive Lyapunov exponents were detected for specific excitation, with the smaller of the two being 245 times smaller than the bigger one, i.e., almost negligible in magnitude. In the present case, all four positive Lyapunov exponents are of the same order of magnitude giving rise, without any doubt, to hyperchaos with quite large dimension.

Acknowledgment

This work was partially supported by Grant Nos. FIRB 2001 and COFIN 2003 of the Italian Ministry for University and Research (MiUR).

References

[1] Reissner, E., 1955, "Nonlinear Effects in Vibrations of Cylindrical Shells," Ramo-Wooldridge Corporation, Rep. No. AM5-6.

[2] Grigolyuk, E. I., 1955, "Vibrations of Circular Cylindrical Panels Subjected to Finite Deflections," Prikl. Mat. Mekh., **19**, pp. 376–382 (in Russian).

[3] Cummings, B. E., 1964, "Large-Amplitude Vibration and Response of Curved Panels," AIAA J., **2**, pp. 709–716.

[4] Leissa, A. W., and Kadi, A. S., 1971, "Curvature Effects on Shallow Shell Vibrations," J. Sound Vib., **16**, pp. 173–187.

[5] Amabili, M., and Païdoussis, M. P., 2003, "Review of Studies on Geometrically Nonlinear Vibrations and Dynamics of Circular Cylindrical Shells and Panels, With and Without Fluid-Structure Interaction," Appl. Mech. Rev., **56**, pp. 349–381.

[6] Chia, C. Y., 1987, "Nonlinear Vibration and Postbuckling of Unsymmetrically Laminated Imperfect Shallow Cylindrical Panels With Mixed Boundary Conditions Resting on Elastic Foundation," Int. J. Eng. Sci., **25**, pp. 427–441.

[7] Fu, Y. M., and Chia, C. Y., 1989, "Multi-Mode Non-Linear Vibration and Postbuckling of Anti-Symmetric Imperfect Angle-Ply Cylindrical Thick Panels," Int. J. Non-Linear Mech., **24**, pp. 365–381.

[8] Yamaguchi, T., and Nagai, K., 1997, "Chaotic Vibration of a Cylindrical Shell-Panel With an In-Plane Elastic-Support at Boundary," Nonlinear Dyn., **13**, pp. 259–277.

[9] Abe, A., Kobayashi, Y., and Yamada, G., 2000, "Non-Linear Vibration Characteristics of Clamped Laminated Shallow Shells," J. Sound Vib., **234**, pp. 405–426.

[10] Amabili, M., 2005, "Nonlinear Vibrations of Circular Cylindrical Panels," J. Sound Vib., **281**, pp. 509–535.

[11] Kobayashi, Y., and Leissa, A. W., 1995, "Large Amplitude Free Vibration of Thick Shallow Shells Supported by Shear Diaphragms," Int. J. Non-Linear Mech., **30**, pp. 57–66.

[12] Amabili, M., 2006, "Theory and Experiments for Large-Amplitude Vibrations of Circular Cylindrical Panels With Geometric Imperfections," J. Sound Vib., **298**, pp. 43–72.

[13] Wolfram, S., 1999, *The Mathematica Book*, 4th ed., Cambridge University Press, Cambridge, UK.

[14] Doedel, E. J., Champneys, A. R., Fairgrieve, T. F., Kuznetsov, Y. A., Sandstede, B., and Wang, X., 1998, "AUTO 97: Continuation and Bifurcation Software for Ordinary Differential Equations (With Homcont)," Concordia University, Montreal, Canada.

[15] Argyris, J., Faust, G., and Haase, M., 1994, *An Exploration of Chaos*, North-Holland, Amsterdam, The Netherlands.

[16] Amabili, M., 2005, "Non-Linear Vibrations of Doubly Curved Shallow Shells," Int. J. Non-Linear Mech., **40**(5), pp. 683–710.

[17] Kurpa, L., Pilgun, G., and Ventsel, E., 2005, "Application of the R-Function Method to Nonlinear Vibrations of Thin Plates of Arbitrary Shape," J. Sound Vib., **284**, pp. 379–392.

[18] Baumgarten, R., and Kreuzer, E., 1996, "Bifurcations and Subharmonic Resonances in Multi-Degree-of-Freedom Panel Models," Meccanica, **31**, pp. 309–322.

[19] Kisliakov, S. D., and Popov, A. A., 1991, "Lyapunov Exponents and a Strange Attractor for the Damped Mathieu Equation," Arch. Appl. Mech., **61**, pp. 49–56.

Dynamic Compressive Response of Stainless-Steel Square Honeycombs

D. D. Radford

G. J. McShane

V. S. Deshpande

N. A. Fleck[1]
e-mail: naf1@eng.cam.ac.uk

Department of Engineering,
University of Cambridge,
Trumpington Street,
Cambridge, CB2 1PZ,
United Kingdom

The dynamic out-of-plane compressive response of stainless-steel square honeycombs has been investigated for impact velocities ranging from quasi-static values to 300 ms^{-1}. *Square-honeycomb specimens of relative density* 0.10 *were manufactured using a slotting technique, and the stresses on the front and back faces of the dynamically compressed square honeycombs were measured using a direct impact Kolsky bar. Three-dimensional finite element simulations of the experiments were performed to model the response and to help interpret the experimental results. The study has identified three distinct factors governing the dynamic response of the square honeycombs: material rate sensitivity, inertial stabilization of the webs against buckling, and plastic wave propagation. Material rate sensitivity and inertial stabilization of the webs against buckling cause the front and back face stresses to increase by about a factor of two over their quasi-static value when the impact speed is increased from 0 to* 50 ms^{-1}. *At higher impact velocities, plastic wave effects cause the front face stress to increase linearly with velocity whereas the back face stress is almost independent of velocity. The finite element predictions are in reasonable agreement with the measurements.* [DOI: 10.1115/1.2424717]

Keywords: honeycomb cores, impact testing, dynamic loads, material rate dependence, dynamic buckling

1 Introduction

Several recent investigations have revealed that metallic sandwich panels have an advantage over monolithic plates of equal mass in blast resistant structural applications [1–4]. The dynamic performance of sandwich panels is strongly dependent upon their core topology and the square-honeycomb core is a promising candidate [1,2]. The present study is a combined experimental and numerical investigation of the dynamic out-of-plane compressive response of square honeycombs in a sandwich configuration.

The out-of-plane compression of aluminium hexagonal honeycombs under low speed impact is fairly well understood. In the automotive industry, the focus of attention has been on the energy absorption capacity of the hexagonal honeycomb of relative density $0.01 \leq \bar{\rho} < 0.04$ subject to impact speeds v_o below 30 ms^{-1}. At these speeds, the dynamic strength enhancement of the honeycombs is primarily due to inertial stabilization of the honeycomb webs against elastic buckling: the honeycombs are in axial equilibrium and, consequently, the forces on the impacted and distal ends are approximately equal; see for example Zhao and Gary [5] and Wu and Jiang [6]. There exists little experimental data on the dynamic compression of honeycombs at high speeds ($v_o > 100$ ms^{-1}) apart from the investigation by Harrigan et al. [7] on the dynamic crushing of aluminium hexagonal honeycombs of relative density $\bar{\rho} \approx 1\%$. They measured the stresses on the impacted end of these specimens and concluded that plastic wave effects lead to an elevation in the dynamic plateau strength.

An optimization study by Xue and Hutchinson [2] has revealed that square honeycombs with relatively densities of about 10%, and made from stainless steel (which displays a strong strain hardening response) are promising for applications in blast resistant sandwich plates. The honeycombs of interest in such blast applications are different from those considered by Harrigan et al. [7]: they have higher relative densities and are made from a material with a high strain hardening capacity. Xue and Hutchinson [8] have recently reported finite element simulations of the dynamic response of such stainless steel square honeycombs subjected to compressive velocities in the range 20–200 ms^{-1}. Their simulations highlighted three factors contributing to the dynamic strength enhancement of the stainless steel square honeycombs: material rate sensitivity, inertial stabilization of the webs against buckling, and plastic wave propagation. For the loading rates of interest in blast applications, they argued that plastic wave propagation and plastic buckling occur over comparable timescales, and substantial axial plastic straining can occur prior to the onset of buckling.

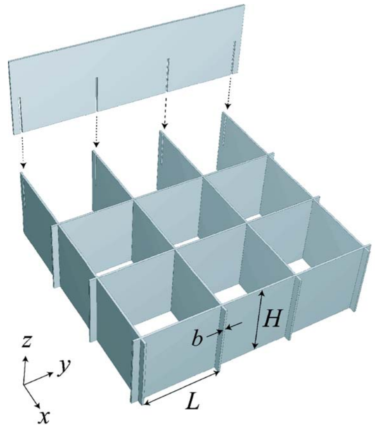
The current study is an experimental validation of the findings of Xue and Hutchinson [8], and has the following scope. The dynamic out-of-plane compressive response of square-honeycomb lattice material is investigated for applied compressive velocities ranging from quasi-static values to 300 ms^{-1}. The stresses on the impacted and distal ends of $\bar{\rho}=0.10$ stainless-steel square honeycombs are measured using a direct impact Kolsky bar, and high-speed photography is employed to observe the dynamic deformation modes. The effect of specimen height is explored by performing tests on specimens of height $H=6$ mm and 30 mm. The experimental measurements are compared with three-dimensional finite element simulations to gauge the fidelity of the simulations and to help interpret the experimental findings.

2 Experimental Investigation

2.1 Specimen Configuration and Manufacture. Square honeycombs have been manufactured from AISI type 304 stainless steel sheets of thickness $b=0.30$ mm using the technique developed by Côté et al. [9]. The sheets were cropped into rectangles of height H equal to 6 mm and 30 mm, and length 21 mm. Cross-slots (Fig. 1) of width $\Delta b=0.31$ mm, and spacing $L=6$ mm, were cut by electro-discharge machining (EDM) and the square honeycomb was assembled as sketched in Fig. 1. A clearance of 10 μm between sheet and slot allowed for easy assembly while providing

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received January 20, 2006; final manuscript received May 9, 2006. Review conducted by Robert M. McMeeking.



$L = 6$ mm
$H = 6$ mm and 30 mm
$b = 0.3$ mm

Fig. 1 Sketch of the manufacturing technique for the square honeycomb

a sufficiently tight fit to assure stability. Brazing was conducted with Ni-Cr 25-P10 (wt %) at a temperature of 1120°C in an atmosphere of dry argon at 0.03–0.1 mbar, and the braze was applied uniformly over the sheets. Capillarity draws the braze into the joints, and results in an excellent bond. All specimens comprised 3×3 cells and had dimensions 21 mm $\times$ 21 mm $\times H$. The relative density of the square-honeycomb is to first order in b/L given by

$$\bar{\rho} = \frac{2b}{L} \tag{1}$$

giving a relative density of 10%. This was confirmed by weighing the specimens.

2.2 Properties of the Constituent Materials. The uniaxial tensile response of the AISI 304 stainless steel used to manufacture the square honeycomb specimens was measured at a nominal strain rate of 10^{-3} s^{-1}, and the true tensile stress versus logarithmic strain curve is plotted in Fig. 2(*a*). This material was tested in the "as-brazed" condition to match that of the as-manufactured specimens, and the measured 0.2% offset yield strength σ_Y and ultimate tensile strength σ_{UTS} were 300 MPa and 700 MPa, respectively. Post-yield, the stainless steel exhibits a linear hardening response with a tangent hardening modulus $E_t \approx 1.4$ GPa.

Stout and Follansbee [10] have investigated the strain-rate sensitivity of the AISI 304 stainless steel for strain rates in the range 10^{-4} s$^{-1} < \dot{\varepsilon} < 10^{4}$ s^{-1}. Their data are replotted in Fig. 2(*b*), where the dynamic strength enhancement ratio R is plotted against the plastic strain rate $\dot{\varepsilon}^p$ for 10^{-3} s$^{-1} < \dot{\varepsilon}^p < 10^{4}$ s^{-1}. Here, R is defined as the ratio of the stress $\sigma_d(\varepsilon^p=0.1)$ at an applied strain rate $\dot{\varepsilon}^p$ to the stress $\sigma_0(\varepsilon^p=0.1)$ at the quasi-static rate $\dot{\varepsilon}^p = 10^{-3}$ s^{-1}. The measured stress versus strain histories presented by Stout and Follansbee [10] indicate that R is reasonably independent of the level of plastic strain ε^p at which R is calculated. Thus, the dynamic strength σ_d versus plastic strain ε^p history can be estimated from the relation

$$\sigma_d(\varepsilon^p) = R(\dot{\varepsilon}^p)\sigma_0(\varepsilon^p) \tag{2}$$

where $R(\dot{\varepsilon}^p)$ is given in Fig. 2(*b*). In the finite element simulations of the experiments presented below, we employ this prescription for the strain-rate sensitivity of the 304 stainless steel, with $\sigma_0(\varepsilon^p)$ given by the measured quasi-static ($\dot{\varepsilon}=10^{-3}$ s^{-1}) stress versus strain history (Fig. 2(*a*)). As an example, the estimated true tensile stress versus logarithmic strain histories of the AISI 304 stainless

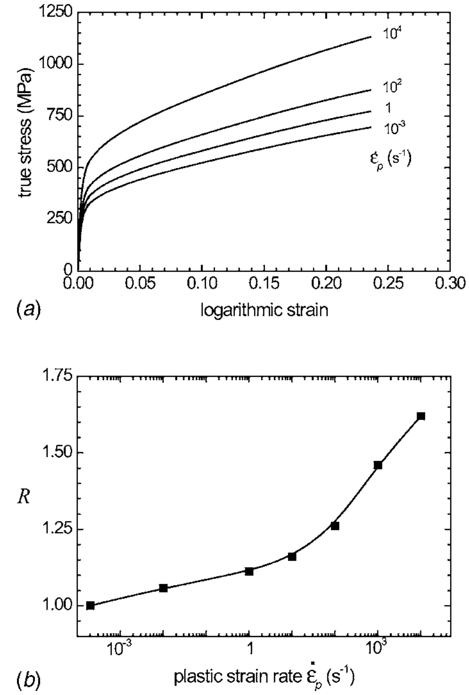


Fig. 2 (*a*) The measured quasi-static tensile stress versus strain response of the AISI 304 stainless steel and the estimated high strain rate response at three additional values of the applied strain rate using the data of Stout and Follansbee [10]. (*b*) The dynamic strength enhancement ratio *R* as a function of plastic strain rate $\dot{\varepsilon}^p$ for the AISI 304 stainless steel at a plastic strain ε^p=0.1 [10].

steel at three selected additional values of applied strain rate are included in Fig. 2(*a*).

2.3 Quasi-static Compressive Response of the Square Honeycombs. The quasi-static compressive response of the square honeycombs was measured in a screw-driven test machine at an applied nominal strain rate of 10^{-3} s^{-1}. A laser extensometer was employed to measure the average nominal compressive strain ε while the nominal stress σ was inferred from the measurements from the load cell of the test machine.

The measured out-of-plane quasi-static compressive responses of the square-honeycomb specimens are plotted in Fig. 3. The specimens display a peak strength followed by softening and, finally, rapid hardening upon densification at a strain $\varepsilon_D \approx 0.6$–0.7. Note that the H=30 mm specimen has a lower peak strength and displays more abrupt softening. Côté et al. [9] have already noted that the quasi-static collapse mode involves tor-

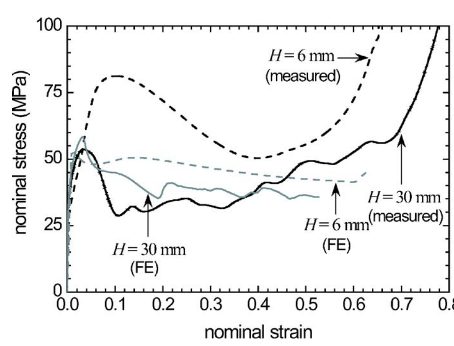


Fig. 3 Quasi-static compressive stress versus strain response of the $\bar{\rho}$=0.10 square honeycombs, of cell height *H*=6 mm and *H*=30 mm

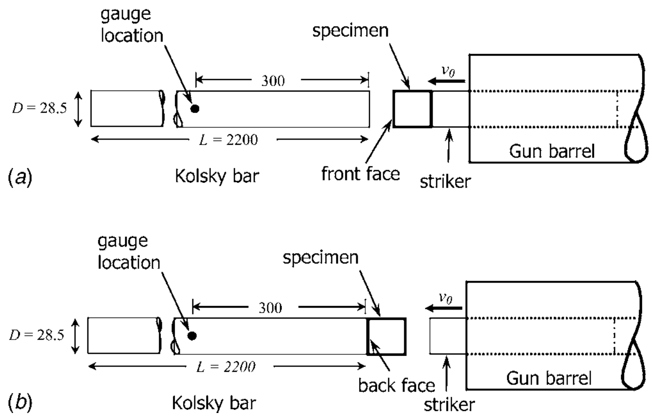


Fig. 4 Sketches of the direct impact Kolsky bar setup for measuring the stress versus time histories in (*a*) front face and (*b*) back face configurations. All dimensions are in mm.

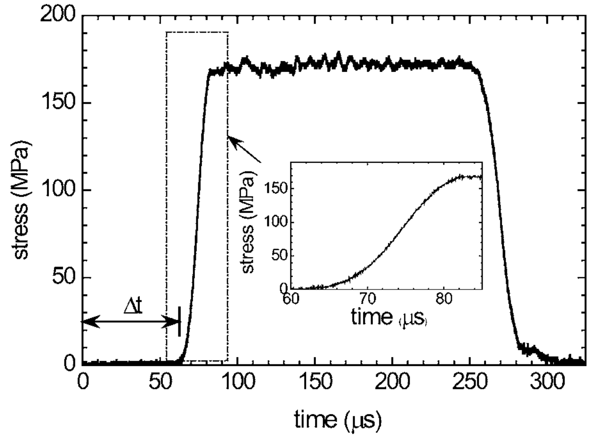


Fig. 5 Stress versus time history measured in the Kolsky bar during a calibration test. A 0.5 m long steel striker is fired at the Kolsky bar at ν_o=9.0 ms^{-1}.

sional plastic buckling of the cells, with the peak static strength σ_s accurately predicted by a plastic bifurcation analysis [9].

2.4 Dynamic Test Protocol. The dynamic out-of-plane compressive response of the square honeycombs was measured from a series of direct impact tests in which the forces on the faces of the honeycomb were measured via a strain-gauged Kolsky bar [11]. Two types of tests were conducted to measure the forces on the impacted and distal ends of the specimens, referred to subsequently as the front and back faces, respectively.

In the front face configuration (Fig. 4(*a*)), the test specimen is attached to one end of the striker bar (sometimes known as the backing mass, [7]) and the combined striker bar and specimen are fired from a gas gun so that the specimen impacts the Kolsky bar normally and centrally. In the back face configuration (Fig. 4(*b*)), the specimen is placed centrally on the stationary Kolsky bar and the striker bar is fired from the gas gun and impacts the specimen. These two independent tests allow for a measurement of the transient force on both the impacted and distal faces of the specimen.

The kinetic energy of the projected striker governs the level of compression attained and the imposed transient velocity at one end of the specimen. We wished to compress the specimens at approximately constant velocity and chose the striker masses accordingly. In the experiments conducted at low velocity ($\nu_o \leqslant 50$ ms^{-1}) and at intermediate velocity (50 ms$^{-1} < \nu_o < 200$ ms^{-1}) strikers of mass M=2.4 kg and 0.5 kg, respectively, were employed. A striker of mass M=0.090 kg sufficed for the high velocity $\nu_o \geqslant 200$ ms^{-1} experiments. The measurements and finite element simulations presented below show that these striker masses are sufficient to provide almost constant velocity compression of the square-honeycomb specimens for nominal compressive strains of up to 40%.

The striker was given the required velocity by firing it from a gas gun of barrel length 4.5 m and diameter 28.5 mm. No sabot was employed as the cylindrical striker had a diameter equal to 28.0 mm. The bursting of copper shim diaphragms formed the breech mechanism of the gun. The impact experiments were performed at velocities ranging from approximately 10 ms^{-1} to 300 ms^{-1}. The velocity of the projectile was measured at the exit of the barrel using laser-velocity gates and the impacted end of the Kolsky bar was placed 100 mm from the open end of the gun barrel.

The setup of the Kolsky pressure bar [11] is standard. A circular cylindrical bar of length 2.2 m and diameter 28.5 mm was made from the maraging steel M-300 (yield strength 1900 MPa). The pressure history on the impacted end of the bar was measured by diametrically opposite axial strain gauges placed approximately 10 diameters from the impact end of the bar. The elastic strain histories in the bars were monitored using the two 120 Ω TML foil gauges of gauge length 1 mm in a half-Wheatstone bridge configuration. A strain bridge amplifier of cutoff frequency 500 kHz was used to provide the bridge input voltage and a digital storage oscilloscope was used to record the output signal. The bridge system was calibrated dynamically over the range of strains measured during the experiments and was accurate to within 3%. The longitudinal elastic wave speed was measured at 4860 ms^{-1}, giving a time window of 800 μs before elastic reflections from the distal end of the bar complicated the measurement of stress.

The response time and accuracy of the measurement system were gauged from a series of calibration tests. We report the results of one such representative test as follows. A maraging steel striker bar of diameter 28.5 mm and length 0.5 m was fired at the Kolsky bar at a velocity ν_o=9.0 ms^{-1}. The stress versus time response measured by the strain gauges on the Kolsky bar is plotted in Fig. 5. With time t=0 corresponding to the instant of impact, the stress pulse arrives at the gauge location at t=66 μs. Elastic wave theory predicts that the axial stress in the bar is $\rho c \nu_o/2$ =175 MPa, where ρ and c are the density and longitudinal elastic wave speed of steel, respectively. The measured peak value of the stress is within 1% of this prediction. However, the measurement system has a finite response time, with the stress rising to this peak value in approximately 10 μs (see the inset in Fig. 5). This rise time places an operational limit on measuring the dynamic response of the square honeycombs. It becomes significant at the higher velocities because significant compression of the specimen is achieved within the first 10 μs. The measured stress in the calibration test drops back to zero at t=270 μs; this is the time for propagation of an axial stress wave down the bar, followed by reflection of the elastic wave from the distal end of the striker bar back to the strain gauge.

3 Experimental Results for the Dynamic Compression of Square Honeycombs

We first present the dynamic compression response of the H =30 mm square-honeycomb specimens and then contrast these measurements with those for the H=6 mm specimens.

3.1 The H=30 mm Square Honeycombs. The measured front and back face axial stress versus normalized time $\bar{t} \equiv \nu_o t/H$ histories for the H=30 mm specimens are presented in Figs. 6(*a*), 7(*a*), and 8(*a*) for impact velocities ν_o=20 ms^{-1}, 50 ms^{-1}, and 240 ms^{-1}, respectively. The time t is measured from the instant of impact and thus the normalized time $\bar{t}$ is a measure of the nominal compressive strain of the square-honeycomb speci-

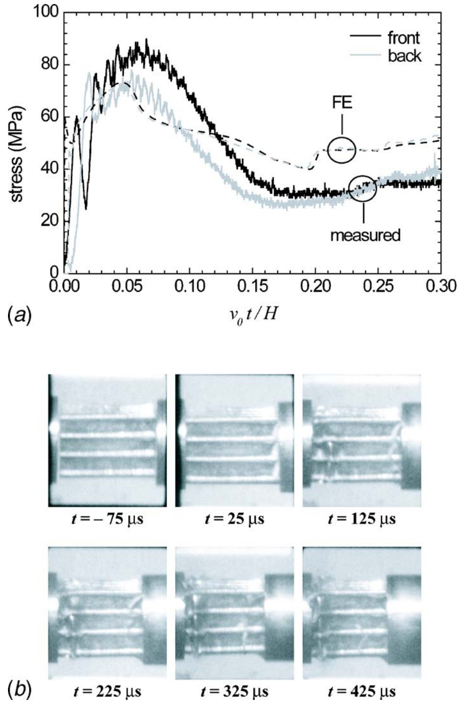


Fig. 6 (*a*) Measured front and back face stress versus normalised time histories in the *H*=30 mm honeycomb specimens impacted at ν_o=20 ms^{-1}; and (*b*) the corresponding high speed photographic sequence of the deformation in the front face configuration at an interframe time of 100 μs. The finite element predictions (constant velocity boundary condition) are included in (*a*).

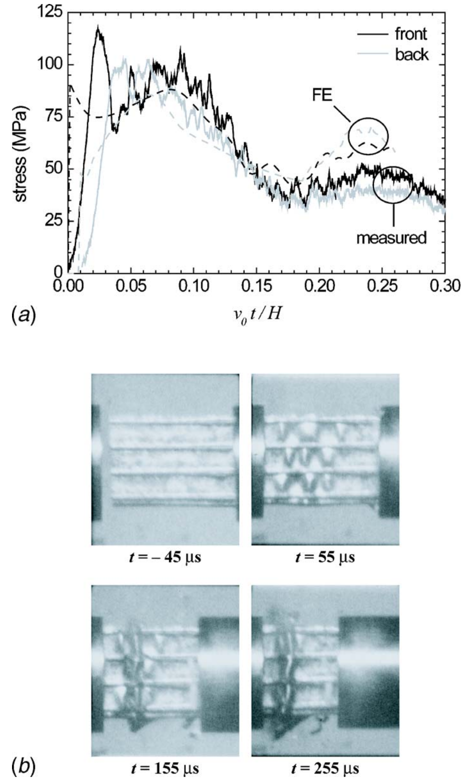


Fig. 7 (*a*) Measured front and back face stress versus normalised time histories in the *H*=30 mm honeycomb specimens impacted at ν_o=50 ms^{-1}; and (*b*) the corresponding high speed photographic sequence of the deformation in the front face configuration at an interframe time of 100 μs. The finite element predictions (constant velocity boundary condition) are included in (*a*).

mens, assuming compression at a uniform velocity ν_o over the deformation history. In these figures, the front and back face stresses are defined from the measured front face force F_f and back face force F_b as

$$\sigma_f \equiv \frac{F_f}{A_o} \tag{4a}$$

and

$$\sigma_b \equiv \frac{F_b}{A_o} \tag{4b}$$

respectively, where $A_o = 21 \times 21$ mm^2 is the cross-sectional area of the square-honeycomb specimens. High-speed photographic sequences of the deformation of the specimens in the front face configuration are given in Figs. 6(*b*), 7(*b*), and 8(*b*). The interframe times in Figs. 6(*b*) and 7(*b*) are 100 μs while the sequences in Fig. 8(*b*) were taken at 40 μs intervals. The exposure time of each photograph is 20% of the interframe time in all cases. The dust clouds in the high-speed photographs are associated with tearing of the square honeycombs along the brazed joints. This was confirmed by post-test visual inspections of the dynamically tested specimens. (Tearing of the joints was also observed in the quasi-static tests.) The dynamic measurements show two qualitatively distinct behaviors: (1) the response for ν_o=20 ms^{-1} and 50 ms^{-1}, and (2) the response for ν_o=240 ms^{-1}.

1. ν_o=20 ms^{-1} and 50 ms^{-1}. The measured front and back face stresses equalize within $t \approx 10$ μs (or $\nu_o t/H$=0.01–0.02). Recall that the response time of the measurement system is 10 μs. Thus, the early differences between the front and back face stresses are likely to be associated with the measurement system and we conclude that the specimens are in axial equilibrium over almost the entire deformation history. Similar to the quasi-static case, the square-honeycomb specimens have a distinct stress peak. These peak stresses increase with impact velocity and at ν_o=50 ms^{-1} they are approximately twice the quasi-static value. Material strain rate sensitivity alone cannot account for the increase in the peak stress and dynamic stabilization of the webs against buckling is expected to play an important role.
2. ν_o=240 ms^{-1}. The measured front face stress is approximately constant over the deformation history (to within "ringing" of the measurements) up to the specimen densification strain of $\nu_o t/H \approx 0.9$. Moreover, the front face stress exceeds the back face stress over the deformation history. This indicates that the specimen is not in equilibrium, with wave propagation effects playing a dominant role. This is substantiated by the high-speed photographs (Fig. 8(*b*)): shortening of the specimen is concentrated near the impacted end, with the distal end of the specimen undergoing only small plastic deformations.

The measured peak front face stress σ_f^p and back face stress σ_b^p are normalized by the peak quasi-static value σ_s and are plotted in Fig. 9 as a function of the impact velocity ν_o. Over the range $0 < \nu_o < 50$ ms^{-1}, the front and back face stresses remain equal and attain double their quasi-static values. For $\nu_o > 50$ ms^{-1}, the back face stress remains approximately constant at its value for ν_o =50 ms^{-1} while the front face stress increases approximately linearly with velocity. These observations suggest that the following three mechanisms provide the dynamic strength enhancements in these experiments.

1. Material strain rate sensitivity. The dynamic strength en-

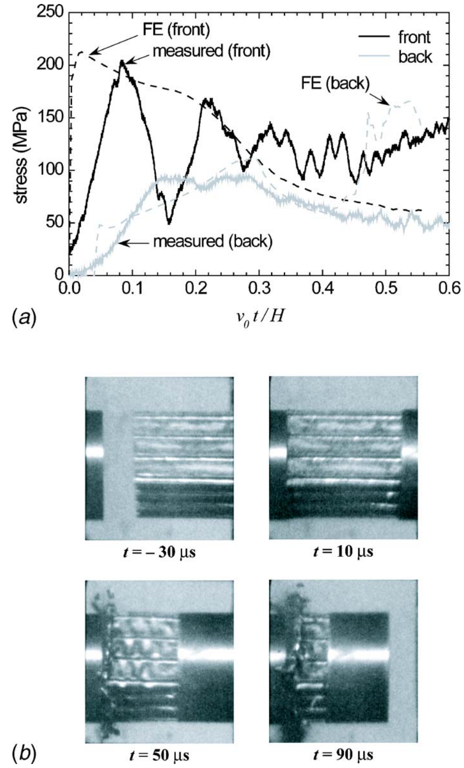


Fig. 8 (*a*) Measured front and back face stress versus normalised time histories in the *H*=30 mm honeycomb specimens impacted at ν_o=240 ms^{-1}; and (*b*) the corresponding high speed photographic sequence of the deformation in the front face configuration at an inter-frame time of 40 μs. The finite element predictions (constant velocity boundary condition) are included in (*a*).

hancement of the square honeycombs due to the material strain rate is estimated from the measurements of Stout and Follansbee [10] by assuming uniform compression of the square honeycombs at a rate $\dot{\varepsilon}^p = \nu_o/H$. The strength enhancement ratio $R(\dot{\varepsilon}^p)$ is plotted versus impact velocity in Fig. 9. Comparisons with the measured dynamic strength enhancement ratios σ_b^p/σ_s and σ_f^p/σ_s suggest that the stress enhancements for $\nu_o < 20$ ms^{-1} are largely due to the material strain rate sensitivity. Material rate sensitivity cannot, however, account for the strength enhancements at the higher velocities.

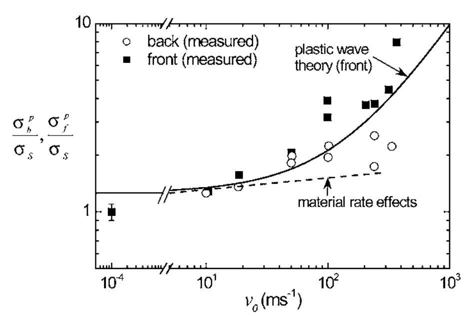


Fig. 9 Measured peak front face stress σ_f^p and back face stress σ_b^p versus impact velocity ν_o in the *H*=30 mm honeycomb specimens. The measured dynamic stresses are normalized by peak quasi-static stress σ_s from Fig. 3. The predictions of the dynamic stresses based on material strain-rate sensitivity and one-dimensional plastic wave propagation are included.

2. Dynamic buckling. As discussed in Section 2.3, the peak quasi-static strength of the square honeycombs is set by torsional plastic buckling of the cells of the square honeycomb. Under dynamic loading, inertial stabilization results in an enhanced buckling strength due to the activation of higher order buckling modes. Abrahmson and Goodier [12] and Calladine and English [13] have elaborated on this in the case for rods which have attained axial equilibrium. This higher order buckling is clearly seen in the high-speed photographs in Fig. 8(*b*). In the velocity range 20 ms$^{-1} < \nu_o < 50$ ms^{-1}, the front and back face stresses are approximately equal but material rate sensitivity alone cannot account for the dynamic strength enhancement. The additional strengthening is attributed to the enhanced dynamic buckling strength of the square honeycombs due to inertial stabilization.
3. Plastic wave propagation. At high impact velocities, a plastic wave moves along the prismatic axis of the honeycomb. Prior to the onset of buckling of the webs of the square honeycombs, one-dimensional elastic–plastic wave theory can be used to estimate dynamic stress on the front face of the honeycomb as the wave propagates through the thickness of the honeycomb. In the small strain, rate-independent limit, the front face stress is given by

$$\sigma_f = \frac{F_f}{A_o} = \sigma_Y \bar{\rho} + \rho_c c_{\rm pl}\left(\nu_o - \frac{\sigma_Y}{\rho_s c_e}\right) \approx \sigma_Y \bar{\rho} + \rho_c c_{\rm pl} \nu_o \qquad (5)$$

where c_e and c_{pl} are the elastic and plastic wave speeds, respectively, of the honeycomb parent material of yield strength σ_Y. The density of the honeycomb is $\rho_c \equiv \bar{\rho}\rho_s$ in terms of the density ρ_s of the parent material. During propagation of the plastic wave to the back face, the back face stress has no inertial contribution and is given by $\sigma_b = \sigma_Y \bar{\rho}$. It remains to specify an appropriate value for σ_Y in order to determine σ_f and σ_b. Here we arbitrarily take σ_Y to be the yield strength of the solid material at a strain rate $\dot{\varepsilon}^p = 10^4$ s^{-1} (the nominal strain rate for a honeycomb of height H=30 mm compressed at ν_o=300 ms^{-1}) in order to illustrate the predictions of the one-dimensional wave model, and plot Eq. (5) with $c_{\rm pl} \equiv \sqrt{E_t/\rho_s} \approx 410$ ms^{-1} in Fig. 9. The good agreement of the front face stress prediction with the experimental measurements for $\nu_o > 50$ ms^{-1} suggests that one-dimensional plastic wave propagation sets the peak front face stresses at these higher velocities.

3.2 The Effect of Specimen Height. The measured front and back face stress versus time histories of the H=30 mm and H=6 mm square-honeycomb specimens for an impact velocity ν_o=100 ms^{-1} are plotted in Figs. 10(*a*) and 10(*b*), respectively. Plastic wave effects play a significant role in the H=30 mm specimen with the front and back face forces equalizing only for $\nu_o t/H > 0.1$. Recall that the plastic wave speed in 304 stainless steel is $c_{\rm pl} \approx 410$ ms^{-1}. Thus, we expect the plastic wave to reach the rear face at the longer time $\nu_o t/H \approx 0.25$. The finite element calculations reported subsequently show that material rate sensitivity results in the specimen attaining axial equilibrium sooner than estimated from a rate independent plastic wave theory due to an "increased" plastic wave speed resulting from smearing out of the plastic shock. (The width of the shock front in a material with a linear viscous rate sensitivity scales as $\eta/(\rho_s c_{\rm pl})$ where ρ_s is the density of the material and η the viscosity; see Kaliski and Wlodarczyk [14].) In contrast, the front and back face stresses in the H=6 mm specimen are approximately equal over the entire deformation history. This is partly due to the smearing of the plastic shock wave and partly a result of the10 μs response time of the measurement system.

The measured peak front and back face stresses (normalized by the corresponding quasi-static peak strength) for the H=6 mm

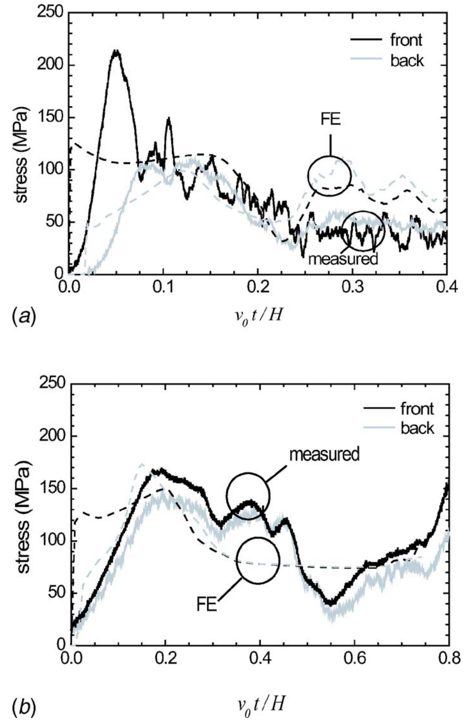


Fig. 10 The measured front and back face stress versus normalised time histories of the (*a*) *H*=30 mm; and (*b*) *H*=6 mm square-honeycomb specimens for an impact velocity v_o =100 ms^{-1}. The finite element predictions (constant velocity boundary condition) are included in the figures.

square honeycomb are plotted in Fig. 11 as a function of the impact velocity v_o. Included in Fig. 11 are estimates of the dynamic front and back face stresses based on the material strain-rate sensitivity and the front face stresses due to plastic wave propagation effects, as discussed in Sec. 3.1. (Again, σ_Y in Eq. (5) is interpreted as the yield strength of the solid material at a strain rate $\dot{\varepsilon}^p=10^4$ s^{-1}.) The basic mechanisms of dynamic strength enhancements in these specimens are similar to those in the H =30 mm specimens. However, in these shorter specimens, material rate sensitivity smears the plastic shock over the entire height (H=6 mm) of the specimen and thus the front and back face

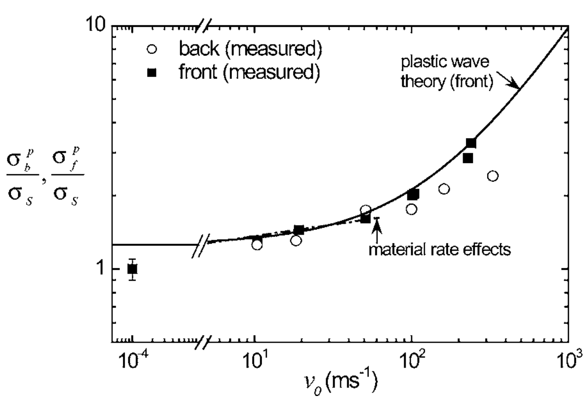


Fig. 11 Measured dynamic peak front (σ_f^p) and back (σ_b^p) face stresses in the *H*=6 mm honeycomb specimens as a function of the impact velocity v_o. The measured dynamic stresses are normalized by peak quasi-static stress σ_s from Fig. 3. The predictions of the dynamic stresses based on material strain-rate sensitivity and one-dimensional plastic wave propagation are included.

stresses are approximately equal over the range of velocities investigated here. Note that the front face stresses for the H =6 mm and H=30 mm specimens are approximately equal for velocities $v_o>50$ ms^{-1} (Figs. 9 and 11). This confirms that velocity rather than applied nominal strain rate is the governing parameter at these higher rates of compression.

4 Finite Element Investigation

A limited finite element (FE) investigation of the dynamic compression of the square-honeycomb specimens has been performed. The aims of this investigation are:

1. To determine the accuracy of three-dimensional finite element calculations in predicting the dynamic compressive response of the square honeycombs;
2. To use the finite element calculations to investigate the effect of striker deceleration on the measured dynamic response; and
3. To demonstrate the effect of the response time of the measurement system on the measured early time dynamic response of the square honeycombs.

4.1 Three-Dimensional FE Simulations of the Dynamic Compression of the Square Honeycombs. All the FE simulations were conducted using the explicit version of the commercial finite element package ABAQUS. The geometry of the honeycomb specimens was identical to that employed in the experimental investigation and the honeycombs were modeled using linear shell elements (S4R in the ABAQUS notation) with b as the thickness of the honeycomb walls. A mesh sensitivity study showed that an element size of $b/2$ sufficed to give a converged solution. All computations reported here employed such a mesh. Rigid, massless plates (discretized using four-noded rigid elements, R3D4 in the ABAQUS notation) were tied to both the front and back faces of the specimens and the general contact option in ABAQUS was employed to provide hard, frictionless contact between all surfaces in the model. The tie constraint is appropriate if negligible sliding occurs at the interface between the honeycomb and striker mass and Kolsky bar. High-speed photographs of the experiments suggest that this is an appropriate assumption for modeling purposes.

Most of the computations were conducted by compressing the specimen at a constant applied velocity as follows. The front rigid plate was constrained to move only in the axial direction (i.e., in the direction of the height H) of the specimen while the back face was fully clamped. A constant velocity v_o in the axial direction was imposed on the front rigid plate and the axial forces on the front and back faces were monitored as a function of time to determine the front and back face stress versus time history.

In a few simulations, the experimentally applied loadings in the front and back face configurations were also mimicked. In the front face configuration, a point mass M was attached to the back face. The specimen, back face and mass M were then given an initial velocity v_o in the axial direction with the front rigid plate fully clamped and the back face and point mass restricted to move only in the axial direction. Similarly, in the back face configuration, the back face was fully clamped and the point mass (now attached to the front face) was given an initial velocity v_o along with the rigid front plate. The point mass and front plate were constrained to move only in the axial direction.

4.2 Material Properties. It was assumed that the square-honeycomb specimens comprised AISI 304 stainless-steel sheets. Unless otherwise specified, the stainless steel was modeled as J2-flow theory rate dependent solid of density $\rho_f=8060$ kg m^{-3}, Young's modulus $E=210$ GPa and Poisson ratio $\nu=0.3$. The uniaxial tensile true stress versus equivalent plastic strain curves at plastic strain rates 10^{-3} s$^{-1}\leqslant\dot{\varepsilon}^p\leqslant10^4$ s^{-1} were tabulated in ABAQUS using the prescription described in Sec. 2.2 and em-

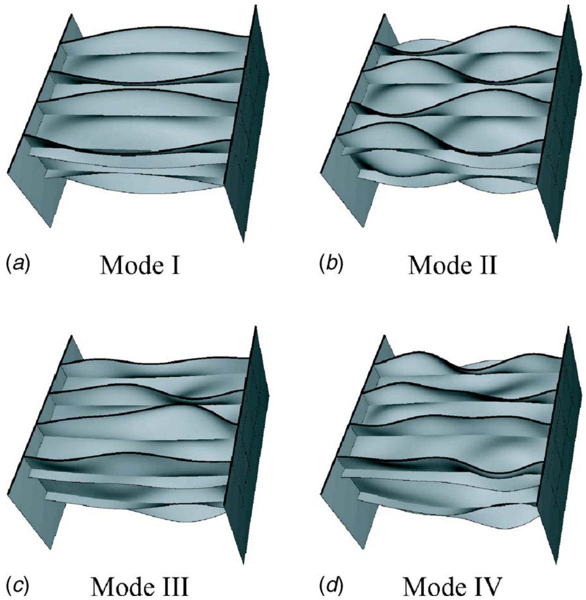


Fig. 12 The four modes of initial imperfections introduced into the FE model of the $H=30$ mm square honeycomb: (*a*) Mode I; (*b*) Mode II; (*c*) Mode III; and (*d*) Mode IV. A section through the midplane of the honeycomb is shown in each case.

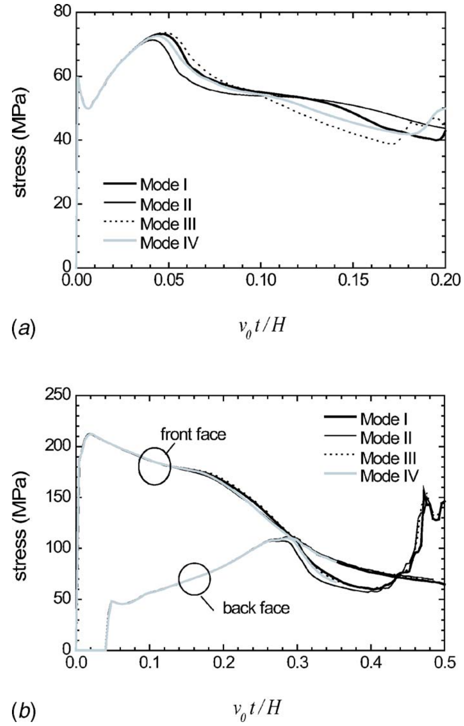


Fig. 13 FE predictions of the front and back face stress versus time histories for the $H=30$ mm compressed at a velocity (*a*) $\nu_o=20\ \mathrm{ms}^{-1}$ (front face); and (*b*) $\nu_o=240\ \mathrm{ms}^{-1}$ (front and back face). In both cases, results are shown for the four modes of initial imperfections with an imperfection amplitude 0.05*b*.

ploying the data of Fig. 2.

A comparison between the quasi-static FE predictions and measurements is included in Fig. 3; see Zok et al. [15] for details of such calculations and reasons for the discrepancies between the measurements and the FE predictions.

4.3 Effect of Initial Imperfections. The effect of initial imperfections on the finite element predictions of the dynamic compressive response was first investigated. Dynamic compression typically results in the activation of higher order buckling modes as seen in the high-speed photographs presented above. The buckling wavelengths are a strong function of the material properties and compression velocity as discussed by Abrahamson and Goodier [12] for the dynamic buckling of rods. Consequently, we investigated the effect of the magnitude and mode of imperfection on the dynamic compressive response of the square honeycombs.

Four modes of initial imperfections were considered. Modes I and II are the two lowest static eigenmodes with the Mode II wavelength half that of Mode I. In order to investigate the effect of a distribution of imperfection wavelengths we also considered modes which are the sum of the first 20 and 40 static eigenmodes (equal maximum amplitude for each mode). These will be subsequently referred to as Modes III and IV, respectively. All four modes are shown in Fig. 12 for the $H=30$ mm honeycomb.

The FE predictions of the stress versus time histories of the $H=30$ mm honeycomb for applied velocities $\nu_o=20\ \mathrm{ms}^{-1}$ and $240\ \mathrm{ms}^{-1}$ are plotted in Figs. 13(*a*) and 13(*b*), respectively. The front and back face stresses are almost equal for $\nu_o=20\ \mathrm{ms}^{-1}$ and thus for the sake of clarity only the front face stresses are shown in Fig. 13(*a*). In each case, results are plotted for the four initial imperfection modes with a maximum imperfection amplitude of 0.05*b*. We observe that the choice of the initial imperfection mode has a negligible effect on the predicted stress versus time histories. In fact, for the range of compressive velocities considered here, the FE calculations predict a response that is insensitive to the mode of the imperfections for imperfection amplitudes in the range $0.02b-0.1b$. This holds for both the $H=30$ mm and $H=6$ mm honeycombs. All the calculations reported subsequently employ the Mode I imperfection with a maximum imperfection amplitude of 0.05*b*.

4.4 Sensitivity of Response to Choice of Loading Condition. The striker mass M in the experiments was chosen to give an approximately uniform compression velocity over a nominal compressive strain of about 40%. Here we employ FE calculations to verify this and also to compare the FE predictions of the striker velocity over the deformation history of the specimens with those inferred from measurements.

FE predictions of the front and back face stresses of the $H=30$ mm square honeycomb with $\nu_o=20\ \mathrm{ms}^{-1}$ are shown in Fig. 14 for constant velocity compression and for projectile impact with a given initial velocity. In the impact simulations, a point mass $M=2.4$ kg was attached to the square honeycomb to represent the striker employed in the experiments. Both sets of FE predictions are nearly identical for $\nu_o t/H<0.2$ with small discrepancies between the two sets of simulations for $\nu_o t/H>0.2$.

We proceed to estimate the velocity reduction of the striker in the experiments and to compare these estimates with the FE pre-

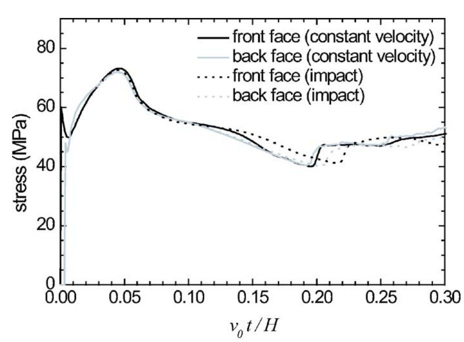


Fig. 14 A comparison between the FE predictions for a constant applied velocity and for impact boundary conditions ($H=30$ mm honeycomb specimen with $\nu_o=20\ \mathrm{ms}^{-1}$)

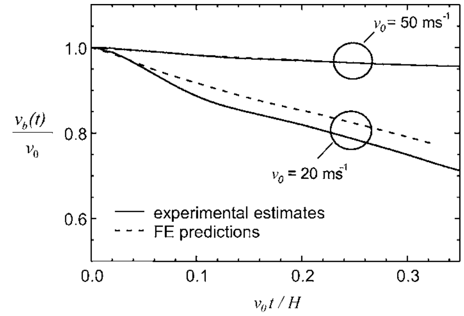


Fig. 15 Experimental estimates and FE predictions of the normalized striker velocity $\nu_b(t)/\nu_o$ as a function of the normalized time $\nu_o t/H$ for the back face configuration with initial striker velocities of ν_o=20 ms^{-1} and 50 ms^{-1}

dictions. This reduction is most severe in the back face configuration since the striker is decelerated by a force associated with driving along the plastic shock wave. The striker velocity versus time relation in the back face configuration is given by

$$\nu_b(t) = \nu_o - \frac{1}{M}\int_0^t F_f\, dt \tag{6}$$

where F_f is the measured force exerted by the square-honeycomb specimens in the corresponding front face configuration. The normalized striker velocity $\nu_b(t)/\nu_o$ (inferred from measurements) in the back face configuration of the $H=30$ mm specimens are plotted in Fig. 15 as a function of the normalized time $\nu_o t/H$ for the $\nu_o=20$ ms^{-1} and 50 ms^{-1} cases. These calculations were performed with $M=2.4$ kg to match the striker mass employed in the experiments. It is clear from Fig. 15 that, while there is a negligible reduction in velocity for the $\nu_o=50$ ms^{-1} case, the velocity of the striker reduces by about 20% for the $\nu_o=20$ ms^{-1} case at $\nu_o t/H\approx 0.35$. The corresponding FE predictions of the striker velocity in the back face configuration are included in Fig. 15 (measured directly from the velocity of the point mass in the FE calculations). The FE calculations predict slightly less velocity reductions than those estimated from the experimental measurements. This is likely to be related to the "ringing" in the experimental measurements which results in an overestimation of the front face force. Similar estimates for the other impact velocities confirmed that negligible striker velocity reductions occurred in all experiments with $\nu_o>50$ ms^{-1}. These experimental estimates and FE predictions confirm that the impact experiments can be regarded as constant velocity experiments over the practical deformation regime $0<\nu_o t/H<0.4$.

4.5 Accuracy of Predictions for the $H=30$ mm Square Honeycombs. FE predictions of the front and back face stresses on the $H=30$ mm specimens are included in Figs. 6(*a*), 7(*a*), and 8(*a*) for a constant velocity of $\nu_o=20$ ms^{-1}, 50 ms^{-1}, and 240 ms^{-1}, respectively. The simulations were terminated within the FE code when large rotations of the shell elements resulted in a loss of accuracy of the simulations. The FE simulations capture the measured stresses in the $\nu_o=20$ ms^{-1} and 50 ms^{-1} experiments to reasonable accuracy. The FE predictions of the deformation mode for the $\nu_o=50$ ms^{-1} case are plotted in Fig. 16 at $t=55$ μs and 155 μs. It is noted that the observed deformation (Fig. 7(*b*)) and predicted deformation modes (Fig. 16) are quite different; a similar discrepancy in deformation mode was observed for $\nu_o=20$ ms^{-1} although the comparison is not shown explicitly. This discrepancy is likely to be associated with the tearing of the brazed joints in the experiments (and not accounted for in the FE calculations). Important differences between the measured and FE predictions of the compressive response at $\nu_o=20$ ms^{-1} and 50 ms^{-1} are as follows.

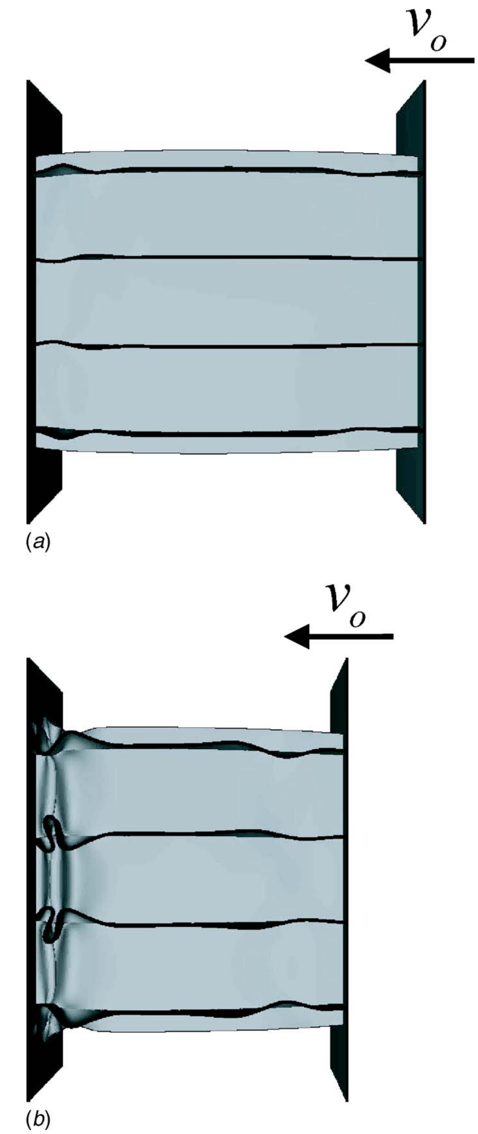
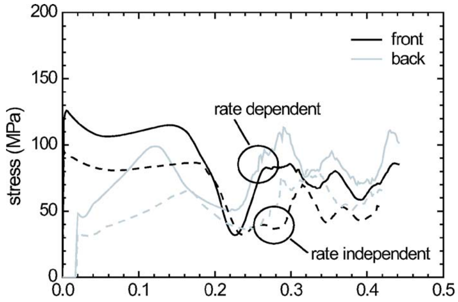


Fig. 16 The FE predictions of the deformation mode of the $H=30$ mm honeycomb specimen, subjected to a constant applied velocity ν_o=50 ms^{-1} at: (*a*) t=55 μs; and (*b*) t=155 μs. These times correspond to the times of the high speed photographs in Fig. 7(*b*). A section through the midplane of the honeycomb is shown.

1. The FE simulations predict a sharp rise of the stresses on the front and back faces due to the presence of the elastic stress wave. Such a sharp rise is not observed in the experiments. This is likely to be related to the fact that the response time of the measurement apparatus is about 10 μs and thus the experimental measurements are unable to capture the instantaneous rise in stress due to the arrival of the elastic wave.

2. At $\nu_o t/H\approx 0.20$, the FE simulations predict a sudden increase in the stresses on the back face due to contact between the cell walls as seen in Fig. 16 for the $\nu_o=50$ ms^{-1} simulation. In reality, a much gentler increase in the measured stresses occurs. We attribute this discrepancy to tearing of the brazed joints of the square honeycombs as evidenced from the dust clouds in the photograph at $t=255$ μs in Fig. 8(*b*). This effect is not included in the FE simulations and is likely to result in an overprediction of the measured stress.

The comparisons between the FE simulations and measurements for the $\nu_o=240$ ms^{-1} case (Fig. 8(*b*)) show larger discrep-



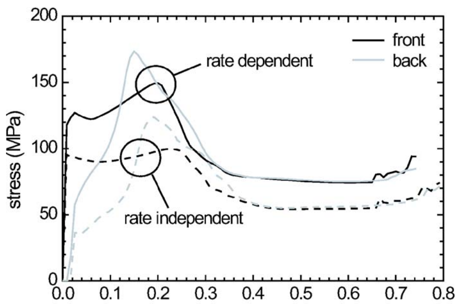


Fig. 17 Comparisons between the rate independent and rate dependent FE predictions (constant applied velocity) of the front and back face stresses: (*a*) $H=30$ mm; and (*b*) $H=6$ mm square-honeycombs, for $\nu_o=100\ \mathrm{ms}^{-1}$

ancies. The differences between the measurements and predictions of the stresses are mainly due to the slow response time of the measurement apparatus, with the predictions and measurements in reasonable agreement for $\nu_o t/H > 0.2$. The oscillations in the measured stress history are due to flexural waves induced in the Kolsky bar by slight misalignment of the impact and are neglected in the FE simulations. Again, the FE calculations predict a rise in the back face stress at $\nu_o t/H \approx 0.4$. This rise is not observed in the experiments due to tearing of the brazed joints of the specimens.

4.6 FE Simulations to Elucidate the Role of Material Rate Sensitivity. The measurements indicate that the plastic wave reaches the distal end of the specimen earlier than predicted by rate independent, small strain plasticity theory. For example, with a plastic wave speed of $c_{pl} \approx 410\ \mathrm{ms}^{-1}$ and an input speed of $\nu_o = 100\ \mathrm{ms}^{-1}$, it is predicted that the plastic wave will increase the stress on the back face to 100 MPa at $\nu_o t/H \approx 0.25$. However, the measurements in Fig. 10 suggest that the plastic wave arrives much earlier, at $\nu_o t/H \approx 0.1$. The FE predictions are in reasonable agreement with the experimental measurements for $t > 10\ \mu$s, that is for $\nu_o t/H > 0.01$ and 0.03 for the $H=30$ mm and 6 mm specimens, respectively. This confirms that the high stress measurements over $0.1 < \nu_o t/H < 0.25$ are not an artefact of the measurement system but are probably related to an increased plastic wave speed associated with the rate sensitivity of the 304 stainless steel.

We proceed to check the effect of material rate sensitivity by performing FE simulations with the 304 stainless steel modeled as a rate independent J2 flow theory solid with a uniaxial tensile stress versus plastic strain response given by the $\dot{\varepsilon}^p = 10^{-3}\ \mathrm{s}^{-1}$ data in Fig. 2(*a*). A comparison between the rate independent and rate dependent FE predictions (constant applied velocity) of the front and back face stresses on $H=30$ mm and 6 mm square honeycombs is shown in Figs. 17(*a*) and 17(*b*), respectively, for $\nu_o = 100\ \mathrm{ms}^{-1}$. The rate dependent and rate independent FE calculations predict that the plastic wave arrives at the back face at $\nu_o t/H \approx 0.15$ and $\nu_o t/H \approx 0.2$, respectively. Thus, the FE calculations demonstrate that material strain rate sensitivity can account for the increased plastic wave speed. The dynamic strength enhancement due to material rate effects are also clearly seen in Fig. 17: the rate dependent FE calculations predict that both the front and back face stresses are about 25% higher than those in the rate independent case.

5 Concluding Remarks

Square-honeycomb specimens of relative density $\bar{\rho}=0.10$ and heights $H=30$ mm and 6 mm were manufactured by slotting together 304 stainless-steel sheets and then brazing together the assembly. The out-of-plane compressive response of these specimens was measured for velocities ranging from quasi-static values to $\nu_o=300\ \mathrm{ms}^{-1}$. The stresses on both the front and back faces of the square honeycombs were measured in the dynamic tests using a direct impact Kolsky bar.

Torsional plastic buckling of the webs is the collapse mode under quasi-static loading. Three distinct mechanisms govern the dynamic response of the square honeycombs: (i) material rate sensitivity; (ii) inertial stabilization of the webs against buckling; and (iii) plastic wave propagation. In the $H=30$ mm specimens, effects (i) and (ii) are dominant for $\nu_o < 50\ \mathrm{ms}^{-1}$ with the front and back face stresses increasing by about a factor of two over their quasi-static values. At higher velocities, plastic wave effects become increasingly important with the back face stress remaining approximately constant at its value for $\nu_o=50\ \mathrm{ms}^{-1}$ and the front face stress increasing approximately linearly with velocity. The peak front face stresses in the $H=6$ mm and $H=30$ mm cases are approximately equal at the higher velocities indicating that velocity rather than strain rate governs the response of these specimens under high rates of compression.

Three-dimensional finite element simulations capture the experimental measurements to reasonable accuracy. Discrepancies between the measurements and predictions are attributed to: (i) a response time of 10 μs by the measurement apparatus; and (ii) tearing of the square honeycombs along the brazed joints. This tearing leads to a significant drop in the transmitted load and in the energy absorbed by the square honeycomb.

Acknowledgment

The authors are grateful to ONR for their financial support through US-ONR IFO Grant No. N00014-03-1-0283 on The Science and Design of Blast Resistant Sandwich Structures and to the Isaac Newton Trust, Trinity College Cambridge.

References

[1] Fleck, N. A., and Deshpande, V. S., 2004, "The Resistance of Clamped Sandwich Beams to Shock Loading," J. Appl. Mech., **71**, pp. 386–401.
[2] Xue, Z., and Hutchinson, J. W., 2004, "A Comparative Study of Blast-Resistant Metal Sandwich Plates," Int. J. Impact Eng., **30**, pp. 1283–1305.
[3] Radford, D. D., Fleck, N. A., and Deshpande, V. S., 2005, "The Response of Clamped Sandwich Beams Subjected to Shock Loading," Int. J. Impact Eng., **32**, pp. 968–987.
[4] McShane, G. J., Radford, D. D., Deshpande, V. S., and Fleck, N. A., 2006, "The Response of Clamped Sandwich Plates With Lattice Cores Subjected to Shock Loading," Eur. J. Mech. A/Solids, **25**, pp. 215–229.
[5] Zhao, H., and Gary, G., 1998, "Crushing Behaviour of Aluminium Honeycombs Under Impact Loading," Int. J. Impact Eng., **21**, pp. 827–836.
[6] Wu, E., and Jiang, W-S., 1997, "Axial Crush of Metallic Honeycombs," Int. J. Impact Eng., **12**, pp. 439–456.
[7] Harrigan, J. J., Reid, S. R., and Peng, C., 1999, "Inertia Effects in Impact Energy Absorbing Materials and Structures," Int. J. Impact Eng., **22**, pp. 955–979.
[8] Xue, Z., and Hutchinson, J. W., 2005, "Crush Dynamics of Square-Honeycomb Sandwich Cores," submitted.

[9] Côté, F., Deshpande, V. S., Fleck, N. A., and Evans, A. G., 2004, "The Out-of-Plane Compressive Behaviour of Metallic Honeycombs," Mater. Sci. Eng., A, **380**, pp. 272–280.
[10] Stout, M. G., and Follansbee, P. S., 1986, "Strain-Rate Sensitivity, Strain Hardening, and Yield Behaviour of 304L Stainless Steel," J. Eng. Mater. Technol., **108**, pp. 344–353.
[11] Dharan, C. K. H., and Hauser, F. E., 1970, "Testing Techniques Based on the Split Hopkinson Bar," Exp. Mech., **10**, pp. 370–376.
[12] Abrahamson, G. R., and Goodier, J. N., 1966, "Dynamic Flexural Buckling of Rods Within an Axial Plastic Compression Wave," J. Appl. Mech., **33**, pp. 241–247.
[13] Calladine, C. R., and English, R. W., 1984, "Strain-Rate and Inertia Effects in the Collapse of Two Types of Energy-Absorbing Structure," Int. J. Mech. Sci., **26**, pp. 689–701.
[14] Kaliski, S., and Wlodarczyk, E., 1967, "On Certain Closed-Form Solutions of the Propagation and Reflection Problem of an Elastic-Visco-Plastic Wave in a Bar," Arch. Mech. Stosow., **19**, pp. 433–452.
[15] Zok, F. W., Rathbun, H., He, M., Ferri, E., Mercer, C., McMeeking, R. M., and Evans, A. G., 2005, "Structural Performance of Metallic Sandwich Panels With Square-Honeycomb Cores," Philos. Mag., **85**, pp. 3207–3234.

The Analysis of Tensegrity Structures for the Design of a Morphing Wing

Keith W. Moored

Hilary Bart-Smith[1]
e-mail: hb8h@virginia.edu

Department of Mechanical and Aerospace Engineering,
University of Virginia,
Charlottesville, VA, 22904

Current attempts to build fast, efficient, and maneuverable underwater vehicles have looked to nature for inspiration. However, they have all been based on traditional propulsive techniques, i.e., rotary motors. In the current study a promising and potentially revolutionary approach is taken that overcomes the limitations of these traditional methods—morphing structure concepts with integrated actuation and sensing. Inspiration for this work comes from the manta ray (Manta birostris) and other batoid fish. These creatures are highly maneuverable but are also able to cruise at high speeds over long distances. In this paper, the structural foundation for the biomimetic morphing wing is a tensegrity structure. A preliminary procedure is presented for developing morphing tensegrity structures that include actuating elements. To do this, the virtual work method has been modified to allow for individual actuation of struts and cables. The actuation response of tensegrity beams and plates are studied and results are presented. Specifically, global deflections resulting from actuation of specific elements have been calculated with or without external loads. Finally, a shape optimization analysis of different tensegrity structures to the biological displacement field will be presented.
[DOI: 10.1115/1.2424718]

Keywords: tensegrity, morphing wing, actuation, force density, biomimetics, optimization

1 Introduction

The family Myliobatidae can achieve large amplitude flapping type of locomotion and have been observed traveling at speeds greater than 1 m/s over long distances. It is these characteristics that make them attractive to study and mimic. By mimicking the movements of these species, a new underwater vehicle design is explored. The goal of this research is to develop a structure that can propel an underwater vehicle with the swift and silent motions of the manta ray. To achieve this goal, a lightweight control surface, manipulated by an active tensegrity structure, with high out-of-plane stiffness and a large range of motion under large restraining moments is being studied. Tensegrity structures are comprised of a set of discontinuous compressed struts held together with a continuous web of tensioned cables. They offer high strength to mass ratios, low mechanical wear in dynamical applications, and high deformability with minimal input energy, which makes these systems excellent candidates for the structural layout of a morphing wing. Actuation of the structure is achieved by replacing passive cables and struts with actuators. Using these structures has the potential to create a new generation of highly efficient, maneuverable air and sea vehicles. Steps towards designing and building a highly deformable and versatile morphing wing, while keeping a high enough stiffness to withstand environmental forces and perturbations, are presented.

2 Tensegrity Background

Around 1963, tensegrity structures (Fig. 1) were originally developed by Emmerich, Fuller, and Snelson, with Fuller coining the word tensegrity as a contraction of the words "tensional integrity." In recent years, tensegrity structures became of engineering interest as their potential in load bearing applications was realized, but still today these structures have not been used in many practical circumstances. To create usable structures researchers have devoted much time to the problem of form finding, which is a procedure used to determine the spatial layout of the structure.

Initial efforts by Fuller [1], Snelson [2], and Kenner [3] focused on using geometrical techniques to solve the problem of form finding. However, the internal self-stress forces of the members must be taken into account in order to have a correct theoretical model for form finding. Pellegrino [4] showed for some polyhedra that the geometric form-finding techniques were not accurate when compared to a physical model. As a result of this discrepancy several methods have been developed to accurately predict the form of a tensegrity structure. They can be categorized into two main groups: (i) kinematical methods and (ii) statical methods [5]. The *kinematical* group includes analytical, nonlinear optimization, and dynamic relaxation techniques. These methods either keep the struts lengths constant while shrinking the cables lengths or vice versa, mimicking the physical assembly of a tensegrity structure. The analytical methods give solutions for n-fold symmetric structures, i.e., prismatic tensegrities. The optimization and relaxation methods can handle generalized structures, but they become computationally intensive when asymmetries or many nodal points are involved. The *statical* techniques encompass analytical solutions, the force density method, the energy minimization method, and the reduced coordinates method. Again the analytical solutions are only viable for simple cases. The force density method, first develop by Schek [6], gives a set of linear equilibrium equations that can analyze large structures as well as asymmetric tensegrities. The energy minimization method is similar to the force density method, however the goal is to find the equilibrium configuration by finding the minimum potential energy state of the members. The reduced coordinates method is an approach that derives the equilibrium equations from the principle of virtual work, giving a model that has more control than the force density or energy minimization methods but requires more extensive calculations. Recently, Masic [7] has developed a form-finding procedure—based on the force density method—that gives

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received January 16, 2006; final manuscript received July 19, 2006. Review conducted by Robert M. McMeeking.

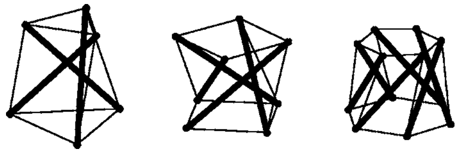


Fig. 1 Three strut, four strut, and six strut tensegrity unit cell structures

adequate control as well as quick computational times. Masic's procedure takes the force density method a step further by adding shape constraints to the structure, allowing one to manipulate the entire shape of the structure. This adapted method presents an opportunity to develop active structures, where the desired morphologies are achieved through the changes in lengths of possibly all of the members.

3 Methods

Before the tensegrity static equilibrium equations are presented some variables and operators must be defined.

DEFINITION 1. *A nodal point*, ν_k, $k=1,\ldots,n_n$, *where* n_n *is the number of nodes*, *is defined as a point where compressive members and tensile members connect. The vector* $\mathbf{p}=[\mathbf{x}^T,\mathbf{y}^T,\mathbf{z}^T]^T$ *is defined as the vector of nodal point locations which is decomposed into the* $\mathbf{x}$, $\mathbf{y}$, *and* $\mathbf{z}$ *coordinates of the nodal points*, *where* $\mathbf{x}\in\mathbb{R}^{nn\times 1}$, $\mathbf{y}\in\mathbb{R}^{nn\times 1}$, $\mathbf{z}\in\mathbb{R}^{nn\times 1}$, *and* $\mathbf{p}\in\mathbb{R}^{3nn\times 1}$.

DEFINITION 2. *Element* $e_i=[\nu_k,\nu_j]$, $k\neq j$, $i=1,\ldots,n_e$ *radiates from node* ν_k *and terminates at node* ν_j. *The direction of* e_i *is arbitrary*, *but once the direction is chosen for a given set of elements*, *then they must be used consistently*.

DEFINITION 3. *The cable connectivity matrix*, $\mathbf{C}_{\mathbf{cables}}\in\mathbb{R}^{nn\times n\text{cables}}$, *is*

$$C_{\text{cables},ji}=\begin{cases}0, & \textit{if } e_i \textit{ does not connect to } \nu_j\\ 1, & \textit{if } e_i \textit{ terminates at } \nu_j\\ -1, & \textit{if } e_i \textit{ radiates from } \nu_j\end{cases}, \quad \begin{array}{c}\textit{where } i=1,\ldots,n_{\text{cables}}\\ j=1,\ldots,n_n\end{array}$$

The strut connectivity matrix, $\mathbf{C}_{\mathbf{struts}}\in\mathbb{R}^{nn\times n\text{struts}}$, is

$$C_{\text{struts},ji}=\begin{cases}0, & \textit{if } e_i \textit{ does not connect to } \nu_j\\ 1, & \textit{if } e_i \textit{ terminates at } \nu_j\\ -1, & \textit{if } e_i \textit{ radiates from } \nu_j\end{cases}, \quad \begin{array}{c}\textit{where } i=1,\ldots,n_{\text{struts}}\\ j=1,\ldots,n_n\end{array}$$

The one-dimensional connectivity matrix, $\mathbf{C}^1\in\mathbb{R}^{nn\times ne}$, is

$$\mathbf{C^1}=[-\mathbf{C}_{\mathbf{cables}}\quad \mathbf{C}_{\mathbf{struts}}] \tag{1}$$

The connectivity matrix, $\mathbf{C}\in\mathbb{R}^{3nn\times 3ne}$, is

$$\mathbf{C}=\begin{bmatrix}\mathbf{C^1} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C^1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{C^1}\end{bmatrix} \tag{2}$$

DEFINITION 4. *The one-dimensional force density vector*, $\lambda^1\in\mathbb{R}^{ne\times 1}$, *is*

$$\lambda_i^1=\frac{f_i}{L_i}=E_iA_i\left(\frac{1}{L_{m,i}}-\frac{1}{L_i}\right) \tag{3}$$

E is the Young's modulus; A is the area of the member; L_m is the unstressed manufacturing length of the member; and L is the final equilibrium length of the member. L is a function of nodal point positions, $\mathbf{p}$, and is the length of a member that is in static equilibrium with the other members of the structure. The force density vector, $\lambda\in\mathbb{R}^{3ne\times 1}$, is

$$\boldsymbol{\lambda}=\begin{bmatrix}\boldsymbol{\lambda}^1\\ \boldsymbol{\lambda}^1\\ \boldsymbol{\lambda}^1\end{bmatrix} \tag{4}$$

DEFINITION 5. *The operator* (^) *is a vector operator that diagonalizes a vector*

$$\hat{\mathbf{x}}=\text{diag}([x_1,x_2,\ldots,x_n]^T)=\begin{bmatrix}x_1 & 0 & 0 & 0\\ 0 & x_2 & 0 & 0\\ 0 & 0 & \ddots & 0\\ 0 & 0 & 0 & x_n\end{bmatrix} \tag{5}$$

These definitions are similar to those presented by Masic in Ref. [7], but they differ due to the lack of member identifiers presented by Masic that describe whether a member is in compression or tension. By not having member identifiers, negative force densities are found for the compressive members and positive force densities for the tensile members. However, for the purposes of this paper the identifiers are not necessary and are therefore not presented in this formulation.

All of the statical form-finding methods for tensegrity structures find a set of equilibrium equations that are either determined by summing forces acting on a structure or using potential energy considerations. The virtual work method (VWM) uses energy considerations and the principle of virtual work to derive the equilibrium equations. The derivation is outlined in Ref. [8]. To obtain the set of equilibrium equations used in this work the virtual work method was employed. Once the set of nonlinear algebraic equilibrium equations are obtained they can be represented in a compact matrix form as the following

$$\mathbf{C}\hat{\boldsymbol{\lambda}}(\mathbf{p})\mathbf{C}^{\mathbf{T}}\mathbf{p}=\mathbf{f}_{\mathbf{ext}} \tag{6}$$

This constitutes a set of $3n_n$ unknowns, $\mathbf{p}$, with the same number of nonlinear equations. This set of equations can be solved numerically using Matlab's *fsolve* function. Since these equations are in Cartesian coordinates, it is now simple to constrain any node to a desired value. In doing this, as can be seen from the virtual work approach, equations that are differentiated by a constrained coordinate are removed. This theoretical model gives control over all of the elements in the structure. When determining the form of a tensegrity structure using Eqs. (6), one must first set the external forces to zero to obtain the following set of equations

$$\mathbf{C}\hat{\boldsymbol{\lambda}}(\mathbf{p})\mathbf{C}^{\mathbf{T}}\mathbf{p}=0 \tag{7}$$

Solving the equilibrium equations with the forces equal to zero guarantees that the structure has adequate self-stress to keep its structural integrity after the external forces have been removed.

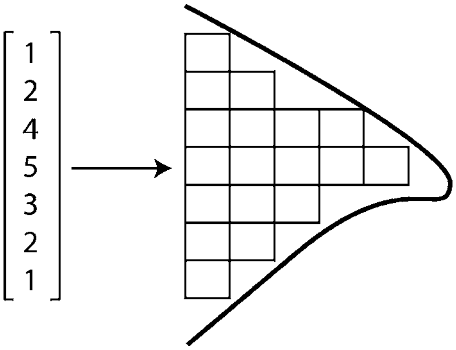


Fig. 2 The configuration vector describes the structural layout of a plate tensegrity structure composed of unit cells. Each square represents a unit cell.

For a more detailed study of the equilibrium equations and the feasibility conditions see Ref. [8].

4 Geometric Construction of a Tensegrity Plate

The initial analysis into the design of a morphing tensegrity wing examines a simple unit cell, specifically a four-strut prismatic structure (Fig. 1). Based on this system, cells are connected together to form a beam, with bar-to-bar connections between the unit cells. Instead of recreating the connectivity matrix whenever the number of cells in the beam is changed, a general connectivity matrix has been developed for any number of cells in a beam made of four-strut bar to bar tensegrities—commonly referred to as a Type 2 tensegrity structure. The generalized connectivity matrix for a Type 2 beam composed of four strut unit cells can be found in Ref. [8]. A cellular plate structure is a structure composed of many unit cells connected together that extend in two different directions. Plate structures can have many different planform shapes defined by the configuration vector (Fig. 2). If the tensegrity plate in composed of four strut unit cells with bar-to-bar connections the whole structure is considered a Type 4 structure since some nodal points have four struts connecting to them. To achieve the type of bar-to-bar connections described, the ratio of the radii of the bottom of the unit cell structure to the top of the unit cell structure must be equal to the square root of two. In order to give the structure a desired level of prestress and to solve for the initial manufacturing lengths of the symmetric unit cell, a simple force balance can be employed and is outlined in Ref. [9].

5 Actuation Mechanism

Based upon the VWM, a technique has been developed to calculate the overall topology of a tensegrity structure that has the ability to actuate strings and/or bars in an asymmetric reconfiguration. In this method, the manufacturing length becomes the actuation variable, so that, for example, a prescribed actuation strain of 20% is defined as a change in the manufacturing length of 20%. It should be pointed out that the final equilibrium length of the cable—after actuation—will not be exactly 20% different from the initial equilibrium length, due to second-order effects.

6 Optimization

The VWM is of great use in determining the global displacement field of a tensegrity structure under external loads with the actuation of individual members. By using the VWM as the foundation for the analysis, a more useful design method has been developed. Up to this point the question asked has been; what is the displacement field due to the actuation of an individual member? Instead, the question that is addressed by the following optimization routine is the inverse; which actuators are necessary to reach a given displacement field? To answer this question a direct search method known as *patternsearch* in Matlab is utilized. This method generates a mesh around an initial point, and the algorithm tests each point for a better functional value than the initial point (Fig. 3). Once a better value is achieved a mesh is generated around that point and the process repeats itself until the optimization function value has converged to a minimum. The variables for the optimization routine are the manufacturing lengths of the cables or struts, whereas the optimization function depends on the location of the nodal points.

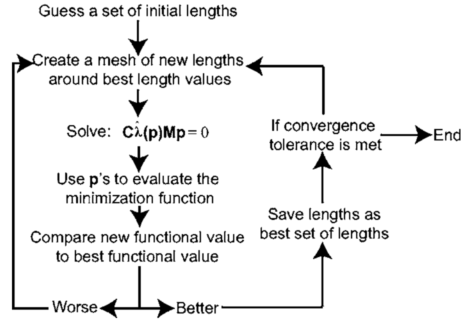


Fig. 3 Flow diagram of patternsearch optimization

The manta ray is the inspiration for designing a highly deformable morphing wing. However, data on the manta ray are rare since they are not easily kept in captivity. An alternative is to study the cownose ray, which is of the same family as the manta ray. The deflections of a ray's wing, as a function of time, is given in Fig. 4 [10]. Although the flapping motion of the cownose ray is asymmetric, these data present a good foundation for an optimization objective function.

The objective function is the difference between the nodal points of the top of the structure and the shape of the cownose ray's deflected wing. To obtain an equation for the shape of the cownose ray's wing an exponential curve was fit to the ray data. The following equation describes the upstroke of the ray

$$z = e^{0.1494x} - 1 \tag{8}$$

It has an R^2 value equal to 0.9901. To make this equation useful for a variety of structures of all different lengths and aspect ratios this curve must be scaled up or down compared to the length of the cownose ray's wing, which is approximately 23 cm. First a size ratio, S, comparing the spanwise length of the structure to the spanwise length of the cownose ray is defined

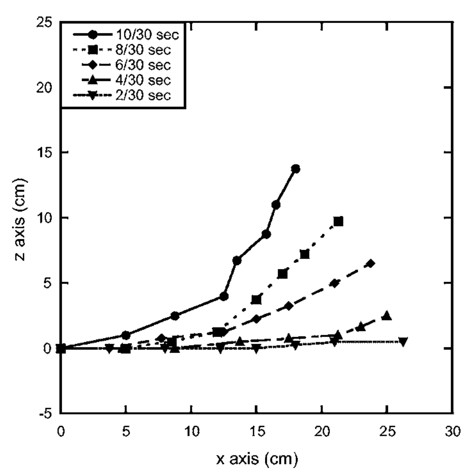


Fig. 4 Cownose ray wing curvature during a flapping cycle at different time steps. 10/30 s is the upward extreme in a normal forward propelling flapping cycle.

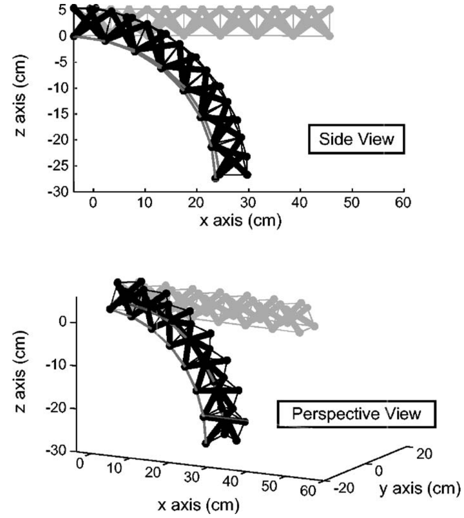


Fig. 5 61% downward deflection of a seven cell beam due to 20% contraction of the spanwise bottom cables

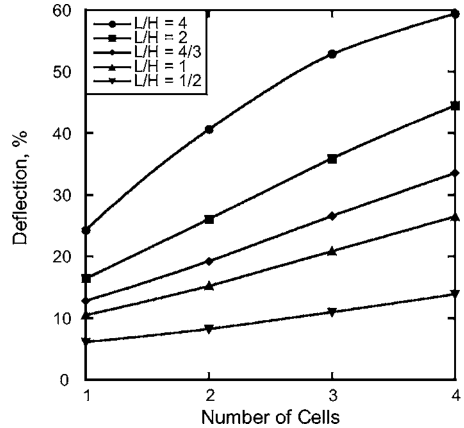


Fig. 6 Graph showing increased deflection capabilities of a beam as a function of number of cells and length to height ratio of the individual cell

$$S = \frac{L_{\text{struct}}}{L_{\text{ray}}} \tag{9}$$

A deflection ratio, D, is empirically found to be 0.6785. This ratio is measured directly from the cownose ray data. It compares the x location of the tip of the deflected wing to the x location of tip of the flat wing. This ratio gives an approximate trajectory from the flat shape to the deflected shape

$$D = \frac{x_{\text{def}}}{x_{\text{flat}}} \tag{10}$$

This allows the displacement equation to be in terms of the x locations of the initial or flat shape and not the deflected shape. By applying the S and D ratios, the curve is scaled to the size of the structure and its x values are in terms of the flat shape. Finally, the curve shifts upward to account for the initial height of the top nodes. When all of these adjustments are made to the shape equation, the following is obtained for an upstroke and downstroke, respectively

$$z = Se^{0.1494(D/S)x} - S + z_0 \tag{11}$$

$$z = -Se^{0.1494(D/S)x} + S + z_0 \tag{12}$$

Using Eqs. (11) and (12) part of the objective function is obtained. The other two parts come from the difference between the y values of the nodes and their initial y states and the x values of the top nodes and the matching deflected x values obtained from the deflection ratio. Thus the objective function is the following for a downstroke

$$\frac{1}{4}\sum_{i=1}^{n\text{top}} |x_i - Dx_{i,\text{flat}}| + \frac{1}{4}\sum_{i=1}^{n} |y_i - y_{i,0}| + \frac{1}{2}\sum_{i=1}^{n\text{top}} |z_i + Se^{0.1494(D/S)x_{i,\text{flat}}} - S_i - z_{i,0}| \tag{13}$$

The z terms of the function must be weighted more than the y terms so that the optimization does not want to converge to the initial shape—this happens because there are more y errors being computed than z errors. Also note that S is in both scalar and vector form, where the vector form is the scalar value multiplied by a ones vector, i.e., $S^{*}[\mathbf{1},\mathbf{1},\mathbf{1}\ldots\mathbf{1}]^{\mathbf{T}}$.

To extend the usefulness of the new optimization method, plate structures have also been studied. For a simple example a three cell by three cell plate structure has been examined. This Type 4 plate structure consisting of 132 members and 40 nodes is actuated into a twisted shape rather than a downward or upward deflection. In order to achieve a twist in the plate, the nodes on the tip of the plate are matched up to a certain degree of twist. The minimization function for the tip nodes is as follows

$$f_{\min} = \frac{1}{9}\sum_{i=1}^{n} |x_i - x_{i,0}| + \frac{4}{9}\sum_{i}^{n\text{tip}} |y_i - y_{\text{goal}}| + \frac{4}{9}\sum_{i}^{n\text{tip}} |z_i - z_{\text{goal}}|$$

$$y_{\text{goal}} = y_{i,0}\cos\theta$$

$$z_{\text{goal}} = y_{i,0}\sin\theta \tag{14}$$

where $x_{i,0}$ and $y_{i,0}$ are initial x and y nodal point positions. The summation from i to $n$tip implies summation over only the tip nodal points and the weights on each summation are somewhat arbitrary, but these values were given to reflect the relative importance of each goal. The angle θ is the prescribed or desired twist angle of the plate.

Either the deflection scenario or the twisting scenario can be cast into the following nonlinear optimization problem:

Given

$$\mathbf{p_{target}}, \mathbf{C}, \mathbf{L_m}, \mathbf{E}, \mathbf{A}, \mathbf{u_s}, \mathbf{u_c}, \mathbf{l_c}$$

$$\min_{\mathbf{L_m}} \quad f_{\min} = \sum |\mathbf{p} - \mathbf{p_{target}}|$$

such that

$$\mathbf{C}\hat{\boldsymbol{\lambda}}\mathbf{C}^{\mathbf{T}}\mathbf{p} = 0$$

$$\hat{\boldsymbol{\gamma}}\mathbf{p} = \hat{\boldsymbol{\gamma}}\mathbf{p_0}$$

$$\mathbf{L_{m,struts}} = \mathbf{u_s}$$

$$\mathbf{l_c} \le \mathbf{L_{m,cables}} \le \mathbf{u_c} \tag{15}$$

where

$$\mathbf{u_s} = \mathbf{L_{m,struts}} \quad \mathbf{l_c} = \hat{\boldsymbol{\alpha}}\mathbf{L_{m,cables}} \quad \mathbf{u_c} = \hat{\boldsymbol{\beta}}\mathbf{L_{m,cables}}$$

In this form-finding problem, γ is a vector of zeros and ones constraining certain nodal points to be fixed to their initial values; α is a vector of values between zero and one; and β is a vector of values between one and infinity. For most of the cases studied in this paper $\alpha = 0.8^{*}\text{ones}(n_{\text{cables}}, 1)$ and $\beta = 1^{*}\text{ones}(n_{\text{cables}}, 1)$. If a subset of strings is to be constrained from actuating, the α's corresponding to the subset can be set to zero that constrains the manufacturing lengths of the strings to stay at their initial values.

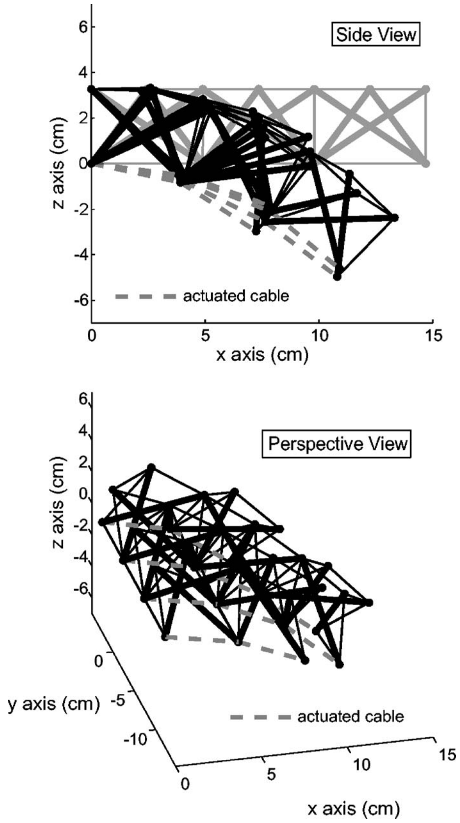


Fig. 7 34% downward deflection of a seven cell elliptical plate due to 20% contraction of the bottom spanwise cables

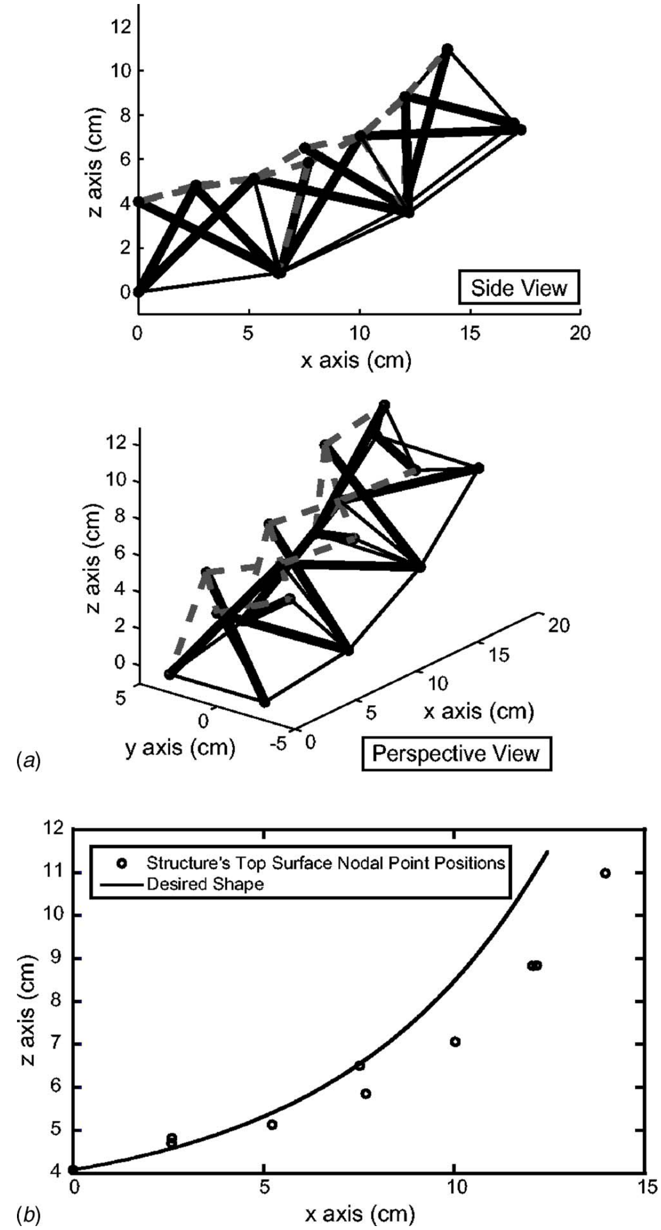


Fig. 9 (*a*) Optimal upward deflection of the unconstrained three cell beam; and (*b*) comparison of the top surface of the structure to the desired shape. With more cells or more allowed actuation strain the desired shape can be easily reached.

Results for the optimization of beams and plates to achieve deflection and twisting requirements will be presented in the following section.

7 Results

The VWM has been used to determine the global deflection of tensegrity beams and plates when individual cables are theoretically actuated. Results from the optimization scheme, developed to determine the optimal locations and contraction amounts for actuating cables, to obtain a desired displacement field, are also presented.

7.1 Beam Structures. For the multiple cell beam case, a seven cell beam was developed that consists of periodic four strut prismatic unit cells connected together bar to bar. This type of structure—based on bar-to-bar connections—is classified as a Type 2 structure. The generalized connectivity matrix was utilized to generate this beam. Three nodes are constrained at a wall such that the connected cells form a cantilever beam configuration, as shown in Fig. 5. There are no external forces acting on the structure and the bottom spanwise cables are contracted 20% each. This causes an overall downward tip deflection of 61% of the span length compared to 55% for the cownose ray, showing that this structure is capable of achieving the biological displacement field to a first-order approximation. One thing to note is that a twisting asymmetry can be seen in the final structural shape. A question that must be addressed is whether these asymmetries will be of importance in developing an actual wing. It can be seen from the seven cell beam structure that the twisting is not large, but it could have a significant effect on the fluid–structure interaction and may necessitate the need to be compensated for through additional actuation. Moreover, this asymmetry highlights the need for an op-

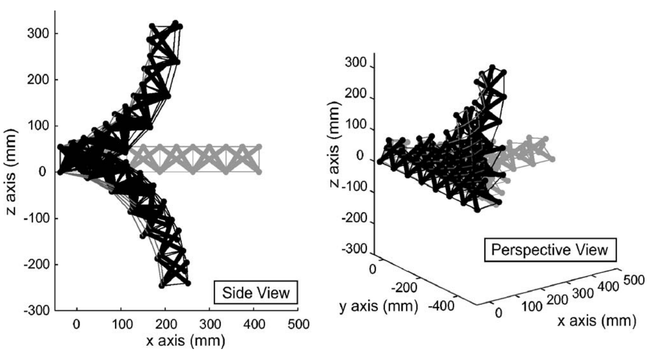


Fig. 8 63% downward and 60% upward deflection of a 19 cell manta ray shaped wing due to 20% contraction of the bottom spanwise cables and 20% contraction of the top cables, respectively

timization method that can determine which actuators to activate in order to minimize the asymmetries in the structure, while reaching the deflection goal.

Beam structures from one to seven cells in length have been studied and the tip deflection resulting from a 20% contraction of the spanwise cables have been compiled for given length to height ratios of the unit cells (Fig. 6). This shows that the addition of cells to the span will give a nonlinear increase in the maximum deflection possible for a fixed amount of contraction. Since the percent deflection is defined as the difference between the deflected tip nodal point and the initial tip nodal point the amount of percent deflection is nonlinear because the structure begins to curl in on itself. This result bodes well for future work on designing tensegrity wings, as the amount of actuation needed to achieve a given deflection decreases with increasing cells. *Deformability* is defined as the amount of tip deflection possible for a given amount of actuation. This can also be controlled by varying the amount of prestress in the structure or by varying the length to height ratio of the beam (Fig. 6).

7.2 Plate Structures. In order to create a morphing structure that has a planform resembling a ray's wing, the beam tensegrity structures must extend outward in the y direction as well as the x direction, forming a plate tensegrity structure. This structure can be thought of as a series of beam structures connected together. To construct a plate tensegrity structure, composed of individual four-strut unit cells with bar-to-bar connections, the generalized connectivity matrix for a beam structure that was previously presented can be used. However, the connections between the beams must be taken into account to construct the correct connectivity matrix. To characterize the configuration of the structure a configuration vector is prescribed, an example of which can be seen in Fig. 2. The configuration vector can be used to construct the full connectivity matrix of the plate. This is done by creating the connectivity matrix for each element of the subvector that represents a beam structure, and then compiling all of the beam connectivity matrices with the added connections between beams. The structure can then be analyzed using Eq. (7).

Two wing configurations have been studied for their actuation capabilities. The first wing configuration (Fig. 7) has an elliptical planform shape with seven four strut unit cells connected together with bar-to-bar connections, classified as a Type 4 tensegrity structure. This planform shape has a configuration vector of $[\mathbf{2}\;\;\mathbf{3}\;\;\mathbf{2}]^{\mathbf{T}}$. To determine the actuation potential of the structure, the bottom cables are contracted by the standard amount of 20% causing a 34% deflection in the–z direction.

Table 1 Errors between structural nodal points and biological data.

	Average X error (%)	Average Y error (%)	Average Z error (%)	Weighted Average error (%)
Upward unconstrained	6.26	1.11	1.06	2.11
Downward unconstrained	N/A[a]	0.75	0.29	0.4
Upward constrained	6.77	1.09	1.28	2.34
Downward constrained	N/A[a]	1.66	1.06	1.21
Designer's choice up	6.05	1.33	1.54	2.4
Designer's choice down	N/A[a]	6.65	2.47	3.52

[a]N/A=not available.

The second wing configuration (Fig. 8) has a planform shape of the cownose ray with 19 four strut cells with bar-to-bar connections, which consists of 279 members and is classified as a Type 4 tensegrity structure. This planform shape has a configuration vector of $[\mathbf{1}\;\;\mathbf{2}\;\;\mathbf{4}\;\;\mathbf{6}\;\;\mathbf{3}\;\;\mathbf{2}\;\;\mathbf{1}]^{\mathbf{T}}$. In this example the bottom cables are contracted by 20% causing a 63% downward deflection and the top cables are also actuated by 20% causing a 60% upward deflection.

The results of this analysis demonstrate the potential for these structures to mimic the kinematics of the cownose ray. However, more needs to be done to accurately mimic the biological displacement field. Moreover, if the manufacturing of one of these structures were to be made practical in terms of power consumption and cost, the structure should be designed with a minimized

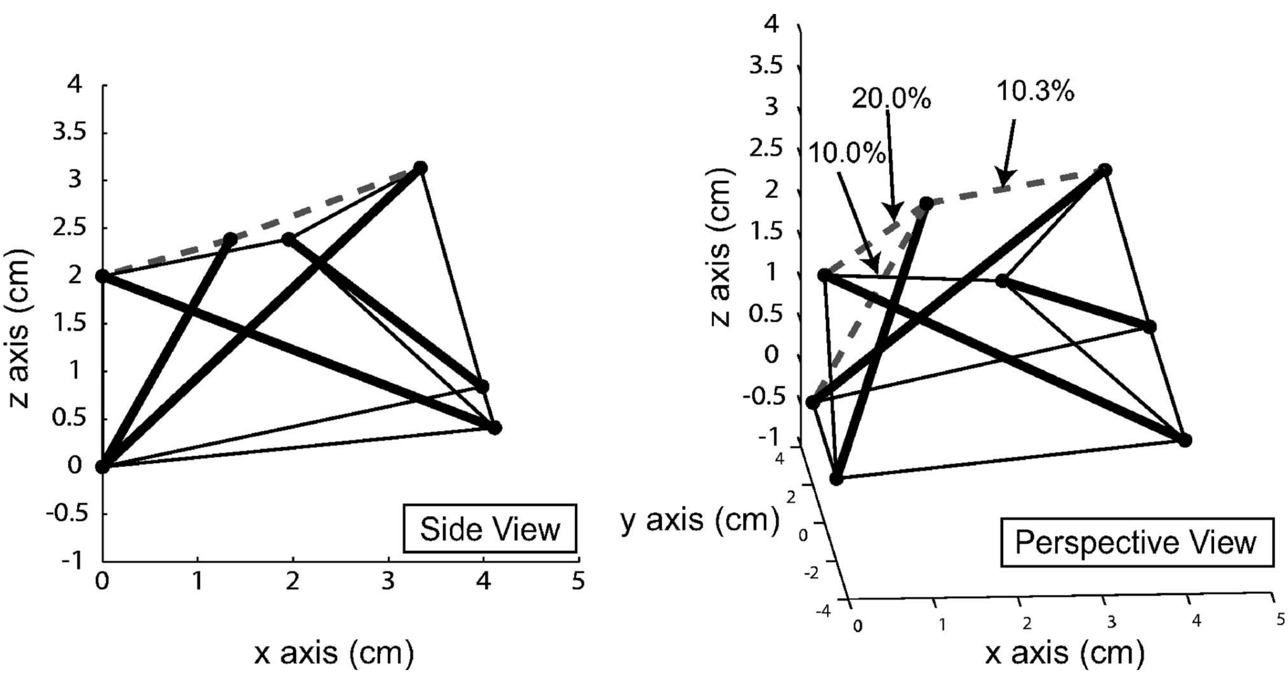
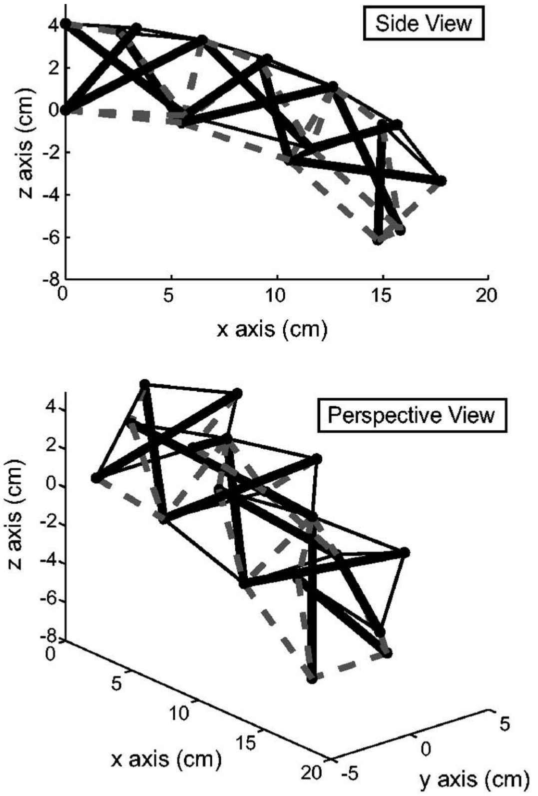
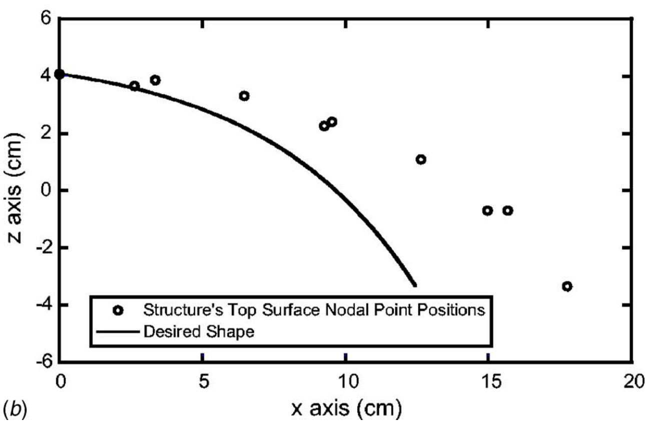


Fig. 10 Contraction amounts of individual cables in unit cell determined by the optimization scheme for upward deflection



(*a*)



(*b*)

Fig. 11 (*a*) Optimal downward deflection of the unconstrained three cell beam; and (*b*) comparison of the top surface of the structure to the desired shape. Since the length of the top of the structure to significantly larger than the length the cownose ray wing in a downward deflection, the structure cannot achieve the same deflection. This accounts for the large error in the *x* direction.

number of actuation elements. To reach the biological displacement field and minimize the number of actuators, the optimization scheme described in Sec. 6 was developed.

7.3 Optimization of Deflected Beams. Using the minimization function described in Eq. (13) cantilever beams, constructed from up to four four-strut unit cells connected together, have been studied. The unit cells are connected bar-to-bar forming a Type 2 tensegrity structure. The maximum allowed contraction percentage is set to 20% of the manufacturing lengths of specified members. This new optimization design tool determines which actuating cables are required to contract and by how much, in order to reach a desired shape or displacement field, subject to predefined constraints. Four distinct cases have been studied. The minimization function for the four optimization cases is strongly weighted to ensure the smallest error occurs for the vertical deflections. As the number of cells in the beam is increased, the structure's ability to achieve and resolve the desired shape strengthens—i.e., the errors get smaller as the number of cells increase. This is an expected consequence as the number of degrees of freedom also increases, allowing for finer shape changes.

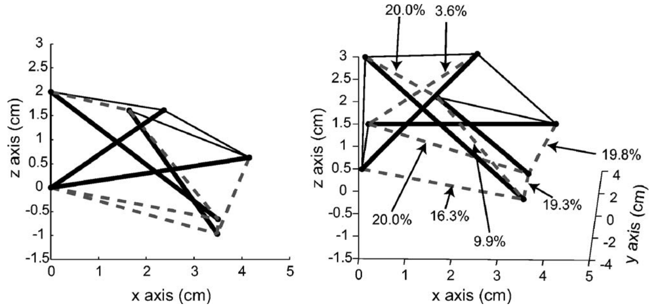


Fig. 12 Contraction amounts of individual cables in unit cell determined by the optimization scheme for downward deflection

The first shape optimization case is an upward deflection where the top nodal points of a structure are matched to the cownose deflected shape and the design space is unconstrained, meaning that all of the cables are possible actuators. The unconstrained problem reaches small minimization functional values, i.e., with all cables being potential actuators; the shape of the structure will be close to the desired shape. The actuation results for a three cell cantilever beam can be seen in Fig. 9. The unconstrained case gives excellent agreement to the cownose data with only three cells connected together with the errors falling to less than 2% in the z direction (Table 1). The greatest source of error is in the x direction which can be reduced by allowing for larger actuation strains than 20% or by increasing the amount of cells in the spanwise direction. An example of the contraction percentages of the cables for the unconstrained case of a single cell beam are given in Fig. 10, for an upward deflection.

For the second case study, the deflected shape of the top nodes of the structure is optimized to achieve the downward deflected shape of the cownose ray, given an unconstrained design space. Figure 11 shows the deflected shape of a three cell beam. The unconstrained problem produces some interesting results in terms of which cables are actuated. As can be seen in Fig. 11, several of the cables connecting the top and bottom layers of the structure are actuated. This is to be expected due to the fact that it is the top surface of the structure that is being matched to the downward deflection field. Again, there is excellent agreement in the unconstrained case with the cownose data (Table 1), except in the x direction. As an example of the contraction percentages the unconstrained case is shown for a single cell beam in Fig. 12, performing a downward deflection.

The x direction error is listed in Table 1 as not applicable because the error between the desired shape and the structural shape cannot be compared in the x direction. This inconsistency arises because the top surface of the structure is matched to the desired shape, while the structure is deflecting downward. In this situation the length of the desired shape curve is significantly shorter than the length of the top surface of the structure leading to a situation where the structure can never reach the desired shape. Since the z direction is the preferential direction in the minimization function the optimization obtains results where the z direction error was very small and the x direction error was very large. If one where to prescribe the x direction as the preferential direction the optimization would obtain results with a small x direction error and a large z direction error. One way to achieve small errors in both the x and z directions would be to scale up the size of the desired shape, however this is not consistent with the cownose ray data

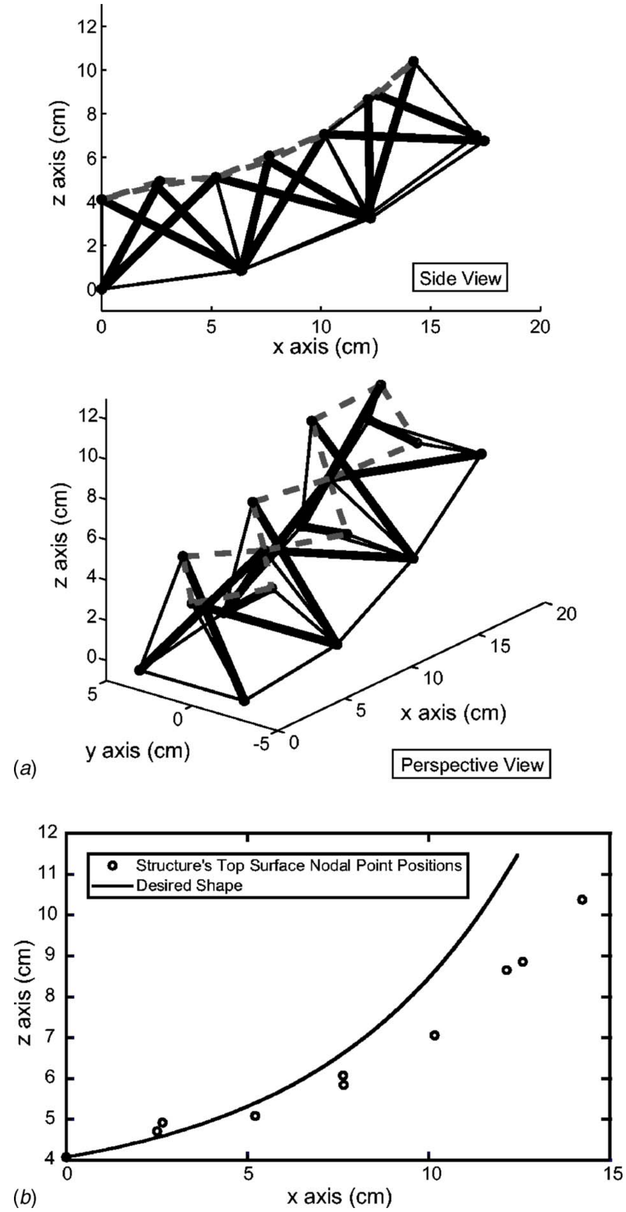


Fig. 13 (*a*) **Optimal upward deflection of the constrained three cell beam; and** (*b*) **comparison of the top surface of the structure to the desired shape**

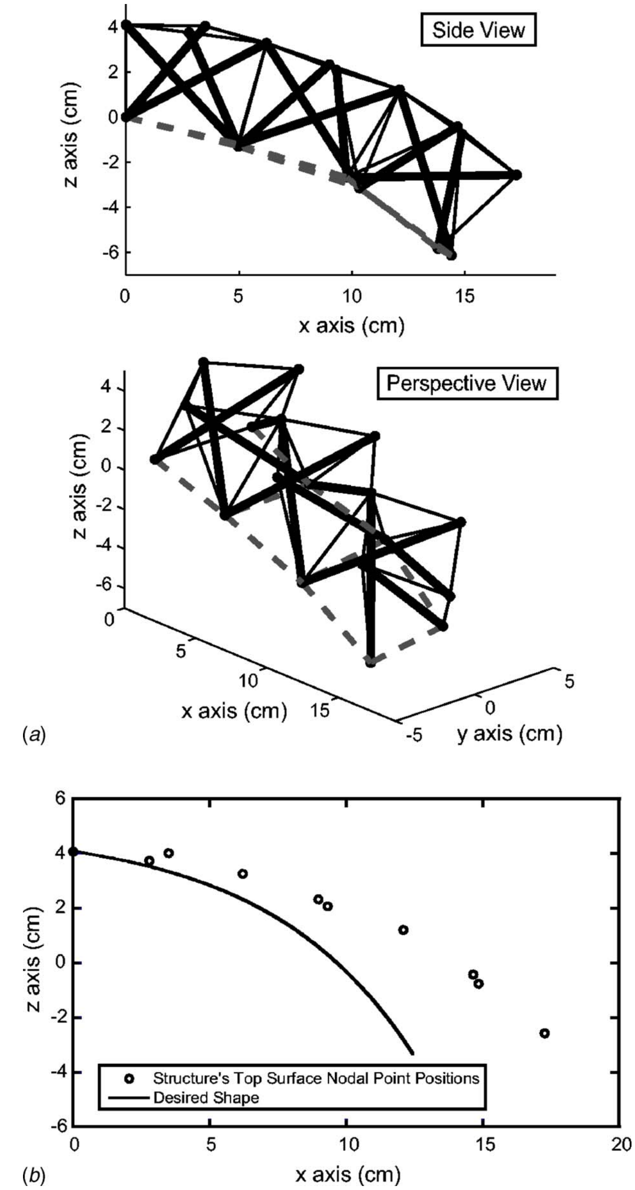


Fig. 14 (*a*) **Optimal downward deflection of the constrained three cell beam; and** (*b*) **comparison of the top surface of the structure to the desired shape**

set. The best way to achieve small error in both directions is to match the bottom surface of the structure to the desired shape in a downward deflection. This relieves the structure of the physical constraint presented in the top surface optimization for a downward deflection.

From the first and second cases the unconstrained problem is shown to be an excellent starting point for determining which sets of actuators are the dominant actuators for a given shape change and even in some cases the unconstrained design space may prove feasible in terms of manufacturability. However, the unconstrained problem typically is not practical since the optimization produces a structure with a large number of active members, making it difficult to build and more expensive to operate. But a constrained optimization case can be used that limits the potential actuators to a certain subset of the members, i.e., the dominant active members from the unconstrained case. The third and fourth cases explore the constrained problem for the upward and downward deflections.

The third shape optimization case is an upward deflection where the top nodal points of the structure are matched to the cownose deflected shape and the design space is constrained to the top cables as potential actuators. The actuation results for a three cell cantilever beam can be seen in Fig. 13. For the constrained case the error between the desired shape and the optimized shape increases slightly over the unconstrained problem, but still gives very good agreement with the biological data (Table 1).

The fourth case studied optimizes the deflected shape of the top nodes of the structure to the downward deflected shape of the cownose ray with a potential actuator space constrained to only the bottom cables. Figure 14 shows the deflected shape of a three cell beam. There is excellent agreement in the constrained case with the cownose data (Table 1) and the errors are only slightly higher than the unconstrained case.

Both the unconstrained and the constrained cases reach much closer to the actual cownose shape than a designer's choice of the actuator locations and amounts (Table 1). In order to evaluate the overall performance of all design choices a weighted average error has been calculated. This error takes into account the weights of the minimization function, which gives the z direction the

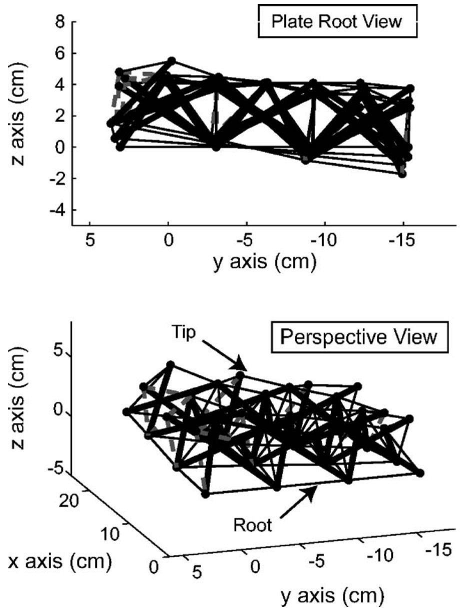


Fig. 15 Nine cell plate optimized for a 15 deg twist

strongest influence. Table 1 highlights the two main reasons for using this optimization method as a design tool: (1) when designing a tensegrity structure to reach a specific shape it is not intuitive which members should be actuators and (2) when the active members are chosen it is not intuitive by how much they should be actuated.

7.4 Optimization of a Twisting Plate. The minimization function presented in Eq. (14) is used to optimize the end nodes of a three cell by three cell plate structure to achieve a prescribed twist angle of 15 deg. The results of this optimization are shown in Fig. 15. The dotted lines represent the active cables, the solid thin lines are the passive cables and the solid thick lines are the struts. The average errors in the x, y, and z directions are 0.67%, 1.36%, and 1.46%, giving an excellent agreement with the desired shape. This example has shown the robustness of the optimization design tool developed in this paper. The method can handle any structural compositions as well as any desired shape by determining a new minimization function for each shape. Material property constraints may also be added once the materials are chosen.

8 Conclusions

This paper applies the virtual work method to the problem of form finding of tensegrity structures. By actuating individual elements using the virtual work method, deformation of single and multiple cell beams were studied. A new optimization design tool was presented which can determine which elements in a structure need to be actuated and by how much in order to reach a desired shape. In particular it was shown that a tensegrity beam structure can match very closely to the biological displacement of the cownose ray with only a few cells connected together. The optimization tool is a necessary step in the design of a morphing wing when the shape of the activated wing must be close to a desired displacement field. As the desired displacement field becomes more complex this optimization method becomes more important. However, any intuitive approach can be improved upon by using this design tool.

Acknowledgment

Keith William Moored and Hilary Bart-Smith gratefully acknowledge the support of the National Science Foundation through Award No. 0348448 and the David and Lucille Packard Foundation through the Packard Fellowship for Science and Engineering.

Nomenclature

$\mathbf{p}$ = nodal point vector
$\mathbf{x}$ = x coordinate vector of nodal points
$\mathbf{y}$ = y coordinate vector of nodal points
$\mathbf{z}$ = z coordinate vector of nodal points
$\mathbf{x_0}$ = initial x positions
$\mathbf{y_0}$ = initial y positions
$\mathbf{z_0}$ = initial z positions
$\mathbf{f}_{\text{ext}}$ = external force at a node
λ = force density in a member
$\boldsymbol{\lambda}^1$ = one-dimensional force density vector
$\boldsymbol{\lambda}$ = three-dimensional force density vector
f = internal force in a member
L = equilibrium length
L_m = unstressed manufacturing length
E = Young's modulus
A = cross-sectional area
n_n = number of nodes
n_e = number of elements
$\mathbf{C}^1$ = one dimensional connectivity matrix
$\mathbf{C}$ = three-dimensional connectivity matrix
$\mathbf{C_{cables}}$ = full cable connectivity matrix
$\mathbf{C_{struts}}$ = full strut connectivity matrix
L_{struct} = characteristic length of the structure
L_{ray} = characteristic length of the cownose ray wing
S = size ratio
D = deflection ratio
x_{def} = deflected x coordinates of a structure
x_{flat} = flat x coordinate of a structure
n_{top} = number of top nodes
n_{tip} = number of tip nodes
$f_{\min}$ = minimization function
θ = desired angle of twist
$\mathbf{p_{target}}$ = desired nodal point positions
$\mathbf{u_s}$ = upper bound of the struts
$\mathbf{u_c}$ = upper bound of the cables
$\mathbf{l_c}$ = lower bound of the cables
$\boldsymbol{\gamma}$ = constraint parameter
$\boldsymbol{\alpha}$ = cable lower bound parameter
$\boldsymbol{\beta}$ = cable upper bound parameter

References

[1] Fuller, R. B., 1962, *Tensile-Integrity Structures*, U.S. Patent and Trademark Office, Washington, D.C.
[2] Snelson, K. D., 1965, *Continuous Tension, Discontinuous Compression Structures*, U.S. Patent and Trademark Office, Washington, D.C.
[3] Kenner, H., 1976, *Geodesic Math and How to Use It*, University of California Press, Berkley, CA.
[4] Pellegrino, S., 1986, *Mechanics of Kinematically Indeterminate Structures*, University of Cambridge Press, Cambridge, UK
[5] Tibert, A. G., and Pellegrino, S., 2003, "Review of Form-Finding Methods for Tensegrity Structures," Int. J. Space Struct., **18**(4), pp. 209–223.
[6] Schek, H. J., 1974, "The Force Density Method for Form Finding and Computation of General Networks," Comput. Methods Appl. Mech. Eng., **3**, pp. 115–134.
[7] Masic, M., Skelton, R. E., and Gill, P. E., 2005, "Algebraic Tensegrity Form-Finding," Int. J. Solids Struct., **42**(16–17), pp. 4833–4858.
[8] Moored, K. W., and Bart-Smith, H., 2005, "The Analysis of Tensegrity Structures for the Design of a Morphing Wing," *Proceedings of IMECE'05*, Orlando, FL, November 8–12, ASME, New York.
[9] Masic, M., and Skelton, R., 2004, "Open Loop Control of Class-2 Tensegrity Towers," Proceedings SPIE's 11th Annual International Symposium on Smart Structures and Materials. San Diego, CA.
[10] Heine, C., 1992, *Mechanics of Flapping Fin Locomotion in the Cownose Ray, Rhinoptera Bonasus (Elasmobranchii, Myliobatidae)*, Duke University Press, Durham, NC.

Carmen Chicone
e-mail: carmen@math.missouri.edu

Michael Heitzman
e-mail: heitzman@missouri.edu

Department of Mathematics,
University of Missouri,
Columbia, MO 65211

Z. C. Feng
Department of Mechanical and Aerospace Engineering,
University of Missouri,
Columbia, MO 65211
e-mail: fengf@missouri.edu

Transient Response of Tapered Elastic Bars

Exact solutions are obtained for a model of the longitudinal displacement along an elastic tapered bar due to a force applied at its blunt end. A formula for velocity amplification is given; it specifies the velocity of the pointed end of the bar shortly after it feels the influence of the force. For a bar with an exponentially decreasing cross-sectional area, the velocity is magnified by twice an exponential function of length. This result has applications in the design of piezoelectric drills. In addition, we discuss the differences between the motions of rigid and elastic bars during the transient before one complete reflection of the wave induced by a force applied to an end of the bar. In this regime, force is proportional to velocity for elastic bars with constant cross-sectional areas. While the force–velocity relationship is more complicated for tapered elastic bars, their exact relationship is determined. [DOI: 10.1115/1.2424719]

1 Introduction

While the motion of a point particle under the influence of a mechanical force is simply described by Newton's laws, the motion of an extended body is complicated by the propagation of that influence throughout the body from the position where the force is applied. In many cases, the propagation speed of the motion within the body is so fast (in comparison to the application of the external force) that it is regarded as being instantaneous and the body is regarded as being rigid. There are, however, situations where understanding the transient response of an object, following the application of a force, is of primary interest. An excellent example is the response of piezoelectric drills, which are used in intracytoplasmic sperm injection (see Refs. [1–5]). In this application, forces of very short duration are generated by a piezoelectric actuator that are used to produce desired motions of the tip of a micropipette. Fast mechanical motion of the pipette tip, which can be induced by such an actuator, greatly facilitates the injection process. Since the duration of the force in this application is comparable to or even shorter than the time it takes for the wave to propagate from the position of the actuator to the tip of the pipette, the compliance of the structure from the force actuator to the pipette tip should be considered in the design of piezo-drill systems.

The design process for compliant materials must take into account the distribution of the compliance. Although this complicates the design analysis, it offers additional options to the designer. While a tapered bar can be used to amplify pressure as in a diamond anvil [6], an elastic tapered bar can be used to amplify *velocity* as in materials testing [7].

We will discuss the amplification of velocity with an eye toward the design of piezoelectric drills that might be used in cell biology [8,9] and needle biopsy devices that are used in cancer diagnosis, where recent research (for example, see Ref. [10]) has shown that the cutting speed of the needle strongly affects the successful penetration of a tumor. In fact, solid tumors require cutting speeds that may not be achievable with existing technology.

We model our structure by a bar with varying cross-sectional area and consider the effect of an axial force that is applied to one end of the bar, called the blunt end. The other end of the bar is called its point. What is the relationship between the force and the motion of the point of a compliant bar? How does the varying cross-sectional area affect the force–motion relationship? The answers to both questions are very simple for a rigid bar: the whole bar has identical acceleration that equals the applied force divided by the total mass of the bar, which is proportional to the length for a bar with a constant cross-sectional area.

Velocity amplification occurs when a longitudinal force is applied to the blunt end of an elastic tapered bar, which we prefer to view as a drill. The purpose of this paper is to determine the drill point velocity in the case where the drill's cross-sectional area is an exponential function of the longitudinal coordinate. More precisely, we consider the longitudinal displacement $u(x,t)$ of an elastic drill with density ρ, cross-sectional area $A(x)$, Young's modulus E, and length l, where the spatial coordinate x resides on the interval $[0,l]$. The input displacement at $x=0$ is given by a function of time $f(t)$ and the drill point is free to move from its initial position at $x=l$. The longitudinal displacement is modeled by the partial differential equation (PDE)

$$\rho A(x)u_{tt} = E[A(x)u_x]_x \tag{1}$$

on the domain $t>0$ and $0<x<l$ with boundary and initial conditions

$$u(0,t)=f(t), \quad u_x(l,t)=0, \quad u(x,0)=0, \quad u_t(x,0)=0$$

Compatibility of the initial and boundary conditions requires the input displacement to satisfy $f(0)=f'(0)=0$. This model can be used to determine the response of the bar to a force applied at $x=0$, which is modeled (using Hooke's law) by prescribing the strain

$$u_x(0,t) = -\frac{\mathcal{F}(t)}{EA(0)}$$

where $\mathcal{F}$ is the force required to cause the input displacement. The model Eq. (1) is physically realistic for the propagation of small amplitude waves in straight bars that do not taper too fast and are thin relative to the wave length of the sound [11].

The differential equation in system (1) has a long history that began in the 17th century [11]. The work of Rayleigh and Webster in the 19th century includes the description of tapered bodies, where the physical application is the propagation of sound. This direction of research continues with the description of loud speakers and other areas of acoustics. The history of the equation, as it relates to tapered drills, goes back to the 1945 patent application of W. P. Mason, where a tapered bar is proposed to magnify velocity using a piezoelectric transducer [12]. Most of the subsequent research (see Refs. [11,13] for a review up to 1966) starts from the hypothesis of sinusoidal motion; that is, the displacement can be expressed in the form

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received February 2, 2006; final manuscript received July 22, 2006. Review conducted by Sridhar Krishnaswamy.

$$u(x,t) = v(x)e^{i\omega t} \tag{2}$$

so that v is the amplitude and ω the circular frequency of vibration. This form of the solution is appropriate for power ultrasonics, where one of the main problems is to determine the resonant frequencies and amplification characteristics of the *oscillating* tapered bar (see Refs. [14], p. 115, [15], Chap. 3, and [16], Chaps. 6–7 for general discussions of longitudinal waves, and Refs. [17], pp. 223–240 for sound waves in loudspeaker horns). With this assumption and a rescaling, the partial differential Eq. (1) leads to the ordinary differential equation (ODE)

$$U'' + \frac{A'(x)}{A(x)}U' + \Omega U = 0 \tag{3}$$

where U is the scaled amplitude and Ω is a constant. This equation cannot be solved explicitly in general; but, the complete solution can be determined if one solution is known [11]. The specialization to the ODE Eq. (3) is useful for understanding the vibrational modes of the tapered bar. Also, this model predicts velocity amplification. For example, in case the cross-sectional area is $A(x)=\alpha e^{\beta x}$ and $\beta<0$, the ODE is an anti-damped harmonic oscillator. So, U grows as a function of position along the bar. On the other hand, this approach is not adequate to model the *transient* motions of tapered bars of interest here; for example, the boundary conditions imposed in the PDE system (1) are incompatible with Eq. (2).

The flexural vibrations of tapered bars have not received as much attention. In most applications in ultrasonics, flexural vibrations are avoided to concentrate ultrasonic energy or amplify velocity at the tip of a tapered bar in the longitudinal direction. On the other hand, the analysis of flexural vibrations is important and recent results are available to model and analyze these motions [18].

Ultrasonic drills are widely used in many different industrial and medical applications. Optimal design of these devices for specialized applications is a topic of current interest. Most applications employ tapered bar devices that operate in the oscillating mode (see, for example, Ref. [19] for a recent design used in drilling and coring in rock and Ref. [20] for a general discussion of ultrasound applications). But, in some important applications (in particular, the piezoelectric drills used in intracytoplasmic sperm injection), the device operates in the transient mode. A fundamental design problem—the motivation for the results presented here—is to determine the cross-sectional area of an elastic bar as a function of the axial coordinate that, for a given input displacement, maximizes the drill-point velocity $u_t(l,t)$ after the tip first begins to move.

The most important practical considerations are the relationships between the input displacement, the input force, the behavior of the displacement of the point of the bar $u(l,t)$, and the system parameters. We establish these relationships—for uniform and exponentially tapered bars—by solving the boundary-initial value problem for the linear PDE Eq. (1).

In Sec. 2, we present the special case of a bar with constant cross section and we contrast these well known results with the behavior of a rigid bar. In Sec. 3, we consider infinite tapered bars and review the standard approach in the literature involving harmonic traveling waves. In Sec. 4, we present our results for a tapered bar whose cross-sectional area is an exponential function of its axial position. Our method gives the expected geometric amplification factor of the input velocity by the tapered bar, and also establishes the relationship between the input force and input displacement in the transient response. We analyze the equation of motion using a variant of the method introduced by Ffowcs Willams and Hawkings in their work on sound generation by the arbitrary motion of surfaces (see Refs. [21] or [22], Sec. 9.4). In Sec. 5, we solve the PDE (1) with a prescribed input force and show that the solution is compatible with our results for prescribed input displacements. This analysis can be used to select actuators based on the desired motion at the tip of the bar. Discussions and conclusions are given in Sec. 6.

2 Constant Cross-Sectional Area

We will obtain the well-known exact solution of our model Eq. (1), with a specified input displacement, for the special case where the cross-sectional area is given by

$$A(x) = \alpha e^{\beta x} \tag{4}$$

with $\beta=0$; that is, the cross-sectional area is constant. The drill point velocity formula Eq. (36) is easy to verify for this case, which serves to illustrate some of our methods. In addition, we will relate our solution of the model equation to the corresponding input force. Finally, we will reconcile the motion of the elastic bar with the motion of the corresponding rigid body in the limit as Young's modulus grows without bound.

2.1 Prescribed Displacement. For $c:=\sqrt{E/\rho}$, the model PDE Eq. (1) reduces to the scalar wave equation

$$u_{tt} - c^2 u_{xx} = 0, \quad t>0, \quad 0<x<l$$

with boundary and initial conditions

$$u(0,t) = f(t), \quad u_x(l,t) = 0, \quad u(x,0) = 0, \quad u_t(x,0) = 0$$

We extend the time domain to the whole real line by setting $f(t)=0$ for $t\leq 0$. D'Alembert's solution

$$u(x,t) = a(t - x/c) + b(t + x/c) \tag{5}$$

satisfies the PDE for all t and $0<x<l$. By applying the boundary values, the functions a and b can be determined explicitly. In fact, a useful representation of the solution (satisfying the initial and boundary conditions) is given by

$$u(x,t) = f(t - x/c) + \sum_{n=1}^{\infty}(-1)^n[f(t - x/c - 2nl/c) - f(t + x/c - 2nl/c)] \tag{6}$$

(see Ref. [23], pp. 508–510). We note that for fixed x and t, the sum contains only a finite number of nonzero terms, since the arguments of f will be negative for sufficiently large n and $f(t)=0$ for $t\leq 0$.

Evaluating Eq. (6) at $x=l$ yields a representation for the displacement of the drill point

$$u(l,t) = 2\sum_{n=0}^{\infty}(-1)^n f[t - (2n+1)l/c]$$

For $t<2l/c$, this expression reduces to

$$u(l,t) = 2f(t - l/c) \tag{7}$$

As we will show, the factor of two in this formula (due to the reflection of the wave at the point of the bar) is present for tapered bars (see the general drill point velocity formula Eq. (36)).

2.2 Prescribed Displacement Versus Force. We relate the prescribed displacement of the previous section, $f(t)$, to $\mathcal{F}(t)$, the force required to cause the displacement.

By Hooke's law

$$-\frac{\mathcal{F}(t)}{\alpha} = Eu_x(0,t) \tag{8}$$

where the negative sign indicates that the stress is compressive when the force $\mathcal{F}(t)$ is positive and $\alpha=A(0)$ (see Eq. (4)).

Differentiating Eq. (6) with respect to x, evaluating the strain $u_x(x,t)$ at $x=0$, and applying Eq. (8) gives

$$\mathcal{F}(t) = \frac{E\alpha}{c} f'(t) + \frac{2E\alpha}{c} \sum_{n=1}^{\infty} (-1)^n f'(t - 2nl/c) \tag{9}$$

Starting with $t < 2l/c$ and "bootstrapping" to larger t, Eq. (9) may be solved for $f'(t)$ (i.e., $u_t(0,t)$) to get (for all t)

$$u_t(0,t) = \frac{c}{E\alpha} \mathcal{F}(t) + \frac{2c}{E\alpha} \sum_{n=1}^{\infty} \mathcal{F}(t - 2nl/c) \tag{10}$$

For $t < 2l/c$, this equation reduces to

$$u_t(0,t) = \frac{c}{E\alpha} \mathcal{F}(t) \tag{11}$$

that is, before the wave is reflected back to $x=0$, the applied force at $x=0$ is proportional to the velocity at $x=0$. This is in stark contrast to a rigid bar, where by Newton's 2nd Law, the force is proportional to the acceleration. These qualitatively different behaviors are reconciled in the next subsection.

2.3 Elastic Versus Rigid Bars. In this section, we show that the response of an elastic uniform bar to a constant force agrees with that of a rigid bar in the limit as Young's modulus (or, equivalently, the wave speed c) approaches infinity.

Let $\mathcal{F}(t) = \mathcal{F}H(t)$ be the force applied at $x=0$, where $\mathcal{F}$ is a constant function and H is the Heaviside function. In view of Eq. (10), the velocity at $x=0$ is given by

$$u_t(0,t) = \frac{\mathcal{F}}{\rho c \alpha} \left[H(t) + 2 \sum_{n=1}^{\infty} H(t - 2nl/c) \right] \tag{12}$$

where we have used the identity $E = \rho c^2$. The graph of $u_t(0,t)$ versus t is simply a step function, where the time length of each step is $2l/c$, and the height change of each step is $2\mathcal{F}/(\rho c \alpha)$, with the first step starting at the height $\mathcal{F}/(\rho c \alpha)$. Thus, an elastic bar under a constant force "accelerates" in a discrete way, moving at constant velocity over a time interval of length $2l/c$ before its speed is suddenly increased by $2\mathcal{F}/(\rho c \alpha)$. This motion is caused by the pressure wave reflecting back and forth within the bar. In the limit as Young's modulus (or equivalently the wave speed c) approaches infinity, the motion of the elastic bar approaches the motion of the corresponding rigid bar. To prove this result, for $t \geq 0$ we rewrite Eq. (12) (keeping the step function graph in mind) as

$$u_t(0,t) = \left[\frac{ct}{2l} \right] \frac{2\mathcal{F}}{\rho c \alpha} + \frac{\mathcal{F}}{\rho c \alpha}$$

where $[]$ indicates the integer part (to count which step occurs at time t). Thus

$$V(t) := \lim_{c \to \infty} u_t(0,t) = \frac{\mathcal{F}t}{\rho \alpha l} = \frac{\mathcal{F}}{M} t$$

where M is the mass of the bar. Differentiating with respect to time gives the acceleration

$$V'(t) = \frac{\mathcal{F}}{M}$$

in agreement with Newton's 2nd Law applied to a rigid bar [24].

3 Infinite Tapered Bars

Before extending the previous analysis of finite uniform bars to finite tapered bars, we review (in this brief section) some known results for a harmonic wave traveling on an infinite bar.

As before, we define $c := \sqrt{E/\rho}$. While c is the wave speed for a bar with constant cross section (where our model equation reduces to the classical wave equation and hence is nondispersive), we note that in what follows c is simply a parameter. Indeed, for a tapered bar (where our model equation allows dispersion), wave speed depends on frequency.

Consider a tapered bar with infinite length (in both the positive and negative x directions) whose cross-sectional area is given by the exponential $A(x) = \alpha e^{\beta x}$. In this case, the model PDE Eq. (1) (without its boundary conditions) reduces to the telegraph equation

$$u_{tt} - c^2 u_{xx} - \beta c^2 u_x = 0 \tag{13}$$

After the change of variables $u(x,t) = e^{-(\beta x/2)} v(x,t)$, Eq. (13) is recast as the Klein–Gordon equation

$$v_{tt} - c^2 v_{xx} + \frac{\beta^2 c^2}{4} v = 0 \tag{14}$$

which occurs in quantum mechanics and can also be used to model the transverse motion of an elastically anchored string (see Ref. [25], p. 109).

A harmonic traveling wave solution

$$v(x,t) = a e^{i(kx - \omega t)}$$

of the Klein–Gordon equation translates to a telegraph equation solution of the form

$$u(x,t) = a e^{-(\beta x/2)} e^{i(kx - \omega t)} \tag{15}$$

both of which give the same dispersion relation

$$\omega^2 = c^2 k^2 + \frac{\beta^2 c^2}{4}$$

Note that k is real if and only if $\omega \geq \beta c/2$. For this reason, $\beta c/2$ is called the cutoff frequency. (The solution $v(x,t)$ is no longer a traveling wave if k is imaginary; indeed, in this case, the spatial dependence is purely exponential rather than oscillatory and the whole bar oscillates in time with frequency ω.) The dispersion relation also gives the phase velocity

$$c_p := \frac{\omega}{k} = c \left(1 - \frac{\beta^2 c^2}{4\omega^2} \right)^{-(1/2)}$$

and the group velocity

$$c_g := \frac{d\omega}{dk} = c \left(1 - \frac{\beta^2 c^2}{4\omega^2} \right)^{(1/2)}$$

Note that $c_p > c$ and $c_g < c$, and both quantities approach c as ω approaches infinity. Also, c_p approaches infinity and c_g approaches 0 as ω approaches $\beta c/2$. The group velocity $c_g(\omega)$ can be viewed as the speed of the envelope of a packet of harmonic traveling waves with frequencies near ω or as the mean velocity of transport of energy (see Ref. [25], p. 111).

The displacement amplification that occurs in exponentially tapered bars (for $\beta < 0$)—in the regime of harmonic traveling waves—is readily apparent from the solution Eq. (15). Indeed, comparing the particle displacements after the wave travels a distance l, we have the ratio

$$\frac{u(x + l, t + l/c_p)}{u(x,t)} = e^{-(\beta l/2)}$$

which is equal to the ratio of the diameter of the bar at x to the diameter at $x + l$. In particular, this amplification ratio is independent of the frequency ω. The same result is obtained for the particle velocity amplification by taking time derivatives. Mason obtained this amplification ratio for a finite tapered bar of exponential cross section for the particular frequency at which the length of the bar is exactly half the wavelength [12].

Unfortunately, as pointed out in the Sec. 1, the boundary conditions for a finite bar subjected to an input displacement which is "turned on" at $t=0$ are not compatible with the solution Eq. (15). A different approach is required to model the transient response.

4 Finite Tapered Bars

In this section, we determine the drill point velocity for the model PDE Eq. (1), with prescribed input displacement, where the cross-sectional area is given by the exponential $A(x)=\alpha e^{\beta x}$ and we relate the prescribed input displacement to the input force.

4.1 Prescribed Input Displacement. As before, we define $c := \sqrt{E/\rho}$ and again note that it is simply a parameter. Using the exponential A, the model equation reduces to the telegraph equation boundary value problem (BVP)

$$u_{tt} - c^2 u_{xx} - \beta c^2 u_x = 0, \quad 0 < x < l, \quad t > 0$$

$$u(x,0) = u_t(x,0) = 0, \quad u(0,t) = f(t), \quad u_x(l,t) = 0$$

After the change of variables $u(x,t)=e^{-(\beta x/2)}v(x,t)$, we have the Klein–Gordon BVP

$$v_{tt} - c^2 v_{xx} + \frac{\beta^2 c^2}{4} v = 0, \quad 0 < x < l$$

$$v(x,0) = v_t(x,0) = 0, \quad v(0,t) = f(t), \quad v_x(l,t) = \frac{\beta}{2} v(l,t)$$

where we assume $f(t)=0$ for $t \leq 0$.

We apply a useful solution method with wide application (see Refs. [21] or [22], Section 9.4). Let H denote the Heaviside function. After multiplying both sides of the Klein–Gordon equation by the product $H(x)H(l-x)$, we have

$$\left(v_{tt} - c^2 v_{xx} + \frac{\beta^2 c^2}{4} v \right) H(x)H(l-x) = 0$$

which holds on the whole real line. Note that

$$[vH(x)H(l-x)]_{tt} = v_{tt}H(x)H(l-x)$$

and

$$[vH(x)H(l-x)]_{xx} = v_{xx}H(x)H(l-x) + v_x(0,t)\delta(x) - v_x(l,t)\delta(l-x) + v(0,t)\delta'(x) + v(l,t)\delta'(l-x)$$

Thus, on the whole real line, we get the PDE

$$w_{tt} - c^2 w_{xx} + \frac{\beta^2 c^2}{4} w = h(x,t), \quad w(x,0) = w_t(x,0) = 0 \tag{16}$$

where $w(x,t)=v(x,t)H(x)H(l-x)$ and

$$h(x,t) = -c^2 \left[v_x(0,t)\delta(x) - \frac{\beta}{2} v(l,t)\delta(l-x) + f(t)\delta'(x) + v(l,t)\delta'(l-x) \right]$$

We note that $w=v$ for $x \in [0,l]$.

The nonhomogeneous Klein–Gordon equation (16) has the well-known solution, for $t \geq 0$, given by

$$w(x,t) = \frac{1}{2c} \int_0^t \int_{x-cs}^{x+cs} J_0\left(\frac{\beta}{2} [c^2 s^2 - (x-\xi)^2] \right) h(\xi, t-s) \, d\xi \, ds \tag{17}$$

where J_0 is the Bessel function defined by

$$J_0(z) = \frac{2}{\pi} \int_0^{\pi/2} \cos(z \sin\theta) d\theta$$

(see, for example, problem 2 on p. 135 in Ref. [26] and apply Duhamel's principle).

For $x \in (0,l)$, we then have

$$\begin{aligned} v(x,t) = {} & \frac{1}{2} f\left(t - \frac{x}{c}\right) + \frac{1}{2} v\left(l, t - \frac{l-x}{c}\right) + \frac{\beta c x}{2} \int_0^{t-x/c} J_0'[g(x,t-s)] f(s) ds \\ & + \int_0^{t-(l-x)/c} \left\{ \frac{\beta c}{4} J_0[g(x-l,t-s)] + \frac{\beta c(l-x)}{2} J_0'[g(x-l,t-s)] \right\} v(l,s) ds \\ & - \frac{c}{2} \int_0^{t-x/c} J_0[g(x,t-s)] v_x(0,s) ds \end{aligned} \tag{18}$$

where

$$g(x,t) := \frac{\beta}{2}(c^2 t^2 - x^2)$$

Assuming v is continuous at $x=l$, we have that

$$\begin{aligned} v(l,t) = {} & f\left(t - \frac{l}{c}\right) + \beta l c \int_0^{t-l/c} J_0'[g(l,t-s)] f(s) ds \\ & + \frac{\beta c}{2} \int_0^t J_0[g(0,t-s)] v(l,s) ds \\ & - c \int_0^{t-l/c} J_0[g(l,t-s)] v_x(0,s) ds \end{aligned} \tag{19}$$

Assuming v_x is continuous at $x=0$, we may differentiate Eq. (18) with respect to x and set $x=0$ to obtain

$$\begin{aligned} v_x(0,t) = {} & -\frac{1}{c} f'(t) + \frac{1}{c} v_t(l, t - l/c) + \beta c \int_0^t J_0'[g(0,t-s)] f(s) ds \\ & + \frac{\beta}{2} v\left(l, t - \frac{l}{c}\right) + \beta c \int_0^{t-l/c} \left\{ \left(\frac{\beta l}{2} - 1 \right) J_0'[g(l,t-s)] + \beta l^2 J_0''[g(l,t-s)] \right\} v(l,s) ds \end{aligned} \tag{20}$$

We have expressed $v_x(0,t)$ using known functions together with $v(l,s)$ and $v_t(l,s)$ for retarded times $s \leq t - l/c$. Thus, Eq. (19) would express $v(l,t)$ using known functions together with $v(l,s)$ and $v_t(l,s)$ for retarded times $s \leq t - l/c$ if it were not for the undesirable third term. For this reason, we recast Eq. (19) in the operator form

$$(I - K)v(l,t) = F(t)$$

where I is the identity operator, K is the operator given by

$$K\psi(t) = \frac{\beta c}{2} \int_0^t J_0[g(0,t-s)] \psi(s) ds \tag{21}$$

and

$$F(t) = f\left(t - \frac{l}{c}\right) + \beta l c \int_0^{t-l/c} J_0'[g(l,t-s)] f(s) ds - c \int_0^{t-l/c} J_0[g(l,t-s)] v_x(0,s) ds$$

We note that $F(t)=0$ for $t \leq l/c$.

Consider $t \leq 2l/c$. To determine $F(t)$, we need to know $v_x(0,s)$ for $s \leq l/c$. But, by Eq. (20) for $s \leq l/c$, we have

$$v_x(0,s) = -\frac{1}{c}f'(s) + \beta c\int_0^s J_0'[g(0,s-\tau)]f(\tau)d\tau$$

Thus, for $t \leq 2l/c$

$$F(t) = f\left(t-\frac{l}{c}\right) + \beta lc\int_0^{t-l/c} J_0'[g(l,t-s)]f(s)ds + \int_0^{t-l/c} J_0[g(l,t-s)]f'(s)ds - \beta c^2\int_0^{t-l/c} J_0[g(l,t-s)]\int_0^s J_0'[g(0,s-\tau)]f(\tau)d\tau\, ds \tag{22}$$

For a more precise definition of the operator K, let $T > 0$ and let $\mathcal{B}$ denote the Banach space of functions that are bounded on $[0,T]$ with the sup norm $\|\cdot\|$ (i.e., for $\psi \in \mathcal{B}$, $\|\psi\| = \sup_{t\in[0,T]}|\psi(t)|$). The operator K is defined on $\mathcal{B}$ by Eq. (21). We let $\|\cdot\|_{\text{op}}$ be the corresponding operator norm (i.e., $\|K\|_{op} = \sup_{\{\psi\in\mathcal{B}:\|\psi\|=1\}}\|K\psi\|$).

Formally, the inverse of the operator $I-K$ is represented by the series

$$(I-K)^{-1} = I + \sum_{j=1}^{\infty} K^j$$

It can be shown that

$$\left\| I + \sum_{j=1}^{\infty} K^j \right\|_{\text{op}} \leq e^{|\beta|cT/2}$$

i.e., the series is a bounded operator and thus converges.

Whenever $F \in \mathcal{B}$, we then have

$$v(l,t) = \left(I + \sum_{j=1}^{\infty} K^j\right)F(t) \tag{23}$$

We note that for $T \leq 2l/c$, we have $F \in \mathcal{B}$ by Eq. (22). By a tedious process of iteration, we can extend this result to as large a value of T as is needed in the evaluation.

As noted before, for $t \leq l/c$, $F(t)=0$ and thus $v(l,t)=0$ and $u(l,t)=0$. Hence the average speed of a "signal" sent from $x=0$ to $x=l$ can be no greater than c. In fact, c is the maximum speed of transmission of a disturbance in any medium governed by a PDE of the form $c^2u_{xx}=u_{tt}+\Phi(u,u_x,u_t)$ (see Ref. [25], pp. 110 and 224).

Note that in the special case $\beta=0$, we have $K=0$ and thus $v(l,t)=F(t)$. In this case, for $t \leq 2l/c$, $F(t)$ simplifies to

$$F(t) = f\left(t-\frac{l}{c}\right) + \int_0^{t-l/c} f'(s)ds = 2f\left(t-\frac{l}{c}\right)$$

This agrees with the wave equation result, as it should, since for $\beta=0$ the Klein–Gordon equation reduces to the wave equation.

To return to the original function u, we multiply Eq. (23) by the factor $e^{-(\beta l/2)}$ to obtain

$$u(l,t) = \left(I + \sum_{j=1}^{\infty} K^j\right)e^{-(\beta l/2)}F(t) \tag{24}$$

where again $F(t)$ is given by Eq. (22) for $t < 2l/c$.

It can be shown (using integration by parts) that for each positive integer n

$$\frac{d}{dt}(K^n\psi)(t) = \frac{\beta c}{2}\psi(0)\left[K^{n-1}J_0\left(\frac{\beta c^2 t^2}{2}\right)\right](t) + (K^n\psi')(t) \tag{25}$$

In particular, since $F(0)=0$, we have

$$\frac{d}{dt}(K^nF)(t) = (K^nF')(t)$$

Thus

$$u_t(l,t) = \left(I + \sum_{j=1}^{\infty} K^j\right)e^{-(\beta l/2)}F'(t) \tag{26}$$

where, by differentiating Eq. (22) and integrating by parts (recalling that $f(0)=0$)

$$F'(t) = 2f'\left(t-\frac{l}{c}\right) + \beta c\int_0^{t-l/c} J_0'[g(l,t-s)]f'(s)[l+c(t-s)]ds - \beta c^2\int_0^{t-l/c} J_0'[g(0,t-l/c-s)]f(s)ds - \beta^2c^4\int_0^{t-l/c}(t-s) \times J_0'[g(l,t-s)]\int_0^s J_0'[g(0,s-\tau)]f(\tau)d\tau ds \tag{27}$$

Note that Eq. (26) can also be obtained by noticing that both u and u_t must satisfy the model Eq. (1).

To see how $u_t(l,t)$ changes from the simple wave equation case ($\beta=0$), we will compute the derivative of $u_t(l,t)$ with respect to β at $\beta=0$. Since the operator K depends on β, we do not use Eq. (26). Instead, we note that $(d/d\beta)F(t)=0$ at $\beta=0$ by Eq. (22) and differentiate

$$(I-K)u(l,t) = e^{-(\beta l/2)}F(t)$$

first with respect to β at $\beta=0$ and then again with respect to t to obtain

$$\left.\frac{d}{d\beta}u_t(l,t)\right|_{\beta=0} = cf(t-l/c) - lf'(t-l/c)$$

Suppose f is a pulse with support in $[0,\delta]$ for some $0<\delta<l/c$. Let $M := \max_{t\in[0,\delta]}f(t)$. There is some time, say $t=\tau$, when $f'(\tau-l/c) > M/\delta$. It follows that $lf'(\tau-l/c) > cf(\tau-l/c)$; hence

$$\left.\frac{d}{d\beta}u_t(l,\tau)\right|_{\beta=0} < 0$$

In fact, since for $\beta=0$, we have $u_t(l,t)=2f'(t-l/c)$, the velocity $u_t(l,t)$ is maximized when $f'(t-l/c)$ is maximized. Thus, if we take τ to be the time at which $u_t(l,t)$ is maximized, then we will have

$$\left.\frac{d}{d\beta}u_t(l,\tau)\right|_{\beta=0} < 0$$

In other words, decreasing β from 0 will increase the maximum velocity of the drill point (at least near $\beta=0$).

Returning to the case of arbitrary β and f, we wish to examine the response $u_t(l,l/c)$. Note that l/c is the time it would take a signal traveling with speed c to go from $x=0$ to $x=l$. Recall that for $t<l/c$, we have $F'(t)=0$ by Eq. (27); hence $u_t(l,t)=0$ on this time interval. For $t=l/c$, Eq. (27) gives $F'(l/c)=2f'(0)$. By Eq. (21), $K^nF'(l/c)=0$ for $n\geq 1$. In fact, it can be easily seen that $K^nF'(l/c+\epsilon)=O(\epsilon^n)$. Thus Eq. (26) yields

$$u_t\left(l,\frac{l}{c}\right) = 2e^{-(\beta l/2)}f'(0) \tag{28}$$

where u_t and f' are right-hand derivatives. This equation implies that decreasing β results in a larger response. Furthermore, it follows that the influence from the initial input travels with speed c, even for nonzero β where dispersion may occur. This fact may be reconciled with the results of Sec. 3 by noting that, for a superposition of harmonic solutions of the form given in Eq. (15) to satisfy the boundary conditions, terms of arbitrarily high frequen-

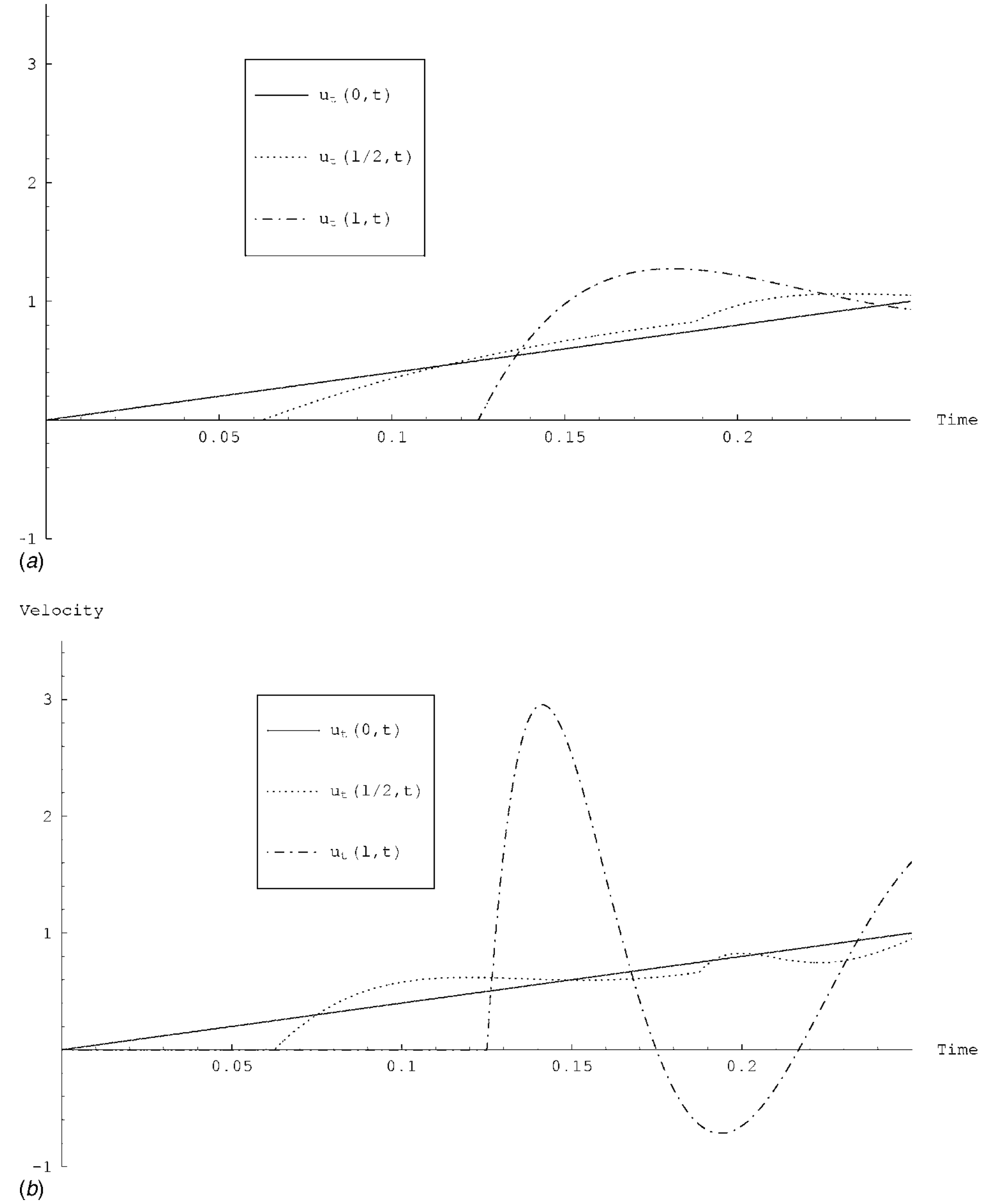


Fig. 1 The velocity response versus time of the elastic tapered bar with $l=1$ and $c=8$ (computed using a finite difference scheme applied to the PDE (1)), with cross-sectional area $A(x)=e^{-4x}$ in the top panel and $A(x)=e^{-8x}$ in the bottom panel, is depicted for the quadratic prescribed input displacement $u(0,t)=f(t)=2t^2$

cies ω would be required, for which the group velocity would approach c. Finally, we note that the velocity $u_t(l,l/c)$ can be made arbitrarily large by adjusting the parameter β. Thus, without further design constraints, no optimal shape exists. Of course, the model (1) may not be valid for large $|\beta|$.

The asymptotic behavior at $(x,t)=(l,l/c)$ is given more precisely by

$$u_t\left(l,\frac{l}{c}+\epsilon\right)=2e^{-(\beta l/2)}\left\{f'(0)+\epsilon\left[f''(0)+\frac{\beta c}{2}f'(0)\right]\right\}+O(\epsilon^2) \tag{29}$$

Figure 1 depicts the velocity response at the blunt end, the middle, and the point of two different tapered bar profiles, clearly showing the velocity amplification. Note that the point velocity, shortly after it begins moving, is much higher than the blunt end velocity after it begins moving, but the duration of this velocity amplification is short.

4.2 Displacement Versus Force.

Recall that, for constant cross-sectional area and $t<2l/c$, the velocity of the blunt end of the bar is proportional to the applied force (see Eq. (11)). In this section we discuss the velocity–force relationship for tapered bars whose cross-sectional areas are given by an exponential function of the length along the bar. The result is the same at the lowest order of approximation. The correction term at the next highest order is determined.

Using Hooke's Law, we relate the force $\mathcal{F}(t)$ and the strain $u_x(0,t)$ as follows

$$-\frac{\mathcal{F}(t)}{\alpha}=Eu_x(0,t) \tag{30}$$

Recalling the change of variables

$$u(x,t)=e^{-(\beta x/2)}v(x,t)$$

and differentiating with respect to x, we find that

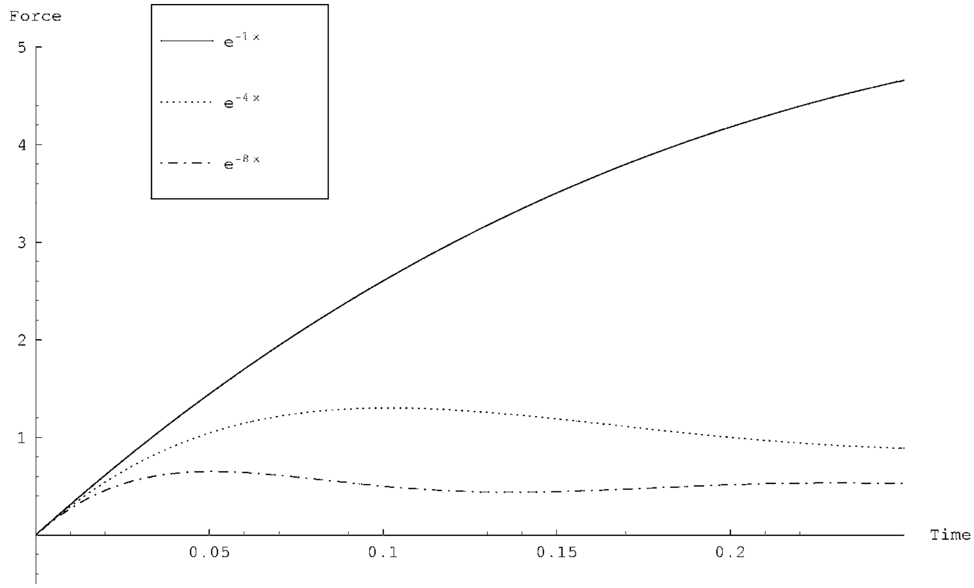


Fig. 2 The force ($-\rho c^2 u_x(0,t)$) versus time (computed using a finite difference scheme applied to the PDE (1)) at the blunt end of the elastic tapered bar (l=1, ρ=1, and c=8) with cross-sectional areas $A(x)=e^{\beta x}$, for $\beta\in\{-1,-4,-8\}$ is depicted for the quadratic prescribed input displacement $u(0,t)=f(t)=2t^2$

$$u_x(x,t) = -\frac{\beta}{2}e^{-(\beta x/2)}v(x,t) + e^{-(\beta x/2)}v_x(x,t)$$

By evaluation of this expression at $x=0$ and an application of the boundary conditions, the strain is given by

$$u_x(0,t) = -\frac{\beta}{2}f(t) + v_x(0,t)$$

For $t<2l/c$, Eq. (20) implies that

$$v_x(0,t) = -\frac{1}{c}f'(t) + \beta c\int_0^t J_0'\left[\frac{\beta}{2}c^2(t-s)^2\right]f(s)ds$$

Hence, for $t<2l/c$

$$u_x(0,t) = -\frac{\beta}{2}f(t) - \frac{1}{c}f'(t) + \beta c\int_0^t J_0'\left[\frac{\beta}{2}c^2(t-s)^2\right]f(s)ds$$

Applying Hooke's Law, in the form of Eq. (30), we get

$$\mathcal{F}(t) = \frac{E\alpha\beta}{2}f(t) + \frac{E\alpha}{c}f'(t) - E\alpha\beta c\int_0^t J_0'\left[\frac{\beta}{2}c^2(t-s)^2\right]f(s)ds \tag{31}$$

We note that the force $\mathcal{F}$ required to cause the displacement f does not depend on the length of the bar, just as one would expect in the transient response.

For $\beta=0$ in Eq. (31), only the middle term survives; and we recover the constant, cross-sectional area result: the input force is proportional to velocity (in the transient regime). For nonzero β, this simple relationship is modified by the first and third terms.

To determine the effect of the modifying terms for small t, we observe that the expansion of Eq. (31) about $t=0$ (recalling that $f(0)=0$) is

$$\mathcal{F}(t) = \frac{E\alpha}{c}\left\{f'(0) + \left[\frac{\beta c}{2}f'(0) + f''(0)\right]t + \left[\frac{\beta c}{2}f''(0) + f'''(0)\right]\frac{t^2}{2}\right\} + O(t^3) \tag{32}$$

In particular, the first-order term is independent of β (even if $f'(0)=0$). Hence, for sufficiently small t, the tapered bar will behave approximately the same as the bar with constant cross-sectional area, that is, the force is approximately proportional to the velocity. Note that every term of the expansion Eq. (32) comes from the first two terms in Eq. (31). The third term in Eq. (31) is of order $O(t^4)$.

Figure 2 shows the computed force at the blunt ends of bars with different cross-sectional area profiles. We also note that—as we have shown (see Eq. (32))—the force is proportional to the velocity for a short time interval; but, the geometry of the bar influences this relationship as time increases.

5 Prescribed Force Boundary Data

In this section, we consider the prescribed force ($\mathcal{F}$) boundary condition at $x=0$ for the model Eq. (1). Recall that the strain at $x=0$ is modeled by

$$u_x(0,t) = -\frac{\mathcal{F}(t)}{EA(0)}$$

where the negative sign indicates that a force acting in the positive x direction will compress the drill, causing a negative strain. The main result here is that the solution of the model equation agrees with the computed force at $x=0$ obtained from the solution of the model equation with prescribed input displacement. We will assume as before that the cross-sectional area is given by $A(x)=\alpha e^{\beta x}$.

5.1 Constant Cross-Sectional Area. In case the cross-sectional area is constant; that is, $\beta=0$ and for $c:=\sqrt{E/\rho}$, the model PDE Eq. (1) reduces to the scalar wave equation

$$u_{tt} - c^2u_{xx} = 0, \quad t>0,\ 0<x<l$$

with boundary and initial conditions

$$u_x(0,t) = -\frac{\mathcal{F}(t)}{E\alpha}, \quad u_x(l,t)=0, \quad u(x,0)=0, \quad u_t(x,0)=0$$

Extending to all t by assuming $\mathcal{F}(t)=0$ for $t\le 0$ and using the D'Alembert solution Eq. (5) and the boundary conditions as in Sec. 2.1, the derivatives of the functions a and b can be determined to obtain a general form of the velocity solution (analogous to the displacement solution Eq. (6)) which for $t<2l/c$ and $x=l$, simplifies to

$$u_t(l,t) = \frac{2c}{E\alpha}\mathcal{F}(t-l/c) \tag{33}$$

5.2 Tapered Bars. In this section, we determine the drill point velocity for the model PDE Eq. (1) with the cross-sectional area given by $A(x)=\alpha e^{\beta x}$ with $\beta \neq 0$, $c := \sqrt{E/\rho}$, and $\mathcal{F}(t)$ the prescribed force at $x=0$.

As before, using the properties of the exponential function A, the model equation reduces to the telegraph equation BVP

$$u_{tt} - c^2 u_{xx} - \beta c^2 u_x = 0, \quad 0<x<l$$

$$u(x,0) = u_t(x,0) = 0, \quad u_x(0,t) = -\frac{\mathcal{F}(t)}{E\alpha}, \quad u_x(l,t) = 0$$

After the change of variables $u(x,t)=e^{-(\beta x/2)}v(x,t)$, we have the Klein–Gordon BVP

$$v_{tt} - c^2 v_{xx} + \frac{\beta^2 c^2}{4} v = 0, \quad 0<x<l$$

$$v(x,0) = v_t(x,0) = 0, \quad v_x(0,t) - \frac{\beta}{2}v(0,t) = -\frac{\mathcal{F}(t)}{E\alpha},$$

$$v_x(l,t) - \frac{\beta}{2}v(l,t) = 0$$

where we assume $\mathcal{F}(t)=0$ for $t \leq 0$.

Following the same general procedure as in Sec. 4.1, we multiply both sides of the Klein–Gordon equation by the product $H(x)H(l-x)$, absorb the Heaviside functions inside the derivatives, and obtain a nonhomogeneous equation analogous to Eq. (16), whose well known solution Eq. (17) yields an integral equation analogous to Eq. (18), involving v and its spatial derivatives at $x=0$ and $x=l$ at various times. Applying the boundary conditions and some manipulations gives, for $t \leq 2l/c$, an expression for $v(l,t)$ in terms of $\mathcal{F}$ and the operator K (analogous to Eq. (23), but much more complicated). Realizing that v_t solves the same homogeneous Klein–Gordon BVP as v with $\mathcal{F}$ replaced by $\mathcal{F}'$, we obtain an expression for $v_t(l,t)$ and hence $u_t(l,t)$ after multiplying by $e^{-(\beta l/2)}$. For $t \leq 2l/c$, the final result can be written as

$$\begin{aligned} u_t(l,t) = e^{-(\beta l/2)}\frac{c}{E\alpha}(I-K)^{-1}\Bigg\{ \int_0^{t-l/c} J_0[g(l,t-s)](I+K)^{-1}\mathcal{F}'(s)\,ds \\ + \frac{2}{\beta c}(I+K)^{-1}K\mathcal{F}'(t-l/c) + 2l\int_0^{t-l/c} J_0'[g(l,t-s)] \\ \times (I+K)^{-1}K\mathcal{F}'(s)\,ds \Bigg\}(t) \end{aligned} \tag{34}$$

The asymptotic behavior near $t=l/c$ is given by

$$\begin{aligned} u_t(l,l/c+\epsilon) = 2e^{-\beta l/2}\frac{c}{E\alpha}\left\{ \mathcal{F}'(0)\epsilon + \left[\mathcal{F}''(0) + \frac{\beta c}{4}\mathcal{F}'(0)\right]\frac{\epsilon^2}{2}\right\} \\ + O(\epsilon^3) \end{aligned} \tag{35}$$

Note the asymptotic behavior of the solution for the constant cross section case given by Eq. (33) is

$$u_t(l,l/c+\epsilon) = \frac{2c}{E\alpha}\mathcal{F}(\epsilon) = \frac{2c}{E\alpha}\left[\mathcal{F}'(0)\epsilon + \mathcal{F}''(0)\frac{\epsilon^2}{2}\right] + O(\epsilon^3)$$

since $\mathcal{F}(0)=0$. Also, note that Eq. (35) is compatible with the prescribed displacement solution Eq. (29) by using Eq. (31).

Again, we see that for the tapered bar, the velocity of the drill tip shortly after the arrival of the wave may be made arbitrarily large by taking β arbitrarily negative. On the other hand, the model Eq. (1) may only be valid for small $|\beta|$.

6 Discussion and Conclusions

Although the wave equation is derived from the pointwise momentum balance and the results are consistent with the predictions from Newton's second law, design principles for elastic bodies in the transient regime are very different from those for rigid bodies. Varying the cross-sectional area of an elastic bar (or drill) allows amplification of the velocity (at the point of the bar) resulting from a force applied at its blunt end. We have found an analytic solution expressing this amplified velocity at the point of the bar in the transient response. At least for bars whose cross-sectional areas are exponential functions of their axial positions, our formula for the drill point velocity is

$$u_t\left(l,\frac{l}{c}+\epsilon\right) = 2e^{-(\beta l/2)}\left\{ f'(0) + \epsilon\left[f''(0) + \frac{\beta c}{2}f'(0)\right]\right\} + O(\epsilon^2) \tag{36}$$

where $c := \sqrt{E/\rho}$, $\epsilon>0$ and $f(t)$ is the input displacement (which is related to the applied force, $\mathcal{F}(t)$ by Eqs. (31) and (32)). A more direct relationship between drill point velocity and $\mathcal{F}(t)$ is given by Eqs. (34) and (35). For compatible boundary and initial conditions, formula (36) reduces to

$$u_t\left(l,\frac{l}{c}+\epsilon\right) = 2\epsilon e^{-(\beta l/2)}f''(0) + O(\epsilon^2) \tag{37}$$

In addition to this formula, we have presented useful analytic solutions of our mathematical model.

We note that if $f''(0) \neq 0$, then $t_* = l/c$, where t_* is the moment before the point of the bar begins to move; that is, the propagation speed of the influence from the input displacement is $\sqrt{E/\rho}$. Also, the velocity of the drill point increases as β decreases. This result shows that the theoretical design problem is delicate; to wit, the drill point velocity grows arbitrarily large as $\beta \to -\infty$. Of course, as mentioned previously, the model (1) may not be valid for large $|\beta|$. On the other hand, as numerical experiments indicate (see Fig. 1), as β decreases the large velocities have short durations. Thus, the properties of the material composition of the drill as well as the desired time interval of drill point action must be taken into account.

In the transient regime (before one complete reflection of the wave produced by the applied force) the *velocity* of a bar (with constant cross-sectional area) at the point of force application is proportional to the applied force. Moreover, this force–velocity relationship is not affected by the length of the bar. For bars with varying cross-sectional areas, the force–velocity relationship is more complicated; but, we have given a formula for the appropriate correction terms. These results are contrary to the behavior of rigid bars. The motions of rigid and elastic bars are reconciled only after an infinite number of reflections of the pressure wave. Our model does not include wave damping, which must be taken into account for long bars.

Acknowledgment

This work is supported in part by Grant No. NSF/DMS-0604331.

References

[1] Huang, T., Kimura, Y., and Yanagimachi, R., 1996, "The Use of Piezo Micromanipulation for Intracytoplasmic Sperm Injection of Human Oocytes," J. Assist Reprod. Genet., **13**(4), pp. 320–328.

[2] Katayose, H., Yanagida, K., Shinoki, T., Kawahara, T., Horiuchi, T., and Sato, A., 1999, "Efficient Injection of Bull Spermatozoa into Oocytes Using a Piezo-driven Pipette," Theriogenology, **52**, pp. 1215–1224.

[3] Kimura, Y., and Yanagimachi, R., 1995, "Intracytoplasmic Sperm Injection in the Mouse," Biol. Reprod., **52**(4), pp. 709–720.
[4] Nakayama, T., Fijiwara, H., Tastumi, K., Fujita, K., Higuchi, T., and Mori, T., 1998, "A New Assisted Hatching Technique Using a Piezo-Micromanipulator," Fertil. Steril., **69**(4), pp. 784–788.
[5] Takeuchi, S., Minoura, H., Shibahara, T., Shen, X., Futamura, N., and Toyoda, N., 2001, "Comparison of Piezo-Assisted Micromanipulation with Conventional Micromanipulation for Intracytoplasmic Sperm Injection into Human Oocytes," Gynecol. Obstet. Invest., **52**, pp. 158–162.
[6] Dunstan, D. J., and Spain, I. L., 1989, "Technology of Diamond Anvil High-Pressure Cells: I. Principles, Design and Construction," J. Phys. E, **22**, pp. 913–923.
[7] *ASTM Standard. ASTM Designation: G32-92. Standard Test Method for Cavitation Erosion Using Vibratory Apparatus*, Philadelphia, PA.
[8] Lal, A., and White, R. M., 1995, "Silicon Microfabricated Horns for Power Ultrasonics", *Proceedings Transducers 95*, Stockholm, Sweden, July 25–29.
[9] Lal, A., and White, R. M., 1995, "Micromachined Silicon Needle for Ultrasonic Surgery," *Proceedings IEEE Ultrasonics Symposium*, Seattle, WA, November 7–10, IEEE, Piscataway, NJ, pp. 1593–1596.
[10] Wendt, O., Siewert, C., Luth, T., Felix, R., and Boenick, U., 2001, "Cutting Speeds and Success of Biopsy with Different Punch Biopsy Instruments," Radiologe, **41**(6), pp. 484–490 in (German).
[11] Eisner, E., 1967, "Complete Solutions of the "Webster" Horn Equation," J. Acoust. Soc. Am., **41**(4), pp. 1126–1967.
[12] Mason, W. P., and Wick, R. F., 1951, "A Barium Titanate Transducer Capable of Large Motion at an Ultrasonic Frequency," J. Acoust. Soc. Am., **23**, pp. 209–214.
[13] Eisner, E., 1963, "Design of Sonic Amplitude Transformers for High Magnification," J. Acoust. Soc. Am., **35**(9), pp. 1367–1377.
[14] Graff, K. F., 1991, *Wave Motion in Elastic Solids*, Reprint, Dover, New York.
[15] Kolsky, H., 1953, *Stress Waves in Solids*, Clarendon Press, Oxford, U.K.
[16] Redwood, M., 1960, *Mechanical Waveguides*, Pergamon, Bristol, U.K.
[17] McLachlan, N. W., 1958, *Ordinary Non-linear Differential Equations in Engineering and Physical Sciences*, 2nd ed., Clarendon, Oxford, U.K.
[18] Zhou, G., and Li, M., 2000, "A Study on Ultrasonic Solid Horns for Flexural Mode," J. Acoust. Soc. Am., **107**(3), pp. 1358–1362.
[19] Sherrit, S., et al., 2004, "Novel Horn Designs for Power Ultrasonics," *Proceedings of the IEEE International Ultrasonics Symposium*, Montreal, August 24–27.
[20] Lynnworth, L. C., 1989, *Ultrasonic Measurements for Process Control: Theory, Techniques, Applications*, Academic, New York.
[21] Ffowcs Williams, J. E., and Hawkings, D. L., 1969, 1969, "Sound Generated by Turbulence and Surfaces in Arbitrary Motion," Philos. Trans. R. Soc. London, Ser. A, **264**, pp. 321–342.
[22] Rienstra, S. W., and Hirschberg, A., 2001, "An Introduction to Acoustics," Rep. No. IWDE 01-03, Eindhoven University of Technology, Eindhoven, The Netherlands.
[23] Courant, R., and Hilbert, D., 1962, *Methods of Mathematical Physics*, Interscience, New York.
[24] Al Kaufman, A., Levshin, A. L., and Larner, K., 2004, "The Role of Wave Propagation in the Motion of an Elastic Body," Eur. J. Phys., **25**, pp. 257–268.
[25] Baldock, G. R., and Bridgeman, T., 1981, *The Mathematical Theory of Wave Motion*, Ellis Horwood Limited, Chichester, UK.
[26] John, F., 1982, *Partial Differential Equations*, 4th ed., Springer, New York.

S. Haq
Faculty of Engineering Sciences,
GIK Institute,
Topi 23640, NWFP Pakistan

A. B. Movchan
Department of Mathematical Sciences,
University of Liverpool,
Liverpool L69 3BX, UK
e-mail: abm@liv.ac.uk

G. J. Rodin
Institute for Computational Engineering and Sciences,
The University of Texas at Austin,
Austin, TX 78712

Lattice Green's Functions in Nonlinear Analysis of Defects

A method for analyzing problems involving defects in lattices is presented. Special attention is paid to problems in which the lattice containing the defect is infinite, and the response in a finite zone adjacent to the defect is nonlinear. It is shown that lattice Green's functions allow one to reduce such problems to algebraic problems whose size is comparable to that of the nonlinear zone. The proposed method is similar to a hybrid finite-boundary element method in which the interior nonlinear region is treated with a finite element method and the exterior linear region is treated with a boundary element method. Method details are explained using an anti-plane deformation model problem involving a cylindrical vacancy. [DOI: 10.1115/1.2710795]

1 Introduction

Analysis of lattices is important for numerous applications involving solid state physics, engineered light-weight materials, and natural porous materials. In this paper, we are concerned with analysis of lattices whose response is linear elastic except for a finite region, where the response is nonlinear. Typically, such a region is adjacent to a defect(s) either inside the lattice or on its boundary. Ordinarily, such a problem gives rise to a sparse system of nonlinear algebraic equations whose size scales with that of the entire lattice. This may lead to significant computational costs for large lattices, and poses conceptual difficulties for infinite lattices, which arise frequently in analysis of defects.

In this paper, we propose a numerical method for infinite lattices that allows one to state the problem as a dense system of nonlinear algebraic equations whose size scales with that of the nonlinear zone. The central idea of the proposed method is to exploit lattice Green's functions for condensing the degrees of freedom in the linear zone to those on the interface between the linear and nonlinear zones. The method also applies to finite lattices. In those cases, the algebraic equations must include the degrees of freedom associated with the outer boundary of the linear zone.

The proposed method is similar to hybrid finite-boundary element methods in which the interior nonlinear region is treated with a finite element method and the exterior linear region is treated with a boundary element method. For lattices, the governing equations are algebraic and therefore discretization is not part of the problem. Furthermore, all basic ideas behind boundary element methods for continuum problems can be extended to problems on lattices [1]. Of course, on lattices, one obtains boundary algebraic equations (BAEs) as opposed to boundary integral equations, and BAE involve lattice Green's functions rather than classical Green's functions.

Lattice Green's functions have been widely used for analyzing defects in lattices. In particular, in Refs. [2–10] lattice Green's functions are used for formulating nonlocal boundary conditions for boundary-value problems associated with nonlinear fine scale models. In contrast to those approaches, the present approach provides a systematic method for constructing various BAE, by exploiting parallels between BAE and boundary integral equations.

In principle, nonlocal boundary conditions can be obtained using simpler continuum Green's functions. However, this approach gives rise to difficulties associated with matching discrete and continuum problems. Furthermore, continuum Green's functions are applicable only if both discrete and nonlinear effects are vanishingly small. In contrast, lattice Green's functions are applicable only if nonlinear effects are vanishingly small. This difference implies that approaches based on continuum Green's functions require larger computational models, which are often prohibitive [11].

To the best of our knowledge, the idea of replacing finite difference equations with BAE was first published by Saltzer in a technical report [12].[1] In the current terminology, Saltzer developed an indirect BAE corresponding to the discrete two-dimensional Laplacian stencil

$$u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j}$$

A numerical example provided by Saltzer does not support the use of BAE because the corresponding dense algebraic problem is more time-consuming than the original finite difference equations. Fifty years later, this conclusion still holds for relatively small problems. In contrast, large dense algebraic problems of size N can be solved with fast iterative solvers that deliver the solution using only $\mathcal{O}(N)$ storage and $\mathcal{O}(N)$ arithmetic operations. At this stage, such methods are well developed for continuum but not discrete problems; for references, see Refs. [13,14]. Nevertheless, asymptotic expansions of lattice Green's functions allow one to extend $\mathcal{O}(N)$ methods for continuum problems to problems on lattices [15].

We present our numerical method using a model scalar-valued problem involving a vacancy in an infinite simple square lattice. In general, mechanical response problems are formulated in three dimensions for the vector-valued displacements. From this perspective, the model problem can be regarded as an anti-plane elasticity problem for a cylindrical vacancy. Of course, the arising mathematical problem also describes conduction through the two-dimensional lattice.

The rest of the paper is organized as follows. In Sec. 2, we formulate the model boundary-value problem in terms of finite difference equations. In Sec. 3, we consider an infinite lattice and formulate the pertinent BAE. In Sec. 4, we derive expressions for the potential energy of a defect in an infinite lattice. In Sec. 5, we present a numerical example that demonstrates how the proposed method works. In Sec. 6, we summarize the results.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received November 11, 2004; final manuscript received August 20, 2006. Review conducted by Sanjay Govindjee.

[1] We are grateful to Proffessor Wolfgang Wendland for bringing this reference to our attention.

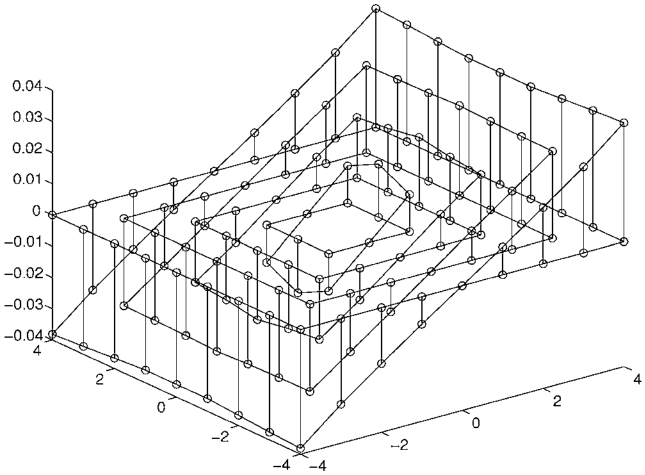


Fig. 1 Deformation of the lattice in the vicinity of the vacancy

2 Model Problem

Consider a *finite* two-dimensional lattice whose elementary cell is a unit square. The lattice contains a vacancy created by removing four links attached to the same node. The lattice is deformed such that the nodal displacements have only the out-of-plane component. The constitutive equation for each link relates the out-of-plane component of the internal force and the relative out-of-plane displacement of the ends δu

$$g = k\mathcal{F}(\delta u) \tag{1}$$

Here k is a constant and the function $\mathcal{F}$ is constructed so that $\mathcal{F}(0)=0$ and $\mathcal{F}'(0)=1$. Thus k can be regarded as the stiffness coefficient for $\delta u \ll 1$. Alternatively, one may regard g as the flux, u as the temperature, and k as the conductivity. The lattice is loaded such that the displacements along its external boundary Γ are prescribed as $u(\mathbf{n})=\boldsymbol{\gamma}\cdot\mathbf{n}$, where $\boldsymbol{\gamma}$ is a constant vector. The nodes along the vacancy (or internal) boundary are force free. A representative example of out-of-plane lattice deformation is shown in Fig. 1.

The governing equilibrium equation for the reference node $\mathbf{n}$ is written in the form

$$\mathcal{A}u(\mathbf{n}) + f(\mathbf{n}) = 0$$

with

$$\mathcal{A}u(\mathbf{n}) := \sum_{p\in\mathcal{N}(n)} g(\mathbf{p},\mathbf{n}) \tag{2}$$

where the sum is over the forces g exerted on the node $\mathbf{n}$ by the links adjacent to this node, and f is the external force. We choose the model problem such that $f=0$. Nevertheless, when we consider free-body diagrams for lattice subdomains, it is natural to regard forces acting on a subdomain as external.

For an internal node

$$\mathcal{N}(\mathbf{n}) = \{\mathbf{n}\pm\mathbf{e}_1, \mathbf{n}\pm\mathbf{e}_2\}$$

where $\mathbf{e}_1=\{1,0\}$, $\mathbf{e}_2=\{0,1\}$. Accordingly, the operator $\mathcal{A}$ can be expanded as

$$\mathcal{A}u(\mathbf{n}) = k\{\mathcal{F}[u(\mathbf{n}+\mathbf{e}_1)-u(\mathbf{n})] + \mathcal{F}[u(\mathbf{n}-\mathbf{e}_1)-u(\mathbf{n})] + \mathcal{F}[u(\mathbf{n}+\mathbf{e}_2)-u(\mathbf{n})] + \mathcal{F}[u(\mathbf{n}-\mathbf{e}_2)-u(\mathbf{n})]\}$$

For a boundary node, the number of the adjacent links is between two and four, and the operator $\mathcal{A}$ takes this into account.

Let us denote by Ω the nodes forming the lattice except for those that belong to Γ (Fig. 2). Then the model boundary-value problem is stated as

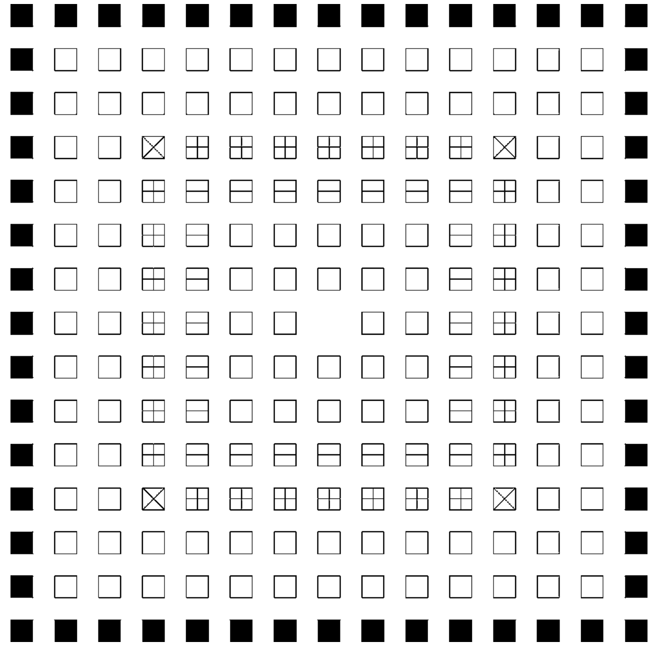


Fig. 2 Node and link sets involved in the analysis. The set Γ_-, formed by the nodes separating the linear and nonlinear zones, is denoted by the squares containing minus sign. The sets Γ_+ and $\mathcal{C}_+$ are denoted by the squares containing the plus and × signs, respectively. The union of these two sets is the first layer of the nodes in the linear zone. The set Γ, formed by the exterior nodes of a finite lattice, is denoted by black squares. All nodes inside Γ_- and Γ_- itself form the set Ω_-; the remaining nodes form the set $\bar{\Omega}_+$ and $\Omega_+=\bar{\Omega}_+\backslash\Gamma$. The links bounded by Γ_- and Γ_+ are denoted by Υ_0, the links bounded by Γ_+ are denoted by Υ_-, and the links bounded by $\Gamma_+\cup\mathcal{C}_+$, and Γ are denoted by Υ_+.

$$\mathcal{A}u(\mathbf{n}) = 0 \quad \mathbf{n}\in\Omega$$

$$u(\mathbf{m}) = \boldsymbol{\gamma}\cdot\mathbf{m} \quad \mathbf{m}\in\Gamma \tag{3}$$

This boundary-value problem has a unique solution as long as $\mathcal{F}'(x)>0$ [16].

There are good reasons to consider problems involving *infinite* lattices. Such problems allow one to isolate defects and study their basic properties. However, infinite lattices pose major difficulties for the stated boundary-value problem simply because in the limit the problem size tends to infinity. One may try to construct an approximate solution by considering a sequence of finite lattices, with the expectation that such a sequence converges sufficiently fast. Here we attack the problem differently, by assuming that $|\boldsymbol{\gamma}|\ll 1$. This allows us to linearize the operator $\mathcal{A}u(\mathbf{n})$ at the nodes far away from the defect:

$$\mathcal{A}u(\mathbf{n}) \approx \mathcal{L}u(\mathbf{n}) := k[u(\mathbf{n}+\mathbf{e}_1)+u(\mathbf{n}-\mathbf{e}_1)+u(\mathbf{n}+\mathbf{e}_2)+u(\mathbf{n}-\mathbf{e}_2) - 4u(\mathbf{n})]$$

In the next section, it is shown that this approximation is sufficient for formulating a *finite* system of equations for an *infinite* lattice.

3 Infinite Lattice

For now, Ω remains finite but sufficiently large, so that it contains a $(2N)\times(2N)$ square centered at the vacancy, such that every link outside of this square exhibits the linear response. We introduce the following definitions (see Fig. 2):

1. Ω_- is the set of nodes forming the square containing the vacancy; $\Omega_+ := \Omega\backslash\Omega_-$ and $\bar{\Omega}_+ := \Omega_+\cup\Gamma$;

2. Γ_- is the set of the $8N+4$ nodes along the outer boundary of Ω_-;
3. Γ_+ is the set of the $8N+4$ nodes along the inner boundary of Ω_+. This set does not include the four corner nodes, which we denote by $\mathcal{C}_+$;
4. Υ_0 is the set of the $8N+4$ links, each having one node in Γ_- and one node in Γ_+. We denote the elements of this set by $\langle \mathbf{m},\mathbf{n}\rangle$ with the provision that $\mathbf{m}\in\Gamma_-$ and $\mathbf{n}\in\Gamma_+$.

Now we can replace Eq. (3) with

$$\mathcal{A}_- u(\mathbf{n}) + f_-(\mathbf{n})\delta(\mathbf{n},\mathbf{m}) = 0 \quad \mathbf{n}\in\Omega_- \quad \mathbf{m}\in\Gamma_- \tag{4}$$

and

$$\mathcal{L}_+ u(\mathbf{n}) + f_+(\mathbf{n})\delta(\mathbf{n},\mathbf{m}) = 0 \quad \mathbf{n}\in\Omega_+ \quad \mathbf{m}\in\Gamma_+$$

$$u(\mathbf{p}) = \boldsymbol{\gamma}\cdot\mathbf{p} \quad \mathbf{p}\in\Gamma \tag{5}$$

where f_- and f_+ are the forces exerted by the links from Υ_0 on Γ_- and Γ_+, respectively. The subscripts of the operators $\mathcal{A}$ and $\mathcal{L}$ denote that the operators are defined with respect to the domains Ω_- and Ω_+, respectively, and δ is the Kronecker's symbol defined by

$$\delta(\mathbf{n},\mathbf{m}) = \begin{cases} 1 & \text{for } \mathbf{n}=\mathbf{m} \\ 0 & \text{for } \mathbf{n}\neq\mathbf{m} \end{cases}$$

By construction, the links from Υ_0 are characterized by the linear constitutive equation and therefore the internal force for the link $\langle \mathbf{m},\mathbf{n}\rangle\in\Upsilon_0$ can be expressed as

$$g(\mathbf{m},\mathbf{n}) = k[u(\mathbf{n}) - u(\mathbf{m})] \quad \mathbf{m}\in\Gamma_-, \quad \mathbf{n}\in\Gamma_+ \tag{6}$$

The equilibrium conditions dictate

$$f_-(\mathbf{m}) = \sum_{\langle \mathbf{m},\mathbf{n}\rangle\in\Upsilon_0} g(\mathbf{m},\mathbf{n}) \tag{7}$$

and

$$f_+(\mathbf{n}) = -g(\mathbf{m},\mathbf{n}) \qquad \langle \mathbf{m},\mathbf{n}\rangle\in\Upsilon_0 \tag{8}$$

Equations (7) and (8) are different because each corner node of Γ_- is connected to two links from Υ_0, while each node of Γ_+ is connected to only one link from Υ_0. We can combine Eqs. (6)–(8) and write

$$f_-(\mathbf{m}) = \sum_{\langle \mathbf{m},\mathbf{n}\rangle\in\Upsilon_0} k[u(\mathbf{n}) - u(\mathbf{m})] \tag{9}$$

and

$$f_+(\mathbf{n}) = -k[u(\mathbf{n}) - u(\mathbf{m})] \qquad \langle \mathbf{m},\mathbf{n}\rangle\in\Upsilon_0 \tag{10}$$

On the surface, the only advantage of the new formulation is that it uses the linear equations in Ω_+, which is a marginal simplification because the equations in Ω_- remain nonlinear. Nevertheless, in what follows we demonstrate that, in the limit as the lattice expands to infinity, one can replace Eq. (5) with a BAE on Γ_+, and as a result reduce the original problem to that defined on Ω_-. This approach does not affect the finite difference equations on Ω_-—it merely supplements them with nonlocal boundary conditions. Alternatively, one can reduce the original problem to that defined on $\Omega_-\cup\Gamma_+$.

The BAE on Γ_+ is formulated using the reciprocity theorem and the fundamental solution, following the standard procedure for the corresponding boundary integral equations (e.g. Ref. [17]). Accordingly, we define the fundamental solution $U(\mathbf{m},\mathbf{n})$ as the displacement at the node $\mathbf{n}$ induced by a unit force applied at the node $\mathbf{m}$ in an infinite perfect lattice

$$U(\mathbf{m},\mathbf{n}) = \frac{1}{16\pi^2 k}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \frac{\exp\{i[\xi_1(m_1-n_1)+\xi_2(m_2-n_2)]\}-1}{\sin^2\dfrac{\xi_1}{2}+\sin^2\dfrac{\xi_2}{2}} d\xi_1\, d\xi_2$$

According to Ref. [18], this expression and its analysis were first presented by Duffin and Shaffer in a technical report; see also Ref. [15]. For large $|\mathbf{m}-\mathbf{n}|$, as $U(\mathbf{m},\mathbf{n})$ approaches the classical fundamental solution [18]

$$U(\mathbf{m},\mathbf{n}) \sim -(2\pi k)^{-1}\log|\mathbf{m}-\mathbf{n}|$$

Now we consider three mechanical states defined on $\bar{\Omega}_+$:

State 1. Subject $\bar{\Omega}_+$ to the boundary displacements $u(\mathbf{n})=\boldsymbol{\gamma}\cdot\mathbf{n}$ on $\partial\bar{\Omega}_+$. By inspection one can establish that these boundary conditions induce spatially uniform forces in the links, in the sense that all links aligned along $\mathbf{e}_1$ ($\mathbf{e}_2$) transmit the same force $k\gamma_1$ ($k\gamma_2$). We denote the boundary displacements and forces of this state by $u^{(0)}$ and $f^{(0)}$, respectively.

State 2. Subject $\bar{\Omega}_+$ to the displacements $u(\mathbf{n})=\boldsymbol{\gamma}\cdot\mathbf{n}$ on Γ, the forces $f_+(\mathbf{n})$ on Γ_+, and $f=0$ on $\mathcal{C}_+$. We denote the boundary displacements and forces of this state by u and f, respectively.

State 3. Subject the infinite lattice to a unit force applied at $\mathbf{n}\in\Gamma_+$, so that the induced displacements are given by $U(\mathbf{m},\mathbf{n})$. Isolate $\bar{\Omega}_+$ from the infinite lattice by replacing the links connecting $\bar{\Omega}_+$ to the rest of the infinite lattice with their internal forces. For $\mathbf{q}\in\Gamma_+$, those forces can be computed as

$$F(\mathbf{q},\mathbf{n}) = k[U(\mathbf{p},\mathbf{n}) - U(\mathbf{q},\mathbf{n})] \quad \langle \mathbf{p},\mathbf{q}\rangle\in\Upsilon_0 \text{ and } \mathbf{n}\in\Gamma_+ \tag{11}$$

Note that $F(\mathbf{q},\mathbf{n})=0$ if $\mathbf{q}\in\mathcal{C}_+$. Also note that the entire set of forces acting on $\partial\bar{\Omega}$ includes $F(\mathbf{q},\mathbf{n})$ and the unit force applied at $\mathbf{n}$.

Upon application of the reciprocity theorem (known as the Betti–Maxwell theorem in structural mechanics) to the first and third states, we obtain the BAE

$$u^{(0)}(\mathbf{q}) + \sum_{n\in\partial\bar{\Omega}_+} F(\mathbf{n},\mathbf{q})u^{(0)}(\mathbf{n}) = \sum_{n\in\partial\bar{\Omega}_+} U(\mathbf{n},\mathbf{q})f^{(0)}(\mathbf{n}) \qquad \mathbf{q}\in\Gamma_+$$

In formulating this equation, the third state was generated by applying a unit force at $\mathbf{q}$. Note that in both states the forces at $\mathcal{C}_+$ are equal to zero. Similarly, for the second and third states we obtain the BAE

$$u(\mathbf{q}) + \sum_{n\in\partial\bar{\Omega}_+} F(\mathbf{n},\mathbf{q})u(\mathbf{n}) = \sum_{n\in\partial\bar{\Omega}_+} U(\mathbf{n},\mathbf{q})f(\mathbf{n}) \qquad \mathbf{q}\in\Gamma_+$$

By subtracting the two BAE we obtain

$$u(\mathbf{q}) - u^{(0)}(\mathbf{q}) + \sum_{n\in\Gamma_+} F(\mathbf{n},\mathbf{q})[u(\mathbf{n}) - u^{(0)}(\mathbf{n})] = \sum_{n\in\Gamma_+} U(\mathbf{n},\mathbf{q})[f_+(\mathbf{n}) - f^{(0)}(\mathbf{n})] + \sum_{n\in\Gamma} U(\mathbf{n},\mathbf{q})[f(\mathbf{n}) - f^{(0)}(\mathbf{n})] \tag{12}$$

Now we can consider the limit as Γ tends to infinity while Γ_+ is fixed. To this end we define the metric of Γ as the radius R of the largest circle Γ can circumscribe. By construction

$$\sum_{n\in\Gamma_+} f(\mathbf{n}) = \sum_{n\in\Gamma_+} f^{(0)}(\mathbf{n})$$

Then, to a leading order, an observer on Γ regards the system of forces $f-f^{(0)}$ applied to Γ_+ as a dipole. For large R, the asymptotic behavior on the lattice is similar to that in a continuum body. Accordingly, $f-f^{(0)}$ on Γ behaves similar to the continuum stress field induced by a dipole, which decays as R^{-2}. This decay is sufficient to overcome the logarithmic growth rate of U and the linear growth rate associated with the summation. As a result in the limit as $R\to\infty$, Eq. (12) can be expressed as

$$u(\mathbf{q}) - u^{(0)}(\mathbf{q}) + \sum_{n\in\Gamma_+} F(\mathbf{n},\mathbf{q})[u(\mathbf{n}) - u^{(0)}(\mathbf{n})] = \sum_{n\in\Gamma_+} U(\mathbf{n},\mathbf{q})[f_+(\mathbf{n}) - f^{(0)}(\mathbf{n})] \quad \mathbf{q}\in\Gamma_+ \tag{13}$$

This equation is the desired BAE on Γ_+. It can be combined with Eqs. (9) and (10) to eliminate f_- from Eq. (4), so that the entire problem can be stated on Ω_-. Alternatively, one may combine Eqs. (4), (9)–(11), and (13) to state the problem on $\Omega_-\cup\Gamma_+$

$$\mathcal{A}_-u(\mathbf{m}) + \sum_{\langle \mathbf{m},\mathbf{n}\rangle\in\Upsilon_0} k[u(\mathbf{n}) - u(\mathbf{m})] = 0 \quad \mathbf{m}\in\Omega_- \tag{14}$$

$$u(\mathbf{q}) - u^{(0)}(\mathbf{q}) + \sum_{n\in\Gamma_+} k[U(\mathbf{p},\mathbf{q}) - U(\mathbf{n},\mathbf{q})][u(\mathbf{n}) - u^{(0)}(\mathbf{n})] = \sum_{n\in\Gamma_+} U(\mathbf{n},\mathbf{q})[ku(\mathbf{p}) - ku(\mathbf{n}) - f^{(0)}(\mathbf{n})] \quad \mathbf{q}\in\Gamma_+,\ \langle\mathbf{p},\mathbf{n}\rangle\in\Upsilon_0$$

Note that it is acceptable to replace Eq. (14) with the expression

$$\mathcal{F}[u(\mathbf{m}+\mathbf{e}_1) - u(\mathbf{m})] + \mathcal{F}[u(\mathbf{m}-\mathbf{e}_1) - u(\mathbf{m})] + \mathcal{F}[u(\mathbf{m}+\mathbf{e}_2) - u(\mathbf{m})] + \mathcal{F}[u(\mathbf{m}-\mathbf{e}_2) - u(\mathbf{m})] = 0 \quad \mathbf{m}\in\Omega_-$$

which does not take into account that the response of the links in Υ_0 is linear.

4 Energy Analysis

Energy analysis plays a central role in mechanics of structures and defects. In particular, using energy concepts, one can analyze structural stability and identify thermodynamic forces driving evolution of defects.

The strain energy of a link subjected to the relative displacement δu is defined as

$$w(\delta u) = \int_0^{\delta u} k\mathcal{F}(\xi)d\xi$$

For the linearized constitutive equation, this expression is simplified to

$$w(\delta u) = \tfrac{1}{2}k\,\delta u^2$$

The strain energy W of a *finite* lattice is defined as the sum of the strain energies of the individual links. If the lattice is subjected to prescribed boundary displacements, then the strain and potential energies are equal: $W=\Pi$. In this section, we restrict our analysis to this case.

The potential energy of a defect is defined as the difference between the potential energy of the lattice containing the defect and the strain energy of the perfect lattice

$$\Delta\Pi = \Pi - \Pi^{(0)} = W - W^{(0)}$$

Here, following the development in the previous section, we assume that $\Pi^{(0)}$ and $W^{(0)}$ are consistent with the linearized constitutive equation, so that, using the virtual work principle, we can express

$$\Pi^{(0)} = W^{(0)} = \frac{1}{2}\sum_{n\in\Gamma} u^{(0)}(\mathbf{n})f^{(0)}(\mathbf{n}) \tag{15}$$

To compute the potential energy of the lattice with the defect, we begin with splitting the links into two nonintersecting sets. The first set Υ_- contains all the links with at least one node in the set Ω_-. The complementary set is denoted by Υ_+. Then

$$\Pi = W = \sum_{l\in\Upsilon_-} w(\mathbf{l}) + \sum_{l\in\Upsilon_+} w(\mathbf{l})$$

The first term in this equation is elementary to compute once the displacements on $\Omega_-\cup\Gamma_+$ have been determined. Using the virtual work principle, the second sum can be expressed in terms the sum over $\Gamma_+\cup\Gamma$, so that

$$\Pi = \sum_{l\in\Upsilon_-} w(\mathbf{l}) + \frac{1}{2}\sum_{n\in\Gamma_+\cup\Gamma} f(\mathbf{n})u(\mathbf{n}) \tag{16}$$

Upon subtraction of Eq. (15) from Eq. (16), we obtain

$$\Delta\Pi = \sum_{l\in\Upsilon_-} w(\mathbf{l}) + \frac{1}{2}\sum_{n\in\Gamma}[f(\mathbf{n})u(\mathbf{n}) - f^{(0)}(\mathbf{n})u^{(0)}(\mathbf{n})] + \frac{1}{2}\sum_{\mathbf{n}\in\Gamma_+} f(\mathbf{n})u(\mathbf{n}) \tag{17}$$

Since, for $\mathbf{n}\in\Gamma$, $u(\mathbf{n})=u^{(0)}(\mathbf{n})$, the sum over Γ can be expressed as

$$\sum_{n\in\Gamma}[f(\mathbf{n})u(\mathbf{n}) - f^{(0)}(\mathbf{n})u^{(0)}(\mathbf{n})] = \sum_{n\in\Gamma}[f(\mathbf{n})u^{(0)}(\mathbf{n}) - f^{(0)}(\mathbf{n})u(\mathbf{n})]$$

The reciprocity theorem allows us to replace the sum on the right-hand side of this expression with a sum over Γ_+

$$\sum_{n\in\Gamma}[f(\mathbf{n})u^{(0)}(\mathbf{n}) - f^{(0)}(\mathbf{n})u(\mathbf{n})] = -\sum_{n\in\Gamma_+}[f(\mathbf{n})u^{(0)}(\mathbf{n}) - f^{(0)}(\mathbf{n})u(\mathbf{n})]$$

and therefore Eq. (17) takes the form

$$\Delta\Pi = \sum_{l\in\Upsilon_-} w(\mathbf{l}) + \frac{1}{2}\sum_{n\in\Gamma_+}[f(\mathbf{n})u(\mathbf{n}) + f^{(0)}(\mathbf{n})u(\mathbf{n}) - f(\mathbf{n})u^{(0)}(\mathbf{n})] \tag{18}$$

This expression does not pose any difficulties in the limit as Ω expands to infinity.

The development under traction-prescribed boundary conditions on Γ follows essentially the same steps, and leads to the same answer.

5 Numerical Examples

In this section, for demonstration purposes, we apply the current method to stability analysis of an infinite lattice containing a vacancy. In the spirit of the classical papers of Frenkel and Kontorova [19], Peierls [20], and Nabarro [21], we adopt the constitutive equation in the form

$$g = k\mathcal{F}(\delta u) = \frac{k}{2\pi}\sin 2\pi\delta u$$

For this equation, the issue of stability arises naturally because $\mathcal{F}'(\delta u)$ changes its sign. For the boundary conditions restricted to the form

$$u(\mathbf{n}) = \boldsymbol{\gamma}\cdot\mathbf{n} = \{\gamma,0\}\cdot\{n_1,n_2\}$$

we seek the smallest value of γ that gives rise to singular algebraic equations; we denote that value by γ_c.

The solution strategy involves the following steps:

1. Set $N=1$;
2. Choose a value for γ;
3. For the chosen γ, compute the solution of Eq. (14);
4. For the computed solution, evaluate the Jacobian $\mathbf{J}$ corresponding to Eq. (14);
5. Adjust the value of γ until one of the eigenvalues of $\mathbf{J}$ becomes close to zero. The corresponding γ provides γ_c for the current N; and

Table 1 Dependence of the critical value γ_c on the number of nonlinear layers *N* around the vacancy

N	1	2	3	4	5	6	7	8	9	10
γ_c	1.124	0.223	0.201	0.194	0.189	0.188	0.186	0.1855	0.1852	0.1851

6. Increase N and repeat steps 2–5 until $|\gamma_c(N+1)-\gamma_c(N)|$ is sufficiently small.

Note that an analytic expression for **J** is straightforward to obtain upon differentiation of Eq. (14).

Computational results are presented in Table 1. They clearly indicate that the computational procedure converges as N increases. Furthermore, they suggest that, for this problem, the nonlinear effects are highly localized, and can be neglected several lattice spacings away from the vacancy.

6 Closure

In this paper, we presented a numerical method for analyzing lattice defects using lattice Green's functions. In contrast to existing approaches, the present method exploits lattice Green's functions following ideas well established for boundary integral equations. This use of lattice Green's functions was proposed by Martinsson and Rodin [1] who referred to the arising equations as BAEs. In this paper, BAEs are used to derive nonlocal boundary conditions on a finite zone surrounding the defect. In a companion paper [22], we considered applications of lattice Green's functions to models of interfacial and linear defects.

BAE can be particularly beneficial for truly large-scale problems involving lattices. First, nonlocal boundary conditions based on lattice Green's functions are more effective than those based on continuum Green's functions because they take into account the discrete lattice structure. Second, BAEs provide a direct link to fast iterative procedures that exploit preconditioners and fast summation methods for matrix–vector multiplication. Furthermore, preliminary computations indicate that BAE are particularly well suited for such methods because they lead to well-conditioned algebraic problems [1].

Acknowledgment

We would like to thank the U.S. National Science Foundation and the Institute for Computational Engineering and Science for providing financial support under Grant Nos. CMS-0219839 and ACI-0220037. We are grateful to the reviewers for their constructive criticism that helped to improve the paper.

References

[1] Martinsson, P. G., and Rodin, G. J., 2003, "Boundary Algebraic Equations for Lattice Problems," IUTAM Symposium on Asymptotics, Singularities and Homogenisation in Problems of Mechanics, Proceedings of the IUTAM Symposium, Liverpool, UK, July 8–11, A. B. Movchan, ed., pp. 191–198.

[2] Masudajindo, K., Tewary, V. K., and Thomson, R., 1991, "Theoretical Study of the Fracture of Brittle Materials: Atomistic Calculations Materials," Mater. Sci. Eng., A, **146**(1–2), pp. 273–289.

[3] Masudajindo, K., Tewary, V. K., and Thomson, R., 1991, "Atomic Theory of Fracture of Brittle Materials: Application to Covalent Semiconductors," J. Mater. Res. **6**(7), pp. 1553–1566.

[4] Thomson, R., Zhou, S. J., Carlsson, A. E., and Tewary, V. K., 1992, "Lattice Imperfections Studied by Use of Lattice Green-Functions," Phys. Rev. B, **46**(17), pp. 10613–10622.

[5] Zhou, S. J., Carlsson, A. E., and Thomson, R., 1993, "Dislocation Nucleation and Crack Stability—Lattice Greens-Function Treatment of Cracks in a Model Hexagonal Lattice," Phys. Rev. B, **47**(13), pp. 7710–7719.

[6] Zhou, S. J., and Curtin, W. A., 1995, "Failure of Fiber Composites: A Lattice Green-Function Model," Acta Metall. Mater., **43**(8), pp. 3093–3104.

[7] Schiotz, J., and Carlsson, A. E., 1997, "Calculation of Elastic Green's Functions for Lattices with Cavities," Phys. Rev. B, **56**(5), pp. 2292–2294.

[8] Rao, S., Hernandez, G., Simmons, J. P., Parthasarathy, T. A., and Woodward, C., 1998, "Green's Function Boundary Conditions in Two-Dimensional and Three-Dimensional Atomistic Simulations of Dislocations," Philos. Mag. A, **77**(1), pp. 231–256.

[9] Masuda-Jindo, K., Menon, M., and Van Hung, V., 2001, "Atomistic Study of Fracture of Nanoscale Materials by Molecular Dynamics and Lattice Green's Function Methods," J. Phys. IV, **11**(PR5), pp. 11–18.

[10] Wang, S., 2002, "Lattice Theory for Structure of Dislocations in a Two-Dimensional Triangular Crystal," Phys. Rev. B, **65**(9), p. 094111.

[11] Saito, Y., 2004, "Elastic Lattice Green's Function in Three Dimensions," J. Phys. Soc. Jpn., **73**, pp. 1816–1826.

[12] Saltzer, C., 1958, "Discrete Potential Theory for Two-Dimensional Laplace and Poisson Difference Equations," Paper No. NACA TN 4086.

[13] Greengard, L., and Rokhlin, V., 2003, "A New Version of the Fast Multipole Method for the Laplace Equation in Three Dimensions," Acta Numerica, **6**, pp. 229–269.

[14] Nishimura, N., 2002, "Fast Multipole Accelerated Boundary Integral Equation Methods," Appl. Mech. Rev., **55**, pp. 299–324.

[15] Martinsson, P. G., and Rodin, G. J., 2003, "Asymptotic Expansions of Lattice Green's Functions," Proc. R. Soc. London, Ser. A, **458**(2027), pp. 2609–2622.

[16] Necas, J., and Hlavacek, J., 1981, *Mathematical Theory of Elastic and Elasto-Plastic Bodies: An Introduction*, Elsevier/North Holland, Amsterdam, Netherlands.

[17] Bonnet, M., 1999, *Boundary Integral Equation Methods for Solids and Fluids*, Wiley, New York.

[18] Duffin, R. J., and Shelly, E. P., 1958, "Difference Equations of Polyharmonic Type," Duke Math. J., **25**, pp. 209–238.

[19] Frenkel, J., and Kontorova, T., 1938, "On the Theory of Plastic Deformation and Twinning," J. Exp. Theor. Phys., **8**, pp. 89–91.

[20] Peierls, R., 1940, "The Size of a Dislocation," Proc. Phys. Soc. London, **52**, pp. 34–37.

[21] Nabarro, F. R. N., 1947, "Dislocations in a Simple Cubic Lattice," Proc. Phys. Soc. London, **59**, pp. 256–272.

[22] Haq, S., Movchan, A. B., and Rodin, G. J., 2006, "Analysis of Lattices With Non-Linear Interphases," Acta Mech. Sin., **22**, pp. 323–330.

Complex Flow Dynamics in Dense Granular Flows—Part II: Simulations

Piroz Zamankhan
Laboratory of Computational Fluid & BioFluid Dynamics,
Lappeenranta University of Technology,
Lappeenranta 53851, Finland
and
Power and Water University of Technology,
School of Energy Engineering,
P.O. Box 16765-1719,
Tehran, Iran
e-mail: qpz002000@yahoo.com

Jun Huang
Laboratory of Computational Fluid & BioFluid Dynamics,
Lappeenranta University of Technology,
Lappeenranta 53851, Finland

By applying a methodology useful for analysis of complex fluids based on a synergistic combination of experiments, computer simulations, and theoretical investigation, a model was built to investigate the fluid dynamics of granular flows in an intermediate regime, where both collisional and frictional interactions may affect the flow behavior. In Part I, experiments were described using a modified Newton's Cradle device to obtain values for the viscous damping coefficient, which were scarce in the literature. This paper discusses detailed simulations of frictional interactions between the grains during a binary collision by employing a numerical model based on finite element methods. Numerical results are presented of slipping, and sticking motions of a first grain over the second one. The key was to utilize the results of the aforementioned comprehensive model in order to provide a simplified model for accurate and efficient granular-flow simulations with which the qualitative trends observed in the experiments can be captured. To validate the model, large scale simulations were performed for the specific case of granular flow in a rapidly spinning bucket. The model was able to reproduce experimentally observed flow phenomena, such as the formation of a depression in the center of the bucket spinning at high frequency of 100 rad/s. *This agreement suggests that the model may be a useful tool for the prediction of dense granular flows in industrial applications, but highlights the need for further experimental investigation of granular flows in order to refine the model.* [DOI: 10.1115/1.2711219]

Keywords: dense granular flows, computer simulations, friction, nonlinear dynamics, rotating bucket

1 Introduction

The behavior of granular materials such as sand or powders is one of the mysterious problems of modern science [1]. As a collection of inelastic particles, the flow of granular media is neither similar to that of solids nor gases nor even liquids. As particle inelasticity is increased and the collisions become more and more inelastic, simple hydrodynamics clearly breaks down and the occurrence of the phenomenon of inelastic collapse can be observed [2].

A most intriguing phenomenon in the mechanics of granular material is size segregation, observed in models of vibrated granular mixtures such as powders or sand. Several mechanisms have been proposed to explain the process of de-mixing the different components of the system under shaking. However, the criteria for predicting segregation in a mixture, which is of great practical importance, are largely unknown.

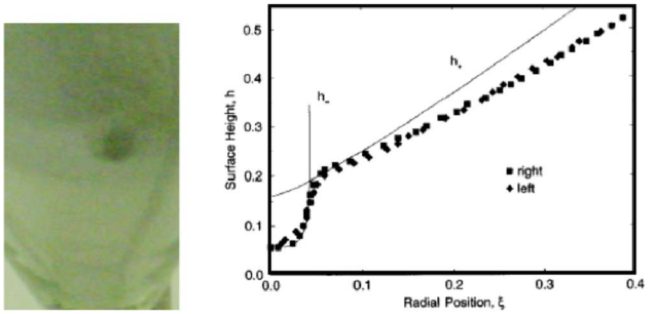
Size segregation is observed in avalanches [3]. Avalanches are shallow, gravity driven, free surface flows of solid particles, which occur when a surface layer of granular material becomes unstable, and may flow at speeds over 200 km/h. Bak [4], inspired by avalanches in a sand pile, discovered the phenomenon of self-organized criticality, which provides a plausible explanation for many natural phenomena. Gray and Hutter [5] have observed that in conjunction with particle size segregation within the flowing avalanche and the occurrence of dispersed shock waves, the avalanche quickly comes to rest.

In silos when a small hole is opened at its center base the grains develop an internal core flow and a V-shaped depression may be formed, as illustrated in Fig. 1(*a*). It is somewhat surprising that Baxter and Yeung [6] reported that at high shear rotation rates in a spinning bucket of sand a depression develops along the rotation axis, as shown in Fig. 1(*b*). According to Gray and Hutter [5], in silos, grains on either side of the core are at rest or the pine tree pattern is preserved there. Apparently, the grains are fed to the core by a sequence of intermittent avalanches that flow down the faces of the depression.

Again inspired by avalanches in a silo, Baxter and Yeung [6] have suggested that at low rotation rates in a rotating bucket, a central region could occur whose slope is significantly less than the critical slope. By utilizing the minimal-ingredient theories of granular surface behavior [7], they, with some degree of success reproduced the central subcritical region observed experimentally at low shear rates. However, this model is based on the fundamental assumption of a thin flowing layer, and hence did not succeed for predicting the flow dynamic behavior of the grains at moderate rotation rates for which the shrinking of the height of the cusp in the central region was observed, even though the rotation rate was held fixed [6]. In addition, it appears that a fundamentally different model from that proposed in Baxter and Yeung [6] is required to reproduce the steep depression around the center observed at high rotation.

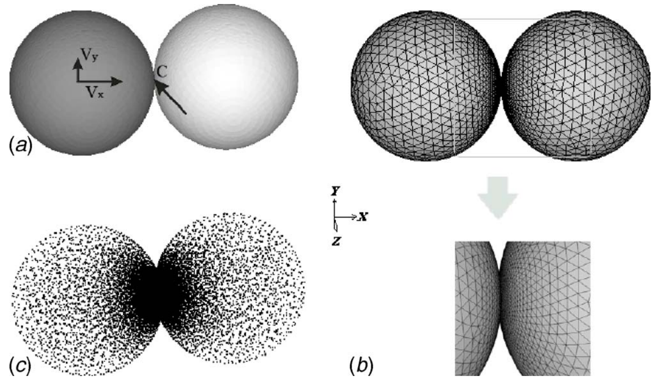
Note that the behavior of granular flow depends on microscopic effects, far or close to the scale which can be observed with the eye. For these systems, it may not even be clear how to begin constructing an approximate theory in a reasonable way. In this light, computer experiments have a valuable role to play in providing essentially exact results for problems of complex fluid. It may be difficult to carry out experiments under different conditions, while a computer experiment of the complex fluids would be perfectly feasible. Quite subtle details of particle motion, structure and other micromechanical effects in granular flows are difficult to probe experimentally, but can be extracted readily from a computer experiment.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received January 2, 2005; final manuscript received September 19, 2006. Review conducted by Kenneth M. Liechti.



(*a*) (*b*)

Fig. 1 (*a*) A steep depression observed in flowing fine grains through a small hole in silos. (*b*) Surface shape at high rotation rate reported in Baxter and Yeung [6].



(*a*) (*c*) (*b*)

Fig. 2 (*a*) Two monosized colliding, rough spheres with diameter of σ at the beginning of the approach period. The sphere on the left with the axial and tangential velocity components of V_x, and V_y, respectively, is brought into contact with the sphere on the right, which is initially stationary. They touch initially at a single point *C*. (*b*) 3D finite element mesh for the spheres as shown in (*a*). More than 4×10^5 tetrahedral elements were used in the numerical treatments. The elements at the vicinity of point of initial contact are magnified in the inset (*c*). The spatial distribution of nodes used in the numerical treatments. Notice the fine zone in the vicinity of contact area where the gradient of stresses and strains are high.

As an example, the surfaces of grains may be characterized as randomly rough with surface roughness on many different length scales. Generally speaking, surface properties such as contact area between the colliding grains in a granular flow are difficult to be measured accurately by experiments. In fact, when two rough grains are brought into contact, the area of real contact is a small fraction of the nominal contact area. The real contact area between the grains has a direct importance for sliding friction. It also has a major influence on adhesive forces between two grains. Therefore it is advantageous to utilize computer experiments to test existing contact models for rough solids and to point the way towards new models.

The objective of this attempt is to develop a model capable of capturing the complex behavior of granular flows. The model will be validated through its application to the above-mentioned unexplained hydrodynamic phenomena that are observed in spinning buckets. The organization of the present paper is as follows. In Sec. 2, the finite element-based model proposed in the previous work [8] was generalized to take into account contact between rough grains in order to improve the accuracy of the estimated contact stresses in the interacting grains. By utilizing the results of the aforementioned finite element-based model, a simplified model was presented with the minimum level of detail for accurate and efficient granular-flow simulations. In Sec. 3, the developed simplified model was used to perform simulations of the specific example of granular flows in a spinning bucket, which is a system of industrial relevance whose dynamic behavior is not well understood. The model results were compared with available experimental data. Finally in Sec. 4, the concluding remarks were presented and recommendations for future study were developed based on the comparisons.

2 Force-Driven Models

The simulation of an impact between bodies is computational complicated due to the dependence of the boundary conditions of the bodies on the solution variables [9]. A collision algorithm consists of two parts, which are a contact search algorithm that identifies penetration between the bodies [10–14], and a general contact algorithm that satisfies the kinematic contact condition as well as calculates the normal and shear stresses on the interacting surfaces [15–17]. A model for colliding rough spheres for accurate granular simulations is presented in the following section.

2.1 Finite Element Model for Colliding Rough Spheres. In a finite element model as presented in Ref. [8], the influence of the surface roughness on collision behavior can be included using a highly detailed model in which the actual geometry of asperities are considered [18]. When magnifying a glass ball surface about 100 times, rough contours, which called asperity, can be seen much lager than molecular dimensions. The detailed model of asperities will be very large and it is likely to be impractical when modeling behavior of colliding glass balls. However, the detailed model of a single asperity could be of use to provide valuable insights in simpler models.

In the present attempt a simplified model is used for which friction is modeled based on a Coulomb formulation. In the model as presented in Fig. 2, micromechanical effects are modeled as nonlinear interface stiffness, k, and a coefficient of friction, μ, which depends on sliding distance, sliding velocity, temperature and pressure.

The finite element contact modeling presented in Ref. [8] is based on a master–slave approach. In this approach, the nodes on the slave surface are not allowed to penetrate the segments of the master surface. The frictional algorithm implemented uses the equivalent of an elasto–plastic spring.

When a slave node penetrates a master segment in a system as illustrated in Fig. 3(*a*) at time t_n, a compatibility normal force, $F_n^{(N)}$, is introduced to the master segment. As outlined in the previous work [8], the normal forces are calculated with the augmented Lagrangian method. By computing the incremental movement of the slave node, δ_e, the interface force is updated to a trial value $F^{(t)}=F_n^{(f)}-k\delta_e$. The interface forces are then updated for all contact nodes that are marked with an (in contact) tag. By checking the yield condition, if $|F^{(t)}|\leq\mu|F_n^{(N)}|$ then the frictional force is updated to $F_{n+1}^{(f)}=F^{(t)}$, and if $|F^{(t)}|>\mu|F_n^{(N)}|$ then $F_{n+1}^{(f)}$ is set to $\mu|F_n^{(N)}|F^{(t)}/|F^{(t)}|$. Here, the coefficient of friction is estimated using an expression given as $\mu=\mu_d+\Delta\mu e^{-b|\delta_e/\Delta t|}$. The interface shear stress that develops as a result of Coulomb friction can be nonuniform whose maximum value is not at the initial point of contact, as illustrated in Fig. 3(*c*). However, the normal stress appears to be symmetric about the initial point of contact as depicted in Fig. 3(*b*). The effective stress, σ_{eff}, defined as $1/\sqrt{2}[(\sigma_{xx}-\sigma_{yy})^2+(\sigma_{xx}-\sigma_{zz})^2+(\sigma_{yy}-\sigma_{zz})^2+6(\sigma_{xy}^2+\sigma_{yz}^2+\sigma_{xz}^2)]^{1/2}$, is a useful quantity with regard to formulation of strain hardening rules. Figure 3(*d*) represents the computed contours of the effective stress around the initial point of contact, *C*.

The numerical results shown in Fig. 3 represent the collision of two monosized, elastic, rough, glass balls with diameter of $\sigma=2$ cm. The material properties chosen for the balls are listed in Table 1. The static and dynamic coefficients of friction are set to $\mu_s=0.85$, $\mu_d=0.7$, respectively, and the decay constant is set to

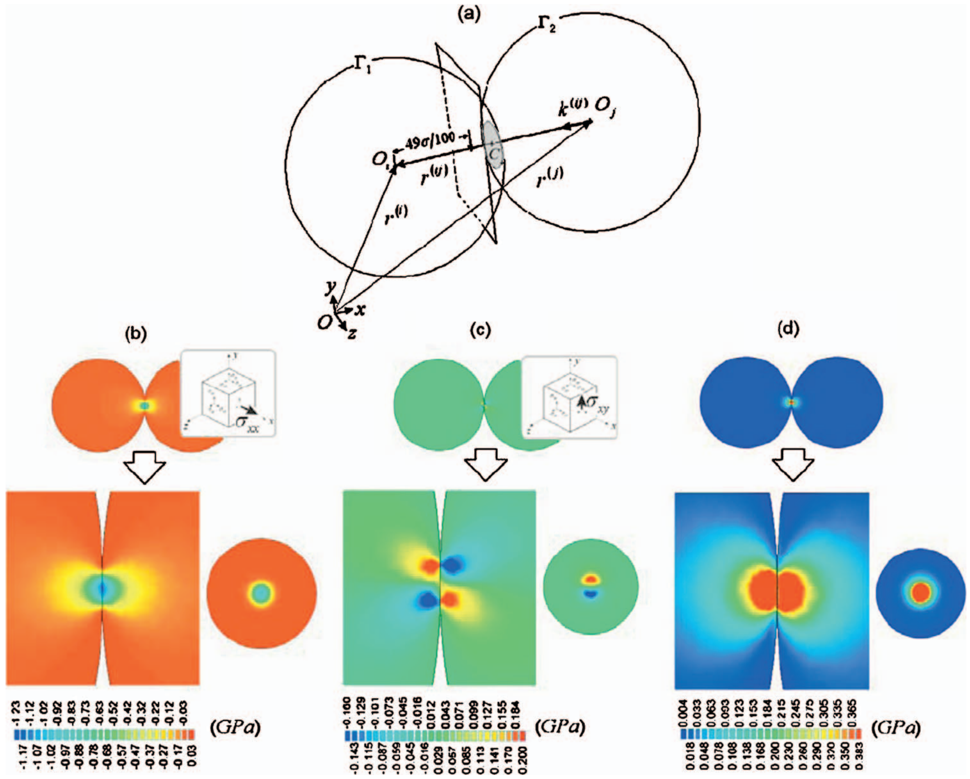


Fig. 3 (*a*) Two elastic balls are deformed in the vicinity of their point of first contact, *C*, under the action of the normal and tangential forces due to collision. The contact area is generally finite though small compared with the dimensions of the balls. Notice that the size of gray color contact area is exaggerated. (*b*) Contour plot of the computed instantaneous normal stress, σ_{xx}, on a cutting *xy* plane passing through the centers of the balls. (Inset) The contact area is magnified to provide a better visualization. (Right) Contour plot of σ_{xx} on a cutting *yz* plane perpendicular to the line of the centers. The position of the cutting *yz* plane is shown in (*a*). (*c*) Contour plot of the computed instantaneous shear stress, σ_{xy}, on a cutting *xy* plane passing through the centers of the balls. (Inset) The contact area is magnified to provide a better visualization. (Right) Contour plot of σ_{xy} on a cutting *yz* plane perpendicular to $r^{(ij)}$: (*d*) Contour plot of the computed instantaneous effective stress, σ_{eff}, on a cutting *xy* plane passing through the centers of the balls with the contact area is magnified in the inset to provide a better visualization. (Right) Contour plot of σ_{eff} on a cutting *yz* plane perpendicular to $r^{(ij)}$.

$b=0$. The initial velocity components of the left ball (before collision) were $V_x=0.7$ m/s, $V_y=-0.1$ m/s, and $V_z=0$. The right ball was initially stationary. The instantaneous results of stresses shown in Fig. 3 are taken after $t\approx 3\times 10^{-5}$ s, when the relative tangential velocity of the initial points of contact on the first and the second spheres reached zero.

The finite element model presented in this section, which is a generalized version of that detailed in Ref. [8], is intended to be used to provide insights into simpler models discussed in the following section.

2.2 Test Cases. The first two test cases presented in this section are those of collisions between two monosized, visco-elastic, rough, glass balls. In a binary collision of particle i and particle j the impulsive force vector excreted by particle i on particle j may be given as

$$J_q=\frac{m(1+e^{(n)})}{2}(g_p^{(ij)}k_p)k_q+\frac{mK(1+e^{(t)})}{2(1+K)}[g_p^{(ij)}-(g_p^{(ij)}k_p)k_q] \quad (1)$$

The first term on the right hand side of Eq. (1) represents the normal impulse, which is in the direction of line of centers. The second term is the tangential impulse, which is in direction perpendicular to the line of centers and lies in the plane of $\mathbf{r}^{(ij)}$, and $\mathbf{J}$. Here, the impact velocity is given by

$$g_p^{(ij)}=(V_p^{(i)}-V_p^{(j)})+\frac{\sigma}{2}\epsilon_{pqm}k_q^{(jp)}(\omega_m^{(j)}+\omega_m^{(p)}) \quad (2)$$

Note that the kinetic energy of the particles is not conserved in a collision between visco-elastic rough balls. In order to describe the degree of plasticity of the collision a coefficient of restitution, $e^{(n)}$, for the collision is introduced. The coefficient of restitution in normal direction is usually defined as the ratio of final to initial relative velocity components of the striking objects in the direction normal to the contact surfaces. Thus the proportionality relation may be given as

Table 1 Material properties used in simulations

Properties	Values
ρ_s	2390 kg/m^3
E	6.3×10^{10} Pa
G_0	2.53×10^{10} Pa
G_∞	0.63×10^{10} Pa
K	4.1×10^{10} Pa
β	9.87×10^{-6} 1/s
υ	0.244

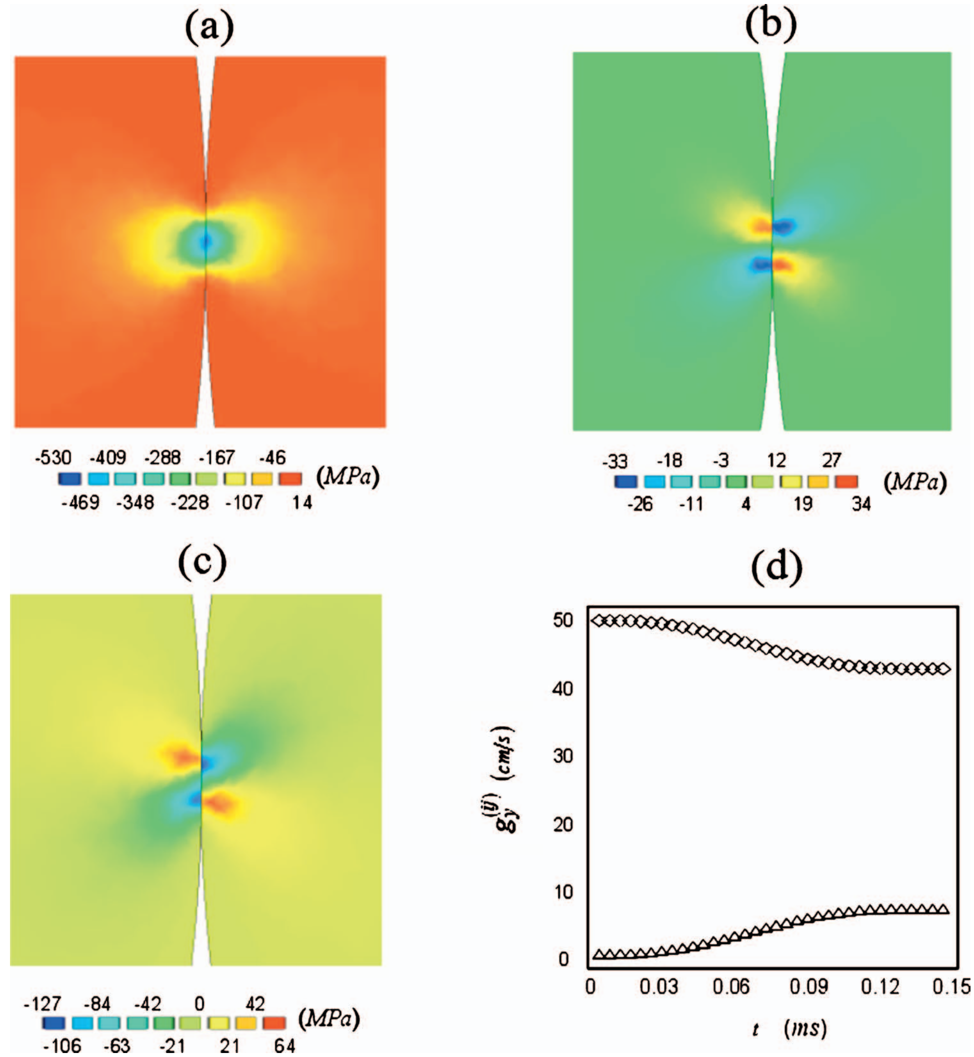


Fig. 4 (*a*) **Computed instantaneous contours of the normal stress, σ_{xx}, on a cutting *xy* plane at t=0.06 ms. (*b*) Computed instantaneous contours shear stress, σ_{xy}, on the same cutting plane as in (*a*), at t=0.06 ms and μ=0. (*c*) Computed instantaneous contours shear stress, σ_{xy} as in (*b*) for μ=0.4. (*d*) The tangential component of impact velocities of the colliding balls versus time. The diamonds represent the tangential impact velocity of the initially moving ball and the gradients are those of the initially stationary ball.**

$$(g_p^{(ij)'}k_q)k_p = -e^{(n)}(g_p^{(ij)}k_q)k_p \tag{3}$$

As mentioned earlier, in this attempt it was assumed that Coulomb friction law [19] describes friction between two colliding balls with a friction coefficient of μ. If the normal impact velocity, $V_{(n)}=g_p^{(ij)}k_p$, is small, then an expression for the tangential coefficient of restitution, which is the ratio of the relative velocity of the pair colliding particles in the tangential direction after collision to that before collision, may be found as

$$e^{(t)} = -1 + \mu(1+e^{(n)})\left(1+\frac{1}{K}\right)\frac{|(g_p^{(ij)}k_p)k_q|}{|(g_q^{(ij)}-(g_p^{(ij)}k_p)k_q)|} \tag{4}$$

The negative values of $e^{(t)}$ in Eq. (4) indicate a reduction in the tangential component of the postcollisional relative velocity (without change in its direction) which represents a slipping motion of a first ball over the second one.

2.3 Slipping Motion. Equation (4) predicts the coefficient of tangential restitution for a collision throughout its lifetime only as much friction will act as is necessary to prevent sliding. To demonstrate the benefit of the detailed finite element model discussed in the preceding section, in the following the numerical results are presented for a binary collision between two monosized rough, visco-elastic glass balls whose material properties are presented in Table 1.

In this case, the moving ball with precollisional velocity components of V_x=0.1 m/s, V_y=0.5 m/s, and V_z=0 collides with a stationary ball. The arrangement of the balls is similar to that as shown in Fig. 2(*a*), and the diameter of identical balls is σ =2 cm.

Figure 4(*a*) represents computed contours of the normal stress, σ_{xx}, on a cutting *xy* plane passing through the centers of the balls at the end of the approaching period, where the contact pressure at the initial point of contact reached a maximum value. The normal stress presses the balls together to give rise to a contact surface. In the absence of friction forces, the computed contours of shear stresses, σ_{xy}, are illustrated in Fig. 4(*b*).

In the presence of friction forces, a tendency to slide exists, and consequently a tangential traction of friction develops in a direction that opposes the relative motion. The traction of friction introduces extra shear stresses into contact surface. Contours of shear stresses are illustrated in Fig. 4(*c*), which may be compared with contours of shear stresses in the absence of friction forces as illustrated in Fig. 4(*b*). The amount of slip is dependent on the motion of the balls, and is also dependent on the deformation and the friction in the contact.

Figure 4(*d*) represents the tangential component of impact velocity of colliding balls as a function of time. It can be seen that there is no change in the direction of the tangential component of the post-collisional relative velocity throughout collision lifetime. This observation also implies that a slipping motion occurred on the first ball over the second ball. Hence, Eq. (4) may be used to estimate the value of coefficient of tangential restitution for the collision whose detailed results are presented in Fig. 4. By substituting the computed value of the coefficient of normal restitution from the finite-element model into Eq. (4), estimations may be made for the coefficient of tangential restitution. The estimated value for the coefficient of tangential restitution is found to be −0.7347.

By using the magnitudes of final velocities of both balls at the end of contact as given in Fig. 4(*c*), the coefficient of tangential restitution predicted by the finite-element model is found to be −0.7309. An excellent agreement between the model prediction for the coefficient of tangential restitution and that calculated using Eq. (4) may establish the validity of the numerical model.

2.4 Sticking Motion. The coefficient of friction for impact phenomena cannot be accurately determined. Consequently, its specification rests upon corresponding values for noncollision processes. The second case as described in the following demonstrates the important role that the coefficient of friction plays in the dynamics of a collision.

By increasing the coefficient of friction for the first case to 0.8, the normal impulse times the coefficient of friction exceeds the magnitude of the tangential impulse. In this case, a sticking motion may occur at impact. As illustrated in Fig. 5(*a*), in sticking contacts not only the reduction of magnitude of the relative velocity in the tangential direction but also the reversal of its direction may occur. Here, no estimations can be made for the coefficient of tangential restitution using Eq. (4), because for a sticking motion $e^{(t)} < -1 + \mu(1+e^{(n)})(1+1/K)|(g_p^{(ij)}k_p)k_m|/|(g_m^{(ij)} - (g_p^{(ij)}k_p)k_m)|$.

Figure 5(*b*) represents the computed contours of the normal stress, σ_{xx}, at the end of the approaching period for a sticking motion of the first ball over the second one. In this case, a tangential impulse less than the limiting friction impulse was applied to the balls in contact so that no initiation of slip could be observed. Consequently, contours of the shear stress, σ_{xy}, as illustrated in Fig. 5(*c*) appear to be different than those representing slipping motion as shown in Fig. 4(*c*). Generally speaking, in a sticking motion, small tangential sliding may occur to avoid infinite traction at the outer edge of the contact area while the contact as a whole does not slide. The stick radius of the contact area is dependent on the value of coefficient of friction among other factors. The present results suggest that the size of the slip region decreases with the magnitude of friction impulse.

To demonstrate more examples of sticking motion, in the following section the behavior of a superball will be described which tends to stick to the surface during a bounce.

2.5 Spin Reversal. When a child's toy called a superball, which is a small ball having the combined properties of a high coefficient of normal restitution and a high surface friction coefficient, collides on a flat surface, the rebound spin, and speed may differ from those values before the collision.

In this section, this rather surprising phenomenon of spin reversal is simulated using the finite-element model discussed in the preceding section. Figure 6(*a*) represents a superball thrown at low speed having a clockwise spin onto the surface, and Fig. 6(*b*) illustrates the grid used in the numerical treatments. A suitable level of mesh refinement is established on the basis of the Hertz elastic contact problem. Here, no consideration is given to the acceleration due to gravity.

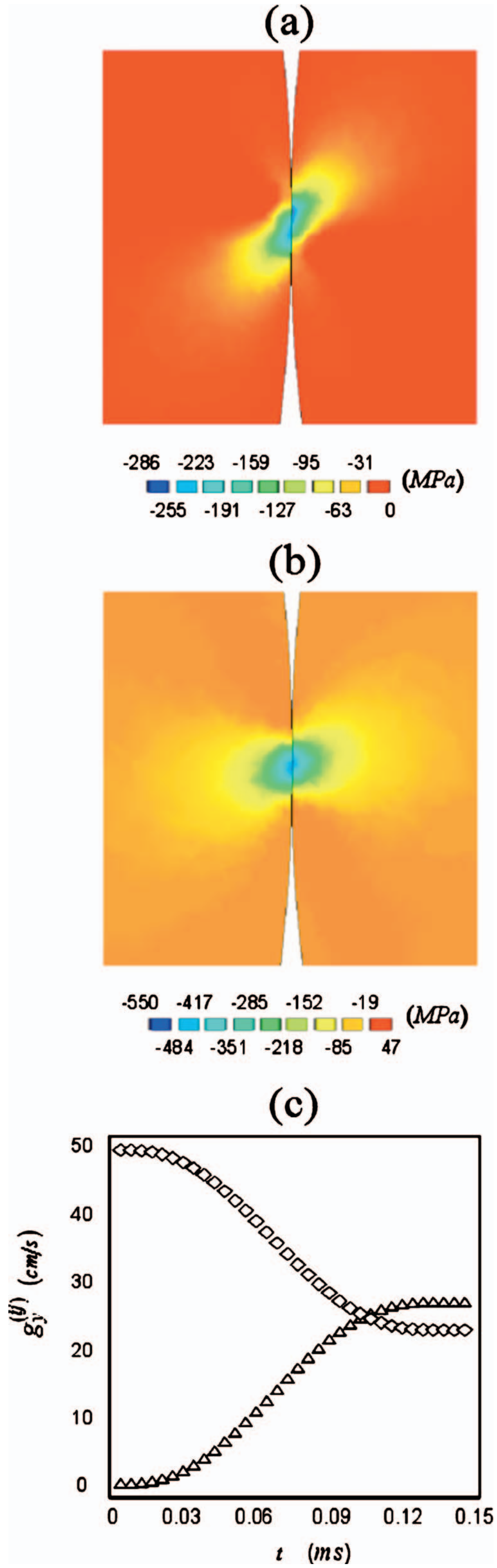


Fig. 5 (*a*) Computed instantaneous contours of the normal stress σ_{xx} on a cutting *xy* plane at $t=0.06$ ms. (*b*) Computed instantaneous contours shear stress σ_{xy} on the same cutting plane as in (*a*) at $t=0.06$ ms. Notice that in this case, a normal contact is sheared by a tangential force, which is insufficient to cause failure. (*c*) The tangential component of impact velocities of the colliding balls versus time. The diamonds represent the tangential impact velocity of the initially moving ball and the gradients are those of the initially stationary ball.

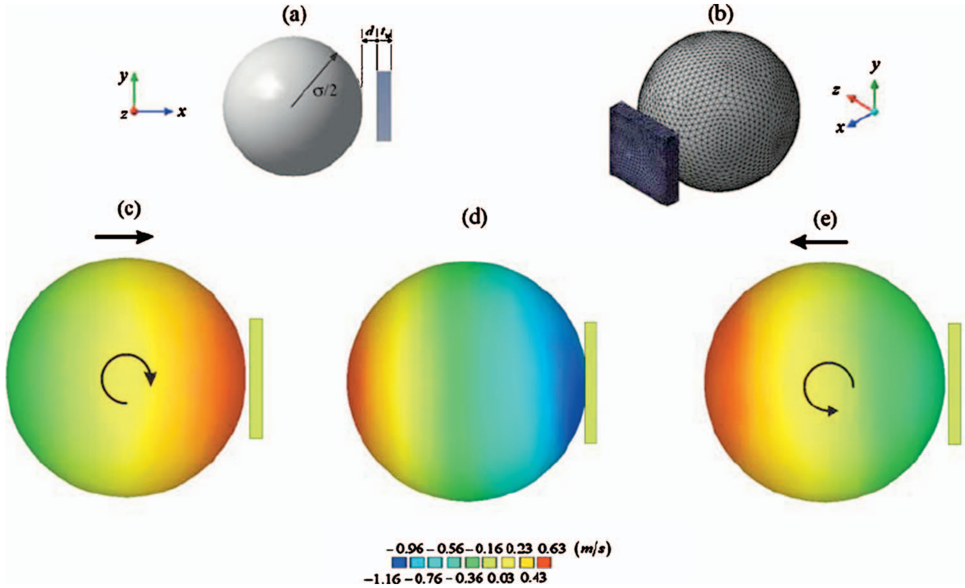


Fig. 6 (***a***) **The arrangement of the super-ball and the wall with some nomenclatures.** (***b***) **3D finite element mesh used in numerical treatments. Here, more than 6×10^4 hexahedral elements were used.** (***c***) **The thrown superball at low speed having a clockwise spin. The initial velocity components of the superball (before collision) were V_x=0.92 m/s, V_y=−0.24 m/s, and V_z=0. Here, the surface of the ball is color coded using the local magnitude of V_y.** (***d***) **The incident of the superball on the front side of the flat wall. The configuration was taken at the end of approaching period when the normal pressure at the initial point of contact reached to the maximum value.** (***e***) **The backwards spinning superball. Notice the spin reversal which is clearly illustrated in** (***c***) **and** (***e***)**. In this simulation the back side of the wall was fixed.**

Figure 6(c) represents the superball thrown with a clockwise topspin at an angular speed around 8.9 rev/s. Figure 6(d) depicts the incident of the superball on the flat wall. Finally, Fig. 6(e) shows the superball spinning backward at around 8.9 rev/s. It can be clearly seen from Fig. 6 that there is the phenomenon of spin reversal. Here, the diameter of the superball is $\sigma=2$ cm. Choosing the appropriate set of material properties for the ball and the flat wall, a coefficient of normal restitution was obtained close to unity, while the coefficient of friction was set to $\mu=0.8$.

2.6 A Simplified Model. Useful results can be obtained using the detailed micromechanics of the collision discussed in the preceding section. However, performing three-dimensional simulations with a large number of balls does not seem to be feasible due to computational intensity which could be involved. The main goal in this section is to utilize the results of the finite-element based model to develop a simplified model by which three-dimensional simulations of a large number of particles can be conducted in more reasonable times.

The simplified model as proposed in the previous work [8] assumes that the grains behave as viscoelastic bodies that interpenetrate during a collision, resulting in the generation of restoring forces that can be characterized by Young's modulus. The short-range interaction of the particles, therefore, may result in a large gradient of the interaction force. The interaction force between the grains in contact should be calculated about 1000 times during a collision to provide accurate results.

As shown in the preceding section, a ball translates and spins, depending on the forces and torques acting on it. In the presence of a gravitational field, the equation of motion of a typical spherical ball j having repulsive interactions with its neighbors, together with normal and tangential forces, may be given as

$$\frac{dV_i^{(j)}}{dt} = -g\,\delta_{iz} + \sum_{p=1}^{N_j} F_i^{(jp)} \tag{5}$$

Here, $F_i^{(jp)}=F_i^{(jp)_n}+F_i^{(jp)_t}$, where $F_i^{(jp)_n}$, and $F_i^{(jp)_t}$ represent the normal and tangential force per unit mass acting on particles j, respectively, which may be conjectured by [8]

$$F_i^{(jp)_n} = \left\{[E/3(1-\nu^2)]\sigma^{1/2}\delta^{(jp)^{3/2}} + K^{(n)}\tau[G_0^2/(G_0 - G_\infty)]\delta^{(jp)}\frac{d\delta^{(jp)}}{dt}\right\}k_i^{(jp)}$$

$$F_i^{(jp)_t} = -\left(k^{(t)}\chi^{(jp)^{3/2}} + K^{(t)}\frac{d\chi^{(jp)}}{dt}\right)e_i^{(t)} \tag{6}$$

where the tangential unit vector is given as

$$e_i^{(t)} = \varepsilon_{iqp}k_q^{(jp)}\varepsilon_{pjk}g_j^{(jp)}k_k^{(jp)}/|g_j^{(jp)}|(1-g_m^{(jp)}k_m^{(jp)}/|g_j^{(jp)}|)$$

The magnitude of $\chi^{(jp)}$, which represents the tangential displacement, is given as , where the tangential displacement $\chi^{(jp)}$, is set initially to zero when a new contact is established and once the contact is broken, all memory of the prior displacement is lost.

Equation (6) represents in a crude manner the complex behavior at real contact as discussed in the preceding section. Recall that when a normal contact is sheared by a tangential force, which is insufficient to cause failure, a region of microslip forms adjacent to the outer perimeter of the contact zone. By increasing the tangential force this region moves inward and two surfaces slip with respect to one another, while in the interior of the contact surfaces remain stuck together. By further increasing the tangential force at which the failure occurs, there is no "stick" region in the interior of the contact zone as illustrated in Fig. 4.

In Eq. (6), Coulomb friction law can be employed to describe friction between two colliding grains with a surface friction coefficient μ. When there is mutual slipping at the point of contact, $|F_i^{(jp)_t}|$ is calculated as necessary to satisfy $|F_i^{(jp)_t}|=\mu|F_i^{(jp)_n}|$. Otherwise, the contact surfaces are considered as stuck while $|F_i^{(jp)_t}| < \mu|F_i^{(jp)_n}|$.

Tangential forces induce torques on particle j, which is defined as $T_i^{(j)}=-\Sigma_{j=1}^{N_j} m\sigma/2\varepsilon_{ijk}k_j^{(jp)}F_k^{(jp)_t}$. Hence, Eq. (6) must be augmented by a torque equation for the rotational motion of particle j which can be written as

$$I^{(j)}\frac{d\omega_i^{(j)}}{dt}=T_i^{(j)} \tag{7}$$

A standard method for solving the ordinary differential Eqs. (6) and (7) is the finite difference approach, through which by giving the particle positions, velocities, and other dynamic information such as impact forces at time t, the positions and velocities at a later time $t+\delta t$ are calculated to a sufficient degree of accuracy. The equations can be solved on a step-by-step basis using the Verlet algorithm [20]. The choice of the time interval δt will depend on the values of the model parameters in Eq. (6).

The contact parameter $k^{(t)}$ in Eq. (6) may be approximated by $k^{(t)} \approx [4E/9(1-v^2)]\sigma^{1/2}$ [21]. However, the contribution of surface roughness in rotational velocity damping must be estimated by comparing the results of the finite-element model and those obtained by solving Eqs. (6) and (7). Applying the material properties used for collision cases as detailed in the preceding section, Eqs. (6) and (7) were integrated using fourth- and fifth-order embedded formulas from Dormand and Prince [22] with $\delta t=3\times 10^{-9}$ s. The initial conditions for Eqs. (6) and (7) are $\delta^{(jp)}(0)=0$, $\chi^{(jp)}(0)=0$, $\dot{\delta}^{(jp)}(0)=0.1$ m/s, and $\dot{\chi}^{(jp)}(0)=0.5$ m/s. The chosen value for $K^{(n)}$ was that suggested in Ref. [8], namely 143.

The variations of ω_z, V_x, V_y, σ_{xx}, and σ_{xy} with time are plotted in Figs. 7 and 8 for the slipping and sticking motions, respectively. The diamonds and the gradients represent the obtained results using the finite-element model for the initially moving and the initially stationary balls, respectively.

The coefficient $K^{(n)}$ in Eq. (6) was considered to be a fit parameter, due to the lack of information concerning the contribution of surface roughness in rotational velocity damping. Choosing $K^{(t)} \approx 4.02\times 10^{14}$ 1/s, the variations of ω_z, V_x, V_y, σ_{xx}, and σ_{xy} with time are also plotted in Figs. 7 and 8 using solid lines, which may be compared with the finite-element model results. As can be seen from Fig. 7(d), the agreement between the predicted results obtained for the normal using two different models is quite satisfactory. The maximum shear stress predicted by the simplified model is almost twice as big as that predicted by the finite-element based model. However, as can be seen from Fig. 4(c), which represents the results of the finite-element based model, the position of the maximum shear stress is not exactly at the initial point of contact. The agreement between the predictions of the two models appears to be improved for the sticking motion as illustrated in Fig. 8(d). Generally speaking, no excellent agreement can be achieved between the predictions of the two models due to the complex behavior of shear stress predicted by the finite-element model as illustrated in Figs. 4(c) and 5(c).

Notice that attempts have been made to use a linear spring dashpot contact law such as that discussed in Ref. [23]. However, no satisfactory agreement could be reached between the results of the finite-element based model and those obtained using the linear model.

Hence, in this section by utilizing the results of the finite-element based model, a simplified model was developed for accurate and efficient granular-flow simulations. An advantage of the simple model is an accurate description of interparticle forces during multiparticle contacting. Interparticle forces determine important physical and mechanical properties of the granular material, which are essential for describing its transport and governing its flow ability. Figure 9(a) shows the variations with time of x, y, and z components, and the magnitude of contact force acting on the darker particle, as depicted in Fig. 9(b). Figure 9(a) reveals the presence of very complex dynamics in a string of rough, viscoelastic particles which could result in the formation of the anisotropic, nonstraight, force chains, even in a partially ordered phase for which jamming could occur. It is worth noting that the lifetime of the chain of particles as illustrated in Fig. 9(a), is quite long time compared to a typical binary collision time of $t_{c,\text{binary}} \sim 10^{-6}$ s. This kind of contact may be termed a long rubbing contact. In this case, particle diameter was $\sigma=1.9$ mm, and the physical properties were those listed in Table 1. In addition, the coefficient of friction was set to $\mu=0.8$.

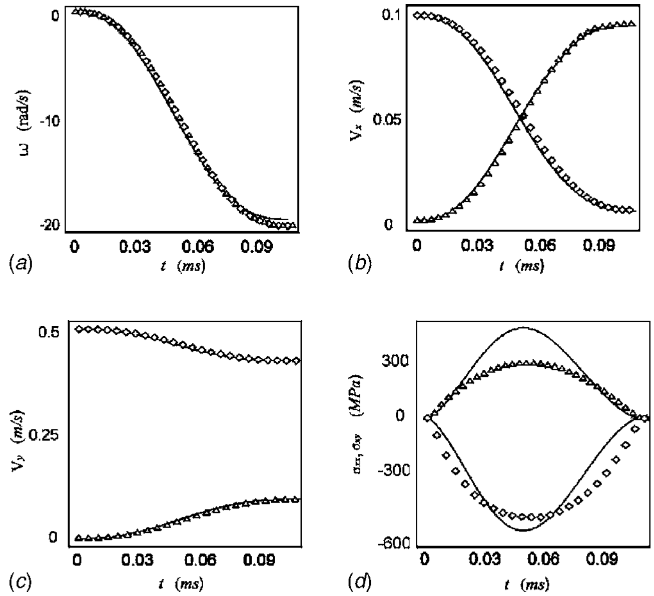


Fig. 7 (*a*) The computed angular velocities, ω_z, versus time for the colliding balls. The diamonds and the gradients represent the results of finite element based model for the initially moving and the initially stationary balls, respectively. The solid lines are the obtained results using the simplified model. As expected, the angular velocity of the identical balls at the end of restitution period is equal to each other. (*b*) The computed normal velocities, V_x, versus time for the colliding balls. The diamonds, the gradients and the solid lines have the same meaning as those in (*a*). (*c*) The computed tangential velocities, V_y, versus time for the colliding balls. The diamonds, the gradients and the solid lines have the same meaning as those in (*a*). (*d*) Variations of the normal stress, σ_{xx}, and the shear stress, σ_{xy}, at the point of initial contact with time for the initially moving ball. The diamonds and the gradients represent the normal and the shear stresses, respectively, obtained using the finite element based model. The lower solid line represents the predictions of the simplified model for the normal stress of the point of contact on the surface of the initially moving ball. The agreement between the results of two models is quite satisfactory. The upper line represents the predictions of the simplified model for the shear stress at the point of contact. The maximum shear stress predicted by the simplified model is almost twice as big as that predicted by the finite element based model. However, as can be seen from Fig. 4(*b*), the position of the maximum shear stress calculated using the finite element based model is not at the initial point of contact.

3 Validation of the Simplifed Model

Cohesionless grains in a spinning bucket exhibit interesting flow dynamics, such as the existence of solid-like and fluid-like qualities side by side, which may result in the formation of circular kinks on the surface of the granular material [24]. Although the

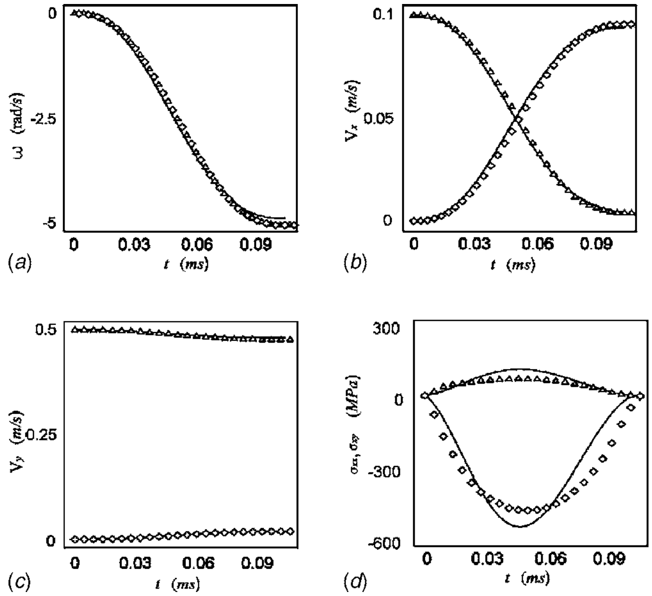


Fig. 8 (*a*) The computed angular velocities, ω_z, versus time for the colliding balls. The diamonds and the gradients represent the results of finite element based model for the initially moving and the initially stationary balls, respectively. The solid lines are the obtained results using the simplified model. As expected, the angular velocity of the identical balls at the end of restitution period is equal to each other. (*b*) The computed normal velocities, V_x, versus time for the colliding balls. The diamonds, the gradients and the solid lines have the same meaning as those in (*a*). (*c*) The computed tangential velocities, V_y, versus time for the colliding balls. The diamonds, the gradients and the solid lines have the same meaning as those in (*a*). (*d*) Variations of the normal stress, σ_{xx}, and the shear stress, σ_{xy}, at the point of initial contact with time for the initially moving ball. The diamonds and the gradients represent the normal and the shear stresses, respectively, obtained using the finite element based model. The lower solid line represents the predictions of the simplified model for the normal stress of the point of contact on the surface of the initially moving ball. The upper line represents the predictions of the simplified model for the shear stress at the point of contact.

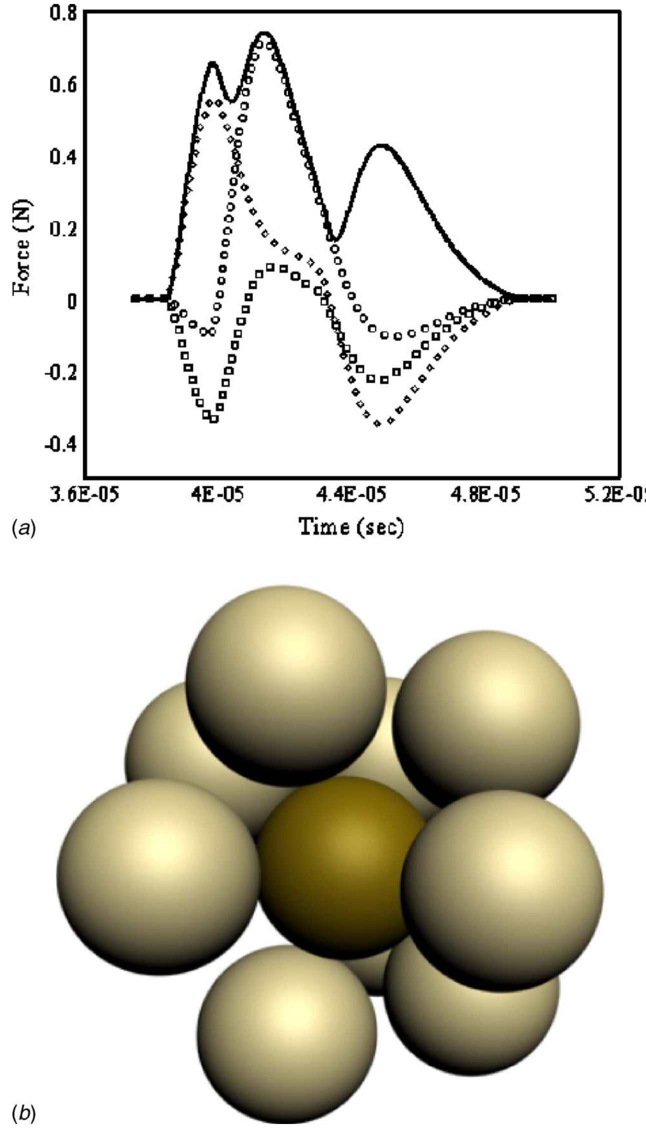


Fig. 9 (*a*) Time series of *x*, *y*, and *z* components of impact force, represented by diamonds, squares, and circles, respectively. As well as the magnitude of impact force on the darker particle, depicted in (*b*), the magnitude of impact force is shown with a solid line. (*b*) Closeup of a cluster of ten particles.

slowly evolving surface shape of a bucket of sand experiencing vertical spinning motion at a fixed rotation rate cannot be completely explained using the Coulomb yield condition [25], the formation of circular kinks, whose radius was reported in Yoon, et al. [24] to be a function of the spinning speed and the tilting angle, remains a mystery. Several researchers [24–26] have proposed theories for the prediction of grain dynamic behavior in a spinning bucket.

In this section, granular flows are studied by means of molecular dynamics simulations in a vertical bucket of sand, rotating around its cylindrical axis at rotational frequencies higher than 50 1/s. The aim is to assess the simplified model developed in the preceding section to predict phenomena such as the observed steep depression [25] around the center of a vertical bucket of sand rotating around its cylindrical axis at high rotational frequencies, namely 100 1/s. To this end, more than 60,000 identical, slightly overlapping, spherical particles with diameter $\sigma = 1.9$ mm were placed randomly in the cylindrical, computational box, as illustrated in Fig. 10. The side and bottom walls of the bucket were spinning at $\omega_0 = 100\mathbf{e}_z$ 1/s. The initial velocities of each particle in the x, y, and z direction are assigned with a magnitude according to a profile $\mathbf{V} = -\omega_0(x\mathbf{e}_x + y\mathbf{e}_y) - \mathbf{e}_z$ m/s, plus a small random number uniformly distributed over the interval $[-0.0025, +0.0025]$ m/s.

The friction is represented by a restoring force characterized by μ, whose value is set to 0.6. This force counteracts mutual sliding motion at contact. The material properties of particles are listed in Table 1. In order to simulate multiple-particle collisions such as that shown in Fig. 9 satisfactorily, the value of time step δt was set to 5×10^{-9} s.

In the experiments [25] a cylindrical container was used, made of polyvinyl chloride (PVC) plastic with density, Young modulus, and Poisson's ratio, and given as $\rho = 1400$ kg/m^3, $E = 3.4 \times 10^9$ Pa, $\upsilon = 0.33$, respectively. The roughness of the wall is much smaller than the particles diameter, which is on the order of millimeters. Hence, a simplified representation of the relevant aspects of the particle wall contacts is chosen, hoping that detailed comparisons between simulations and experiments can then be used to establish the quality of the approximation. The interaction between the particles and the surface is described by a simple friction law, in which the coefficient of friction is a constant set to $\mu_w = 0.8$.

The particle–wall frictional interaction will likely perturb the flow field. In fact, by decreasing the particle-to-wall friction the

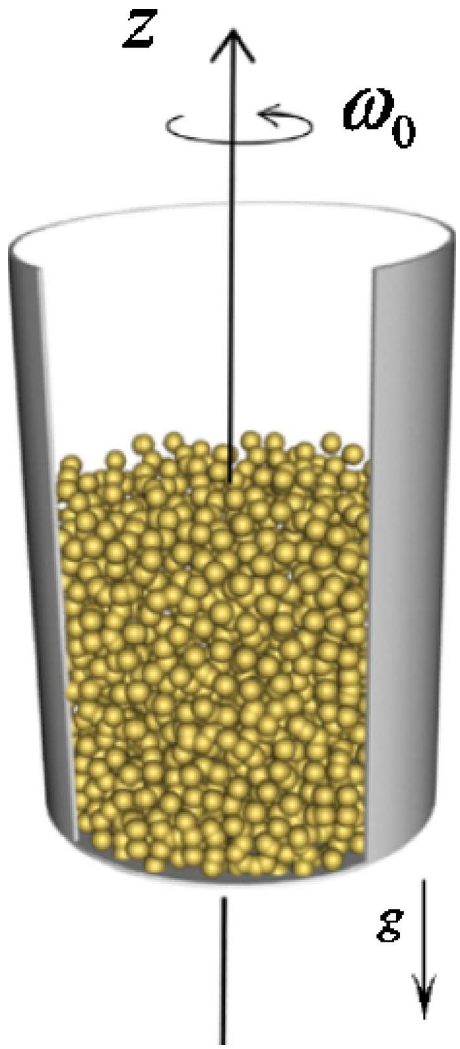


Fig. 10 A cutaway view of the spinning bucket. Note that gravity acts in the negative *z* direction.

particle wall contact increases, leading to denser and more ordered structures in the wall region. On the other hand, by increasing particle-to-wall friction ordering can be delayed [27], which affects the ability of the particle to slide on the walls. Thus an important question is the degree to which the present model is able to reproduce the effects of the wall friction on the bulk flow field. Here, the above-mentioned value of μ_w is chosen for the simulations with the anticipation of inhibiting sliding motion.

The normal elastic constants between particles, a flat bottom wall, $k^{(n)iw}$, and cylindrical side walls, $k^{(n)ic}$, are calculated using expressions presented below, for Hertzian contacts. They are

$$k^{(n)iw} = \frac{8}{15}\frac{1}{m\left(\frac{1-\nu_p^2}{E_p}+\frac{1-\nu_w^2}{E_w}\right)}\sqrt{\frac{\sigma}{2}} \tag{8}$$

and

$$k^{(n)ic} = \frac{8}{15}\frac{(\varphi_1+\varphi_2)^{1/2}}{m\left(\frac{1-\nu_p^2}{\pi E_p}+\frac{1-\nu_w^2}{\pi E_w}\right)\sqrt{\frac{2}{\sigma}}}\left[\int_0^\infty \frac{d\zeta}{\sqrt{[(1+\zeta)(1+\zeta)\zeta]}}\right]^{3/2} \tag{9}$$

The parameter $k^{(t)_w}$ in tangential force model is set to $2/3k^{(n)iw}$, and an arbitrary but reasonable value is chosen for the viscous damping constant, $k^{(n)_w}=100$. Due to lack of information, no contribution of surface roughness in rotational velocity damping is assumed during the particle–wall interaction.

The equations of motion presented in this section are solved using the Verlet algorithm [5], with a time step of $\delta t=5\times10^{-9}$ s. In the present simulations, the values of normal elastic constant are large enough to avoid grain interpenetration, and the time step $\delta t < t_{c,\text{binary}}/100$ is small enough to assure an accurate simulation. In addition, regardless of momentum exchanged between the particle and the wall, the rotational rate of the bucket is assumed to be constant

In the present attempt, two-fifths of a spinning bucket of radius $R=0.0516$ m and height $L=0.15$ m is initially filled with monosized, rough, viscoelastic, spherical, glass particles. The apparent density of the system was 1155 kg/m^3. The free surface, located at $z_0\approx7.5$ cm, was nearly flat, as illustrated in Fig. 11(*a*), before the spinning begins.

Figure 11(*b*) represents the configuration of glass balls in the bucket rotating at rate of rotation of $\omega_0=100$ 1/s, after five complete rotations. Figure 11(*c*) illustrates the time smoothing of the volume-averaged solids fraction. In order to obtain the local description of solids fraction, first the bucket is divided into an appropriate number of sampling volume using criteria given in Ref. [28]. Then, the local values are calculated as averages over the particles whose instantaneous configurations are illustrated in Fig. 11(*d*). Finally, the timesmoothing is carried out by averaging 500 configurations each separated by $t=10^{-5}$ s. As it can be seen, the free surface becomes curvy, but azimuthally symmetric with a depression in the center of the bucket. The inset of Fig 11(*d*) clearly shows the formation of a steep depression whose diameter is about $D_h=0.15$ R.

The stress tensor for a system as illustrated in Fig. 11 may be calculated as the sum over all particles i within a sampling volume V, which is a sector of the cylinder, given by Allen and Tildesley [5]. That is

$$P_{\alpha\beta} = -\frac{1}{V}\left(\sum_i \prod_i\right) \tag{10}$$

where $\Pi_i=m_i(V_{i\alpha}-\bar{V}_\alpha)(V_{i\beta}-\bar{V}_\beta)+\Sigma_i\Sigma_{j>i}r_{ij\alpha}F_{ij\beta}$.

If $\mathbf{n}$ is the unit outer normal to the side wall of the bucket, then the normal stress acting on the sidewall is given by

$$P_n = P_{\alpha\beta}n_\alpha n_\beta \tag{11}$$

Figure 11(*e*) represents the glass balls color coded with the value of normal component of Π_i. Figure 11(*f*) illustrates contour plots of the time smoothing of the volume-averaged value of the normal component of, $P_{\alpha\beta}$, which characterizes the granular pressure. The results of simulations such as those shown in Figs. 11(*c*) and 11(*f*) can be very useful in order to avoid arbitrary treatments or unnecessary assumptions for developing a set of continuum equations for granular flows in an intermediate regime.

Notice that by using $\mu_w=0.8$, ordering was avoided in the wall region as evidenced from Fig. 11(*e*).

By setting the rotational rate of the bucket to $\omega_0=50$ 1/s after six complete rotations, a cusp in the central section was developed as clearly depicted in Fig. 12(*a*). The time-smoothing of the volume-averaged value of the solids fraction is shown in Fig. 12(*b*). The present results are in excellent agreement with the observations reported in Ref. [25]. They provide the benchmark for developing and testing theories for granular materials. The details such as formation of a depression as illustrated in Fig. 11, improve the resolution significantly compared to that reported in Ref. [25].

4 Conclusions and Outlook

In this and the previous parts [8] of this attempt, a methodology based on a synergistic combination of experiments and simulations was applied to improve the understanding of granular flows.

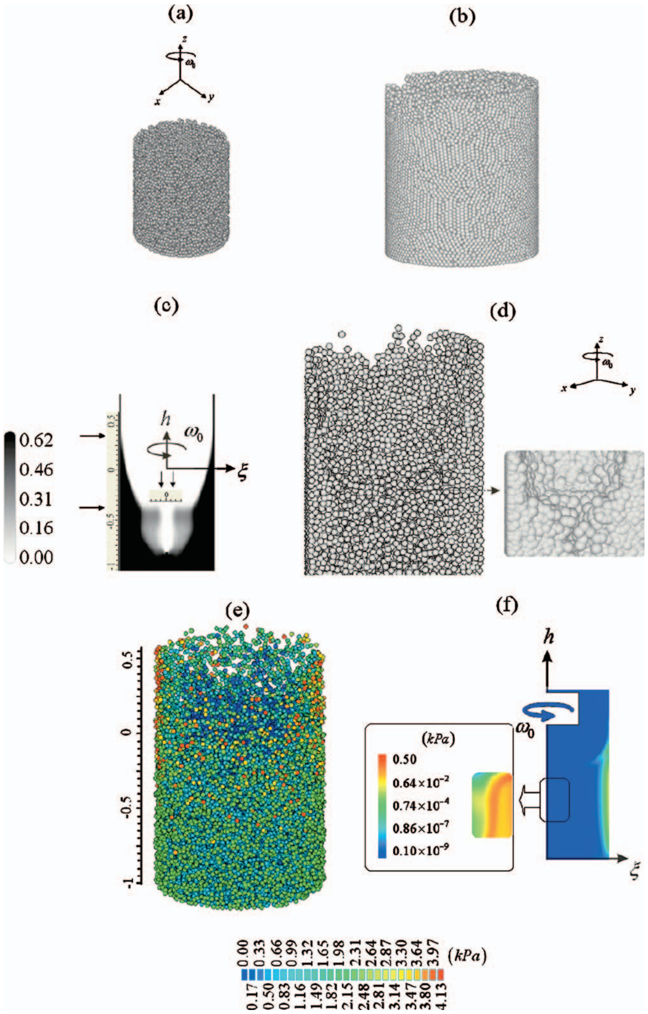


Fig. 11 (*a*) Initial configuration of the spherical balls with a flat free surface before the spinning begins. (*b*) Configuration of the spherical balls after five complete rotations at the rate of ω_0=100 1/s. (*c*) Contour plot of the time smoothing of the volume averaged solids after five complete rotations of the bucket. (*d*) A typical instantaneous configuration of spherical balls in a cutaway view of the spinning bucket after several rotations at the rate of ω_0=100 1/s. To provide a clear picture of depression the central part of the bucket is magnified in the inset. Notice that the diameter of identical particles is much smaller than that shown in the inset. (*e*) An instantaneous configuration of the spherical balls. The balls are color coded using the local value of the normal component of Π. (*f*) Time smoothing of the volume averaged of normal component of $P_{\alpha\beta}$. (Inset) Time smoothing of the volume averaged of normal component of $P_{\alpha\beta}$ in the central part of bucket using a finer scale.

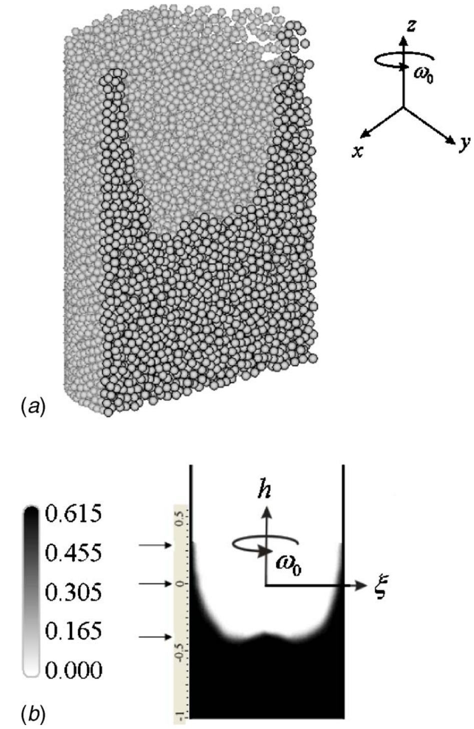


Fig. 12 (*a*) A typical instantaneous configuration of spherical balls in a cutaway view of the spinning bucket after several rotations at the rate of ω_0=50 1/s. Notice the formation of a cusp in the central region. (*b*) Contour plot of the time-smoothing of the volume averaged solids after six complete rotations of the bucket at the rate of ω_0=50 1/s.

A model was developed to describe granular flow dynamic behavior in the intermediate regime where both collisional and frictional interactions between particles may occur. For this case, experiments were carried out to provide a general overview of impact analysis of real glass balls.

A detailed simulation of frictional interactions between the grains was provided by employing a three-dimensional (3D) finite-element model. Numerical results were presented of slipping, and sticking motions during a binary collision. The results of this analysis were used in the development of a practical methodology for a large-scale, 3D, molecular dynamics type simulation of dense granular material consisting of glass particles.

The model was applied to examine the continuous flow of grains in a spinning bucket at rotational frequencies higher than 50 1/s. Formation of a steep depression with diameter of about D_h=0.15R near the axis of rotation was predicted at rotation rate of 100 L/s. In addition, at lower rate of rotation of ω_0=50 1/s a cusp in the central section of bucket was observed.

Based on the ability of the model to reproduce poorly understood phenomena in complex granular systems, the present model appears to represent a significant advance in state-of-the-art prediction of granular flows. However, only a rather limited amount of experimental data are available for model validation. Hence it is recommended for performing additional measurements in simple geometries useful for testing and for comparing theories of granular materials.

One of the future challenges would be to develop a set of continuum equations for further exploration of dynamics of granular flows in an intermediate regime where both collisional and frictional interactions between particles should be taken into account. The results of simulations presented in this part would be of use in order to avoid arbitrary treatments or unnecessary assumptions in future theoretical investigations. Recall that the current understanding of the dynamics of granular flows comes from two rather disjointed models, namely continuum models such as those used in soil mechanics, and kinetic theory models. Therefore, a future attempt is directed at providing greater insight toward the explanation of poorly understood hydrodynamic phenomena in the field of granular flows.

Nomenclature

A_{seg} = area of contact segment
b = decay constant
d = initial distance of super ball from the wall
D_h = size of depression
$(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ = unit vectors in the x, y, and z direction

$e^{(n)}$ = normal coefficient of restitution
$e^{(t)}$ = tangential coefficient of restitution
E = Young modulus
$F_n^{(f)}$ = friction force at time t_n
$F_n^{(N)}$ = normal force at time t_n
$F^{(t)}$ = trial value of interface force
$F_i^{(jp)}$ = force per unit mass acting on particle j
$F_i^{(jp)_n}$ = normal force per unit mass acting on particle j
$F_i^{(jp)_t}$ = tangential force per unit mass acting on particle j
G_0 = instantaneous (glassy) shear modulus
G_∞ = long time shear modulus
g = acceleration due to gravity
$g_p^{(ij)}$ = (indicial rotation) relative velocity vector at the point of initial contact for a pair particles i and j
h = dimensionless vertical distances ($h=z/2R$)
$I^{(j)}$ = moment of inertia
$\mathbf{J}$ = (vector notation) impulsive force
J_q = (indicial notation) impulsive force
k_n = (indicial notation) unit vector directed from the center of particle i to the center of particle j at impact defined as $k_n=r_n^{(ij)}/\sigma$ with $\mathbf{r}^{(ij)}=\mathbf{r}^{(j)}-\mathbf{r}^{(i)}$ as depicted in Fig. 3(a)
K = bulk modulus
K = gyration radius
$K^{(t)}$ = parameter which characterizes the contribution of surface roughness in rotational velocity damping
$K^{(n)_w}$ = damping rotational velocity parameter for walls
k = contact stiffness ($\mathrm{k}=0.1A_{\mathrm{seg}}K/V_{\mathrm{seg}}$)
$k^{(t)}$ = contact parameter
$k^{(t)_w}$ = contact parameter for walls
$k^{(n)_{ic}}$ = normal elastic constant between the particles and the cylindrical side walls
$k^{(n)_{iw}}$ = normal elastic constant between the particles and the base of bucket
L = height of bucket
m = mass of particle
N_j = number of particles in contact with particle j at time t
$\mathbf{n}$ = (vector notation) unit outer normal to the side walls of the bucket
$P_{\alpha\beta}$ = mean stress tensor for the particle phase
R = radius of bucket
r = radial distances
$\mathbf{r}^{(ij)}$ = the line of centers of the ith and the jth spheres (vector notation)
$T_i^{(j)}$ = (indicial notation) torque on sphere j
t = time
t_w = wall thickness
$t_{c,\mathrm{binary}}$ = typical collision time of a binary collision
t_{col} = collision time
$V^{(n)}$ = magnitude of normal impact velocity
$V_i^{(j)}$ = (indicial notation) velocity of the center of mass of sphere j
$V_i^{(jp)}$ = (indicial notation) relative velocity of the center of mass of spheres j and p ($V_i^{(jp)}=V_i^{(j)}-V_i^{(p)}$)
V = sampling volume
V_{seg} = volume of segments on solid element
V_x = x component of the velocity vector
V_y = y component of the velocity vector
V_z = z component of the velocity vector
$\mathbf{V}$ = initial velocity of particles in the bucket
x = distances in the x direction
y = distances in the y direction
z = distances in the z direction
z_0 = initial location of free surface

Greek Symbols

β = reciprocal of relaxation time
χ = tangential displacement
δ_e = movement of slave nodes
δ_{iz} = Kronecker delta
δt = time step
ε_{imn} = alternating tensor
φ_1 = elliptic integral given as $\varphi_1=\int_0^\infty d\zeta/\sqrt{[(1+\zeta)^3(1+\zeta)\zeta]}$
φ_2 = elliptic integral given as $\varphi_2=\int_0^\infty d\zeta/\sqrt{[(1+\zeta)(1+\zeta)^3\zeta]}$
Γ = boundary of a sphere
Π = particle stress tensor (defined in Eq. (10))
μ = coefficient of friction
μ_s = static coefficient of friction
μ_d = dynamic coefficient of friction
μ_w = coefficient of wall friction
$\Delta\mu$ = ($\Delta\mu=\mu_s-\mu_d$)
ν = Poission's ratio
Δt = time step size ($\Delta t=t_{n+1}-t_n$)
ρ_s = particle material density
σ = particle diameter
σ_{xx} = normal stress
σ_{xy} = shear stress
σ_{eff} = effective stress
$\omega_m^{(j)}$ = (indicial notation) spin velocity vector of spheres j
ω_0 = rate of rotation of bucket of sand
ω_z = spin velocity in z direction
ξ = dimensionless radial position ($\xi=r/R$)

Superscripts

prime = post-collisional value
overbar = mean values

References

[1] Jaeger, H. M., Nagel, S. R., and Behringer, R. P., 1996, "The Physics of Granular Materials," Phys. Today, **49**, pp. 32–39.
[2] Zhou, T., and Kadanoff, L. P., 1996, "Inelastic Collapse of Three Particles," Phys. Rev. E, **54**, pp. 623–628.
[3] Savage, S. B., and Lun, C. K. K., 1988, "Particle Size Segregation in Inclined Chute Flow of Dry Cohesionless Granular Solids," J. Fluid Mech., **189**, pp. 311–335.
[4] Bak, P., *How nature Works. A Science of Self-Organized Criticality*, Copernicus, New York.
[5] Gray, J. M. N. T., and Hutter, K., 1997, "Pattern Formation in Granular Avalanches," Continuum Mech. Thermodyn., **9**, pp. 341–345.
[6] Baxter, G. W., and Yeung, C., 1999, "The Rotating Bucket of Sand: Experiment and Theory," Chaos, **9**, pp. 631–638.
[7] Bouchaud, J. P., Cates, M. E., Prakash, J. R., and Edwards, S. F., 1995, "Hysteresis and Metastability in a Continuum Sandpile Model," Phys. Rev. Lett., **74**, pp. 1982–1985.
[8] Zamankhan, P., and Boardbar, H. M., 2006, "Complex Flow Dynamics in Dense Granular Flows—Part I, Experimentation," J. Appl. Mech. **73** (4), pp. 648–657.
[9] Goldsmith, W., 1960, *Impact, the Theory and Physics Behaviour of Colliding Solids*, Dover, New York.
[10] Hallquist, J. O., 1979, "NIKE2D—A Vectorzed, Implicit, Finite Deformation, Finite Elements Code for Analyzing the Static and Dynamic Response Of 2D Solids," Reports No. UCRL-52678, LLNL, Livemore, CA.
[11] Benson, D. J., and Hallquist, J. O., 1990, "A Single Surface Contact Algorithm for Post-Bucking Analysis of Shell Structure," Comput. Methods Appl. Mech. Eng., **78**, pp. 141–163.
[12] Belytschko, T., and Neal, M. O., 1991, "Contact-Impact by the Pinball Algorithm with Penalty And Lagrangian Methods," Int. J. Numer. Methods Eng., **31**, pp. 547–572.
[13] Nilsson, L., Zhong, Z.-H., and Oldenburg, M., 1989, "Analysis of Shell Struc-

tures Subjected to Contact Impacts," *Proceedings of the Symposium on Analytical and Computational Models of Shells*, Vol. ASME CED-3, ASME, New York, p. 457.
[14] Zhong, Z.-H., and Nilsson, L., 1996, "A Unified Contact Algorithm Based on the Territory Concept," Comput. Methods Appl. Mech. Eng., **130**, pp. 1–16.
[15] Oden, J. T., and Pires, E. B., 1983, "Nonlocal and Nonlinear Friction Laws and Variational Principles for Contact Problems in Elasticity," ASME J. Appl. Mech., **50**, pp. 67–76.
[16] Bathe, K. J., and Chaudhary, A., 1985, "A Solution Method for Planar and Axismetric Contact Problem," Int. J. Numer. Methods Eng., **21**, pp. 65–88.
[17] Guerra, F. M., and Browning, R. V., 1983, "Comparison of Two Slideline Methods Using Adina," Comput. Struct., **17**, pp. 819–834.
[18] Sellgran, U., Björklund, S., and Andersson, S., 2003, "A Finite Element-Based Model of Normal Contact Between Rough Surfaces," Wear, **254**, pp. 1180–1188.
[19] Coulomb, C. A., 1785, "Theorie des Machines Simples," Memoires De Mathematique Et De Physique De L' Academie Des Sciences, Paris, **10**, pp. 161–331.
[20] Allen, M. P., and Tildesley, D. J., 1997, *Computer Simulation of Liquids*, Clarendon, Oxford, UK.
[21] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, UK.
[22] Dormand, J. R., and Price, P. J., 1980, "A Family of Embedded Runge-Kutta Formulae," J. Comput. Appl. Math., **6**, pp. 19–26.
[23] Silbert, L. E., Etras, D., Grest, G. S., Hasley, T. C., Levine, D., and Plimpton, S. J., 2001, "Granular Flow Down an Inclined Plane: Bagnold Scaling and Rheology," Phys. Rev. E, **64**, p. 051303.
[24] Yoon, S., Eom, B., Lee, J., and Yu, I., 1999, "Circular Kinks on the Surface of Granular Material Rotated in a Tilted Spinning Bucket," Phys. Rev. Lett., **82**, pp. 4639–4642.
[25] Vavrek, M. E., and Baxter, G. W., 1994, "Surface Shape of a Spinning Bucket of Sand," Phys. Rev. E, **50**, pp. R3353–R3356.
[26] Yeung, C., 1998, "Metastability of a Granular Surface In A Spinning Bucket," Phys. Rev. E, **57**, pp. 4528–4534.
[27] Zamankhan, P., Tynjala, T, Polashenski, W., Jr., Zamankhan, P., and Sarkomaa, P., 1999, "Stress Fluctuations in Continuously Sheared Dense Granular Materials," Phys. Rev. E, **60**, pp. 7149–7156.
[28] Polashenski, W., Jr., Zamankhan, P., Makiharju, S., and Zamankhan, P., 2002, "Fine Structures in Sheared Granular Flows," Phys. Rev. E, **66**, p. 021303.

Kinking of Transversal Interface Cracks Between Fiber and Matrix

Federico París
e-mail: paris@esi.us.es

Elena Correa
e-mail: correa@esi.us.es

Vladislav Mantič
e-mail: mantic@esi.us.es

School of Engineering,
Group of Elasticity and Strength of Materials,
University of Seville,
Camino de los Descubrimientos s/n,
41092 Sevilla, Spain

Under loads normal to the direction of the fibers, composites suffer failures that are known as matrix or interfiber failures, typically involving interface cracks between matrix and fibers, the coalescence of which originates macrocracks in the composite. The purpose of this paper is to develop a micromechanical model, using the boundary element method, to generate information aiming to explain and support the mechanism of appearance and propagation of the damage. To this end, a single fiber surrounded by the matrix and with a partial debonding is studied. It has been found that under uniaxial loading transversal to the fibers direction the most significant phenomena appear for semidebonding angles in the interval between 60 deg *and* 70 deg. *After this interval the growth of the crack along the interface is stable (energy release rate (ERR) decreasing) in pure Mode II, whereas it is plausibly unstable in mixed mode (dominated by Mode I for semidebondings smaller than* 30 deg*) until it reaches the interval. At this interval the direction of maximum circumferential stress at the neighborhood of the crack tip is approximately normal to the applied load. If a crack corresponding to a debonding in this interval leaves the interface and penetrates into the matrix then: (a) the growth through the matrix is unstable in pure Mode I; (b) the value of the ERR reaches a maximum (in comparison with other debonding angles); and (c) the ERR is greater than that released if the crack continued growing along the interface. All this suggests that it is in this interval of semidebondings* (60–70 deg) *that conditions are most appropriate for an interface crack to kink. Experiments developed by the authors show an excellent agreement between the predictions generated in this paper and the evolution of the damage in an actual composite.* [DOI: 10.1115/1.2711220]

Keywords: composites, matrix failure, fiber-matrix debonding, interface crack, kinked crack, boundary element method

1 Introduction

Failure criteria of fibrous composites are still an open matter. A world wide failure exercise has been recently carried out [1,2] to validate the predictions of several approaches over different laminates. A point of significant difference in the predictions of failure is that associated with the failure in the matrix also referenced as interfiber failure, Hashin and Rotem [3] being the first to give a differentiating character to this type of failure. Although the major field of application of composites is that in which there is a predominant direction of loading, it is obvious that in general one lamina of a laminate will also have to suffer external actions originating stresses perpendicular to the fibers, these stresses then being capable of producing a matrix dominated failure.

The mechanism of damage of this failure can be quite well described, particularly under tension loads, by the appearance of interface cracks growing between fiber and matrix which under certain conditions leave the interface and penetrate into the matrix, generating the coalescence of these interface cracks and giving rise to a macrocrack that is associated to the failure of the lamina under consideration. This is perfectly illustrated by Fig. 1 [4], where the interface cracks can be clearly observed.

The problem of an elastic circular (in two dimension (2D)) or cylindrical (in 3D) inclusion embedded in an elastic matrix with a partial debond at their interface (modeled like an interface crack) subjected to a uniaxial tension at infinity, perpendicular to the debond, has received considerable attention in the past. For the sake of brevity, only a few of these works, which are closely related to the objectives of the present work, will be cited.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received July 25, 2005; final manuscript received September 27, 2006. Review conducted by K. Ravi-Chandar.

In analytical works based on an application of the Kolosov–Muskhelisvili complex potentials to solve the planar problem of partially debonded circular inclusion, faces of the interface crack were assumed to be traction free according to the open model of interface cracks introduced by Williams [5]. Developing the pioneering works of England [6] and Perlman and Sih [7], Toya [8] established, in his remarkable contribution, a theoretical basis for any other analysis of this problem. Toya deduced an analytical expression of the total energy release rate (ERR) as a function of the debond angle and applied it in a fracture criterion to assess the debond growth along the interface. Toya also proposed a strength based criterion for competition between a debond extent and a kink out of the interface, the kink angle having been determined by the maximum circumferential stress (MCS) criterion. It is worth mentioning that Toya's strength criterion for an interface crack extent can be considered equivalent to the following simple ERR based criterion currently used [9], $G^{\text{int}}=G_c^{\text{int}}=G_{1c}^{\text{int}}(1+\tan^2\psi_K)$, where the critical value of ERR depends on the fracture mode mixity given by the phase angle ψ_K (see Sec. 2 for notation used).

Zhang et al. [10] presented experimental results for transverse single-fiber specimens, and Varna et al. [11] studied the measured debond growth along the interface modifying Toya's [8] ERR based fracture criterion in order to take into account increasing participation of the shear fracture mode when the debond grows. Varna et al. [11] assumed a linear variation of the critical ERR value with the debond angle up to a maximum angle with a negligible contact zone at the crack tip and a constant critical ERR for larger debonding angles.

Recently, Prasad and Simha [12] applied the complex potentials theory, assuming the open model of interface cracks, to compute the two components of the complex stress intensity factor (SIF) as functions of the debond angle, and also to apply the MCS criterion

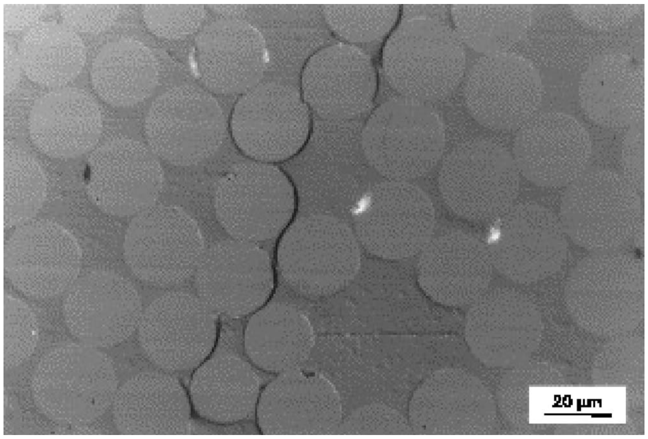


Fig. 1 Debonding cracks in a fibrous composite material under transversal loading

for interface crack predictions. Nevertheless, they did not specify the corresponding small reference distances to the crack tip (in SIF evaluation and the MCS criterion application) which are crucial for the interpretation of the open model results due to the oscillatory character of the near-tip elastic solution.

Chao and Laws [13] extended the Comninou [14] formulation of the contact model of interface cracks, which assumes a near-tip contact zone, to circular arc interface cracks. According to their numerical parametric study the extent of the near-tip contact zone can be regarded as approximately independent of the Dundurs parameter α under assumption of an inclusion more rigid than the matrix.

París et al. [15] and Varna et al. [16] compared different aspects of Toya's [8] solution and the elastic solution obtained by boundary element method (BEM) using a contact algorithm, noticeable differences between these two solutions (e.g., in relative opening displacements, radial stresses in the near-tip bonded zone, and in the total ERR) being observed for semidebonding angles greater than 60 deg where a physically relevant near-tip contact zone appears due to a change in the relative orientation of the crack tip with respect to the remote load. Thus, the limited scope of the Toya's [8] work to small debonds is fundamentally associated with the appearance of this near-tip contact zone starting from semidebonding angles of about 60 deg. Varna et al. [11] nevertheless used Toya's results to estimate the critical value of ERR associated with the shear fracture mode of fiber–matrix debonds with semidebond angles over 60 deg, not considering the contribution of the opening mode.

The purpose of this paper is to generate more knowledge about the conditions under which the cracks at micromechanical level, under macrotensile stresses, first grow along the interface and then kink into the matrix. To this end a numerical model based on BEM [17] is generated, involving one long fiber with a debonding embedded into a large matrix. This model is the base to reproduce the mechanism of failure represented in Fig. 1. The model allows separation of the lips of the cracks to occur as well as contact between the debonded lips of the cracks. When a contact zone is detected, the frictionless case is assumed based on the following considerations. First, when the damage is reduced to a crack between the fiber and the matrix it will be shown that the significant phenomena take place for semidebonding angles up to the interval between 60 deg and 70 deg. For semidebondings of this level the contact zone is very small and the contribution of frictional dissipation of energy can be assumed to be negligible in view of the objective of the present work. This is in accordance with the observations of Varna et al. [16] that the presence of friction does not alter the problem qualitatively. Second, when the damage is envisaged by a kinked crack, daughter of the interface crack, it will be shown that the lips of the kinked crack run into the matrix along the direction suitable for working in Mode I, which will prevent the appearance of a contact zone.

Focusing attention on damage originated at long fibers embedded into the matrix and taking into account that experimentally observed debonds, e.g., by Zhang et al. [10], are typically much larger in the axial direction than in the arc direction, the plane strain state in linear elastic formulation is assumed in the present work. The present fracture problems are characterized by the ERR of a crack running along the interface or kinking and penetrating into the matrix.

It has been considered necessary to include a concise review of the results and notation of the interface fracture mechanics used in the present analysis of fiber matrix interface crack propagation in order to make the paper self-contained, this review being performed in Sec. 2. The configurations of the two problems considered, interface crack and kinked crack, are presented in Sec. 3. The fiber–matrix interface crack is studied, in terms of the ERR, fracture mode mixity, and fracture criterion in Sec. 4. The stress state at the neighborhood of the fiber–matrix interface crack tip is studied, in order to predict the kinking of the interface crack, in Sec. 5 using the BEM solution. The study of the ERR of the kinked crack is carried out in Sec. 6. The conclusions are finally presented in Sec. 7.

2 Models and Crack Propagation Criteria of the Interface Fracture Mechanics

The two principal models of interface cracks typically used in their analysis and growth predictions, hereinafter referred to as open model and frictionless contact model, are briefly reviewed. In the open model [5] the interface crack is assumed to be open (with usually traction free crack faces) whereas in the contact model [14] the faces are assumed to be in contact near the crack tip under the load application. The reason for employing these two models is that, as will be seen, neither of them is free of inconsistencies and/or difficulties in its application. There are situations where only one of these models is adequate. However, there are also cases, like that considered in the present study, where both models are required along the fracture process that appears in the problem. In particular, in the present study a switch from the open model to the contact model is required when analyzing the growth of the fiber–matrix debonding. The approaches based on the SIF and ERR concepts can be applied to both models and are introduced in what follows. Finally, some growth criteria for propagation along the interface and for possible kinking out of it will be introduced.

Let us first of all introduce briefly the notation and some definitions valid for both models. Consider two homogeneous isotropic linear elastic materials (denoted 1 and 2), which are perfectly bonded along a surface except for a debonding region referred to as interface crack, subjected to the plane strain state. In the present analysis of the near-tip singular elastic field the interface is considered to be locally flat at the crack tip, which is justified due to an analysis of circular interface cracks by England [6] and Toya [8], and curved ones by Yuan and Yang [18]. Let the local Cartesian system (x,y) and polar coordinate system (r,θ) be defined at the crack tip as shown in Fig. 2. Let E_k and ν_k denote Young elasticity modulus and Poisson ratio, respectively, of material $k=1,2$. The Dundurs bimaterial mismatch parameters α and β are defined as

$$\alpha=\frac{\mu_1(\kappa_2+1)-\mu_2(\kappa_1+1)}{\mu_1(\kappa_2+1)+\mu_2(\kappa_1+1)}=\frac{E_1'-E_2'}{E_1'+E_2'}$$

and

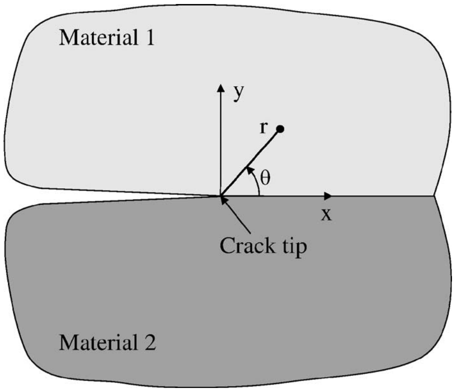


Fig. 2 Local coordinate systems at the interface crack tip

$$\beta=\frac{\mu_1(\kappa_2-1)-\mu_2(\kappa_1-1)}{\mu_1(\kappa_2+1)+\mu_2(\kappa_1+1)} \tag{1}$$

where $\mu_k=E_k/2(1+\nu_k)$ is the shear modulus; $\kappa_k=3-4\nu_k$ is the Kolosov's constant; and $E'_k=E_k/(1-\nu_k^2)$ is the effective elasticity modulus in plane strain.

Finally, the parameter ϵ (typically called oscillatory index due to its relevant role in the open model) is used in expressions of both open and contact models

$$\epsilon=\frac{1}{2\pi}\ln\frac{1-\beta}{1+\beta} \tag{2}$$

Note that $|\alpha|\leq 1$, $|\beta|\leq 0.5$, and $|\epsilon|\leq \ln 3/2\pi=0.175$.

2.1 Near-Tip Solution of the Open Model. Starting from Williams [5] asymptotic expansion, near-tip singular tractions acting along the bonded part of the interface and near-tip relative displacement of crack faces, $\Delta u_i(r)=u_i(r,\theta=\pi)-u_i(r,\theta=-\pi)$ are, respectively, approximated for $r\rightarrow 0$ (r being the distance from the tip) by

$$\sigma_{yy}(r,0)+i\sigma_{xy}(r,0)=\frac{K(r/l)^{i\epsilon}}{\sqrt{2\pi r}} \tag{3}$$

$$\Delta u_y(r)+i\Delta u_x(r)=\frac{8}{1+2i\epsilon}\frac{K(r/l)^{i\epsilon}}{\cosh(\pi\epsilon)E^*}\sqrt{\frac{r}{2\pi}} \tag{4}$$

where i is the imaginary unit; $K=K_1+iK_2$ is the complex SIF; l is a reference length scale introduced by Rice [19]; and E^* is the harmonic mean of the effective Young moduli

$$\frac{1}{E^*}=\frac{1}{2}\left(\frac{1}{E'_1}+\frac{1}{E'_2}\right) \tag{5}$$

The term $(r/l)^{i\epsilon}=\cos(\epsilon\ln r/l)+i\sin(\epsilon\ln r/l)$, for $\beta\neq 0$ and equivalently $\epsilon\neq 0$, is responsible for an oscillatory behavior (including sign changes) in each traction and relative displacements component superimposed over the well-known square root behavior of these components when $r\rightarrow 0$. Due to these oscillations in relative displacements, interpenetrations between crack faces are predicted by this solution in a zone close to the crack tip [20,21]. Let r_i denote the distance of the first interpenetration from the crack tip. Associated with these oscillations in interface tractions, unbounded normal (tensional and compressive), and shear stresses are predicted close to the crack tip independently of the problem configuration, in particular of the far-field load applied. Therefore, SIF components K_1 and K_2, respectively, are not associated with pure opening and shear fracture modes, and consequently no separation of fracture modes is possible here.

Define r_p as the extent of the nonlinear zone, including plastic and fracture process zones, i.e., the zone subjected to very high loads where the fracture process takes place (material separation occurs), and where the elastic solution cannot be regarded as a realistic description of the stress state [22]. In practice, the predicted zones of crack face interpenetrations and traction oscillations can be frequently considered to be physically nonrelevant, r_i being smaller than r_p, it being quite typical that r_i is of an atomic size. The concept of small-scale contact (SSC), introduced by Rice [19] to characterize such situations, provides the theoretical base for application of the open model to interface crack predictions. From a practical point of view, following Rice [19], SSC conditions are associated with situations where r_i is less than 1% of the smallest characteristic length of the problem, e.g., the crack length.

The fracture mode mixity is (in both SIF and ERR approaches) a crucial question to address. One measure of fracture mode mixity, based on the SIF concept, is given by the local phase angle ψ_K ($-\pi\leq\psi_K\leq\pi$) defined by $K=|K|e^{i\psi_K}$ or equivalently, in view of Eqs. (3) and (4), by

$$\psi_K=\arg K=\arg(\sigma_{yy}(l,0)+i\sigma_{xy}(l,0))=\arg(\Delta u_y(l)+i\Delta u_x(l))+\arctan(2\epsilon) \tag{6}$$

where arg is the argument function of a complex number and a sufficiently small length l is considered. Note that $|K|$ is independent of l, whereas according to Eq. (6) ψ_K is an l-dependent measure of fracture mode mixity. Thus, values of ψ_K should always be provided with the associated reference length l. According to Rice [19], ψ_K has only a weak dependence on l for many real bimaterials characterized by a small value of ϵ.

The singular oscillatory term in the asymptotic expansion of the near tip stresses can be expressed as

$$\sigma_{ij}(r,\theta)=\frac{1}{\sqrt{2\pi r}}\{\mathrm{Re}[K(r/l)^{i\epsilon}]\sigma_{ij}^{I}(\theta,\epsilon)+\mathrm{Im}[K(r/l)^{i\epsilon}]\sigma_{ij}^{II}(\theta,\epsilon)\}\quad -\pi\leq\theta\leq\pi \tag{7}$$

where the universal dimensionless functions σ_{ij}^{m} ($m=I,II$) were presented by Rice et al. [23] in polar coordinates.

With reference to the ERR approach, the virtual crack closure method by Irwin [24] when applied to an interface crack considering a small but finite length Δa of a virtual crack extension along the interface, gives the total specific available energy, called ERR, associated with this crack extension (see Sun and Jih [25], Raju et al. [26], and Toya [27])

$$G^{\mathrm{int}}(\Delta a)=G_I^{\mathrm{int}}(\Delta a)+G_{II}^{\mathrm{int}}(\Delta a) \tag{8}$$

where Mode I and II components $G_I^{\mathrm{int}}(\Delta a)$ and $G_{II}^{\mathrm{int}}(\Delta a)$, respectively, correspond to energy released by normal stresses acting through crack opening displacements and shear stresses acting through crack face sliding displacements at the interface crack

$$G_I^{\mathrm{int}}(\Delta a)=\frac{1}{2\Delta a}\int_0^{\Delta a}\sigma_y(r,0)\Delta u_y(\Delta a-r)dr \tag{9}$$

$$G_{II}^{\mathrm{int}}(\Delta a)=\frac{1}{2\Delta a}\int_0^{\Delta a}\sigma_{xy}(r,0)\Delta u_x(\Delta a-r)dr \tag{10}$$

It has to be mentioned that although formally the values of the stresses and displacements in Eqs. (9) and (10) correspond to two different configurations (original length of the crack for the stresses and this length plus Δa for the displacements), both sets of values can be taken, originating no noticeable differences in the ERR for sufficiently small Δa, from the original configuration of the crack, taking stresses and displacements from the opposite sides of the crack tip. This will be the procedure followed in this paper.

Whereas the total ERR G^{int} associated with an infinitesimal virtual crack extension can be written in terms of K as (Malyshev and Salganik [28])

$$G^{\text{int}} = \lim_{\Delta a \to 0} G^{\text{int}}(\Delta a) = \frac{|K|^2}{\cosh^2(\pi\epsilon)E^*} \tag{11}$$

the components of ERR $G^{\text{int}}_{I,II}(\Delta a)$ oscillate as functions of Δa, due to the oscillatory character of the near-tip elastic field, and consequently their limit does not exist as $\Delta a \to 0$. Analyzing a result by Toya [27], Mantič and París [29] showed that

$$G^{\text{int}}_{I,II}(\Delta a) = 0.5G^{\text{int}}(1 \pm F(\epsilon)\cos\{2[\psi_K + \psi_0(\Delta a/l,\epsilon)]\}) \tag{12}$$

where the amplitude function $F(\epsilon)$ and the phase shift angle $\psi_0(\Delta a/l,\epsilon)$ can be evaluated using the following series expansions

$$F(\epsilon) = 1 + (\pi^2/3 - 2)\epsilon^2 + O(\epsilon^4)$$

and

$$\psi_0(\Delta a/l,\epsilon) = \epsilon \ln \Delta a/4el + [\zeta(3) + 4/3]\epsilon^3 + O(\epsilon^5) \tag{13}$$

$\zeta(3)=1.202$ being the Apéry's constant and $e=2.718$ being the base of the natural logarithm. The particular value of ratio $\Delta a/l$ for which the phase shift ψ_0 vanishes is quite independent of ϵ, such values of $\Delta a/l$ being placed between 10.1169 (for $|\epsilon| =0.175$) and 10.8731 (for $\epsilon \approx 0$).

A consequence of oscillations in $G^{\text{int}}_{I,II}(\Delta a)$ is that the energetic phase angle ψ_G ($0 \leqslant \psi_G \leqslant \pi/2$) defined as

$$\tan^2 \psi_G = \frac{G^{\text{int}}_{II}(\Delta a)}{G^{\text{int}}_{I}(\Delta a)}$$

or equivalently

$$\cos 2\psi_G = \frac{G^{\text{int}}_{I}(\Delta a) - G^{\text{int}}_{II}(\Delta a)}{G^{\text{int}}(\Delta a)} \tag{14}$$

represents a Δa-dependent measure of fracture mode mixity. Thus, values of ψ_G should always be provided with the associated length of the virtual crack extension Δa. Notice that when $\epsilon=0$ then $\psi_G=|\psi_K|$.

The following fundamental relation [29] between ψ_K and ψ_G, defined in Eqs. (6) and (14), respectively, can be obtained by substituting Eq. (12) into Eq. (14)

$$\cos 2\psi_G = F(\epsilon)\cos[2(\psi_K + \psi_0)] \tag{15}$$

which implies that

$$\psi_G = 0.5 \arccos\{F(\epsilon)\cos[2(\psi_k + \psi_0)]\}$$

and

$$\psi'_K = 0.5 \arccos[F(\epsilon)^{-1}\cos(2\psi_G)] \tag{16}$$

where $\psi'_K=|\psi_K+\psi_0+n\pi|$, n being an integer number (usually $n=0,\pm 1$) giving $0 \leqslant \psi'_K \leqslant \pi/2$. For small values of ϵ, values of ψ'_K are closely approximated by values of ψ_G except for ψ'_K very close to 0 or $\pi/2$.

2.2 Near-Tip Solution of the Contact Model. The problems already mentioned associated with the open model suggest the possible presence of a contact zone between the lips of an interface crack near the crack tip. Thus, instead of an infinite number of zones where crack face overlapping occurs in the open model at the crack tip, for $\beta \neq 0$, one connected near-tip contact zone appears in this model [14], the extent of this contact zone being denoted as r_c. Due to the presence of a near-tip contact, no fracture Mode I SIF arises, the interface crack growing in Mode II exclusively. Thus, the near-tip singular state is uniparametric, being governed by one multiplicative constant represented by the fracture Mode II SIF: K^C_{II} (C referring to the Comninou contact model). Hence, for a particular bimaterial, relations between values of singular stresses are independent of the far-field load.

According to Comninou [14] near-tip singular tractions and the near-tip relative displacement of crack faces, respectively, are approximated for $r \to 0$ by

$$\sigma_{xy}(r,0) = \frac{K^C_{II}}{\sqrt{2\pi r}}, \quad \sigma_y(r,\pm\pi) = -\frac{\beta K^C_{II}}{\sqrt{2\pi r}} \tag{17}$$

$$\Delta u_x(r) = \frac{8K^C_{II}}{\cosh^2(\pi\epsilon)E^*}\sqrt{\frac{r}{2\pi}} \tag{18}$$

A crucial consequence of the inequality in Eq. (17), implied by a requirement of near-tip compressive stresses between crack faces, is that

$$\beta K^C_{II} \geqslant 0$$

and considering also Eq. (18)

$$\beta \Delta u_x(r) \geqslant 0 \quad \text{for } r \to 0 \tag{19}$$

Thus the sign of both the SIF and the allowed direction of the near-tip relative slip depends only on β, being independent of the far-field load direction. When the global imposed shear loading agrees with this intrinsically allowed slip direction, a relatively large near-tip contact zone may occur. However, when the applied global load tends to originate slip opposite to the allowed near-tip slip direction, only an extremely small contact zone, typically of subatomic size, is predicted at this tip. Thus, in the latter configuration no near-tip contact zone would be observable in experiments, the (locally) open model being adequate for analysis and predictions of crack behavior in this case.

The singular term in the asymptotic expansion of the near-tip stresses can be expressed as

$$\sigma_{ij}(r,\theta) = \frac{K^C_{II}}{\sqrt{2\pi r}}\sigma^C_{ij}(\theta,\beta), \quad -\pi \leqslant \theta \leqslant \pi \tag{20}$$

where the universal dimensionless functions σ^C_{ij} were presented by Comninou [14] in polar coordinates.

Starting from Eqs. (8), (9), and (10) (which are obviously valid in the contact model), taking into account that due to the near-tip contact $G^{\text{int},C}_{I}(\Delta a)=0$ for any sufficiently small Δa, and applying the first expression in Eq. (17) and Eq. (18), the total ERR $G^{\text{int},C}$ associated with an infinitesimal virtual crack extension can be written in terms of K^C_{II} as [14]

$$G^{\text{int},C} = \lim_{\Delta a \to 0} G^{\text{int},C}(\Delta a) = \lim_{\Delta a \to 0} G^{\text{int},C}_{II}(\Delta a) = \frac{(K^C_{II})^2}{\cosh^2(\pi\epsilon)E^*} \tag{21}$$

2.3 Notes on Application of the Interface Crack Models. Although the solution of the contact model, as opposed to the open model solution, is strictly and locally speaking the unique physically correct solution (within the context of linear elasticity) of the interface crack problem regardless of the geometry and the loading conditions, this model is nevertheless not always adequate to characterize fracture. The adequacy of the contact model to characterize an interface crack growth basically depends on the relation between the near-tip contact zone extent r_c (note that r_c is typically of the same order and slightly smaller than the extent of the zone of crack face overlapping in the open model r_i) and the nonlinear zone size r_p [19,30].

SSC conditions are applicable to a situation where r_c is sufficiently smaller than r_p. Then the open model and the contact model solutions approximately coincide outside the nonlinear zone, and the near-tip singular solution of the open model contains all the relevant information to characterize the fracture process. Thus, the singular term of the open model solution, governed by

$K = K_1 + iK_2$, is suitable for representing a fracture mode mixity at the crack tip under SSC conditions. By contrast, the singular term of the contact solution, which is governed by only one real parameter K_{II}^C, is not able to represent any fracture mode mixity, other asymptotically nonsingular terms contributing significantly to the solution value at small but physically relevant distances from the crack tip in the contact model (see e.g., Aravas and Sharma [31]). Finally, note that under SSC conditions $|K| \cong K_{II}^C$ and equivalent by $G^{\mathrm{int}} \cong G^{\mathrm{int},C}$.

On other hand, when r_c is significantly larger than r_p, the open model solution and the contact solution differ significantly outside the nonlinear zone, only the contact solution being able to provide useful information relevant to the process of fracture in the shear mode present at the crack tip. Thus, in such situations the contact model is the appropriate one to analyze and predict interface crack growth.

Therefore, in a general practical numerical procedure for interface crack analysis both models must be taken into account. The open model would be of application when SSC conditions hold, otherwise the contact model is applied. Both models, following this rule, will be used in the analysis carried out in this paper.

2.4 Fracture Path Selection: Criteria for Interface Crack Growth and Kinking. Consider a stationary interface crack subjected to a load. This crack may grow by its further extension along the interface or kink out of the interface. The minimum value of the total ERR G^{int} (or $G^{\mathrm{int},C}$) that originates an interface decohesion is called critical interface ERR and is denoted as G_c^{int} (in the open model assuming SSC conditions), or $G_c^{\mathrm{int},C}$ (in the contact model assuming a relevant near-tip contact zone). It is believed that the cracking path is defined by the local singular stress state at the parent crack tip and by the relation between the critical ERR of the interface and of the material toward which the kink is directed G_c^{kink}.

The competition between interface crack extension and kinking (assuming Mode I propagation after kink) can be formulated on the energetic basis comparing ratios of the corresponding ERRs associated with a load level, G^{int} (or $G^{\mathrm{int},C}$) and G^{kink}, and the critical ERRs for extension and kinking [32]

$$\frac{G^{\mathrm{int}}}{G_c^{\mathrm{int}}} > \frac{G^{\mathrm{kink}}}{G_c^{\mathrm{kink}}} \Rightarrow \text{extension}, \quad \frac{G^{\mathrm{int}}}{G_c^{\mathrm{int}}} < \frac{G^{\mathrm{kink}}}{G_c^{\mathrm{kink}}} \Rightarrow \text{kink} \tag{22}$$

Note that G^{kink} corresponds to a kink angle θ_{kink} predicted by a suitable criterion.

When G_c^{int} is relatively small in comparison with G_c^{kink}, the first inequality in Eq. (22) implies that the interface crack may be trapped at the interface and propagate along it either under SSC conditions in mixed fracture mode (characterized by ψ_K or ψ_G, see Eqs. (6) and (14)) or, in the presence of a physically relevant near-tip contact zone, in pure shear mode.

In the latter case, a single value G_c^{int} is used, the criterion for onset of an interface crack extension being expressed by inequality

$$G^{\mathrm{int},C} \geqslant G_c^{\mathrm{int},C} \tag{23}$$

becoming an equality for quasi-static propagation.

In the former case (SSC conditions), a strong dependence on the mode mixity of G_c^{int} (ψ_K) has been observed in extensive experiments by Evans et al. [33], Liechti and Chai [34], Banks-Sills and Askhenazi [9], and others. Thus, ψ_K (or ψ_G) is an important parameter governing interface crack growth under SSC conditions. From several phenomenological laws for G_c^{int} (ψ_K) suggested in the past, the following family [35], is considered to be representative of a large number of bimaterial systems

$$G_c^{\mathrm{int}}(\psi_K) = G_{Ic}^{\mathrm{int}}[1 + \tan^2(1-\lambda)\psi_K] \tag{24}$$

where G_{Ic}^{int} is the Mode I critical interface ERR (associated with the minimum value of G_c^{int} (ψ_K)) and λ is a fracture mode-sensitivity parameter, e.g., the typical range $0.2 \leqslant \lambda \leqslant 0.3$ characterizes interfaces with moderately strong fracture mode dependence. The interface fracture criterion is expressed now by inequality

$$G^{\mathrm{int}} \geqslant G_c^{\mathrm{int}}(\psi_K) \tag{25}$$

When G_c^{int} is relatively large in comparison with G_c^{kink}, the second inequality in Eq. (22) implies that the interface crack will kink out of the interface. It can be assumed that the kink angle θ_{kink} is determined by the near-tip stress field of the parent interface crack. Note that for application of Eq. (22), a prediction of θ_{kink} by a certain criterion is in fact required.

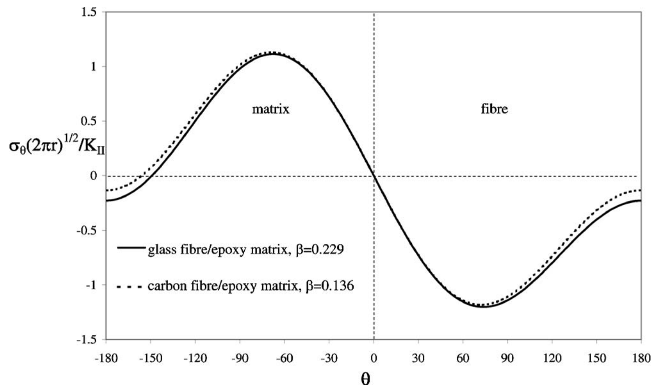


Fig. 3 Angular distribution of σ_θ for $\beta = 0.229$ and $\beta = 0.136$ following Comninou [14]

There are several classical criteria for θ_{kink} predictions, the most simple and widely applied being the MCS criterion by Erdogan and Sih [36]. In comparison with the other popular criterion of the maximum ERR (MERR), discussed by the same authors, which requires kink crack modeling; the MCS criterion requires only knowledge of the near-tip stress solution of the parent interface crack (either that given by Eq. (7) or by Eq. (20)), predicting θ_{kink} in the direction where σ_θ is maximized.

The difficulty with the application of the MCS or MERR criteria to interface crack kinking predictions under SSC conditions at the parent crack when $\beta \neq 0$ is associated with the oscillatory character of the singular near-tip solution for the parent crack in the open model (see Eq. (7)). Due to this oscillatory character, these criteria do not predict a unique value of θ_{kink}, different distances to the tip or kink lengths implying different θ_{kink} values predicted by these criteria (see He and Hutchinson [32] and Geubelle and Knauss [37]). In particular, with reference to MERR criterion, there is no limit of G^{kink} at a certain angle $\theta \neq 0$ for vanishing kink crack length. Thus, a characteristic distance to the tip in the MCS criterion and a characteristic kink crack length in MERR criterion for which these criteria are evaluated must be specified. It should be stressed that although the MCS criterion is applied for θ_{kink} prediction, the nonuniquely defined value of G^{kink} at a certain angle θ associated with an infinitesimal kink crack makes an application of Eq. (22) difficult, an assumption of the characteristic length of the kink crack being necessary.

Application of the MCS or MERR criteria to interface crack kinking predictions in the presence of a physically relevant size of the contact zone at the parent interface crack tip is simplified due to the nonoscillatory and uniparametric character of the singular near-tip solution of the contact model, see Eq. (20). With reference to the MCS criterion, from typical angular variations of σ_θ^C, shown in Fig. 3 (for the two particular values of parameter $\beta = 0.229$ and 0.136, respectively, associated with the bimaterials glass–epoxy and carbon–epoxy studied in the present work), it can be observed that $\sigma_\theta^C < 0$ in the stiffer material (material 1 for $\beta > 0$) for all angles θ. Therefore, only kinking toward the more

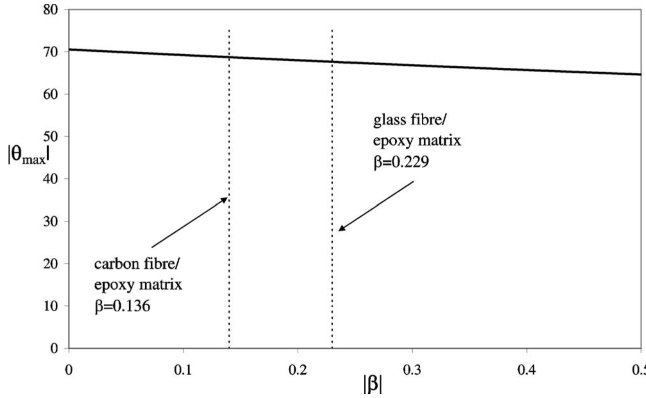


Fig. 4 Angle of maximum σ_θ as a function of β

compliant material is expected (material 2 for $\beta>0$). Applying condition $d\sigma_\theta^C/d\theta=0$ the following expression of θ_{kink} predicted by the MCS criterion is obtained from Eq. (20)

$$\theta_{\text{kink}} = -2\,\text{sgn}(\beta)\arccos\sqrt{\frac{2+|\beta|}{3+|\beta|}}, \quad \text{for } \beta \neq 0 \tag{26}$$

where sgn(·) gives the sign of a real number. Notice that θ_{kink} in Eq. (26) is a unique function of β, thus being independent of the global problem configuration. As can be observed in Fig. 4, θ_{kink} in Eq. (26) does not vary substantially as a function of β, the range of the predicted θ_{kink} values (mentioned already by Hayashi and Nemat-Nasser [38]) being 64.6 deg $\leqslant|\theta_{\text{kink}}|\leqslant$ 70.5 deg. For the particular bimaterials considered in the present work θ_{kink} $=-67.6$ deg (glass–epoxy) and -68.8 deg (carbon–epoxy). Note that angles θ_{kink} observed in experiments by Comninou [39] agree reasonably well with Eq. (26).

An application of Eq. (22) requires the evaluation of G^{kink} for an infinitesimal kink crack, which is well defined in the contact model. After the kink, the whole crack is assumed to be an ordinary one, i.e., open in the zone adjacent to the tip in a homogeneous material, ordinary SIFs $k_{I,II}$ being associated with this tip. Limits of these SIFs for vanishing kink crack length can be expressed as a function of K_{II}^C of the parent crack, according to Leblond and Frelat [40], in terms of the universal dimensionless functions C_{i2}^C as

$$k_I(\theta) = C_{12}^C(\theta,\alpha,\beta)K_{II}^C, \quad k_{II}(\theta) = C_{22}^C(\theta,\alpha,\beta)K_{II}^C \tag{27}$$

Then the ERR at an infinitesimal kinked crack is expressed as

$$G^{\text{kink}}(\theta) = \frac{k_I^2(\theta)+k_{II}^2(\theta)}{E'_{\text{kink}}} \tag{28}$$

where E'_{kink} is the effective elasticity modulus for the material toward which the kink is directed. Note that Eq. (28) can be used not only in Eq. (22), after defining θ_{kink} by Eq. (26), but also in application of MERR criterion to predict values of θ_{kink}, which according to Eqs. (27) and (28) will depend on both bimaterial parameters α and β. Finally, it should be stressed that taking into account Eqs. (21) and (28) with Eq. (27) and the fact that θ_{kink} depends only on the bimaterial, $G^{\text{int},C}/G^{\text{kink}}(\theta_{\text{kink}})$ is also fixed for a bimaterial, being independent of the problem geometry, loadings, and, in particular, independent of K_{II}^C.

Note that the last result may be significant in competition between further crack extension and kinking (when this is considered to be governed by Eq. (22)) in situations where during an interface crack growth the contact zone, originally negligible, comes to be physically relevant. Thus, the crack will not kink, if it has not already kinked before a significant contact zone has arisen. In other words, there are no reasons for a crack to kink once the contact model controls the crack extension process. All this considering that the singular term of the Comninou contact model controls the initiation of the kink. This fact will have implications in the problem considered in this paper.

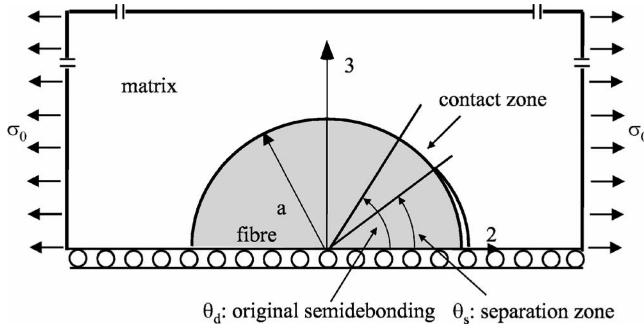


Fig. 5 Model of the single fiber debonded from the matrix

3 Interface and Kinked Cracks Problems

The numerical analysis is performed by BEM [17]. The basic model used in the analysis carried out is shown in Fig. 5 and represents the case of a crack that grows along the interface, a similar configuration having already been studied by París et al. [15]. Due to the symmetry only one half of the problem needs to be studied. This basic model is used in Secs. 4 and 5. The number of boundary elements modeling the fiber is 83 and that corresponding to the matrix is 115. A strongly refined BEM mesh toward the crack tip is applied, the size of the smallest element located at the crack tip being $7\cdot10^{-7}a$, in order to achieve a very high accuracy of the numerical results obtained. All the boundaries have been modeled in this study with continuous linear elements [17].

As several aspects of kinking will be analyzed in the different sections of this paper, slight variations on the basic model will be performed later on. The model represented in Fig. 5 permits the development of a contact zone $(\theta_d-\theta_s)$ between the debonded surfaces of the fiber and the matrix to be taken into consideration, θ_d being the original semidebonding and θ_s being the separation zone corresponding to this semidebonding.

To characterize the problem from the fracture mechanics point of view, the ERR will be used. The expression employed, in the virtual crack closure technique, when the crack propagates from a certain angle α to $\alpha+\Delta\alpha(\Delta\alpha\ll\alpha)$, is

$$G(\alpha,\Delta\alpha) = \frac{1}{2\delta}\int_\alpha^{\alpha+\Delta\alpha}[(\sigma_{rr})_\alpha(u_r)_{\alpha+\Delta\alpha}+(\sigma_{r\theta})_\alpha(u_\theta)_{\alpha+\Delta\alpha}]d\theta \tag{29}$$

where σ_{rr} and $\sigma_{r\theta}$ represent, respectively, radial and shear stresses along the interface and u_r and u_θ the associated relative displacements of the crack lips. The two modes of fracture, I (associated to σ_{rr}) and II (associated to $\sigma_{r\theta}$), are obviously considered in Eq. (29), which corresponds to Eqs. (8), (9), and (10) adapted to the polar coordinate system used in this geometry. The value of $\Delta\alpha$ used in the present calculations is 0.5 deg. The chosen value of $\Delta\alpha$ is sufficiently small to be representative of the value of the total ERR. Substantially smaller values of $\Delta\alpha$ would make the model incoherent with the hypothesis of continuum media assumed. Recall that when the SSC conditions are fulfilled, the components of ERR depend weakly on the value of $\Delta\alpha$ taken (see Eq. (12)), and their limits for $\Delta\alpha\to0$ do not exist. Notice that the superscripts used in Sec. 2 ("int" and "kink") are omitted in what follows for the sake of simplicity.

The presence or not of a physically relevant contact zone between the debonded faces of matrix and fiber will in any case define the character of the fracture mode [15]. When the present

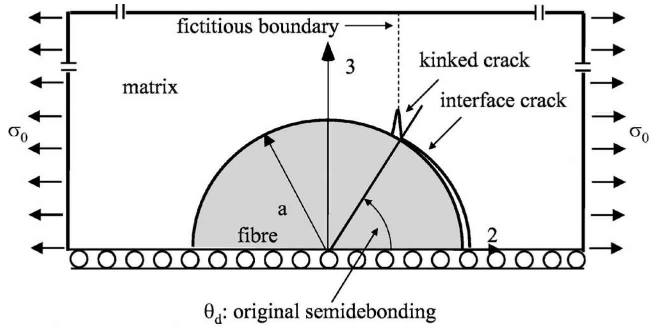


Fig. 6 Model of the single fiber after kinking of the debonding crack

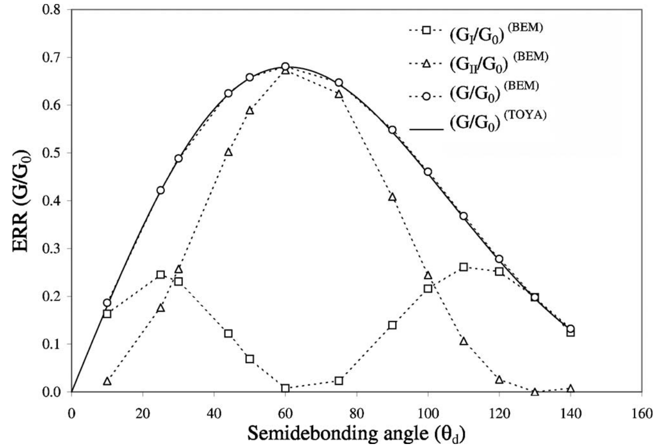


Fig. 7 Values of ERR for the fiber–matrix interface crack under remote tension from BEM (open model) and from Toya's analytical model

numerical model does not detect such a contact zone, both stresses σ_{rr} and $\sigma_{r\theta}$ behave in the model as singular and the crack is working, in accordance with the numerical model developed, in a mixed mode. When a physically relevant contact zone is detected, only $\sigma_{r\theta}$ (with reference to the components of the stress vector beyond the crack tip) reaches a singular value, the crack then working in a pure shear mode.

When modeling the kinked part of the crack becomes necessary, which happens in Sec. 6, the former model is altered, as shown in Fig. 6, in order to represent the case of a crack that has first grown along the interface and then progressed through the matrix.

The original discretization of the basic model is maintained and artificial new boundaries through the matrix are created in order to model the kinked crack, the distribution of elements at the neighborhood of the kinked crack tip being the same as that employed in the interface crack of the basic model. Nodes of the two fictitious boundaries ahead of the crack tip are considered bonded during the numerical calculations, thus representing the continuity of the matrix.

When the ERR for the kinked crack growing in the matrix, modeled as a straight line, is calculated (Sec. 6) the classical Irwin [24] virtual crack closure technique (developed originally for a crack in a homogeneous material) is used in a similar way as described for the previous model.

Two bimaterial systems, those most typical in fiber reinforced composites, have been considered. One corresponds to a glass–epoxy system and the other to a carbon–epoxy system. The properties are listed below as follows:

- Poisson coefficient of the fiber: $\nu^f=0.22$ (glass)/$\nu^f_{12}=0.22$, $\nu^f_{23}=0.25$ (carbon);
- Poisson coefficient of the matrix: $\nu^m=0.33$ (epoxy);
- Young modulus of the fiber: $E^f=7.08\times10^{10}$ Pa (glass)/$E^f_1=2.01\times10^{11}$ Pa, $E^f_2=1.35\times10^{10}$ Pa (carbon); and
- Young modulus of the matrix: $E^m=2.79\times10^9$ Pa (epoxy).

Plane strain state has been considered, effective elasticity properties for the isotropic in-plane problem being evaluated as follows:

- for isotropic materials (glass and epoxy) $E'=E/(1-\nu^2)$, $\nu'=\nu/(1-\nu)$; and
- for transversely isotropic material (carbon) $E'=E_2/(1-\nu_{12}\nu_{21})$, $\nu'=(\nu_{23}+\nu_{12}\nu_{21})/(1-\nu_{12}\nu_{21})$.

Thus, Dundurs parameters, α and β (in Eq. (1)), and the oscillatory index, ϵ (in Eq. (2)), have the following values for both bimaterial systems:

- $\alpha_{\text{glass–epoxy}}=0.919$, $\alpha_{\text{carbon–epoxy}}=0.624$;
- $\beta_{\text{glass–epoxy}}=0.229$, $\beta_{\text{carbon–epoxy}}=0.136$; and
- $\epsilon_{\text{glass–epoxy}}=-0.074$, $\epsilon_{\text{carbon–epoxy}}=0.044$.

In order to make the results corresponding to both systems more easily comparable, the same radius of the fiber, $a=7.5\cdot10^{-6}$ m, has been taken for both fibers (feasible for carbon and glass). In any case, dimensionless results for ERR, G, will be presented in all cases. These dimensionless values of ERR are obtained, based on Toya's [8] approach, dividing the dimensional results by $G_0=[(1+\kappa^m)/8\mu^m]\sigma_0^2 a\pi$, where $\kappa^m=3-4\nu^m$, μ^m is the shear modulus of the matrix, and σ_0 is the modulus of the applied tension.

4 The Fiber–Matrix Interface Crack

Kinking, as considered here, is a micromechanical phenomenon that is supposed to take place once the crack has grown a certain length along the interface. Thus, it seems worthwhile to study first the evolution of the crack along the interface as a previous step to undertaking a kinking study.

The case of a crack along the interface under uni and bidirectional loads has already been studied for a glass fiber and epoxy matrix system by París et al. [41], some of the results affecting the present paper (case of a tension normal to the macromechanical plane of failure) being refreshed and rounded-off in what follows.

First of all, and in order to prove the accuracy of the BEM model employed, a comparison between the numerical and analytical predictions under remote tension is performed. The analytical solution for the open model of interface cracks (Sec. 2) is taken from Toya [8], and due to the fact that it does not prevent interpenetrations, the contact procedure has, just for this comparison, been removed from the numerical analysis. This means that the open model is taken into account in the BEM analysis. The comparison is performed just on the glass fiber–epoxy resin system.

Figure 7 represents the evolution of the ERR values with θ_d. The excellent agreement between the numerical and analytical predictions in the total value of the ERR (components I and II of the numerical solution are also represented) can be seen. It is necessary to insist on the validation character that the results shown in Fig. 7 have over the numerical tool employed. As will be seen, some of these results (e.g., values of G_I for $\theta_d>60$ deg in the presence of interpenetrations) have nevertheless no physical meaning.

Numerically, the smallest defect modeled corresponds to a value of θ_d of 10 deg. To start from this value of the semidebonding can be justified by exploring the initiation of the damage in the interface between the fiber and the matrix, assumed originally in

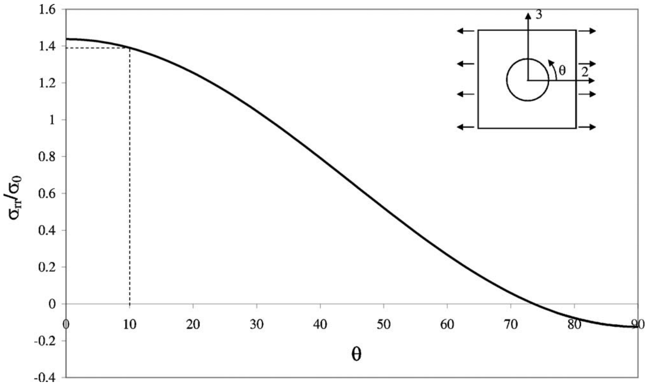


Fig. 8 Distribution of the radial stress between fiber and matrix with no damage at the interface

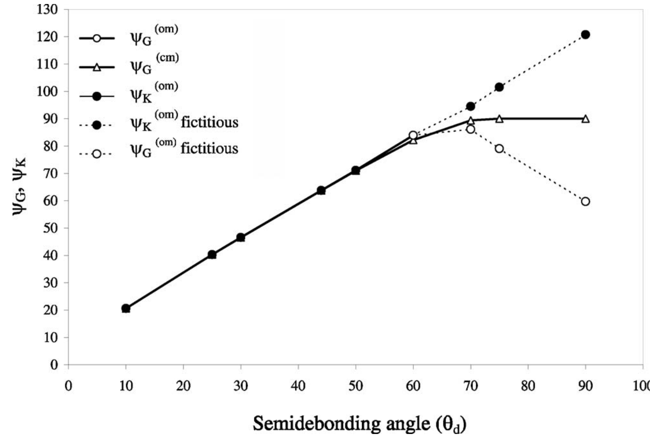


Fig. 10 Evolution of the fracture mode mixity with the semidebonding angle

perfect conditions. To this end, the distribution of the radial stress along the interface, assuming perfect bonding between fiber and matrix, is represented in Fig. 8.

It is plausible to assume that in an undamaged interface, the start of the damage will be controlled by the value of the radial stress in tension. It can be seen from Fig. 8 that the radial stress is quite constant in a zone corresponding to a value of θ_d of 10 deg. Thus, it is quite reasonable to assume that when the radial stress reaches the tensile strength of the interface, it will produce a defect associated to a value of θ_d of an order of 10 deg (or more), starting from which interface fracture mechanics (see Sec. 2) is able to control the growth of such a defect.

Now, once the numerical solution has been validated, as well as the use of interface fracture mechanics, the actual evolution (considering the possibility of appearance of contact between the debonded parts of fiber and matrix) of the ERR with the debonding angle in the presence of a remote tension is calculated. The results shown in Fig. 9 correspond to two different bimaterial systems: glass–epoxy and carbon–epoxy. When a contact zone is detected, these results correspond (as has already been mentioned) to the frictionless case.

First of all, the results represented in Fig. 9 show a clear similarity in the behavior of the dimensionless total ERR, as well as of its components, as a function of θ_d, for both bimaterial systems studied. It indicates that the geometrical features of the problem play a more important role in the variable analyzed (and consequently in the fracture process) than the properties of the materials involved. In what follows, and based on this similarity, the results will be shown only for the glass–epoxy system.

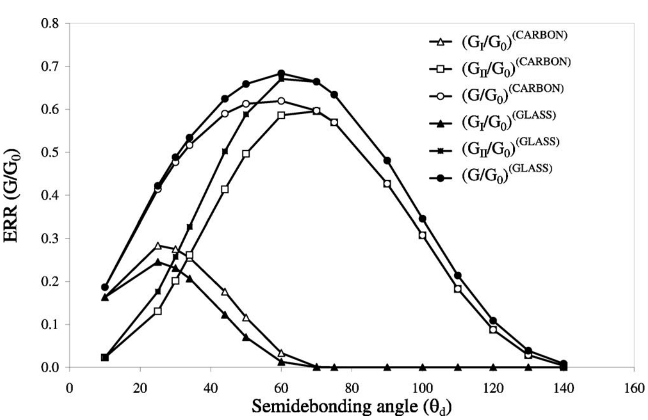


Fig. 9 Values of ERR for the fiber–matrix interface crack under remote tension for two bimaterial systems: carbon–epoxy and glass–epoxy

For the two bimaterial systems studied, a maximum value of ERR appears for θ_d along the interval between 60 deg and 70 deg coinciding with the appearance of a physically relevant contact zone between the debonded parts of the fiber and the matrix [15]. Thus, with reference to the interface fracture mechanic models developed, the open model would be of applicability for values of θ_d under this interval, whereas the contact model would be valid for θ_d over those of the interval. Along the interval of θ_d between 60 deg and 70 deg both models produce numerically similar acceptable values. Along this interval the adequacy goes from the open model over to the contact model (see Sec. 2). The maximum value of ERR appears (for the values of θ_d considered) for $\theta_d = 60$ deg although it is roughly constant for both bimaterial systems along the interval of θ_d between 60 deg and 70 deg.

It should be mentioned that an appearance of a physically relevant near-tip contact zone for large debondings corresponds to the fact that the direction of the globally imposed shear, originated by an inclination of the crack tip zone with respect to the direction of the nominal tension, agrees with the intrinsically allowed slip direction specified by Eq. (19).

G_I is only significant versus G_{II} for small values of θ_d, where the crack tips are oriented almost perpendicular to the loading direction. As the debonding starts to be significant, Mode I starts to disappear and Mode II starts to be absolutely dominant, which happens for θ_d in Fig. 9 greater than 30 deg, coinciding with the appearance of an extremely small near-tip contact zone detected by the present numerical model.

Looking at the evolution of the value of ERR with the debonding, it is clear that as soon as the hypothetical growth of the crack takes place in a pure Mode II (which happens when a physically relevant contact zone is detected) this growth is carried out, in accordance with the decreasing character of ERR observed in Fig. 9 for θ_d over the interval 60–70 deg in a stable manner. However, to describe how the damage (the debonding) may grow starting from an assumed existing damage, formerly evaluated in $\theta_d = 10$ deg, up to the 60–70 deg, interval of θ_d, in a mixed mode of fracture, it is necessary to have an estimate of the critical value of ERR, G_c. As explained in Sec. 2.4, G_c is a function of the fracture mode mixity, which evolve with θ_d, as can be deduced from Fig. 9. Then, first of all the knowledge of the evolution of the fracture mode mixity, characterized by the local phase angles ψ_K and ψ_G, with the value of the debonding is required. For the reasons invoked formerly this will be done for the glass–epoxy system only.

Figure 10 represents the evolution of both phase angles, ψ_G (both for the open model, "om," and contact model, "cm") and ψ_K, as a function of θ_d. ψ_G is calculated using Eq. (14) for $\Delta\alpha$

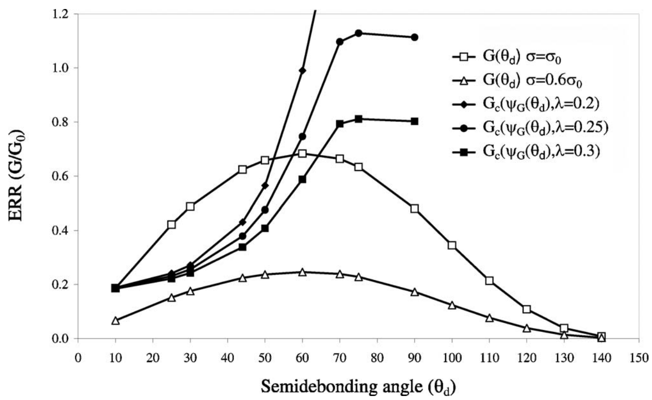


Fig. 11 Evolution of G and G_c as a function of the semidebonding angle

=0.5 deg (see also Eq. (29)), and ψ_K using Eq. (16), taking $\Delta\alpha/l$ (l in terms of an angle) equal to 10.7244, in order to obtain the vanishing phase-shift angle ψ_0 in Eq. (13).

It can be observed that ψ_G and ψ_K take similar values for debondings where the open model is valid. However, as soon as a physically relevant contact zone is detected the three results represented in Fig. 10 start to diverge. ψ_K continues to increase and ends up having physical meaning for ψ_K greater than 90 deg, associated with the presence of a compression ahead of the crack tip. These fictitious values of ψ_K are represented by a dashed line in Fig. 10. The values of ψ_G obtained with an open model also end up having physical meaning when a physically relevant contact zone is detected, the values of ψ_G obtained allowing contact between the interface crack faces being the unique representative, remaining in a 90 deg value when a physically relevant contact zone, and consequently a pure Mode II, appears.

In the zone of interest, up to a semidebonding corresponding to 60–70 deg, the open model is valid and the values of the fracture mode mixity can be used in a prediction of the evolution of the value of G_c with the mixity. In this study the proposal of Hutchinson and Suo [35], Eq. (24) where the phase angle ψ_K is replaced by ψ_G (obtained allowing contact between the interface crack faces), is followed to perform such a prediction. Figure 11 represents the evolution of both G and a hypothetical G_c as a function of θ_d.

For the case of $G_c(\psi_G)$ according to Eq. (24), three values of the parameter λ (0.2, 0.25, and 0.3) have been considered to cover the range of typically accepted values. The curves take constant values when a pure Mode II is detected. For the case of G two curves associated with different fractions (0.6 and 1) of the applied stress σ_0 have been considered. It has been assumed that the stress would increase up until a moment in which the value of G corresponding to θ_d=10 deg reaches the value of G_c. It has been assumed to happen for a value of $\sigma=\sigma_0$.

It can be seen from Fig. 11 that once G reaches G_c for θ_d =10 deg, for all the values of λ considered, the growth of the crack will be unstable but not trespass the interval of semidebondings between 60 deg and 70 deg. Thus, it is plausible (and in accordance with experimental evidence (Zhang et al. [10]) that the semidebonding will not initially pass the interval 60–70 deg and over these values the growth, as previously reasoned, will be carried out in a stable manner.

The change in the character of the growth of the crack from unstable to stable in the neighborhood of θ_d=60 deg makes this zone a favorable one for the appearance of another mechanism of failure.

To summarize the section, it has been proven that a defect originated by the radial stresses and of a size according to the distribution of these radial stresses, will grow in an unstable manner,

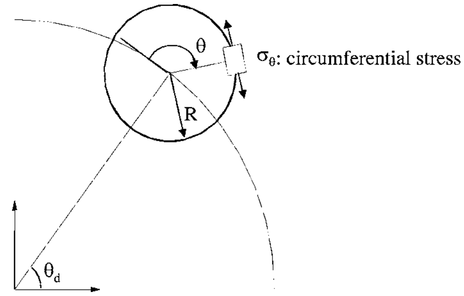


Fig. 12 Configuration of circumferential stresses at the neighborhood of the interface crack tip

stopping the growth in the interval of θ_d between 60 deg and 70 deg. The growth of the interface crack after this interval is stable, an increment in the value of the applied stress being required for an extension of the crack. The mentioned interval of θ_d between 60 deg and 70 deg represents a configuration where the damage is stabilized and another mechanism of damage may take place. This will be studied in the next section.

5 Stress State at the Neighborhood of the Fiber–Matrix Interface Crack Tip: Numerical BEM Solution

In this section the first singular terms of the asymptotic expansions of the elastic solutions at the straight interface crack tip presented in Sec. 2, given by Eqs. (3), (4), and (7) for the open model and Eqs. (17) and (18) and Eq. (20) for the contact model, will be taken as a reference for an analysis of the present situation of the curved interface crack between fiber and matrix. It is considered that the characteristic features of these local solutions, comprehensively discussed in Sec. 2, are of crucial importance when an interface crack propagation out from this interface by kinking is analyzed.

For example, the local distribution of circumferential stress σ_θ in the contact model (see Fig. 3) predicts compressions in the fiber, whereas in the matrix compressions are only in directions which are very inclined and backward with respect to the interface crack propagation, the maximum σ_θ being achieved inside the matrix at the angle θ_d=67.6 deg with respect to the direction tangent to the fiber ahead of the interface crack, obtained by Eq. (26) for β=0.229.

The first question to be considered in the numerical study using BEM is the search for the expected direction of kinking. In this sense it can be assumed that, if the crack growing along the interface changes its direction of growth to penetrate into the matrix, the direction of growth, given by the angle θ_{kink} of Sec. 2.4, will be one along which the circumferential stress (Fig. 12) is maximum. This supposition is based on the fact that Mode I is, in general terms, the main cause of an existing crack growing or kinking, and it is along the MCS direction that the effect of this mode is more dominant.

Thus, the direction of the MCS (defined by an angle $\theta=\theta_{\text{kink}}$, Fig. 12) at the neighborhood of the crack tip corresponding to different semidebondings θ_d at which kinking is assumed to take place, will first be determined. The selection of the characteristic distance to the crack tip, referred to as radius R (Fig. 12) at which the study is going to be carried out, must be performed carefully. On one hand the value of R must be small enough for the stresses to control the possible change of direction of the crack, but on the other hand must be large enough to maintain the physical meaning of the matrix as a continuum medium. Two values of R (R =0.1% and 1% of the fiber radius a) satisfying these conditions have been considered. The results are shown in Fig. 13, where the angle φ=90 deg+θ_d−θ_{kink}, used by Toya [8], represents the MCS direction with respect to the load (horizontal) direction.

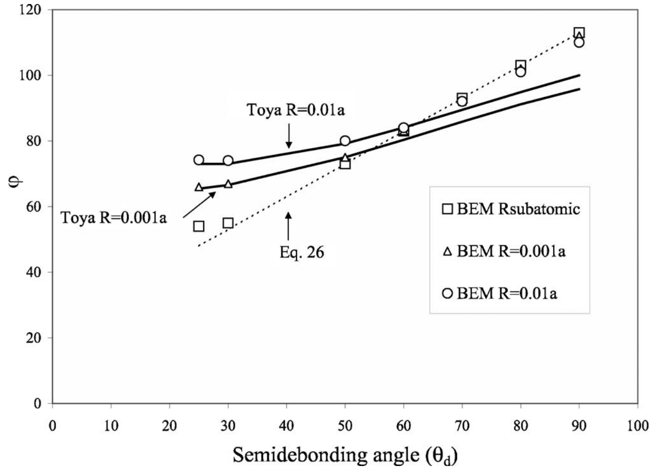


Fig. 13 Value of the angle φ defining the MCS direction as a function of the semidebonding angle

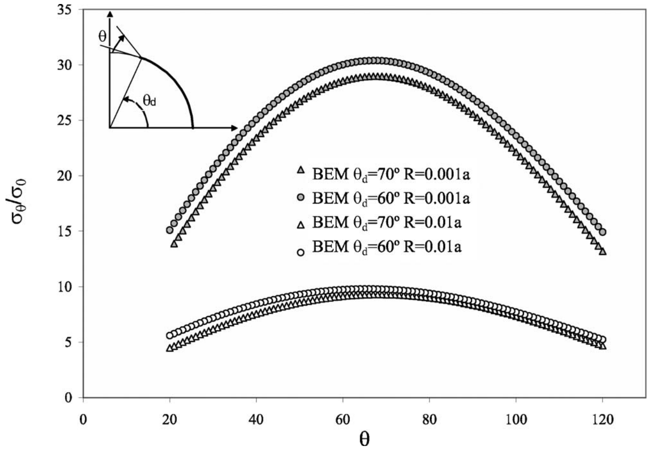


Fig. 15 Distribution of the circumferential stresses around the crack tip of the interface crack

In addition to the numerical results, those associated to the analytical solutions by Toya [8] and by Eq. (26) obtained from the singular asymptotic term of the Comninou contact solution have also been included. It can be observed that Toya's results, represented by continuous lines in Fig. 13, agree excellently with BEM results in the interval of debondings where there is no contact zone, whereas in the presence of a contact zone BEM results are very close to the values predicted by Eq. (26). Due to the different properties of fiber and matrix, Comninou contact model solution always predicts a near-tip contact zone, although of a very small size for θ_d below 50–60 deg (being of subatomic size for θ_d about 40 deg and smaller). Thus, taking a value of R at a subatomic scale, BEM predictions can still adjust predictions by Eq. (26), as can be observed in Fig. 13 for values of θ_d of an order of 30 deg.

Once the numerical solution has been checked, it can be observed from Fig. 13 that the evolution of the angle of MCS with the semidebonding, for the two values of R studied, is quite consistent. It is in any case instructive, in order to have a physical meaning for these results, to represent them on the scheme of a fiber as a function of the semidebonding (Fig. 14). An interval of directions is represented when there is a variation in the angle of MCS for different radii R of inspection between 0.01 and 0.001 of fiber radius. The angle predicted by Eq. (26) is also represented for comparison, it being possible to observe that the value of this angle is placed in the interval of directions numerically predicted by the MCS criterion for θ_d over 50 deg. As mentioned above, this is due to the fact that for these debonds, a non-negligible contact zone appears at the interface crack tip and the singular asymptotic term of the Comninou contact solution starts to govern the near-tip solution at physically meaningful distances, such as those $R=0.01a$ and $R=0.001a$ considered.

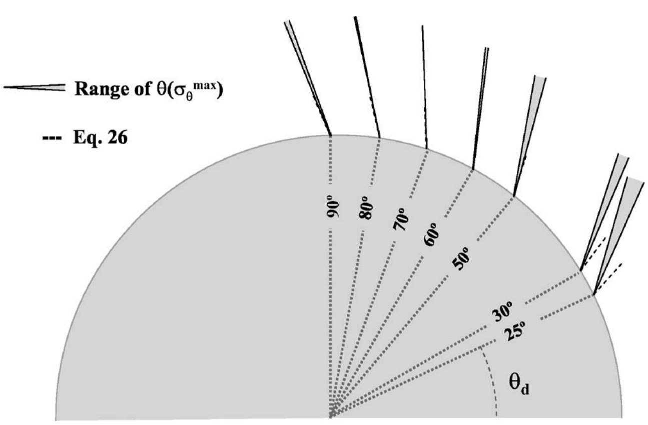


Fig. 14 Graphical description of the evolution of the angle of MCS with the semidebonding angle

It is clear that the angle of MCS is approximately oriented normal to the nominal load for values of θ_d in the interval between 60 deg and 70 deg, which again makes this interval as the candidate to change the damage from debonding to kinking.

The distribution of the circumferential stresses along a circumference of radius R centered at the crack tip is shown in Fig. 15 for the two different debondings θ_d defining the interval of interest and for the two values of R already mentioned.

First, the state of circumferential stress is quite similar for the extremes of the interval where it is assumed that the debonding crack stops. This similarity refers to both the values of the stress and the position where the maximum is reached (as was already observed in Figs. 13 and 14). Also, it has to be mentioned that the shape of the distribution is not affected by the radius taken for the inspection, which permits these distributions to be taken as representative of the stress state controlling the kinking of the debonding crack.

It is now of interest to extend the study carried out for $\theta_d=60$ deg and 70 deg to a wider range of θ_d in order to find out, for θ_d apart from the mechanically expected range of θ_d at which kinking may occur, how the value of MCS evolves. After the similarity in the results obtained in Fig. 14 for the two radii considered, now only the case of $R=0.01a$ will be considered. Thus, Fig. 16 shows for this radius the circumferential stresses at the neighborhood of the crack tip versus the circumferential coordinate.

It can be perfectly observed that the circumferential coordinate for which the circumferential stress reaches its maximum does not change too much. Additionally, it can also be observed in the figure that for each θ_d there is a certain interval of values of θ (of an order of 10 deg for significant semidebondings of $\theta_d=60$–70 deg) at which the circumferential stresses are quite similar to the maximum value. Finally, it can also be noticed that there are no values of MCS significantly greater than those corresponding to the interval of θ_d between 60 deg and 70 deg (in fact the maximum values correspond to $\theta_d=60$ deg).

6 Energy Release Rate at Kinking

In the previous section the direction of MCS at the neighborhood of the crack tip for different lengths of the interface crack has been studied. The purpose of that study was to determine the direction in which the crack would penetrate into the matrix, if

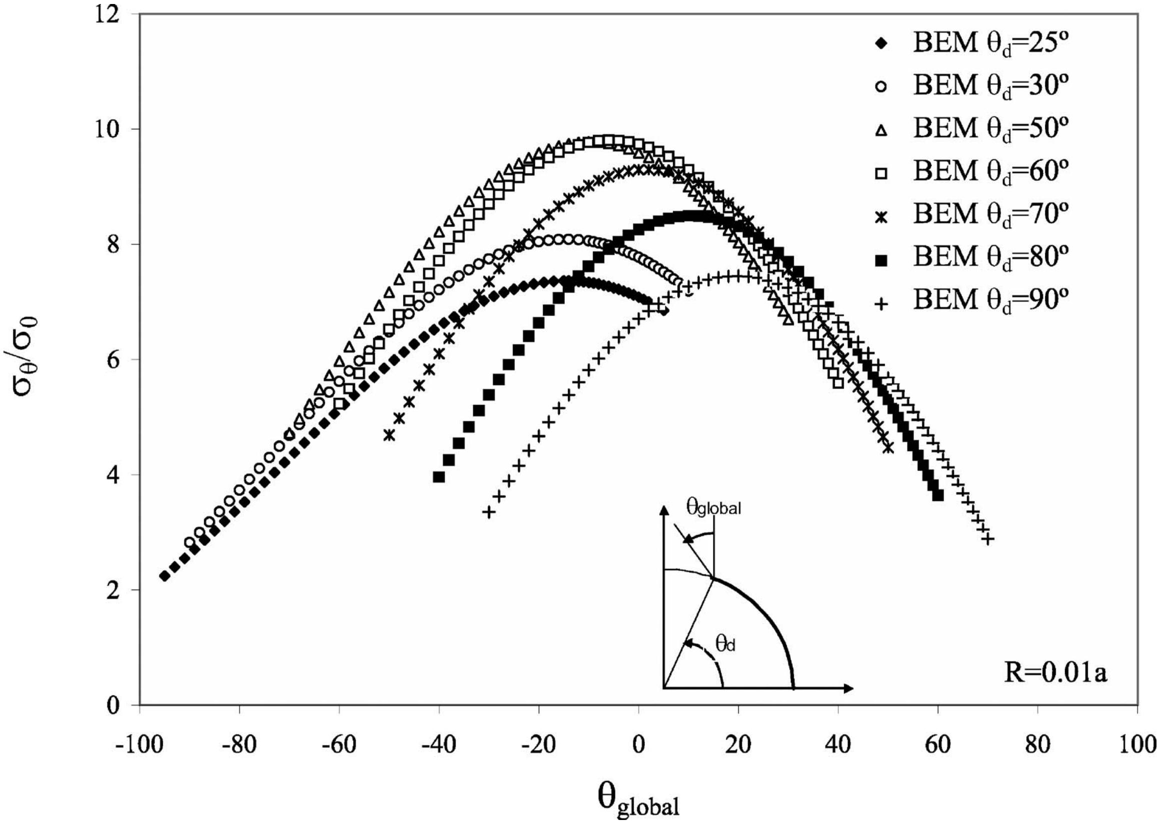


Fig. 16 Distribution of circumferential stresses at the neighborhood of the crack tip, for different semidebondings

appropriate circumstances (similar to those considered in this study) caused this to happen, in accordance with the MCS criterion, Sec. 2.4.

This section will evaluate how possible it is for a crack that is growing along the interface to turn into the matrix following the kinking direction previously determined, using the criterion for competition between interface crack extension and kinking described in Eq. (22). To this end an ERR analysis at kinking is going to be performed.

It has to be pointed out that in this paper only a comparison of the ERR of a crack growing along the interface versus a crack leaving the interface and penetrating into the matrix is going to be performed. This comparison cannot be completely conclusive regarding the appearance of kinking due to the fact that the critical values of G, G_c, of the interface and the matrix, appearing in Eq. (22), would also be involved in the phenomenon.

A description of the configuration of the complete crack when kinking appears is first of all performed. The description is with reference to the case of $\theta_d=70$ deg. The kinked part of the crack is always open independently of the length of the crack. This is coherent with the almost normal orientation of the kinked part of the crack with respect to the load applied. When the length of the kinked crack is very small, the previous contact zone existing in the interface crack before the kinking appears still remains, although being of a very small size. When the length of the modeled kinked crack is not very small, the previous contact zone at the interface crack completely disappears when the load is applied. In view of all this, the whole crack (the interface and the kinked parts) was represented as fully open in Fig. 6.

The ERR by a kinked crack associated with $\theta_d=70$ deg and penetrating into the matrix along the vertical direction (almost coincident with MCS) is now calculated. The results of this case are shown in Fig. 17, where ERR values appear, as well as their components, versus the length of the kinked crack.

The unequal contribution of the two modes of failure to the total ERR is first of all noticeable. While contribution of Mode I, G_I, is very important, the contribution of Mode II ERR component, G_{II}, is almost nonexistent. This result was foreseeable, in view of the completely transversal position of the applied load in relation to the direction of the kinked crack, for any length of the kinked crack.

Referring to the evolution of the ERR, it can be observed that it increases with the crack length. Thus, taking into account that Mode I completely dominates the growth, it can be concluded that the crack propagation is unstable. In this situation, once kinked, no additional load increase would be necessary for the crack to continue growing.

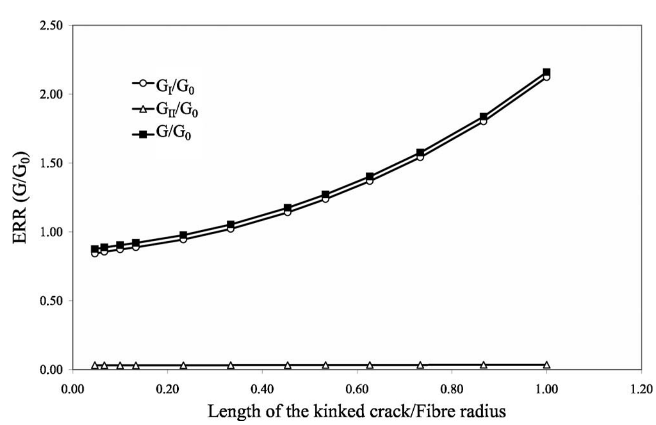


Fig. 17 Values of the ERR and its components for a kinked crack

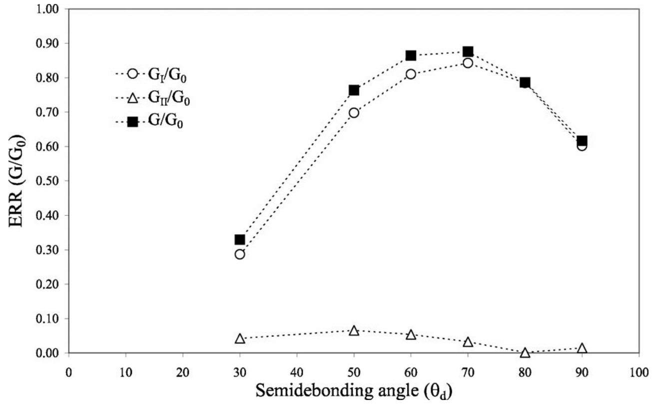


Fig. 18 Values of ERR for the shortest kinked cracks corresponding to different semidebondings

The asymptotic tendency shown in Fig. 17 by the evolution of the ERR at the origin (vanishing length of the kinked crack) is quite noticeable. It appears clear from the figure that ERR does not go to zero when the length of the kinked crack is very short. On the contrary, there is for a vanishing kinked crack a finite value of the ERR, in agreement with the analysis associated with Eq. (28), Sec. 2.4. This situation is coherent with the fact that the total length of the crack is not only the length of the kinked crack but also includes that of the interface crack and, consequently, is not zero when the kinked crack vanishes. Moreover, this finite limit value of the ERR of the kinked crack in general does not coincide with the value of the ERR for the crack continuing to grow along the interface, as will be seen later on.

Having studied the case of θ_d=70 deg, the behavior for different θ_d is now investigated in order to evaluate the variation with θ_d of the ERR of the kinked crack. To this end the results obtained for different θ_d are shown together in Fig. 18. In all cases, vertical kinked cracks, as in the case studied above, have been considered. The purpose being to show tendencies of how ERR of the kinked crack changes with θ_d, only the values for the shortest kinked cracks studied, whose size is $0.05a$, have been included in the representation. In any case, it has to be mentioned that the rising character of the ERR with the length of the kinked crack was observed for all the values of θ_d considered.

As can be observed from Fig. 18, a dashed line instead of a continuous one has been used to link points belonging to the same group of values of ERR (Mode I, Mode II, and total), aiming to emphasize the existent disconnection between them. All points plotted in the graph represent an ERR (Mode I or Mode II component or total) associated with a kinked crack of the same size but coming from a different interface crack (characterized by θ_d), so that values of the ERR of the same group (Mode I, Mode II, or total) correspond to cracks that have kinked from different places of the interface, a physical evolution from one crack to another then being impossible.

The tendency shown in Fig. 18 for the values of ERR as a function of the debonding favors the idea that, if kinking appears, the most plausible values of θ_d are those between 60 deg and 70 deg, where G_I and G reach a maximum.

In order to have more information about the plausibility of such a phenomenon, according to the competition criterion Eq. (22) it is necessary (although not sufficient to conclude whether the kinking appears because it involves the G_c values of the interface and matrix) to compare the ERR of the crack when initiating the propagation through the matrix, shown in Fig. 18, with that released when continuing to grow along the interface. Figure 19 represents this comparison and for the sake of completeness the values of the ERR of a kinked crack growing in the direction of MCS have also been included, in addition to the case of vertical kinked cracks predominantly studied here.

First of all, the two cases of the kinked cracks considered show a coherent evolution with θ_d. The values coincide at θ_d=70 deg where the direction of MCS is approximately normal to the load, the values of ERR at the neighborhood of 70 deg being very close to each other. The main discrepancies appear for small angles where, in agreement with Fig. 14, the MCS direction is quite far from the normal direction to the load, the effect of Mode I, which was the dominant contributor to ERR, then being less important.

With reference to the comparison of the ERRs for the kinked crack versus the crack continuing to grow along the interface, the significantly greater values of the ERR of the kinked crack in the interval of interest are quite apparent. It should also be remem-

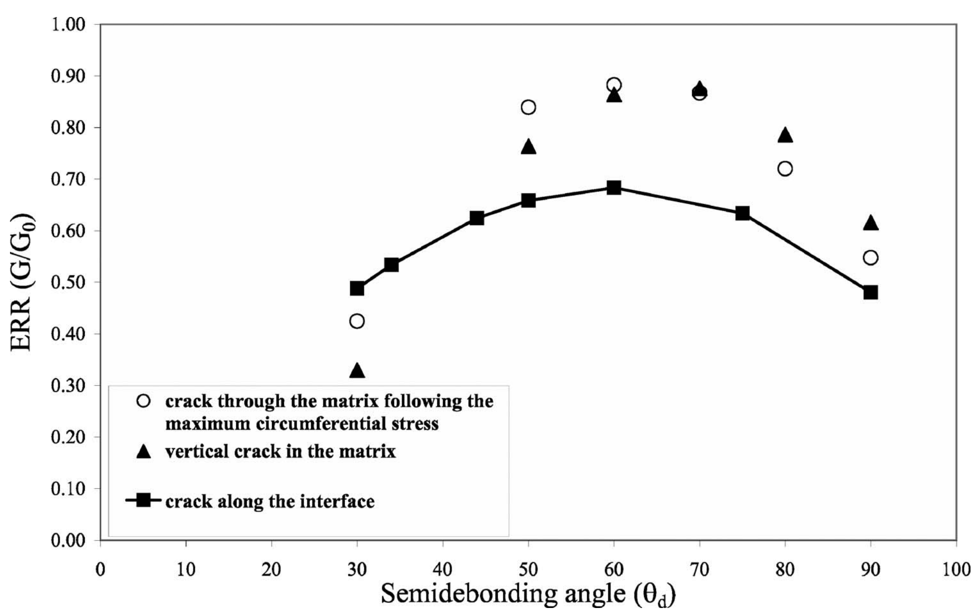


Fig. 19 Comparison between ERR of an interface crack and kinked cracks growing normal to the load and along the direction of MCS

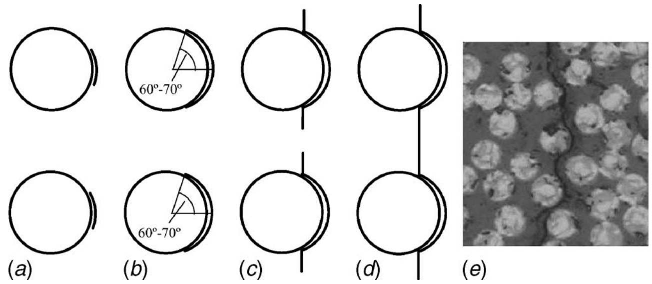


Fig. 20 Schematic development of the damage: (*a*) damage in the form of the small cracks originated by the radial stress, Fig. 8; (*b*) unstable growth of the crack until a semidebonding of 60–70 deg is reached, Fig. 11; (*c*) kinking of the interface crack, Figs. 13–15 and 18; (*d*) unstable growth of the kinked crack, Fig. 17, originating a macrocrack; and (*e*) actual damage in a fibrous composite under transversal load

bered that the presumably unstable growth of the interface crack (after Figs. 9 and 11) up to a θ_d at the neighborhood of 60–70 deg where the stable growth clearly starts, would theoretically prevent the crack (unless interference with another crack arose) from separating from the interface. At θ_d between 60 deg and 70 deg the maximum differences appear between the ERR of the crack continuing to grow along the interface and of the kinked crack, it having been made clear previously that whereas the growth of the crack along the interface becomes stable, the growth of the crack penetrating the matrix is unstable.

In addition, and although as was previously stated, the prediction of growth of the crack along two alternative paths would imply the knowledge of the fracture resistance parameters G_c for the interface and for the matrix, it should be remembered at this point that whereas the growth along the interface is in Mode II the growth along the matrix is in Mode I, the values of G_c for Mode I being for a determined configuration (material or interface) significantly smaller than the values of G_c for Mode II.

Finally, under the assumption that the Comninou frictionless contact model governs the debonding growth for larger debonding angles, the analysis presented after Eq. (28) implies that, in the case of a weak interface (relatively small value of G_c for the interface), if the debonding crack does not kink for θ_d in the interval between 60 deg and 70 deg, where the near-tip contact zone becomes physically relevant, it will not kink for larger debondings.

All this supports the idea that, if kinking takes place, the most plausible values of θ_d at which kinking can appear are in the interval between 60 deg and 70 deg, which does not prevent the existence, without kinking of interface cracks corresponding to semidebonding angles greater than those indicated.

7 Conclusions

Numerical evidence has been generated by means of a planar micromechanical model for the growing and kinking of an interface crack between fiber and matrix under load transversal to the fiber. The information generated can be summarized in the following points, which will refer to Fig. 20 where a schematic development of the damage between two adjacent fibers transversally placed to the orientation of the load is shown:

(a) Assuming that initially the interface between fiber and matrix is undamaged, the failure is assumed to start controlled by the radial interface tension. In accordance with Fig. 8, it is reasonable to assume, due to the distribution of the radial stresses, that the minimum possible damage can be envisaged as a semidebonding θ_d of about 10 deg, as represented in Fig. 20(*a*).

(b) Assuming the formerly mentioned semidebonding θ_d, the crack will start to grow when the ERR equals the fracture toughness of the interface. In accordance with Fig. 11, it will make the crack grow to a θ_d in the interval between 60 deg and 70 deg as represented in Fig. 20(*b*). Special mention must be made of the assumption of variation of the fracture toughness with the fracture mode mixity, which changes in turns with the length of the interface crack.

(c) Exploring the most plausible direction of kinking in the formerly mentioned semidebondings interval, it has been found that the direction of MCS corresponds with that oriented normal to the direction of application of the load (see Figs. 13–15). It is along this direction that the energy released by the kinked crack is maximum (Fig. 19). All this supports the appearance of a kinked crack as represented in Fig. 20(*c*).

(d) The growth of the kinked crack (Fig. 17) is unstable and, because of the relative orientation of the kinked crack with respect to the load, under almost pure Mode I. This supports the idea that a complete macrocrack will appear by the coalescence of two adjacent kinked interface cracks (Fig. 20(*d*)).

When comparing the ERRs for an interface crack in the range of θ_d=60–70 deg for the two alternative situations of continuing to grow along the interface or penetrating into the matrix, it has been found that the ERR in the case of kinking is greater (by an order of 25%) than in the case of continuing to grow along the interface. This fact has to be added to the previous conclusions that the growth of the interface crack after 60 deg is stable, whereas that of the kinked crack is unstable.

The described mechanism of failure can be clearly connected with that observed in experiments run in the laboratory with a carbon fiber composite Z-19.775, which is represented in Fig. 20(*e*). In view of the connections between the predictions generated in this paper using interfacial fracture mechanics and the observed behavior, it can be concluded that interfacial fracture mechanics has been shown to be efficient in explaining the generation and propagation of damage in the interfiber failure of a fibrous composite under transversal load.

Interfacial fracture mechanics was already used by the authors to explain the effect of bidirectional loading transversal to the fibers direction [41]. Now under study is the application of interfacial fracture mechanics to explain the apparent macroscopic angle of failure under transversal compressive loading. All the knowledge generated using this tool may help in the near future to generate physically based failure criteria of fibrous composites.

Acknowledgment

The work was supported by the Spanish Ministry of Education and Culture under Grant No. MAT2000-1115. The authors thank Dr. E. Graciani (University of Seville) and Mr. R. Ávila (Airbus España) for their BEM code used to solve elasticity problems in this study and to Dr. E. K. Gamstedt (KTH Solid Mechanics, Sweden) for a photograph from his Ph.D. thesis.

References

[1] Hinton, M. J., Soden, P. D., and Kaddour, A. S., guest eds., 1998, "Failure Criteria in Fiber-Reinforced-Polymer Composites (Special Issue)," Compos. Sci. Technol., **58**(7).
[2] Hinton, M. J., Soden, P. D., and Kaddour, A. S., guest eds., 2002, "World Wide Failure Exercise on Failure Predictions in Composites (Special Issue)," Compos. Sci. Technol., **62**(12–13).
[3] Hashin, Z., and Rotem, A., 1973, "A Fatigue Failure Criterion for Fiber Reinforced Materials," J. Compos. Mater., **7**, pp. 448–464.
[4] Gamstedt, E. K., and Sjögren, B. A., 1999, "Micromechanisms in Tension-Compression Fatigue of Composite Plies Containing Transverse Plies," Compos. Sci. Technol., **59**, pp. 167–178.
[5] Williams, M. L., 1959, "The Stress Around a Fault of Crack in Dissimilar Media," Bull. Seismol. Soc. Am., **49**, pp. 199–204.
[6] England, A. H., 1966, "An Arc Crack Around a Circular Elastic Inclusion," J.

Appl. Mech., **33**, pp. 637–640.
[7] Perlman, A. B., and Sih, G. C., 1967, "Elastostatic Problems of Curvilinear Cracks in Bonded Dissimilar Materials," Int. J. Eng. Sci., **5**, pp. 845–867.
[8] Toya, M., 1974, "A Crack Along the Interface of a Circular Inclusion Embedded in an Infinite Solid," J. Mech. Phys. Solids, **22**, pp. 325–348.
[9] Banks-Sills, L., and Ashkenazi, D., 2000, "A Note on Fracture Criteria for Interface Fracture," Int. J. Fract., **103**, pp. 177–188.
[10] Zhang, H., Ericson, M. L., Varna, J., and Berglund, L., A., 1997, "Transverse Single-Fiber Test for Interfacial Debonding in Composites: 1. Experimental Observations," Composites, Part A, **28A**, pp. 309–315.
[11] Varna, J., Berglund, L. A., and Ericson, M. L., 1997, "Transverse Single Fiber Test for Interfacial Debonding in Composites 2: Modelling," Composites, Part A, **28**, pp. 317–326.
[12] Prasad, P. B. N, and Simha, K. R. Y., 2003, "Interface Crack Around Circular Inclusion: SIF, Kinking, Debonding Energetics," Eng. Fract. Mech., **70**, pp. 285–307.
[13] Chao, R., and Laws, N., 1997, "The Fiber-Matrix Interface Crack," J. Appl. Mech., **64**, pp. 992–999.
[14] Comninou, M., 1977, "The Interface Crack," J. Appl. Mech., **44**, pp. 631–636.
[15] Paris, F., del Caño, J. C., and Varna, J., 1996, "The Fiber-Matrix Interface Crack—A Numerical Analysis Using Boundary Elements," Int. J. Fract., **82**, pp. 11–29.
[16] Varna, J., Paris, F., and del Caño, J. C., 1997, "The Effect of Crack-Face Contact and Fibre/Matrix Debonding in Transverse Tensile Loading," Compos. Sci. Technol., **57**, pp. 523–532.
[17] Paris, F., and Cañas, J., 1997, *Boundary Element Method. Fundamentals and Applications*, Oxford University Press, Oxford, UK.
[18] Yuan, F. G., and Yang, S., 1997, "The Curved Interfacial Crack Between Dissimilar Isotropic Solids," Int. J. Solids Struct., **34**, pp. 641–660.
[19] Rice, J. R., 1988, "Elastic Fracture Mechanics Concepts for Interfacial Cracks," J. Appl. Mech., **55**, pp. 98–103.
[20] England, A. H., 1965, "A Crack Between Dissimilar Media," J. Appl. Mech., **32**, pp. 400–402.
[21] Erdogan, F., 1965, "Stress Distribution Bonded Dissimilar Materials with Cracks," J. Appl. Mech., **32**, pp. 403–410.
[22] Broberg, K. G., 1999, *Cracks and Fracture*, Academic Press, London.
[23] Rice, J. R., Suo, Z., and Wang, J. S., 1990, "Mechanics and Thermodynamics of Brittle Interfacial Failure in Bimaterial Systems," *Metal-Ceramics Interfaces*, M. Rühle et al., eds., Pergamon, New York, pp. 269–294.
[24] Irwin, G. R., 1957, "Analysis of Stresses and Strain Near the End of a Crack Transversing a Plate," J. Appl. Mech., **24**, pp. 361–364.
[25] Sun, S. T., and Jih, C. J., 1987, "On Strain Energy Release Rates for Interfacial Cracks in Bi-material Media," Eng. Fract. Mech., **28**, pp. 13–20.
[26] Raju, I. S., Crews, J. H., and Aminpour, M. A., 1988, "Convergence of Strain Energy Release Rate Components for Edge-Delaminated Composite Laminates," Eng. Fract. Mech., **30**, pp. 383–396.
[27] Toya, M., 1992, "On Mode I and Mode II Energy Release Rates of an Interface Crack," Int. J. Fract., **56**, pp. 345–352.
[28] Malyshev, B. M., and Salganik, R. L., 1965, "The Strength of Adhesive Joints Using the Theory of Cracks," Int. J. Fract. Mech., **1**, pp. 114–128.
[29] Mantič, V., and París, F., 2004, "Relation Between SIF and ERR Based Measures of Fracture Mode Mixity in Interface Cracks," Int. J. Fract., **130**, pp. 557–569.
[30] Hills, D. A., and Barber, J. R., 1993, "Interface Cracks," Int. J. Mech. Sci., **35**, pp. 27–37.
[31] Aravas, N., and Sharma, S. M., 1991, "An Elastoplastic Analysis of the Interface Crack With a Contact Zones," J. Mech. Phys. Solids, **39**, pp. 311–344.
[32] He, M.-Y., and Hutchinson, J. W., 1989, "Kinking of a Crack Out of an Interface," J. Appl. Mech., **56**, pp. 270–278.
[33] Evans, A. G., Rühle, M., Dalgleish, B. J., and Charalambides, P. G., 1990, "The Fracture Energy of Bimaterial Interfaces," Metall. Trans. A, **21A**, pp. 2419–2429.
[34] Liechti, K. M., and Chai, Y. S., 1992, "Asymmetric Shielding in Interfacial Fracture under In-plane Shear," J. Appl. Mech., **59**, pp. 295–304.
[35] Hutchinson, J. W., and Suo, Z., 1992, "Mixed Mode Cracking in Layered Materials," Adv. Appl. Mech., **29**, pp. 63–191.
[36] Erdogan, F., and Sih, G. C., 1965, "On the Crack Extension in Plates Under Plane Loading and Transverse Shear," J. Basic Eng., **85**, pp. 519–527.
[37] Geubelle, P. H., and Knauss, W. G., 1994, "Crack Propagation at and Near Bimaterial Interfaces: Linear Analysis," J. Appl. Mech., **61**, pp. 560–566.
[38] Hayashi, K., and Nemat-Nasser, S., 1981, "On Branched, Interface Cracks," J. Appl. Mech., **48**, pp. 529–533.
[39] Comninou, M., 1990, "An Overview of Interface Cracks," Eng. Fract. Mech., **37**, pp. 197–208.
[40] Leblond, J. B., and Frelat, J., 2001, "Crack Kinking from an Interface Crack with Initial Contact Between the Crack Lips," Eur. J. Mech. A/Solids, **20**, pp. 937–951.
[41] París, F., Correa, E., and Cañas, J., 2003, "Micromechanical View of Failure of the Matrix in Fibrous Composite Materials," Compos. Sci. Technol., **63**, pp. 1041–1052.

Residual Elastic Strains in Autofrettaged Tubes: Variational Analysis by the Eigenstrain Finite Element Method

Alexander M. Korsunsky
e-mail: alexander.korsunsky@eng.ox.ac.uk

Gabriel M. Regino

Department of Engineering Science,
University of Oxford,
Parks Road,
Oxford OX1 3PJ, UK

Autofrettage is a treatment process that uses plastic deformation to create a state of permanent residual stress within thick-walled tubes by pressurizing them beyond the elastic limit. The present paper presents a novel analytical approach to the interpretation of residual elastic strain measurements within slices extracted from autofrettaged tubes. The central postulate of the approach presented here is that the observed residual stress and residual elastic strains are secondary parameters, in the sense that they arise in response to the introduction of permanent inelastic strains (eigenstrains) by plastic deformation. The problem of determining the underlying distribution of eigenstrains is solved here by means of a variational procedure for optimal matching of the eigenstrain finite element model to the observed residual strains reported in the literature by Venter et al., 2000, J. Strain Anal., ***35****, p. 459. The eigenstrain distributions are found to be particularly simple, given by one-sided parabolas. The relationship between the measured residual strains within a thin slice to those in a complete tube is discussed.*
[DOI: 10.1115/1.2711222]

1 Introduction

Residual stresses are present in all manufactured products, at a higher or lower level, by design or as a side effect of some other processing operation. Residual stresses affect the performance of many different devices and components, from strained layers and optical components in microelectronics, to microcantilever sensors in biological applications, to large engineering components such as the aero-engine fan and turbine blades.

Residual stresses not only affect the deformation behavior and performance of components and structures, but also respond themselves to externally applied thermal and mechanical loads, be it during component manufacture, assembly, or in service. The interaction between residual stresses and fatigue performance is particularly interesting for many applications.

The present study described an application of a broad and generic concept of eigenstrain theory of residual stresses. The fundamental postulate of this theory is that residual stresses and residual elastic strains, quantities that are simply related to each other through generalized Hooke's law, are in fact secondary parameters that arise within the material as a consequence of elastic equilibration following the introduction of permanent inelastic strains, or *eigenstrains*. Elastic equilibration here means the process of generation of additional elastic strains that preserve strain compatibility, but also give rise to residual stresses that satisfy the equilibrium conditions within the body, and traction-free conditions at boundaries.

Some readers may perceive a conceptual difficulty here: how can one be entirely certain that, as a consequence of the introduction of some inelastic strains into an object, no plastic deformation or creep would take place? The way to overcome this difficulty is to realise that the separation of strains into eigenstrains and response strains is the prerogative of the person performing the analysis: *by definition* one chooses to call *all* the inelastic strains arising from various processes (plastic deformation, creep, transformation, cutting and joining, etc.) by the collective term eigenstrain. All the strains that remain must then necessarily be elastic. Note that residual stresses only arise in connection and in numerical correspondence to this elastic part of the total strain within the object. Of course, one must not disregard the possibility of additional inelastic strains occurring during subsequent deformation of the object. Nevertheless, at any particular instant in time it is possible to point out unequivocally the inelastic part of strain, and to designate it eigenstrain, and the elastic part of strain, and to associate this part with the elastic equilibration response, and the attendant residual stress distribution.

The knowledge of eigenstrain, or inelastic strain distribution is a better quality, *more fundamental knowledge* than that of residual stresses. The basis of this statement is as follows. Consider a residually stressed body that contains a distribution of eigenstrain. If this distribution is known the complete residual elastic strain and residual stress states can be readily reconstructed through the single application of the theory of elasticity. This so-called *direct* problem of eigenstrain theory is simple, since it is linear and does not require iterative solution. The answer can be obtained by the application of elasticity theory with a perturbation in the Saint-Venant strain compatibility condition [1], or from a special formulation of the finite element method that makes an allowance for the existence of eigenstrain distributions [2,3].

If, on the other hand, the residual stress is known, then the underlying eigenstrain distribution cannot be readily obtained from elastic theory. It appears, in fact, that the solution of this *inverse* problem of eigenstrain theory would require iterative procedures. Nevetheless, it is possible to develop an efficient regularization of the inverse problem that allows an approximation to the unknown eigenstrain distribution to be determined without iteration. This framework will be presented below.

A further advantage flows from knowing the eigenstrain distribution. If the object of study is subjected to a nondisturbing sectioning operation (imagine separating the object into two halves by using an electric discharge machining tool), then the prior knowledge of residual stress state provides little help in deducing the residual stress states in the resulting pieces. If, on the other hand, the eigenstrain distribution were known prior to sectioning,

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received September 30, 2005; final manuscript received August 15, 2006. Review conducted by Sanjay Govindjee.

then the newly created residual stress states can be readily computed, each by the application of an easy procedure for solving the direct eigenstrain problem. Furthermore, knowledge of the underlying eigenstrain distributions provides the link between residual states before and after sectioning, thus overcoming the principal difficulty in dealing with residual stress states that often baffles researchers. In fact, using eigenstrain as the basis for the modeling approach allows utilization of the information on residual stresses prior to and after sectioning to achieve more accurate determination of eigenstrain than each of those states separately.

Finally, and additional advantage of the eigenstrain approach lies in the fact that residual stress can be assessed not only at locations where it was originally measured, but *everywhere* within the object. One has to exercise certain care in interpreting the results, of course, since the reconstructed residual stress state is only accurate within the state space spanned by the functional form chosen for the eigenstrains. Nevertheless, the eigenstrain approach provides a neat way of incorporating the constraints imposed by the requirements of continuum mechanics: traction-free boundary conditions, strain compatibility, etc. In generic terms, the present residual stress reconstruction approach seeks to approximate the entire *field*, while the traditional methods rely on *pointwise* interpretation.

The above discussion outlines the broad context of the present study as the interpretation of experimental measurements of residual elastic strains (or their changes) with the purpose of reconstructing the complete residual stress field, and of identifying the underlying inelastic strain distribution that acts as the source of the residual stress state. More specifically, the present study addresses the analysis of axially symmetric inelastic deformation of thick-walled tubes subjected to internal pressure exceeding the elastic limit of the tube. As the pressure is increased, it induces plastic deformation within a hollow cylinder with the external boundary that progresses outward from the inner bore toward the outer surface of the tube.

The original analysis of the axisymmetric deformation of thick-walled tube is attributed to Lamé's work published in 1852 [4], and can be found for example in Soutas-Little [5]. The elasto–plastic axially symmetric deformation of thick-walled tubes has been reviewed by Den Hartog [6] and solved under various specific assumptions by Gao et al. [7–9].

The experimental data used in the present study were published by Venter et al. [10]. Two specimens were considered designated by the authors of the strain measurement study as specimen B and specimen C, respectively. We present the results of eigenstrain modeling of residual elastic strains that were measured by neutron diffraction in both cases.

Approximate matching of conventional direct finite element (FE) calculations to residual elastic strain measurements by neutron diffraction has already been carried out and reported by Venter et al. [10]. This was repeated by the present authors as the first tentative step of the analysis. However, the conditions of autofrettage loading and strength properties of the material of the tube (e.g., complex hardening behavior) are not known in sufficient detail. It is therefore impossible to carry out the *direct* FE solution to obtain a prediction of the "correct" residual stress state. A key advantage of the proposed eigenstrain approach is that the knowledge of yield behavior of the material is not required for matching—it suffices to know the elastic properties only.

2 Description of Experimental Data

Experimental results of strain evaluation in slices obtained from autofrettaged tubes were described by Venter et al. [10], who used three different techniques: neutron diffraction, Sachs boring [11], and the compliance method [12]. We focus particular attention on the neutron diffraction method, since the authors of that paper conclude that this method provides the most detailed and reliable assessment of the residual elastic strains.

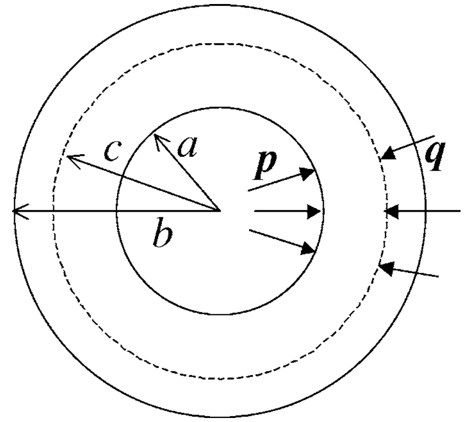


Fig. 1 Schematic illustration for the description of axisymmetric deformation of a thick-walled tube of internal radius *a* and external radius *b* under internal pressure *p*. Parameter *c* indicated the radius of the elastic–plastic boundary, and *q* is the pressure transmitted across this interface.

Strain evaluation by diffraction is achieved by monitoring the shift in the position of the scattered peak center. This can be done either in the angular dispersive configuration, by scanning the scattering angle 2θ from a monochromatic beam, or in the energy dispersive mode, by monitoring the energy of scattered photons using an energy-resolving detector and a white beam.

It is then possible to deduce the lattice orientation specific elastic strain using Bragg's law

$$2d_{hkl} \sin \theta = \lambda = \frac{hc}{E} \tag{1}$$

using the following formula

$$\varepsilon_{hkl} = \frac{d_{hkl} - d_{hkl}^0}{d_{hkl}^0} \tag{2}$$

The selection of the most appropriate diffraction peak is a separate research topic in experimental strain analysis. Here we contend ourselves with following the conclusions of Venter et al. [10] that the (211) peak in steel presents the best choice for the purposes of estimating the engineering macroscopic average strain.

The experimental data of Ref. [10] were collected using monochromatic neutron diffraction on the residual strain instrument of the SAFARI-1 research reactor operated by NECSA in Pretoria, South Africa.

Figure 1 illustrates the geometry of the annular "slices" obtained from autofrettaged tubes by Venter et al. [10]. Included in this diagram is the putative position of the elastic–plastic boundary. This boundary is not known a priori and depends not only on the maximum pressure applied during autofrettage treatment, but also on the mechanical response characteristics of the tube material.

Figure 2 illustrates a possible arrangement for diffraction measurement and indicates the positions of the incident and diffracted beams. Diffraction peak is obtained from the scattering volume occupying a cube of approximately 4 mm side length, thus providing significant averaging.

The markers in Figures 3 and 4 indicate the values of the radial and hoop residual elastic strains, respectively, measured in specimen B [10]. The continuous curves in these diagrams correspond to the predictions from eigenstrain modeling, and will be discussed in some detail below.

The markers in Fig. 5 and 6 indicate the values of the radial and hoop residual elastic strains, respectively, measured in specimen C [10]. Once again, continuous curves in these diagrams correspond to the predictions from eigenstrain modeling.

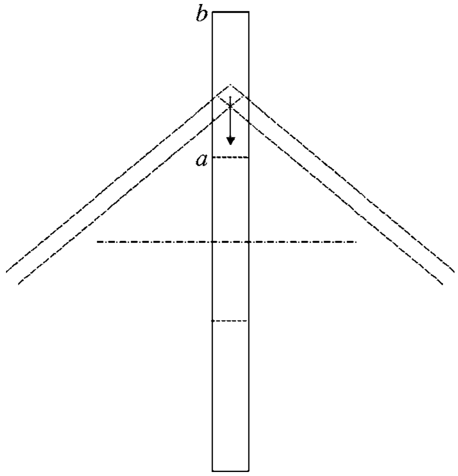


Fig. 2 A possible arrangement of autofrettaged tube slices with respect to the incident and diffracted beams. The dashed lines indicate the incident and diffracted beams; the arrow shows the scattering vector that indicates the orientation of the strain component being measured (radial in the present example).

3 The Direct Problem of Autofrettage

We use the eigenstrain finite element (eFE) to compute the residual stress fields within the transverse annular slices of the autofrettaged tubes. The commercial ABAQUS finite element package (Version. 6.3) was used in the axisymmetric formulation with a simple rectangular uniform mesh. Eigenstrains were represented using anisotropic temperature dependent pseudo-thermal strains [2].

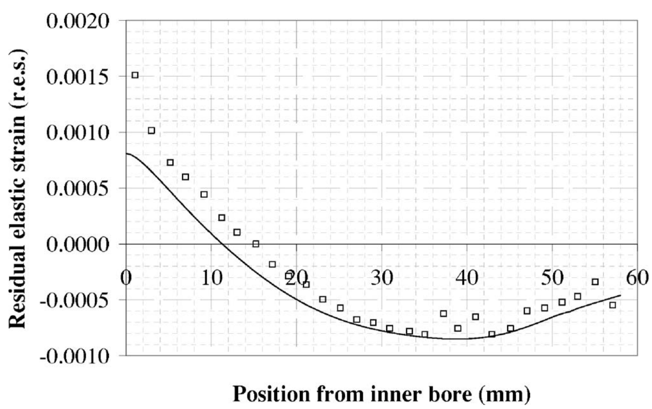


Fig. 3 Radial residual elastic strain in specimen *B*: experimental measurements (markers) and eigenstrain model prediction (continuous curve).

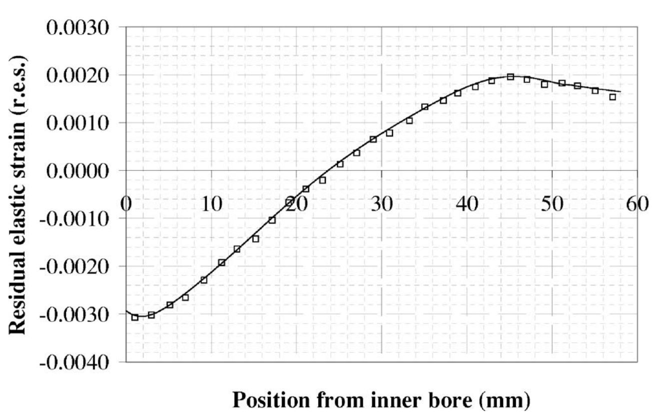


Fig. 4 Hoop residual elastic strain in specimen *B*: experimental measurements (markers) and eigenstrain model prediction (continuous curve)

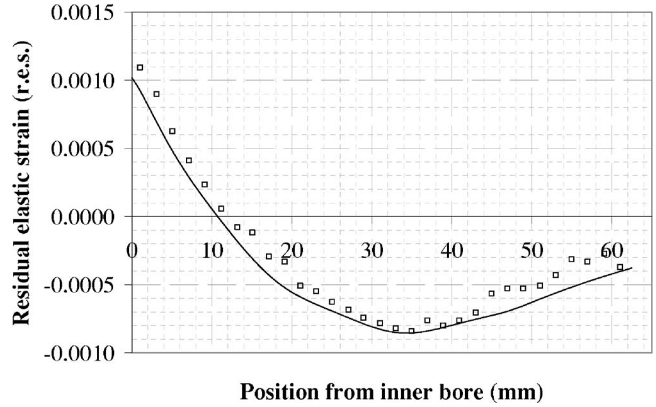


Fig. 5 Radial residual elastic strain in specimen *C*: experimental measurements (markers) and eigenstrain model prediction (continuous curve)

The axial eigenstrain effects are ignored in the present study, on the basis of the assumption of plane strain persisting within long cylindrical objects loaded transversely to their generator, and following the experiments carried out within the simulation. The radial and hoop eigenstrains are assumed to be equal and opposite in order to enforce the plastic incompressibility condition.

The unknown eigenstrain distribution may therefore be readily described by a single function of the radial position within the cylinder: we ignore any possibility of circumferential strain variation. Moreover, in the present problem it is immediately clear that the extent of the eigenstrain distribution is limited by the inner bore of the tube, on the one hand, and the outer radius, a priori unknown, of the plastic zone.

In order to develop a solution for the direct problem of eigenstrain theory for the present problem we use an axially symmetric finite element model and introduce the eigenstrains by way of anisotropic pseudo-thermal strains represented by a truncated series of basis functions, so that

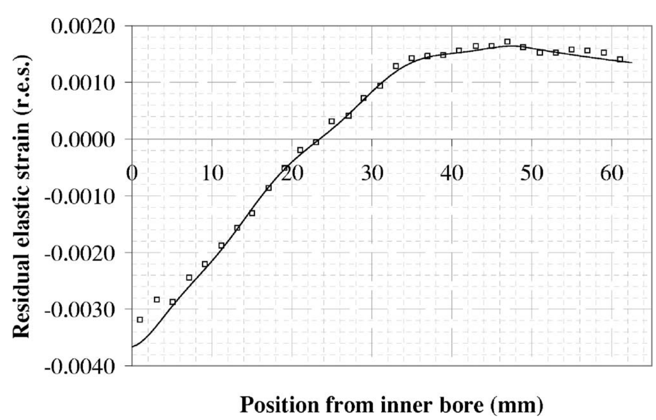


Fig. 6 Hoop residual elastic strain in specimen *C*: experimental measurements (markers) and eigenstrain model prediction (continuous curve)

$$-\varepsilon^*_{rr}(r) = \varepsilon^*_{\theta\theta}(r) = \sum_{i=1}^{N} c_i \xi_i(r) \tag{3}$$

Here N is the total number of basis distributions used in the model. In practice a suitable choice of the basis functions could be made in the form of Chebyshev polynomials, so that

$$\xi_i(r) = T_{i-1}(x) \tag{4}$$

where the variable x is introduced to provide a normalization of the radial range of eigenstrain distribution to interval $[-1,1]$ of conventional definition of Chebyshev polynomials. We choose the eigenstrain distribution range to exceed the likely correct extent of the plastic zone, and allow the variational procedure described below to determine the real extent over which the eigenstrain must be introduced.

It must be noted here that the solution of the direct eigenstrain problem can be readily obtained for any eigenstrain distribution by an essentially elastic calculation within the eFE model. In particular we note that this task is easily accomplished for each of the basis functions $\xi_i(r)$.

We further note that due to the problem's linearity, the solution of the direct problem described by a linear combination of individual eigenstrain basis functions $\xi_i(r)$ with coefficients c_i is given by the linear superposition of solutions with the same coefficients.

This observation provides a basis for formulating an efficient variational procedure for solving the *inverse* problem about the determination of underlying eigenstrain distribution. This procedure is introduced in the following section.

4 Eigenstrain Inverse Problem Formulation

The problem that we wish to address in the present study stands in an inverse relationship to the one described in the previous section. In practice most frequently the residual elastic strain distribution is known, usually partially, e.g., from diffraction measurement. The details of the preceding deformation process need to be found, such as the depth of the plastic zone. Alternatively, in the absence of nondestructive measurements of residual elastic strain, changes in the elastic strain can be monitored, e.g., using strain gauges, in the course of material removal. In all cases the purpose is to determine the unknown parameters of the deformation.

In practice the residual elastic strain, or its increments, can only ever be measured at a finite number of points. We are therefore seeking to find the unknown deformation parameters, such as the plastic zone outer radius, c, by matching the residual elastic strain distributions predicted by the elasto–plastic model of the previous section, to the finite data set of measured values.

Questions arise regarding the invertibility of the problem; its uniqueness; the regularity of solution, i.e., whether the solution depends smoothly on the unknown parameters. Although we give no answer to these questions here, we present a constructive inversion procedure that can subsequently be evaluated in terms of its uniqueness and regularity.

Consider a set of experimental data consisting of the values of *radial* residual elastic strains (r.e.s.) ε^j_{rr} and *hoop* residual elastic strains (r.e.s.) $\varepsilon^j_{\theta\theta}$ collected at positions r_j, $j=1,\ldots,m$. Thus, we assume that the data were collected from a one-dimensional scan in coordinate r, but that two components of residual elastic strain were measured at each point. It is worth noting, however, that the approach presented below is not in any way limited to problems arising from one-dimensional scans, and can be readily generalized to two- and three-dimensional cases.

Now denote by $e_{rr}(r)$ and $e_{\theta\theta}(r)$ the *predicted*, or *modeled* distributions of, respectively, the radial and hoop components of the residual elastic strain. Evaluating $e_{rr}(r)$ and $e_{\theta\theta}(r)$ at each of the measurement points gives the *predicted values* $e^j_{rr}=e_{rr}(r_j)$ and $e^j_{\theta\theta}=e_{\theta\theta}(r_j)$. In order to measure the goodness of the prediction we form a functional J given by the sum of squares of differences between actual measurements and the predicted values, with weights

$$J = \sum_{j=1}^{m} [w_j(\varepsilon^j_{rr} - e^j_{rr})^2 + w_j(\varepsilon^j_{\theta\theta} - e^j_{\theta\theta})^2] \tag{5}$$

The choice of weights w_j can be made on the basis of additional information available; for example, they could be chosen based on the accuracy of individual measurements being interpreted. However, in the present analysis we shall ascribe equal unit weights to all squared differences appearing in Eq. (5).

Minimization of functional J provides a rational variational basis for selecting the most suitable model to match the measurements, in terms of the overall goodness of fit. Given a set of fixed measurements ε^j_{rr} and $\varepsilon^j_{\theta\theta}$, J can be thought of as a function of the residual stress state, or any set of parameters that describes its generation. One important possibility is to use *eigenstrains* to define the residual stress and residual elastic strain state. These are inelastic strains of any origin that are responsible for producing residual stress. In principle, a continuous distribution of eigenstrains must be described by (at worst) three-dimensional variation of six scalars (or one symmetric second rank tensor). One way to reduce the complexity of the problem is to represent the unknown source eigenstrain distribution as a linear combination, i.e., a truncated sum of basis functions with unknown coefficients in Eq. (3).

The results of the previous section contain the procedure for the solution of the direct problem, i.e., the determination of the residual elastic strain distribution that arises in response to an arbitrary eigenstrain distribution $\varepsilon^*(x)$. This procedure can now be applied to each of the N *basis* distributions $\xi_i(x)$ in turn. As a result, a family of residual elastic strain solutions $E_i(x)$ is obtained.

Due to the linearity of the direct problem, the predicted values e^j_{rr} and $e^j_{\theta\theta}$ of the radial and hoop residual elastic strains due to the eigenstrain distribution $\varepsilon^*(x)$ of the equation can themselves be written in the form of a superposition of responses to the *basis* eigenstrain distributions, namely

$$e^j_{rr} = \sum_{i=1}^{N} c_i E^i_{rr}(x_j) = \sum_{i=1}^{N} c_i e^{ij}_{rr} \tag{6}$$

with the same coefficients c_i as in Eq. (3).

The inverse problem of determining the unknown eigenstrain distribution $\varepsilon^*(x)$ has now been reduced to the problem of determination of N unknown coefficients c_i that deliver a minimum to the functional J in Eq. (5), which may now be rewritten as

$$J = \sum_{j=1}^{m} \left[w_j\left(\varepsilon^j_{rr} - \sum_{i=1}^{N} c_i e^{ij}_{rr}\right)^2 + w_j\left(\varepsilon^j_{\theta\theta} - \sum_{i=1}^{N} c_i e^{ij}_{\theta\theta}\right)^2 \right] \tag{7a}$$

For the purposes of avoiding laborious algebraic manipulations the above equation is rewritten as

$$J = \sum_{j=1}^{m} w_j \left(\sum_{i=1}^{2N} c_i e_{ij} - y_j\right)^2 \tag{7b}$$

This equation can be seen as a shorthand notation for Eq. (7*a*), where e_{ij} is taken to denote e^{ij}_{rr} for $1<i<N$, $e^{ij}_{\theta\theta}$ for $N<i<2N$, and y_j is taken to denote ε^j_{rr} for $1<i<N$, $\varepsilon^j_{\theta\theta}$ for $N<i<2N$.

The expression in equation (7*b*) is quadratic and positive definite in the unknown coefficients c_i. It follows that the functional has a unique minimum that is found by satisfying the condition

$$\nabla_c J = 0, \quad \text{or } \partial J/\partial c_i = 0, \quad i=1,\ldots,N \tag{8}$$

Due to the quadratic nature of the functional in Eq. (7*a*), the system of equations in Eq. (8) is linear. Therefore, the solution for the unknown coefficients c_i can be readily found *without iteration* by inverting the linear system arising in Eq. (8). This system is written out explicitly below.

The partial derivative of J with respect to the coefficient c_i can be written explicitly as

$$\partial J/\partial c_i = 2\sum_{j=1}^{m} w_j e_{ij}\left(\sum_{k=1}^{N} c_k e_{kj} - y_j\right)$$
$$= 2\left(\sum_{k=1}^{N} c_k \sum_{j=1}^{m} w_j e_{ij} e_{kj} - \sum_{j=1}^{m} w_j e_{ij} y_j\right) = 0 \qquad (9)$$

For purposes of illustration, let us now assume that the weights are equal to unity, so that Eq. (9) simplifies to

$$\partial J/\partial c_i = 2\left(\sum_{k=1}^{N} c_k \sum_{j=1}^{m} e_{ij} e_{kj} - \sum_{j=1}^{m} e_{ij} y_j\right) = 0 \qquad (10)$$

We introduce the following matrix and vector notation

$$\mathbf{E} = \{e_{ij}\}, \quad \mathbf{y} = \{y_j\}, \quad \mathbf{c} = \{c_i\} \qquad (11)$$

Noting that notation e_{kj} corresponds to the transpose of matrix $\mathbf{E}$, the entities appearing in Eq. (10) can be written in matrix form as

$$\mathbf{A} = \sum_{j=1}^{m} e_{ij} e_{kj} = \mathbf{E}\mathbf{E}^T, \quad \mathbf{b} = \sum_{j=1}^{m} e_{ij} y_j = \mathbf{E}\mathbf{y} \qquad (12)$$

Hence Eq. (10) assumes the form

$$\nabla_c J = 2(\mathbf{A}\mathbf{c} - \mathbf{b}) = 0 \qquad (13)$$

The solution of the inverse problem has thus been reduced to the solution of the linear system

$$\mathbf{A}\mathbf{c} = \mathbf{b} \qquad (14)$$

for the unknown vector of coefficients $c=\{c_i\}$.

Whenever the solution of an inverse problem is sought, questions arise concerning the existence and uniqueness of the solution, and also concerning the well posedness of the problem, i.e., the continuity of the dependence of the solution on the problem parameters, the choice of the basis functions, the number of terms N in the truncated series, etc.

Within the present regularized formulation of the problem, for an *arbitrary* choice of the family of basis functions and *arbitrary* number of basis functions N, a unique solution is guaranteed to exist. This is a consequence of the positive definiteness of the quadratic functional J. Furthermore, it is clear that increasing the number of terms N is guaranteed to deliver a sequence of monotonically nonincreasing values of J, i.e., the goodness of approximation will not be diminished.

An interesting question concerns the convergence of the solution, e.g., in terms of eigenstrain distribution $\varepsilon^*(x)$, to the "true" solution, in the limit $N \to \infty$. Similarly, the continuity in the behavior of the solution with the choice of basis functions deserves to be discussed. While it must be emphasized that these questions are clearly fundamental and ought to be addressed, the focus is currently placed on the development of a practical tool for residual strain analysis. In sofar as this is the aim of the present study, the proposed framework offers an efficient "one shot" approach to the solution the of the inverse problem. Furthermore, the choice of moderate values N, compared to the number of measurements m, offers a rational procedure for smoothing the data.

5 Results and Discussion

Results of the application of the variational procedure for the determination of the underlying eigenstrain distribution for specimens B and C of Venter et al. [10] are presented in Fig. 3–8.

In Fig. 3 the markers indicate the experimentally measured residual elastic strains in the radial direction reported in Ref. [10]. The continuous curve represents the residual elastic strain distribution predicted by the variational eigenstrain model. It is important to note that the reconstructed residual elastic strain profile shown in this figure is *not* the result of trying to obtain the best approximation to the radial strain data alone. Instead the model prediction must be considered together with the result shown in Figure 4, where the experimentally measured values of the *hoop* strain are compared with model predictions. Note that the functional J was introduced in the previous section as the measure of the goodness of match to *both* the radial and hoop strain distributions *simultaneously*. As a consequence the procedure minimizes the total mismatch between the model and experiment measured as the sum over all of the available data. It is worth noting in passing that if some additional data (e.g., another set of measurement points) were to become available, the analysis procedure can be readily repeated, and thus the additional data would be incorporated in the analysis.

Figures 5 and 6 contain the experimental data for the radial and hoop residual elastic strains, respectively, obtained by Venter et al. [10] from specimen C. Once again, the markers indicate experimental measurement data, and continuous curves show the model predictions.

In both cases the model provides a satisfactory approximation to the experimental data, although the agreement is not perfect. There are several possible reasons for this lack of agreement, most notable being the uncertainty in the determination of the reference unstrained lattice spacing d_0 that serves as the basis for strain determination; the possibility of deviation of the reported single peak diffraction strain from the macroscopic strain values needed

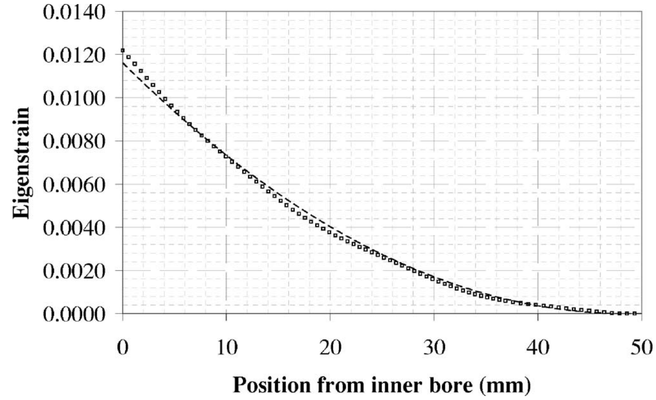


Fig. 7 Eigenstrain profile in specimen *B*: the distribution determined by variational eigenstrain analysis (markers) and a parabolic fit (dashed curve)

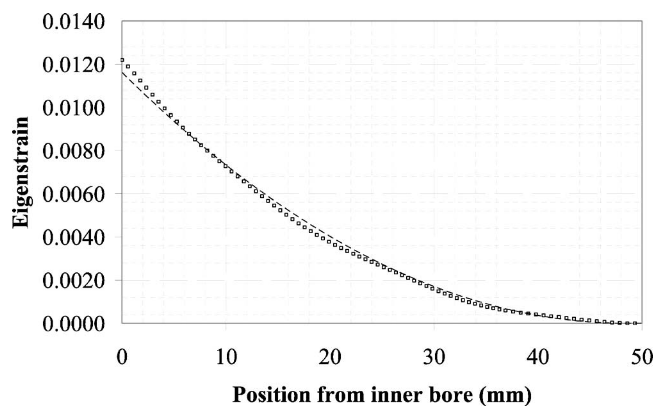


Fig. 8 Eigenstrain profile in specimen *C*: the distribution determined by variational eigenstrain analysis (markers) and a parabolic fit (dashed curve)

for continuum deformation analysis, and the possibility of large grain effects on diffraction strain scatter. It is important to note that the procedure described here is guaranteed to provide the best possible fit to the experimental data within the functional space introduced by Eq. (3).

The solution to the problem is obtained in terms of coefficients c_i, that in turn determine the eigenstrain distributions shown in Figs. 7 and 8. It is apparent that the eigenstrain distributions decay to zero at the radial positions c of 46 mm and 50 mm, respectively. This fact can be established using the generic polynomial functional basis, e.g., using Chebyshev polynomials. Once the boundary of the eigenstrain distribution zone (equivalent to the plastic zone boundary) has been found in this way, the quality and rate of convergence to the solution can be improved. This can be achieved by amending the choice of the basis functions so as to reflect the nature of the eigenstrain variation. For example, using a quadratic multiplier $(r-c)^2$ in the expression

$$\xi_i(r) = \begin{cases} (r-c)^2 \sum_{i=0}^{N} c_i \xi_i(r), & r \leqslant c \\ 0, & r \geqslant c \end{cases} \tag{15}$$

ensures smooth decay of the eigenstrain distribution to zero at the boundary $r=c$.

An interesting observation on the nature of the eigenstrain distributions arising in autofrettage can be made on the basis of Figs. 7 and 8, where the variation of eigenstrain with the radial position from the inner bore is presented for specimens B and C, respectively. The markers indicate the eigenstrain values obtained from the variational eigenstrain determination procedure, while dashed curves show parabolic fits to these values. It is apparent that in both cases the eigenstrain, varies approximately proportionally to the square of the radial distance from the elasto–plastic boundary.[1] Thus it appears that the choice of $N=1$ in Eq. (15) is sufficient to obtain an adequate approximation to the measured data. Experimentation with the number of terms in Eq. (15) showed that increasing the number of coefficients to about $N=7$ or 8 leads to insignificantly small improvements in the quality of fit, but increasing the number of terms to beyond 10 leads to nonphysical oscillatory behavior of the solution.

[1]In the case of elastic–ideally plastic deformation closer analysis of autofrettage in thick tubes shows that plastic strain distribution has the form $\varepsilon^*(r)=A-D/r^2$. This distribution, however, may also be closely approximated by a parabolic distribution.

This result suggests that the elasto–plastic problem solution can be expressed very simply in terms of plastic strain: the process of progressive yielding of the tube under increasing pressure can be thought of as the progressive and self-similar development of eigenstrain distribution. It also suggests that a simple formulation can be developed for this problem in the form of an elastic problem with a perturbation term in the form of incompatibility due to plastic strain. This subject is being studied further by the present authors.

A final remark is due on the observation by Venter et al. [10] that the measured distributions indicate the presence of reverse yielding and Bauschinger effect in the tube material. In the present study no evidence of such reverse yielding was found. This observation is substantiated by the simple and smooth form of the eigenstrain distributions in Figs. 7 and 8. The behavior of residual elastic strain profiles attributed to reverse yielding may in fact be associated with the finite and small thickness of the annular slices studied here: no such behavior was found in the long tube model of autofrettage when the same eigenstrain distributions found from variational analysis were introduced into such a model.

References

[1] Korsunsky, A. M., 2005, “On the Modelling of Residual Stresses due to Surface Peening Using Eigenstrain Distributions,” J. Strain Anal. Eng. Des., **40**(8), pp. 817–824.

[2] Korsunsky, A. M., Regino, G. M., Nowell, D., and Liu, J., 2006, “Eigenstrain Analysis of Synchrotron X-Ray Diffraction Measurement of Residual Strains in Machined Nickel Alloy Plates,” J. Strain Anal. Eng. Des., **41**, pp. 381–395.

[3] Hill, R., 1950, *The Mathematical Theory of Plasticity*, Clarendon, Oxford, UK.

[4] Lamé, G., 1852, *Leçons Sur la Théorie de L'Élasticité*, Gauthier-Villars, Paris, France.

[5] Soutas-Little, R. W., 1973, *Elasticity*, Dover, New York.

[6] Den Hartog, J. P., 1949, *Strength of Materials*, Dover, New York.

[7] Gao, X.-L., 1993, “An Exact Elasto Plastic Solution for a Closed-End Thick-Walled Cylinder of Elastic Linear-Hardening Material With Large Strains,” Int. J. Pressure Vessels Piping, **56**, pp. 331–350.

[8] Gao, X.-L., 1992, “An Exact Elasto-Plastic Solution for an Open-Ended Thick-Walled Cylinder of a Strain-Hardening Material,” Int. J. Pressure Vessels Piping, **52**, pp. 129–144.

[9] Gao, X.-L., 2003, “Elasto-Plastic Analysis of an Internally Pressurized Thick-Walled Cylinder Using a Strain Gradient Plasticity Theory,” Int. J. Solids Struct., **40**, pp. 6445–6455.

[10] Venter, A. M., de Swardt, R. R., and Kyriacou, S., 2000, “Comparative Measurements on Autofrettaged Cylinders with Large Bauschinger Reverse Yielding Zones,” J. Strain Anal. Eng. Des., **35**, pp. 459–469.

[11] Sachs, G., 1927, “Der Nachweis Inner Spannungen in Stangen und Rohren,” Z. Metallkd., **72**, pp. 352–357.

[12] Cheng, W., and Finnie, I., 1986, “Measurement of Residual Hoop Stresses in Cylinders Using the Compliance Method,” ASME J. Eng. Mater. Technol., **108**, pp. 87–92.

Micromechanics of a Compressed Fiber Mass

Mårten Alkhagen

Staffan Toll
e-mail: staffan.toll@me.chalmers.se

Department of Applied Mechanics,
Chalmers University of Technology,
SE-412 96 Gothenburg, Sweden
and
School of Textiles,
University College of Borås,
SE-50190 Borås, Sweden

A theory is presented for the rate modeling of flexible granular solids based on affine average motion of interparticle contacts. We allow contacts to form and break continually but assume the existence of a finite friction coefficient rendering contacts force free as they form or break. The resulting constitutive equations are of the hypoelastic type. A specific model for the deformation of a fiber mass is then developed. The model improves on previous theories for fiber masses in at least two respects: First, it is more general in that it is not restricted to uniaxial compression, although it is restricted to predominantly compressive deformations histories, due to neglect of frictional dissipation. Second, by allowing torsion as well as bending of fibers, this theory covers a larger deformation range. Compression experiments are performed on carded slivers of PA6 fibers under various conditions. The measured response is found to be in close agreement with that predicted by the model. [DOI: 10.1115/1.2711223]

Keywords: fiber networks, granular materials, fiber packing, fiber orientation, fiber contact, porous materials

Introduction

The mechanical properties of a fiber mass are important in many fields of engineering such as paper making and composite manufacturing, and lot of effort has been spent to describe and model these properties. Schofield [1] described the compressive behavior of fiber masses with three-dimensional (3D) random orientation, e.g., cotton and wool, and he proposed an empirical equation where the force response is proportional to the volume fraction cubed. Van Wyk [2] pioneered the mechanistic analysis of the compressibility of 3D random fiber masses. Van Wyk regarded the fiber mass as a system of bending units consisting of fiber beam elements between adjacent fiber contacts and he ignored twisting, slip, and extension of the fibers. His key assumptions were that the mean contact spacing is proportional to the reciprocal of the fiber volume fraction and that the segments deform as in bending of straight slender beams. The result of his analysis is a simple power law for the compressive stress as

$$P = ke(\Phi^3 - \Phi_0^3) \tag{1}$$

where k is a structure dependent constant; e is the Young's modulus of the fibers; and Φ_0 is the limiting fiber volume fraction below which $P=0$. Van Wyk found that his analysis overestimated the value of k by two orders of magnitude compared to experimental results. He also found that Eq. (1) only holds for moderate values of $\Phi(<10\%)$. Most of the work following van Wyk has accepted the form (Eq. (1)) and focused on finding an appropriate expression for k by extending the theories describing the structure development during deformation. Because the earlier theories were based on simple regular geometries they cannot be applied to real structures. Corte and Kallmes [3] suggested a statistical geometry for a two-dimensional fiber mass, e.g., paper. By extending the two-dimensional theory, Komori and Makishima [4] derived the number of contact points for a three-dimensional fiber mass assuming infinite fiber length but without the limitation of randomness. Lee and Lee [5] used this to derive a model that allows for structure development. By incorporating slip at contacts Pan and Carnaby [6] could to some extent explain the compression hysteresis exhibited by fiber masses. Komori and Itoh [7] modified the beam element so that the bending length depended on the beam element orientation rather than being a universal mean length. From this modified fiber contact theory and by using curved beams instead of straight ones, Komori et al. [8] derived an equation describing uniaxial compression of a random fiber mass. Neckář [9] derived an expression describing bidimensional deformation of transversely isotropic fiber masses. Toll [10] suggested a more general power law

$$P = ke(\Phi^n - \Phi_0^n) \tag{2}$$

where the exponent, n, was shown to take the value of 3 for the 3D and 5 for the 2D random cases. By adjusting the exponent, Eq. (2) was shown to fit experimental data for materials where the contacts between fibers are lines rather than points, e.g., fiber bundles. It should be mentioned that considerable effort has been put into deriving the correct number of contact points in a fiber mass by introducing concepts of forbidden volume or hindrance factors [11–13]. Nevertheless, since the assumptions used in those analyses seem unsubstantiated we choose here to use the idealized model where the number of contacts per fiber is directly proportional to the fiber volume fraction [14].

The above theories are more or less limited to uniaxial or isotropic compression. Although there have been some attempts to describe the shear response of fiber masses, e.g., Komori and Itoh [11], there seems to be no model for general deformations, e.g., combinations of shear and compression. A parallel field of research is granular solids. Although fiber masses belong to this category, in the sense that they consist of solid particles interacting by contact forces, the two fields have developed quite independently. Slender particles are special in two important ways: (1) they form a network already at very low packing densities, typically as soon as the average particle is engaged in a few contacts; and (2) their deformation is due to bending and torsion of particles rather than the local deformation close to the contact zones. Fiber masses are thus generally much more compressible than low aspect ratio granular solids. Compaction of a fiber mass generates an ever-increasing amount of contact points, and the topology of the assembly evolves in the process of deformation by contacts continually forming and breaking. For this reason their response is strongly nonlinear and highly strain path dependent. As the topology of the network keeps changing, the local stiffnesses governing the rate of loading of each contact point change as well. This

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received May 31, 2005; final manuscript received October 23, 2006. Review conducted by Ellen M. Arruda.

makes the statistical modeling of the contact force distribution at a given instant very difficult and would require following the entire strain history. The statistical modeling of the contact force *rate* is much more feasible as this can be done entirely based on the current state variables, e.g., the current volume fraction, orientation distribution, velocity gradient, etc.

This paper therefore proposes a framework for the rate modeling of flexible granular solids. The theory is based on affine average motion of the contact points. Contacts are allowed to form and break continually, but we assume the existence of a finite friction coefficient rendering contacts force free as they form or break. The resulting constitutive equations turn out to be of the hypoelastic type with internal variables (structure tensors), i.e., the stress is not obtained from an energy function.

A specific model is then developed for the deformation of a fiber mass. The work improves on previous theories for fiber masses in at least two respects: First, the present theory is more general in that it is not restricted to uniaxial compression, although it is restricted to predominantly compressive deformations histories, due to neglect of frictional dissipation. Second, by including multiple deformation mechanisms, this theory covers a larger deformation range. By allowing torsion as well as bending of fibers we succeed in modeling the fiber mass to a higher volume fraction compared to previous results. We finally specialize to uniaxial compression and compare the results to compressive experiments on carded polyamide-6 fibers of different diameters. The compression experiments were performed in a novel triaxial rheometer for soft compressible solids, developed in-house [15]. The rheometer is particularly suitable for materials having a large characteristic length, such as fiber masses, in that edge effects at the perimeter of the sample are effectively eliminated.

Granular Solids

A granular solid consists of solid particles which interact by mechanical contact; all stress being transmitted across contact surfaces from one particle to another. Since the contact areas are usually small compared to particle dimensions, the contact surface traction $\mathbf{n}\cdot\boldsymbol{\sigma}$ is approximately concentrated to a contact *point*

$$\mathbf{n}\cdot\boldsymbol{\sigma} = \mathbf{p}\,\delta(\mathbf{x}-\mathbf{r}) \tag{3}$$

where $\mathbf{n}$ is the unit normal of a contact point; $\boldsymbol{\sigma}$ is the contact stress; $\mathbf{r}$ is the position; and $\mathbf{p}$ is the contact force. The Dirac delta function, δ, has the properties, $\delta(t)=0$ for $t\neq 0$, $\lim_{t\to 0}\delta(t)=\infty$ and $\int_0^\infty \delta(t)dt=1$. Each particle is thus subjected to a set of discrete forces exerted on its surface by its contacting neighbors and the equilibrium conditions are

$$\sum \mathbf{p} = \mathbf{0}$$

and

$$\sum \mathbf{r}\times\mathbf{p} = \mathbf{0} \tag{4}$$

where the sums are taken over all contacts of a particle. The continuum description of a granular solid presumes a representative volume, V, small enough that the macroscopic velocity gradient is constant (within some acceptable error) on the scale of V. Thus placing an origin somewhere inside V, the macroscopic velocity field within V may be written

$$\mathbf{V}(\mathbf{x}) = \mathbf{V}(\mathbf{0}) + \mathbf{L}\cdot\mathbf{x} \tag{5}$$

where $\mathbf{L}=(\boldsymbol{\nabla}\mathbf{V})^t$ is the spatial velocity gradient. In the quasistatic conditions given by Eq. (4), as has been shown by many authors, e.g., Love [16], the average stress $\boldsymbol{\Sigma}$ (macroscopic Cauchy stress) is

$$\boldsymbol{\Sigma} = \frac{1}{V}\,\mathrm{sym}\sum \mathbf{pr} \tag{6}$$

where the summation is carried out over all contact points on all particles in V. Notice that each contact gives rise to two contact points, one on each particle, with associated contact force and normal $(\mathbf{p},\mathbf{n})$ and $(-\mathbf{p},-\mathbf{n})$, respectively. It also should be pointed out that the contact vector $\mathbf{r}$ need not refer a common origin, but may be defined locally for each particle.

Most theories take the view that the contact mechanics are primarily controlled by the orientation of the contact plane, and express the stress in terms of the distribution of the normal vector $\mathbf{n}$

$$\boldsymbol{\Sigma} = N\,\mathrm{sym}\oint \overline{\mathbf{pr}}(\mathbf{n})\zeta(\mathbf{n})d\mathbf{n} \tag{7}$$

where N is the number of contacts per unit volume; $\zeta(\mathbf{n})$ is the probability density of $\mathbf{n}$; and the overbar denotes the average over all contact points of orientation $\mathbf{n}$. One usually makes the critical assumption [17] that $\mathbf{p}$ and $\mathbf{r}$ are uncorrelated (for $\mathbf{n}$ fixed), allowing the separation

$$\overline{\mathbf{pr}}(\mathbf{n}) = \overline{\mathbf{p}}(\mathbf{n})\overline{\mathbf{r}}(\mathbf{n}) \tag{8}$$

whereupon the problem reduces to modeling the expectations of contact force and contact vector at a fixed unit normal, $\overline{\mathbf{p}}=\overline{\mathbf{p}}(\mathbf{n})$ and $\overline{\mathbf{r}}=\overline{\mathbf{r}}(\mathbf{n})$. Typically, this is done by linking $\overline{\mathbf{r}}(\mathbf{n})$ directly to macrostrain or $\overline{\mathbf{p}}(\mathbf{n})$ to macrostress [18], and applying some nonlinear contact law, $f(\overline{\mathbf{r}},\overline{\mathbf{p}},\mathbf{n})=0$, at the contact points, such as Emeriault and Cambou [18] or Norris and Johnson [19], or even a linear relation, e.g., Alzebdeh and Ostoja-Starzewski [20]. Norris and Johnson [19] also discuss the fundamental strain path dependence of granular solids. Path dependence may obviously result from frictional slip but, curiously, it may also occur without dissipation. Indeed if a granular solid is loaded to a given strain along a strain path which is consistently dominated by compression, the strain energy will depend on the strain path, yet unloading by retracing the same path, may still return all the energy. This type of response is essentially *hypoelastic* [21] (although Norris and Johnsson call it elastic) and hence requires a rate theory.

A Rate Theory for Granular Solids

Consider a contact between two particles, one of which is considered as *test particle*. The dynamics of the contact will depend on the entire geometric configuration in a volume enclosing the two contacting particles and several others. Since the complete configuration cannot be specified the contact configuration must be described by a limited number of *contact variables*, such as the contact normal $\mathbf{n}$, and the contact force $\mathbf{p}$ will be a stochastic function of those variables. The contact variables should include at least the contact vector $\mathbf{r}$, defined as the position of the contact point with respect to an origin (e.g., the centroid) of the test particle, and the contact normal $\mathbf{n}$, defined as a unit vector along the outward normal to the test particle surface at the contact point. Naturally, neither of these need be included explicitly, provided that they are uniquely defined by the contact variables used.

Now make the following assumptions:

1. The contact force vanishes as the normal force goes to zero; and
2. The expectation of the contact displacement rate $\dot{\mathbf{r}}$ for a given set of values of the contact variables is affine

$$\overline{\dot{\mathbf{r}}} = \overline{\dot{\mathbf{r}}}(\mathbf{r},\mathbf{L}) = \mathbf{L}\cdot\mathbf{r} \tag{9}$$

where the overbar denotes the conditional expectation given the contact variables $(\mathbf{r},\mathbf{n},\cdots)$.

The first assumption implies ideal friction in the sense that the contact surfaces are smooth and adhesion free. The second assumption is a weaker, and more realistic, form of the common affine assumption, $\dot{\mathbf{r}}=\mathbf{L}\cdot\mathbf{r}$. Next we write Eq. (6) as

$$\boldsymbol{\Sigma} = N\,\mathrm{sym}\langle\bar{\mathbf{p}}\mathbf{r}\rangle_c \tag{10}$$

where N is the number of contacts per unit volume (defined such that each contact point is counted twice; once for each particle involved) and the angle brackets $\langle\cdot\rangle_c$ denote an average over all possible values of the contact variables. Notice here that since $\overline{\mathbf{pr}}=\bar{\mathbf{p}}\mathbf{r}$, not by assumption but because $\mathbf{r}$ is one of the contact variables, the assumption in Eq. (8) is not needed. Presently, with the restricted set of contact variables $(\mathbf{r},\mathbf{n},\cdots)$, the average is

$$\langle\cdot\rangle_c = \oint\oint(\cdot)\zeta(\mathbf{r},\mathbf{n},\cdots)d\mathbf{r}\,d\mathbf{n}\cdots \tag{11}$$

$\zeta(\mathbf{r},\mathbf{n},\cdots)$ being the probability density of the configuration $(\mathbf{r},\mathbf{n},\cdots)$, subject to the normalization condition

$$\oint\oint\zeta(\mathbf{r},\mathbf{n},\cdots)d\mathbf{r}\,d\mathbf{n}\cdots = 1 \tag{12}$$

Now time differentiation of Eq. (6) yields

$$\frac{d\boldsymbol{\Sigma}}{dt} = \mathrm{sym}\left[-\frac{1}{V^2}\frac{dV}{dt}\sum\mathbf{pr}+\frac{1}{V}\frac{d}{dt}\left(\sum\mathbf{pr}\right)\right] \tag{13}$$

The first term on the right-hand side, in view of Eq. (6), is just $(\boldsymbol{\delta}:\mathbf{L})\boldsymbol{\Sigma}$. Due to the assumed frictional nature of the contact forces (Assumption 1), contacts in the process of being formed or broken must be force free, and hence the differential operator in the second right-hand side term may be moved inside the summation sign. Finally Assumption 2 implies that $\overline{\mathbf{p}\dot{\mathbf{r}}}=\bar{\mathbf{p}}\mathbf{r}\cdot\mathbf{L}^t$. The general affine rate equation thus takes the general form

$$\dot{\boldsymbol{\Sigma}}+(\boldsymbol{\delta}:\mathbf{L})\boldsymbol{\Sigma} = N\,\mathrm{sym}(\langle\bar{\mathbf{p}}\mathbf{r}\rangle_c\cdot\mathbf{L}^t+\langle\bar{\dot{\mathbf{p}}}\mathbf{r}\rangle_c) \tag{14}$$

As $\bar{\dot{\mathbf{p}}}$ is the average of $\dot{\mathbf{p}}$ over *all* contact points with a given configuration $(\mathbf{r},\mathbf{n},\cdots)$, it may only depend on $\mathbf{r},\mathbf{n},\cdots$ and macroscopic field variables such as $\mathbf{L}$

$$\bar{\dot{\mathbf{p}}} = \boldsymbol{\varphi}(\mathbf{r},\mathbf{n},\cdots;\mathbf{L},\cdots) \tag{15}$$

The force rate function $\boldsymbol{\varphi}$ constitutes the kernel of the constitutive equation, Eq. (14). This function must be a three-dimensional vector and have the dimension of force per unit time.

We must now ensure that the function $\boldsymbol{\varphi}(\mathbf{r},\mathbf{n},\cdots;\mathbf{L},\cdots)$ is indifferent to an arbitrary time dependent rotation (change of the observer's motion) $\mathbf{Q}^t=\mathbf{Q}^t(t)$ applied to all objective quantities

$$\mathbf{r}^* = \mathbf{Q}\cdot\mathbf{r} \tag{16}$$

$$\mathbf{p}^* = \mathbf{Q}\cdot\mathbf{p} \tag{17}$$

$$\mathbf{n}^* = \mathbf{Q}\cdot\mathbf{n} \tag{18}$$

$$\dot{\mathbf{p}}^* = \mathbf{Q}\cdot\dot{\mathbf{p}}+\dot{\mathbf{Q}}\cdot\mathbf{p} \tag{19}$$

$$\mathbf{L}^* = \dot{\mathbf{Q}}\cdot\mathbf{Q}^t+\mathbf{Q}\cdot\mathbf{L}\cdot\mathbf{Q}^t \tag{20}$$

where the asterisk denotes a change of reference frame. This frame indifference may be expressed as

$$\bar{\dot{\mathbf{p}}}^* = \bar{\dot{\mathbf{p}}}(\mathbf{r}^*,\mathbf{n}^*,\cdots;\mathbf{L}^*,\cdots) \tag{21}$$

Combining all the above yields the condition,

$$\begin{aligned}&\bar{\dot{\mathbf{p}}}(\mathbf{Q}\cdot\mathbf{r},\mathbf{Q}\cdot\mathbf{n},\cdots;\dot{\mathbf{Q}}\cdot\mathbf{Q}^t+\mathbf{Q}\cdot\mathbf{L}\cdot\mathbf{Q}^t,\cdots)\\&\quad-\mathbf{Q}\cdot\bar{\dot{\mathbf{p}}}(\mathbf{r},\mathbf{n},\cdots;\mathbf{L},\cdots)-\dot{\mathbf{Q}}\cdot\bar{\mathbf{p}}=\mathbf{0}\end{aligned} \tag{22}$$

It may also be verified that the resulting constitutive equation (Eq. (14)) is material frame indifferent if and only if this condition is satisfied.

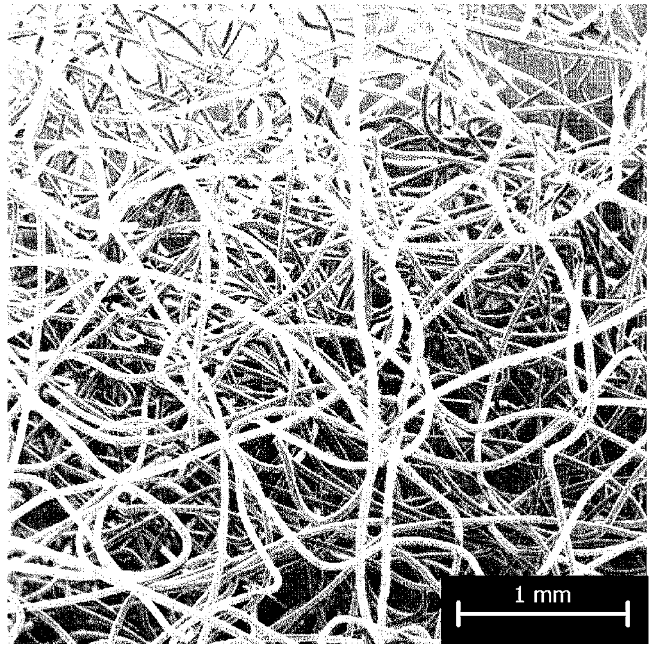


Fig. 1 SEM micrograph of carded polyamide-6 fibers with a diameter of 50 μm

Fiber Mass Model

Having formulated the rate theory, we are in a position to construct a specific model of a fiber mass. Figure 1 shows a scanning electron microscope (SEM) graph of a carded sliver of 50 μm diameter PA6 fibers. In general one would have to treat each fiber as a *particle* (in the sense of the previous section). When the fibers are slender and crooked, such a description would be very complex indeed. Instead we assume that the fiber is divisible into roughly straight segments, (Fig. 2), which can be treated as independent particles. The validity of such a description requires that the contact spacing along a fiber be smaller than the length scale on which the fiber can be considered straight (the *crimp spacing* $2b$ in Fig. 2), so that the stress due to the contact forces on the segment is much larger than that due to the load from the rest of the fiber. This may not seem to be the case judging from Fig. 1, but it should be appreciated that the material in the micrograph is only about 1% volume fraction. At 10% volume fraction the number of contacts between crimps is about eight. Several other objections can be raised against this proposition; perhaps most importantly it presumes independent rotation of segments belonging to the same fiber. Furthermore, slip is considered absent and the degree of alignment is assumed to be moderate.

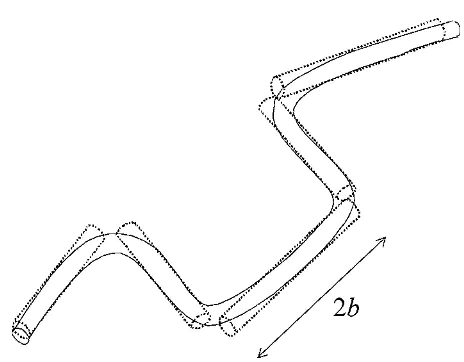


Fig. 2 Segmentation of a fiber where $2b$ is the crimp spacing

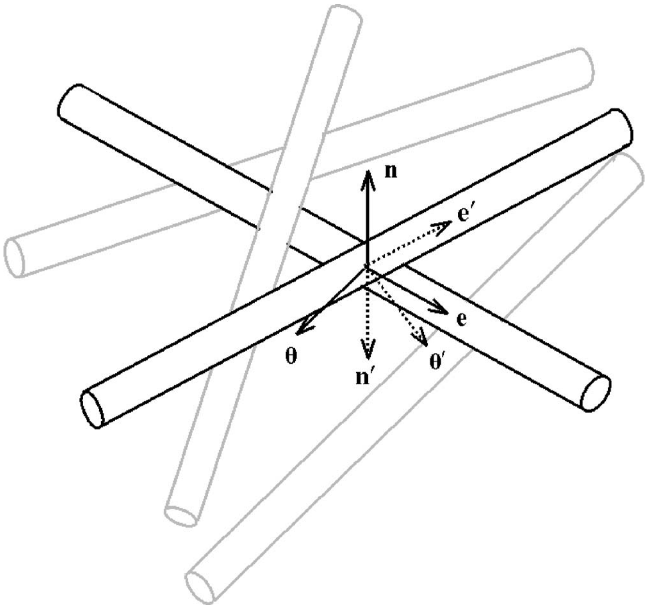


Fig. 3 Schematic of a contact point with the local basis vectors e, n, and $\boldsymbol{\theta}$ indicated. The primed basis vectors refer to the contacting fiber.

Introducing a set of local orthonormal basis vectors for each contact point $\mathbf{n}$, $\mathbf{e}$, and $\boldsymbol{\theta}$ relative to the fiber axis and contact normal (Fig. 3), and resolving $\mathbf{r}$ and $\mathbf{p}$ on this basis yields

$$\mathbf{r} = r_n\mathbf{n} + r_e\mathbf{e} + r_\theta\boldsymbol{\theta} \tag{23}$$

$$\mathbf{p} = p_n\mathbf{n} + p_e\mathbf{e} + p_\theta\boldsymbol{\theta} \tag{24}$$

Our choice of contact variables will be the axis orientations $\mathbf{e}$ and $\mathbf{e}'$ and the position vectors $\mathbf{r}$ and $\mathbf{r}'$ of the two contacting fibers. Notice that $\mathbf{n}$ is not needed as a contact variable, since it is determined by the other ones. The expectation $=(\mathbf{e},\mathbf{e}',\mathbf{r},\mathbf{r}')$ then refers to fixed values of these variables. We now make the simplest choice of convected force rate that satisfies the frame-indifference condition Eq. (22):

$$\overline{\dot{\mathbf{p}}} = \left(\overline{\dot{p}_n}\mathbf{n} + \overline{\dot{p}_e}\mathbf{e} + \overline{\dot{p}_\theta}\boldsymbol{\theta}\right) + \mathbf{L}\cdot\overline{\mathbf{p}} \tag{25}$$

In fact this choice is inconsistent with maintaining orthonormality of the basis vectors, adding an unphysical convected-force contribution to the force rate, which may be considerable in shearing deformations. However, it will offer considerable simplification, and it should work as long as the deformation is predominantly compressive. Introducing this in Eq. (14), we obtain

$$\overset{\Delta}{\boldsymbol{\Sigma}} = N\langle\overline{\dot{p}_n}r_n\mathbf{nn} + \overline{\dot{p}_e}r_e\mathbf{ee} + \overline{\dot{p}_\theta}r_\theta\boldsymbol{\theta\theta}\rangle_c\mathbf{:L} \tag{26}$$

where

$$\overset{\Delta}{\boldsymbol{\Sigma}} = \dot{\boldsymbol{\Sigma}} + (\boldsymbol{\delta}\mathbf{:L})\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\cdot\mathbf{L}^t - \mathbf{L}\cdot\boldsymbol{\Sigma} \tag{27}$$

is an objective stress rate, known as the Truesdell stress rate [22]. In Eq. (26), since the probability density $\zeta(\mathbf{e},\mathbf{e}',\mathbf{r},\mathbf{r}')$ and $\dot{p}_n$, $\dot{p}_e$, $\dot{p}_\theta$ are even functions of $\mathbf{e}$, $\mathbf{e}'$, $\mathbf{n}$, and $\boldsymbol{\theta}$, and r_n, r_e, and r_θ are positive, all odd terms such as $\langle\dot{p}_n r_e\mathbf{ne}\rangle_c$vanish.

Based on the neglect of slip, we now relate the force rates $\dot{p}_n$, $\dot{p}_e$, and $\dot{p}_\theta$ to the principal displacement rates through

$$\overline{\dot{p}_n} = \frac{\overline{\dot{r}_n}}{\langle\overline{s_n}\rangle_c} \tag{28}$$

$$\overline{\dot{p}_e} = \frac{\overline{\dot{r}_e}}{\langle\overline{s_e}\rangle_c} \tag{29}$$

$$\overline{\dot{p}_\theta} = \frac{\overline{\dot{r}_\theta}}{\langle\overline{s_\theta}\rangle_c} \tag{30}$$

where s_n, s_e, and s_θ are compliances of a given contact, defined as

$$s_n = \frac{\partial r_n}{\partial p_n}, \quad s_e = \frac{\partial r_e}{\partial p_e}$$

and

$$s_\theta = \frac{\partial r_\theta}{\partial p_\theta} \tag{31}$$

We thus let the force rates be controlled by the average compliance of all contacts rather than the local compliance of a particular contact. In other words $dr_i/dp_i \approx \langle\overline{s_i}\rangle_c \neq \partial r_i/\partial p_i = s_i$. Notice here the importance of assuming affine expectations of motion $\overline{\dot{\mathbf{r}}} = \mathbf{L}\cdot\mathbf{r}$ (Eq. (9)) rather than affine motion $\dot{\mathbf{r}} = \mathbf{L}\cdot\mathbf{r}$. If we were to make the latter assumption, the freedom to specify Eqs. (28)–(30) would be lost, as these relations would necessarily have to be $\dot{p}_i = \dot{r}_i\overline{s_i}^{-1}$ and the model would become severely overstiff. Moreover the stiffer contacts would contribute excessively to the total average. In the case of a stiffer contact surrounded by more compliant contacts, the latter are likely to take a major part of the deformation.

By inserting Eqs. (28)–(30) into Eq. (26) we have

$$\overset{\Delta}{\boldsymbol{\Sigma}} = N\left(\frac{\langle r_n\overline{\dot{r}_n}\mathbf{nn}\rangle_c}{\langle\overline{s_n}\rangle_c} + \frac{\langle r_e\overline{\dot{r}_e}\mathbf{ee}\rangle_c}{\langle\overline{s_e}\rangle_c} + \frac{\langle r_\theta\overline{\dot{r}_\theta}\boldsymbol{\theta\theta}\rangle_c}{\langle\overline{s_\theta}\rangle_c}\right)\mathbf{:L} \tag{32}$$

The principal displacement rates are now assumed to be equal to the stretch rates in the $\mathbf{n}$, $\mathbf{e}$, and $\boldsymbol{\theta}$ directions, respectively

$$\overline{\dot{r}_n} = r_n\mathbf{nn:L} \tag{33}$$

$$\overline{\dot{r}_e} = r_e\mathbf{ee:L} \tag{34}$$

$$\overline{\dot{r}_\theta} = r_\theta\boldsymbol{\theta\theta}\mathbf{:L} \tag{35}$$

yielding

$$\overset{\Delta}{\boldsymbol{\Sigma}} = N\left[\frac{\langle r_n^2\mathbf{nnnn}\rangle_c}{\langle\overline{s_n}\rangle_c} + \frac{\langle r_e^2\mathbf{eeee}\rangle_c}{\langle\overline{s_e}\rangle_c} + \frac{\langle r_\theta^2\boldsymbol{\theta\theta\theta\theta}\rangle_c}{\langle\overline{s_\theta}\rangle_c}\right]\mathbf{:L} \tag{36}$$

Finally using the assumption that r_n, r_e, and r_θ are all approximately independent of $\mathbf{n}$, $\mathbf{e}$, and $\boldsymbol{\theta}$ (this is exact for random and planar random orientation), one obtains

$$\overset{\Delta}{\boldsymbol{\Sigma}} = N\left[\frac{\langle r_n^2\rangle_c}{\langle\overline{s_n}\rangle_c}\langle\mathbf{nnnn}\rangle_c + \frac{\langle r_e^2\rangle_c}{\langle\overline{s_e}\rangle_c}\langle\mathbf{eeee}\rangle_c + \frac{\langle r_\theta^2\rangle_c}{\langle\overline{s_\theta}\rangle_c}\langle\boldsymbol{\theta\theta\theta\theta}\rangle_c\right]\mathbf{:L} \tag{37}$$

where $\langle\mathbf{nnnn}\rangle_c$, $\langle\mathbf{eeee}\rangle_c$, and $\langle\boldsymbol{\theta\theta\theta\theta}\rangle_c$ are fourth-order structure tensors.

Structure Tensors

There are several possible ways of forming a fiber network, and the network structure may depend on this route. Here the assembly is assumed to be statistically homogeneous, and disperse in the sense that there is no correlation between the spatial and orientational distributions. We adopt the approach of Toll [10] and consider a network formed by random placement of fibers in space, with no restriction of interpenetrating fibers by bending, thus turning each interpenetration into a contact point [10,14]. Thus the contact points in the network of interpenetrating fibers are modeled as the volume intersections in a random network of interpenetrating ones. The benefit of this model is that the probabilities of such volume intersections can be obtained exactly.

Now consider a fiber population having the overall probability density $\psi(\mathbf{e})$ of the fiber axis orientation vector $\mathbf{e}$. Since the probability density of intersection of two fibers of orientation $\mathbf{e}$ and $\mathbf{e}'$ must be proportional to $|\mathbf{e}\times\mathbf{e}'|\psi(\mathbf{e})\psi(\mathbf{e}')$ and independent of $\mathbf{r}$ and $\mathbf{r}'$ it immediately follows from the normalization condition, Eq. (12), that

$$\oint\oint \zeta(\mathbf{e},\mathbf{e}',\mathbf{r},\mathbf{r}')d\mathbf{r}\,d\mathbf{r}' = \frac{|\mathbf{e}\times\mathbf{e}'|\psi(\mathbf{e})\psi(\mathbf{e}')}{\oint\oint |\mathbf{e}\times\mathbf{e}'|\psi(\mathbf{e})\psi(\mathbf{e}')d\mathbf{e}\,d\mathbf{e}'} \tag{38}$$

Hence the average over all contact configurations of a quantity $q(\mathbf{e},\mathbf{e}')$ that is independent of $\mathbf{r}$ and $\mathbf{r}'$ may be written as

$$\begin{aligned}\langle q\rangle_c &= \oint\oint\oint\oint q(\mathbf{e},\mathbf{e}')\zeta(\mathbf{e},\mathbf{e}',\mathbf{r},\mathbf{r}')d\mathbf{e}\,d\mathbf{e}'\,d\mathbf{r}\,d\mathbf{r}' \\ &= \frac{1}{f}\oint\oint q(\mathbf{e},\mathbf{e}')|\mathbf{e}\times\mathbf{e}'|\psi(\mathbf{e})\psi(\mathbf{e}')d\mathbf{e}\,d\mathbf{e}'\end{aligned} \tag{39}$$

where

$$f = \oint\oint |\mathbf{e}\times\mathbf{e}'|\psi(\mathbf{e})\psi(\mathbf{e}')d\mathbf{e}\,d\mathbf{e}' \tag{40}$$

In particular, for 3D random orientation $f^{3\mathrm{D}}=\pi/4$ and for 2D random orientation $f^{2\mathrm{D}}=2/\pi$. Now the average of $\mathbf{eeee}$ taken over all contacts is

$$\langle\mathbf{eeee}\rangle_c = \frac{1}{f}\oint\oint \mathbf{eeee}|\mathbf{e}\times\mathbf{e}'|\psi(\mathbf{e})\psi(\mathbf{e}')d\mathbf{e}\,d\mathbf{e}' \tag{41}$$

This tensor should not be confused with the so-called fourth-order orientation tensor; $\langle\mathbf{eeee}\rangle_c$ is averaged over contact points, whereas orientation tensors are averaged over fiber orientations. Similarly the fourth order average of the contact normal $\mathbf{n}$ is

$$\langle\mathbf{nnnn}\rangle_c = \frac{1}{f}\oint\oint \mathbf{nnnn}|\mathbf{e}\times\mathbf{e}'|\psi(\mathbf{e})\psi(\mathbf{e}')d\mathbf{e}\,d\mathbf{e}' \tag{42}$$

where $\mathbf{n}$ is given by

$$\mathbf{n} = \pm\frac{\mathbf{e}\times\mathbf{e}'}{|\mathbf{e}\times\mathbf{e}'|} \tag{43}$$

The fourth-order average of transverse tangential vector $\boldsymbol{\theta}$ is

$$\langle\boldsymbol{\theta\theta\theta\theta}\rangle_c = \frac{1}{f}\oint\oint \boldsymbol{\theta\theta\theta\theta}|\mathbf{e}\times\mathbf{e}'|\psi(\mathbf{e})\psi(\mathbf{e}')d\mathbf{e}\,d\mathbf{e}' \tag{44}$$

where $\boldsymbol{\theta}$ is given by

$$\boldsymbol{\theta} = \pm\frac{\mathbf{e}\times\mathbf{n}}{|\mathbf{e}\times\mathbf{n}|} \tag{45}$$

Notice that all these averages are based on the orientation distribution function, $\psi(\mathbf{e})$. Hence provided that $\psi(\mathbf{e})$ is obtainable in some form, e.g., discretized or as an approximate function, it will be straightforward to compute the structure tensors $\langle\mathbf{nnnn}\rangle_c$, $\langle\mathbf{eeee}\rangle_c$, and $\langle\boldsymbol{\theta\theta\theta\theta}\rangle_c$.

The orientation distribution will be modeled here by assuming affine rotation of the fiber axes

$$\dot{\mathbf{e}} = \mathbf{L}\cdot\mathbf{e} - \mathbf{eee}:\mathbf{L} \tag{46}$$

For a given reference configuration, this may be written in terms of the deformation gradient, $\mathbf{F}=\partial\mathbf{X}/\partial\mathbf{X}_r$

$$\mathbf{e} = \frac{\mathbf{F}\cdot\mathbf{e}_r}{\sqrt{\mathbf{e}_r\cdot\mathbf{F}^t\cdot\mathbf{F}\cdot\mathbf{e}_r}} \tag{47}$$

where the index r refers to the reference configuration. We will always start our computations from a state of very small packing fraction and isotropic orientation, assuming that the real fiber mass structures can be obtained in this way.

Contact Compliances

To estimate the compliances defined by Eq. (31) we identify two sources of contact force: fiber bending and fiber torsion. The normal compliance s_n is governed by bending of the test fiber (unprimed fiber), which is described as a generic Euler beam

$$\bar{s}_n = \overline{\frac{\partial r_n}{\partial p_n}} = \frac{\overline{\lambda^3}}{6\pi k_b e a^4} \tag{48}$$

where e is the fiber Young's modulus; a is the fiber radius; and λ is the contact spacing, i.e. the fiber length between two adjacent contacts. The geometric constant k_b is unity if the beam is loaded at its midsection and fixed at its end sections. The transverse compliance is governed by torsion and bending of the test fiber

$$\bar{s}_\theta = \overline{\frac{\partial r_\theta}{\partial p_\theta}} = \frac{\overline{16\lambda}}{3\pi k_t e a^2} + \frac{\overline{\lambda^3}}{6\pi k_b e a^4} \tag{49}$$

where the geometric constant k_t is unity for a simple torsion bar loaded by a couple ap_θ at its midsection and fixed at its end sections. The axial compliance is governed by torsion and bending of the *contacting* fiber

$$\bar{s}_e = \overline{\frac{\partial r_e}{\partial p_e}} = \frac{\overline{16\lambda'}}{3\pi k_t e' a'^2} + \frac{\overline{\lambda'^3}}{6\pi k_b e' a'^4} \tag{50}$$

where the primed quantities refer to the contacting fiber. In all the three above equations, it is only the contact spacing, λ, that is not a contact variable or a constant.

The magnitude of the elements of the position vector is assumed to be

$$r_n \sim r_\theta \sim a \tag{51}$$

and, assuming a random distribution of contacts along a fiber

$$r_e \sim xb \tag{52}$$

where b is half the crimp spacing (Fig. 2) and x is a stochastic variable randomly distributed in the interval $0\leqslant x\leqslant 1$. In order to allow for a fiber size dependence the crimp spacing will be assumed to be directly proportional to the fiber size a

$$b = \beta a \tag{53}$$

where β will be called the crimp ratio. Due to the random distribution of contacts, the distribution of the contact spacing λ is exponential

$$f(\lambda) = \frac{1}{\bar{\lambda}}e^{-\frac{\lambda}{\bar{\lambda}}} \tag{54}$$

and the third moment of λ is

$$\overline{\lambda^3} = \bar{\lambda}^{-1}\int_0^\infty \lambda^3 e^{-\frac{\lambda}{\bar{\lambda}}}d\lambda = 6\bar{\lambda}^3 \tag{55}$$

Now, the compliance related averages in Eq. (37) are

$$\frac{\langle r_n^2\rangle_c}{\langle\bar{s}_n\rangle_c} = 6\pi k_b\frac{\langle a^2\rangle_c}{\langle e^{-1}a^{-4}\overline{\lambda^3}\rangle_c} = \pi k_b\frac{\langle a^2\rangle_c}{\langle e^{-1}a^{-4}\bar{\lambda}^3\rangle_c} \tag{56}$$

$$\frac{\langle r_\theta^2\rangle_c}{\langle\bar{s}_\theta\rangle_c} = \frac{\langle a^2\rangle_c}{\frac{16}{3\pi}k_t^{-1}\langle e^{-1}a^{-2}\bar{\lambda}\rangle_c + \frac{1}{\pi}k_b^{-1}\langle e^{-1}a^{-4}\bar{\lambda}^3\rangle_c} \tag{57}$$

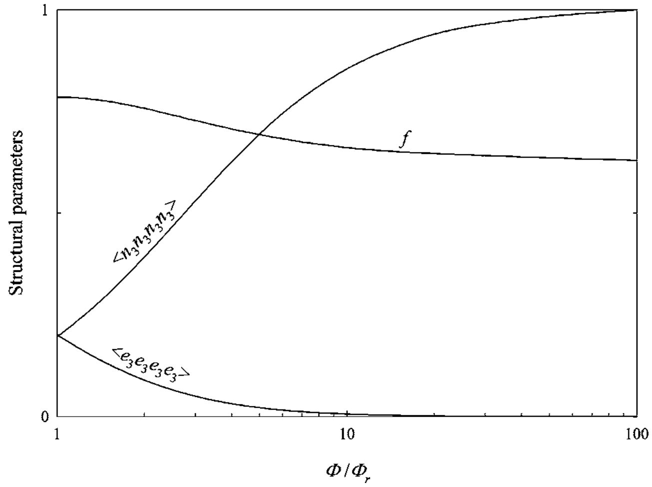


Fig. 4 Evolution of some structural parameters for an initially 3D random fiber mass during uniaxial compression

$$\frac{\langle r_e^2\rangle_c}{\langle \bar{s}_e\rangle_c} = \frac{\beta^2\langle a^2\rangle_c}{\frac{16}{\pi}k_t^{-1}\langle e^{-1}a^{-2}\bar{\lambda}\rangle_c + \frac{3}{\pi}k_b^{-1}\langle e^{-1}a^{-4}\bar{\lambda}^3\rangle_c} \tag{58}$$

For constant fiber size a, the expected contact spacing $\bar{\lambda}$, in terms of the fiber volume fraction Φ, is (e.g., Ref. [10])

$$\bar{\lambda}(\mathbf{e}) = \frac{\pi a}{4f(\mathbf{e})\Phi} \tag{59}$$

where $f(\mathbf{e}) = \oint|\mathbf{e}\times\mathbf{e}'|\psi(\mathbf{e}')d\mathbf{e}'$. The number of contacts per unit volume is (e.g., Ref. [10])

$$N = \frac{4f\Phi^2}{\pi^2a^3} \tag{60}$$

where f is defined by Eq. (40). Multiplying Eqs. (56)–(58) with N and assuming that $\langle f(\mathbf{e})^{-1}\rangle_c \approx f^{-1}$ and $\langle f(\mathbf{e})^{-3}\rangle_c \approx f^{-3}$ (this is mostly true, see Fig. 4, and is exact for random, planar random, and unidirectional orientation)

$$N\frac{\langle r_n^2\rangle_c}{\langle \bar{s}_n\rangle_c} = \frac{256k_be}{\pi^4}f^4\Phi^5 \tag{61}$$

$$N\frac{\langle r_\theta^2\rangle_c}{\langle \bar{s}_\theta\rangle_c} = \frac{e}{\left(\frac{3}{\pi^2}k_tf^2\Phi^3\right)^{-1} + \left(\frac{256}{\pi^4}k_bf^4\Phi^5\right)^{-1}} \tag{62}$$

$$N\frac{\langle r_e^2\rangle_c}{\langle \bar{s}_e\rangle_c} = \frac{e\beta^2}{\left(\frac{1}{\pi^2}k_tf^2\Phi^3\right)^{-1} + \left(\frac{256}{3\pi^4}k_bf^4\Phi^5\right)^{-1}} \tag{63}$$

Clearly, provided that $\beta^2 \gg 1$ and $\langle\mathbf{eeee}\rangle_c \sim \langle\boldsymbol{\theta\theta\theta\theta}\rangle_c$, the tangential component may be neglected. Hence the final result is

$$\overset{\Delta}{\mathbf{\Sigma}} = \left[\frac{256k_be}{\pi^4}f^4\Phi^5\langle\mathbf{nnnn}\rangle_c + \frac{e\beta^2}{\left(\frac{1}{\pi^2}k_tf^2\Phi^3\right)^{-1} + \left(\frac{256}{3\pi^4}k_bf^4\Phi^5\right)^{-1}}\langle\mathbf{eeee}\rangle_c\right]:\mathbf{L} \tag{64}$$

where $\overset{\Delta}{\mathbf{\Sigma}}$ is given by Eq. (27). Perhaps the greatest uncertainty of this model is in the loading conditions of the beam elements, therefore the constants, k_b and k_t, will be used rather as correction parameters and will be determined experimentally.

Computation for Uniaxial Compression

Equation (64) can be applied to various kinds of fiber mass and deformation. Here, however, we will focus on uniaxial compression, because we suspect that violating the orthonormality of the basis vectors (Eq. (25)) is unrealistic in shearing deformations but acceptable in compressive deformations. In the uniaxial case, only one component of each structure tensor is needed: $\langle n_3n_3n_3n_3\rangle_c$ and $\langle e_3e_3e_3e_3\rangle_c$. Moreover, our parallel-plate instrument allows us to measure the true stress response in uniaxial compression.

For uniaxial compression in the 3-direction, the macroscopic velocity gradient may be expressed in terms of the rate of change of the fiber volume fraction

$$L_{33} = -\frac{\dot{\Phi}}{\Phi} \tag{65}$$

To examine the resulting stress component $\Sigma_{33} = -P$ we specialize Eq. (64) accordingly

$$\frac{dP}{d\Phi} + \frac{P}{\Phi} = ef^4\Phi^4\left[\frac{256k_b}{\pi^4}\langle n_3n_3n_3n_3\rangle_c + \frac{\beta^2\langle e_3e_3e_3e_3\rangle_c}{\left(\frac{1}{\pi^2}k_tf^{-2}\Phi^{-2}\right)^{-1} + \left(\frac{256}{3\pi^4}k_b\right)^{-1}}\right] \tag{66}$$

In the special case of a planar fiber mass, we have $\langle n_3n_3n_3n_3\rangle_c = 1$ and $\langle e_3e_3e_3e_3\rangle_c = 0$, which yields

$$P = \frac{128}{3\pi^4}k_bef^4\Phi^5 \tag{67}$$

This result coincides with an earlier result for planar fiber masses [23] with $k_b \approx 2.4$.

A nonrandom initial orientation is modeled by a fictitious initial deformation $\mathbf{F}_0$ relative to a reference configuration $\mathbf{F}_r = \boldsymbol{\delta}$ in which the orientation is taken to be random ($\Psi_r = \pi/4$). In this way the initial orientation distribution Ψ_0 is simply controlled by the choice of initial deformation $\mathbf{F}_0$ or, equivalently, by the choice of reference configuration. For the uniaxial compression of a fiber mass with random in-plane orientation we thus have

$$F_{11} = F_{22} = 1$$

and

$$F_{33} = \frac{\Phi_r}{\Phi} \tag{68}$$

At any given deformation, $\mathbf{F}$, the orientation distribution is obtained by applying Eq. (47) to each fiber in a reference set of 3000 fibers, having random orientation. The structure tensors are then numerically computed from this set using Eqs. (41) and (42). In Fig. 4 the computed values for $\langle n_3n_3n_3n_3\rangle_c$, $\langle e_3e_3e_3e_3\rangle_c$ and f, starting from a 3D random fiber mass, are plotted versus the degree of compression, in this case indicated as Φ/Φ_r. As expected, when $\Phi/\Phi_r = 1$ and $\Phi/\Phi_r \to \infty$ the computed orientation is 3D random and 2D random, respectively.

The result using $k_b = k_t = 1$ for an initially 3D random fiber mass ($\Phi_r = \Phi_0$) is plotted in Fig. 5 for some different values of the crimp ratio ($\beta = b/a$). Compared to the van Wyk equation (Eq. (1) with $k = 0.01$), the shape of Eq. (66) differs above a fiber volume fraction of approximately 0.1. This is probably due to the transition from a 3D to a planar structure, where $P \propto \Phi^5$ is expected [10].

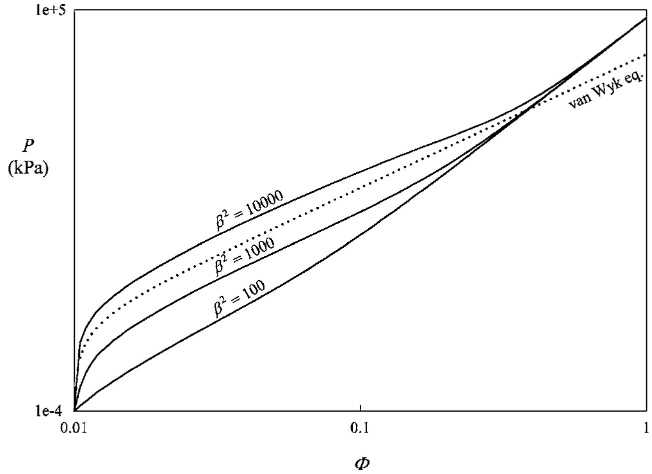


Fig. 5 Uniaxial compressive stress $P=-\Sigma_{33}$ versus volume fraction Φ. The dotted line is the van Wyk equation with $k=0.01$.

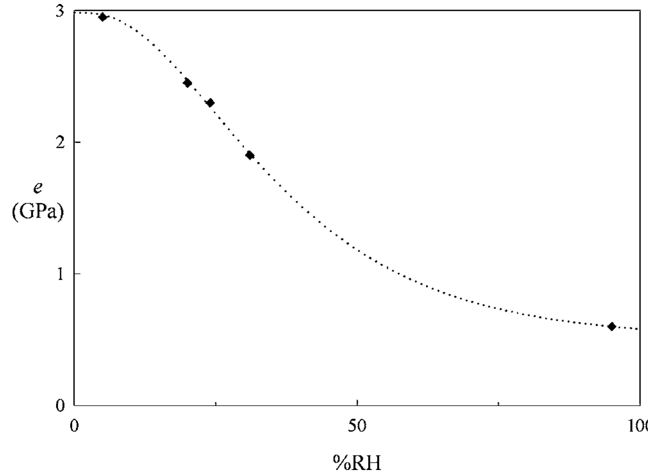


Fig. 7 The RH dependence of the elastic modulus, *e*, for PA-6 monofilaments. The dotted line is an arbitrary curve fit.

Experiment

Measuring stress–strain relations with conventional parallel plate techniques is associated with inaccuracies due to free edge effects [15]. This is especially true for materials with a coarse microstructure. Using the fact that the free edge effect decays away from an edge, we have constructed a parallel-plate instrument [15] where the stress is measured by means of a local stress transducer in a region away from the sample edges, where the strain field is uniform (Fig. 6). The newly developed stress transducer is triaxial, measuring the in-plane shear stresses as well as the compressive stress. Here, however, only the compression axis is used.

The sample is fixed between the transducer head and an x, y, z table and then deformed by imposing a horizontal and/or a vertical displacement on one plate relative the other. The relative plate displacement in the z direction is detected by means of high accuracy laser sensors and the resolution of the resulting strain is well within ±0.1% for a sample thickness of at least 1 mm. The output signals from the load cell and the laser sensor are digitized and recorded by a computer. The load cell sensitivity is about 10^{-1} kPa in compression.

Four separate fiber masses, consisting of polyamide-6 fibers of different fiber diameters 27, 35, 50, and 67 μm, were studied. The material was carded in a laboratory carding machine, CORMATEX model CC/400, to a pile. The fiber diameter was uniform within each pile. The surface weight of the piles was $0.06\ \mathrm{g/cm^2} \pm 10\%$. The samples were cut out of the piles using a punching machine. To erase the stress history from the carding and storage, the samples were compressed to a pressure of $P=100$ kPa and then unloaded to $P=80$ kPa. Compression data were sampled in the parallel-plate instrument at a constant displacement rate of 50 mm/min, a constant temperature of 21.0°C, and a relative humidity of 31%. To study the effect of the relative humidity on the Young's modulus, e, of the monofilaments, their tensile behavior was measured in a RheoMetrics RSA II at five different humidity levels, 5%, 21%, 24%, 32%, and 95% relative humidity (RH), at a constant strain rate of $\dot{\varepsilon}=10^{-3}\ \mathrm{s}^{-1}$. The fibers were conditioned for at least 2 h. Counting the number of crimps, n, on fibers of a total length, l, gave a mean crimp spacing of $2b=l/n$.

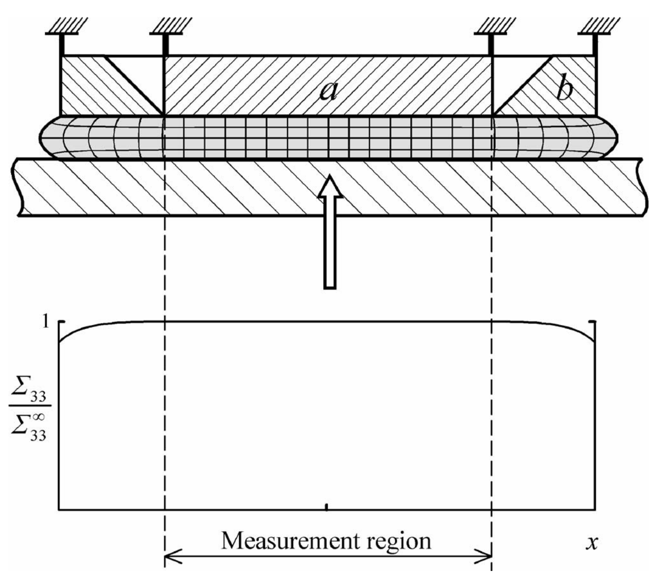


Fig. 6 Schematic of sample deformation and the resulting stress distribution $\Sigma_{33}/\Sigma_{33}^{\infty}$ in compression. The objective is to measure the asymptotic stress, Σ_{33}^{∞}, far from the edge.

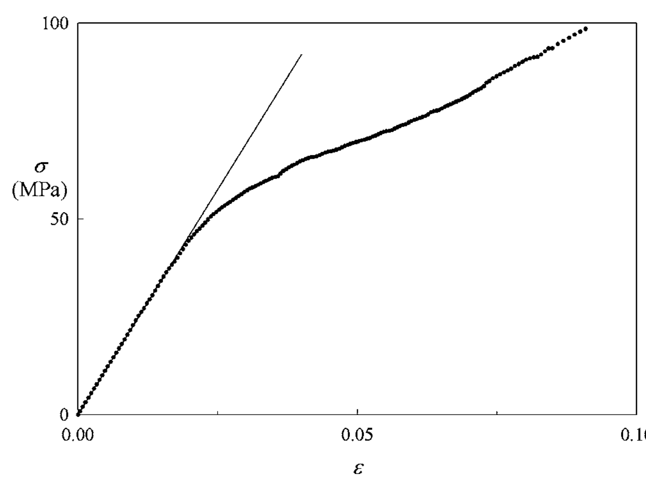


Fig. 8 Tensile data for PA6 monofilament at 24% RH. The straight line indicates the initial elastic modulus, *e*.

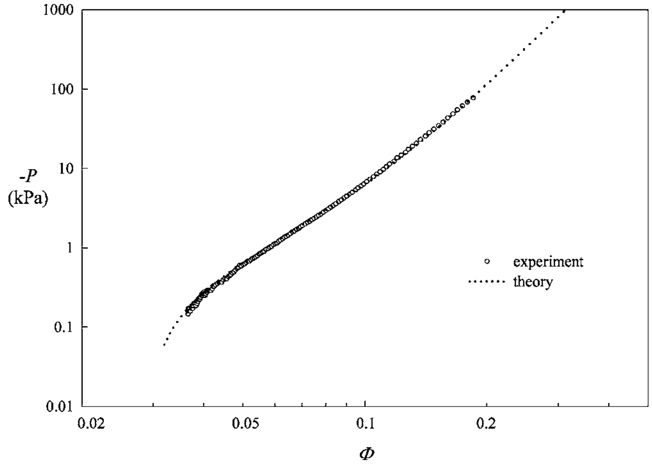


Fig. 9 Compression data for a fiber mass consisting of fibers with $d=35\ \mu$m

Results

The results of the monofilament elastic modulus measurements at different humidity levels are summarized in Fig. 7. The value of e decays exponentially with increasing humidity, as should be expected. Different monofilament diameters did not have any influence on the measured values of e. Figure 8 shows a typical tensile stress–strain curve (average of five separate monofilaments).

Figures 9 and 10 present data from the compression experiments. Each plot represents the average of ten different samples. The dotted lines are fits of Eq. (66) to the experimental data. The resulting values of the parameters are presented in Table 1. As discussed in the previous section, the structural parameters were computed from a reference state with 3D random orientation. Image analysis of the pile surface after carding revealed a very close to isotropic orientation of fiber *segments*. Hence we simply let the reference volume fraction Φ_r be the volume fraction of the pile as carded. It is difficult to determine any precise volume fraction of the unloaded pile, but this volume fraction appeared to be proportional to the limiting volume fraction Φ_0. The latter is taken as the volume fraction where a nonzero compressive stress is first registered, and is straightforward to determine from experimental data (Figs. 9 and 10). Hence, we chose to hold the ratio constant at $\Phi_0/\Phi_r=23/4$, corresponding to $\Phi_r\approx 0.8\%$. The initial condition of the numerical computation of the pressure response was $P=0$ at $\Phi=\Phi_0$.

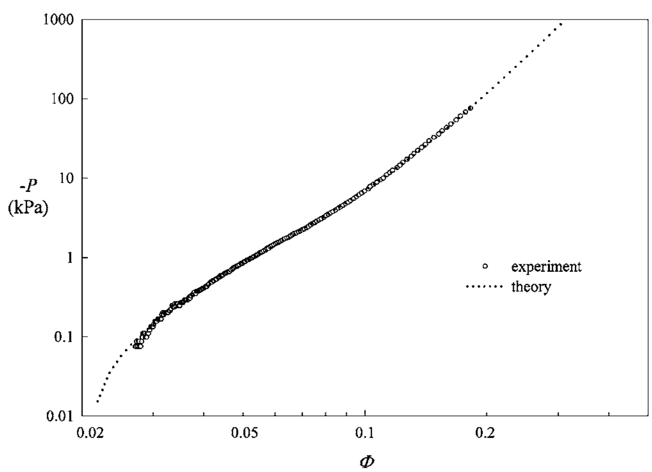


Fig. 10 Compression data for a fiber mass consisting of fibers with $d=50\ \mu$m

Table 1 Fitted parameters

$2a$ (μm)	27	35	50	67
Φ_0	0.022	0.023	0.022	0.024
k_b	2.2	2.2	2.2	2.2
k_t	0.05	0.05	0.05	0.05
$2b$ (mm)	0.8	1.0	1.5	1.9

Table 2 Measured crimp spacing

$2a$ (μm)	27	35	50	67
$2b$ (mm)	0.9	1.2	1.7	2.1

As anticipated, the fiber mass response in compression at higher volume fractions ($>12\%$) is independent of the fiber diameter, i.e., k_b is constant, provided that the fiber diameter is uniform and the fibers are slender. This suggests (as the Young's modulus is constant) that the fiber orientation at a certain fiber volume fraction is close to identical for the different samples used here. The experimental value of the bending correction parameter is $k_b \approx 2.2$, irrespective of fiber radius. At lower volume fractions ($<8\%$) the main contribution to the total fiber mass response comes from the last term in Eq. (66), which involves all adjustable parameters (k_b, k_t, β). The resulting experimental value of the torsional correction factor was $k_t\approx 0.05$ for all tested materials. Finally, the crimp ratio, β, seems to be constant irrespective of fiber radius, suggesting that the assumption of proportionality between crimp spacing and fiber radius is correct. The experimental value of the ratio, $\beta\approx 29$, corresponds to a crimp spacing of 0.8–1.9 mm for the tested materials (Table 1). These values of b are in very good agreement with the measured values of the crimp spacing (Table 2). Above all, the theoretical estimate according to Eq. (66) is in excellent agreement with the experimental data.

A benefit of using polyamide fibers was that their elastic modulus, e, could be varied by controlling their moisture content. In this way it was possible to study the effect of changing e without changing the network structure (ignoring the volume change of the fibers). A change in the Young's modulus results in a linear shift of the compression curves, again in agreement with the theory.

Discussion

When comparing the model predictions with the experimental data in Figs. 9 and 10, it should be appreciated that there is really only one adjustable parameter needed: k_b. The value of k_t plays a minor role, and only matters in the region where the curve increases its slope (around $\Phi\approx 0.1$). The limiting volume fraction Φ_0 only affects the initial rise of the curve and, finally, the crimp spacing $2b$ may be taken as measured from the micrographs.

The magnitude of the correction factor k_b is quite close to unity, thus supporting the modeling of a deformation element as a beam with fixed ends and central loading. The assumption of fixed end sections probably overestimates the stiffness while loading at the midsection probably underestimates it. The torsional stiffness factor k_t, on the other hand, is far below unity. The most likely explanation is that the fixed-end condition is inadequate in this case. Because the actual torsional compliance is so high, the contribution of this mechanism to the overall response is nearly negligible.

A critical assumption in our derivation is the no-slip condition at contacts. In compressive deformations of a close to planar fiber mass, the normal force will be large compared to the tangential force on the vast majority of contacts. In such circumstances, slip is an unlikely event, and frictional dissipation may be neglected.

However, in shearing and tensile deformations slip will occur whenever the tangential contact force overcomes the friction. The overall stress at which a certain contact point will slip depends on the contact configuration.

The assumption of affine rotation of the fiber axes, Eq. (46), is a strong simplification. It presumes that all fiber segments having the same orientation rotate equally when the fiber mass is deformed, independently of the fiber segment surroundings. First this implies no steric hindrance from neighbors, which may be realistic for low fiber volume fractions but unrealistic at higher fractions. However, in our experiments, the orientation at higher volume fractions is close to planar and the orientational parameters are either close to constant or close to zero. In this way the effect of steric hindrance is probably reduced. Second it implies that no restrictions are imposed on a fiber segment by the rest of the fiber to which it is linked. This is a clearly unphysical effect, implicit in all current models that account for change of orientation [5–8,10,11]. Since the orientation evolution is computed separately in our approach, a more advanced model for this could easily be introduced. In terms of modeling the stress at any given orientation state, the assumption of independent straight fiber segments should be valid as long as the fiber crimp length is large enough that each segment holds sufficiently many contacts that the forces transmitted across segment ends may be neglected.

Introducing the assumption that the expectation of the contact displacement rate $\overline{\dot{\mathbf{r}}}$ is affine (Eq. (9)) implies that the average displacement of contacts with a certain configuration only depends on the macroscopic velocity gradient. The alternative of describing the motion of each contact is clearly not feasible within an analytical approach.

Conclusions

The proposed model has been tested for three-dimensionally oriented fiber masses in uniaxial compression. The model successfully describes the response over a wider range of compressive strain than earlier models. It coincides with earlier models, such as the van Wyk model, in the appropriate regime. The model involves four adjustable parameters, but all of those have a clear physical meaning, and can be estimated independently. The experimental values of the correction factors k_b and k_t are reasonably close to their theoretical estimates and thus advocate our choice of deformation mechanisms, bending, and torsion of the fibers between contacts. The close correlation between the measured values of the segmental length b and the values obtained by fitting the model to compression experiments further support this choice.

Our treatment of fiber orientation is rudimentary and leaves ample room for further development. The simplifying assumption of affine rotation of the fiber segment axes is crude, but the perhaps most important is to determine the actual initial orientation of the fiber mass. In the present, we had to assume that the initial orientation of the fiber mass is approximately random.

Our possibly most serious assumption is the unrealistic choice of convected force rate (Eq. (25)) which does not maintain orthonormality of the basis vectors. Apparently, from the experimental results, this is not a problem in the case of uniaxial compression. Nevertheless it would most likely add an inaccurate convected force contribution to the contact point force rate in shearing deformations. To describe properly more general deformations one would need a more realistic (and more complicated) convected force rate which conserves the orthonormality of the basis vectors. In addition it may be necessary to incorporate frictional dissipation at the contact points, which is ignored in the present. The latter is likely to be a considerable challenge.

References

[1] Schofield, J., 1938, “Research on Wool Felting,” J. Text. Inst., **29**, pp. T239–T252.

[2] van Wyk, C. M., 1946, “Note on the Compressibility of Wool,” J. Text. Inst., **37**, pp. T285–T292.

[3] Corte, H., and Kallmes, O., 1962, *Formation and Structure of Paper*, 1st ed., Tech. Sect. Brit. Papers and Board Makers Assn., London, Vol. 1.

[4] Komori, T., and Makishima, K., 1977, “Numbers of Fiber-to-Fiber Contacts in General Fiber Assemblies,” Text. Res. J., **47**, pp. 13–17.

[5] Lee, D. H., and Lee, J. K., 1985, “Initial Compressional Behaviour of Fibre Assembly,” *Objective Measurements: Application to Product Design and Process Control*, S. Kawabata, R. Postle, and M. Niwa, eds., The Textile Machinery Society of Japan, Osaka, pp. 613–622.

[6] Carnaby, G. A., and Pan, N., 1989, “Theory of the Compression Hysteresis of Fibrous Assemblies,” Text. Res. J., **59**(5), pp. 275–284.

[7] Komori, T., and Itoh, M., 1991, “A New Approach to the Theory of the Compression of Fiber Assemblies,” Text. Res. J., **61**(7), pp. 420–428.

[8] Komori, T., Itoh, M., and Takaku, A., 1992, “Model Analysis of the Compressibility of Fiber Assemblies,” Text. Res. J., **62**(10), pp. 567–574.

[9] Neckár, B., 1997, “Compression and Packing Density of Fibrous Assemblies,” Text. Res. J., **67**(2), pp. 123–130.

[10] Toll, S., 1998, “Packing Mechanics of Fiber Reinforcements,” Polym. Eng. Sci., **38**, pp. 1337–1350.

[11] Komori, T., and Itoh, M., 1997, “Analyzing the Compressibility of a Random Fiber Mass Based on the Modified Theory of Fiber Contact,” Text. Res. J., **67**(3), pp. 204–210.

[12] Komori, T., and Itoh, M., 1994, “A Modified Theory of Fiber Contact in General Fiber Assemblies,” Text. Res. J., **64**(9), pp. 519–528.

[13] Pan, N., 1993, “A Modified Analysis of the Microstructural Characteristics of General Fiber Assemblies,” Text. Res. J., **63**(6), pp. 336–345.

[14] Toll, S., 1993, “Note: On the Tube Model for Fiber Suspensions,” J. Rheol., **37**(1), pp. 123–125.

[15] Alkhagen, M., and Toll, S., 2002, “A Triaxial Rheometer for Soft Compressible Solids,” J. Rheol., **46**(1), pp. 31–47.

[16] Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, 4th ed., Cambridge University Press, Cambridge, UK.

[17] Ouadfel, H., and Rothenburg, L., 2001, “Stress-Force-Fabric’ Relationship for Assemblies of Ellipsoids,” Mech. Mater., **33**(4), pp. 201–221.

[18] Emeriault, F., and Cambou, B., 1996, “Micromechanical Modelling of Anisotropic Non-Linear Elasticity of Granular Medium,” Int. J. Solids Struct., **33**(18), pp. 2591–2607.

[19] Norris, A. N., and Johnson, D. L., 1997, “Nonlinear Elasticity of Granular Media,” Trans. ASME, J. Appl. Mech., **64**(1), pp. 39–49.

[20] Alzebdeh, K., and Ostoja-Starzewski, M., 1999, “On a Spring-Network Model and Effective Elastic Moduli of Granular Materials,” Trans. ASME, J. Appl. Mech., **66**(1), pp. 172–180.

[21] Truesdell, C., and Noll, W., 1992, *The Non-Linear Field Theories of Mechanics*, 2nd ed., Springer, Berlin.

[22] Holzapfel, G. A., 2000, *Nonlinear Solid Mechanics, A Continuum Approach for Engineering*, 1st ed., Wiley, Chichester, UK, Vol. 1.

[23] Toll, S., and Manson, J.-A. E., 1995, “Elastic Compression of a Fiber Network,” Trans. ASME, J. Appl. Mech., **62**(1), pp. 223–226.

A Preconditioning Mass Matrix to Avoid the Ill-Posed Two-Fluid Model

Angel L. Zanotti
e-mail: azanotti@intec.unl.edu.ar

Carlos G. Méndez

Norberto M. Nigro

Mario Storti

International Center for Computational Methods in Engineering (CIMEC),
INTEC-Universidad Nacional del Litoral-CONICET,
S3000GLN,
Güemes 3450,
Santa Fe, Argentina

Two-fluid models are central to the simulation of transport processes in two-phase homogenized systems. Even though this physical model has been widely accepted, an inherently nonhyperbolic and nonconservative ill-posed problem arises from the mathematical point of view. It has been demonstrated that this drawback occurs even for a very simplified model, i.e., an inviscid model with no interfacial terms. Much effort has been made to remedy this anomaly and in the literature two different types of approaches can be found. On one hand, extra terms with physical origin are added to model the interphase interaction, but even though this methodology seems to be realistic, several extra parameters arise from each added term with the associated difficulty in their estimation. On the other hand, mathematical based-work has been done to find a way to remove the complex eigenvalues obtained with two-fluid model equations. Preconditioned systems, characterized as a projection of the complex eigenvalues over the real axis, may be one of the choices. The aim of this paper is to introduce a simple and novel mathematical strategy based on the application of a preconditioning mass matrix that circumvents the drawback caused by the nonhyperbolic behavior of the original model. Although the mass and momentum conservation equations are modified, the target of this methodology is to present another way to reach a steady-state solution (using a time marching scheme), greatly valued by researchers in industrial process design. Attaining this goal is possible because only the temporal term is affected by the preconditioner. The obtained matrix has two parameters that correct the nonhyperbolic behavior of the model: the first one modifies the eigenvalues removing their imaginary part and the second one recovers the real part of the original eigenvalues. Besides the theoretical development of the preconditioning matrix, several numerical results are presented to show the validity of the method. [DOI: 10.1115/1.2711224]

Keywords: two-fluid model, ill-posed problems, multiphase flow

1 Introduction

A multi-phase flow denotes a continuum where more than a single phase is present, e.g., gas bubbles rising in a liquid, droplets of fluid moving in a gas, steam–water flows in a boiler, pipeline transport of gas, and oil mixtures and oil–gas–water flows in an oil well. These examples are found in a great variety of industrial and technological applications such as chemical reactors, boilers, combustion chambers, steelmaking plants, and nuclear power plants devices. According to the geometry of the interface, a two-phase flow can be classified into three types (see Ishii [1]): *separated flows, transitional or mixed flows*, and *dispersed flows*, with significant differences in their behavior.

When modeling a biphasic flow, it is necessary to know what phenomena, effects, and flow structures are important. In some cases, the exact structure or position of the interface is important, while in other cases only some kind of average quantity is needed for the flow analysis.

Models for two-phase flows can be categorized into two different groups. The first group is the so called interface tracking models (ITMs), which track the interface between the two phases, and are well suited for separated flows. The most frequently employed Eulerian-based ITM for predicting certain classes of multiphase flows are the volume of fluid (VOF) method [2–7], the front tracking (FT) or immersed boundary method [8,9], the level set (LS) methods [10–13], and the phase field (PF) methods [14–16].

Contributed by the Applied Mechanics Division of ASME for publication in JOURNAL OF APPLIED MECHANICS. Manuscript received June 27, 2005; final manuscript received October 10, 2006. Review conducted by Gustavo C. Buscaglia.

Ideally, one would like to track the interface between the phases at all times, which is similar to solving all relevant scales of turbulent single-phase flow direct numerical simulation (DNS). However, this is often computationally too expensive and sometimes redundant for dispersed flows. Thus, in the second group of models, the exact position of the interface is not followed accurately and only spatial distributions of volume-averaged quantities such as the void fraction and the velocity field for the dispersed phases are calculated. Dispersed flows are usually modeled using models from this second group.

Two generic approaches can be used for modeling dispersed flows: the *Lagrangian* and the *Eulerian* formulation. For particulate (or particle-like) flows, it is possible to build methods based on the marker and cell (MAC) scheme (see Ref. [17]). The general idea is to follow each particle of the dispersed phase as it is transported by the continuous phase. This approach, in which the continuous phase is calculated in a Eulerian reference frame, is referred to as a *Lagrangian–Eulerian formulation.*

A different way of modeling dispersed flows is to treat both phases as a continuum. This is generally referred to as the *Eulerian–Eulerian formulation* or the *two-fluid model*. In this model, discussed in detail by Ishii [1], each phase has its own set of balance equations and the interaction between phases is represented via interface transfer terms that arise from the constitutive relations. Therefore, both phases are present in each point of the domain, each one with an associated volume fraction.

In this study, the two-fluid model is chosen to simulate dispersed biphasic flows and a mathematical analysis of this model is presented in the next sections. It is well known (see Ref. [18]) that in a mathematical sense this model is *ill posed* because its *hyper-*

bolic nature may not be warranted for all the flow parameters. Furthermore, the model is questionable for its *nonconservative* formulation [19] and *nonlinear* terms make the problem more complex for the analysis.

Numerical solutions of ill-posed two-fluid problems have two drawbacks: excessive numerical diffusion and instabilities. These situations are common for ill-posed initial boundary value problems (see, e.g., Ref. [20]) and, therefore, the success of any method depends on the requirement that such systems be well posed.

It is important to make a distinction between *nonhyperbolic problems* considered in the present work and another class of problems that address loss of hyperbolicity due to a change of type in the partial differential equations. Change of type problems are a separate subject of mathematical analysis, which is related to how the information propagates between the hyperbolic and elliptic domains [20].

A lot of effort has been made to remedy the nonhyperbolic anomaly of the two-fluid model and two different approaches can be found. One is based on the physics of the problem and consists of adding extra terms to the interfacial interaction, thus searching for a physical solution to the nonhyperbolic behavior. Even though this methodology seems to be realistic, several extra parameters arise from each added term with the associated difficulty in their estimation [21]. One of the usually added terms, which produces significant improvement in the parameter range when the problem is well posed, is the interfacial pressure. This enhanced model relaxes the assumption of equal pressure for both phases and the introduction of two pressures is justified noting that the pressure of the continuous phase, computed by the two-fluid model, is far away from the unresolved flow around a bubble. This seems logical since the two-fluid model aims are not to solve the details of the flow close to individual bubbles. In this sense, Lahey [22] has reached very promising results for air–water systems finding well posedness for all void fractions with very large densities ratios using as a necessary condition that $C_p > 0.166$ (pressure coefficient in the pressure jump at interface). However, this lower bound may be in conflict with the results obtained by Drew and Passman [18] for highly viscous flow, where they found that C_p should be negative. Other contributions using different pressures for each phase that allows us to extend the range of parameters for which the problem is hyperbolic are available in the literature (see, e.g., Refs. [23–28]).

Another important and different contribution has been made by Stadtke et al. [29]. In their work, they split the interfacial momentum coupling terms in a viscous and nonviscous part. Drag forces are representative for the former. For the latter, a series of terms have been introduced in order to compensate for the information lost during the averaging procedure. These terms contain only space and time derivatives of major dependent parameters, including phasic velocities, void fraction, and phasic densities, using only one pressure. The authors enumerate the criteria for the design of the model and finally demonstrate that they achieve a full set of real eigenvalues with a complete set of independent eigenvectors, warranting that the problem will be well posed [30].

From a mathematical point of view, the target is to find a way to remove the complex eigenvalues obtained with the two-fluid conservation equations. Following Hadamard [31], a well posed problem is defined in these terms: "In order for a problem involving a partial differential equation (PDE) to be well-posed, the solution to the problem must exist and be unique, and the solution must depend continuously upon the initial and boundary data." An equivalent definition, more suitable for numerical evaluation, may be found through the strong hyperbolic character of a given equation system. The two-fluid model is a first order in a time system of equations with first- and second-order spatial derivatives for the convective and diffusive terms, respectively. In addition, there are some zero-order terms for sources coming from interfacial terms. Transforming these equations into a first-order system, a necessary requirement for the well posedness of the problem, is that this first-order system be strongly hyperbolic (see Refs. [32,33]), which means that the system be diagonalizable with real eigenvalues. The equivalent first-order differential equations represent the time evolution of the system and the nonguaranteed real eigenvalues suffice to prove the lack of causality of the solution and the final blowup. This means that in the effort of getting a numerical solution through a time marching scheme, fatal instabilities appear and forbid not only the computation of the time evolution of the variables but also the knowledge of the steady state of the system. Even though real problems have a temporal evolution, the knowledge of the steady solution is very useful for the industrial design of processes with multiphase flow involved. This fact has motivated the idea of recovering at least the steady solution in trying to modify the time marching scheme by a time preconditioner.

Preconditioned systems, characterized as a projection of the complex eigenvalues over the real axis, may be one of the possible choices. The main disadvantage of this alternative is that the original equations should be modified so that they lose some conservation properties. The present paper introduces a preconditioning mass matrix method that makes the two-fluid model hyperbolic, recovering the conservation properties for the final solution, and providing an alternative to obtain steady solutions, for multiphase flow problems, which will be of help for process design.

The next sections are organized as follows. Section 2 presents the two-fluid model for a one-dimensional problem with equal phase pressures. Section 3 develops a characteristic analysis of this particular problem (see Ref. [34]) with emphasis on the lack of real eigenvalues in order to motivate the proposal of a preconditioning mass matrix. The paper then presents the mass matrix used as a preconditioner, analyzing the role of the two parameters included in its definition: the first one, β, is for making the model hyperbolic, and the second one, γ, is used for recovering the propagation wave speed corresponding to the original model. Finally, representative problems are solved numerically with the purpose of checking that the preconditioning method circumvents the instabilities showed by most multifluid models when using a large density ratio, high void fractions, and large relative velocities among the phases.

2 Two-Fluid Model

The complex nature of multiphase flows, characterized by changes in the geometrical configuration of the different phases, makes it extremely difficult to find models that reproduce the physics of the system and that are at the same time numerically tractable at a reasonable computational cost. Mathematical models based on averaged fields of the phases (see, e.g., Refs. [1,18,35]) associated with experimental correlations seem to be one of the best alternatives. These multifield models are widely used to model and to simulate the transport of multiphase flow systems. They treat each phase as an interpenetrating continuum (field), and conservation laws are applied to each one of them. In this approximation, constitutive laws have to be provided to represent the interaction between fields. In Fig. 1 we can see a schematic representation of two-fluid models.

Two-phase flow averaged equations are presented below (see Refs. [1,36]), which have been obtained as a result of temporal and/or space averaging of the instantaneous local balance equations. Without loss of generality, *the simplest model* proposed by Drew and Passman [18] is considered. In this one-dimensional model the flow is assumed to be inviscid, isothermal, and without phase change. Supposing for concreteness that one phase is gaseous and the other one liquid, we call α_g and α_l their corresponding volume fractions. Thus, mass and momentum balance equations can be written in the following way

$$\frac{\partial(\alpha_g \rho_g)}{\partial t} + \frac{\partial(\alpha_g \rho_g v_g)}{\partial x} = 0 \qquad (1)$$

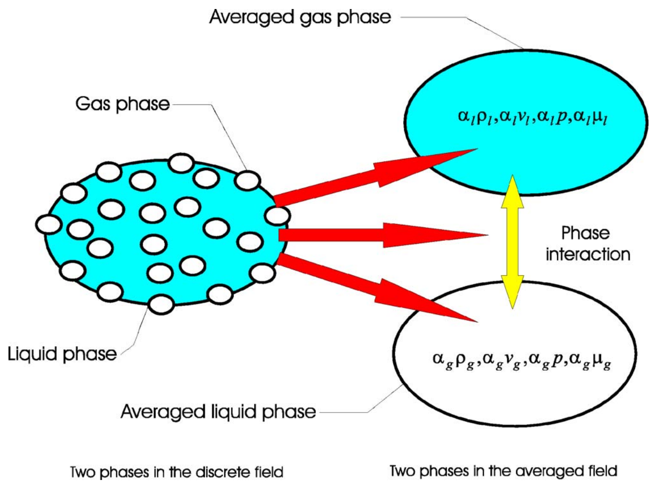


Fig. 1 Schematic representation of two-fluid model

$$\frac{\partial(\alpha_l\rho_l)}{\partial t} + \frac{\partial(\alpha_l\rho_l\nu_l)}{\partial x} = 0 \tag{2}$$

$$\alpha_g\rho_g\frac{\partial\nu_g}{\partial t} + \alpha_g\rho_g\nu_g\frac{\partial\nu_g}{\partial x} = -\alpha_g\frac{\partial p}{\partial x} + \alpha_g\rho_g g - F_I \tag{3}$$

$$\alpha_l\rho_l\frac{\partial\nu_l}{\partial t} + \alpha_l\rho_l\nu_l\frac{\partial\nu_l}{\partial x} = -\alpha_l\frac{\partial p}{\partial x} + \alpha_l\rho_l g + F_I \tag{4}$$

where ρ, ν, and p denote density, velocity, and pressure, respectively. In this model, the pressure p is assumed the same for both phases. Terms containing g represent the gravitational force, and F_I models the interaction between the phases. Several physical effects may be included in interphase force F_I [21], but certainly the most common of them is the drag force. For dispersed flows this force can be modeled as [37]

$$F_I = \frac{3}{8}C_d\rho_l(\nu_g-\nu_l)|\nu_g-\nu_l|\frac{\alpha_g}{r_b} \tag{5}$$

where r_b is the mean radius of the dispersed phase and C_d is the drag coefficient for which empirical correlations are available as a function of the Reynolds number.

3 Characteristics Analysis

Examining the characteristic values of the governing equations, it can be determined if a model is properly formulated. Taking into account the constraint $\alpha_g+\alpha_l=1$, we can define $\Phi=(\alpha_g,p,\nu_g,\nu_l)$ as the unknown state vector and Eqs. (1)–(4) can be written in vector form as

$$\mathbf{A}\frac{\partial\Phi}{\partial t} + \mathbf{B}\frac{\partial\Phi}{\partial x} + \mathbf{C} = 0 \tag{6}$$

If incompressibility of each phase is assumed, matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & \alpha_g\rho_g & 0 \\ 0 & 0 & 0 & (1-\alpha_g)\rho_l \end{pmatrix} \tag{7}$$

$$\mathbf{B} = \begin{pmatrix} \nu_g & 0 & \alpha_g & 0 \\ -\nu_l & 0 & 0 & 1-\alpha_g \\ 0 & \alpha_g & \alpha_g\rho_g\nu_g & 0 \\ 0 & 1-\alpha_g & 0 & (1-\alpha_g)\rho_l\nu_l \end{pmatrix} \tag{8}$$

and

$$\mathbf{C} = \begin{pmatrix} 0 \\ 0 \\ -\alpha_g\rho_g g + F_I \\ -(1-\alpha_g)\rho_l g - F_I \end{pmatrix} \tag{9}$$

The local linear dynamic character of Eq. (6) can be examined by linearizing the system about an initial state Φ_0 (from now on, we assume that all matrices and derivatives are evaluated in this state). The linear differential equation for the behavior of a perturbation $\delta\Phi=\Phi-\Phi_0$ is

$$\mathbf{A}\frac{\partial\delta\Phi}{\partial t} + \mathbf{B}\frac{\partial\delta\Phi}{\partial x} + \left(\frac{\partial\mathbf{A}}{\partial\Phi}\cdot\frac{\partial\Phi}{\partial t} + \frac{\partial\mathbf{B}}{\partial\Phi}\cdot\frac{\partial\Phi}{\partial x} + \frac{\partial\mathbf{C}}{\partial\Phi}\right)\delta\Phi = 0 \tag{10}$$

A solution in the form of a traveling wave is assumed

$$\delta\Phi = \delta\Phi_0\exp[i(kx-\omega t)] \tag{11}$$

where $\delta\Phi_0$ represents the initial amplitude of the perturbation. The imaginary part ω_I of ω will govern growth or decay depending on its sign and the real part ω_R yields the speed of propagation. Substituting Eq. (11) into Eq. (10), the compatibility condition that $\delta\Phi_0$ must satisfy is

$$-i\omega\mathbf{A}\delta\Phi_0 + ik\mathbf{B}\delta\Phi_0 + \left(\frac{\partial\mathbf{A}}{\partial\Phi}\cdot\frac{\partial\Phi}{\partial t} + \frac{\partial\mathbf{B}}{\partial\Phi}\cdot\frac{\partial\Phi}{\partial x} + \frac{\partial\mathbf{C}}{\partial\Phi}\right)\delta\Phi_0 = 0 \tag{12}$$

For an initial uniform steady state $\partial\Phi/\partial t$ and $\partial\Phi/\partial x$ are zero. Defining $\lambda=\omega/k$ and $\mathbf{D}=\partial\mathbf{C}/\partial\Phi$, the condition under which nontrivial solutions for $\delta\Phi_0$ exist is given by

$$\det\left(\mathbf{A}\lambda - \mathbf{B} + \frac{i}{k}\mathbf{D}\right) = 0 \tag{13}$$

In the limit as $k\to\infty$, Eq. (13) reduces to the *characteristic equation* corresponding to Eq. (6)

$$\det(\mathbf{A}\lambda-\mathbf{B}) = 0 \tag{14}$$

and the values that λ can take are the *characteristic values*. Note that algebraic terms like gravitational or drag forces, which do not contain derivatives of the unknowns, do not affect the characteristic analysis.

The characteristic values of the simplest two-fluid model, Eqs. (6)–(9), are given by

$$\lambda = \left[\infty,\infty,\frac{1}{d}(r\pm s^{1/2})\right] \tag{15}$$

where

$$r = \alpha_g\rho_l\nu_l + (1-\alpha_g)\rho_g\nu_g$$

$$s = -\alpha_g(1-\alpha_g)\rho_g\rho_l(\nu_g-\nu_l)^2$$

$$d = (1-\alpha_g)\rho_g + \alpha_g\rho_l \tag{16}$$

We can observe that, except for the case when $\nu_g=\nu_l$, there are two complex conjugated values for the characteristic λ. Thus, since we are working in the limit $k\to\infty$, the imaginary part of omega $\omega_I=k\lambda_I$ can take arbitrarily large values. Consequently, as can be seen from Eq. (11), the perturbation will grow without limit even for a small increment in time. In other words, a little disturbance of the initial state will diverge instantaneously. This is in contradiction to the third Hadamard condition for a well-posed problem because small perturbations are not reflected as small (or at least finite) changes in the solution. Therefore, the solution does not depend continuously on its data and the problem is said to be ill-posed.

It is known that a well-posed problem can be guaranteed if all the characteristic values are real and distinct (strong hyperbolic system) [32,33]. As in the one-phase case, the degeneration of the two infinite values of Eq. (15) can be removed if a finite sound velocity for each phase is considered. However, the other two complex values are not so easy to avoid. In the next section, we will show a method to solve this difficulty.

4 Preconditioning Mass Matrix

We have seen in the previous section that the simplest two-fluid model has complex characteristic roots and is therefore ill posed as an initial value problem except for the trivial case of equal phase velocities [18]. In this section, we propose a simple method (already used in one-phase flows [38,39]) that permits us to hyperbolize the two-fluid model. This method consists of premultiplying the matrix **A** of Eq. (6) by another matrix **M**. From Eq. (6) we can see that the *preconditioning matrix* **M** only affects terms with temporal derivatives. Since these terms vanish when the steady-state solution is achieved, this solution is not changed by the preconditioning matrix.

The preconditioning mass matrix **M** has two parameters. The first, β, only affects the inertia of the mass equation and allows the hyperbolization of the model (no complex characteristic values). The second, γ, affects each one of the temporal terms of the balance equations, its purpose being to correct the speed of propagation of the waves. Thus, the form of the proposed matrix **M** is

$$\mathbf{M} = \gamma \begin{pmatrix} \beta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{17}$$

and defining $\tilde{\mathbf{A}} = (\mathbf{MA})$, the new characteristic equation is given by

$$\det(\tilde{\mathbf{A}}\lambda - \mathbf{B}) = 0 \tag{18}$$

Developing the determinant, we arrive to the general expression

$$I_0\lambda^2 + I_1\lambda + I_2 = 0 \tag{19}$$

where

$$I_0 = -(\rho_l\alpha_g + \beta\rho_g\alpha_l) \tag{20}$$

$$I_1 = (2\rho_l\nu_l\alpha_g + (\beta+1)\rho_g\nu_g\alpha_l)/\gamma \tag{21}$$

$$I_2 = -(\rho_l\nu_l^2\alpha_g + \rho_g\nu_g^2\alpha_l)/\gamma^2 \tag{22}$$

The roots of the characteristic equation, Eq. (19), are

$$\lambda_{1,2} = -\frac{I_1}{2I_0} \pm \sqrt{\left(\frac{I_1}{2I_0}\right)^2 - \frac{I_2}{I_0}} \tag{23}$$

and defining $C_1 = I_1/(2I_0)$ and $C_2 = I_2/I_0$ with

$$C_1 = -\frac{2\rho_l\nu_l\alpha_g + (\beta+1)\rho_g\nu_g\alpha_l}{2\gamma(\rho_l\alpha_g + \beta\rho_g\alpha_l)} \tag{24}$$

$$C_2 = \frac{\rho_l\nu_l^2\alpha_g + \rho_g\nu_g^2\alpha_l}{\gamma^2(\rho_l\alpha_g + \beta\rho_g\alpha_l)} \tag{25}$$

we arrive at the following expression for the roots

$$\lambda_{1,2} = -C_1 \pm \sqrt{C_1^2 - C_2} \tag{26}$$

Since we want to assure hyperbolicity, the roots should not have imaginary components and therefore we ask for

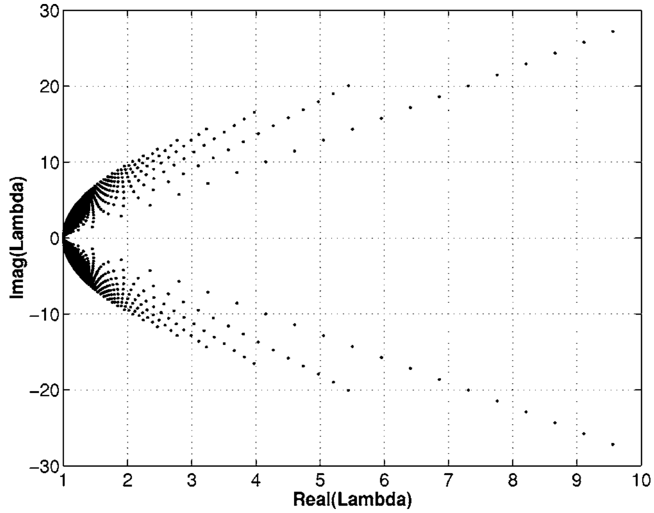


Fig. 2 Eigenvalues for a sweeping in alpha (0.01:0.01:0.99) and velocity relations (1:5:100), without preconditioning

$$\Delta = C_1^2 - C_2 \geqslant 0 \tag{27}$$

If the discriminant is equalized to zero, β_{crit} values are found. Calling $A = \alpha_l/\alpha_g$, $B = \rho_l/\rho_g$, and $C = \nu_l/\nu_g$ we obtain

$$\Delta = C_1^2 - C_2 = A\beta_{\text{crit}}^2 - 2(2BC(C-1) + A)\beta_{\text{crit}} + A + 4B(C-1) = 0 \tag{28}$$

After computing the β_{crit} from the previous equation and substituting it into Eq. (26) we reach the following expression for the characteristic roots with $\beta = \beta_{\text{crit}}$

$$\lambda_{1,2} = \frac{(D \pm \sqrt{E \cdot D})\nu_g}{(\rho_l\alpha_g((\nu_l - \nu_g)^2 + \nu_l^2) + \rho_g\nu_g^2\alpha_l \pm 2\sqrt{E \cdot D})\gamma}$$

$$D = \rho_l\nu_l^2\alpha_g + \rho_g\nu_g^2\alpha_l$$

$$E = \rho_l\alpha_g(\nu_l - 2\nu_g)^2 \tag{29}$$

In Fig. 2 we have represented the eigenvalues of the problem without preconditioning which arise as a result of a sweeping in the void fraction and velocity relations. As it can be seen, they possess imaginary components except for velocity relations $C = \nu_l/\nu_g = 1$. Figure 3 corresponds to the problem with preconditioning for $\gamma = 1$. In this case, although the characteristics do not take complex value, the propagation velocity has been modified for the parameter β. Note that not only has the maximum value of the real component been modified, but there are also negative propagation velocities.

The parameter β that permits hyperbolizing the differential two-fluid model modifies the real part of characteristic values. Therefore, the temporal evolution of the problem is different from the original one. With the purpose of recovering the temporal behavior, the preconditioning matrix contains a parameter γ whose determination is next developed.

From Eq. (26) for the characteristic values, we can observe that the original system, $\beta = 1$ and $\gamma = 1$, presents imaginary characteristics. The real part, which determines the propagation velocity of the information in the medium, is given by

$$\text{Re}(\lambda_{1,2}) = -C_1(\beta = 1; \gamma = 1) \tag{30}$$

Parameter γ is obtained from the relation between the characteristic for $\beta = \beta_{\text{crit}}$ and the real part $\text{Re}(\lambda_{1,2})$. Therefore

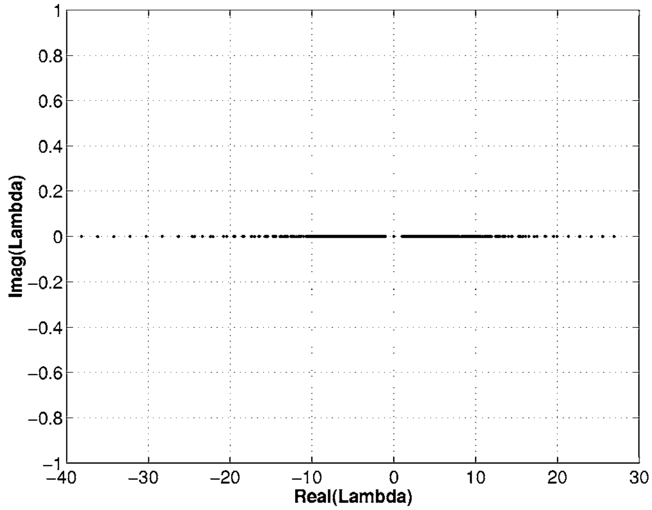


Fig. 3 Eigenvalues for a sweeping in alpha (0.01:0.01:0.99) and velocity relations (1:5:100), with preconditioning and γ=1. All eigenvalues are real, but are different from the real part of eigenvalues of Fig. 2.

$$\gamma_{1,2} = -\frac{\rho_l\alpha_g+\rho_g\alpha_l}{\rho_l\nu_l\alpha_g+\rho_g\nu_g\alpha_l}\frac{(D\pm\sqrt{E\cdot D})\nu_g}{\rho_l\alpha_g((\nu_l-\nu_g)^2+\nu_l^2)+\rho_g\nu_g^2\alpha_l\pm 2\sqrt{E\cdot D}} \tag{31}$$

In Fig. 4 we can observe the characteristic values with the parameter γ determined by the last expression. These values do not present complex components and they recover the value of the real component corresponding to the original problem without preconditioning.

5 Numerical Results

In this section we verify numerically the validity of the preconditioning method using a time marching scheme to solve as a first example a well-known problem, the so called *water faucet*, which has an analytical solution. The ill posedness of the two-fluid model without preconditioning causes in numerical implementations a tendency to develop instabilities that grow up and propagate through the domain. These instabilities are more likely to occur at high void fractions and at large density and velocity ratios. Thus, we propose another example with unfavorable initial conditions (with a tendency to instabilities) that we call the *wave-traveling problem*. Without preconditioning, even the robust numerical scheme described in the next subsection fails to solve this second example.

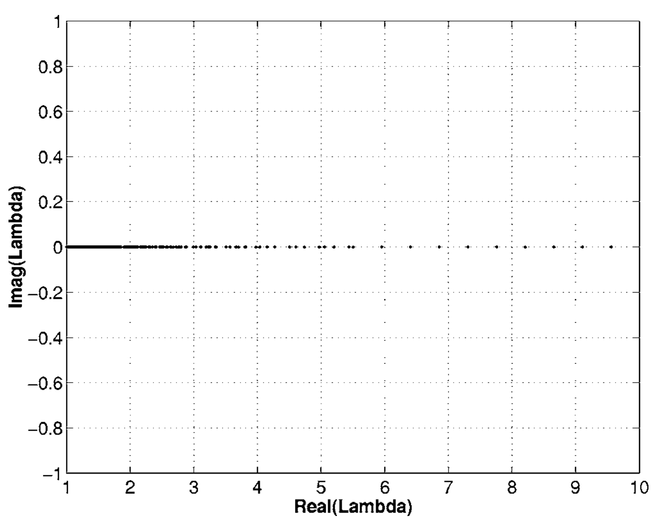


Fig. 4 Eigenvalue for a sweeping in alpha (0.01:0.01:0.99) and velocity relations (1:5:100) with preconditioning and $\gamma\neq$1. All eigenvalues are real and are equal to the real part of eigenvalues of Fig. 2.

5.1 Discretization. The implemented numerical discretization is based on a semi-implicit scheme with donor or upwind cell differencing for the convective terms. A staggered spatial nodalization is used and thus scalar variables α, ρ, and p are determined at the center of the control volumes (j index), while velocity variables ν_g and ν_l are located at the edges ($j+1/2$ index). The discretized form of Eqs. (1)–(4) are then

$$\frac{1}{\Delta t}[(\alpha_g\rho_g)_j^{n+1}-(\alpha_g\rho_g)_j^n]+\frac{1}{\Delta x}[(\hat{\alpha}_g\hat{\rho}_g)_{j+1/2}^n(\nu_g)_{j+1/2}^{n+1} - (\hat{\alpha}_g\hat{\rho}_g)_{j-1/2}^n(\nu_g)_{j-1/2}^{n+1}]=0 \tag{32}$$

$$\frac{1}{\Delta t}[(\alpha_l\rho_l)_j^{n+1}-(\alpha_l\rho_l)_j^n]+\frac{1}{\Delta x}[(\hat{\alpha}_l\hat{\rho}_l)_{j+1/2}^n(\nu_l)_{j+1/2}^{n+1} - (\hat{\alpha}_l\hat{\rho}_l)_{j-1/2}^n(\nu_l)_{j-1/2}^{n+1}]=0 \tag{33}$$

$$\frac{1}{\Delta t}(\hat{\alpha}_g\hat{\rho}_g)_{j+1/2}^n(\nu_g^{n+1}-\nu_g^n)_{j+1/2} + \frac{1}{\Delta x}(\hat{\alpha}_g\hat{\rho}_g\nu_g)_{j+1/2}^n[(\nu_g)_{j+1/2}^n-(\nu_g)_{j-1/2}^n] = -\frac{1}{\Delta x}(\alpha_g)_{j+1/2}^n(P_{j+1}-P_j)^{n+1}+(\alpha_g\rho_g)_{j+1/2}^n g-(F_l)_{j+1/2}^{n+1} \tag{34}$$

and

$$\frac{1}{\Delta t}(\hat{\alpha}_l\hat{\rho}_l)_{j+1/2}^n(\nu_l^{n+1}-\nu_l^n)_{j+1/2} + \frac{1}{\Delta x}(\hat{\alpha}_l\hat{\rho}_l\nu_l)_{j+1/2}^n[(\nu_l)_{j+1/2}^n-(\nu_l)_{j-1/2}^n] = -\frac{1}{\Delta x}(\alpha_l)_{j+1/2}^n(P_{j+1}-P_j)^{n+1}+(\alpha_l\rho_l)_{j+1/2}^n g-(F_l)_{j+1/2}^{n+1} \tag{35}$$

Scalar variables α and ρ at $j+1/2$ are the average value between j and $j+1$ and, calling θ to α or ρ, hat variables are defined as

$$\hat{\theta}_{j+1/2}=\frac{1}{2}(\theta_j+\theta_{j+1})+\frac{1}{2}\frac{\nu_{j+1/2}}{|\nu_{j+1/2}|}(\theta_j-\theta_{j+1}) \tag{36}$$

where ν is the velocity corresponding to the phase considered.

This numerical model is the one used in the RELAP5 code for multiphase flows. More details, like the implementation of an automatic control of the time step based on the Courant number, can be found in Refs. [28,40]. The inclusion of the preconditioning matrix **M** within this scheme does not present difficulties. It is just to multiply the corresponding terms in the balance equations by the parameters β and γ according to Eq. (17). In order to ensure strong hyperbolicity (real and distinct characteristic values), large but finite sound velocities as well as β parameters slightly greater than β_{crit} were used.

5.2 Water Faucet Problem. Due to the fact that it has an analytical solution, the water faucet problem devised by Ransom [41] is widely used to validate two-phase flow models [42–44]. The problem consists of a 12-m-long vertical tube where there is initially a uniform volume fraction (α_l^0=0.6) of water (ρ_l =1000 kg/m^3) that is moving at constant velocity (ν_l^0=15 m/s) in a stagnant air annulus. No interaction between phases is considered. When the simulation starts (t=0), a gravity field (g =9.8 m/s^2) is applied and this causes the water column to accelerate. At the top of the tube (inlet), water volume fraction and velocity are kept unchanged ($\alpha_l^{\text{inlet}}=\alpha_l^0$ and $\nu_l^{\text{inlet}}=\nu_l^0$), and at the

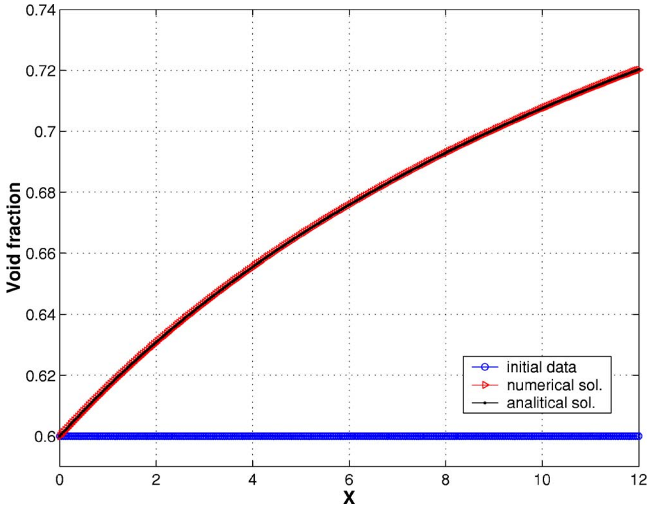


Fig. 5 The water faucet problem. Void fraction with preconditioning at steady state, for a mesh of 320 uniform lineal elements. Comparison between numerical and analytical solutions.

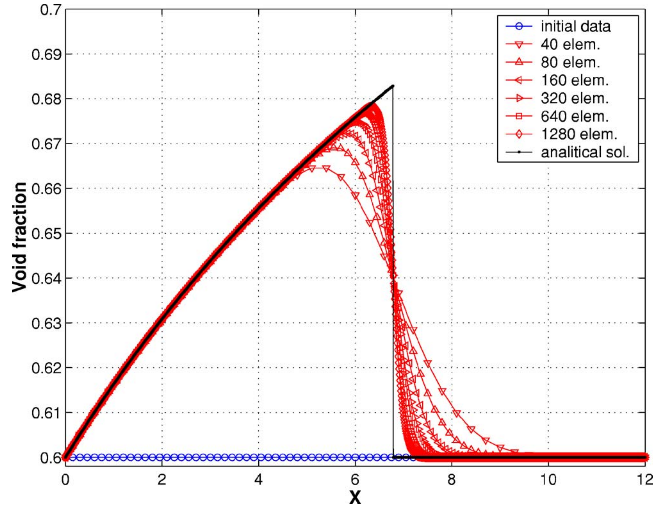


Fig. 7 The water faucet problem. Void fraction with preconditioning at t=0.4 s, for six different meshes of 40, 80, 160, 320, 640, and 1280 uniform lineal elements.

bottom (outlet) a constant pressure is maintained ($p^{\text{outlet}}=10^5$ Pa). Due to the acceleration, a contact discontinuity propagates downward until a steady state is reached when the discontinuity arrives at the outlet.

Neglecting pressure gradient in both fluids, the analytical solution to the water faucet problem is given by

$$\text{If } x \leqslant \nu_l^0 t + \frac{1}{2} g t^2 \begin{cases} \alpha_g = 1 - \dfrac{\alpha_l^0 \nu_l^0}{\nu_l} \\ \nu_l = [(\nu_l^0)^2 + 2gx]^{1/2}; \end{cases}$$

$$\text{and otherwise } \alpha_g = 1 - \alpha_l^0$$

Using the preconditioning mass matrix, this problem was simulated with six different meshes of 40, 80, 160, 320, 640, and 1280 uniform lineal elements. In Figs. 5 and 6 numerical as well as analytical solutions for the void fraction and water velocity are shown. These figures correspond to a time ($t > 2$ s) in which the steady state was reached. It can be seen that, as mentioned above, the preconditioning method does not affect the steady state.

Figures 7 and 8 show the void fraction and liquid velocity, respectively, at time (t=0.4 s). At this time, the discontinuity is still within the tube. Although agreement between numerical and analytical solutions in a transient state is not expected (we are using $\gamma=1$), these figures show that the velocity of propagation of the discontinuity is well reproduced. Besides, numerical solutions for water velocity and void fraction tend to analytical values when the number of elements is increased, capturing the contact discontinuity very well.

5.3 Wave Traveling Problem. In this problem, unfavorable conditions are set up in order to reflect in the numerical simulation the ill-posed character of the differential equations. The domain is one dimensional and its length L=0.4 m is discretized in 100 elements. The density ratio ρ_l/ρ_g is 1000:1 and all the fields are initially uniform (ν_l=1 m/s, ν_g=10 m/s, p=0) except for the void fraction that is a sinusoidal perturbation [α_g=0.5 +0.45 sin($4\pi x/L$)]. Periodic boundary conditions are imposed. Thus, waves generated by the void fraction perturbation can

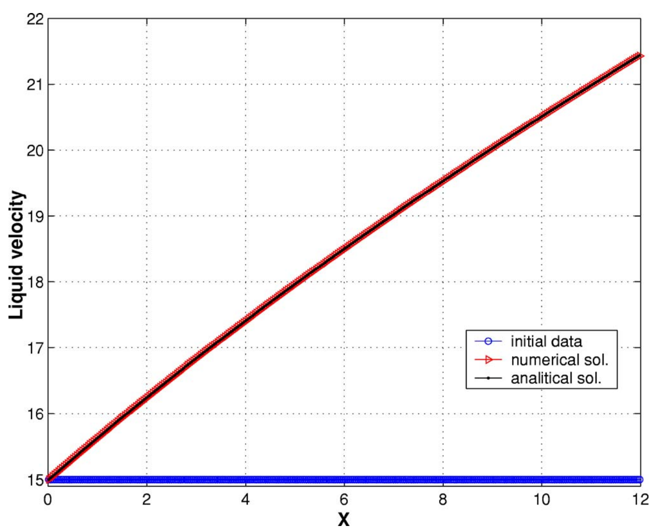


Fig. 6 The water faucet problem. Liquid velocity with preconditioning at steady state, for a mesh of 320 uniform lineal elements. Comparison between numerical and analytical solutions.

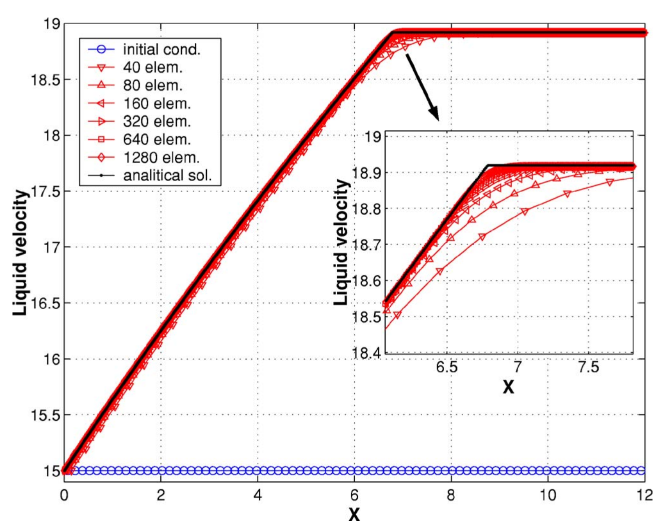


Fig. 8 The water faucet problem. Liquid velocity with preconditioning at t=0.4 s, for six different meshes of 40, 80, 160, 320, 640, and 1280 uniform lineal elements.

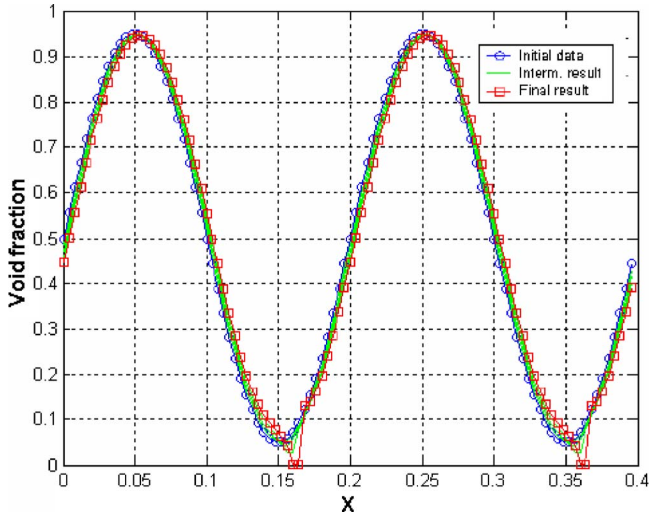


Fig. 9 The wave traveling problem. Void fraction without preconditioning at four time steps of 0.001 s.

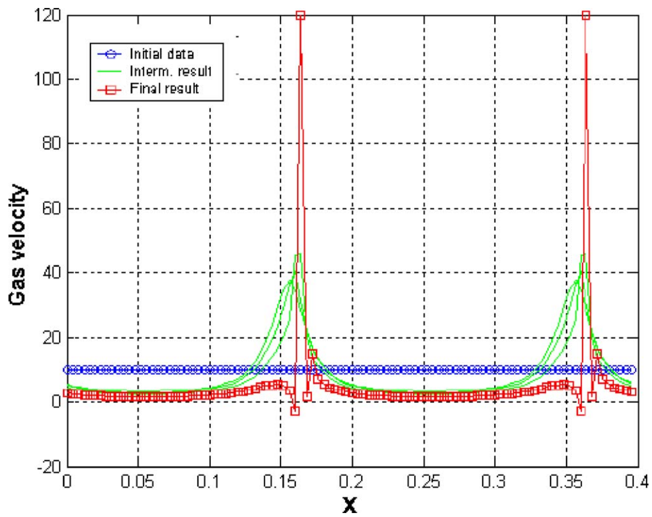


Fig. 10 The wave traveling problem. Gas velocity without preconditioning at four time steps of 0.001 s.

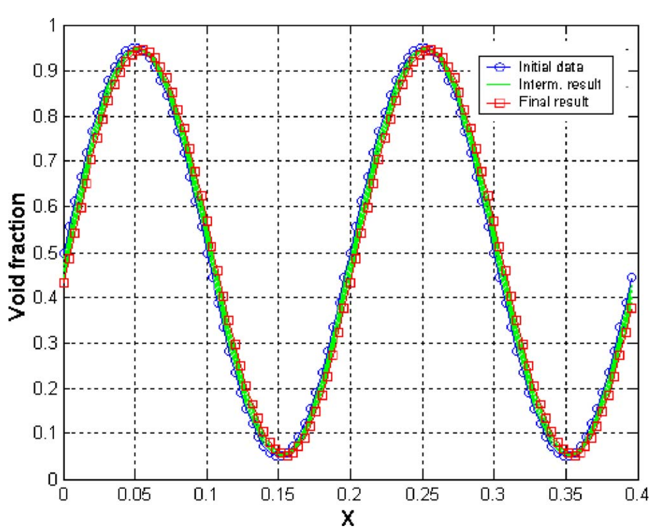


Fig. 11 The wave traveling problem. Void fraction with preconditioning at five time steps of 0.001 s.

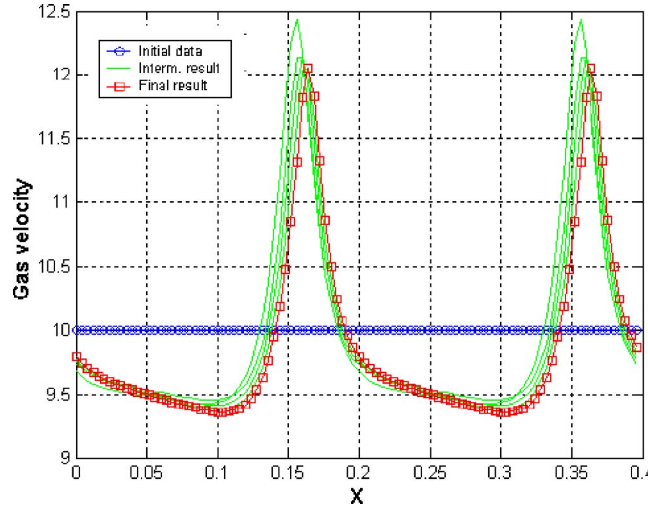


Fig. 12 The wave traveling problem. Gas velocity with preconditioning at five time steps of 0.001 s.

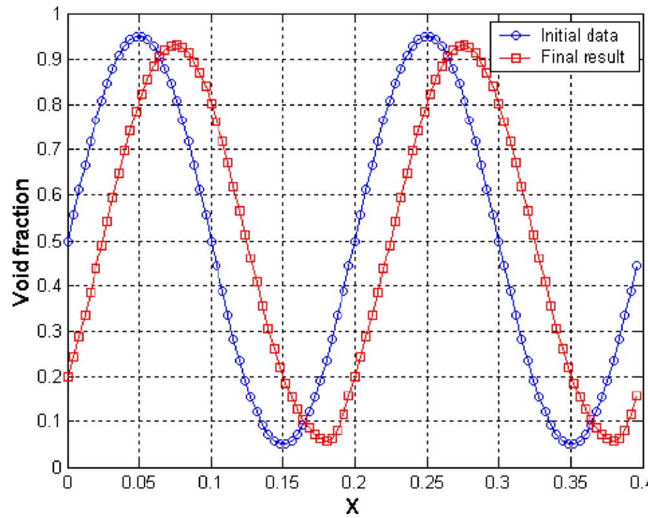


Fig. 13 The wave traveling problem. Void fraction with preconditioning at 25 time steps of 0.001 s.

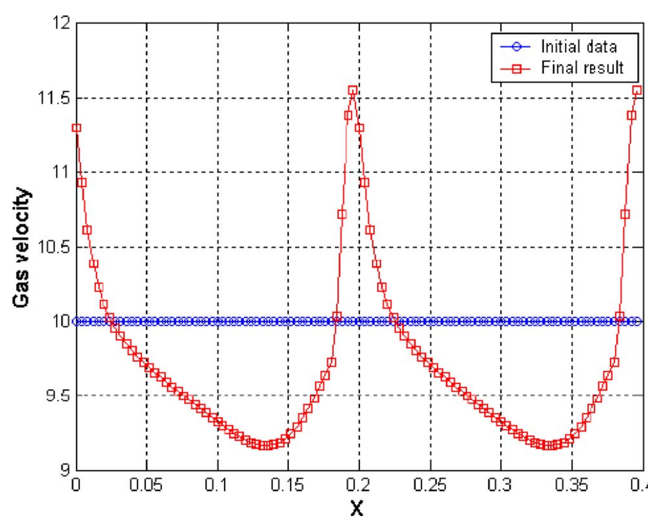


Fig. 14 The wave traveling problem. Gas velocity with preconditioning at 25 time steps of 0.001 s.

propagate freely through the domain. As in the water faucet problem, no interaction between phases is considered in this case. Although the initial conditions can seem rather extreme, these conditions (or even worse ones) appear for example in the steelmaking industry when argon is injected at the bottom of the ladle to produce stirring in the liquid steel.

Figure 9 shows the void fraction obtained *without* preconditioning just before the numerical implementation becomes unstable at five time steps of 0.001 s. This instability is more evident in Fig. 10, where the gas velocity is plotted. Figures 11 and 12 were obtained *with* preconditioning and they correspond to five time steps of 0.001 s. As can be observed, using the preconditioning method the solution does not diverge and is smooth. This is not surprising since now we are working with a well-posed system of differential equations. The simulation can be continued and no sign of instability is found as can be seen in Figs. 13 and 14 where we show the same curves after 25 time steps of 0.001. It is also worth noting that the same problem has been solved with the commercial code CFX [45], and the solution diverges after seven time steps of 0.001 s.

6 Conclusions

We have developed a preconditioning method for the mass matrix of the two-fluid model in order to make the model well posed. The preconditioning matrix has two parameters: the first one β is for making the model hyperbolic, and the second one γ is for recovering the wave velocity propagation corresponding to the original model. Both parameters depend on the void fraction and on velocity and density relations between phases.

The characteristic values obtained with the preconditioning method do not show any imaginary component and share the same real part with the original model. To test the proposed method we have modified a well-established numerical scheme to include the preconditioning. Two examples were analyzed using $\gamma=1$. One of them, the so-called water faucet problem, is a well known benchmark widely used for two-phase flows. Using the preconditioning method, the numerically obtained steady-state solution agrees with the analytical solution. Besides, although the time evolution cannot be guaranteed using $\gamma=1$, the method reproduces transient solutions very well. The second example was chosen in order to evidence numerically the ill-posed nature of the two-fluid model. This example, a large amplitude traveling wave, allowed us to see how the ill posedness is reflected as instabilities in the numerical simulation and shows that, as expected, these instabilities do not appear when the differential equations are made well posed by means of the preconditioning matrix.

Acknowledgment

This work has received financial support from Consejo Nacional de Investigaciones Científícas y Técnicas (CONICET, Argentina, Grant Nos. PIP 02552/00 and PIP 5271/05); Universidad Nacional del Litoral (UNL, Argentina, Grant No. CAI+D 2005-10-64) and Agencia Nacional de Promoción Científíca y Tecnológica (ANPCyT, Argentina, Grant Nos. PICT 12-14573/2003 and PME 209/2003).

The authors made extensive use of freely distributed software as GNU/Linux OS, MPI, PETSc, GCC compilers, Octave, OpenDX, Perl, Python, among many others. Also, many ideas from these packages have been inspirational to them.

References

[1] Ishii, M., 1975, *Thermofluid Dynamic Theory of Two-Phase Flow*, Eyrolles, Paris.

[2] Hirt, C. W., and Nichols, B. D., 1981, "Volume of Fluid (VOF) Method for the Dynamics of Free Boundaries," J. Comput. Phys., **39**, pp. 201–225.

[3] Kothe, D. B., and Mjolsness, R. C., 1992, "RIPPLE: A New Model for Incompressible Flows With Free Surfaces," AIAA J., **30**, pp. 2694–2700.

[4] Kothe, D. B., Rider, W. J., Mosso, S. J., Brock, J. I., and Hochstein, J. S., 1996, "Volume Tracking of Interfaces Having Surface Tension in Two and Three Dimensions," Technical Report AIAA 96-0859, AIAA.

[5] Richards, J. R., Beris, A. N., and Lenho, A. M., 1995, "Drop Formation in Liquid-Liquid Systems Before and After Jetting," Phys. Fluids, **7**, pp. 2617–2630.

[6] Rider, W. J., Kothe, D. B., Mosso, S. J., Cerutti, J., and Hochstein, J. S., 1995, "Volume Tracking of Interfaces Having Surface Tension in Two and Three Dimensions," Technical Report AIAA 95-0699, AIAA.

[7] Rider, W. J., and Kothe, D. B., 1998, "Reconstructing Volume Tracking," J. Comput. Phys., **141**, pp. 112–152.

[8] Juric, D., and Tryggvason, G., 1998, "Computations of Boiling Flows," Int. J. Multiphase Flow, **24**, pp. 387–410.

[9] Unverdi, S. O., and Tryggvason, G., 1992, "A Front-Tracking Method for Viscous Incompressible Multi-Fluid Flows," J. Comput. Phys., **100**, pp. 25–37.

[10] Fedkiw, R. P., Aslam, T., Merriman, B., and Osher, S., 1999, "A Non-Oscillatory Eulerian Approach to Interfaces in Multimaterial Flows (the Ghost Fluid Method)," J. Comput. Phys., **152**(2) pp. 457–492.

[11] Osher, S., and Sethian, J. A., 1988, "Fronts Propagating With Curvature Dependent Speed: Algorithm Based on Hamilton-Jacobi Formulations," J. Comput. Phys., **79**, pp. 12–49.

[12] Osher, S., and Fedkiw, R. P., 2001, "Level Set Methods: An Overview and Some Recent Results," J. Comput. Phys., **169**, pp. 463–502.

[13] Sussman, M., and Smereka, P., 1997, "Axisymmetric Free Boundary Problems," J. Fluid Mech., **341**, pp. 269–294.

[14] Anderson, D. M., McFadden, G. B., and Wheeler, A. A., 1998, "Diffuse Interface Methods in Fluid Mechanics," Annu. Rev. Fluid Mech., **30**, pp. 139–165.

[15] Jacqmin, D., 1999, "Calculation of Two-Phase Navier-Stokes Flows Using Phase-Field Modeling," J. Comput. Phys., **155**, pp. 96–127.

[16] Lowengrub, J., and Truskinovsky, L., 1998, "Quasi-Incompressible Cahn-Hilliard Fluids and Topological Transitions," Proc. R. Soc. London, Ser. A, **454**, pp. 2617–2654.

[17] Harlow, F. H., and Welch, J. E., 1965, "Numerical Calculation of Time-Dependent Viscous Incompressible Flow of Fluids With A Free Surface," Phys. Fluids, **8**, pp. 2182–2189.

[18] Drew, D. A., and Passman, S. L., 1998, *Theory of Multicomponent Fluids*, Applied Mathematical Sciences, Springer, New York, Vol. 135.

[19] Keyfitz, B. L., 2000, "Mathematical Properties of Nonhyperbolic Models for Incompressible Two-Phase Flow," *Proceedings of the Fourth International Conference on Multiphase Flow*, New Orleans (report).

[20] Joseph, D. D., and Saut, J. C., 1990, "Short Wave Instabilities and Ill Posed Initial Value Problems," Journal of Theoretical and Computational Fluid Mechanics, **1**, pp. 191–227.

[21] Méndez, C. G., Nigro, N., and Cardona, A., 2005, "Drag and Non-Drag Force Influences in Numerical Simulations of Metallurgical Ladles," J. Mater. Process. Technol., **160**, pp. 296–305.

[22] Lahey, R. T., Jr., Yin, J., and Tiwari, P., 2005, "Analysis of Void Wave Propagation and Sonic Velocity Using a Two-Fluid Model," Multiphase Sci. Technol., **17**, pp. 293–320.

[23] Chung, M., Chang, K., and Lee, S., 2002, "Numerical Solution of Hyperbolic Two-Fluid Two-Phase Flow Model With Non-Reflecting Boundary Conditions," Int. J. Eng. Sci., **40**, pp. 789–803.

[24] Micaelli, J. C., 1988, *Document Syntheses CATHARE* (CATHARE An Advanced Best-Estimate Code for PWR Safety Analysis), SETh/LEML-EM/88-129, Grenoble Cedex.

[25] Bestion, D., 1990, "The Physical Closure Laws in CATHARE Code," Nucl. Eng. Des., **124**, 229–245.

[26] Coquel, F., Gallouët, T., Hérard, J., and Seguin, N., 2002, "Closure Laws for a Two-Fluid Two-Pressure Model," C. R. Acad. Sci., Ser. I: Math., **334**, pp. 927–932.

[27] Lee, S., Chang, K., and Kim, K., 1998, "Pressure Wave Speeds From the Characteristics of Two Fluids, Two-Phase Hyperbolic Equation System," Int. J. Multiphase Flow, **24**, pp. 855–866.

[28] Pokharna, H., Mori, M., and Ransom, V. H., 1997, "Regularization of Two-Phase Flow Models: A Comparison of Numerical and Differential Approaches," J. Comput. Phys., **134**, pp. 282–295.

[29] Städtke, H., Franchello, G., and Worth, B., 1997, "Numerical Simulation of Multi-Dimensional Two-Phase Flow Based on Flux Vector Splitting," Nucl. Eng. Des., **177**, pp. 199–213.

[30] Städtke, H., and Holtbecker, R., 1992, "Hyperbolic Model for One-Dimensional Inhomogeneous Two-Phase Flow," *Proceedings 29th Meeting of European Two-Phase Flow Group*, Stockholm, Sweden, June 2–3.

[31] Anderson, A., and York, J., 1999, "Fixing Einstein's Equations," Phys. Rev. Lett., **82**, pp. 4384–4387.

[32] Reula, O., 2004, "Strongly Hyperbolic Systems," *General Relativity and Quantum Cosmology*, e-print 0403007.

[33] Gundlach, C., and Martin-García, J., 2005, "Hyperbolicity of Second Order in Space Systems of Evolution Equations," *General Relativity and Quantum Cosmology*, e-print 0506037.

[34] Jones, A. V., and Prosperetti, A., 1985, "On the Stability of First-Order Differential Models for Two-Phase Flow Prediction," Int. J. Multiphase Flow, **11**, pp. 133–148.

[35] Zhang, D. Z., and Prosperetti, A., 1994, "Ensemble Phase Averaged Equations for Bubbly Flows," Phys. Fluids, **6**, pp. 2956–2970.

[36] Lahey, R. T., Jr., and Drew, D. A., 1988, "The Three-Dimensional Time and Volume Averaged Conservation Equations of Two-Phase Flows," Adv. Nucl. Sci. Technol, **20**, pp. 1–69.

[37] Ishii, M., and Zuber, N., 1979, "Relative Motion and Interfacial Drag Coeffi-

cient in Dispersed Two-Phase Flow of Bubbles, Drops and Particles," AIChE J., **25**, pp. 843–855.
[38] Chorin, A. J., 1967, "A Numerical Method for Solving Incompressible Viscous Problems," J. Comput. Phys., **2**, pp. 12–26.
[39] Turkel, E., 1999, "Preconditioning Techniques in Computational Fluid Dynamics," Annu. Rev. Fluid Mech., **31**, pp. 385–416.
[40] Ransom, V. H., et al., 1985, *RELAP5/MOD2 Code Manual, 1, Code 2, Structure System Models, and Solution Methods*, NUREG/CR-4312.
[41] Shieh, A. S., Ransom, V. H., and Krishnamurthy, R., 1994, *RELAP5/MOD3 Code Manual, Validation of Numerical Techniques in RELAP5/MOD3.0*, NUREG/CR-5535, Vol. 6.
[42] Coquel, F., Amine, K. El., Godlewski, E., Perthame, B., and Rascle, P., 1997, "A Numerical Method Using Upwind Schemes for the Resolution of Two-Phase Flows," J. Comput. Phys., **136**, pp. 272–288.
[43] Niu, Y. Y., 2001, "Advection Upwinding Splitting Method to Solve a Compressible Two-Fluid Model," Int. J. Numer. Methods Fluids, **36**, pp. 351–371.
[44] Trapp, J. A., and Riemke, R. A., 1986, "A Nearly-Implicit Hydrodynamic Numerical Scheme for Two-Phase Flows," J. Comput. Phys., **66**, pp. 62–82.
[45] AEA Technology, 2004, *User Manual to CFX Version 5.7.*, AEA Technology plc, Harwell, UK.

D. Bigoni[1]
Dipartimento di Ingegneria Meccanica e Strutturale,
Università di Trento,
Via Mesiano 77,
38050 Povo, Trento, Italy
e-mail: bigoni@ing.unitn.it

W. J. Drugan
Dipartimento di Ingegneria Meccanica e Strutturale,
Università di Trento,
Via Mesiano 77,
38050 Povo, Trento, Italy
and
Department of Engineering Physics,
University of Wisconsin-Madison,
1500 Engineering Drive,
Madison, WI 53706-1687
e-mail: drugan@engr.wisc.edu

Analytical Derivation of Cosserat Moduli via Homogenization of Heterogeneous Elastic Materials

Why do experiments detect Cosserat-elastic effects for porous, but not for stiff-particle-reinforced, materials? Does homogenization of a heterogeneous Cauchy-elastic material lead to micropolar (Cosserat) effects, and if so, is this true for every type of heterogeneity? Can homogenization determine micropolar elastic constants? If so, is the homogeneous (effective) Cosserat material determined in this way a more accurate representation of composite material response than the usual effective Cauchy material? Direct answers to these questions are provided in this paper for both two- (2D) and three-dimensional (3D) deformations, wherein we derive closed-form formulae for Cosserat moduli via homogenization of a dilute suspension of elastic spherical inclusions in 3D (and circular cylindrical inclusions in 2D) embedded in an isotropic elastic matrix. It is shown that the characteristic length for a homogeneous Cosserat material that best mimics the heterogeneous Cauchy material can be derived (resulting in surprisingly simple formulae) when the inclusions are less stiff than the matrix, but when these are equal to or stiffer than the matrix, Cosserat effects are shown to be excluded. These analytical results explain published experimental findings, correct, resolve and extend prior contradictory theoretical (mainly numerical and limited to two-dimensional deformations) investigations, and provide both a general methodology and specific results for determination of simple higher-order homogeneous effective materials that more accurately represent heterogeneous material response under general loading conditions. In particular, it is shown that no standard (Cauchy) homogenized material can accurately represent the response of a heterogeneous material subjected to a uniform plus linearly varying applied traction, while a homogenized Cosserat material can *do so (when inclusions are less stiff than the matrix).* [DOI: 10.1115/1.2711225]

Keywords: homogenization, Cosserat-elasticity, dilute suspension of elastic spheres, nonlocal constitutive equations, micropolar effects

1 Introduction

There is a long-standing debate in the solid mechanics community concerning the possibility of predicting micropolar elastic (Cosserat) behavior from Cauchy-elastic materials containing inhomogeneities or microstructures. In fact, although the motivation leading to Cosserat effects seems to be very intuitive, theoretical results in the literature are often contradictory and no definitive conclusion is available (see Appendix A for details). Moreover, experimental results support Cosserat effects for porous materials (like bone or foam [1–5]), but find an absence of these effects for reinforced materials [6,7].

In the present paper we provide a general methodology for the determination of the moduli for a homogeneous Cosserat-elastic material that best approximates a heterogeneous Cauchy-elastic material. We apply this methodology to the specific cases of three-dimensional (3D) deformations of a dilute suspension of (Cauchy, linear, and isotropic) elastic spherical inclusions, and two-dimensional (2D) deformations of circular cylindrical inclusions, in a (Cauchy, linear, and isotropic) elastic matrix. With reference to a Cosserat (linear and isotropic) material, it is shown that:

1. Cosserat effects are predicted for spherical or cylindrical inclusions less stiff than the matrix, but are excluded for inclusions having stiffness equal to or greater than that of the matrix;
2. simple, closed-form formulae give the Cosserat characteristic length (and the other effective Cosserat moduli) as a function of the inclusion radius, volume fraction, and the elastic contrast of the constituent phases; and
3. the characteristic length that results for three-dimensional deformations of a matrix with spherical inclusions is significantly smaller than that resulting for two-dimensional deformations of a matrix with circular cylindrical inclusions.

Conclusion (1) rigorously explains experimental evidence demonstrating micropolar effects for porous material, but displaying an opposite trend, or in the words of Gauthier [7] "an anti-micropolar phenomenon," for inclusions stiffer than the matrix.

A closely related issue is that standard homogenization results for linear elastic materials provide overall or effective elastic moduli that relate (*uniform*) average stress to (*uniform*) average strain. This means that standard homogenization results give a homogeneous "effective" material that is able to represent well the overall response of the actual heterogeneous elastic material *when the applied loading is uniform*. However, when the applied loading deviates from uniformity, the homogeneous "effective" material less accurately represents the overall response of the actual heterogeneous material. This fact is important, since of course in general composite materials are employed in applications where the applied loading is not uniform.

We show in this paper, for situations in which the applied loading on a heterogeneous material varies sufficiently slowly that it admits a Taylor series expansion, that whereas the standard homogenization results provide a homogeneous "effective" material that can accurately represent the actual heterogeneous one only

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received February 10, 2006; final manuscript received April 20, 2006. Review conducted by Robert M. McMeeking.

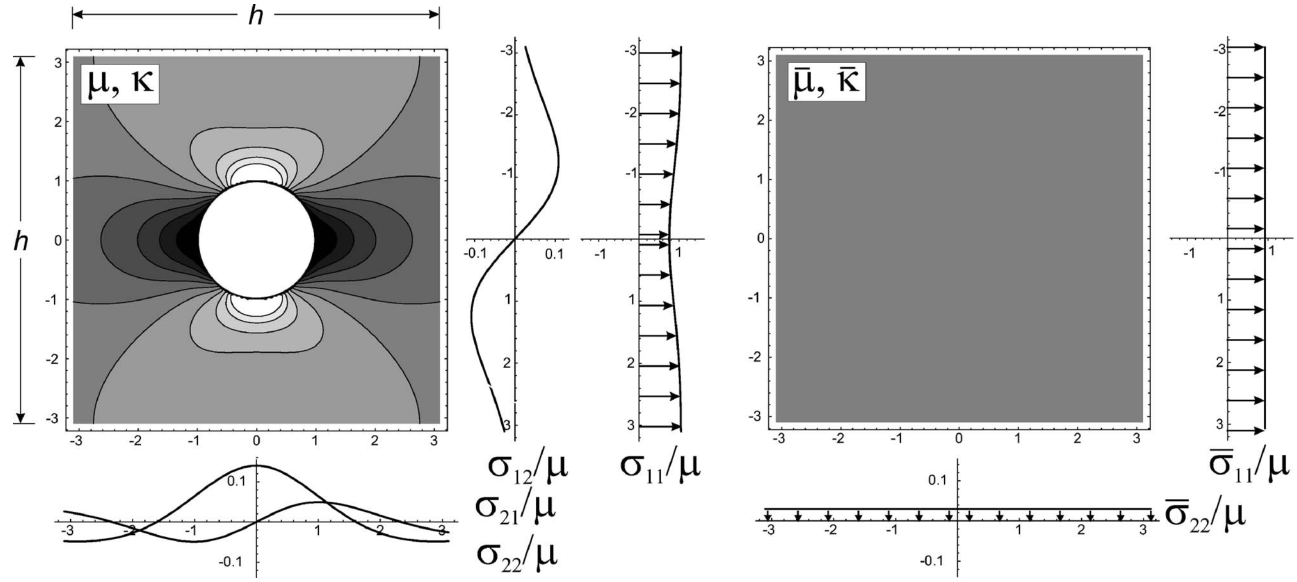


Fig. 1 Procedure of homogenization of a material containing a dilute distribution of circular voids. Heterogeneous material (left) is an $h \times h$ prism removed from an infinite sheet that is subjected to uniform, uniaxial far-field stress; homogeneous material (right) is subjected to the mean stresses calculated from the heterogeneous prism. For the plane strain problem, level sets of σ_{11} are shown; note that the values of σ_{12}, σ_{21}, and σ_{22}, shown parallel to the edges, are less than 1/10 the maximum value of σ_{11} at a distance from inclusion center equal to three times the radius of the inclusion.

when the leading-order (uniform) term in the Taylor series is retained, a homogeneous "effective" Cosserat material can do so when two terms in the Taylor series are retained, when the material heterogeneities are less stiff than the matrix material. The result is a simple homogeneous material model that more accurately represents actual (compliant-inclusion-type) heterogeneous material response under slowly varying applied loading.

2 Review of Homogenization Results for Uniform Applied Loading

Here we briefly summarize well-known results for the effective moduli of a homogeneous, isotropic linear elastic matrix containing a dilute suspension of homogeneous, isotropic linear elastic inclusions having in general different moduli than the matrix; the inclusions are either cylinders (for plane strain deformations) or spheres (for three-dimensional deformations). As noted in Sec. 1, one approach for deriving such moduli is to require that they relate average (uniform) stress and strain in the same way that these quantities are related in the actual heterogeneous material. An alternative, equivalent approach for their derivation is to require that the total elastic energy in the uniform "effective" medium equals that in the actual heterogeneous medium under uniform applied loading. We will employ this energy approach in the present work.

When the composite is dilute, as considered here, we may employ the solution for an infinite body containing a single inclusion and subjected to uniform far-field loading. From this solution, we select a finite region containing the inclusion, and calculate the mean stresses acting on it. The effective moduli may then be calculated by equating, through first order in volume fraction, the elastic energy contained in the selected finite region calculated from the actual heterogeneous material solution with that calculated from a homogeneous effective body of the same size subjected to the mean stresses calculated from the infinite-body solution. We define the effective shear modulus as $\bar{\mu}$ and bulk modulus as $\bar{\kappa}$ (for 3D, whereas $\bar{\kappa}=3-4\bar{\nu}$, with $\bar{\nu}$ denoting in-plane Poisson's ratio for plane strain 2D). A sketch of this procedure is shown in Fig. 1, for plane strain deformation of an infinite plane with a circular hole.

Eshelby [8] and independently Hashin [9] have obtained the following effective elastic moduli for the three-dimensional problem of a matrix containing spherical inclusions (here retaining terms through first order in the volume fraction f of the inclusion phase)

$$\bar{\mu} = \mu_m + f\frac{5\mu_m(\mu_i-\mu_m)(3\kappa_m+4\mu_m)}{2(\mu_i+\mu_m)(3\kappa_m+4\mu_m)+\mu_m(3\kappa_m+4\mu_i)}$$

$$\bar{\kappa} = \kappa_m + f(\kappa_i-\kappa_m)\frac{3\kappa_m+4\mu_m}{3\kappa_i+4\mu_m} \tag{1}$$

where subscripts m and i denote matrix and inclusion, respectively.

In two-dimensional (plane strain) elasticity, the spheres are replaced by parallel infinite circular cylinders and the effective-modulus formulae through $O(f)$ are [10]

$$\bar{\mu} = \mu_m + f(1+\kappa_m)\mu_m\frac{\mu_i-\mu_m}{\kappa_m\mu_i+\mu_m}$$

$$\bar{\kappa} = \kappa_m + f(1+\kappa_m)\left[(\kappa_m-1)\frac{\mu_i-\mu_m}{\kappa_m\mu_i+\mu_m} - \frac{(\kappa_m-1)\mu_i-(\kappa_i-1)\mu_m}{2\mu_i+(\kappa_i-1)\mu_m}\right] \tag{2}$$

where now $\kappa=3-4\nu$, with ν denoting (in-plane) Poisson's ratio.

3 Homogenization Under Nonuniform Applied Loading

3.1 Taylor Series Representation of Slowly Varying Applied Loading. Let us consider an infinite body of composite material with a *dilute* distribution of inclusions, subjected to arbitrary but slowly varying far-field ("boundary") conditions. The far-field displacement field $\mathbf{u}(\mathbf{x})$ can then be expanded in a Taylor series about the location of the center of an inclusion (chosen as the origin of coordinates). Through second order, the most general representation for this is

$$u_i = \alpha_{ij}x_j + \beta_{ijk}x_jx_k \tag{3}$$

where α_{ij} and β_{ijk} are constant coefficients, the latter having the obvious symmetry $\beta_{ijk}=\beta_{ikj}$ (since x_j and x_k play the same role), indices range between 1 and 3 (1 and 2 for plane strain), and the usual summation convention for repeated indices is employed here and throughout the paper except where noted. Although coefficients α_{ij} are unrestricted, the quadratic part of the displacement field must satisfy the Navier equations of equilibrium without body forces, resulting in the following three (two for plane strain) restrictions

$$\beta_{kki} = -(1-2\nu_m)\beta_{ikk} \tag{4}$$

As is well known, the homogeneous effective Cauchy-elastic material, Eqs. (1) and (2), accurately mimics the response of a heterogeneous Cauchy material when this is subjected to a linearly varying displacement (uniform applied loading). However, in most practical situations, a composite material is subjected to a spatially varying applied loading. How well does the homogeneous effective Cauchy material mimic the actual heterogeneous one in this case, and can a homogeneous Cosserat material do better? Let us consider plane strain and three-dimensional deformations separately.

3.2 Plane Strain. Employing the constraint Eq. (4) and explicitly exhibiting the plane-strain bending contributions, the quadratic terms in the remote displacement field Eq. (3) become

$$u_1 = \left[\tilde{\beta}_{13} - \frac{\nu_m}{2(1-\nu_m)R_{23}}\right]x_1^2 + \frac{x_1x_2}{R_{13}} - \left(2\frac{1-\nu_m}{1-2\nu_m}\tilde{\beta}_{13} + \frac{1}{2R_{23}}\right)x_2^2$$

$$u_2 = \left[\tilde{\beta}_{23} - \frac{\nu_m}{2(1-\nu_m)R_{13}}\right]x_2^2 + \frac{x_1x_2}{R_{23}} - \left(2\frac{1-\nu_m}{1-2\nu_m}\tilde{\beta}_{23} + \frac{1}{2R_{13}}\right)x_1^2,$$

$$u_3 = 0 \tag{5}$$

where coefficients $\tilde{\beta}_{13}$, $\tilde{\beta}_{23}$ (index 3 denotes the out-of-plane direction and the others the nonnull displacement component directions) and bending curvatures R_{13} and R_{23} (index 3 again denotes the out-of-plane direction, while the other indices denote the directions of the normal components of bending stress) are arbitrary. Displacements Eqs. (5) a priori satisfy the Navier equations and thus represent the most general equilibrium plane-strain quadratic displacement field.

The problem of an infinite sheet containing a circular hole and subjected to far-field bending was solved by Muskhelishvili [11], and by Sendeckyj [12] in the general case of a circular elastic inclusion. The elastic fields produced by the far-field loading modes associated with $\tilde{\beta}_{13}$ and $\tilde{\beta}_{23}$ in an infinite sheet containing a circular elastic inclusion are determined in Appendix B (where the bending solution is also included for completeness). The important point with respect to our upcoming accurate modeling of effective material response is that *these solutions show* that the displacement field Eq. (5), valid exactly for a homogeneous material, is perturbed by the inclusion, in the material outside the inclusion, only by terms of $O(f^2)$.

3.3 Three-Dimensional Deformations. The most general quadratic equilibrium remote displacement field can be written as, using Eq. (3) with Eq. (4) (summation *not* implied for repeated indices)

$$u_i = \frac{x_ix_j}{R_{ik}} + \frac{x_ix_k}{R_{ij}} - \frac{1}{2R_{jk}}\left(x_j^2 + \frac{\nu_m}{1-\nu_m}x_i^2\right) - \frac{1}{2R_{kj}}\left(x_k^2 + \frac{\nu_m}{1-\nu_m}x_i^2\right)$$

$$+ (\Theta_j - \Theta_k)x_jx_k + (\tilde{\beta}_{ik} + \tilde{\beta}_{ij})x_i^2 - 2\frac{1-\nu_m}{1-2\nu_m}(\tilde{\beta}_{ik}x_j^2 + \tilde{\beta}_{ij}x_k^2) \tag{6}$$

where indices i, j, k are cyclic permutations of 1, 2, 3 (i.e., 1,2,3; 2,3,1; 3,1,2), illustrating the fact that the kinematics are the sum of six plane strain modes (defined by bending curvatures R_{ij} and additional free coefficients $\tilde{\beta}_{ij}$ (where i denotes the direction of the bending stress or non-null displacement component and j the out-of-plane direction)) and three torsional angles of twist/length Θ_i ($i=1,2,3$). Therefore, the plane-strain displacement field Eq. (5) can be obtained from Eq. (6) by taking $1/R_{12}=1/R_{21}=1/R_{32}=1/R_{31}=\tilde{\beta}_{12}=\tilde{\beta}_{21}=\tilde{\beta}_{32}=\tilde{\beta}_{31}=\Theta_1=\Theta_2=\Theta_3=0$.

The problem of an infinite elastic matrix containing a spherical void and subjected to remote bending loading (a particular case of Eq. (6) in which all $\tilde{\beta}_{ij}$ and Θ_i are zero) has been solved by Sen [13], and by Das [14] for the general case of a spherical elastic inclusion. These solutions show that the bending displacement field, valid exactly for a homogeneous material, is perturbed by the inclusion in the region outside the inclusion by terms of $O(f^{5/3})$. The fact that the perturbation remains at $O(f^{5/3})$ for the general quadratic displacement field Eq. (6) is shown in Appendix C, where the solution for a spherical elastic inclusion in an infinite elastic matrix, subject to the remote displacement field Eq. (6) is obtained. Appendix C also shows that the Das [14] solution is incomplete, and that it can be expressed purely in terms of simple functions.

3.4 Conclusion. The *quadratic* part of the displacement field Eq. (3), together with equilibrium requirements Eq. (4), which is valid exactly for a homogeneous material, is perturbed by a cylindrical or spherical inclusion in the region outside the inclusion by terms of $O(f^2)$ for two-dimensional elasticity and $O(f^{5/3})$ for three-dimensional elasticity.

In other words, in an asymptotic expansion in inclusion volume fraction f of the displacement field solution outside the inclusion, through order f the inclusion is neutral under remote quadratic displacement conditions. Therefore, the effective moduli determined under the remote quadratic displacement conditions are identical (to first order in f) with the moduli of the matrix material.

4 Standard Homogenized Material is in Error for Quadratic Applied Displacements

Now we are in a position to face the main problem, namely: under an applied *linear* remote displacement field (*uniform* applied remote stress) the perturbation induced by the inclusion in the displacement field solution in the surrounding matrix material is $O(f)$, while under an applied *quadratic* remote displacement field (applied *linear* remote stress field) the perturbation in the displacement field solution in the matrix becomes $O(f^{5/3})$ for 3D and $O(f^2)$ for 2D elasticity.

Therefore, *the effective material defined by Eqs. (1) and (2) is stiffer (more compliant) for linearly varying applied loading than the actual heterogeneous material for inclusions stiffer (more compliant) than the matrix.* That is, if the heterogeneous material (matrix with inclusion) is represented in the usual way in composite materials theory—by a homogeneous material with effective moduli given by Eqs. (1) or (2)—this representation works well for uniform applied loading, but for linearly varying applied stress, it is in error by terms of $O(f)$.

To better elucidate this point, let us consider a cube of edge h, composed of a homogeneous effective material having properties Eqs. (1) or (2) and subject to the quadratic displacement field Eqs. (5) or (6). The total elastic energy in such a cube is obtained by calculating the strain energy density from Eqs. (5) or (6) and then integrating this over the cube.

The total elastic energy in the cube, $\mathcal{E}$, is, for plane strain

$$
\begin{aligned}
\mathcal{E}_{\text{Cauchy}} &= \frac{h^5\bar{\mu}}{12(1-\bar{\nu})}\left(\frac{1}{R_{13}^2}+\frac{1}{R_{23}^2}\right)+\frac{h^5\bar{\mu}(1-\bar{\nu})(3-4\bar{\nu})}{3(1-2\bar{\nu})^2}(\tilde{\beta}_{13}^2+\tilde{\beta}_{23}^2)\\
&= h^5\mu_m\left(\frac{1}{R_{13}^2}+\frac{1}{R_{23}^2}\right)\left(\frac{1}{12(1-\nu_m)}-\Xi_R f\right)+h^5\mu_m(\tilde{\beta}_{13}^2+\tilde{\beta}_{23}^2)\left[\frac{(1-\nu_m)(3-4\nu_m)}{3(1-2\nu_m)^2}-\Xi_\beta f\right]+O(f^2)
\end{aligned}
\tag{7}
$$

where

$$
\Xi_R = \frac{1}{12(1-\nu_m)}\left[\frac{3\mu_m+\mu_i(1-4\nu_m)}{\mu_m+\mu_i(3-4\nu_m)}-\frac{2\mu_i(1-\nu_m)}{\mu_i+\mu_m(1-2\nu_i)}\right]
$$

$$
\begin{aligned}
\Xi_\beta = -\frac{(1-\nu_m)(3-4\nu_m)}{3(1-2\nu_m)^2}\Bigg(&\frac{2\mu_i(1-\nu_m)(5-6\nu_m)}{(1-2\nu_m)(3-4\nu_m)[\mu_i+\mu_m(1-2\nu_i)]}\\
&+\frac{\mu_i\{-13+2\nu_m[9+4(3-4\nu_m)\nu_m]\}+\mu_m\{-7+2\nu_m[13-8\nu_m(3-2\nu_m)]\}}{(1-2\nu_m)(3-4\nu_m)[\mu_m+\mu_i(3-4\nu_m)]}\Bigg)
\end{aligned}
\tag{8}
$$

while for three-dimensional deformation it is

$$
\begin{aligned}
\mathcal{E}_{\text{Cauchy}} &= \frac{h^5\bar{\mu}}{12(1-\bar{\nu})}\sum_{\substack{i,j=1\\ i\neq j}}^{3}\left(\frac{1}{R_{ij}^2}+\frac{\bar{\nu}}{R_{ij}R_{ji}}\right)+\frac{h^5\bar{\mu}}{12}\left(\sum_{i=1}^{3}\Theta_i^2-\frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{3}\Theta_i\Theta_j\right)+\frac{2h^5\bar{\mu}(1-\bar{\nu})}{3(1-2\bar{\nu})}\left[\tilde{\beta}_{12}\tilde{\beta}_{13}+\tilde{\beta}_{21}\tilde{\beta}_{23}+\tilde{\beta}_{31}\tilde{\beta}_{32}+\frac{3-4\bar{\nu}}{2(1-2\bar{\nu})}\sum_{\substack{i,j=1\\ i\neq j}}^{3}\tilde{\beta}_{ij}^2\right]\\
&= h^5\mu_m\sum_{\substack{i,j=1\\ i\neq j}}^{3}\frac{1}{R_{ij}^2}\left[\frac{1}{12(1-\nu_m)}-\Xi_R f\right]+h^5\mu_m\sum_{\substack{i,j=1\\ i\neq j}}^{3}\frac{1}{R_{ij}R_{ji}}\left[\frac{\nu_m}{12(1-\nu_m)}-\Xi_{RR} f\right]+h^5\mu_m\left(\sum_{i=1}^{3}\Theta_i^2-\frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{3}\Theta_i\Theta_j\right)\left(\frac{1}{12}-\Xi_\Theta f\right)\\
&\quad+h^5\mu_m\sum_{\substack{i,j=1\\ i\neq j}}^{3}\tilde{\beta}_{ij}^2\left[\frac{(1-\nu_m)(3-4\nu_m)}{3(1-2\nu_m)^2}-\Xi_\beta f\right]+h^5\mu_m(\tilde{\beta}_{12}\tilde{\beta}_{13}+\tilde{\beta}_{21}\tilde{\beta}_{23}+\tilde{\beta}_{31}\tilde{\beta}_{32})\left[\frac{2(1-\nu_m)}{3(1-2\nu_m)}-\Xi_{\beta\beta} f\right]+O(f^{5/3})
\end{aligned}
\tag{9}
$$

where

$$
\Xi_R = \frac{1}{2(1-\nu_m)}-\frac{28\mu_i^2+34\mu_i\mu_m+13\mu_m^2}{4(2\mu_i+\mu_m)[\mu_m(7-5\nu_m)+2\mu_i(4-5\nu_m)]}-\frac{\mu_m(2\mu_i-\mu_m)(1-2\nu_i)+2\mu_i^2(1+\nu_i)}{4(2\mu_i+\mu_m)[\mu_i(1+\nu_i)+2\mu_m(1-2\nu_i)]}
$$

$$
\Xi_{RR} = \frac{1}{2(1-\nu_m)}-\frac{26\mu_i^2+38\mu_i\mu_m+11\mu_m^2}{4(2\mu_i+\mu_m)[\mu_m(7-5\nu_m)+2\mu_i(4-5\nu_m)]}-\frac{(\mu_i+\mu_m)\mu_i(1+\nu_i)+\mu_m^2(1-2\nu_i)}{4(2\mu_i+\mu_m)[\mu_i(1+\nu_i)+2\mu_m(1-2\nu_i)]}
$$

$$
\Xi_\Theta = \frac{5(\mu_m-\mu_i)(1-\nu_m)}{4\mu_m(7-5\nu_m)+8\mu_i(4-5\nu_m)}
$$

$$
\begin{aligned}
\Xi_{\beta\beta} = -&\frac{2(1-\nu_m)}{[2\mu_m(1-2\nu_i)+\mu_i(1+\nu_i)](1-2\nu_m)^2[\mu_m(7-5\nu_m)+\mu_i(8-10\nu_m)]}\{2\mu_i^2(1+\nu_i)(1-2\nu_m)(3-5\nu_m)+\mu_i\mu_m(1-5\nu_m)[3-4\nu_m\\
&-\nu_i(9-14\nu_m)]-\mu_m^2(1-2\nu_i)[9-\nu_m(26-25\nu_m)]\}
\end{aligned}
$$

$$
\begin{aligned}
\Xi_\beta = &\frac{(1-\nu_m)}{[\mu_m(7-5\nu_m)+2\mu_i(4-5\nu_m)](1-2\nu_m)^3[2\mu_m(1-2\nu_i)+\mu_i(1+\nu_i)]}\{2\mu_i^2(1+\nu_i)(-10+53\nu_m-91\nu_m^2+50\nu_m^3)+\mu_m^2(1-2\nu_i)(25\\
&-104\nu_m+181\nu_m^2-110\nu_m^3)-\mu_i\mu_m[5-73\nu_m+164\nu_m^2-100\nu_m^3+\nu_i(5+149\nu_m-454\nu_m^2+320\nu_m^3)]\}
\end{aligned}
\tag{10}
$$

For two-dimensional deformations the terms Ξ_R and Ξ_β, while for three-dimensional deformations the terms $(\Xi_R-|\Xi_{RR}|)$, $(2\Xi_\beta-|\Xi_{\beta\beta}|)$, and Ξ_Θ, are *all negative* for inclusions stiffer than the matrix (i.e., when the energy of a composite specimen is higher than that of the same specimen comprised of purely matrix material), all zero when they have the same stiffness, and *all positive* for inclusions less stiff than the matrix. This will be in a sense used as our definition of "inclusion stiffer than the matrix."

If the elastic energy Eq. (9) (or Eq. (7)) is compared to that evaluated for an identical prism now comprised of matrix material and containing a spherical (cylindrical in 2D) inclusion, ideally removed from an infinite body that is subjected to the far-field quadratic displacements Eq. (6) (or Eq. (5)), there is a mismatch of the linear terms in f, so that homogenization yields a material stiffer (more compliant) than the heterogeneous solution, for an inclusion stiffer (more compliant) than the matrix.

5 Comparison With Cosserat Material

The key point in the above discussion is that the results for the heterogeneous material are compared to a homogeneous linear elastic material, providing the effective properties. While a homogeneous material with appropriately chosen effective moduli can successfully mimic the composite material when uniform stress fields are applied, we showed that it cannot do so when the simplest nonuniform (i.e., uniform plus linearly varying) stress field is applied. What happens now if this comparison is made between a

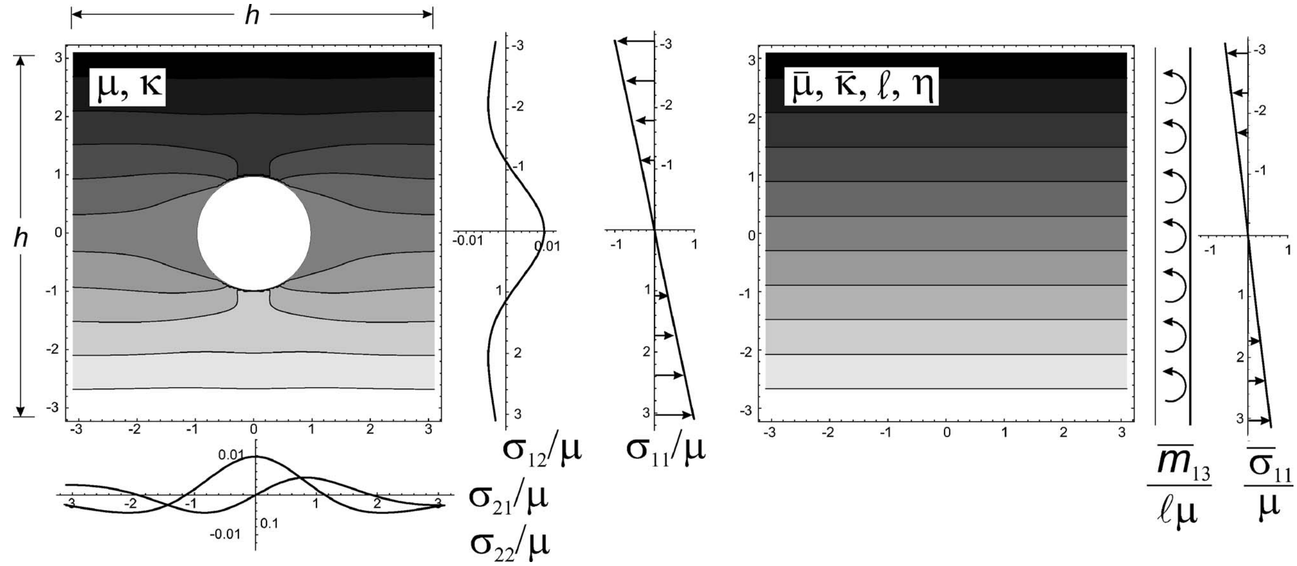


Fig. 2 Procedure of homogenization of a material containing a dilute distribution of circular voids and subject to a far-field bending stress distribution. Heterogeneous material (left) is an $h\times h$ prism removed from an infinite sheet that is subjected to uniaxial, linearly varying far-field stress; homogeneous Cosserat-elastic material (right) subject to the same mean moment (produced by $\bar{m}_{13}$ and $\bar{\sigma}_{11}$) calculated from the heterogeneous prism. For the plane strain problem (where η does not appear), level sets of σ_{11} are shown; note that the values of σ_{12}, σ_{21}, and σ_{22}, shown parallel to the edges, are less than 1/100 of the maximum value of σ_{11} at a distance from inclusion center equal to three times the radius of the inclusion (contrast this with the order of the effect in Fig. 1).

composite material and a homogeneous Cosserat or micropolar material? Note that this question has fundamental—as opposed to empirical—motivation: the assumption leading to standard Cauchy elasticity—that surface resultant moments/area vanish as the Cauchy tetrahedron becomes vanishing small—is a sensible approximation for materials with extremely small-scale microstructure, but is not in general otherwise justifiable. Absent this assumption, a Cosserat-type constitutive framework arises.

5.1 Simplest Cosserat Constitutive Model. We begin for simplicity with constrained-rotation micropolar materials (the simplest Cosserat constitutive model), for which the constitutive equations are [15]

$$\sigma_{ij} = 2\mu\left(\varepsilon_{ij} + \frac{\nu}{1-2\nu}\varepsilon_{kk}\delta_{ij}\right), \quad m_{ij} = 4\mu\ell^2(\chi_{ji} + \eta\chi_{ij}) \quad (11)$$

where σ_{ij} is the symmetric part of the force-stress tensor; ε_{ij} is the infinitesimal strain tensor; m_{ij} is the deviator of the couple-stress tensor; and χ_{ij} is the torsion-flexure tensor. The kinematical quantities are defined in terms of the displacement field u_i as

$$\varepsilon_{ij} = \tfrac{1}{2}(u_{i,j} + u_{j,i}), \quad \chi_{ij} = \omega_{i,j} = \tfrac{1}{2}e_{ihk}u_{k,hj} \quad (12)$$

where e_{ihk} is the Ricci (permutation) tensor; ω_i is the macrorotation axial vector; and a subscript comma denotes partial differentiation with respect to subsequent indices. The material parameters ν and μ appearing in Eq. (11) are the usual (Poisson and shear) elastic moduli (subject to the usual restrictions), whereas material parameters ℓ and η define the Cosserat behavior; in particular, the former is a characteristic length of the material and the latter is dimensionless and subject to the restriction $-1 < \eta < 1$ for positive definiteness of the strain energy.

Let us consider now two ideal material elements: a cube of edges h of Cauchy-elastic material containing an inclusion, ideally removed from an infinite body that is subjected to far-field loading, and the same cube instead composed of a homogeneous, constrained-rotation Cosserat material, Eqs. (11). We wish to determine the values of the effective Cosserat moduli $\bar{\mu}$, $\bar{\nu}$, ℓ, and η so that the homogeneous Cosserat material best mimics the heterogeneous Cauchy material under general slowly varying applied loading (Fig. 2, illustrating for simplicity a bending stress distribution).

5.2 Matching With the Uniform Stress Field. For *uniform* applied stress (and zero applied couple stress) the effective modulus values Eqs. (1) and (2), identical to those obtained for Cauchy-elastic material, are found for the Cosserat material. The reason for this is simply that for a uniform applied stress on the Cosserat material, a homogeneous deformation with null deformation-curvature tensor is produced, so that the Cosserat effects disappear (i.e., the moduli ℓ and η do not enter the solution).

5.3 Matching With Linearly Varying Remote Stress Field. For a linearly varying remote applied stress on the Cosserat material, Cosserat effects are present and, as will be shown, for inclusions less stiff than the matrix, they permit minimization, and for certain deformations elimination, of the mismatch in the strain energy between the actual composite material and the homogeneous effective Cosserat material.

Boundary conditions for a Cosserat solid and a Cauchy-elastic solid are not equivalent. For instance, in a purely kinematic approach, for a Cosserat material we can prescribe displacements Eqs. (6) (or Eqs. (5)) along a side of the prism, but the two tangential components of the rotation must also be specified, the latter not being necessary in a Cauchy solid. Following the kinematic approach, we assume displacements Eqs. (6) (or Eqs. (5)), and *the rotations deduced from these displacements*, to be prescribed along all sides of the prism for the Cosserat material. (For the Cauchy material, only the displacements Eqs. (6) (or Eqs. (5)) are prescribed on the boundary, but the resulting solution has rotations there identical to those prescribed for the Cosserat material, so the Cosserat and Cauchy material solutions correspond to exactly the same problem.) The solution to this boundary value problem for pure bending of the Cosserat material was given by Koiter ([15], his Secs. 6.2 and 6.3).

Generalizing the Koiter solution, for the displacement field Eq. (6) (or Eq. (5)), with ν_m replaced by $\bar{\nu}$, the non-null kinematical quantities become

$$\varepsilon_{ii} = \frac{x_j}{R_{ik}} + \frac{x_k}{R_{ij}} - x_i \frac{\bar{\nu}}{1-\bar{\nu}}\left(\frac{1}{R_{jk}} + \frac{1}{R_{kj}}\right) + 2x_i(\tilde{\beta}_{ij} + \tilde{\beta}_{ik})$$

(indices not summed; i,j,k cyclic),

$$\varepsilon_{ij} = -2\frac{1-\bar{\nu}}{1-2\bar{\nu}}(x_j\tilde{\beta}_{ik} + x_i\tilde{\beta}_{jk}) - e_{ijk}x_k\frac{\Theta_i - \Theta_j}{2}$$

(indices not summed and all different),

$$\chi_{ij} = \frac{1}{2}\delta_{ij}\left(3\Theta_i - \sum_{k=1}^{3}\Theta_k\right) + e_{jik}\left(\frac{1}{R_{ji}} + 2\tilde{\beta}_{ki}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right) \tag{13}$$

(indices not summed).

The total strain energy in the cube is thus

$$\mathcal{E} = \mathcal{E}_{\text{Cauchy}} + 2h^3\bar{\mu}\ell^2(\chi_{ij}\chi_{ij} + \eta\chi_{ji}\chi_{ij}) \tag{14}$$

where

$$\chi_{ij}\chi_{ij} = \frac{3}{2}\left(\sum_{i=1}^{3}\Theta_i^2 - \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{3}\Theta_i\Theta_j\right) + \sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{3}\left(\frac{1}{R_{ij}} + 2\tilde{\beta}_{kj}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right)^2 \tag{15}$$

and

$$\chi_{ji}\chi_{ij} = \frac{3}{2}\left(\sum_{i=1}^{3}\Theta_i^2 - \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{3}\Theta_i\Theta_j\right) - \sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{3}\left(\frac{1}{R_{ij}} + 2\tilde{\beta}_{kj}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right)\left(\frac{1}{R_{ji}} + 2\tilde{\beta}_{ki}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right) \tag{16}$$

which, for plane-strain deformations in the x_1, x_2 plane become

$$\chi_{ij}\chi_{ij} = \left(\frac{1}{R_{23}} + 2\tilde{\beta}_{13}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right)^2 + \left(\frac{1}{R_{13}} + 2\tilde{\beta}_{23}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right)^2, \quad \chi_{ji}\chi_{ij} = 0 \tag{17}$$

5.4 Result 1. The nonpolar (i.e., standard effective Cauchy) case is obtained from the strain energy Eq. (14) by setting the internal length equal to zero, $\ell=0$; therefore, since ℓ enters Eq. (14) only as ℓ^2, and since its coefficient cannot be negative for allowable modulus values, *the strain energy for the effective Cosserat material is never less than the strain energy for the effective Cauchy material.* This means that the introduction of Cosserat effects can only *increase* the strain energy of the effective material and therefore can only be useful when coefficients Ξ_R and Ξ_β are positive in plane strain (in Eqs. (7)) or when $\Xi_R - |\Xi_{RR}| > 0$, $2\Xi_\beta - |\Xi_{\beta\beta}| > 0$, and $\Xi_\Theta > 0$, in 3D (in Eqs. (9)), i.e., *for inclusions less stiff than the matrix*. In the case of an inclusion stiffer than the matrix, Cosserat effects make the homogenized material even stiffer than the already overly stiff effective Cauchy material resulting from homogenization for uniform stress. For these situations the simple Cosserat effective material cannot provide an improvement to the standard Cauchy effective material.

5.5 Result 2 for 2D Deformations. Let us begin with the two-dimensional (plane strain) formulation, where there is only one remaining undetermined parameter, the internal characteristic length ℓ, in the elastic energy, Eq. (14) (parameter η only enters the elastic energy in the three-dimensional case). We seek the ℓ value that permits minimization of the elastic energy difference through $O(f)$, for arbitrary equilibrium quadratic displacement remote boundary conditions, between the heterogeneous Cauchy material (whose energy has no $O(f)$ term) and the homogeneous effective Cosserat material:

$$\mathcal{E}_{\text{Cauchy}}(\mu_m,\nu_m) - [\mathcal{E}_{\text{Cauchy}}(\bar{\mu},\bar{\nu}) + 2h^3\bar{\mu}\ell^2\chi_{ij}\chi_{ij}] \tag{18}$$

which is (having divided by h^5)

$$\begin{aligned}&\frac{\mu_m}{12(1-\nu_m)}\left(\frac{1}{R_{13}^2} + \frac{1}{R_{23}^2}\right) + \frac{\mu_m(1-\nu_m)(3-4\nu_m)}{3(1-2\nu_m)^2}(\tilde{\beta}_{13}^2 + \tilde{\beta}_{23}^2)\\ &- \left\{\frac{\bar{\mu}}{12(1-\bar{\nu})}\left(\frac{1}{R_{13}^2} + \frac{1}{R_{23}^2}\right) + \frac{\bar{\mu}(1-\bar{\nu})(3-4\bar{\nu})}{3(1-2\bar{\nu})^2}(\tilde{\beta}_{13}^2 + \tilde{\beta}_{23}^2)\right.\\ &\left.+ 2\bar{\mu}\frac{\ell^2}{h^2}\left[\left(\frac{1}{R_{23}} + 2\tilde{\beta}_{13}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right)^2 + \left(\frac{1}{R_{13}} + 2\tilde{\beta}_{23}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right)^2\right]\right\}\end{aligned} \tag{19}$$

We wish to use ℓ to increase the elastic energy of the effective Cosserat material in such a way that this becomes closer to the correct value $\mathcal{E}_{\text{Cauchy}}(\mu,\nu)$, but without exceeding this value for any value of the *free parameters* defining the deformation modes: $1/R_{13}$, $1/R_{23}$, $\tilde{\beta}_{13}$, and $\tilde{\beta}_{23}$. Therefore, employing Eq. (7), Eq. (19) can be written as, retaining only terms through $O(f)$

$$\begin{aligned}&\left(\frac{1}{R_{13}^2} + \frac{1}{R_{23}^2}\right)\Xi_R f + (\tilde{\beta}_{13}^2 + \tilde{\beta}_{23}^2)\Xi_\beta f - 2\frac{\ell^2}{h^2}\left[\left(\frac{1}{R_{23}} + 2\tilde{\beta}_{13}\frac{1-\nu_m}{1-2\nu_m}\right)^2\right.\\ &\left.+ \left(\frac{1}{R_{13}} + 2\tilde{\beta}_{23}\frac{1-\nu_m}{1-2\nu_m}\right)^2\right] \geqslant 0\end{aligned} \tag{20}$$

and the problem is to find an ℓ^2/h^2 such that Eq. (20) is satisfied for all $1/R_{13}$, $1/R_{23}$, $\tilde{\beta}_{13}$, and $\tilde{\beta}_{23}$, coming as close to equality as possible. Note that, since the term multiplying ℓ^2 is always negative, inequality (20) can be satisfied only for inclusions less stiff than the matrix, i.e., when Ξ_R and Ξ_β are both positive.

Now, problem (20) can be transformed into the form $\mathbf{xAx} \geqslant 0$, with vector $\{\mathbf{x}\} = \{1/R_{13}, \tilde{\beta}_{23}, 1/R_{23}, \tilde{\beta}_{13}\}$, so that it becomes equivalent to the requirement of positive semi-definiteness of the 4×4 matrix $\mathbf{A}$ (which is composed of two identical 2×2 blocks, while all other entries are null). This matrix has two distinct eigenvalues with double multiplicity; requiring that the smaller eigenvalue be zero yields

$$\frac{\ell^2}{h^2} = \frac{f}{\left(\frac{1-\nu_m}{1-2\nu_m}\right)^2\frac{8}{\Xi_\beta} + \frac{2}{\Xi_R}} \tag{21}$$

valid only for both Ξ_R and Ξ_β positive.

Obviously, the meaning of negative values of ℓ^2 is merely that the inclusion is stiffer than the matrix and no (real) value exists for the characteristic length that will permit the elastic energies to match. In such cases, $\ell=0$ gives the smallest difference between the energies. Using $f=\pi a^2/h^2$, Eq. (21) becomes

$$\ell = a\sqrt{\frac{\pi}{\frac{8(1-\nu_m)^2}{(1-2\nu_m)^2\Xi_\beta} + \frac{2}{\Xi_R}}} \tag{22}$$

valid only for both Ξ_R and Ξ_β positive. Note from Eq. (22) that $\ell=0$ when $a=0$, but that ℓ/a is independent of f (under our assumption of small f). Note that in the limit of an incompressible matrix, $\nu_m=1/2$, Eq. (22) reduces to

$$\ell = a\sqrt{\frac{\pi\Xi_R}{2}} \tag{23}$$

showing that the corresponding applied deformation mode is a pure bending. In this case, in other words, the characteristic length Eq. (23) provides an exact match between the energies of the actual heterogeneous solid and the homogenized one under arbitrary uniform plus pure bending applied loading.

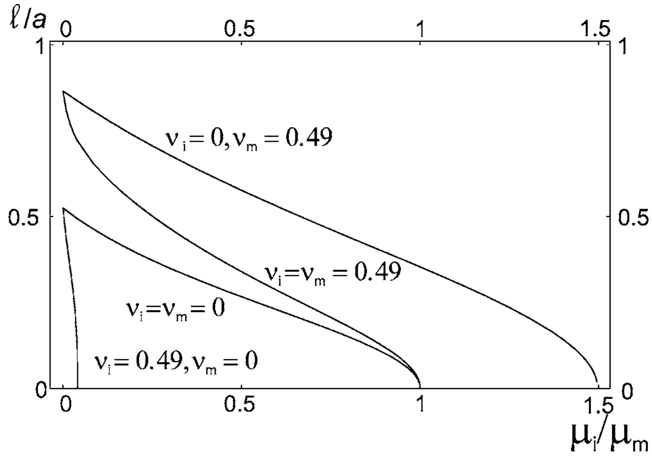


Fig. 3 Characteristic length divided by circular cylindrical inclusion radius for a homogeneous Cosserat material deduced from homogenization of a matrix containing a dilute distribution of parallel, infinite circular cylindrical inclusions (plane strain, Eq. (22))

In the extreme case when the inclusion is a void, Eq. (22) becomes

$$\ell = a\sqrt{\frac{\pi}{24(1-\nu_m)\left\{\dfrac{1}{3}+\dfrac{1-2\nu_m}{7-2\nu_m[13-8\nu_m(3-2\nu_m)]}\right\}}} \tag{24}$$

where the radical in Eq. (24) is always positive.

The characteristic length divided by the radius of the inclusion, ℓ/a, is plotted in Fig. 3 versus the contrast in the inclusion/matrix shear moduli, μ_i/μ_m. A null contrast corresponds to a void, Eq. (24). The different curves in the figure refer to different values of Poisson's ratios. The values of the curves at $\mu_i/\mu_m=0$ depend only on ν_m; curves are plotted for ν_m and ν_i each having values 0.49 and 0. Note also that for $\nu_m=\nu_i$, $\ell=0$ results for $\mu_i=\mu_m$, as it should.

For a sufficiently compliant inclusion, a positive characteristic length for an effective Cosserat material is always found, which decreases to zero at sufficiently high inclusion stiffness.

5.6 Result 2 for 3D Deformations. Let us now consider three-dimensional deformations. By introducing the symbol

$$T^2 = \sum_{i=1}^{3}\Theta_i^2 - \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{3}\Theta_i\Theta_j \tag{25}$$

the three-dimensional version of non-negativity of the energy difference Eq. (18) becomes

$$\begin{aligned}
&\sum_{\substack{i,j=1\\ i\neq j}}^{3}\frac{1}{R_{ij}^2}\Xi_R + \sum_{\substack{i,j=1\\ i\neq j}}^{3}\frac{1}{R_{ij}R_{ji}}\Xi_{RR} + T^2\Xi_\Theta + \sum_{\substack{i,j=1\\ i\neq j}}^{3}\tilde{\beta}_{ij}^2\Xi_\beta + (\tilde{\beta}_{12}\tilde{\beta}_{13}\\
&\quad + \tilde{\beta}_{21}\tilde{\beta}_{23} + \tilde{\beta}_{31}\tilde{\beta}_{32})\Xi_{\beta\beta} - 2\frac{\bar{\mu}\ell^2}{\mu_m f h^2}\left[(1+\eta)\frac{3}{2}T^2 + \sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{3}\left(\frac{1}{R_{ij}}\right.\right.\\
&\quad \left. + 2\tilde{\beta}_{kj}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right)^2 - \eta\sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{3}\left(\frac{1}{R_{ij}} + 2\tilde{\beta}_{kj}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right)\left(\frac{1}{R_{ji}}\right.\\
&\quad \left.\left. + 2\tilde{\beta}_{ki}\frac{1-\bar{\nu}}{1-2\bar{\nu}}\right)\right] \geqslant 0
\end{aligned} \tag{26}$$

Equation (26) depends on the arbitrary deformation modes. These are coupled in groups of four (each group entering in exactly the same way), plus T; for example, $1/R_{13}$, $1/R_{31}$, $\tilde{\beta}_{21}$, $\tilde{\beta}_{23}$ are coupled. Thus it is sufficient to consider these four parameters together with T, and take all others equal to zero. Doing this, Eq. (26) becomes, retaining only leading-order terms in f

$$\begin{aligned}
&\left(\frac{1}{R_{13}^2}+\frac{1}{R_{31}^2}\right)\Xi_R + \frac{2}{R_{13}R_{31}}\Xi_{RR} + T^2\Xi_\Theta + (\tilde{\beta}_{21}^2+\tilde{\beta}_{23}^2)\Xi_\beta\\
&\quad + \tilde{\beta}_{21}\tilde{\beta}_{23}\Xi_{\beta\beta} - 2\frac{\ell^2}{fh^2}\left[(1+\eta)\frac{3T^2}{2} + \left(\frac{1}{R_{13}} + 2\tilde{\beta}_{23}\frac{1-\nu_m}{1-2\nu_m}\right)^2\right.\\
&\quad + \left(\frac{1}{R_{31}} + 2\tilde{\beta}_{21}\frac{1-\nu_m}{1-2\nu_m}\right)^2 - 2\eta\left(\frac{1}{R_{13}} + 2\tilde{\beta}_{23}\frac{1-\nu_m}{1-2\nu_m}\right)\left(\frac{1}{R_{31}}\right.\\
&\quad \left.\left. + 2\tilde{\beta}_{21}\frac{1-\nu_m}{1-2\nu_m}\right)\right] \geqslant 0
\end{aligned} \tag{27}$$

Equation (27) involves a quadratic form, so that it can be represented in matrix form as

$$\left[\Xi_\Theta - 3\frac{\ell^2}{fh^2}(1+\eta)\right]T^2 + \mathbf{xAx} \geqslant 0 \tag{28}$$

where vector $\{\mathbf{x}\}=\{1/R_{13}, 1/R_{31}, \tilde{\beta}_{21}, \tilde{\beta}_{23}\}$. Matrix $\mathbf{A}$ is a 4×4 block and the Condition (26), viewed as the condition of positive semi-definiteness of $\mathbf{A}$ (since the coefficient of T^2 must be $\geqslant 0$), yields non-negativeness of four eigenvalues, plus non-negativity of the coefficient of T^2. Two of these conditions can be shown to be contained within the other two, from which two values of Cosserat length ℓ can be obtained to ensure positive semi-definiteness of $\mathbf{A}$. The minimum among these two lengths, plus that obtained considering T, yields the Cosserat length for Condition (26) to be satisfied

$$\ell(\eta) = af^{1/6}\left(\frac{4\pi}{3}\right)^{1/3} g(\mu_i,\mu_m,\nu_i,\nu_m,\eta) \tag{29}$$

where

$$g(\mu_i,\mu_m,\nu_i,\nu_m,\eta) = \min\left[\frac{\Xi_\Theta}{3(1+\eta)}, \frac{(1-2\nu_m)^2}{2(1+\eta)\left(\dfrac{8(1-\nu_m)^2}{2\Xi_\beta - \Xi_{\beta\beta}} + \dfrac{(1-2\nu_m)^2}{\Xi_R - \Xi_{RR}}\right)}, \frac{(1-2\nu_m)^2}{2(1-\eta)\left(\dfrac{8(1-\nu_m)^2}{2\Xi_\beta + \Xi_{\beta\beta}} + \dfrac{(1-2\nu_m)^2}{\Xi_R + \Xi_{RR}}\right)}\right]^{1/2} \tag{30}$$

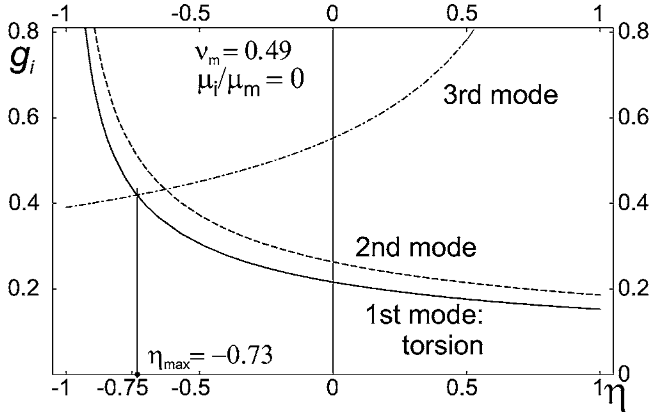


Fig. 4 The three functions g_i appearing in Eq. (30), among which the minimum is selected for given values of η

in which all terms are always non-negative for inclusions less stiff than the matrix. Equation (29) applies for given values of μ_i, μ_m, ν_i, ν_m, and η.

In Eq. (30), the minimum among the three functions (call them g_i) is taken. These functions have the typical dependence on η shown in Fig. 4, drawn for $\mu_i/\mu_m=0$ (so that the inclusion is a void) and $\nu_m=0.49$ (a case that will also be considered later). In this figure, one of the g_i's corresponds to the torsion mode, while the other two modes involve both bending and the modes described by the $\tilde{\beta}_{ij}$.

Since η is a constitutive parameter which can be chosen so that the effective Cosserat material best mimics the actual heterogeneous material's response, it is optimal to choose it so that the *Cosserat effective material matches the actual heterogeneous one for two modes of deformation*, which corresponds to the intersection of the two lower curves in Fig. 4, that is, to the largest of the minima (i.e., the supremum) of the three g_i's (corresponding to $\eta_{\max}$ in the figure). Therefore

$$\ell = af^{1/6}\left(\frac{4\pi}{3}\right)^{1/3} \sup_{\eta\in(-1,1)} g(\mu_i,\mu_m,\nu_i,\nu_m,\eta) \tag{31}$$

The case of an incompressible matrix ($\nu_m \to 1/2$) is worth noting. In this case, Eq. (30) becomes

$$g(\mu_i,\mu_m,\nu_i,\nu_m,\eta) = \min\left[\frac{\Xi_\Theta}{3(1+\eta)}, \frac{\Xi_R - \Xi_{RR}}{2(1+\eta)}, \frac{\Xi_R + \Xi_{RR}}{2(1-\eta)}\right]^{1/2} \tag{32}$$

showing that bending and torsion are the only modes entering the formula. In this case, in other words, the characteristic length ℓ and parameter η found from Eq. (31), in which Eq. (32) is used for function g, provide an exact match between the energies of the actual heterogeneous solid and the homogenized Cosserat one under arbitrary uniform plus bending and torsion applied loading.

The limit $\mu_i \to 0$ of Eq. (31) yields the case of a spherical void

$$\ell = af^{1/6}\frac{\sqrt[3]{4\pi/3}}{\sqrt{7-5\nu_m}} \max_{\eta\in(-1,1)} \min\left(\frac{5(1-\nu_m)}{12(1+\eta)}, \frac{5(1-\nu_m)[4-\nu_m(11-15\nu_m)]}{4[13-\nu_m(37-40\nu_m)](1+\eta)}, \frac{(1+\nu_m)\{17-\nu_m[74-\nu_m(129-80\nu_m)]\}}{2(1-\nu_m)\{21-\nu_m[78-\nu_m(121-80\nu_m)]\}(1-\eta)}\right)^{1/2} \tag{33}$$

In the case of matrix incompressibility, $\nu_m \to 1/2$, Eq. (33) becomes

$$\ell = af^{1/6}\frac{\sqrt[3]{4\pi/3}}{6}\sqrt{\frac{41}{6}}, \quad \eta = -\frac{31}{41} \tag{34}$$

in which case both the bending and the torsion modes are simultaneously matched.

We emphasize with respect to all the above cases that when the far-field applied loading is such that our "optimal" choice of the Cosserat parameter does not provide an exact match between the effective Cosserat material's energy and that of the actual heterogeneous material, our optimal effective Cosserat material will still be an improvement over the standard effective Cauchy material for *all* equilibrium uniform plus linear far-field applied loadings (for compliant-inclusion-type composites).

The characteristic length divided by the radius of the inclusion multiplied now by the volume fraction to the power $-1/6$, i.e., $f^{-1/6}\ell/a$, is plotted in Fig. 5 versus the contrast in the inclusion/matrix shear moduli, μ_i/μ_m, so that a null contrast corresponds to a void, Eq. (33). The different curves in the figures refer to different values of Poisson ratios, the same investigated for plane strain (ν_m and ν_i each having values 0.49 and 0). The values of the curves at $\mu_i/\mu_m=0$ depend only on ν_m.

The figures show that the qualitative behavior is the same for the two-dimensional and three-dimensional cases: for a sufficiently compliant inclusion, a positive characteristic length for an effective Cosserat material is always found, which decreases to zero at sufficiently high inclusion stiffness. However, there are also important differences between the 2D and the 3D cases:

1. For all values of the Poisson ratios of the matrix and inclusion, Ξ_Θ vanishes when $\mu_i=\mu_m$ and then becomes negative for $\mu_i>\mu_m$. Therefore, due to the effect of the torsion mode and in contrast to the 2D case, it is always impossible to produce a positive characteristic length ℓ for $\mu_i>\mu_m$, regardless of the values of the Poisson ratios, so that $\ell=0$ always results for $\mu_i \geqslant \mu_m$ (and not only for the special case $\nu_m=\nu_i$);
2. The curve for ℓ for the case $\nu_i=0$ and $\nu_m=0.49$ for 3D displays a jump to zero (not found for 2D deformations) at $\mu_i=\mu_m$. This behavior, occurring when $\nu_m>\nu_i$, is related to torsion and to the fact that η simultaneously tends to the limit -1. This means that the quantity $\ell^2(1+\eta)$, related to the characteristic length in torsion, is not discontinuous and correctly approaches zero when μ_i/μ_m tends to 1; and
3. *Result* 3. The characteristic length is substantially smaller in three dimensions than in two. This is partially due to the fact that $\ell \propto af^{1/6}$ in three dimensions, whereas $\ell \propto a$ in two dimensions. The figures show that the largest characteristic length (strongest Cosserat effect) occurs for an incompressible matrix containing voids ($\nu_m=0.5$, $\mu_i=0$), in which case

$$\ell \approx 0.702af^{1/6}, \quad \ell = \frac{\sqrt{\pi}}{2}a \approx 0.886a \tag{35}$$

for 3D and 2D, respectively. For example, if $f=0.1$, Eqs. (35) show ℓ/a in 3D to be 54% of that in 2D.

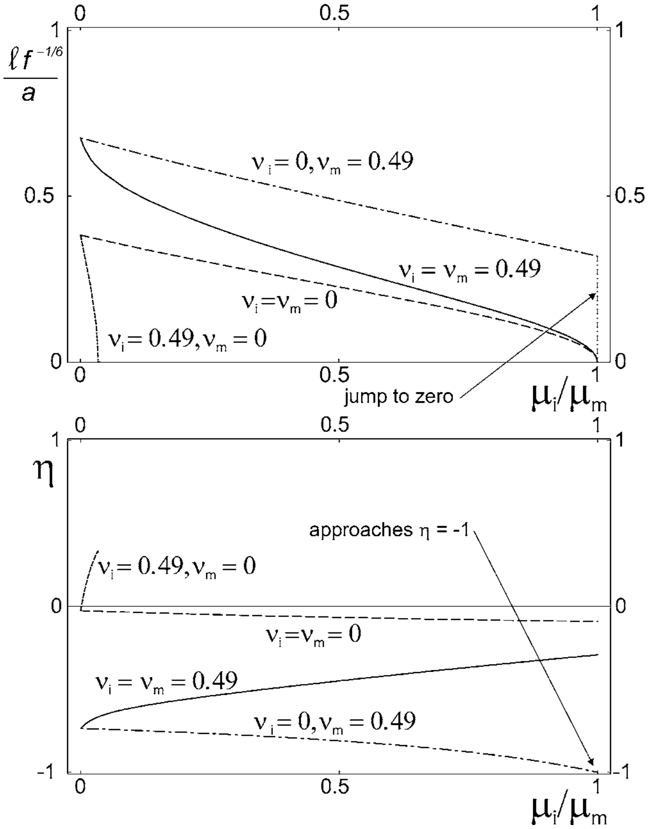


Fig. 5 Characteristic length divided by spherical inclusion radius and multiplied by $f^{-1/6}$ (top) and parameter η (bottom) for a homogeneous Cosserat material deduced from homogenization of a matrix containing a dilute distribution of spherical inclusions (Eq. (31))

6 Unconstrained Cosserat Materials do not Change Results 1, 2, and 3

At this point we are in a position to address the following question: can Result 1, stating that Cosserat effects only arise for inclusions less stiff than the matrix, be changed by making recourse to a more general theory of micropolar behavior than the constrained-rotation theory of Eq. (11)? Moreover, does Result 2, providing a closed-form formula for the characteristic length ℓ, and consequent Result 3, change if a general theory of micropolar behavior is assumed? The answers to these questions turn out to be negative, but they require a digression.

A general isotropic, linear micropolar material is characterized by the following constitutive equations [16,2]

$$\Sigma_{ij} = \lambda\varepsilon_{kk}\delta_{ij} + 2\mu\varepsilon_{ij} + \gamma e_{ijk}(\omega_k - \phi_k),$$

$$\mu_{ij} = \alpha\phi_{k,k}\delta_{ij} + 4\mu\ell^2(\phi_{j,i} + \eta\phi_{i,j}) \tag{36}$$

where Σ_{ij} and μ_{ij} are the asymmetric force-stress and couple-stress tensors, respectively, and ω_k and ϕ_i are the macro- and micro-axial rotation vectors, respectively. Constants λ and μ play the role of the usual Lamé moduli of Cauchy elasticity, and α, η, γ, and ℓ are new material constants.

The important point is to note that Eqs. (11) are obtained from Eqs. (36) by taking $\phi_k = \omega_k$; then Σ_{ij} and μ_{ij} reduce to σ_{ij} (the symmetric part of the stress tensor) and m_{ij} (the deviator of the couple-stress tensor), respectively, and the terms containing α and γ in Eqs. (36) become identically zero.

Now we note that in the unconstrained theory, kinematical boundary conditions must involve prescription of displacements, macrorotations, and microrotations. If we make the sensible choice that the microrotations are identical to the macrorotations *on the boundary* and these are those arising from displacements Eqs. (6), then a (unique) solution to the full unconstrained theory produces the same energy Eq. (14). The same results for ℓ, Eqs. (22) and (31), are obtained. Now, however, parameters α and γ remain undetermined. Thus we find no advantage to use of the more complex unconstrained Cosserat model in the homogenization problem, and indeed we find the constrained-rotation model employed by Koiter [15] to have the great advantages of simplicity and physical transparency.

7 Experiments and Applications

We have already reported that our results explain and confirm the Gauthier [7] experimentally based claim that "an anti-micropolar phenomenon" is found for inclusions stiffer than the matrix. For inclusions less stiff than the matrix, our theory provides Cosserat parameters ℓ and η (only ℓ for plane strain) for the effective material which exactly match two quadratic deformation modes (one in plane strain), so that these parameters would be found in an ideal experiment performed on a specimen, when the boundary conditions corresponding to those modes are imposed. With the exception of an incompressible matrix material, the quadratic modes correspond to a combination of bending, torsion, and other modes, which are usually not experimentally investigated.

7.1 Bending and Torsion Experiments, and Applications Involving Pure Bending and Torsion Loading. Common experiments involve bending (usually bending of a plate deformed in plane strain) and torsion (usually of a bar with circular cross section). Performing such experiments will not in general (again, with the exception of plane-strain bending of a composite with an incompressible matrix material) yield our Cosserat parameters. This is because we have selected these to give the greatest possible improvement over the effective Cauchy material for all possible imposed linear plus quadratic displacement fields, such that the effective Cosserat material is never stiffer than the actual heterogeneous one. If, however, the applied loading of interest is known to consist of uniform plus pure bending loading in 2D, or uniform plus pure bending and pure torsion loading in 3D, the effective Cosserat parameters can be chosen to produce an exact energy match between the effective Cosserat material and the actual heterogeneous one.

In particular, for plane-strain deformations of a slab containing a dilute distribution of *cylindrical inclusions* (with axis parallel to the depth)

$$\ell_{\text{2D-bending}} = a\sqrt{\frac{\pi\Xi_R}{2}} \tag{37}$$

with Ξ_R given by Eq. $(8)_1$, provides an exact match for a plane strain bending experiment.

For plane-strain deformations of a slab containing a dilute distribution of *spherical inclusions* (note that, due to the plane strain constraint, parameter η does not enter)

$$\ell_{\text{3D plane-strain bending}} = af^{1/6}\left(\frac{4\pi}{3}\right)^{1/3}\sqrt{\frac{\Xi_R}{2}} \tag{38}$$

where Ξ_R is given by Eq. $(10)_1$, gives an exact match for a plane-strain bending experiment.

For torsion of a cylindrical specimen (of circular cross section) containing a dilute distribution of *spherical inclusions*

$$(\ell\sqrt{1+\eta})_{\text{torsion cylindrical bar}} = af^{1/6}\left(\frac{4\pi}{3}\right)^{1/3}\sqrt{\frac{\Xi_\Theta}{3}} \qquad (39)$$

where Ξ_Θ is given by Eq. $(10)_3$, gives an exact match. Obviously, ℓ and η can be chosen to satisfy Eqs. (38) and (39) simultaneously.

7.2 A Comparison With Existing Experimental Results. It is interesting now to compare our results with experiments performed on material containing compliant inclusions, for instance voids. In particular, our results indicate that the most effective experimental setting to display Cosserat effects would be a material containing cylindrical voids deformed in plane strain, with a matrix Poisson's ratio tending to the limit value 0.5; for instance, a rubber block with cylindrical holes. Unfortunately, nothing like this experimental setup is available in the literature and also nothing pertaining to *dilute* suspensions of *spherical voids*.

The only results that we were able to find are those by Lakes [2] pertaining to two foams with nearly spherical voids. Specifically, one material is a syntactic foam consisting of hollow glass microbubbles embedded in an epoxy matrix, for which the mean diameter of voids is 0.125 mm and the volume fraction is 0.468. The second material is a high-density rigid polyurethane closed-cell foam, for which the mean diameter of voids is 0.1 mm and the volume fraction is 0.690. Within the general Cosserat framework Eqs. (36), Lakes [2] finds $\ell=0.032$ mm for the first material and $\ell=0.327$ mm for the second. Lakes also determines the quantity $\ell\sqrt{2(1+\eta)}$, which he estimates to be 0.065 mm and 0.62 mm, respectively.

There are several difficulties in attempting to compare our results with these materials:

1. The void volume fraction is so high that the dilute approximation is almost certainly not directly applicable;
2. The mechanical properties of the matrix material are not available;[2] and
3. The voids in the first material are coated by a glass shell of unknown stiffness.

Since these factors make a precise comparison impossible, we simply employ our model results with $\nu_m=1/2$, Eqs. (34), to make an order-of-magnitude comparison. Thus Eqs. (34) give $\ell=0.039$ mm and $\ell\sqrt{2(1+\eta)}=0.030$ mm for the first material and $\ell=0.033$ mm and $\ell\sqrt{2(1+\eta)}=0.025$ mm for the second. These results are only in qualitative agreement with the experimental findings; however, they are consistent with the fact that our model, based on the dilute approximation, underestimates the characteristic length ℓ for the given high values of the pore volume fractions. The fact that the characteristic length is better predicted for the first material than for the second is probably a consequence of the presence of the glass shell coating the voids, providing a stiffness, which strongly decreases ℓ.

8 Summary of General Methodology

Here we summarize the general methodology proposed in this paper and employed in the specific cases of a matrix containing a dilute suspension of spherical or circular cylindrical inclusions. We emphasize that our general methodology is not restricted to composites consisting of a matrix containing a dilute concentration of another phase. To determine the effective moduli for a homogeneous Cosserat-elastic material that best approximates a heterogeneous Cauchy-elastic material under general applied loadings, one first determines the effective Cauchy-elastic moduli in the standard manner (i.e., using the most accurate approach available from standard composite materials theory. We emphasize that we regard the uniform loading as the primitive case, so that this initial determination will not be affected by subsequent calculations). One then needs to compute the elastic energy in the heterogeneous material of interest when this is subjected to a general equilibrium linearly varying applied traction (or quadratically varying displacements) on the boundary. One then compares this energy to the energy computed for the homogeneous Cosserat material (whose Cauchy moduli have already been determined via the standard homogenization approach) subjected to the same quadratically varying displacements and rotations as in the Cauchy solution, and one chooses the Cosserat parameters so that these two energies are in closest possible agreement. In the specific cases analyzed in this paper, the Cosserat length is nonzero when the heterogeneous material is less stiff than its predominant phase, and zero otherwise.

9 Conclusions

It has been shown that a dilute dispersion of elastic isotropic spherical inclusions in a 3D composite (and infinitely long, parallel circular cylindrical inclusions in a 2D one) produce Cosserat effects when the inclusions are less stiff than the matrix. The effects induce a characteristic length in three dimensions

$$\ell \propto af^{1/6}$$

and one in two dimensions

$$\ell \propto a$$

where a is the inclusion radius and f the volume fraction of the inclusion material. The maximum characteristic length occurs when the inclusions are cavities, and the matrix material is incompressible; this length is substantially larger in 2D versus 3D for cavities having the same radius. Cosserat effects are on the other hand excluded for the opposite situation of inclusions having stiffnesses equal to or greater than that of the matrix.

An important practical implication of our findings is that the response of a composite material containing inclusions less stiff than the matrix and subjected to nonuniform stressing can be more accurately represented by a homogeneous Cosserat material with appropriately chosen moduli than by a standard (Cauchy) effective material.

Acknowledgment

We thank Professor R. S. Lakes (University of Wisconsin-Madison) for helpful discussions. D.B. acknowledges financial support from MURST-Cofin 2005 (Prot. No. 2005085973–002). W.J.D. thanks Dr. W. G. Wolfer for his interest in this work and for support from Lawrence Livermore National Laboratory under Contract W-7405-ENG-48 with the U.S. Department of Energy. W.J.D. also acknowledges support from the Italian Ministry Research Grant "Rientro dei cervelli" MIUR Grant No. 26/1/2001.

Appendix A: The State of the Art on Cosserat Effects as Deduced From Elastic, Inhomogeneous Media

The literature on Cosserat effects arising from heterogeneous media is rife with conflicting views. Berglund [17], claiming that previous results [18–20] were inconsistent, provides two theoretical arguments to *disprove micropolar effects*, employing both a discrete structural model of a cubic lattice and a framework for homogenization of a heterogeneous continuum. These appear to be far from conclusive, since the former invokes reduction of structural dimensions to zero (which is inconsistent with the fact that Cosserat effects should be related to some non-null characteristic microstructural length) and the latter does indeed predict some micropolar effects, which are then argued to be negligible. On the contrary, Cosserat behavior was found by Wang and Stronge [21] for a hexagonal lattice. Moreover, certain theoretical arguments *in favor* of Cosserat behavior have been provided by Achenbach and Hermann [22] and Beran and McCoy [23], but the

[2]Only ν_m is needed to determine ℓ. However, the knowledge of μ_m would allow us to determine $\bar{\mu}$ and $\bar{\kappa}$ from Eqs. (1), which compared to experimental results by Lakes would permit an assessment of the quality of the estimate.

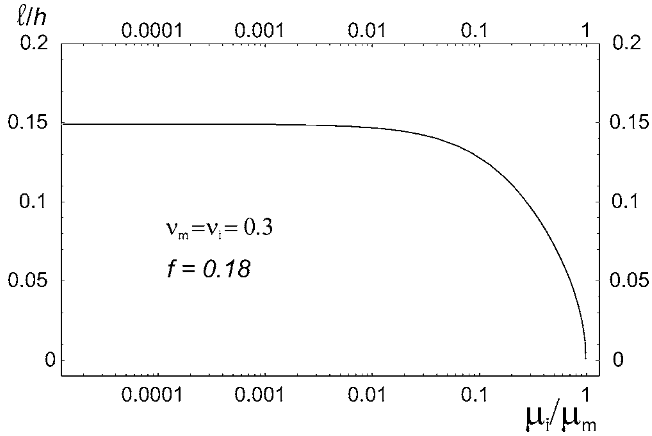


Fig. 6 Characteristic length divided by the cell size for volume fraction of disperse phase f=0.18, for a Cosserat material deduced from homogeneization of a matrix containing a dilute distribution of parallel, infinite circular cylindrical inclusions (plane strain, Eq. (21))

former holding only in certain circumstances involving dynamical effects and the latter apparently finally disproving the effects for composites with homogeneous and isotropic statistics of inclusions. Recently, Forest [24], Ostoja-Starzewski et al. [25], and Bouyge et al. [26] provided numerical finite element investigations supporting Cosserat effects in heterogeneous materials. Forest treats an anisotropic composite with an unusual microstructure, and does not directly provide values for the Cosserat characteristic length. The latter two papers treat plane problems of a matrix containing a dispersion of circular inclusions; they find a nonzero Cosserat length both for inclusions stiffer and more compliant than the matrix, a fact contradicted previously by experiments [6,7], and now by the analytical results derived in the present paper.

When Eq. (21) is plotted using a semi-logarithmic scale, such as that employed in Ref. [26] for their parameter values of $\nu_i = \nu_m = 0.3$ and $f = 0.18$, we obtain the graph shown in Fig. 6. The numerical values at high contrast are similar to those found in Ref. [26] (their Fig. 8), but our results: (1) correctly approach zero when the elastic mismatch disappears (while a nonzero characteristic length is found in Ref. [25] even for zero mismatch); and (2) show that Cosserat effects are excluded for mismatch greater than 1 (in which case ℓ would be imaginary).

Appendix B: Plane-Strain Solution of an Elastic Circular Inclusion in an Infinite Elastic Matrix, Subject to Remote Displacements Field Eqs. (5)

We use the Kolosov–Muskhelishvili [11] complex potentials representation of the general solution for plane problems in homogeneous isotropic linear elastostatics, which in polar coordinates is

$$u_r + \mathrm{i}u_\vartheta = \frac{1}{2\mu} e^{-\mathrm{i}\vartheta}[\kappa\varphi(z) - z\overline{\varphi'(z)} - \overline{\psi(z)}] \tag{B1}$$

$$\sigma_{rr} + \sigma_{\vartheta\vartheta} = 4\,\mathrm{Re}[\varphi'(z)]$$

$$\sigma_{\vartheta\vartheta} - \sigma_{rr} + 2i\sigma_{r\vartheta} = 2e^{2\mathrm{i}\vartheta}[\bar{z}\varphi''(z) + \psi'(z)] \tag{B2}$$

where $z = x_1 + ix_2 = re^{i\vartheta}$, $\varphi(z)$ and $\psi(z)$ are analytic functions, Re[] denotes the real part, and $\kappa = 3 - 4\nu$ for plane strain.

First, we consider a pure bending far-field applied loading, corresponding to

$$\sigma_{22} = mx_1, \quad \sigma_{11} = \sigma_{12} = 0 \quad \text{for } r \to \infty \tag{B3}$$

or, in terms of complex potentials

$$\varphi(z) = \psi(z) = \frac{m}{8}z^2, \quad \text{for } |z| \to \infty \tag{B4}$$

The solution for a matrix containing an inclusion of radius a is

$$\varphi(z) = \frac{m}{8}z^2 + \frac{\mu_i - \mu_m}{2(\kappa_m\mu_i + \mu_m)}\frac{ma^4}{4z^2} \tag{B5}$$

$$\psi(z) = \frac{m}{8}z^2 + \frac{\kappa_m\mu_i - \mu_m\kappa_i}{\mu_i + \kappa_i\mu_m}\frac{ma^4}{8z^2} + \frac{\mu_i - \mu_m}{\kappa_m\mu_i + \mu_m}\frac{ma^6}{4z^4} \tag{B6}$$

in material outside the inclusion, and

$$\varphi(z) = \frac{(\kappa_m + 1)\mu_i}{\mu_i + \kappa_i\mu_m}\frac{m}{8}z^2 - \frac{\mu_i}{\mu_m}\frac{\mu_i + \mu_m(\kappa_i - \kappa_m - 1)}{\kappa_i(\mu_i + \kappa_i\mu_m)}\frac{ma^2}{4} \tag{B7}$$

$$\psi(z) = \frac{(\kappa_m + 1)\mu_i}{\kappa_m\mu_i + \mu_m}\frac{m}{8}z^2 \tag{B8}$$

in material inside the inclusion.

Second, we consider a quadratic far-field applied displacement field, corresponding to

$$u_1 = \tilde{\beta}_{13}\left(x_1^2 - \frac{\kappa_m + 1}{\kappa_m - 1}x_2^2\right), \quad u_2 = u_3 = 0 \quad \text{for } r \to \infty \tag{B9}$$

or, in terms of complex potentials

$$\varphi(z) = \frac{\mu_m\tilde{\beta}_{13}}{\kappa_m - 1}z^2, \quad \psi(z) = -\frac{\mu_m\kappa_m\tilde{\beta}_{13}}{\kappa_m - 1}z^2 \quad \text{for } |z| \to \infty \tag{B10}$$

The solution is

$$\varphi(z) = \frac{\mu_m\tilde{\beta}_{13}}{\kappa_m - 1}z^2 + \frac{\mu_m - \mu_i}{\kappa_m\mu_i + \mu_m}\frac{a^4\mu_m\kappa_m\tilde{\beta}_{13}}{z^2(\kappa_m - 1)} \tag{B11}$$

$$\psi(z) = -\frac{\mu_m\kappa_m\tilde{\beta}_{13}}{\kappa_m - 1}z^2 + \frac{\kappa_m\mu_i - \mu_m\kappa_i}{\mu_i + \kappa_i\mu_m}\frac{a^4\mu_m\tilde{\beta}_{13}}{z^2(\kappa_m - 1)} + \frac{\mu_m - \mu_i}{\kappa_m\mu_i + \mu_m}\frac{2a^6\mu_m\kappa_m\tilde{\beta}_{13}}{z^4(\kappa_m - 1)} \tag{B12}$$

in material outside the inclusion, and

$$\varphi(z) = \frac{(\kappa_m + 1)\mu_i}{\mu_i + \kappa_i\mu_m}\frac{\mu_m\tilde{\beta}_{13}}{\kappa_m - 1}z^2 - \frac{\mu_i}{\mu_m}\frac{\mu_i + \mu_m(\kappa_i - \kappa_m - 1)}{\kappa_i(\mu_i + \kappa_i\mu_m)}\frac{2a^2\mu_m\tilde{\beta}_{13}}{\kappa_m - 1} \tag{B13}$$

$$\psi(z) = -\frac{(\kappa_m + 1)\mu_i}{\kappa_m\mu_i + \mu_m}\frac{\mu_m\kappa_m\tilde{\beta}_{13}}{\kappa_m - 1}z^2 \tag{B14}$$

in material inside the inclusion.

Appendix C: Three-Dimensional Solution of a Spherical Elastic Inclusion in an Infinite Elastic Matrix, Subject to Remote Displacements Field Eqs. (6)

C.1 Torsion Prescribed at Infinity

First, we consider an applied far-field torsion, consisting of a different (in general) angle of twist/length applied about each of the three Cartesian axes. This corresponds to the equilibrium displacement field

$$u_1 = (\Theta_2 - \Theta_3)x_2x_3, \quad u_2 = (\Theta_3 - \Theta_1)x_3x_1,$$
$$u_3 = (\Theta_1 - \Theta_2)x_1x_2 \quad \text{for } r \to \infty \tag{C1}$$

or, in spherical coordinates

$$u_r = 0, \quad u_\vartheta = -(\Theta_1 - \Theta_2)r^2 \sin\phi\cos\phi\sin\vartheta$$

$$u_\phi = -\frac{\Theta_1 + \Theta_2 - 2\Theta_3 + (\Theta_1 - \Theta_2)\cos 2\phi}{4} r^2 \sin 2\vartheta \quad \text{for } r \to \infty \tag{C2}$$

The solution to this applied far-field, satisfying equilibrium everywhere, and displacement and traction continuity across the inclusion-matrix boundary $r=a$, is

$$u_r = 0$$

$$u_\vartheta = -(\Theta_1 - \Theta_2)\sin\phi\cos\phi\sin\vartheta\left[r^2 + \frac{a^5(\mu_m - \mu_i)}{r^3(4\mu_m + \mu_i)}\right]$$

$$u_\phi = -\frac{\Theta_1 + \Theta_2 - 2\Theta_3 + (\Theta_1 - \Theta_2)\cos 2\phi}{4}\sin 2\vartheta\left[r^2 + \frac{a^5(\mu_m - \mu_i)}{r^3(4\mu_m + \mu_i)}\right] \tag{C3}$$

in material outside the inclusion, and

$$u_r = 0$$

$$u_\vartheta = -5\mu_m r^2 \frac{\Theta_1 - \Theta_2}{4\mu_m + \mu_i}\sin\phi\cos\phi\sin\vartheta$$

$$u_\phi = -5\mu_m r^2 \frac{\Theta_1 + \Theta_2 - 2\Theta_3 + (\Theta_1 - \Theta_2)\cos 2\phi}{4(4\mu_m + \mu_i)}\sin 2\vartheta \tag{C4}$$

in material inside the inclusion.

C.2 Bending and the Other Equilibrium Quadratic Displacement Modes Prescribed at Infinity

Second, we consider the applied far-field equilibrium displacement field

$$u_1 = -\frac{1}{2R_{23}}\left(x_2^2 + \frac{\nu_m}{1-\nu_m}x_1^2\right) - \frac{1}{2R_{32}}\left(x_3^2 + \frac{\nu_m}{1-\nu_m}x_1^2\right) + (\tilde{\beta}_{13} + \tilde{\beta}_{12})x_1^2 - 2\frac{1-\nu_m}{1-2\nu_m}(\tilde{\beta}_{13}x_2^2 + \tilde{\beta}_{12}x_3^2)$$

$$u_2 = \frac{x_2 x_1}{R_{23}}, \quad u_3 = \frac{x_3 x_1}{R_{32}} \quad \text{for } r \to \infty \tag{C5}$$

from which the general representation Eq. (6) can be obtained by using superposition and adding torsion. The bending we treat is plane-strain bending, while Sen [13] and Das [14] have considered a pure (uniaxial-stress) bending. Their case is recovered by redefining coefficients $1/R_{23}$ and $1/R_{32}$ as follows

$$\frac{1}{R_{23}} = -\frac{\nu_m A}{E_m} + \frac{C}{E_m} \quad \text{and} \quad \frac{1}{R_{32}} = \frac{A}{E_m} - \frac{\nu_m C}{E_m} \tag{C6}$$

where A and C are arbitrary constants and E_m is the elastic modulus of the matrix material. The case $C=0$ is that analyzed in Refs. [13,14], and this is sufficient to solve the general case Eq. (C6) via superposition. We note also that the modes defined by coefficients $\tilde{\beta}_{ij}$ can be redefined in a way similar to Eq. (C6), and again by superposition it is sufficient to solve for the case

$$\tilde{\beta}_{13} = -\frac{1-2\nu}{3-4\nu}\tilde{\beta}_{12} \tag{C7}$$

In polar coordinates, the far-field representation Eq. (C5) with Eq. (C6) (taking $C=0$ and all other coefficients null) has the same structure as Eq. (C5) with Eq. (C7) (with all other coefficients null). This is

$$u_r = Br^2\cos\phi\sin\vartheta(c_1 + c_2\cos 2\vartheta)$$

$$u_\vartheta = Br^2\cos\phi\cos\vartheta(c_3 + c_2\cos 2\vartheta)$$

$$u_\phi = Br^2\sin\phi(c_4 + c_5\cos 2\vartheta) \tag{C8}$$

where

$$B = \frac{A}{4E_m}, \quad c_1 = c_4 = c_3 + 4 = 1 - \nu_m, \quad c_2 = c_5 = 1 + \nu_m \tag{C9}$$

for bending, while

$$B = -\frac{2(1-\nu_m)\tilde{\beta}_{12}}{(1-2\nu_m)(3-4\nu_m)}, \quad c_1 = c_3 = -c_4 = 1 - \nu_m,$$

$$c_2 = -c_5 = 2 - 3\nu_m \tag{C10}$$

for the mode defined by coefficients $\tilde{\beta}_{ij}$.

The solution to this applied far-field displacement field that satisfies equilibrium everywhere, and displacement and traction continuity across the inclusion-matrix boundary $r=a$, is

$$u_r = B\cos\phi\sin\vartheta\left[r^2(c_1 + c_2\cos 2\vartheta) + \frac{a^5(k_1 + k_2\cos 2\vartheta)}{r^3} + \frac{a^7(k_3 + k_4\cos 2\vartheta)}{r^5}\right]$$

$$u_\vartheta = B\cos\phi\cos\vartheta\left[r^2(c_3 + c_2\cos 2\vartheta) + \frac{a^5(k_5 + k_6\cos 2\vartheta)}{r^3} + \frac{a^7(k_7 + k_8\cos 2\vartheta)}{r^5}\right]$$

$$u_\phi = B\sin\phi\left[r^2(c_4 + c_5\cos 2\vartheta) + \frac{a^5(k_9 + k_{10}\cos 2\vartheta)}{r^3} + \frac{a^7(k_{11} + k_{12}\cos 2\vartheta)}{r^5}\right] \tag{C11}$$

in material outside the inclusion, and

$$u_r = Br^2\cos\phi\sin\vartheta\left(\frac{c_0}{r^2} + m_1 + m_2\cos 2\vartheta\right)$$

$$u_\vartheta = Br^2\cos\phi\cos\vartheta\left(\frac{c_0}{r^2} + m_3 + m_2\cos 2\vartheta\right)$$

$$u_\phi = Br^2\sin\phi\left(-\frac{c_0}{r^2} + m_4 + m_5\cos 2\vartheta\right) \tag{C12}$$

in material inside the inclusion. (The Sen [14] solution violates displacement continuity across $r=a$ since it is missing the c_0 terms in Eq. (C12).) All coefficients appearing in the above Eqs. (C11) and (C12) are dimensionless and are defined as

$$k_1 = \frac{12k_8(1-4\nu_m) - 5[4c_1(2-3\nu_m) - 3c_2 - 3m_1(1-4\nu_m)]}{15(1-4\nu_m)} - \frac{15c_2 - 25m_1 + k_8(22-28\nu_m)}{15(1-4\nu_i)}$$

$$k_2 = \frac{14k_8(3-2\nu_m)}{15}, \quad k_3 = -\frac{4k_8}{5}, \quad k_4 = -\frac{4k_8}{3},$$

$$k_5 = -\frac{k_1}{2} - k_{10} + \frac{k_2(1+2\nu_m)}{2(3-2\nu_m)}$$

$$k_6 = k_2 - \frac{5k_2}{2(3-2\nu_m)}, \quad k_7 = -\frac{7k_8}{15}$$

$$k_8 = \frac{15c_2[E_m(1+\nu_i) - E_i(1+\nu_m)]}{2E_i(1+\nu_m)(11-14\nu_m) + E_m(13-7\nu_m)(1+\nu_i)}$$

$$k_9 = \frac{k_1}{2} - \frac{k_2(1-\nu_m)}{3-2\nu_m}$$

$$k_{10} = \frac{E_m(1+\nu_i) - E_i(1+\nu_m)}{E_i(1+\nu_m) + 4E_m(1+\nu_i)}\left\{\frac{2\nu_m(c_1+3c_2+2c_3) - 3c_1 - c_3}{1-4\nu_m} + \frac{c_2[E_i(27-\nu_m-28\nu_m^2) + 4E_m(2+7\nu_m)(1+\nu_i)]}{2E_i(1+\nu_m)(11-14\nu_m) + E_m(1+\nu_i)(13-7\nu_m)}\right\}$$

$$k_{11} = -\frac{k_8}{5}, \quad k_{12} = -\frac{k_8}{3}$$

$$c_0 = \frac{-5(5c_1-3c_2)}{15(1-4\nu_m)} - \frac{5(3c_2-5m_1) + k_8(22-28\nu_m)}{15(1-4\nu_i)}$$

$$m_1 = \frac{5(1-\nu_m)[26c_1 - 3c_2 - 14(c_1+3c_2)\nu_m]}{3(1-4\nu_m)(13-7\nu_m)} - \frac{(5c_1-3c_2)(1-\nu_m)[5E_m + 2E_i(1+\nu_m)]}{3(1-4\nu_m)[2E_m(2-3\nu_i) + E_i(1+\nu_m)]} - \frac{42c_2E_i(1-\nu_m)(1+\nu_m)(11-14\nu_m)}{(13-7\nu_m)[2E_i(1+\nu_m)(11-14\nu_m) + E_m(1+\nu_i)(13-7\nu_m)]}$$

$$m_2 = c_2 + \frac{2k_8(11-14\nu_m)}{15}$$

$$m_3 = -k_{10} + \frac{1}{2}\left(c_1 + 2c_3 + \frac{5c_1-3c_2}{1-4\nu_m} + \frac{3c_2}{1-4\nu_i}\right) - \frac{m_1(3-2\nu_i)}{1-4\nu_i} + \frac{k_8[27+24\nu_i - 28\nu_m(1+2\nu_i)]}{15(1-4\nu_i)}$$

$$m_4 = \frac{m_1(3-2\nu_i) - 2m_2(1-\nu_i)}{1-4\nu_i},$$

$$m_5 = -m_3 + \frac{m_2(1+2\nu_i) - m_1(3-2\nu_i)}{1-4\nu_i}$$

where E_m, E_i, and ν_m, ν_i are the elastic moduli and Poisson's ratios of the matrix and inclusion materials, respectively.

References

[1] Lakes, R. S., 1983, "Size Effects and Micromechanics of a Porous Solid," J. Mater. Sci., **18**, pp. 2572–2581.

[2] Lakes, R. S., 1986, "Experimental Microelasticity of Two Porous Solids," Int. J. Solids Struct., **22**, pp. 55–63.

[3] Lakes, R. S., 1995, "Experimental Methods for Study of Cosserat Elastic Solids and Other Generalized Continua," *Continuum Models for Materials with Micro-Structure*, H. B. Muhlhaus, ed., Wiley, New York, pp. 1–22.

[4] Yang, J. F. C., and Lakes, R. S., 1981, "Transient Study of Couple Stress in Compact Bone: Torsion," J. Biomech. Eng., **103**, pp. 275–279.

[5] Yang, J. F. C., and Lakes, R. S., 1982, "Experimental Study of Micropolar and Couple Stress Elasticity in Bone in Bending," J. Biomech., **15**, pp. 91–98.

[6] Gauthier, R. D., and Jahsman, W. E., 1975, "A Quest for Micropolar Elastic Constants," ASME J. Appl. Mech., **42**, pp. 369–374.

[7] Gauthier, R. D., 1982, "Experimental Investigation on Micropolar Media," *Mechanics of Micropolar Media*, O. Brulin and R. K. T. Hsieh, eds., CISM Lecture Notes, World Scientific, Singapore, pp. 395–463.

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

[9] Hashin, Z., 1959, "The Moduli of an Elastic Solid Containing Spherical Particles of Another Elastic Material," *Non-Homogeneity in Elasticity and Plasticity*, W. Olszak, ed., Pergamon, New York, pp. 463–478.

[10] Hashin, Z., and Rosen, W. B., 1964, "The Elastic Moduli of Fiber-Reinforced Materials," J. Appl. Mech., **31**, pp. 223–232.

[11] Muskhelishvili, N. I., 1953, *Some Basic Problems of the Mathematical Theory of Elasticity*, Noordhoff, Groningen, Holland.

[12] Sendeckyj, G. P., 1970, "Elastic Inclusion Problems in Plane Elastostatics," Int. J. Solids Struct., **6**, pp. 1535–1543.

[13] Sen, B., 1933, "On the Concentration of Stresses Due to a Small Spherical Cavity in a Uniform Beam Bent by Terminal Couples," Bull. Calcutta Math. Soc., **25**, pp. 107–114.

[14] Das, S. C., 1953, "On the Stresses Due to a Small Spherical Inclusion in a Uniform Beam Under Constant Bending Moment," Bull. Calcutta Math. Soc., **45**, pp. 55–63.

[15] Koiter, W. T., 1964, "Couple-Stresses in the Theory of Elasticity, Parts I and II," Proc. K. Ned. Akad. Wet., Ser. B: Phys. Sci., **67**, pp. 17–44.

[16] Nowacki, W., 1986, *Theory of Asymmetric Elasticity*, Pergamon, Oxford.

[17] Berglund, K., 1982, "Structural Models of Micropolar Media," *Mechanics of Micropolar Media (CISM Lecture Notes)*, O. Brulin and R. K. T. Hsieh, eds., World Scientific, Singapore, pp. 35–86.

[18] Dean, D. L., and Urgate, C. P., 1968, "Field Solutions for Two-Dimensional Frameworks" Int. J. Mech. Sci., **10**, pp. 315–339.

[19] Bažant, Z. P., and Christensen, M., 1972, "Analogy Between Micropolar Continuum and Grid Frameworks Under Initial Stress," Int. J. Solids Struct., **8**, pp. 327–346.

[20] Banks, C. B., and Sokolowski, U., 1968, "On Certain Two-Dimensional Applications of Couple-Stress Theory," Int. J. Solids Struct., **4**, pp. 15–29.

[21] Wang, X. L., and Stronge, W. J., 1999, "Micropolar Theory for Two-Dimensional Stresses in Elastic Honeycomb," Proc. R. Soc. London, Ser. A, **445**, pp. 2091–2116.

[22] Achenbach, J. D., and Herrmann, G., 1968, "Dispersion of Free Harmonic Waves in Fibre-Reinforced Composites," AIAA J., **6**, pp. 1832–1836.

[23] Beran, M. J., and McCoy, J. J., 1970, "Mean Field Variations in a Statistical Sample of Heterogeneous Linearly Elastic Solids," Int. J. Solids Struct., **6**, pp. 1035–1054.

[24] Forest, S., 1998, "Mechanics of Generalized Continua: Construction by Homogenization," J. Phys. IV, **8**, pp. 39–48.

[25] Ostoja-Starzewski, M., Boccara, S., and Jasiuk, I., 1999, "Couple-Stress Moduli and Characteristic Length of Composite Materials," Mech. Res. Commun., **26**, pp. 387–397.

[26] Bouyge, F., Jasiuk, I., and Ostoja-Starzewski, M., 2001, "A Micromechanically Based Couple-Stress Model of an Elastic Two-Phase Composite," Int. J. Solids Struct., **38**, pp. 1721–1735.

T. Monprapussorn
Department of Civil Engineering,
South-East Asia University,
Bangkok 10160, Thailand

C. Athisakul

S. Chucheepsakul[1]
Mem. ASME
e-mail: somchai.chu@kmutt.ac.th

Department of Civil Engineering,
King Mongkut's University of Technology
Thonburi,
Bangkok 10140, Thailand

Nonlinear Vibrations of an Extensible Flexible Marine Riser Carrying a Pulsatile Flow

The influence of transported fluid on static and dynamic behaviors of marine risers is investigated. The internal flow of the transported fluid could have a constant, a linear, or a wave velocity. The riser pipe may possibly experience the conditions of high extensibility, flexibility, and large displacements. Accordingly, the mathematical riser models should be governed by the large strain formulations of extensible flexible pipes transporting fluid. Nonlinear hydrodynamic dampings due to ocean wave–pipe interactions implicate the high degree of nonlinearity in the riser vibrations, for which numerical solutions are determined by the state–space–finite-element method. It is revealed that the impulsive acceleration of internal flow could seriously relocate the vibrational equilibrium positions of the riser pipe. The fluctuation of the pulsatile flow relatively introduces the expansion of amplitudes and the reduction of frequencies of the riser vibrations. The pulsatile frequencies of the internal flow in wave aspect could reform the oscillation behavior of the conveyor pipe. [DOI: 10.1115/1.2711226]

Keywords: internal pulsating flow, large strain formulations, extensible flexible risers, Poisson's ratio effect

1 Introduction

The employment of flexible marine risers as the transporters of crude resource fluids drilled from beneath the deep ocean has been operated over the past decade. In this time, the analysis and design of those have been developed progressively to accumulate accuracy in predicting and controlling mechanical behavior under serviceability requirements [1–4]. Nevertheless, because of various sources of nonlinearity to which the riser may be subjected, the development so far has never been enough, and still must be continued to achieve the proper way to treat an assortment of situations that the riser may encounter. In the applied mechanics research area, the flexible riser may be assorted as follows:

1. The tensioned pipe undergoing large displacements;
2. The extensible elastica under fluid pressure fields; and
3. The largely sagged conveyer of the pulsating flow.

Based on the following literature review, it is shown that these nonlinear mechanics have not yet been elucidated comprehensively, whether under onshore or in offshore environments. The tensioned pipes were studied by numerous research works [5–10]; however, they were restricted to the small displacement theory. For the analysis of elasticas, the extensible elastica theories have been well established [11,12]; however, the elastica behavior under fluid pressure fields versus the Poisson's ratio effect has never been imparted. Finally, the pulsatile flows induced vibrations of pipe structures were investigated by a number of researchers [13–15], but were limited to the cases of straight or taut pipes. Based on the literature of marine risers [16–24], the combination effects of the riser extensibility and the pulsating flow have never been studied. Almost all of the research works concerning the internal flow effects on the riser behavior are focused on steady flow, inextensibility condition, and otherwise slug flow.

Adding to the literature, this paper therefore aims to explore the effects of internal pulsating flow on static and dynamic behaviors of the extensible riser pipe. The factors considering the importance of the effects were discussed in the paper by Chucheepsakul et al. [25]. The following assumptions are stipulated in the present analysis.

1. The pipe materials are linearly elastic;
2. At the undeformed state, the pipes are initially straight, and have no residual stresses;
3. The pipes are sufficiently thick walled to suppose that, ideally, their cross sections remain circular after change of cross-sectional size due to the effects of fluid pressures and Poisson's ratio, so that the elastic rod theories and the apparent tension concept are usable for handling the effects of large and radial displacements of pipe wall, and Brazier's effect or flattening of bent tubes is negligible;
4. Longitudinal strain is large, but shear strain is insignificant for elastic rods with high slenderness ratio;
5. Plane sections of the pipes remain plane at all states;
6. The internal and external fluids are inviscid, incompressible, and irrotational. Their densities are uniform along arc lengths of the pipes;
7. The internal flow is the one-dimensional plug laminar flow;
8. Morison's equation is adopted for evaluating external hydrodynamic forces of external fluid. The distributed couple induced by a flow asymmetry due to vortex shedding is neglected;
9. The effect of rotary inertia is negligible; and
10. Although a realistic characterization of pulsatile flow velocity in practice might require a statistical description, to gain insight into the problem, the flow velocity is represented herein as

$$V_i = V_{io} + V_{id} = V_{io} + a_{io}t + V_{ia}\cos\omega_i t \tag{1}$$

where V_{io} and V_{id} are the static and dynamic parts of the internal flow velocity, respectively.

The static velocity V_{io} is not constant, and thus represents the nonuniform internal flow. It varies along the pipe's arc length s_0

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received February 2, 2004; final manuscript received November 22, 2006. Review conducted by Sanjay Govindjee.

due to change of the cross-sectional size of the pipe element under the gradient of pressure fields. In the updated Lagrangian control volume [25], it is expressed that $V_{io}=\bar{V}_i/(1-\varepsilon_o)$, in which $\bar{V}_i$ is the constant mean flow velocity of a fully developed flow, and ε_o is the statical axial strain varying along the arc length of the pipe.

The dynamic velocity $a_{io}t$ represents a transient state of internal flow, where the constant step of the internal flow acceleration a_{io} is applied. The oscillatory flow of the transported fluid caused by the operation of a high-pressure pump may be described by the pulsatile velocity $V_{ia}\cos\omega_i t$, where V_{ia} is the fluctuation amplitude, and ω_i the pulsation frequency.

Once differentiating Eq. (1) is to yield the internal flow acceleration

$$a_i = \frac{\partial V_{id}}{\partial t} = a_{io} - V_{ia}\omega_i \sin \omega_i t \tag{2}$$

To undertake the effect of high extensibility of the risers, the mathematical formulations based upon the work-energy principles and the extensible elastica theory [25] are adopted.

2 Mathematical Formulations and Solution Methods

Based upon the large strain formulations [25] of extensible flexible marine pipes transporting fluid as shown in Fig. 1, the total governing equations for nonlinear dynamic, large amplitude vibrations of the risers could be obtained in the updated Lagrangian Cartesian coordinates (x,y) as

$$\mathbf{m}\ddot{\mathbf{x}} + \mathbf{c}\dot{\mathbf{x}} + \mathbf{g}\dot{\mathbf{x}}' + (\mathbf{k}_{b1}\mathbf{x}'')'' + (\mathbf{k}_{b2}\mathbf{x}'')' + \mathbf{k}_{t1}\mathbf{x}'' + \mathbf{k}_{t2}\mathbf{x}' = \mathbf{f} \tag{3}$$

where $\mathbf{x}$ is the position vector; $\mathbf{m}$ the total mass matrix; $\mathbf{c}$ the hydrodynamic damping matrix; $\mathbf{g}$ the gyroscopic matrix; $\mathbf{k}_{b1}$ and $\mathbf{k}_{b2}$ the bending stiffness matrices; $\mathbf{k}_{t1}$ and $\mathbf{k}_{t2}$ the axial stiffness matrices; and $\mathbf{f}$ the external load vector (see Ref. [25] for more details).

For the vibrations with infinitesimal amplitudes, the axial strain ϵ in the stiffness matrices can be approximated by the two-term binomial expansion

$$\varepsilon = \frac{s' - \bar{s}'}{s_0'} = \varepsilon_o + \left(\frac{s'}{s_0'} - 1\right) = \underbrace{\varepsilon_o}_{\text{static strain}} + \underbrace{(\sqrt{1+2\gamma_d} - 1)}_{\text{dynamic strain } \varepsilon_d} \approx \varepsilon_o + \gamma_d \tag{4}$$

where $\bar{s}'$, s_0', and s' are the differential arc lengths at the undeformed, the equilibrium, and the displaced states; and γ_d is the dynamic updated Green strain that can be expressed as

$$\gamma_d = \frac{1}{s_0'^2}\left(x_o'u' + y_o'v' + \frac{u'^2}{2} + \frac{v'^2}{2}\right) \approx \frac{1}{s_0'^2}(x_o'u' + y_o'v') \tag{5}$$

in which (x_o,y_o) is the Cartesian coordinates at the equilibrium state; and (u,v) the displacement vector. Therefore, the apparent tension [25] corresponding to the axial strain, included in the axial stiffness matrices, could be written as

$$\begin{aligned} T_a &= T + 2\nu(p_e A_{eo} - p_i A_{io}) = EA_{P0}\varepsilon \\ &\approx \underbrace{T_{ao}}_{\text{static tension}} + \underbrace{\frac{EA_{P0}}{s_0'^2}(x_o'u' + y_o'v')}_{\text{dynamic tension}} \end{aligned} \tag{6}$$

where E is the elastic modulus; ν Poisson's ratio; A_{eo}, A_{io}, A_{P0} the outside, inside, and cross-sectional area of the pipe at the equilibrium state; p_e and p_i the external and internal fluid pressures; and T is the true wall tension of which determination is given in the Appendix .

By separation of variables, the displacement vector is assumed as

$$\{\mathbf{d}\} = \{u \quad v\}^T = [\mathbf{N}(y_o)]\{\mathbf{d}_n(t)\} \tag{7}$$

where the generalized coordinates of the nodal displacements for each element are

$$\{\mathbf{d}_n\} = \{u_1 \quad u_1' \quad u_1'' \quad v_1 \quad v_1' \quad v_1'' \,|\, u_2 \quad u_2' \quad u_2'' \quad v_2 \quad v_2' \quad v_2''\}^T \tag{8}$$

and the shape function matrix at the displaced state is

$$[\mathbf{N}] = \left[\begin{array}{cccccc|cccccc} N_{51} & N_{52} & N_{53} & 0 & 0 & 0 & N_{54} & N_{55} & N_{56} & 0 & 0 & 0 \\ 0 & 0 & 0 & N_{51} & N_{52} & N_{53} & 0 & 0 & 0 & N_{54} & N_{55} & N_{56} \end{array}\right] \tag{9}$$

Note that N_{5i} is the coefficient of the fifth-order polynomial shape functions.

Substituting Eqs. (1), (2), (6), and (7) into Eq. (3) together with neglecting the higher-order terms, eliminating the time-independent terms, and following the standard procedures of the Galerkin finite element method, the system of partial differential Eq. (3) could be transformed into the system of ordinary differential equations

$$[\mathbf{m}^{(e)}]\{\ddot{\mathbf{d}}_n\} + ([\mathbf{c}^{(e)}] + [\mathbf{g}^{(e)}])\{\dot{\mathbf{d}}_n\} + [\mathbf{k}^{(e)}]\{\mathbf{d}_n\} = \{\mathbf{f}^{(e)}\} \tag{10}$$

where the element mass matrix [25] is

$$[\mathbf{m}^{(e)}] = \int_\alpha \left\{[\mathbf{N}]^T s_0'(m_{P0} + m_{io} + C_{ao}^*)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}[\mathbf{N}]\right\} d\alpha \tag{11a}$$

the element hydrodynamic damping matrix [25] is

$$[\mathbf{c}^{(e)}] = \int_\alpha \left\{[\mathbf{N}]^T s_0' \begin{bmatrix} C_{eqxo}^* & C_{eqxyo}^* \\ C_{eqxyo}^* & C_{eqyo}^* \end{bmatrix}[\mathbf{N}]\right\} d\alpha \tag{11b}$$

the element gyroscopic matrix [25] is

$$[\mathbf{g}^{(e)}] = \int_\alpha \left\{[\mathbf{N}]^T m_{io} V_{io} \begin{bmatrix} 2 - \dfrac{x_o'^2}{s_0'^2} & -\dfrac{x_o' y_o'}{s_0'^2} \\ -\dfrac{x_o' y_o'}{s_0'^2} & 2 - \dfrac{y_o'^2}{s_0'^2} \end{bmatrix}[\mathbf{N}']\right\} d\alpha \tag{11c}$$

the element stiffness matrix [25] is

$$[\mathbf{k}^{(e)}] = [\mathbf{k}_{b1}^{(e)}] + [\mathbf{k}_{b2}^{(e)}] + [\mathbf{k}_{t1}^{(e)}] + [\mathbf{k}_{t2}^{(e)}] \tag{11d}$$

in which the bending stiffness matrix of the fourth-order derivative [25] is

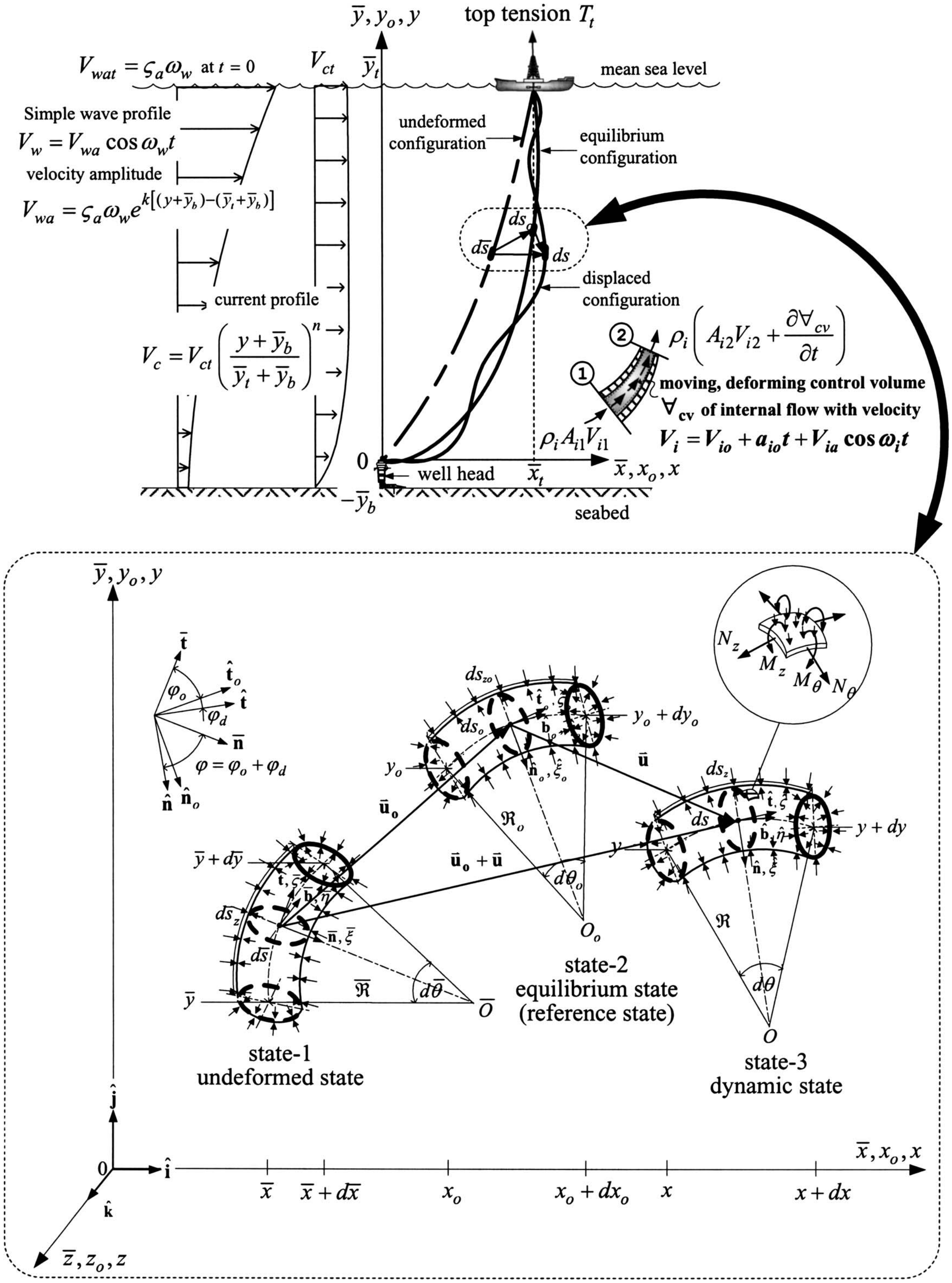


Fig. 1 An extensible flexible marine riser carrying a pulsatile flow (*a*); and schematic of deformations (*b*)

$$[\mathbf{k}_{b1}^{(e)}] = \int_\alpha \left\{ [\mathbf{N}'']^T \frac{B_o}{s_0'^5} \begin{bmatrix} y_o'^2 & -x_o' y_o' \\ -x_o' y_o' & x_o'^2 \end{bmatrix} [\mathbf{N}''] \right\} d\alpha \quad (11e)$$

the bending stiffness matrix of the third-order derivative [25] is

$$[\mathbf{k}_{b2}^{(e)}] = \int_\alpha \left\{ [\mathbf{N}']^T \frac{B_o \kappa_o}{s_0'^4} \begin{bmatrix} 2x_o' y' & y_o'^2 - x_o'^2 \\ y_o'^2 - x_o'^2 & -2x_o' y_o' \end{bmatrix} [\mathbf{N}''] \right\} d\alpha \quad (11f)$$

the axial stiffness matrix of the second-order derivative [25] is

$$[\mathbf{k}_{t1}^{(e)}] = \int_\alpha \left\{ \begin{array}{c} [\mathbf{N}']^T\left(\dfrac{T_{ao}-m_{io}V_{io}^2}{s_0'}\right)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}[\mathbf{N}'] \\ +[\mathbf{N}']^T\dfrac{EA_{P0}}{s_0'^3}\begin{bmatrix} x_o'^2 & x_o'y_o' \\ x_o'y_o' & y_o'^2 \end{bmatrix}[\mathbf{N}'] \end{array} \right\} d\alpha \quad (11g)$$

the axial stiffness matrix of the first-order derivative [25] is

$$[\mathbf{k}_{t2}^{(e)}] = \int_\alpha \left\{ [\mathbf{N}]^T\left(\frac{m_{io}V_{io}V_{io}'}{s_0'^2}\right)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}[\mathbf{N}'] \right\} d\alpha \qquad (11h)$$

and the element hydrodynamic excitation vector [25] is

$$\{\mathbf{f}^{(e)}\} = \int_\alpha [\mathbf{N}]^T s_0' \left\{ \begin{array}{c} C_{Dxo}^*(2V_cV_w+V_w^2)+C_{Mo}^*\dot{V}_w-\dfrac{m_{io}x_o'a_i}{s_0'} \\ C_{Dxy1o}^*(2V_cV_w+V_w^2)-\dfrac{m_{io}y_o'a_i}{s_0'} \end{array} \right\} d\alpha \qquad (11i)$$

Please note that the expressions in detail of the coefficients of added mass $m_{Po}+m_{io}+C_{ao}^*$, of inertia C_{Mo}^*, of drag $\{C_{Dxo}^*, C_{Dxy1o}^*\}$, and of equivalent damping $\{C_{eqxo}^*, C_{eqxyo}^*, C_{eqyo}^*\}$, the bending rigidity $B_o\kappa_o$, the current and wave velocities $\{V_c, V_w\}$ are given in the paper by Chucheepsakul et al. [25]

Assembling the element equations, the global system of finite element equations can be obtained as

$$[\mathbf{M}]\{\ddot{\mathbf{D}}_n\}+([\mathbf{C}]+[\mathbf{G}])\{\dot{\mathbf{D}}_n\}+[\mathbf{K}]\{\mathbf{D}_n\}=\{\mathbf{F}\} \qquad (12)$$

where

$$\{\mathbf{D}_n\} = \sum_{i=1}^{n\text{elem}} [\mathbf{d}_n] \qquad (13a)$$

$$[\mathbf{M}] = \sum_{i=1}^{n\text{elem}} [\mathbf{m}^{(e)}] \qquad (13b)$$

$$[\mathbf{C}] = \sum_{i=1}^{n\text{elem}} [\mathbf{c}^{(e)}] \qquad (13c)$$

$$[\mathbf{G}] = \sum_{i=1}^{n\text{elem}} [\mathbf{g}^{(e)}] \qquad (13d)$$

$$[\mathbf{K}] = \sum_{i=1}^{n\text{elem}} [\mathbf{k}^{(e)}] \qquad (13e)$$

$$[\mathbf{F}] = \sum_{i=1}^{n\text{elem}} [\mathbf{f}^{(e)}] \qquad (13f)$$

are the global nodal displacement, the total mass matrix, the total hydrodynamic damping matrix, the total gyroscopic matrix, the total stiffness matrix, and the total hydrodynamic excitation vector, respectively, in which $n$elem is the number of finite elements. The second-order model of Eq. (12) has to be transformed to the first-order model via the state space formulation [26] as

$$\{\dot{\mathbf{X}}_n\} = [\mathbf{A}]\{\mathbf{X}_n\}+\{\mathbf{B}\} \qquad (14)$$

where

$$\{\mathbf{X}_n\} = \begin{Bmatrix} \mathbf{D}_n \\ \mathbf{V}_n \end{Bmatrix} \qquad (15a)$$

$$[\mathbf{A}] = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}(\mathbf{C}+\mathbf{G}) \end{bmatrix}, \quad \{\mathbf{B}\} = \begin{Bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}\mathbf{F} \end{Bmatrix} \qquad (15b)$$

are the state vector of displacement $\mathbf{D}_n$ and velocity $\mathbf{V}_n$, the state transition, and the deterministic input matrices. The initial-value problem of the highly nonlinear state Eq. (14) in association with the initial condition equations can be solved to determine the nonlinear responses by the explicit time integration. The drawback to the explicit methods dealing with the conditionally stable state of the critical step size of time is overcome by the Gear's stiff method that includes the automatically adaptive time-step-size algorithm [27]. This algorithm automatically improves the time-step size during the integration process so that the absolute error criterion is achieved.

3 Numerical Studies and Results

The basic equilibrium geometries of the riser pipes commonly found in marine riser operations are concerned with both the vertical-straight lines and the inclined-catenary curves with top end offset. The vertical-straight line configuration is mostly set for static equilibrium of the high strength steel, rigid riser, which is supplied by a high tension and a mooring system at the top end support (on a floating vessel). The inclined catenary is the initial profile of the composite flexible riser, which allows the top end excursion and the large displacements of the pipe.

In this section, the effects of internal flow of transported fluid on both rigid and flexible risers are demonstrated. However, the rigid riser results are referred to merely to validate the numerical results, because of its well-known linear behavior under infinitesimal static displacements. The advanced mechanics of the effect of internal pulsating flow will be studied parametrically through the flexible riser analysis.

The parameters utilized for the flexible riser analysis are given in Table 1. Some of them are manipulated into dimensionless forms for the sake of comprehension in the parametric study of the effects of pipe's extensibility and internal flow. For the reason that the risers as the tensioned pipes are subjected to the large top tension as compared with the flexural rigidity, it is convenient to use the applied tension at the top end of the riser T_t as the basis for the parametric normalization to obtain the following dimensionless quantities

$$\hat{V}_{io} = \bar{V}_i\sqrt{\frac{\bar{m}_i}{T_t}} \qquad (16a)$$

$$\hat{a}_{io} = a_{io}\left(\frac{\bar{m}_i L}{T_t}\right) \qquad (16b)$$

$$\hat{V}_{ia} = V_{ia}\sqrt{\frac{\bar{m}_i}{T_t}} \qquad (16c)$$

$$\hat{\omega}_i = \omega_i L\sqrt{\frac{\bar{m}_i}{T_t}} \qquad (16d)$$

where $\bar{m}_i$ is the transported fluid mass per unit undeformed length of the pipe. The parameters $\hat{V}_{io}$, $\hat{a}_{io}$, $\hat{V}_{ia}$, and $\hat{\omega}_i$ denote the effects of the mean flow velocity, the constant-step acceleration, the fluctuation amplitude, and the pulsation frequency of transported fluid, respectively.

3.1 Validation of Numerical Results. In order to verify the validity of the present model, the specimen of a vertical production riser given by Moe and Chucheepsakul [17] is adopted for the comparative study. The parameters used in the calculation are shown in Table 1. The natural frequency of the rigid production riser tends to decrease along with the elevation of internal flow speed, as shown in Table 2. For the inextensible analysis exclud-

Table 1 Input parameters of the marine riser specimens

Parameter	Production risers	Flexible risers
Elastic modulus E (N/m^2)	0.207×10^{12}	0.207×10^{12}
External diameter of the pipe $\bar{D}_e$ (m)	0.26	0.26
Internal diameter of the pipe $\bar{D}_i$ (m)	0.20	0.20
Density of pipe material ρ_P (kg/m^3)	7850	7850
Density of external fluid ρ_e (kg/m^3)	1025	1025
Density of internal fluid ρ_i(kg/m^3)	998	998
Poisson's ratio ν	0.50	0.50
Static offset of the vessel $\bar{x}_t$ (m)	0	70
Vertical depth of risers $\bar{y}_t$ (m)	300	300
Height of the bottom support over seabed $\bar{y}_b$ (m)	~0	~0
Applied top tension T_t (N)	476,200	(basis)[a]
Normal drag coefficient C_{Dn}	0.70	0.70
Tangential drag coefficient C_{Dt}	0.03	0.03
Current velocity at mean sea level V_{ct} (m/s)	0	0.20
Internal flow velocity $\bar{V}_i$ (m/s)	0.00	(nondim.)[b]
Added mass coefficient C_a	1.00	1.00
Wave amplitude ζ_a (m)	6	6
Wave frequency ω_w (rad/sec)	0.6	0.6
Wave number k	0.03	0.03
Linear acceleration of int. flow a_{io} (m/s)	0	(nondim.)[b]
Wave velocity amplitude of int. flow V_{ia} (m/s)	0	(nondim.)[b]
Internal flow frequency ω_i (rad/sec)	0	(nondim.)[b]

[a]Basis denotes the basic dimension used for the parametric normalization to obtain the dimensionless quantities.
[b]Nondim denotes the nondimensional quantity to be varied for the parametric study.

ing bending rigidity of the pipe, the analytical solution [17] based upon using the first term of an asymptotic series for the Bessel functions yields the expression of the natural frequency of the riser as

$$\omega_n = \frac{n\pi}{2L}\frac{\left(\sqrt{T_e\big|_{y_o=\bar{y}_t} - m_i\bar{V}_i^2} + \sqrt{T_e\big|_{y_o=0} - m_i\bar{V}_i^2}\right)}{\sqrt{m_P + m_i + C_a^*}} \tag{17}$$

in which n is the mode number; L the span length of the riser; $T_e = T + p_eA_e - p_iA_i$ the effective tension; m_P and m_i the masses of pipe and transported fluid; $C_a^* = \rho_eA_eC_a$ the added mass of external fluid; and $\bar{V}_i$ the constant mean flow velocity of a fully developed flow. The analytical results computed from Eq. (17) are shown in Table 2 as well as the numerical solutions using 20-finite elements [17].

Table 2 The fundamental natural frequencies ω_1 (rad/s) of the vertical production riser conveying fluid with various speeds of internal flow $\bar{V}_i$ (m/s)[a]

	Moe and Chucheepsakul[b] (IA, EBR)		This study (20-finite elements) (EA)	
$\bar{V}_i$ (m/s)	Analytical solution	Numerical solution	EBR	IBR
0	0.2878	0.2890	0.2891	0.3001
5	—	—	0.2881	0.2994
10	0.2838	0.2853	0.2853	0.2972
15	—	—	0.2804	0.2934
20	0.2706	0.2730	0.2731	0.2880
25	—	—	0.2627	0.2809
30	0.2413	0.2478	0.2478	0.2717
35	—	—	0.2224	0.2603
38	(unstable)	0.1704	0.1710	0.2522
40	—	(unstable)	(unstable)	0.2461
45	—	—	—	0.2282
50	—	—	—	0.2052
55	—	—	—	0.1738
60	—	—	—	0.1250
65	—	—	—	(unstable)

[a]IA=inextensible analysis; EA=extensible analysis; EBR=excluding bending rigidity; IBR=including bending rigidity.
[b]See Ref. [17].

The extensible analysis is also carried out using 20-finite elements. In Table 2, it is shown that the results of the extensible analysis (EA) converge nearly to the inextensible analysis (IA) results, and thus indicate that the effect of the pipe's extensibility is quite low for the rigid production riser. The finite element solutions for the step increments of the internal flow velocity $\bar{V}_i$ reveal that the fundamental mode of divergence instability reaches the critical velocity $\bar{V}_{i(\mathrm{cr})}=38.2$ m/s, which yields the negative combined tension at the bottom portion of the riser. In the case where the bending rigidity (B) of the pipe is taken into account in the analysis, the natural frequency of the pipe augments significantly, and the first mode instability takes place at the higher critical flow speed such as $\bar{V}_{i(\mathrm{cr})}=64$ m/s. These results are understandable, showing that bending rigidity contributes the bending stiffness to the system, and therefore enhances the stability criteria of the system.

Figure 2 plots the natural frequencies of the rigid production riser under various internal flow velocities. It is noticed that the shapes of modal vibrations should modify the curvature mode after the critical internal flow velocity. For example, prior to the critical velocity ($\bar{V}_i=0-63$ m/s), the fundamental mode of the transverse vibration u_n holds the explicit single-curvature mode shape. When the critical velocity 64 m/s is approached, the single curvature develops to the implicit double curvature, and maybe bifurcates onto the explicit double curvature mode shape in the second vibration mode.

From $\bar{V}_i=65$ m/s to 83 m/s, the divergent instability overwhelms the single curvature mode, and the double curvature mode occupies the fundamental mode instead. The second bifurcation occurs at the second critical velocity 84 m/s in the same manner that the vibration tends to bifurcate from the double curvature

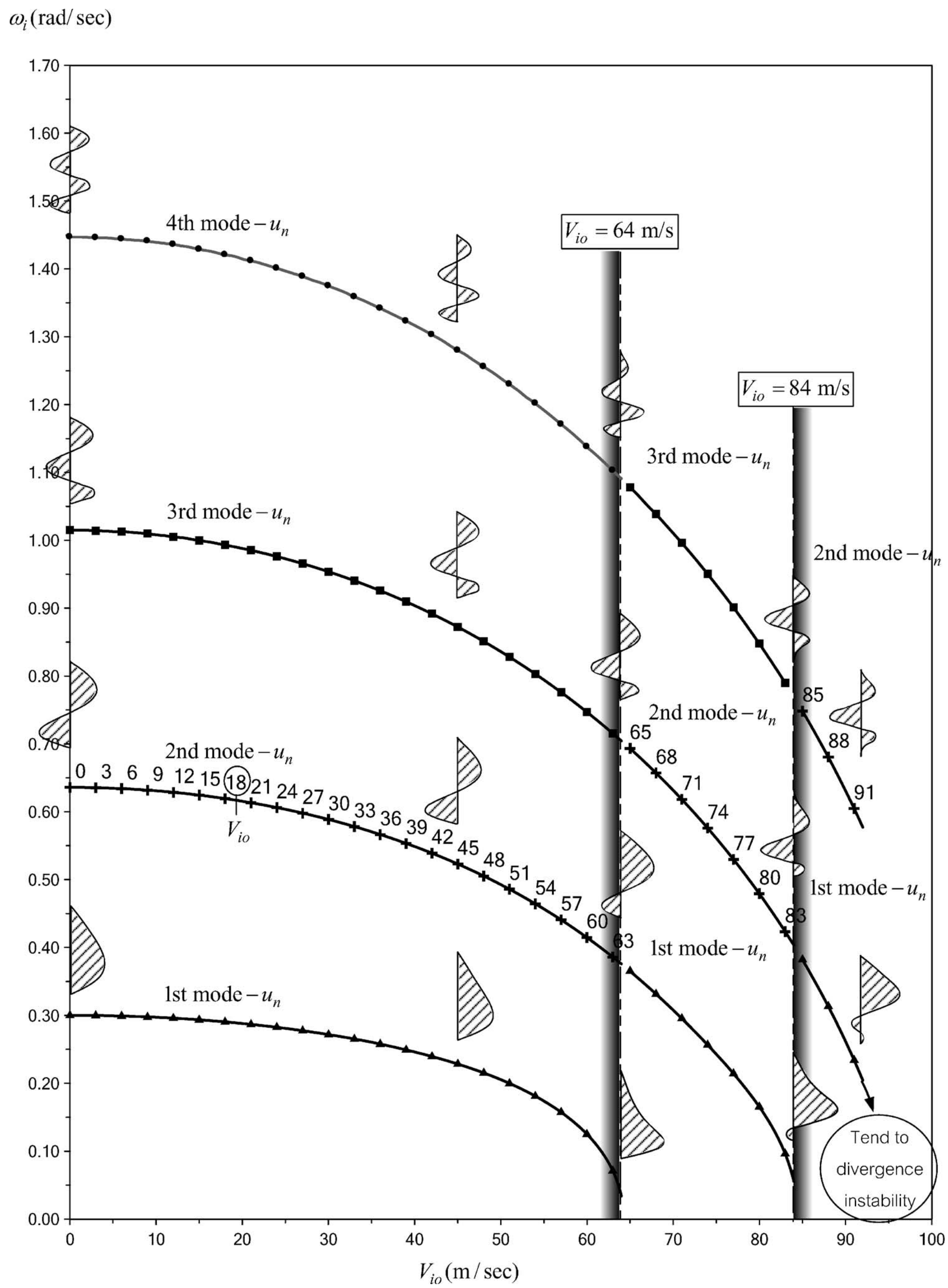


Fig. 2 The effect of fluid transportation rate on natural frequencies and mode shapes of the vertical production risers

mode to the triple curvature mode, which will become the next fundamental vibrational mode instead. The summary of alterations of the curvature modes against the critical state for each vibration mode is given in Table 3.

3.2 The Effects of the Mean Flow Velocity of Transported Fluid. In Fig. 3, it is clearly seen that the steady flow of transported fluid expands the amplitudes of nonlinear vibrations of the riser. On the phase planes in Fig. 3(*c*), the steady internal flow also induces the small increase of the orbit difference of the trajectory to reduce the orbital stability of the pipe motion. The oscillation phase is slightly shifted by the steady internal flow as shown in the time history in Fig. 3(*d*). It agrees with the analytical

Table 3 Summary of alterations of the curvature modes versus the critical state

	The curvature modes (counted by the number of explicit curvatures) for			
$V_{io}=\bar{V}_i$ (m/s)	The first vibrational mode	The 2nd vibrational mode	The third vibrational mode	The fourth vibrational mode
0–63	1	2	3	4
64	1–2	2–3	3–4	4–5
65–83	2	3	4	5
84	2–3	3–4	4–5	5–6

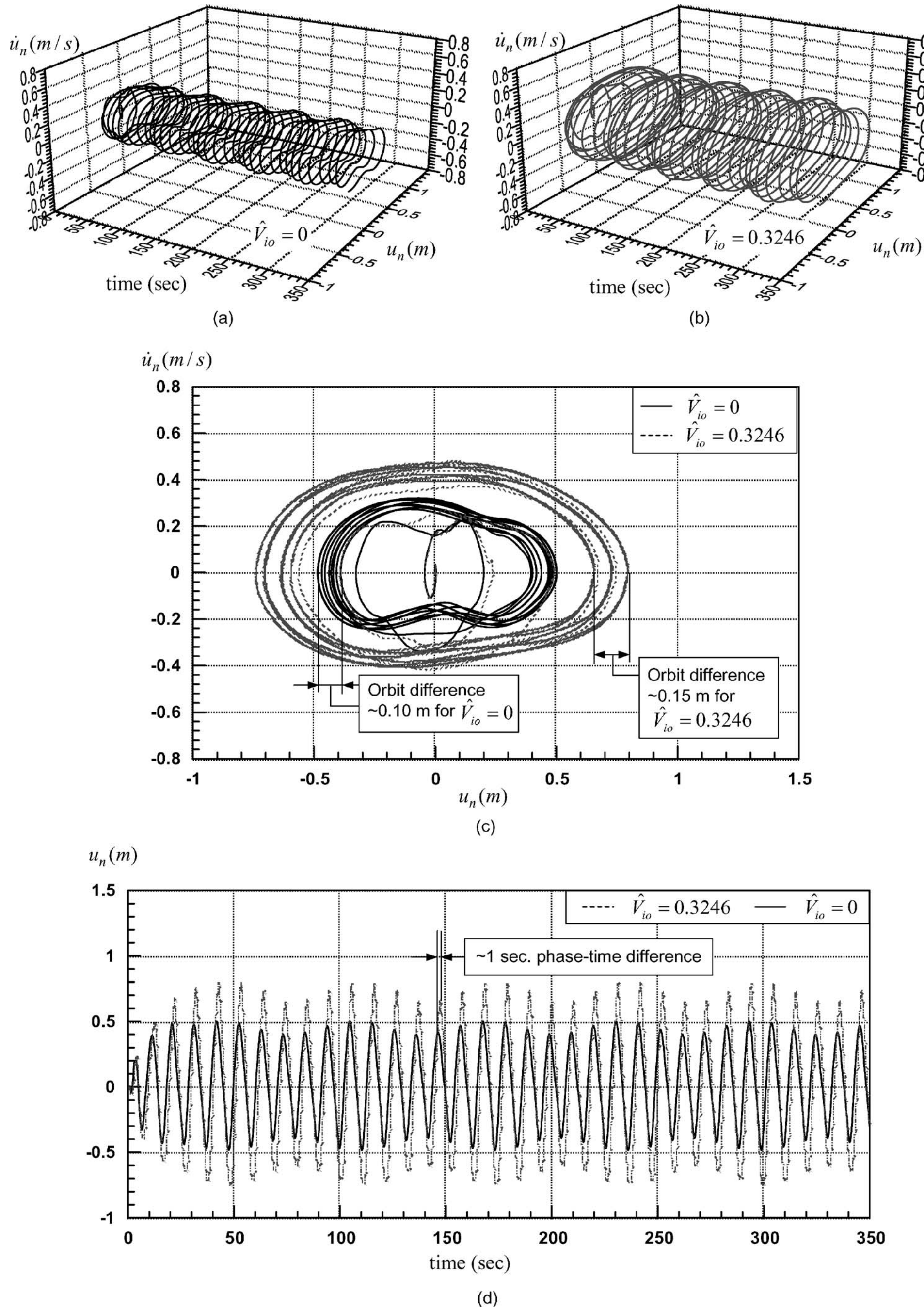


Fig. 3 The effect of the mean flow velocity of transported fluid $\hat{V}_{io}$ on (*a*); (*b*) phase spaces; (*c*) phase planes; (*d*) time histories of the dynamic normal displacement u_n of the flexible riser at y_o=150 m for $\hat{a}_{io}$=0, $\hat{V}_{ia}$=0.0406, and $\hat{\omega}_i$=1.25

proof by Moe and Chucheepsakul [17] that the oscillation phase of the riser could be shifted by the Coriolis effect of internal flow $2\rho_i A_i \bar{V}_i$.

Since the steady internal flow provides loading to the system by the Coriolis and the centrifugal effects as manifested in Eqs. (11*c*), (11*g*), and (11*h*), the steady internal flow effect enlarges the transverse and longitudinal dynamic displacements as shown in Figs. 4(*a*) and 4(*b*). The equivalent of the centrifugal force of the steady internal flow to the compressive internal force $-m_{io}V_{io}^2$ in Eq. (11*g*) reveals that the steady internal flow reduces the dynamic tension of the riser, and contribute the secondary effect on increasing bending moment as shown in Figs. 4(*c*) and 4(*d*), respectively.

3.3 The Effects of the Step Acceleration of Transported Fluid. Practically, the acceleration of internal flow should be inputted into the transporting system to improve a level of transportation rate in some operation periods whether while starting or during pumping process. Consider the internal flow acceleration Eq. (2) corresponding to the assumption (j). The internal flow rate could be modified by the adjustments of the following:

1. The constant rates of flow acceleration $\hat{a}_{io}$;
2. The fluctuation amplitude $\hat{V}_{ia}$; and
3. The pulsatile frequency $\hat{\omega}_i$ of the flow pulsation.

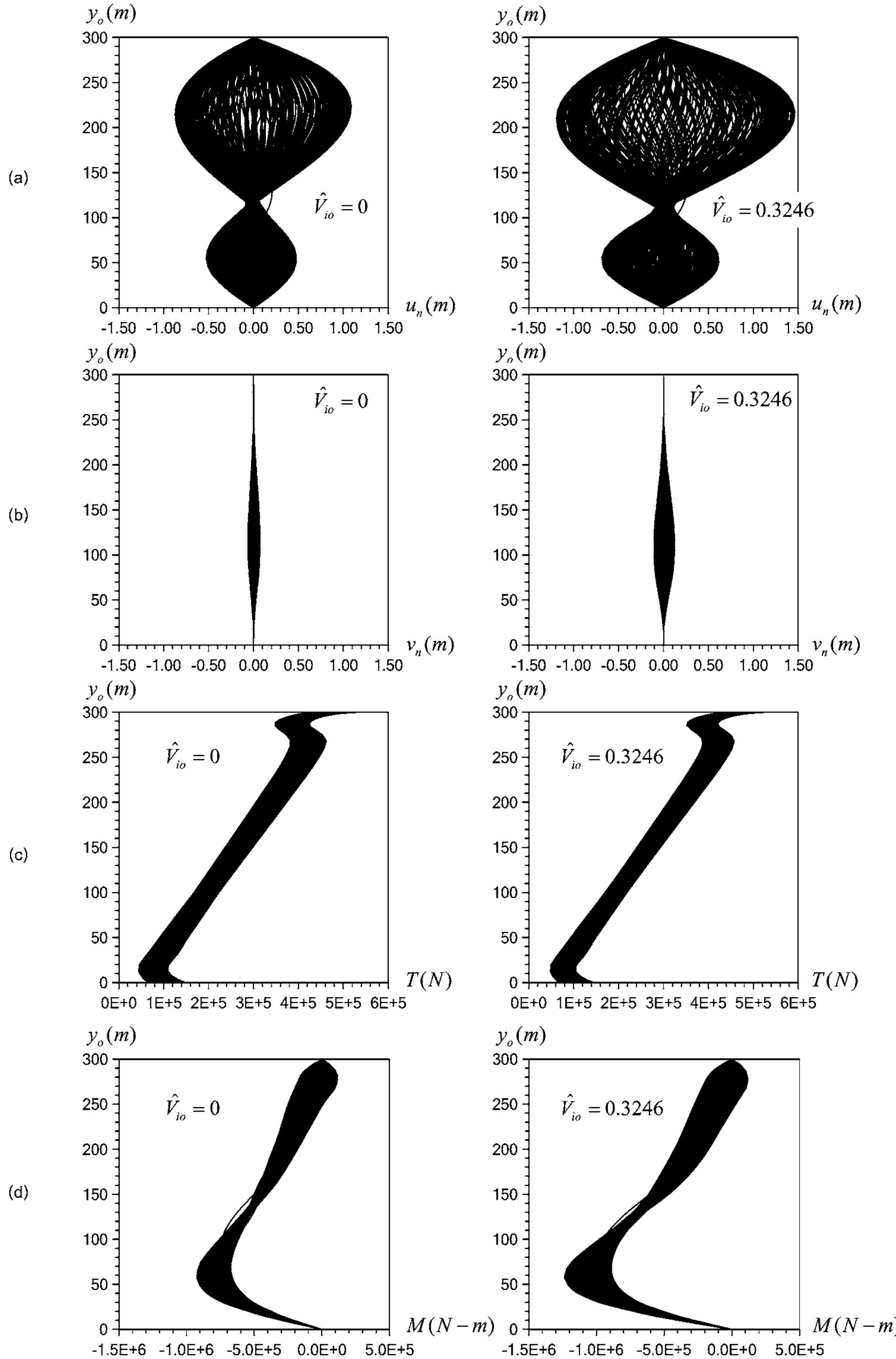


Fig. 4 The effect of the mean flow velocity of transported fluid $\hat{V}_{io}$ on envelopes of: (*a*) the dynamic normal displacement u_n; (*b*) the dynamic tangential displacement v_n; (*c*) the axial force *T*; and (*d*) the bending moment *M* for $\hat{a}_{io}=0$, $\hat{V}_{ia}=0.0406$, $\hat{\omega}_i=1.25$

In this section, the improvement of the internal flow rate by adapting (a) the constant acceleration rates is examined in the condition where the fluctuation amplitude and frequency of the pulsatile flow are nontrivial and invariable. The other ways of the flow-rate improvement such as (b) and (c) will be studied subsequently. Although a constant input step of the acceleration is to represent a linear change of the internal flow velocity, their multisteps in the appropriate multiple rise-time intervals may approximately converge to the linear or nonlinear acceleration effects.

To perceive the results obviously, Figs. 5(*c*) and 5(*d*) demonstrate the numerical results under the three steps of the internal flow acceleration. First the riser conveys fluid with the zero step acceleration: $\hat{a}_{io}=0$, in association with the constant pulsatile per-

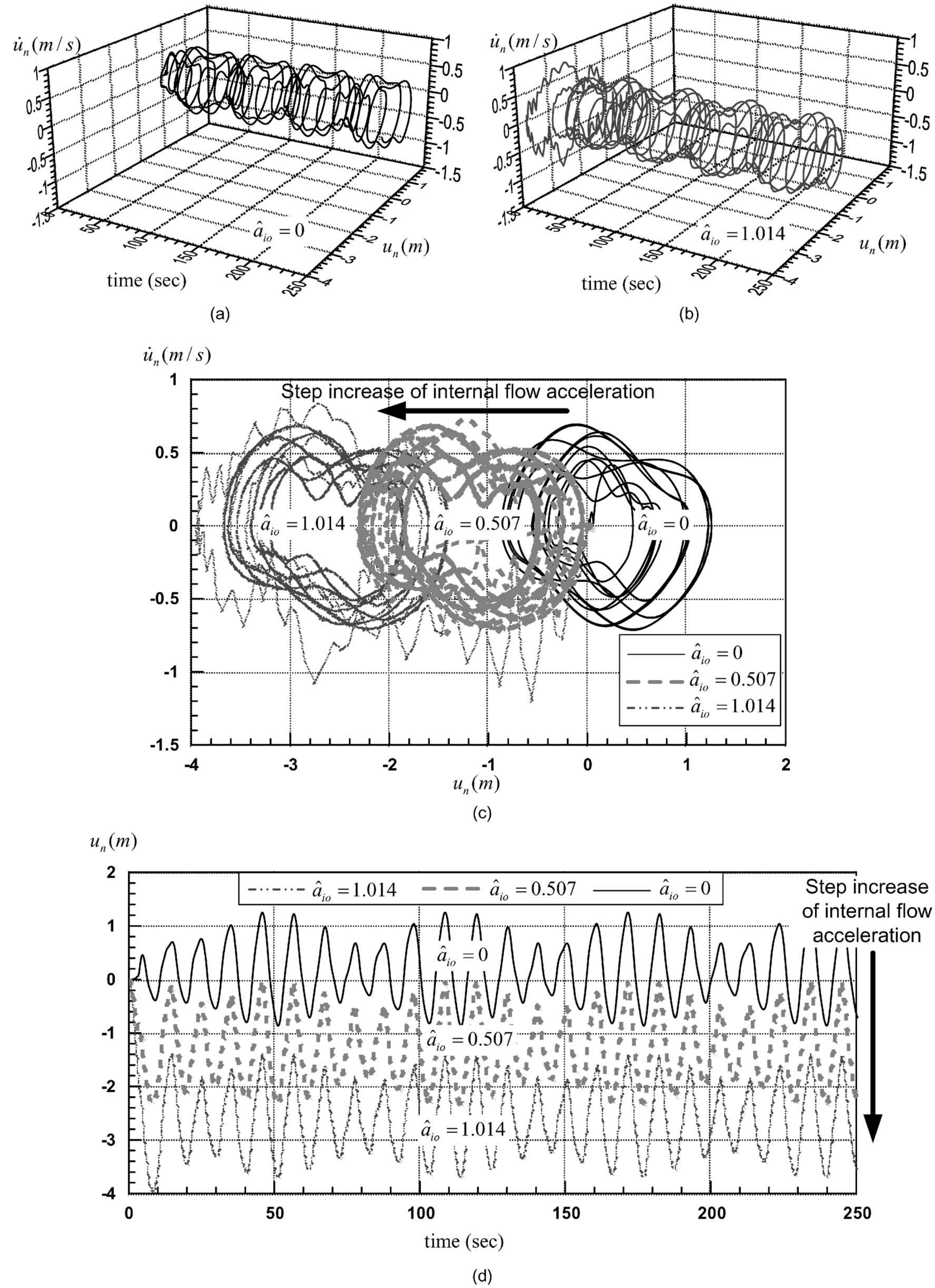


Fig. 5 The effect of the constant step acceleration of transported fluid $\hat{a}_{io}$ on (*a*); (*b*) phase spaces; (*c*) phase planes; and (*d*) time histories of the dynamic normal displacement u_n of the flexible riser at y_o=150 m for $\hat{V}_{io}$=0, $\hat{V}_{ia}$=0.0406, $\hat{\omega}_i$=1.25

turbations due to $\hat{V}_{ia}=0.0406$ and $\hat{\omega}_i=1.25$. It is discovered that the riser oscillation develops to travel obitally around the equilibrium point ($u_n=0.2$ m, $\dot{u}_n=0$ m/s) as seen on the phase plane in Fig. 5(*c*).

Anytime afterward, if the transportation rate is improved by the linear change of the internal flow velocity such as by increasing from $\hat{a}_{io}=0$ to $\hat{a}_{io}=0.507$ or to $\hat{a}_{io}=1.014$, the impulsive effect will overshoot the riser oscillation traveling to the new orbits of the equilibrium points (u_n=−1.2 m, $\dot{u}_n$=0 m/s) or (u_n=−2.6 m, $\dot{u}_n$=0 m/s), respectively.

Whenever, the upgrade of transportation rate is satisfactory enough, the current internal flow velocity will be maintained steadily, and thus the acceleration $\hat{a}_{io}$ will suddenly return to zero again. As a result, the riser motion will travel back to the original orbit around the equilibrium point (u_n=0.2 m, $\dot{u}_n$=0 m/s).

The multistep alterations of the transportation rate could be explained in the same way. If the internal flow speed is twice-stepped accelerated: from $\hat{a}_{io}$=0 to $\hat{a}_{io}$=0.507 and then from $\hat{a}_{io}$ =0.507 to $\hat{a}_{io}$=1.014, the trajectory and the time history of the riser vibration under $\hat{a}_{io}$=0 will be first overshot to the others under $\hat{a}_{io}$=0.507, and then second overshot to the others under $\hat{a}_{io}$=1.014 at the impulsive times as shown in Figs. 5(*c*) and 5(*d*). The deceleration of the internal flow speed should yield the contrary impulses to skip the trajectory and the time history back to the current condition of $\hat{a}_{io}$.

In conclusion, the constant step of the internal flow acceleration undertakes a small effect on the oscillation amplitudes, but drastically disturbs the static equilibrium locations of oscillations. As shown on the phase planes and phase spaces in Figs. 5(*a*)–5(*c*), although the initial conditions are identical, the pipes with and without the step acceleration $\hat{a}_{io}$ have the different positions of the static equilibrium. The internal flow acceleration $\hat{a}_{io}$ removes the beating orbits far away from the initial condition, but still maintains the stability of motion by some increase of the limit cycle size.

Consider the impulsive effect from Eqs. (2) and (11*i*). The negative tangential inertial forces due to the constant step of the internal flow acceleration $\hat{a}_{io}$

$$\{\mathbf{f}^{(e)}\} = \int_{\alpha} [\mathbf{N}]^T \begin{Bmatrix} -m_{io}x'_o a_{io} \\ -m_{io}y'_o a_{io} \end{Bmatrix} d\alpha \tag{18}$$

which could be considered as an ideal step input, is applied instantaneously at the initial condition.

As a result, the static displacement position is shifted from the displacement levels u_n=0.2−−1.2 m and to u_n=−2.6 m, and there is the slight overshooting of oscillation about the static equilibrium position, as shown in Fig. 5(*d*). The shifting of the static equilibrium position and the small overshoot of the oscillation increases the large displacements in the transverse and the longitudinal directions, the maximum dynamic tension, and the bending moment of the riser as shown in Figs. 6(*a*)–6(*d*), respectively.

3.4 The Effects of the Fluctuation Amplitude of Transported Fluid. Consider Eqs. (2) and (11*i*) to recognize the mechanics of the harmonically perturbed flow. The harmonic pulsatile flow components generate the positive tangential excitation forces

$$\{\mathbf{f}^{(e)}\} = \int_{\alpha} [\mathbf{N}]^T \begin{Bmatrix} m_{io}x'_o(V_{ia}\omega_i \sin \omega_i t) \\ m_{io}y'_o(V_{ia}\omega_i \sin \omega_i t) \end{Bmatrix} d\alpha \tag{19}$$

It is clearly seen that the fluctuation amplitude V_{ia} joins a part of the harmonic excitation amplitude of internal flow. Therefore, it does the external virtual work to yield an increase of the oscillation amplitudes in both horizontal and vertical directions as shown in Figs. 7(*a*)–7(*d*), 8(*a*), and 8(*b*).

The larger fluctuation amplitude would expand the sizes of limit cycle and beating amplitude, but lower the frequency of the riser oscillation. The extension of orbital motion degrades the closeness of the neighborhood limit cycles; as a result the stability of pipe motion is decayed based upon the orbital stability in the sense of Poincaré [28].

The fluctuation amplitude amplifies significantly the transverse motion and the dynamic tension in the upper region of the pipe, as shown in Figs. 8(*a*) and 8(*c*). It also expands the longitudinal displacements and the bending moment in the largely sagged portion of the pipe, where the greater curvature of pipe bending is developed, as shown in Figs. 8(*b*) and 8(*d*).

3.5 The Effects of the Pulsation Frequency of Transported Fluid. Consider Eq. (19). It is seen that the pulsatile frequency of the internal flow ω_i affects nonlinearly both the amplitude and the frequency of the excitation forces. Since, the variation of the pulsatile frequency may induce a type of instability such as parametric resonances of the riser oscillation, it is unnecessary for the pulsatile frequency to always expand the oscillation amplitude of the riser akin to the effect of the fluctuation amplitude of the internal flow.

As demonstrated in Figs. 9 and 10, the increase of pulsation frequency from 1.25 to 12.50 could provide the complex limit cycle associated with the subharmonic responses of the pipe's velocity, as shown in Figs. 9(*a*)–9(*c*). The alteration of the pulsation frequency eliminates the occurrence of beat phenomenon to become the steady-state response, as shown in Fig. 9(*d*). Consequently, the orbit of trajectory is reduced, and the stability of transverse motion of the pipe is revitalized, as shown in Fig. 9(*c*).

After the close to resonance vanishes, the maximum transversal oscillation reduces to occupy a magnitude at the steady state. Nevertheless, this usually does not mean that the pulsation frequency would reduce the pipe's responses. In fact if the resonance or beating is not involved, the pulsation frequency will take some effect on increasing the response frequency and the transverse vibration of the pipe according to Eq. (19). The additional case is studied by setting the pulsation frequency as 6.25, and it is found that there is no substantial difference between the time histories for $\hat{\omega}_i$=6.25 and for $\hat{\omega}_i$=12.50. More noticeably, the pulsation frequency would increase the longitudinal response, the dynamic tension, and the bending moment, as shown in Figs. 10(*a*)–10(*d*). Therefore, the stiffer axial and bending strength would be required in the pipe design to resist the excitation of the higher pulsation frequency.

4 Concluding Remarks

The mean flow velocity of transported fluid increases various responses, and reduces the stability of motion of the pipe. The constant step acceleration of internal flow does shift the static equilibrium position of the pipe to gain large displacement behavior and the extension of internal forces. The harmonic excitations of the pulsatile internal flow produced by the fluctuation amplitude have a similar influence on the effect of steady flow velocity, but the wave characteristic of the flow velocity has an extraordinary effect on the dynamic tension. The pulsatile frequency also affects the oscillation behavior of the pipe. The phenomena of beating, resonance, and steady-state vibrations may be disturbed, depending upon the variation of the pulsation frequency, which induces modification of the subharmonic frequency of the pipe velocity.

Acknowledgment

The authors gratefully acknowledge the support of the Thailand Research Fund (TRF) under Contract No. MRG4780133. The first author would like to express sincere gratitude to the Department of Civil and Environmental Engineering, the Engineering College of the University of Texas at Arlington (UTA) for providing a visiting scholarship over the fall semester 2000.

Appendix: Determination of the True Wall Tension

The true wall tension T is the actual stress used in the design of the riser pipe section. Hence, It is useful to manipulate Eq. (6) to determine its expression such that

$$T = EA_{P0}\varepsilon - 2\nu(p_e A_{e0} - p_i A_{i0}) \tag{A1}$$

where the axial strain ε is solved from Eq. (4).

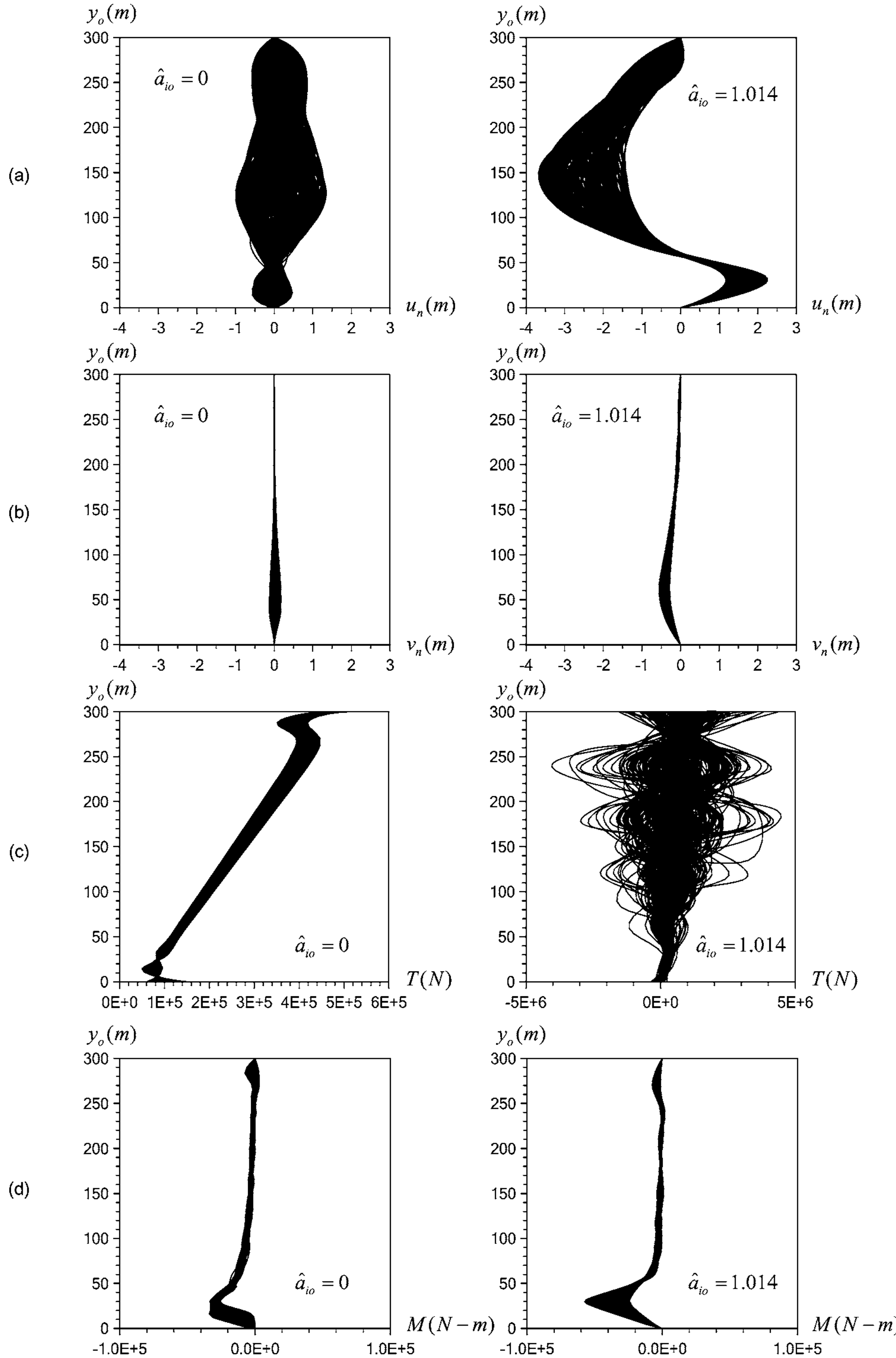


Fig. 6 The effect of the constant step acceleration of transported fluid $\hat{a}_{io}$ on envelopes of: (*a*) the dynamic normal displacement u_n; (*b*) the dynamic tangential displacement v_n; (*c*) the axial force T; and (*d*) the bending moment M for $\hat{V}_{io}$=0, $\hat{V}_{ia}$=0.0406, and $\hat{\omega}_i$=1.25

The pressure forces of internal and external fluids p_iA_{io} and p_eA_{eo} may be rendered from the tangential equilibrium conditions of the transported fluid and pipe as follows [25]

$$(p_iA_{io})' = \tau_w s' - (m_{io}g\cos\theta + m_{io}a_{Ft})s' \tag{A2}$$

$$(T + T_{tri} + p_eA_e)' = \tau_w s' - Q\theta' - [f_{Ht} - (m_P - m_e)g\cos\theta - m_P a_{Pt}]s' \tag{A3}$$

where τ_w is the wall-shear friction; a_{Ft} the tangential acceleration of the transported fluid [25]; $T_{\text{tri}}=EA_P\varepsilon_{\text{tri}}=(2\nu-1)(p_eA_{eo}-p_iA_{io})$ is the tension induced by triaxial pressures [25]; Q is the shear

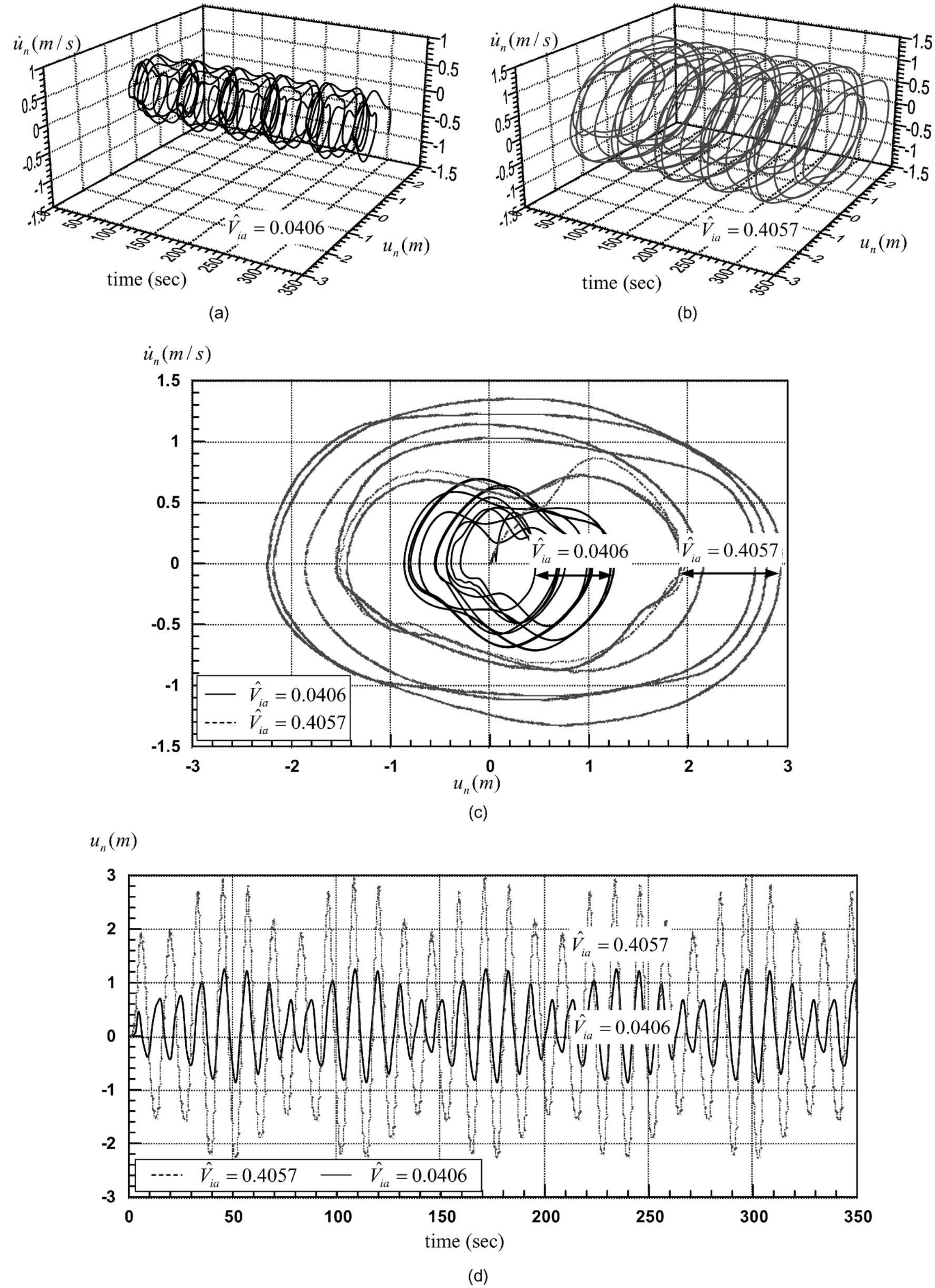


Fig. 7 The effect of the fluctuation amplitude of transported fluid $\hat{V}_{ia}$ on: (*a*); (*b*) phase spaces; (*c*) phase planes; (*d*) time histories of the dynamic normal displacement u_n of the flexible riser at y_o=150 m for $\hat{V}_{io}$=0, $\hat{a}_{io}$=0, and $\hat{\omega}_i$=1.25

force; and a_{Pt} is the tangential acceleration of the pipe. The three variables: T, p_e, and p_i could be determined by solving the system of three Eqs. (A1)–(A3).

Denoting

$$\text{total internal fluid pressure } p_i = \text{static pressure } p_{is} + \text{dynamic pressure } p_{id} \tag{A4}$$

and

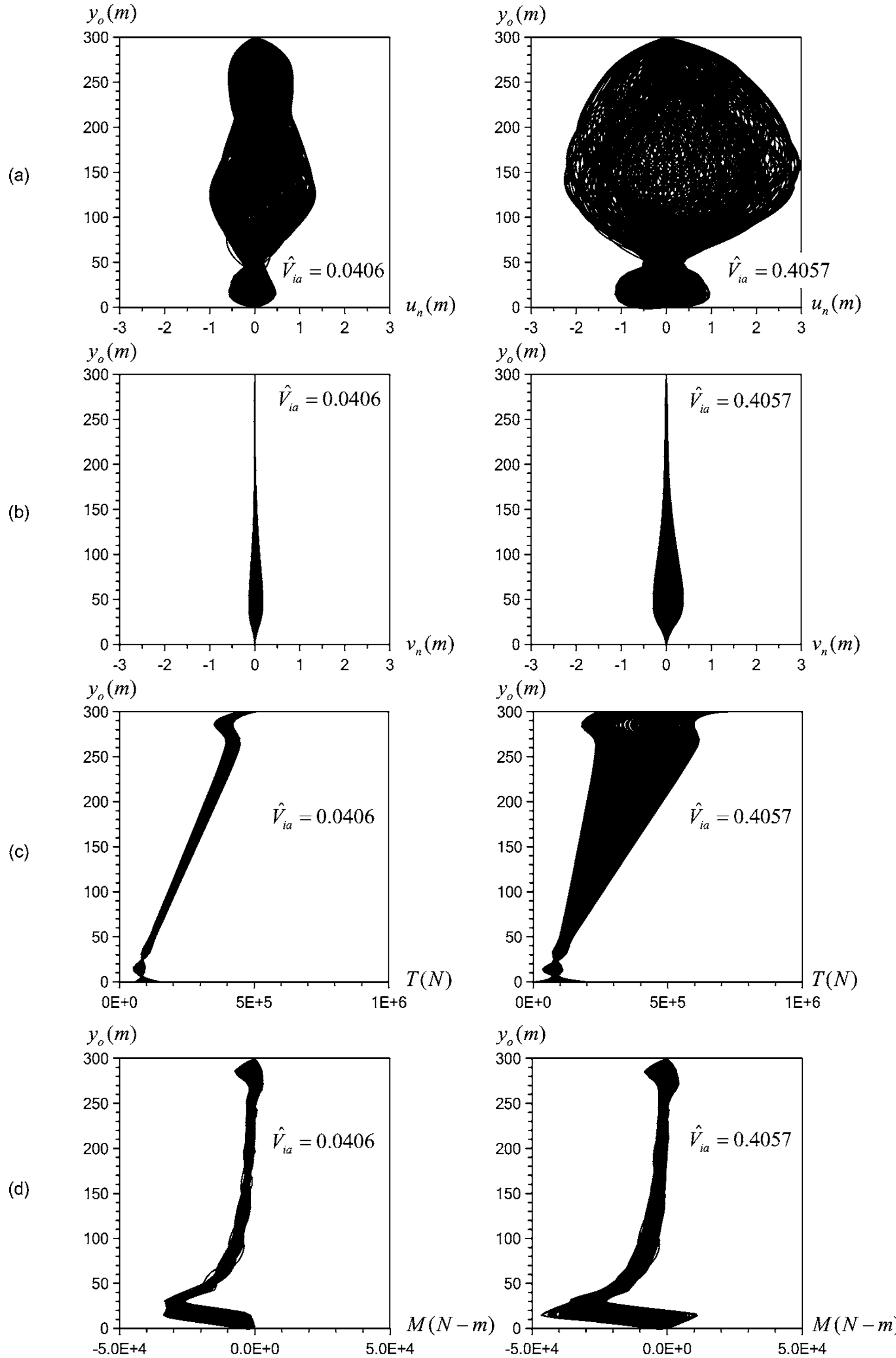


Fig. 8 The effect of the fluctuation amplitude of transported fluid $\hat{V}_{ia}$ on envelopes of: (*a*) the dynamic normal displacement u_n; (*b*) the dynamic tangential displacement v_n; (*c*) the axial force *T*; and (*d*) the bending moment *M* for $\hat{V}_{io}=0$, $\hat{a}_{io}=0$, and $\hat{\omega}_i=1.25$

$$\text{total external fluid pressure } p_e = \text{static pressure } p_{es} + \text{dynamic pressure } p_{ed} \tag{A5}$$

decompositions of Eqs. (A2) and (A3) between the static and dynamic states of fluids yield

$$(p_{is}A_{io})' = -(m_{io}g\cos\theta)s' \tag{A6a}$$

$$(p_{id}A_{io})' = \tau_w s' - (m_{io}a_{Ft})s' \tag{A6b}$$

$$(p_{es}A_{eo})' = -(m_{eo}\,g\cos\theta)s' \tag{A7a}$$

$$(p_{ed}A_{ed})' = \tau_w s' - Q\theta' - (f_{Ht} - m_P g\cos\theta - m_P a_{Pt})s' - (T + T_{\text{tri}})' \tag{A7b}$$

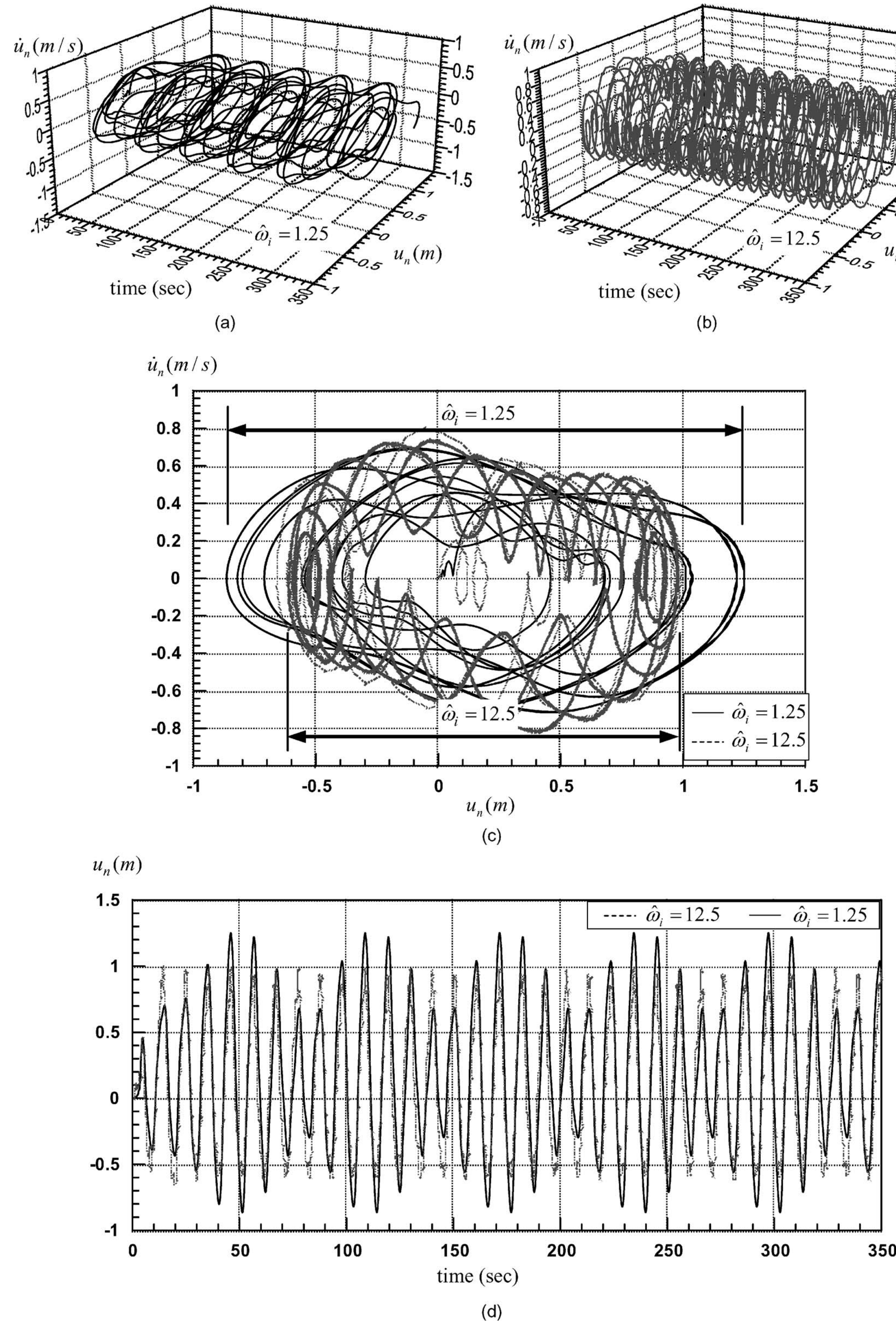


Fig. 9 The effect of the pulsation frequency of transported fluid $\hat{\omega}_i$ on (*a*); (*b*) phase spaces; (*c*) phase planes; (*d*) time histories of the dynamic normal displacement u_n of the flexible riser at y_o=150 m for $\hat{V}_{io}$=0, $\hat{a}_{io}$=0, $\hat{V}_{ia}$=0.0406

As well known in the subjects of fluid mechanics, Eqs. (A6*a*) and (A7*a*) give the classical solutions of the static fluid pressures $p_{is}(\alpha)$ and $p_{es}(\alpha)$ such that

$$p_{is} = m_{io} g(\bar{y}_t - y) \tag{A8a}$$

$$p_{es} = m_{eo} g(\bar{y}_t - y) \tag{A8b}$$

The excitation functions of the dynamic pressures $p_{id}(\alpha,t)$ and $p_{ed}(\alpha,t)$ to induce the other dynamic tension effect may be solved based upon the system of Eqs. (A1), (A6*b*), and (A7*b*). However,

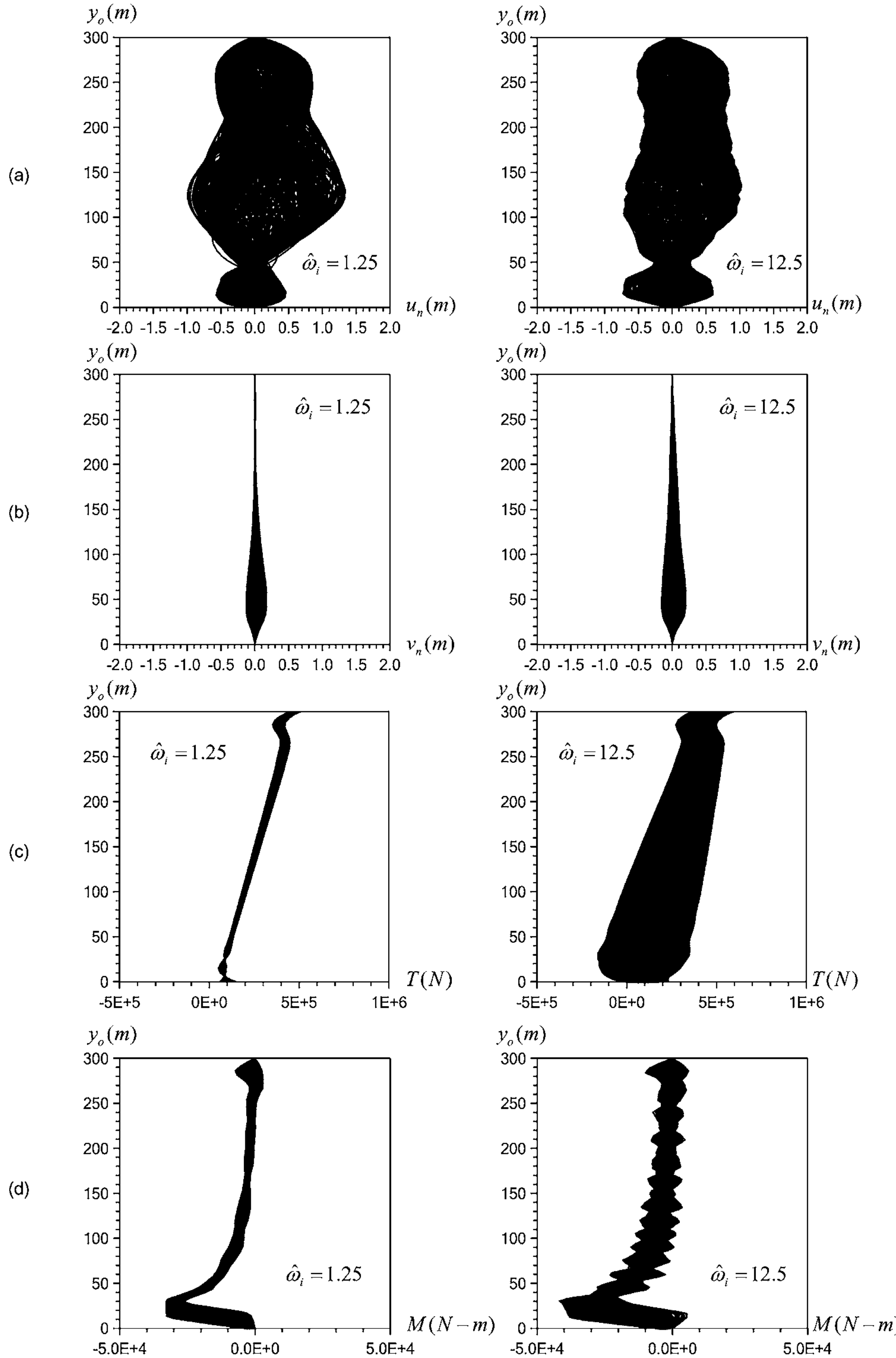


Fig. 10 The effect of the pulsation frequency of transported fluid $\hat{\omega}_i$ on envelopes of: (*a*) the dynamic normal displacement u_n; (*b*) the dynamic tangential displacement v_n; (*c*) the axial force T; and (*d*) the bending moment M for $\hat{V}_{io}$=0, $\hat{a}_{io}$=0, $\hat{V}_{ia}$=0.0406

it is quite cumbersome to recognize the exact value of the wall-shear friction τ_w, which varies upon various kinds of factors such as Reynolds' number, fluid density, fluid velocity, flow type, etc. The experimental information may be necessary to evaluate the mean value of the wall-shear friction in numerous situations. Thus, the effect of dynamic pressure fields of fluid flows on the dynamic true wall tension is neglected herein and still left as a future research question.

References

[1] Chakrabarti, S. K., and Frampton, R. E., 1982, "Review of Riser Analysis Techniques," Appl. Ocean. Res., **4**, pp. 73–90.

[2] Ertas, A., and Kozik, T. J., 1987, "A Review of Approaches to Riser Modeling," J. Energy Resour. Technol., **109**, pp. 155–160.

[3] Jain, A. K., 1994, "Review of Flexible Risers and Articulated Storage Systems," Ocean Eng., **21**, pp. 733–750.

[4] Patel, M. H., and Seyed, F. B., 1995, "Review of Flexible Riser Modelling and Analysis Techniques," Eng. Struct., **17**, pp. 293–304.

[5] Skop, R. A., and Clark, A. V., 1972, "The Bending of a Neutrally Buoyant, Tensioned Pipe in a Steady Current," Ocean Eng., **2**, pp. 141–157.
[6] Fang, J., and Lyons, G. J., 1996, "Structural Damping of Tensioned Pipes With Reference to Cables," J. Sound Vib., **193**, pp. 891–907.
[7] Lee, S. Y., and Mote, Jr., C. D., 1997, "A Generalized Treatment of the Energetics of Translating Continua, Part I: Strings and Second Order Tension Pipes," J. Sound Vib., **204**, pp. 717–734.
[8] Öz, H. R., and Boyaci, H., 2000, "Transverse Vibrations of Tensioned Pipes Conveying Fluid With Time-Dependent Velocity," J. Sound Vib., **236**, pp. 259–276.
[9] Öz, H. R., 2001, "Non-linear Vibrations and Stability Analysis of Tensioned Pipes Conveying Fluid With Variable Velocity," Int. J. Non-Linear Mech., **36**, pp. 1031–1039.
[10] Öz, H. R., and Evrensel, C. A., 2002, "Natural Frequencies of Tensioned Pipes Conveying Fluid and Carrying a Concentrated Mass," J. Sound Vib., **250**, pp. 368–377.
[11] Antman, S. S., 1995, *Nonlinear Problems of Elasticity*, Springer, New York.
[12] Atanackovic, T. M., 1997, *Stability Theory of Elastic Rods*, World Scientific, Singapore.
[13] Chen, S. S., 1971, "Dynamic Stability of Tube Conveying Fluid," J. Engrg. Mech. Div., **97**, pp. 1469–1485.
[14] Ginsberg, J. H., 1973, "The Dynamic Stability of a Pipe Conveying a Pulsatile Flow," Int. J. Eng. Sci., **11**, pp. 1013–1024.
[15] Gorman, D. G., Reese, J. M., and Zhang, Y. L., 2000, "Vibration of a Flexible Pipe Conveying Viscous Pulsating Fluid Flow," J. Sound Vib., **230**, pp. 379–392.
[16] Irani, M. B., Modi, V. J., and Weit, F., 1987, "Riser Dynamics With Internal Flow and Nutation Damping," *Proceedings 6th International Offshore Mechanics and Arctic Engineering Symposium*, 3, Houston, TX, March 1–5, pp. 119–125.
[17] Moe, G., and Chucheepsakul, S., 1988, "The Effect of Internal Flow on Marine Risers," *Proceedings 7th International Offshore Mechanics and Arctic Engineering Symposium*, 1, Houston, TX, February 7–12, pp. 375–382.
[18] Patel, H. M., and Seyed, F. B., 1989, "Internal Flow-Induced Behavior of Flexible Risers," Eng. Struct., **11**, pp. 266–280.
[19] Wu, M. C., and Lou, J. Y. K., 1991, "Effects of Rigidity and Internal Flow on Marine Riser Dynamics," Appl. Ocean. Res., **13**, pp. 235–244.
[20] Seyed, F. B., and Patel, H. M., 1992, "Mathematics of Flexible Risers Including Pressure and Internal Flow Effects," Mar. Struct., **5**, pp. 121–150.
[21] Huang, T., 1993, "Kinematics of Transported Mass Inside Risers and Pipes," *Proceedings 3rd International Offshore and Polar Engineering ISOPE*, 2, Singapore, June 6–11, pp. 331–336.
[22] Chucheepsakul, S., and Huang, T., 1994, "Influence of Transported Mass on the Equilibrium Configuration of Risers," *Proceedings 4th International Offshore and Polar Engineering ISOPE*, 2, Osaka, Japan, April 10–15, pp. 246–249.
[23] Moe, G., Stromsem, K. C., and Fylling, I., 1994, "Behavior of Risers With Internal Flow Under Various Boundary Conditions," *Proceedings 4th International Offshore and Polar Engineering ISOPE*, 2, Osaka, Japan, April 10–15, pp. 258–262.
[24] Bar-Avi, P., 2000, "Dynamic Response of Risers Conveying Fluid," J. Offshore Mech. Arct. Eng., **122**, pp. 188–193.
[25] Chucheepsakul, S., Monprapussorn, T., and Huang, T., 2003, "Large Strain Formulations of Extensible Flexible Marine Pipes Transporting Fluid," J. Fluids Struct., **17**, pp. 185–224.
[26] Monprapussorn, T., Chucheepsakul, S., and Huang, T., 2004, "The Coupled Radial-Axial Deformations Analysis of Flexible Pipes Conveying Fluid," Int. J. Numer. Methods Eng., **59**, pp. 1399–1452.
[27] Gear, C. W., 1971, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Clifs, NJ.
[28] Meirovitch, L., 1997, *Principles and Techniques of Vibrations*, Prentice-Hall, Englewood Cliffs, NJ.

The Eshelby Tensors in a Finite Spherical Domain—Part I: Theoretical Formulations

Shaofan Li[1]
e-mail: li@ce.berkeley.edu

Roger A. Sauer

Gang Wang

Department of Civil and Environmental Engineering,
University of California,
Berkeley, CA 94720

This work is concerned with the precise characterization of the elastic fields due to a spherical inclusion embedded within a spherical representative volume element (RVE). The RVE is considered having finite size, with either a prescribed uniform displacement or a prescribed uniform traction boundary condition. Based on symmetry and group theoretic arguments, we identify that the Eshelby tensor for a spherical inclusion admits a unique decomposition, which we coin the "radial transversely isotropic tensor." Based on this notion, a novel solution procedure is presented to solve the resulting Fredholm type integral equations. By using this technique, exact and closed form solutions have been obtained for the elastic disturbance fields. In the solution two new tensors appear, which are termed the Dirichlet–Eshelby tensor and the Neumann–Eshelby tensor. In contrast to the classical Eshelby tensor they both are position dependent and contain information about the boundary condition of the RVE as well as the volume fraction of the inclusion. The new finite Eshelby tensors have far-reaching consequences in applications such as nanotechnology, homogenization theory of composite materials, and defects mechanics. [DOI: 10.1115/1.2711227]

1 Introduction

One of the corner stones of contemporary micromechanics and nanomechanics is Eshelby's inclusion theory [1–3]. Eshelby's ellipsoidal inclusion solution was obtained based on the assumption that an inclusion is embedded in unbounded ambient space. This is a good approximation if the size effect of the inclusion is negligible, i.e., the size of the inclusion is small compared to the size of the representative volume element. In engineering applications, the size of the representative volume element (RVE) is finite. Therefore, certain approximations have to be made in order to utilize Eshelby's classical solution in homogenization. This limitation becomes obvious, when size effects and interfacial boundary effects of a second phase in a composite, or the size effect and boundary effects of an inhomogeneity, become prominent issues, which is one of main focuses of the nanocomposite mechanics and materials, e.g., Refs. [4,5]. Today, there is a call for the solution of the inclusion problem in a finite domain.

Inclusion problems in a finite domain have been considered before, e.g., Refs. [6–9]. A common approach adopted is to first find the Green's function of Navier's equation for a finite domain, and then to find the solution of the corresponding inclusion problem. However, attempts based on this approach have been futile, we believe, because of the mathematical difficulties involved in obtaining a closed form solution of the finite Green's function. This is true even for a highly symmetrical spherical domain. In fact, the Green's function of Navier's equation for a finite spherical domain has not been found yet. To the best of the authors' knowledge, there has never been any exact, closed form solution of the inclusion problem in a finite domain published in the literature. A solution has been obtained by Luo and Weng [10], which coincides with our solution in a special case. Their solution, however, is not in closed form and lacks expressions for the Eshelby tensors.

In this paper, which is the first part of a series, we present the exact solution of the finite Eshelby tensors of a spherical inclusion embedded concentrically within a finite spherical RVE. The following section illustrates the two boundary value problems (BVPs) we are considering and their resulting integral equations. In Sec. 3 the notion of a transversely isotropic tensor is discussed, which is used in Sec. 4 to solve the two integral equations. Section 5 concludes this part. The Eshelby tensors derived in this paper have some profound consequences for both homogenization and the study of inhomogeneities in finite elastic solids. In the second part of this work, applications to homogenization of composites are discussed [11].

2 The Inclusion Problem

We consider Eshelby's homogeneous inclusion problem in a finite domain. Figure 1 shows a spherical inclusion Ω_I with radius a embedded at the center of a spherical representative volume element Ω with radius A. Consider two arbitrary points $\mathbf{x}\in\Omega$, $\mathbf{y}\in\Omega$, and let $\mathbf{r}=\mathbf{y}-\mathbf{x}$. Each vector $\mathbf{x}$, $\mathbf{y}$, $\mathbf{r}$ can be expressed as its length multiplied by a unit direction vector. We shall denote them as $\mathbf{x}=|\mathbf{x}|\bar{\mathbf{x}}$, $\mathbf{y}=|\mathbf{y}|\bar{\mathbf{y}}$ and $\mathbf{r}=r\bar{\mathbf{r}}$, with $r=|\mathbf{r}|$. Note that if $\mathbf{y}\in\partial\Omega$ we have $|\mathbf{y}|=A$ and $\bar{\mathbf{y}}=\mathbf{n}$, i.e., the direction of $\mathbf{y}$ is equal to the outward surface normal $\mathbf{n}$. Furthermore we define the ratios $\rho=a/|\mathbf{x}|$, $\rho_0=a/A$ and $t=|\mathbf{x}|/A=\rho_0/\rho$ to allow for a nondimensional description. Suppose that a constant eigenstrain field is prescribed inside the inclusion

$$\epsilon^*_{ij}(\mathbf{x})=\begin{cases}\epsilon^*_{ij}, & \mathbf{x}\in\Omega_I\\ 0, & \mathbf{x}\in\Omega_E=\Omega/\Omega_I\end{cases} \tag{1}$$

The infinitesimal elastic strain equals the total strain subtracting the eigenstrain

$$e_{ij}=\epsilon_{ij}-\epsilon^*_{ij}$$

with

$$\epsilon_{ij}=\tfrac{1}{2}(u_{i,j}+u_{j,i}) \tag{2}$$

where $u_{i,j}$ denotes the spatial differentiation $\partial u_i/\partial x_j$. We assume that the RVE is a linear elastic medium, i.e.,

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 6, 2006; final manuscript received June 13, 2006. Review conducted by Subrata Mukherjee. Paper presented at the 2005 ASME International Mechanical Engineering Congress (IMECE2005), Orlando, FL, USA, November 5–11, 2005.

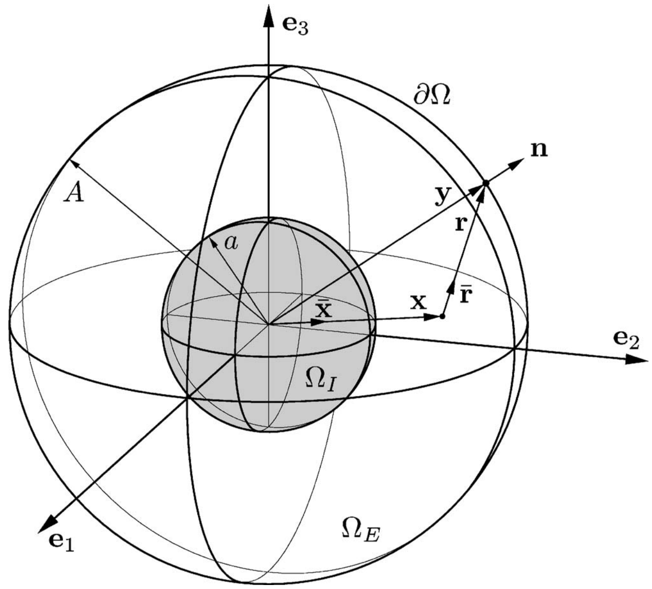


Fig. 1 A spherical representative element containing a spherical inclusion

$$\sigma_{ij} = \mathbb{C}_{ijk\ell} e_{k\ell}, \quad \forall\, \mathbf{x} \in \Omega \tag{3}$$

where σ_{ij} are the components of the Cauchy stress tensor, and $\mathbb{C}_{ijk\ell}$ is the elasticity tensor.

On the boundary of the RVE, two types of boundary conditions (BC), a prescribed displacement (Dirichlet) or a prescribed traction (Neumann) boundary condition, are considered

Dirichlet BC

$$u_i = \epsilon_{ij}^0 x_j, \quad \forall\, \mathbf{x} \in \partial\Omega \tag{4}$$

Neumann BC

$$t_i = \sigma_{ij}^0 n_j, \quad \forall\, \mathbf{x} \in \partial\Omega \tag{5}$$

where ε_{ij}^0 and σ_{ij}^0 are the background strain and stress fields. The elastic fields inside the RVE can be decomposed into the background field from the remote boundary loads and a disturbance field, which arises due to the presence of the inclusion. Thus the displacement and traction can be written as

$$u_i = u_i^0 + u_i^d, \quad t_i = t_i^0 + t_i^d \tag{6}$$

so that we obtain the following two homogeneous boundary conditions

Dirichlet BC

$$u_i^d = 0, \quad \forall\, \mathbf{x} \in \partial\Omega \tag{7}$$

Neumann BC

$$t_i^d = 0, \quad \forall\, \mathbf{x} \in \partial\Omega \tag{8}$$

for the disturbance fields. The solution for the background field depends on the macro problem. Here we are concerned with the solution of the disturbance fields. Considering the equilibrium equations $\sigma_{ji,j}^d = 0$, we obtain either the Dirichlet–Eshelby BVP

$$\mathbb{C}_{ijk\ell} u_{k,\ell j}^d(\mathbf{x}) - \mathbb{C}_{ijk\ell} \epsilon_{k\ell,j}^*(\mathbf{x}) = 0, \quad \forall\, \mathbf{x} \in \Omega$$

$$u_i^d(\mathbf{x}) = 0, \quad \forall\, \mathbf{x} \in \partial\Omega \tag{9}$$

or the Neumann–Eshelby BVP

$$\mathbb{C}_{ijk\ell} u_{k,\ell j}^d(\mathbf{x}) - \mathbb{C}_{ijk\ell} \epsilon_{k\ell,j}^*(\mathbf{x}) = 0, \quad \forall\, \mathbf{x} \in \Omega$$

$$t_i^d(\mathbf{x}) = n_j \mathbb{C}_{ijk\ell} u_{k,\ell}^d(\mathbf{x}) = 0, \quad \forall\, \mathbf{x} \in \partial\Omega \tag{10}$$

Let us denote the Green's function, $G_{mi}^\infty(\mathbf{x}-\mathbf{y})$, as the solution of the following Navier's equation in unbounded space

$$\mathbb{C}_{ijk\ell} G_{mk,\ell j}^\infty(\mathbf{x}-\mathbf{y}) + \delta_{mi}\delta(\mathbf{x}-\mathbf{y}) = 0, \quad \forall\, \mathbf{x},\mathbf{y} \in \mathbb{R}^3, \quad i = 1,2,3 \tag{11}$$

For an isotropic linear elastic space, the Green's function is (e.g., Ref. [12])

$$G_{ij}^\infty(\mathbf{x}-\mathbf{y}) = \frac{1}{16\pi\mu(1-\nu)}\left[\frac{(x_i - y_i)(x_i - y_i)}{r^3} + (3-4\nu)\frac{\delta_{ij}}{r}\right] \tag{12}$$

where $r = \sqrt{(x_i - y_i)(x_i - y_i)}$; μ is the shear modulus; and ν is Poisson's ratio. By using Somigliana's identity [12], the displacement field solution of BVPs Eqs. (9) and (10) may be expressed as

$$\begin{aligned} u_m^d(\mathbf{x}) = &- \int_\Omega \mathbb{C}_{ijk\ell} G_{im,j}^\infty(\mathbf{x}-\mathbf{y}) \epsilon_{k\ell}^*(\mathbf{y}) d\Omega_y \\ &+ \int_{\partial\Omega} \mathbb{C}_{ijk\ell} u_{k,\ell}^d(\mathbf{y}) G_{im}^\infty(\mathbf{x}-\mathbf{y}) n_j(\mathbf{y}) dS_y \\ &+ \int_{\partial\Omega} \mathbb{C}_{ijk\ell} u_k^d(\mathbf{y}) G_{im,j}^\infty(\mathbf{x}-\mathbf{y}) n_\ell(\mathbf{y}) dS_y \end{aligned} \tag{13}$$

where we have denoted $G_{im,j}^\infty := \partial G_{im}^\infty / \partial x_j = -\partial G_{im}^\infty / \partial y_j$. For the Dirichlet–Eshelby problem, this integral equation becomes

$$\begin{aligned} u_m^d(\mathbf{x}) = &- \int_\Omega \mathbb{C}_{ijk\ell} G_{im,j}^\infty(\mathbf{x}-\mathbf{y}) \epsilon_{k\ell}^*(\mathbf{y}) d\Omega_y \\ &+ \int_{\partial\Omega} \mathbb{C}_{ijk\ell} u_{k,\ell}^d(\mathbf{y}) G_{im}^\infty(\mathbf{x}-\mathbf{y}) n_j(\mathbf{y}) dS_y \end{aligned} \tag{14}$$

and for the Neumann–Eshelby problem, Eq. (13) reduces to

$$\begin{aligned} u_m^d(\mathbf{x}) = &- \int_\Omega \mathbb{C}_{ijk\ell} G_{im,j}^\infty(\mathbf{x}-\mathbf{y}) \epsilon_{k\ell}^*(\mathbf{y}) d\Omega_y \\ &+ \int_{\partial\Omega} \mathbb{C}_{ijk\ell} u_k^d(\mathbf{y}) G_{im,j}^\infty(\mathbf{x}-\mathbf{y}) n_\ell(\mathbf{y}) dS_y \end{aligned} \tag{15}$$

In case of the Dirichlet–Eshelby BVP, the disturbance strain field follows from the displacement Eq. (14) as

$$\begin{aligned} \epsilon_{ij}^d(\mathbf{x}) = &-\frac{1}{2}\epsilon_{mn}^* \int_{\Omega_I} \mathbb{C}_{k\ell mn}[G_{ki,\ell j}^\infty(\mathbf{x}-\mathbf{y}) + G_{kj,\ell i}^\infty(\mathbf{x}-\mathbf{y})] d\Omega_y \\ &+ \frac{1}{2}\int_{\partial\Omega} \mathbb{C}_{k\ell pq} \epsilon_{pq}^d(\mathbf{y}) [G_{ki,j}^\infty(\mathbf{x}-\mathbf{y}) + G_{kj,i}^\infty(\mathbf{x}-\mathbf{y})] n_\ell(\mathbf{y}) dS_y \end{aligned} \tag{16}$$

For the Dirichlet–Eshelby BVP we solve Eq. (16) which is an integral equation for the unknown strain field ϵ_{ij}^d. In case of the Neumann–Eshelby BVP we can directly solve Eq. (15) which is an integral equation for the unknown displacement field u_i^d. In passing, we note that Eq. (16) becomes a hypersingular integral equation if $\mathbf{x} \in \partial\Omega$.

To illustrate our solution procedure, we re-examine the classical Eshelby tensors. For inclusion problems in unbounded space, the boundary term in Eqs. (14)–(16) drops out. One can then find the disturbance strain fields in terms of the Eshelby tensors [1,2],

$$\epsilon_{ij}^d(\mathbf{x}) = \mathbb{S}_{ijk\ell}^{\cdot,\infty}(\mathbf{x}) \epsilon_{k\ell}^*, \quad \forall\, \mathbf{x} \in \mathbb{R}^3 \tag{17}$$

where the superscript $\cdot$ represents the interior solution ($\cdot = I$) or the exterior solution ($\cdot = E$), depending on the location of $\mathbf{x}$, i.e.,

$$\mathbb{S}_{ijk\ell}^{\cdot,\infty}(\mathbf{x}) = \begin{cases} \mathbb{S}_{ijk\ell}^{I,\infty}(\mathbf{x}), & \forall\, \mathbf{x} \in \Omega_I \\ \mathbb{S}_{ijk\ell}^{E,\infty}(\mathbf{x}), & \forall\, \mathbf{x} \in \mathbb{R}^3/\Omega_I \end{cases} \tag{18}$$

For spherical inclusions in an infinite elastic medium, the Eshelby tensors have the elementary form (e.g., Refs. [13,14]

1. Interior solution

$$\mathbb{S}_{ijmn}^{I,\infty}(\mathbf{x}) = \frac{(5\nu-1)}{15(1-\nu)}\delta_{ij}\delta_{mn} + \frac{(4-5\nu)}{15(1-\nu)}(\delta_{im}\delta_{jn} + \delta_{in}\delta_{jm}), \quad \mathbf{x} \in \Omega_I \tag{19}$$

2. Exterior solution

$$\begin{aligned}\mathbb{S}_{ijmn}^{E,\infty}(\mathbf{x}) = \frac{\rho^3}{30(1-\nu)}&[(3\rho^2+10\nu-5)\delta_{ij}\delta_{mn} + (3\rho^2-10\nu+5)\\ &\times(\delta_{im}\delta_{jn}+\delta_{in}\delta_{jm}) + 15(1-\rho^2)\delta_{ij}\bar{x}_m\bar{x}_n + 15(1-2\nu\\ &-\rho^2)\delta_{mn}\bar{x}_i\bar{x}_j + 15(\nu-\rho^2)(\delta_{im}\bar{x}_j\bar{x}_n + \delta_{jm}\bar{x}_i\bar{x}_n + \delta_{in}\bar{x}_j\bar{x}_m\\ &+ \delta_{jn}\bar{x}_i\bar{x}_m) + 15(7\rho^2-5)\bar{x}_i\bar{x}_j\bar{x}_m\bar{x}_n], \quad \mathbf{x} \in \mathbb{R}^3/\Omega_I\end{aligned} \tag{20}$$

where $\rho := a/|\mathbf{x}|$ and $|\mathbf{x}| = \sqrt{x_i x_i}$, $i=1,2,3$.

Inspired by Eq. (17), we postulate the following form of the two considered BVPs

$$\epsilon_{ij}^d(\mathbf{x}) = \mathbb{S}_{ijk\ell}^{\cdot,\star}(\mathbf{x})\epsilon_{k\ell}^*, \quad \forall\, \mathbf{x} \in \Omega \tag{21}$$

where $\mathbb{S}_{ijk\ell}^{\cdot,\star}(\mathbf{x})$ is an unknown fourth-order tensor (the finite Eshelby tensor), we are seeking to obtain. As before the superscript $\cdot$ represents the interior solution ($\cdot = I$) or the exterior solution ($\cdot = E$). The superscript $\star$ stands for the Dirichlet–Eshelby tensor ($\star = D$) or the Neumann–Eshelby tensor ($\star = N$). As a special case we expect to obtain the original infinite Eshelby tensor ($\star = \infty$).

In principle, one may be able to use spherical harmonics to represent the general solution of Eqs. (14) and (15) based on symmetry, but the solution procedure is very much involved. In fact, no explicit solution has been found as shown in a particular case worked out by Luo and Weng [10].

3 The Radial Isotropic Tensor

To solve Eqs. (14) and (15), we first introduce a novel concept of the *radial transversely isotropic tensor*, or in short, the *radial isotropic tensor*.

It may be observed from the expressions of $\mathbb{S}_{ijmn}^{I,\infty}(\mathbf{x})$ and $\mathbb{S}_{ijmn}^{E,\infty}(\mathbf{x})$ above that there appear six independent tensorial bases, which can be arranged in an array as follows

$$\boldsymbol{\Theta}_{ijmn}(\bar{\mathbf{x}}) := \begin{bmatrix} \delta_{ij}\delta_{mn} \\ \delta_{im}\delta_{jn} + \delta_{in}\delta_{jm} \\ \delta_{ij}\bar{x}_m\bar{x}_n \\ \delta_{mn}\bar{x}_i\bar{x}_j \\ \delta_{im}\bar{x}_j\bar{x}_n + \delta_{in}\bar{x}_j\bar{x}_m + \delta_{jm}\bar{x}_i\bar{x}_n + \delta_{jn}\bar{x}_i\bar{x}_m \\ \bar{x}_i\bar{x}_j\bar{x}_m\bar{x}_n \end{bmatrix} \tag{22}$$

We term this array as the circumference basis of the Eshelby tensor. By using $\boldsymbol{\Theta}_{ijmn}(\bar{\mathbf{x}})$, both the original interior and exterior Eshelby tensor can be recast into a canonical form, the dot product of two arrays, i.e.,

$$\begin{aligned}\mathbb{S}_{ijmn}^{\cdot,\infty}(\mathbf{x}) &= S_1^{\cdot,\infty}(t)\delta_{ij}\delta_{mn} + S_2^{\cdot,\infty}(t)(\delta_{im}\delta_{jn}+\delta_{in}\delta_{jm}) + S_3^{\cdot,\infty}(t)\delta_{ij}\bar{x}_m\bar{x}_n\\ &\quad + S_4^{\cdot,\infty}(t)\delta_{mn}\bar{x}_i\bar{x}_j + S_5^{\cdot,\infty}(t)(\delta_{im}\bar{x}_j\bar{x}_n + \delta_{in}\bar{x}_j\bar{x}_m + \delta_{jm}\bar{x}_i\bar{x}_n\\ &\quad + \delta_{jn}\bar{x}_i\bar{x}_m) + S_6^{\cdot,\infty}(t)\bar{x}_i\bar{x}_j\bar{x}_m\bar{x}_n\\ &= \boldsymbol{\Theta}_{ijmn}^T(\bar{\mathbf{x}})\mathbf{S}^{\cdot,\infty}(t).\end{aligned} \tag{23}$$

The arrays, $\mathbf{S}^{I,\infty}(t)$ and $\mathbf{S}^{E,\infty}(t)$, are termed the radial basis of the infinite Eshelby tensor. In accordance to Eqs. (19) and (20) they are given as

$$\mathbf{S}^{I,\infty}(t) = \frac{1}{15(1-\nu)}\begin{bmatrix} 5\nu-1 \\ 4-5\nu \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\mathbf{S}^{E,\infty}(t) = \frac{\rho_0^3/t^3}{30(1-\nu)}\begin{bmatrix} 3\rho_0^2/t^2+10\nu-5 \\ 3\rho_0^2/t^2-10\nu+5 \\ 15(1-\rho_0^2/t^2) \\ 15(1-2\nu-\rho_0^2/t^2) \\ 15(\nu-\rho_0^2/t^2) \\ 15(7\rho_0^2/t^2-5) \end{bmatrix} \tag{24}$$

where $t = |\mathbf{x}|/A = \rho_0/\rho$, with $\rho = a/|\mathbf{x}|$, and $\rho_0 = a/A$.

The above heuristic discussion reveals an important fact, that the Eshelby tensor is a so-called "radial isotropic tensor," which is a generalization of an isotropic tensor. Here, we define the *radial isotropic tensor* as a transversely isotropic tensor along a given radial direction, i.e., a tensor whose properties in all directions perpendicular to the radial direction, $\bar{\mathbf{x}}$, are the same. In general, the radial isotropic tensor, depending on $\mathbf{x} = tA\bar{\mathbf{x}}$, can be expressed in the following canonical form

$$\begin{aligned}\mathbb{S}_{ijmn}(\mathbf{x}) &= S_1(t)\delta_{ij}\delta_{mn} + S_2(t)(\delta_{im}\delta_{jn}+\delta_{in}\delta_{jm}) + S_3(t)\delta_{ij}\bar{x}_m\bar{x}_n\\ &\quad + S_4(t)\delta_{mn}\bar{x}_i\bar{x}_j + S_5(t)(\delta_{im}\bar{x}_j\bar{x}_n + \delta_{in}\bar{x}_j\bar{x}_m + \delta_{jm}\bar{x}_i\bar{x}_n\\ &\quad + \delta_{jn}\bar{x}_i\bar{x}_m) + S_6(t)\bar{x}_i\bar{x}_j\bar{x}_m\bar{x}_n\\ &= \boldsymbol{\Theta}_{ijmn}^T(\bar{\mathbf{x}})\mathbf{S}(t)\end{aligned} \tag{25}$$

This canonical form decomposes $\mathbb{S}_{ijmn}$ into the circumference basis $\boldsymbol{\Theta}_{ijmn}$, which is only a function of the direction vector $\bar{\mathbf{x}}$, and into the radial basis $\mathbf{S}$, which is only a function of the dimensionless radial distance t.

It is well known that a transversely isotropic tensor has the similar symmetric properties (e.g., Ref. [15]). Using the definition

$$a_{ij} := \delta_{ij} - r_i r_j \tag{26}$$

$$b_{ij} := r_i r_j \tag{27}$$

which are the idempotent parts of a second-order unit tensor,

$$\delta_{ij} = a_{ij} + b_{ij} \tag{28}$$

one can show that the following six bases [15]

$$\mathbb{E}_{ijmn}^1 = \tfrac{1}{2}a_{ij}a_{mn} \tag{29}$$

$$\mathbb{E}_{ijmn}^2 = b_{ij}b_{mn} = r_i r_j r_m r_n \tag{30}$$

$$\mathbb{E}_{ijmn}^3 = \tfrac{1}{2}(a_{im}a_{jn} + a_{jm}a_{in} - a_{ij}a_{mn}) \tag{31}$$

$$\mathbb{E}_{ijmn}^4 = \tfrac{1}{2}(a_{im}b_{jn} + a_{jn}b_{jm} + a_{jn}b_{im} + a_{jm}b_{in}) \tag{32}$$

$$\mathbb{E}_{ijmn}^5 = a_{ij}b_{mn} \tag{33}$$

$$\mathbb{E}_{ijmn}^6 = b_{ij}a_{mn} \tag{34}$$

form a finite non-Abelian group. Furthermore

$$\mathbf{E}^p:\mathbf{E}^q = \mathbf{E}^p \quad \text{if } p=q, \quad \mathbf{E}^p:\mathbf{E}^q = 0 \quad \text{if } p \neq q, \quad p,q = 1,2,3,4 \tag{35}$$

where $\mathbf{E}^p = \mathbb{E}_{ijmn}^p \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_m \otimes \mathbf{e}_n$. Subsequently, for $p = 1,2,3,4,5,6$, one can find the "less congenial multiplication table" shown in Ref. [15].

Nevertheless, to the best of the authors' knowledge, we are the first to show that the circumference basis $\boldsymbol{\Theta}_{ijmn}$ of a spherical inclusion in a finite domain is a transversely isotropic tensor.[2] Instead of using the partially idempotent canonical form to represent the circumference basis $\boldsymbol{\Theta}_{ijmn}$ of a radial transversely isotropic tensor, we use an equivalent but different description introduced in Eq. (22).

Based on the symmetry of the problem, we now postulate that the Eshelby tensor for a finite RVE, $\mathbb{S}^{\cdot,\star}_{ijmn}$, should also be a radial isotropic tensor. It therefore admits the following multiplicative decomposition

$$\mathbb{S}^{\cdot,\star}_{ijmn}(\mathbf{x}) = \boldsymbol{\Theta}^T_{ijmn}(\bar{\mathbf{x}})\mathbf{S}^{\cdot,\star}(t) \tag{36}$$

with the superscripts, $\cdot = I$ or E and $\star = D$ or N. Here $\boldsymbol{\Theta}_{ijmn}(\bar{\mathbf{x}})$ is the circumference basis according to Eq. (22) and $\mathbf{S}^{\cdot,\star}(t)$ is the radial basis given as

$$\mathbf{S}^{\cdot,\star}(t) = [S_1^{\cdot,\star}(t), S_2^{\cdot,\star}(t), S_3^{\cdot,\star}(t), S_4^{\cdot,\star}(t), S_5^{\cdot,\star}(t), S_6^{\cdot,\star}(t)]^T \tag{37}$$

The scalar entries $S_J^{\cdot,\star}(t)$, $J=1,2,3,4,5,6$, are unknown functions of the nondimensional radial variable $t=|\mathbf{x}|/A$, which are to be determined.

The postulate above is motivated by the following two considerations. Due to the concentric and spherical symmetry of inclusion and RVE the tensorial basis of the finite Eshelby tensor can only depend on the radial direction vector $\bar{x}_i$ (and the second-order identity δ_{ij}). Therefore its tensorial basis, can only consist of combinations of zeroth-, second-, and fourth-order homogeneous functions of $\bar{\mathbf{x}}$. Furthermore due to the symmetry of the strain tensor the finite Eshelby tensor must have minor symmetries. Its tensorial basis can therefore only admit the six tensorial bases listed in $\boldsymbol{\Theta}_{ijmn}(\bar{\mathbf{x}})$. We note that one should expect more than six bases for problems described by more that one vector, such as ellipsoidal inclusions or non-concentrically placed inclusions within the RVE. Such problems may also be solvable with a similar procedure to ours. Due to the postulate the search for the finite Eshelby tensors reduces to the search for their radial basis $\mathbf{S}^{\cdot,\star}(t)$. We will see in the subsequent section, that the two solutions we obtain satisfy the governing equations exactly, thereby justifying postulate Eq. (36).

In analogy to Eq. (21), we can express the disturbance displacement field as

$$u_i^d(\mathbf{x}) = \begin{cases} \mathbb{U}^{I,\star}_{imn}(\mathbf{x})\epsilon^*_{mn}, & \forall \mathbf{x} \in \Omega_I \\ \mathbb{U}^{E,\star}_{imn}(\mathbf{x})\epsilon^*_{mn}, & \forall \mathbf{x} \in \Omega_E \end{cases} \tag{38}$$

where $\mathbb{U}^{I,\star}_{imn}(\mathbf{x})$ is a third-order radial isotropic tensor, whose relation to $\mathbb{S}^{I,\star}_{ijmn}(\mathbf{x})$ is discussed next. The disturbance strain is linked to the displacement field by the relation

$$\begin{aligned}\epsilon^d_{ij}(\mathbf{x}) &= \tfrac{1}{2}[u^d_{i,j}(\mathbf{x}) + u^d_{j,i}(\mathbf{x})] = \tfrac{1}{2}[\mathbb{U}^{\cdot,\star}_{imn,j}(\mathbf{x}) + \mathbb{U}^{\cdot,\star}_{jmn,i}(\mathbf{x})]\epsilon^*_{mn} \\ &= \mathbb{S}^{\cdot,\star}_{ijmn}(\mathbf{x})\epsilon^*_{mn}\end{aligned} \tag{39}$$

It can be shown that $\mathbb{U}^{\cdot,\star}_{imn}(\mathbf{x})$ can only admit the following multiplicative decomposition, so that the related Eshelby tensors $\mathbb{S}^{I,\star}_{ijmn}(\mathbf{x})$ are radial isotropic tensors

$$\mathbb{U}^{I,\star}_{imn}(\mathbf{x}) = \boldsymbol{\Xi}^T_{imn}(\bar{\mathbf{x}})\mathbf{U}^{I,\star}(t), \quad \forall\, \mathbf{x} \in \Omega_I \tag{40}$$

$$\mathbb{U}^{E,\star}_{imn}(\mathbf{x}) = \boldsymbol{\Xi}^T_{imn}(\bar{\mathbf{x}})\mathbf{U}^{E,\star}(t), \quad \forall\, \mathbf{x} \in \Omega_E \tag{41}$$

with the appearing arrays defined as

$$\mathbf{U}^{I,\star}(t) = \begin{bmatrix} U_1^{I,\star}(t) \\ U_2^{I,\star}(t) \\ U_3^{I,\star}(t) \end{bmatrix}, \quad \mathbf{U}^{E,\star}(t) = \begin{bmatrix} U_1^{E,\star}(t) \\ U_2^{E,\star}(t) \\ U_3^{E,\star}(t) \end{bmatrix}$$

and

$$\boldsymbol{\Xi}_{imn}(\bar{\mathbf{x}}) = \begin{bmatrix} \bar{x}_i\delta_{mn} \\ \bar{x}_m\delta_{in} + \bar{x}_n\delta_{im} \\ \bar{x}_i\bar{x}_m\bar{x}_n \end{bmatrix} \tag{42}$$

Here $\mathbf{U}^{I,\star}(t)$ and $\mathbf{U}^{E,\star}(t)$ are the radial basis arrays of the displacement field. $\boldsymbol{\Xi}_{imn}(\bar{\mathbf{x}})$ is the circumference basis array of the displacement field, whose third-order tensorial entries can only be first- or third-order homogeneous function of $\bar{\mathbf{x}}$. Hence the disturbance displacement field has the following canonical form

$$u_i^d(\mathbf{x}) = u_i^d(\bar{\mathbf{x}},t) = \begin{cases} \epsilon^*_{mn}\boldsymbol{\Xi}^T_{imn}(\bar{\mathbf{x}})\mathbf{U}^{I,\star}(t), & \forall \mathbf{x} \in \Omega_I \\ \epsilon^*_{mn}\boldsymbol{\Xi}^T_{imn}(\bar{\mathbf{x}})\mathbf{U}^{E,\star}(t), & \forall \mathbf{x} \in \Omega_E \end{cases} \tag{43}$$

Furthermore, the kinematic relation (39) yields the following differential mapping, which uniquely determines the relationship between the radial basis array of the strain field and the radial basis array of the displacement field

$$\mathbf{S}^{\cdot,\star}(t) = \mathcal{D}(t)\mathbf{U}^{\cdot,\star}(t) \tag{44}$$

where $\mathcal{D}(t)$ is a differential operator that is defined in matrix form

$$\mathcal{D}(t) = \frac{1}{A}\begin{bmatrix} \frac{1}{t} & 0 & 0 \\ 0 & \frac{1}{t} & 0 \\ 0 & 0 & \frac{1}{t} \\ -\frac{1}{t} + \frac{d}{dt} & 0 & 0 \\ 0 & -\frac{1}{2t} + \frac{1}{2}\frac{d}{dt} & \frac{1}{2t} \\ 0 & 0 & -\frac{3}{t} + \frac{d}{dt} \end{bmatrix}^{6\times 3} \tag{45}$$

Likewise, if $\mathbf{S}$ is given $\mathbf{U}$ can be determined from

$$\mathbf{U}^{\cdot,\star}(t) = \mathfrak{I}(t)\mathbf{S}^{\cdot,\star}(t) \tag{46}$$

where $\mathfrak{I}(t)$ is the integration operator

$$\mathfrak{I}(t) = tA\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}^{3\times 6} \tag{47}$$

We note that the displacements are only uniquely determinable up to the rigid body motion, which is set to zero here.

4 Eshelby Tensors for Finite Domains

For simplicity, in the rest of the paper, we term the Eshelby tensor for a finite domain as the finite Eshelby tensor.

4.1 The Dirichlet–Eshelby Tensor. We first consider the Dirichlet BVP Eq. (9) in which case $\star = D$. Substituting Eq. (21) into Eq. (16), one obtains a tensorial integral equation for the unknown finite Eshelby tensor

$$\begin{aligned}\mathbb{S}^{\cdot,D}_{ijmn}(\mathbf{x}) = \mathbb{S}^{\cdot,\infty}_{ijmn}(\mathbf{x}) + \frac{1}{2}\int_{\partial\Omega}[G^\infty_{ik,j}(\mathbf{x}-\mathbf{y}) + G^\infty_{jk,i}(\mathbf{x}-\mathbf{y})] \\ \times n_\ell(\mathbf{y})\mathbb{C}_{k\ell pq}\mathbb{S}^{E,D}_{pqmn}(\mathbf{y})d\Omega_y\end{aligned} \tag{48}$$

This integral equation has two different forms, depending on whether $\mathbf{x}$ is inside or outside the inclusion

$$\begin{aligned}\mathbb{S}^{I,D}_{ijmn}(\mathbf{x}) = \mathbb{S}^{I,\infty}_{ijmn}(\mathbf{x}) + \frac{1}{2}\int_{\partial\Omega}[G^\infty_{ik,j}(\mathbf{x}-\mathbf{y}) + G^\infty_{jk,i}(\mathbf{x}-\mathbf{y})] \\ \times n_\ell(\mathbf{y})\mathbb{C}_{k\ell pq}\mathbb{S}^{E,D}_{pqmn}(\mathbf{y})d\Omega_y \quad \forall\, \mathbf{x} \in \Omega_I\end{aligned} \tag{49}$$

[2]We first derived the result in a 2004 manuscript that was submitted to Proceedings of Royal Society of London.

$$\mathbb{S}^{E,D}_{ijmn}(\mathbf{x}) = \mathbb{S}^{E,\infty}_{ijmn}(\mathbf{x}) + \frac{1}{2}\int_{\partial\Omega}[G^{\infty}_{ik,j}(\mathbf{x}-\mathbf{y}) + G^{\infty}_{jk,i}(\mathbf{x}-\mathbf{y})] \times n_{\ell}(\mathbf{y})\mathbb{C}_{k\ell pq}\mathbb{S}^{E,D}_{pqmn}(\mathbf{y})d\Omega_y \quad \forall\, \mathbf{x}\in\Omega_E \tag{50}$$

Since $\mathbf{y}$ lies on the boundary, $\mathbf{y}\in\partial\Omega$, and since $\mathbf{y}=tA\bar{\mathbf{y}}$ with $t=1$, $\bar{\mathbf{y}}=\mathbf{n}$, we can use Eq. (36) to write $\mathbb{S}^{E,D}_{ijmn}(\mathbf{y})=\mathbf{\Theta}^{T}_{ijmn}(\mathbf{n})\mathbf{S}^{E,D}(1)$. The postulate, that the circumference basis of the Eshelby tensors for a finite spherical RVE is the same as for the Eshelby tensors in an infinite domain, is true only if the circumference basis is invariant under the boundary integral in Eqs. (49) and (50). This means that the Dirichlet boundary integral, which we denote by $\mathbb{S}^{B,D}_{ijmn}$, can be expressed in terms of the canonical form

$$\begin{aligned}\mathbb{S}^{B,D}_{ijmn}(\mathbf{x}) &= \frac{1}{2}\int_{\partial\Omega}\mathbb{C}_{k\ell st}[G^{\infty}_{ik,j}(\mathbf{x}-\mathbf{y}) \\ &\quad + G^{\infty}_{jk,i}(\mathbf{x}-\mathbf{y})]n_{\ell}(\mathbf{y})\mathbf{\Theta}^{T}_{stmn}(\mathbf{n})\mathbf{S}^{E,D}(1)dS_y \\ &= \int_{\partial\Omega}\mathbb{G}_{ijmn}(\mathbf{x},\mathbf{y})dS_y = \mathbf{\Theta}^{T}_{ijmn}(\bar{\mathbf{x}})\mathbf{S}^{B,D}(t)\end{aligned} \tag{51}$$

Here $\mathbb{G}_{ijmn}$ is the integrand of the boundary integral which follows as

$$\begin{aligned}\mathbb{G}_{ijmn}(\mathbf{x},\mathbf{y}) &= \frac{1}{2}\mathbb{C}_{k\ell st}[G^{\infty}_{ik,j}(\mathbf{x}-\mathbf{y}) + G^{\infty}_{jk,i}(\mathbf{x}-\mathbf{y})] \\ &\quad \times n_{\ell}(\mathbf{y})\mathbf{\Theta}^{T}_{stmn}(\mathbf{n})\mathbf{S}^{E,D}(1) \\ &= \frac{-1}{16\pi(1-\nu)\mu r^2}[T_1\bar{r}_k n_k\delta_{ij}\delta_{mn} + T_1(2\nu-1)(\bar{r}_i n_j \\ &\quad + \bar{r}_j n_i)\delta_{mn} + T_2(\bar{r}_m n_n + \bar{r}_n n_m)\delta_{ij} + T_2(2\nu-1)(\delta_{im}\bar{r}_j n_n \\ &\quad + \delta_{in}\bar{r}_j n_m + \delta_{jm}\bar{r}_i n_n + \delta_{jn}\bar{r}_i n_m) - 3T_1\bar{r}_i\bar{r}_j\bar{r}_k n_k\delta_{mn} \\ &\quad + T_3\bar{r}_k n_k n_m n_n\delta_{ij} - 3T_2(\bar{r}_i\bar{r}_j\bar{r}_m n_n + \bar{r}_i\bar{r}_j\bar{r}_n n_m) + T_3(2\nu \\ &\quad - 1)(\bar{r}_i n_j n_m n_n + \bar{r}_j n_i n_m n_n) - 3T_3\bar{r}_i\bar{r}_j\bar{r}_k n_k n_m n_n]\end{aligned} \tag{52}$$

where $\bar{r}_i=(y_i-x_i)/r$, $r=|\mathbf{y}-\mathbf{x}|$ and $n_i(\mathbf{y})=y_i/|\mathbf{y}|$ for $\mathbf{y}\in\partial\Omega$. The coefficient vector

$$\mathbf{T} = [T_1, T_2, T_3]^T \tag{53}$$

is a stress projection vector (see Sec. 4.3), which follows as $\mathbf{T}=\mathbf{K}_1\mathbf{S}^{E,D}(1)$ where

$$\mathbf{K}_1 = \mu\begin{bmatrix} \frac{2+2\nu}{1-2\nu} & \frac{4\nu}{1-2\nu} & 0 & \frac{2(1-\nu)}{1-2\nu} & 0 & 0 \\ 0 & 2 & 0 & 0 & 2 & 0 \\ 0 & 0 & \frac{2+2\nu}{1-2\nu} & 0 & \frac{4}{1-2\nu} & \frac{2(1-\nu)}{1-2\nu}\end{bmatrix} \tag{54}$$

With the aid of the following integrals (see the Appendix)

$$\text{(I)} \quad \int_{\partial\Omega}\frac{1}{r^2}\bar{r}_k n_k\, dS_y = 4\pi \tag{55}$$

$$\text{(I)} \quad \int_{\partial\Omega}\frac{1}{r^2}\bar{r}_i n_j\, dS_y = \frac{4\pi}{3}\delta_{ij} \tag{56}$$

$$\text{(III)} \quad \int_{\partial\Omega}\frac{1}{r^2}\bar{r}_i\bar{r}_j\bar{r}_k n_k\, dS_y = \frac{4\pi}{3}\delta_{ij} \tag{57}$$

$$\text{(IV)} \quad \int_{\partial\Omega}\frac{1}{r^2}\bar{r}_k n_k n_m n_n\, dS_y = \frac{4\pi}{15}(5-3t^2)\delta_{mn} + \frac{12\pi}{5}t^2\bar{x}_m\bar{x}_n \tag{58}$$

$$\text{(V)} \quad \int_{\partial\Omega}\frac{1}{r^2}\bar{r}_i\bar{r}_j(\bar{r}_m n_n + \bar{r}_n n_m)dS_y = \frac{8\pi}{15}(\delta_{ij}\delta_{mn} + \delta_{im}\delta_{jn} + \delta_{jm}\delta_{in}) \tag{59}$$

$$\begin{aligned}\text{(VI)} \quad &\int_{\partial\Omega}\frac{1}{r^2}(\bar{r}_i n_j + \bar{r}_j n_i)n_m n_n\, dS_y \\ &= \frac{\pi}{105}[(56-24t^2)(\delta_{ij}\delta_{mn} + \delta_{im}\delta_{jn} + \delta_{jm}\delta_{in}) + 120t^2\delta_{ij}\bar{x}_m\bar{x}_n \\ &\quad - 48t^2\bar{x}_i\bar{x}_j\delta_{mn} + 36t^2(\delta_{im}\bar{x}_j\bar{x}_n + \delta_{in}\bar{x}_j\bar{x}_m \\ &\quad + \delta_{jm}\bar{x}_i\bar{x}_n + \delta_{jn}\bar{x}_i\bar{x}_m)]\end{aligned} \tag{60}$$

$$\begin{aligned}\text{(VII)} \quad &\int_{\partial\Omega}\frac{1}{r^2}\bar{r}_i\bar{r}_j\bar{r}_k n_k n_m n_n\, dS_y \\ &= \frac{\pi}{105}[(28-20t^2)(\delta_{ij}\delta_{mn} + \delta_{im}\delta_{jn} + \delta_{jm}\delta_{in}) + 100t^2\delta_{ij}\bar{x}_m\bar{x}_n \\ &\quad + 16t^2\bar{x}_i\bar{x}_j\delta_{mn} - 12t^2(\delta_{im}\bar{x}_j\bar{x}_n + \delta_{in}\bar{x}_j\bar{x}_m \\ &\quad + \delta_{jm}\bar{x}_i\bar{x}_n + \delta_{jn}\bar{x}_i\bar{x}_m)]\end{aligned} \tag{61}$$

we obtain the boundary contribution by explicit integration as

$$\mathbb{S}^{B,D}_{ijmn}(\mathbf{x}) = \mathbf{\Theta}^{T}_{ijmn}(\bar{\mathbf{x}})\mathbf{S}^{B,D}(t) \tag{62}$$

with

$$\mathbf{S}^{B,D}(t) = \mathbf{K}_2(t)\mathbf{K}_1\mathbf{S}^{E,D}(1)$$

and

$$\mathbf{K}_2(t) = \frac{-1}{420(1-\nu)} \times \begin{bmatrix} 70(2\nu-1) & 28 & 4\nu(7-3t^2) \\ 0 & 28(5\nu-4) & 7(4\nu-5)+3t^2(7-4\nu) \\ 0 & 0 & 6t^2(10\nu-7) \\ 0 & 0 & -24\nu t^2 \\ 0 & 0 & 18\nu t^2 \\ 0 & 0 & 0 \end{bmatrix} \tag{63}$$

The integral Eqs. (49) and (50) can then be reduced to a pair of algebraic evolution equations,

$$\mathbf{\Theta}^{T}_{ijmn}(\bar{\mathbf{x}})\mathbf{S}^{I,D}(t) = \mathbf{\Theta}^{T}_{ijmn}(\bar{\mathbf{x}})[\mathbf{S}^{I,\infty}(t) + \mathbf{K}_3(t)\mathbf{S}^{E,D}(1)] \tag{64}$$

$$\mathbf{\Theta}^{T}_{ijmn}(\bar{\mathbf{x}})\mathbf{S}^{E,D}(t) = \mathbf{\Theta}^{T}_{ijmn}(\bar{\mathbf{x}})[\mathbf{S}^{E,\infty}(t) + \mathbf{K}_3(t)\mathbf{S}^{E,D}(1)] \tag{65}$$

where $\mathbf{K}_3(t) := \mathbf{K}_2(t)\mathbf{K}_1$. The matrix $\mathbf{K}_2(t)$ maps the effect of the boundary traction onto the domain of the RVE.

Eliminating the circumference basis from Eqs. (64) and (65), we derive the following parametric algebraic equations

$$\mathbf{S}^{I,D}(t) = \mathbf{S}^{I,\infty}(t) + \mathbf{K}_3(t)\mathbf{S}^{E,D}(1), \quad 0\leq t<\rho_0 \tag{66}$$

$$\mathbf{S}^{E,D}(t) = \mathbf{S}^{E,\infty}(t) + \mathbf{K}_3(t)\mathbf{S}^{E,D}(1), \quad \rho_0\leq t<1 \tag{67}$$

Let us assume that $\mathbf{S}^{E,D}(t)$ continuously depends on t so that

$$\lim_{t\to 1}\mathbf{S}^{E,D}(t) \to \mathbf{S}^{E,D}(1) \tag{68}$$

Now let $t\to 1$ in Eq. (67), so that we can obtain $\mathbf{S}^{E,D}(1)$ by solving Eq. (67), i.e.,

$$
\mathbf{S}^{E,D}(1) = [\mathbf{I} - \mathbf{K}_3(1)]^{-1}\mathbf{S}^{E,\infty}(1) = \rho_0^3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ \dfrac{7 - 28\nu + 20\nu^2 - 7(1-2\nu)\rho_0^2}{2(1-\nu)(7-10\nu)} \\ \dfrac{10\nu - 7\rho_0^2}{2\ (7-10\nu)} \\ \dfrac{7\rho_0^2 - 5}{2(1-\nu)} \end{bmatrix} \tag{69}
$$

The boundary contribution $\mathbf{S}^{B,D}(t) = \mathbf{K}_3(t)\mathbf{S}^{E,D}(1)$ then becomes

$$
\mathbf{S}^{B,D}(t) = -\frac{\rho_0^3}{15(1-\nu)} \begin{bmatrix} 5\nu - 1 \\ 4 - 5\nu \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \frac{\rho_0^3(1-\rho_0^2)}{20(1-\nu)(7-10\nu)} \begin{bmatrix} 2(7-10\nu t^2) \\ 7(5t^2-3) - 20\nu t^2 \\ -10t^2(7-10\nu) \\ -40\nu t^2 \\ 30\nu t^2 \\ 0 \end{bmatrix} \tag{70}
$$

With this the considered Dirichlet–Eshelby problem of a spherical RVE is fully solved. The radial basis arrays of the Dirichlet–Eshelby tensors is given by

$$
\mathbf{S}^{I,D}(t) = \mathbf{S}^{I,\infty}(t) + \mathbf{S}^{B,D}(t), \quad 0 \leqslant t < \rho_0 \tag{71}
$$

$$
\mathbf{S}^{E,D}(t) = \mathbf{S}^{E,\infty}(t) + \mathbf{S}^{B,D}(t), \quad \rho_0 \leqslant t \leqslant 1 \tag{72}
$$

and from Eq. (36) we finally obtain the Dirichlet–Eshelby tensors. The interior solution is

$$
\begin{aligned}
\mathbb{S}^{I,D}_{ijmn}(\mathbf{x}) = \frac{1}{1-\nu}\Bigg\{ & \left[\frac{5\nu - 1}{15}(1-\rho_0^3) + \frac{7 - 10\nu t^2}{10(7-10\nu)}\rho_0^3(1-\rho_0^2)\right]\delta_{ij}\delta_{mn} \\
& + \left[\frac{4-5\nu}{15}(1-\rho_0^3) + \frac{7(5t^2-3) - 20\nu t^2}{20(7-10\nu)}\rho_0^3(1-\rho_0^2)\right] \\
& \times(\delta_{im}\delta_{jn} + \delta_{in}\delta_{jm}) - \frac{t^2}{2}\rho_0^3(1-\rho_0^2)\delta_{ij}\bar{x}_m\bar{x}_n - \frac{2\nu t^2}{7-10\nu}\rho^3(1 \\
& - \rho_0^2)\delta_{mn}\bar{x}_i\bar{x}_j + \frac{3\nu t^2}{2(7-10\nu)}\rho_0^3(1-\rho_0^2)(\delta_{im}\bar{x}_j\bar{x}_n + \delta_{in}\bar{x}_j\bar{x}_m \\
& + \delta_{jm}\bar{x}_i\bar{x}_n + \delta_{jn}\bar{x}_i\bar{x}_m)\Bigg\}
\end{aligned} \tag{73}
$$

and the exterior solution is

$$
\begin{aligned}
\mathbb{S}^{E,D}_{ijmn}(\mathbf{x}) = \frac{\rho_0^3}{1-\nu}\Bigg\{ & \left[\frac{3\rho_0^2/t^2 + 10\nu - 5}{30t^3} - \frac{5\nu-1}{15} + \frac{7-10\nu t^2}{10(7-10\nu)}(1-\rho_0^2)\right]\delta_{ij}\delta_{mn} \\
& + \left[\frac{3\rho_0^2/t^2 - 10\nu + 5}{30t^3} - \frac{4-5\nu}{15} + \frac{7(5t^2-3) - 20\nu t^2}{20(7-10\nu)}(1-\rho_0^2)\right](\delta_{im}\delta_{jn} + \delta_{in}\delta_{jm}) \\
& - \left[\frac{\rho_0^2/t^2 - 1}{2t^3} + \frac{t^2}{2}(1-\rho_0^2)\right]\delta_{ij}\bar{x}_m\bar{x}_n - \left[\frac{\rho_0^2/t^2 + 2\nu - 1}{2t^3} + \frac{2\nu t^2}{7-10\nu}(1-\rho_0^2)\right]\delta_{mn}\bar{x}_i\bar{x}_j \\
& - \left[\frac{\rho_0^2/t^2 - \nu}{2t^3} - \frac{3\nu t^2}{2(7-10\nu)}(1-\rho_0^2)\right](\delta_{im}\bar{x}_j\bar{x}_n + \delta_{in}\bar{x}_j\bar{x}_m + \delta_{jm}\bar{x}_i\bar{x}_n + \delta_{jn}\bar{x}_i\bar{x}_m) + \frac{7\rho_0^2/t^2 - 5}{2t^3}\bar{x}_i\bar{x}_j\bar{x}_m\bar{x}_n\Bigg\}
\end{aligned} \tag{74}
$$

We can see that both the interior and the exterior Dirichlet Eshelby tensor are neither constant nor isotropic. The dependency on the position $\mathbf{x}$ is captured by the dependency on $\bar{\mathbf{x}}$ and t. Furthermore, both tensors depend explicitly on the ratio ρ_0 between inclusion and RVE. If we let $\rho_0 \to 0$ we recover the original infinite Eshelby tensors exactly since the boundary contribution then vanishes. To visualize the Dirichlet–Eshelby tensors the profiles of the components of the radial basis arrays $\mathbf{S}^{\cdot,\infty}(t)$, $\mathbf{S}^{\cdot,D}(t)$ and $\mathbf{S}^{B,D}(t)$ are shown in Fig. 2. Here the relative size of the inclusion is chosen as $\rho_0=0.4$, so that the volume fraction becomes $\rho_0^3=0.064$. Poisson's ratio of the matrix phase is picked as $\nu=0.3$. One can clearly observe that the boundary term $\mathbf{S}^{B,D}$, which can be understood as a correction of Eshelby's original result, is substantial. It can also be noted that there is a discontinuity across the interface between the inclusion and the matrix.

The disturbance displacement field $u_i^d(\mathbf{x})$ is now given by Eqs. (38), (40), and (41), i.e.,

$$
u_i^d(\mathbf{x}) = \mathbf{\Xi}^T_{imn}(\bar{\mathbf{x}})\mathbf{U}^{\cdot,D}(t)\epsilon^*_{mn}, \quad \forall\, \mathbf{x} \in \Omega \tag{75}
$$

where the arrays $\mathbf{\Xi}_{imn}(\bar{\mathbf{x}})$, $\mathbf{U}^{I,D}(t)$, and $\mathbf{U}^{E,D}(t)$ follow from Eqs. (42) and (46). Applying operator (47) to Eqs. (71), (72), (24), and (70) we easily obtain

$$
\mathbf{U}^{I,D}(t) = \mathbf{U}^{I,\infty}(t) + \mathbf{U}^{B,D}(t), \quad 0 \leqslant t < \rho_0 \tag{76}
$$

$$
\mathbf{U}^{E,D}(t) = \mathbf{U}^{E,\infty}(t) + \mathbf{U}^{B,D}(t), \quad \rho_0 \leqslant t \leqslant 1 \tag{77}
$$

with

$$
\mathbf{U}^{I,\infty}(t) = \frac{tA}{15(1-\nu)}\begin{bmatrix} 5\nu - 1 \\ 4 - 5\nu \\ 0 \end{bmatrix} \tag{78}
$$

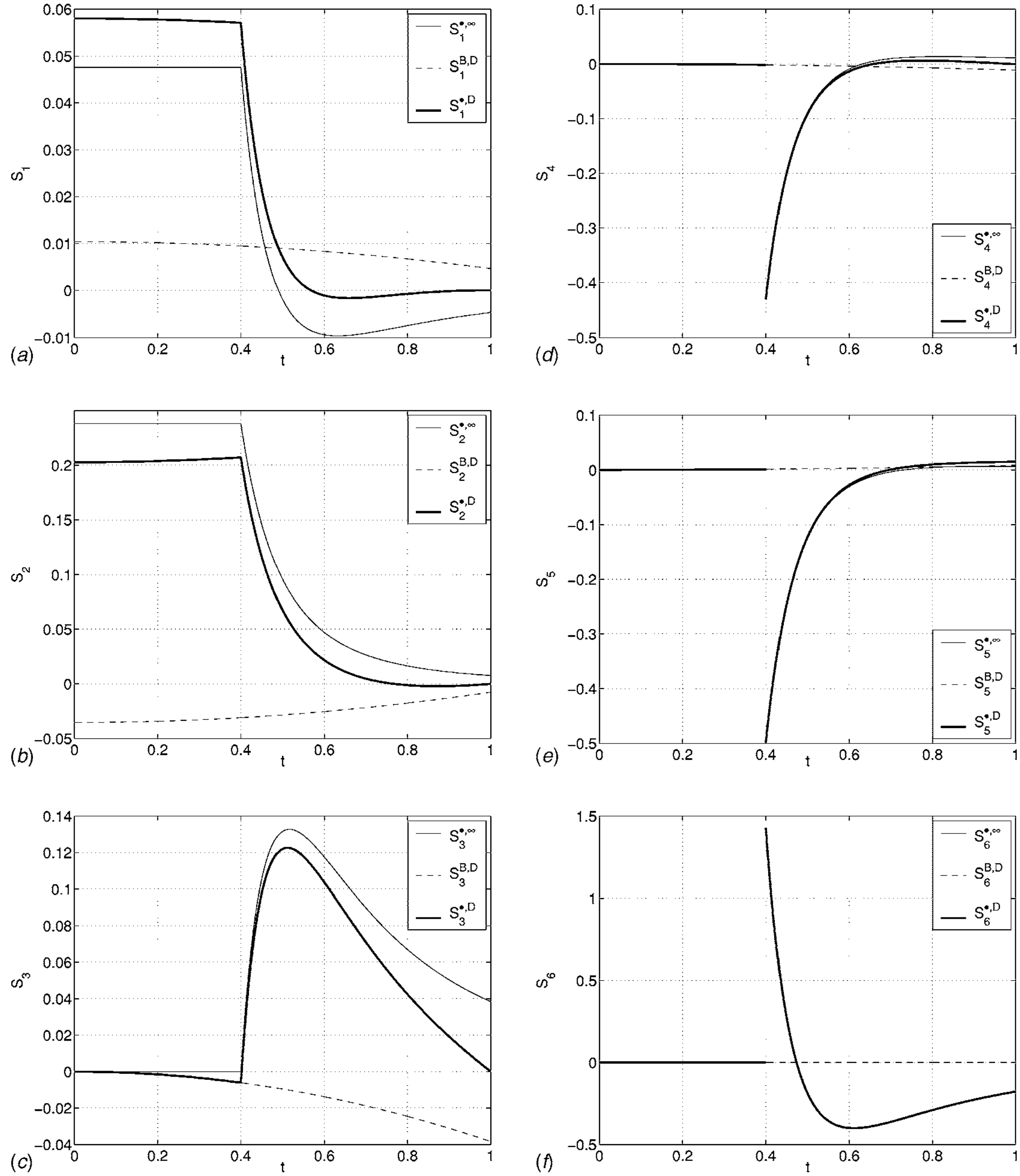


Fig. 2 The components of the radial basis arrays $\mathbf{S}^{\cdot,\infty}$, $\mathbf{S}^{B,D}$, and $\mathbf{S}^{\cdot,D}$

$$\mathbf{U}^{E,\infty}(t) = \frac{\rho_0^3 A}{30t^2(1-\nu)} \begin{bmatrix} 3\rho_0^2/t^2 + 10\nu - 5 \\ 3\rho_0^2/t^2 - 10\nu + 5 \\ 15 - 15\rho_0^2/t^2 \end{bmatrix} \tag{79}$$

and

$$\mathbf{U}^{B,D}(t) = -\frac{\rho_0^3 tA}{15(1-\nu)} \begin{bmatrix} 5\nu - 1 \\ 4 - 5\nu \\ 0 \end{bmatrix} + \frac{\rho_0^3(1-\rho_0^2)tA}{20(1-\nu)(7-10\nu)} \begin{bmatrix} 2(7-10\nu t^2) \\ 7(5t^2-3) - 20\nu t^2 \\ -10t^2(7-10\nu) \end{bmatrix} \tag{80}$$

Here $\mathbf{U}^{I,\infty}(t)$ and $\mathbf{U}^{E,\infty}(t)$ are the radial basis array of Eshelby's classical solution in unbounded space and $\mathbf{U}^{B,D}(t)$ is the radial basis contribution from the Dirichlet boundary of the RVE.

We remark that u_i^d given by the equations above satisfies the Fredholm-type integral equation of the Dirichlet BVP Eq. (14) exactly. Furthermore it is readily verified that when $t=1$

$$\mathbf{U}^{E,D}(1) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \to u_i^d(\mathbf{y}) = \epsilon_{mn}^* \mathbf{\Xi}_{imn}^T(\mathbf{n})\mathbf{U}^{E,D}(1) = 0, \quad \forall\, \mathbf{y} \in \partial\Omega \tag{81}$$

This confirms that the obtained displacement solution does indeed satisfy the Dirichlet boundary condition. The coefficients of the radial bases $\mathbf{U}^{\cdot,\infty}$, $\mathbf{U}^{B,D}$, and $\mathbf{U}^{\cdot,D}$ are displayed in Figs. 3(*a*), 3(*c*), and 3(*e*), where we have chosen $\rho_0=0.4$ and $\nu=0.3$. Again, we observe that the boundary correction is substantial. Further, one can see that the Dirichlet solution satisfies the zero displacement boundary condition exactly.

4.2 The Neumann–Eshelby Tensor. The solution of the Neumann–Eshelby problem (now $\star=N$) is different from the

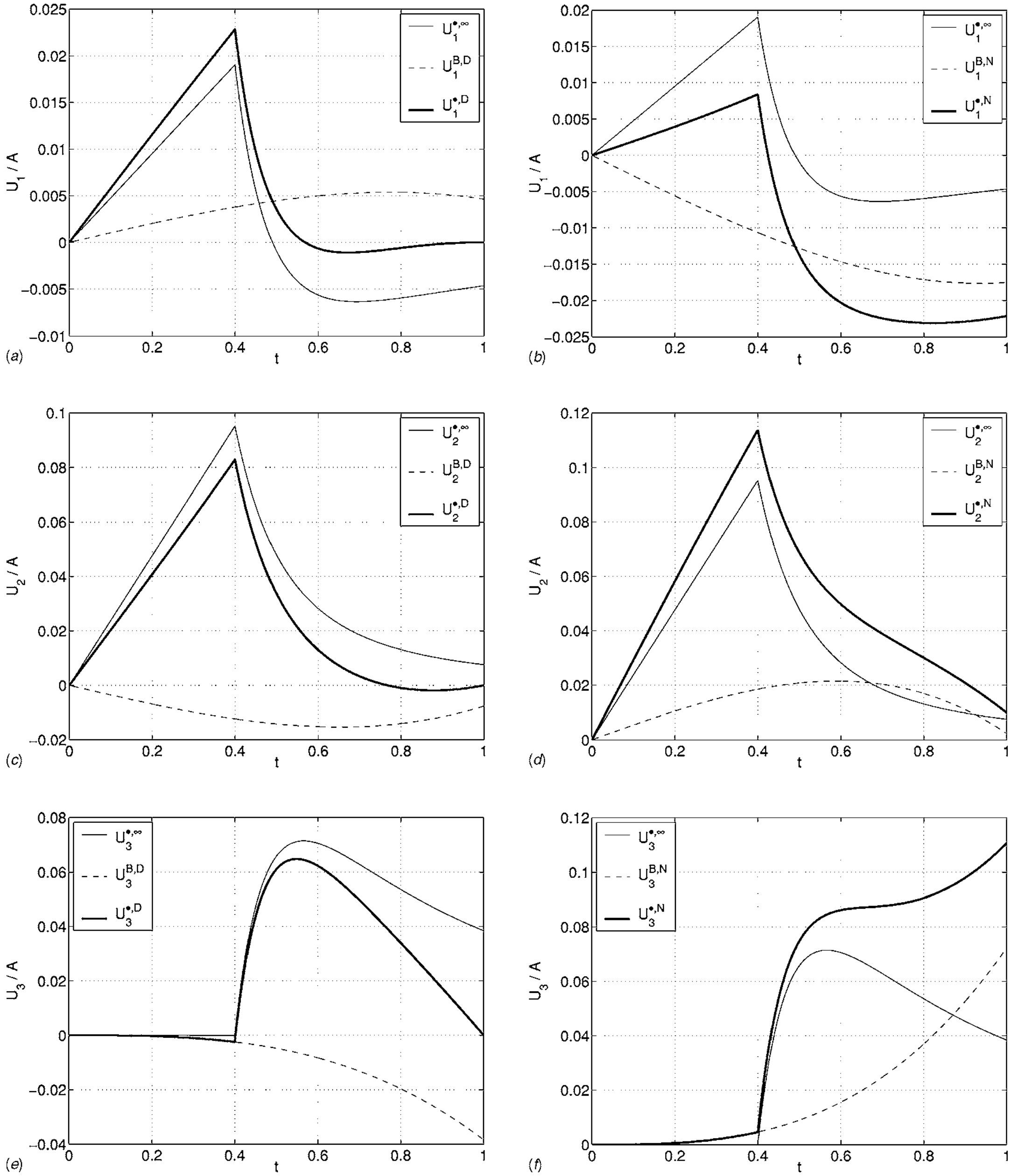


Fig. 3 The components of the radial basis arrays $\mathbf{U}^{\cdot,\infty}$, $\mathbf{U}^{B,D}$, $\mathbf{U}^{\cdot,D}$, $\mathbf{U}^{B,N}$, and $\mathbf{U}^{\cdot,N}$

Dirichlet–Eshelby problem tensor; here the solution is based on the displacement field. For the Neumann BVP Eq. (10), the displacements on the boundary of the RVE ($t=1$) are nonzero and according to Eq. (43) we have

$$u_k^d(\mathbf{y}) = \epsilon_{mn}^* \mathbf{\Xi}_{kmn}^T(\mathbf{n})\mathbf{U}^{E,N}(1), \quad \forall\, \mathbf{y} \in \partial\Omega \tag{82}$$

By substituting Eq. (82) into the integral equation corresponding to the Neumann BVP Eq. (15), we obtain an equation for the unknown radial basis, $\mathbf{U}^{\cdot,N}(t)$

$$\begin{aligned}\epsilon_{mn}^* \mathbf{\Xi}_{imn}^T(\bar{\mathbf{x}})\mathbf{U}^{\cdot,N}(t) &= -\epsilon_{mn}^* \int_{\Omega_e} \mathbb{C}_{pqmn} G_{pi,q}^{\infty}(\mathbf{x}-\mathbf{y})\, d\Omega_y \\ &\quad + \epsilon_{mn}^* \int_{\partial\Omega} \mathbb{C}_{pqk\ell} G_{pi,q}^{\infty}(\mathbf{x}-\mathbf{y}) \mathbf{\Xi}_{kmn}^T(\mathbf{n})\mathbf{U}^{E,N}(1) n_\ell(\mathbf{y})\, dS_y\end{aligned} \tag{83}$$

where $\cdot = I$, or E. Depending on whether $\mathbf{x}$ is inside or outside the inclusion, the domain integral in Eq. (83) has two different forms, which can be expressed in the canonical form

$$-\int_{\Omega_e} \mathbb{C}_{pqmn} G_{pi,q}^{\infty}(\mathbf{x}-\mathbf{y})\, d\Omega_y = \begin{cases} \mathbf{\Xi}_{imn}^T(\bar{\mathbf{x}})\mathbf{U}^{I,\infty}(t), & \forall\, \mathbf{x} \in \Omega_I \\ \mathbf{\Xi}_{imn}^T(\bar{\mathbf{x}})\mathbf{U}^{E,\infty}(t), & \forall\, \mathbf{x} \in \Omega_E \end{cases} \tag{84}$$

Here $\mathbf{U}^{I,\infty}(t)$ and $\mathbf{U}^{E,\infty}(t)$ are the radial basis arrays of Eshelby's classical solution for unbounded space (see Eqs. (78) and (79)). In analogy to the Dirichlet case (see Eq. (51)), we stipulate that a similar canonical form holds for the Neumann boundary contribution in Eq. (83)

$$\int_{\partial\Omega} \mathbb{C}_{pqk\ell} G^{\infty}_{pi,q}(\mathbf{x}-\mathbf{y}) \boldsymbol{\Xi}^T_{kmn}(\mathbf{n}) \mathbf{U}^{E,N}(1) n_\ell(\mathbf{y}) dS_y$$

$$= \boldsymbol{\Xi}^T_{imn}(\bar{\mathbf{x}}) \mathbf{U}^{B,N}(t), \quad \forall\, \mathbf{x} \in \Omega \tag{85}$$

where $\mathbf{U}^{B,N}(t)$ denotes the radial basis array arising from the Neumann boundary. Substituting Eqs. (84) and (85) into Eq. (83) and eliminating ϵ^*_{mn} and the circumference basis $\Xi^T_{imn}(\bar{\mathbf{x}})$, one may reduce Eq. (83) into a pair of parametric, algebraic equations for the radial basis arrays, $\mathbf{U}^{\cdot,N}(t)$, i.e.,

$$\mathbf{U}^{I,N}(t) = \mathbf{U}^{I,\infty}(t) + \mathbf{U}^{B,N}(t), \quad 0 \le t \le \rho_0 \tag{86}$$

$$\mathbf{U}^{E,N}(t) = \mathbf{U}^{E,\infty}(t) + \mathbf{U}^{B,N}(t), \quad \rho_0 \le t \le 1 \tag{87}$$

Here $\mathbf{U}^{I,\infty}(t)$ and $\mathbf{U}^{E,\infty}(t)$ are the radial basis vectors of Eshelby's classical solution of unbounded space (see Eqs. (78) and (79)). The boundary contribution, $\mathbf{U}^{B,N}(t)$ follows directly from Eq. (85) as

$$\boldsymbol{\Xi}^T_{imn}(\bar{\mathbf{x}}) \mathbf{U}^{B,N}(t) = \int_{\partial\Omega} \mathbb{H}_{imn}(\mathbf{x},\mathbf{y}) dS_y \tag{88}$$

with the integrand

$$\begin{aligned}
\mathbb{H}_{imn} &= \mathbb{C}_{pqk\ell} G^{\infty}_{pi,q}(\mathbf{x}-\mathbf{y}) \boldsymbol{\Xi}^T_{kmn}(\mathbf{n}) \mathbf{U}^{E,N}(1) n_\ell(\mathbf{y}) \\
&= \frac{1}{8\pi(1-\nu) r^2} \{ U^{E,N}_1(1)[(1-2\nu)(2 n_i n_p \bar{r}_p \delta_{mn} - \bar{r}_i \delta_{mn}) \\
&\quad + 3\bar{r}_i n_p \bar{r}_p n_q \bar{r}_q \delta_{mn}] + U^{E,N}_2(1)[(1-2\nu)(n_m n_p \bar{r}_p \delta_{in} + n_n n_p \bar{r}_p \delta_{im} \\
&\quad + n_i n_m \bar{r}_n + n_i n_n \bar{r}_m - 2 n_m n_n \bar{r}_i) + 3 n_m \bar{r}_i \bar{r}_n n_p \bar{r}_p + 3 n_n \bar{r}_i \bar{r}_m n_p \bar{r}_p] \\
&\quad + U^{E,N}_3(1)[(1-2\nu)(2 n_i n_m n_n \bar{r}_p n_p - n_m n_n \bar{r}_i) \\
&\quad + 3 n_m n_n \bar{r}_i n_p \bar{r}_p n_q \bar{r}_q]\}
\end{aligned} \tag{89}$$

With the aid of following integrals (see the Appendix)

$$\text{(VIII)} \quad \int_{\partial\Omega} \frac{1}{r^2} \bar{r}_i \, dS_y = 0 \tag{90}$$

$$\text{(IX)} \quad \int_{\partial\Omega} \frac{1}{r^2} n_i n_k \bar{r}_k \, dS_y = \frac{8\pi}{3} t \bar{x}_i \tag{91}$$

$$\text{(X)} \quad \int_{\partial\Omega} \frac{1}{r^2} n_i n_j \bar{r}_k \, dS_y = \frac{4\pi}{15} t (3\bar{x}_i \delta_{jk} + 3\bar{x}_j \delta_{ik} - 2\bar{x}_k \delta_{ij}) \tag{92}$$

$$\text{(XI)} \quad \int_{\partial\Omega} \frac{1}{r^2} n_i \bar{r}_j \bar{r}_k n_p \bar{r}_p \, dS_y = \frac{4\pi}{15} t (4\bar{x}_i \delta_{jk} - \bar{x}_j \delta_{ik} - \bar{x}_k \delta_{ij}) \tag{93}$$

$$\begin{aligned}
\text{(XII)} \quad \int_{\partial\Omega} \frac{1}{r^2} n_i n_j n_k n_p \bar{r}_p \, dS_y &= \frac{\pi}{105} [t(56 - 48t^2)(\bar{x}_i \delta_{jk} + \bar{x}_j \delta_{ik} \\
&\quad + \bar{x}_k \delta_{ij}) + 240 t^3 \bar{x}_i \bar{x}_j \bar{x}_k]
\end{aligned} \tag{94}$$

$$\text{(XIII)} \quad \int_{\partial\Omega} \frac{1}{r^2} \bar{r}_i n_p \bar{r}_p n_q \bar{r}_q \, dS_y = 0 \tag{95}$$

$$\begin{aligned}
\text{(XIV)} \quad \int_{\partial\Omega} \frac{1}{r^2} n_i n_j \bar{r}_k n_p \bar{r}_p n_q \bar{r}_q \, dS_y &= \frac{\pi}{105} [t(84 - 80t^2)(\bar{x}_i \delta_{jk} + \bar{x}_j \delta_{ik}) \\
&\quad - t(56 - 32t^2) \bar{x}_k \delta_{ij} + 64 t^3 \bar{x}_i \bar{x}_j \bar{x}_k]
\end{aligned} \tag{96}$$

Eq. (88) can be integrated exactly. After some manipulations, the final result can be expressed in a succinct form

$$\mathbf{U}^{B,N}(t) = \mathbf{K}_4(t) \mathbf{U}^{E,N}(1) \tag{97}$$

where

$$\mathbf{K}_4(t) = \frac{t}{1-\nu} \begin{bmatrix} \dfrac{2(1-2\nu)}{3} & \dfrac{2(1-5\nu)}{15} & \dfrac{-2\nu(7-4t^2)}{35} \\ 0 & \dfrac{7-5\nu}{15} & \dfrac{7(5-\nu)+6t^2(4\nu-7)}{105} \\ 0 & 0 & \dfrac{4(7-10\nu)t^2}{35} \end{bmatrix} \tag{98}$$

Equation (97) represents the boundary contribution or "the image contribution" to the disturbance displacement field inside the RVE. Now, the parametric algebraic equations are solely in terms of the displacement radial basis array $\mathbf{U}^{\cdot,N}(t)$

$$\mathbf{U}^{I,N}(t) = \mathbf{U}^{I,\infty}(t) + \mathbf{K}_4(t) \mathbf{U}^{E,N}(1), \quad 0 \le t \le \rho_0 \tag{99}$$

$$\mathbf{U}^{E,N}(t) = \mathbf{U}^{E,\infty}(t) + \mathbf{K}_4(t) \mathbf{U}^{E,F}(1), \quad \rho_0 \le t < 1 \tag{100}$$

We assume that the radial basis array, $\mathbf{U}^{E,N}(t)$, depends continuously on t so that

$$\lim_{t\to 1} \mathbf{U}^{E,N}(t) = \mathbf{U}^{E,N}(1) \tag{101}$$

One can then solve for $\mathbf{U}^{E,N}(1)$ by letting $t=1$ in Eq. (100), i.e.,

$$\mathbf{U}^{E,N}(1) = [\mathbf{I} - \mathbf{K}_4(1)]^{-1} \mathbf{U}^{E,\infty}(1) \tag{102}$$

which gives

$$\mathbf{U}^{E,N}(1) = \frac{\rho_0^3 A}{2(7+5\nu)} \begin{bmatrix} 7(\rho_0^2 - 1) \\ 5\nu + 7\rho_0^2 \\ 35(1-\rho_0^2) \end{bmatrix} \tag{103}$$

Substituting Eq. (103) into Eq. (97), one can evaluate the radial basis array due to the boundary or image contribution

$$\begin{aligned}
\mathbf{U}^{B,N}(t) &= \frac{\rho_0^3 t A}{30(1-\nu)} \begin{bmatrix} 2-10\nu \\ 7-5\nu \\ 0 \end{bmatrix} \\
&\quad - \frac{\rho_0^3(1-\rho_0^2) t A}{5(1-\nu)(7+5\nu)} \begin{bmatrix} 2(7-10\nu t^2) \\ 7(5t^2-3) - 20\nu t^2 \\ -10t^2(7-10\nu) \end{bmatrix}
\end{aligned} \tag{104}$$

Note the similarity between the two boundary contributions $\mathbf{U}^{B,N}(t)$ and $\mathbf{U}^{B,D}(t)$ (see Eq. (80)). With the above result we can now find $\mathbf{U}^{I,N}(t)$ and $\mathbf{U}^{E,N}(t)$ from Eqs. (86) and (87).

With the radial basis arrays of the displacement field given, one can apply the differential operator Eq. (45) to obtain the radial basis array of the strain field, i.e.,

$$\mathbf{S}^{I,\infty}(t) = \mathcal{D}(t) \mathbf{U}^{I,\infty}(t), \quad \mathbf{S}^{E,\infty}(t) = \mathcal{D}(t) \mathbf{U}^{E,\infty}(t)$$

and

$$\mathbf{S}^{B,N}(t) = \mathcal{D}(t) \mathbf{U}^{B,N}(t) \tag{105}$$

$\mathbf{S}^{I,\infty}(t)$ and $\mathbf{S}^{E,\infty}(t)$ follow as given in Eq. (24); furthermore by differentiation we find

$$\mathbf{S}^{B,N}(t) = \frac{\rho_0^3}{30(1-\nu)}\begin{bmatrix} 2-10\nu \\ 7-5\nu \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \frac{\rho_0^3(1-\rho_0^2)}{5(1-\nu)(7+5\nu)}\begin{bmatrix} 2(7-10\nu t^2) \\ 7(5t^2-3)-20\nu t^2 \\ -10t^2(7-10\nu) \\ -40\nu t^2 \\ 30\nu t^2 \\ 0 \end{bmatrix} \tag{106}$$

In analogy to Eqs. (86) and (87), the radial basis arrays $\mathbf{S}^{\cdot,N}$ of the Neumann–Eshelby tensors now follows from

$$\mathbf{S}^{I,N}(t) = \mathbf{S}^{I,\infty}(t) + \mathbf{S}^{B,N}(t), \quad 0 \leqslant t \leqslant \rho_0 \tag{107}$$

$$\mathbf{S}^{E,N}(t) = \mathbf{S}^{E,\infty}(t) + \mathbf{S}^{B,N}(t), \quad \rho_0 \leqslant t \leqslant 1 \tag{108}$$

The Neumann–Eshelby tensors for a spherical inclusion embedded in a spherical RVE under the prescribed traction boundary condition can now be obtained from Eq. (22) which yields the following exact and elementary expressions. The interior solution is

$$\begin{aligned}\mathbb{S}^{I,N}_{ijmn}(\mathbf{x}) = \frac{1}{1-\nu}\Bigg\{&\left[\frac{5\nu-1}{15}(1-\rho_0^3) - \frac{2(7-10\nu t^2)}{5(7+5\nu)}\rho_0^3(1-\rho_0^2)\right]\delta_{ij}\delta_{mn} \\ &+ \left[\frac{1-\nu}{2} + \frac{5\nu-7}{30}(1-\rho_0^3) - \frac{7(5t^2-3)-20\nu t^2}{5(7+5\nu)}\rho_0^3(1 - \rho_0^2)\right](\delta_{im}\delta_{jn}+\delta_{in}\delta_{jm}) + \frac{2t^2(7-10\nu)}{7+5\nu}\rho_0^3(1-\rho^2)\delta_{ij}\bar{x}_m\bar{x}_n \\ &+ \frac{8\nu t^2}{7+5\nu}\rho^3(1-\rho_0^2)\delta_{mn}\bar{x}_i\bar{x}_j - \frac{6\nu t^2}{7+5\nu}\rho_0^3(1-\rho_0^2)(\delta_{im}\bar{x}_j\bar{x}_n \\ &+ \delta_{in}\bar{x}_j\bar{x}_m + \delta_{jm}\bar{x}_i\bar{x}_n + \delta_{jn}\bar{x}_i\bar{x}_m)\Bigg\}\end{aligned} \tag{109}$$

and the exterior solution is

$$\begin{aligned}\mathbb{S}^{E,N}_{ijmn}(\mathbf{x}) = \frac{\rho_0^3}{1-\nu}\Bigg\{&\left[\frac{3\rho^2/t^2+10\nu-5}{30t^3} + \frac{5\nu-1}{15} - \frac{2(7-10\nu t^2)}{5(7+5\nu)}(1-\rho_0^2)\right]\delta_{ij}\delta_{mn} \\ &+ \left[\frac{3\rho_0^2/t^2-10\nu+5}{30t^3} + \frac{7-5\nu}{30} - \frac{7(5t^2-3)-20\nu t^2}{5(7+5\nu)}(1-\rho_0^2)\right](\delta_{im}\delta_{jn}+\delta_{in}\delta_{jm}) \\ &- \left[\frac{\rho_0^2/t^2-1}{2t^3} - \frac{2t^2(7-10\nu)}{7+5\nu}(1-\rho_0^2)\right]\delta_{ij}\bar{x}_m\bar{x}_n - \left[\frac{\rho_0^2/t^2+2\nu-1}{2t^3} - \frac{8\nu t^2}{7+5\nu}(1-\rho_0^2)\right]\delta_{mn}\bar{x}_i\bar{x}_j \\ &- \left[\frac{\rho_0^2/t^2-\nu}{2t^3} + \frac{6\nu t^2}{7+5\nu}(1-\rho_0^2)\right](\delta_{im}\bar{x}_j\bar{x}_n + \delta_{in}\bar{x}_j\bar{x}_m + \delta_{jm}\bar{x}_i\bar{x}_n + \delta_{jn}\bar{x}_i\bar{x}_m) + \frac{7\rho_0^2/t^2-5}{2t^3}\bar{x}_i\bar{x}_j\bar{x}_m\bar{x}_n\Bigg\}\end{aligned} \tag{110}$$

Figure 4 shows a comparison of the Neumann–Eshelby tensor with the original Eshelby tensor for $\rho_0=0.4$ and $\nu=0.3$. Here we display the six coefficients of the radial basis arrays of the finite Eshelby tensors, $\mathbf{S}^{\cdot,N}$ and the original Eshelby tensors, $\mathbf{S}^{\cdot,\infty}$. One can see that there are significant differences in the first three coefficients.

A display of the displacement bases, $\mathbf{U}^{\cdot,N}$, $\mathbf{U}^{\cdot,\infty}$, and $\mathbf{U}^{B,N}$ for $\nu=0.3$, is shown in Figs. 3(*b*), 3(*d*), and 3(*f*). One can observe that the difference between the Neumann and the original solution is large, even though the volume fraction is only $\rho_0^3=0.4^3=0.064$. Figure 3 also illustrates different characters of the Dirichlet and the Neumann solution.

Remark 4.1. The volumetric part of the disturbance strain ϵ^d_{ij} is related to the volumetric part of the eigenstrain ϵ^*_{ij} by a scalar coefficient

$$\mathbb{S}^{\cdot,\star}_{iijj}(\mathbf{x}) = \mathbf{\Theta}_{iijj}\mathbf{S}^{\cdot,\star}(t)$$

with

$$\mathbf{\Theta}_{iijj} = [9,6,3,3,4,1]^T \tag{111}$$

From this, one can find some interesting relationships of the finite Eshelby tensors. First we have

$$\mathbb{S}^{\cdot,\star}_{iijj}(\mathbf{x}) = 9\mathbf{S}^{\cdot,\star}_1 + 6\mathbf{S}^{\cdot,\star}_2 + 3\mathbf{S}^{\cdot,\star}_3 + 3\mathbf{S}^{\cdot,\star}_4 + 4\mathbf{S}^{\cdot,\star}_5 + \mathbf{S}^{\cdot,\star}_6 = \text{const.}\ \forall\ \mathbf{x} \in \Omega \tag{112}$$

This implies that even though the finite Eshelby tensors derived here are functions of the position vector $\mathbf{x}$, the dilatational part of the Eshelby tensor is a constant. In particular, the following dilatational contractions have elementary forms,

$$\mathbb{S}^{I,D}_{iijj} = \frac{(1-f)(1+\nu)}{1-\nu}, \quad \mathbb{S}^{E,D}_{iijj} = -\frac{f(1+\nu)}{1-\nu} \tag{113}$$

$$\mathbb{S}^{I,N}_{iijj} = \frac{(1+\nu)+2f(1-2\nu)}{1-\nu}, \quad \mathbb{S}^{E,N}_{iijj} = \frac{2f(1-2\nu)}{1-\nu} \tag{114}$$

where $f=\rho_0^3$ is the volume fraction of the inclusion phase. Second it is interesting to note that the Dirichlet and the Neumann Eshelby tensors follow the ordering

$$\mathbb{S}^{I,N}_{iijj} \geqslant \mathbb{S}^{I,D}_{iijj}, \quad \mathbb{S}^{E,N}_{iijj} \geqslant \mathbb{S}^{E,D}_{iijj}, \quad \text{“ = ” holds iff } f=0 \tag{115}$$

and that the difference between interior and exterior solution is

$$\mathbb{S}^{I,D}_{iijj} - \mathbb{S}^{E,D}_{iijj} = \mathbb{S}^{I,N}_{iijj} - \mathbb{S}^{E,N}_{iijj} = \frac{1+\nu}{1-\nu} = \mathbb{S}^{I,\infty}_{iijj} \tag{116}$$

In classical theory, the dilatational eigenstrain has some special properties, e.g., the dilatational eigenstrain due to a dilating inclusion is constant. It appears that some of these properties are still preserved in the finite spherical inclusion solution. This not only validates the present theory, but also indicates that the present theory may have some important applications, because dilatational eigenstrains are usually associated with, for example, thermal expansion, lattice mismatch in quantum dots, and misfit strain in phase transformation.

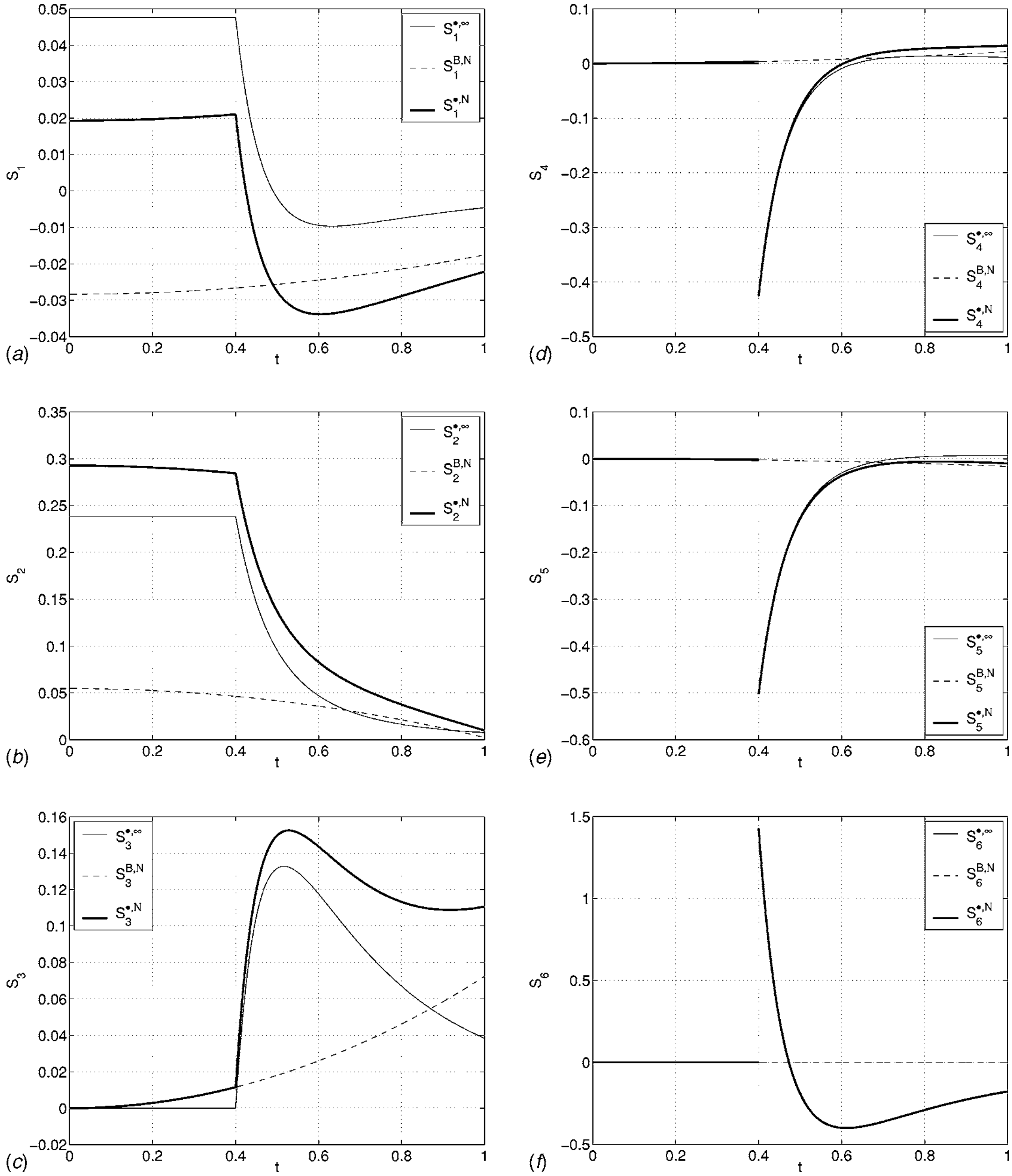


Fig. 4 The components of the radial basis arrays $\mathbf{S}^{\cdot,\infty}$, $\mathbf{S}^{B,N}$, and $\mathbf{S}^{\cdot,N}$

4.3 Traction Distributions. Next, we examine the radial projection of the disturbance stress field. The physical meaning of such a stress projection field is a set of parametric traction fields on the surfaces of successive concentric spheres. Any point, $\mathbf{x}$, inside the spherical RVE lies on a spherical surface whose normal $\bar{\mathbf{x}}$ is along the direction of the position vector $\mathbf{x}$. Thus the parametric traction field is defined as

$$t_i^d(\mathbf{x}) = \sigma_{ji}^d(\mathbf{x})\bar{x}_j(\mathbf{x}) \tag{117}$$

which can be expressed in terms of the eigenstrain

$$t_i^d(\mathbf{x}) = \begin{cases} \bar{x}_j(\mathbf{x})\mathbb{C}_{ijk\ell}[\mathbb{S}^{I,\star}_{k\ell mn}(\mathbf{x}) - \mathbb{I}^s_{k\ell mn}]\epsilon^*_{mn}, & \forall \mathbf{x} \in \Omega_I \\ \bar{x}_j(\mathbf{x})\mathbb{C}_{ijk\ell}\mathbb{S}^{E,\star}_{k\ell mn}(\mathbf{x})\epsilon^*_{mn}, & \forall \mathbf{x} \in \Omega_E \end{cases} \tag{118}$$

Here $\mathbb{I}^s_{k\ell mn}$ is the fourth-order symmetric identity tensor, which also falls into our definition of a fourth-order radial isotropic tensor, i.e.,

$$\mathbb{I}^s_{k\ell mn} = \mathbf{\Theta}^T_{k\ell mn}(\mathbf{r})\mathbf{I}^s \tag{119}$$

where $\mathbf{I}^s = [0, 1/2, 0, 0, 0, 0]^T$. One may then rewrite Eq. (118) as

$$t_i^d(\mathbf{x}) = \begin{cases} \bar{x}_j(\mathbf{x})\mathbb{C}_{ijk\ell}\mathbf{\Theta}^T_{k\ell mn}(\bar{\mathbf{x}})[\mathbf{S}^{I,\star}(t) - \mathbf{I}^s]\epsilon^*_{mn}, & \forall \mathbf{x} \in \Omega_I \\ \bar{x}_j(\mathbf{x})\mathbb{C}_{ijk\ell}\mathbf{\Theta}^T_{k\ell mn}(\bar{\mathbf{x}})\mathbf{S}^{E,\star}(t)\epsilon^*_{mn}, & \forall \mathbf{x} \in \Omega_E \end{cases} \tag{120}$$

In analogy to the displacement field (see Eq. (43)) the disturbance traction can also be written as

$$t_i^d(\mathbf{x}) = \mathbf{\Xi}^T_{imn}(\bar{\mathbf{x}})\mathbf{T}^{\cdot,\star}(t)\epsilon^*_{mn} \tag{121}$$

where $\mathbf{T}^{\cdot,\star}$ is the radial basis array of the traction field and $\mathbf{\Xi}_{imn}$ is given by Eq. (42). The preceding two equations establishes a relation between the arrays $\mathbf{T}^{\cdot,\star}$ and $\mathbf{S}^{\cdot,\star}$. We find that

$$\mathbf{T}^{I,\star}(t) = \mathbf{K}_1[\mathbf{S}^{I,\star}(t) - \mathbf{I}^s] \tag{122}$$

$$\mathbf{T}^{E,\star}(t) = \mathbf{K}_1\mathbf{S}^{E,\star}(t) \tag{123}$$

where $\mathbf{K}_1$ is given by Eq. (54). In view of Eqs. (71) and (72) we can write

$$\mathbf{T}^{I,\star}(t) = \mathbf{T}^{I,\infty}(t) + \mathbf{T}^{B,\star}(t) - \mathbf{T}^{*}(t), \quad 0 \leqslant t < \rho_0 \tag{124}$$

$$\mathbf{T}^{E,\star}(t) = \mathbf{T}^{E,\infty}(t) + \mathbf{T}^{B,\star}(t), \quad \rho_0 \leqslant t \leqslant 1 \tag{125}$$

where the individual pieces are as follows. Corresponding to the original Eshelby problem we have

$$\mathbf{T}^{I,\infty}(t) = \mathbf{K}_1\mathbf{S}^{I,\infty}(t) = \frac{2\mu}{15(1-\nu)}\begin{bmatrix} (1-12\nu+5\nu^2)/(2\nu-1) \\ 4-5\nu \\ 0 \end{bmatrix} \tag{126}$$

$$\mathbf{T}^{E,\infty}(t) = \mathbf{K}_1\mathbf{S}^{E,\infty}(t) = \frac{\mu\rho_0^3/t^3}{15(1-\nu)}\begin{bmatrix} -12\rho_0^2/t^2 + 10(1-\nu) \\ -12\rho_0^2/t^2 + 5(1+\nu) \\ 60(\rho_0^2/t^2 - 1) \end{bmatrix} \tag{127}$$

and the Dirichlet and Neumann boundary contributions are

$$\mathbf{T}^{B,D}(t) = \mathbf{K}_1\mathbf{S}^{B,D}(t) = -\frac{2\mu\rho_0^3}{15(1-\nu)}\begin{bmatrix} (1-12\nu+5\nu^2)/(2\nu-1) \\ 4-5\nu \\ 0 \end{bmatrix} + \frac{\mu\rho_0^3(1-\rho_0^2)}{10(1-\nu)(7-10\nu)}\begin{bmatrix} 2(7+5\nu t^2) \\ 7(5t^2-3)+10\nu t^2 \\ 10t^2(7-5\nu) \end{bmatrix} \tag{128}$$

$$\mathbf{T}^{B,N}(t) = \mathbf{K}_1\mathbf{S}^{B,N}(t) = -\frac{\mu\rho_0^3}{15(1-\nu)}\begin{bmatrix} 2(1+5\nu) \\ 7-5\nu \\ 0 \end{bmatrix} - \frac{2\mu\rho_0^3(1-\rho_0^2)}{5(1-\nu)(7+5\nu)}\begin{bmatrix} 2(7+5\nu t^2) \\ 7(5t^2-3)+10\nu t^2 \\ 10t^2(7-5\nu) \end{bmatrix} \tag{129}$$

The final contribution, arising from the eigenstrains, is

$$\mathbf{T}^{*}(t) = \mathbf{K}_1\mathbf{I}^{s} = \begin{bmatrix} 2\nu/(1-2\nu) \\ 1 \\ 0 \end{bmatrix} \tag{130}$$

It is readily verified that for $t=1$ the traction basis corresponding to the Neumann–Eshelby problem is

$$\mathbf{T}^{E,N}(1) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{131}$$

which assures $t_i(\mathbf{x}) \equiv 0$ for $\forall \mathbf{x} \in \partial\Omega$. Therefore the prescribed Neumann boundary condition is indeed satisfied by the solution presented. This fact can also be clearly observed in Fig. 5, which shows the three components of $\mathbf{T}^{\cdot,\infty}$, $\mathbf{T}^{\cdot,\star}$ and $\mathbf{T}^{B,\star}$ for both the Dirichlet problem (a),(c),(e), and the Neumann problem (b),(d),(f). Here we choose $\rho_0=0.4$ and $\nu=0.3$. We observe that the components of $\mathbf{T}^{E,N}(t)$ go to zero at the boundary of the RVE ($t=1$). It can also be seen that the boundary corrections $\mathbf{T}^{B,D}$ and $\mathbf{T}^{B,N}$ are substantial even though the volume fraction is small, i.e., $\rho_0^3=0.064$.

5 Closure

In this paper, the elastic fields due to a spherical inclusion subjected to prescribed eigenstrains and embedded in a finite spherical RVE are studied. On the outer surface of the RVE, uniform boundary conditions are prescribed, which are either a displacement (Dirichlet) boundary condition or a prescribed traction (Neumann) boundary condition.

The notion of a *radial isotropic tensor* is introduced, which is a generalization of the isotropic tensor. It has been argued that if a spherical inclusion is placed concentrically within a spherical RVE, the finite Eshelby tensors, which map the prescribed eigenstrain to the disturbance strain field, are radial isotropic tensors. In other words, the tensorial circumference basis for the finite Eshelby tensors is the same as the basis for the Eshelby tensors in unbounded space.

By utilizing this property, we have solved a pair of Fredholm type integral equations, and we have obtained, for the first time, the exact, closed form solutions for both the interior and exterior Eshelby tensors for an inclusion in a finite, three-dimensional RVE. It has been revealed that the finite Eshelby tensors depend on both the location and the volume fraction of the inclusion, which accurately captures both the size effect of the inclusion and the boundary image contribution to the original inclusion problem.

One of advantages of the present solution procedure is that it circumvents the use of a finite Green's function. As a matter of fact, the solution of Green's function of Navier's equation for a finite spherical domain is a more difficult problem, which is still open. On the other hand, we hope that this work may shed some light on the search for the finite Green's function, however, we believe, not without some added difficulties. We further note that, by using our solution technique, one may be able to extend the present solution to the elliptical inclusion problem in a finite domain. The difficulty then will be how to find the symmetry group of the circumference basis of the elliptical geometry, which has to be also invariant under the integral equation that involves the boundary integrals Eqs. (51) and (85).

We also would like to mention that the spherical RVE may be subjected to general boundary conditions. Nevertheless, the two fundamental solutions corresponding to the Dirichlet and the Neumann boundary conditions form a basis for the finite Eshelby tensors under a general boundary value problem. This issue will be further discussed in detail in a separate paper [16]. It should also be pointed out that even though the two basic finite Eshelby tensors obtained here are the solutions of the homogeneous inclusion problems, they are two fundamental elements for the finite Eshelby tensors of a general RVE with more complex microstructures. By using superposition, they can be readily used to construct the solutions for the n-inclusion ($n \geqslant 2$) problem, and they can be used to solve various homogenization problems as well as the problem of inhomogeneity induced elastic fields in a finite spherical domain.

To illustrate such applications, in the second part of this work [11], we apply the finite Eshelby tensors to evaluate the effective material properties of composites. It has been shown that the method employing the finite Eshelby tensor provides remarkably accurate predictions in simple homogenization procedures. Furthermore they furnish new variational bounds, and lead to a new class of general homogenization methods.

Acknowledgment

This work is supported by a grant from National Science Foundation (Grant No. CMS-0239130), which is greatly appreciated.

Appendix: Integration Formulas

In this Appendix the solution of the fourteen integrals listed in Eqs. (55)–(61) and (90)–(96) is given. The procedure is similar to the two dimensional case reported in Li et al. [17] and Wang et al. [18]. Considering $\mathbf{x} + r\bar{\mathbf{r}} = \mathbf{y}$, where $\mathbf{y} \in \partial\Omega$, we have

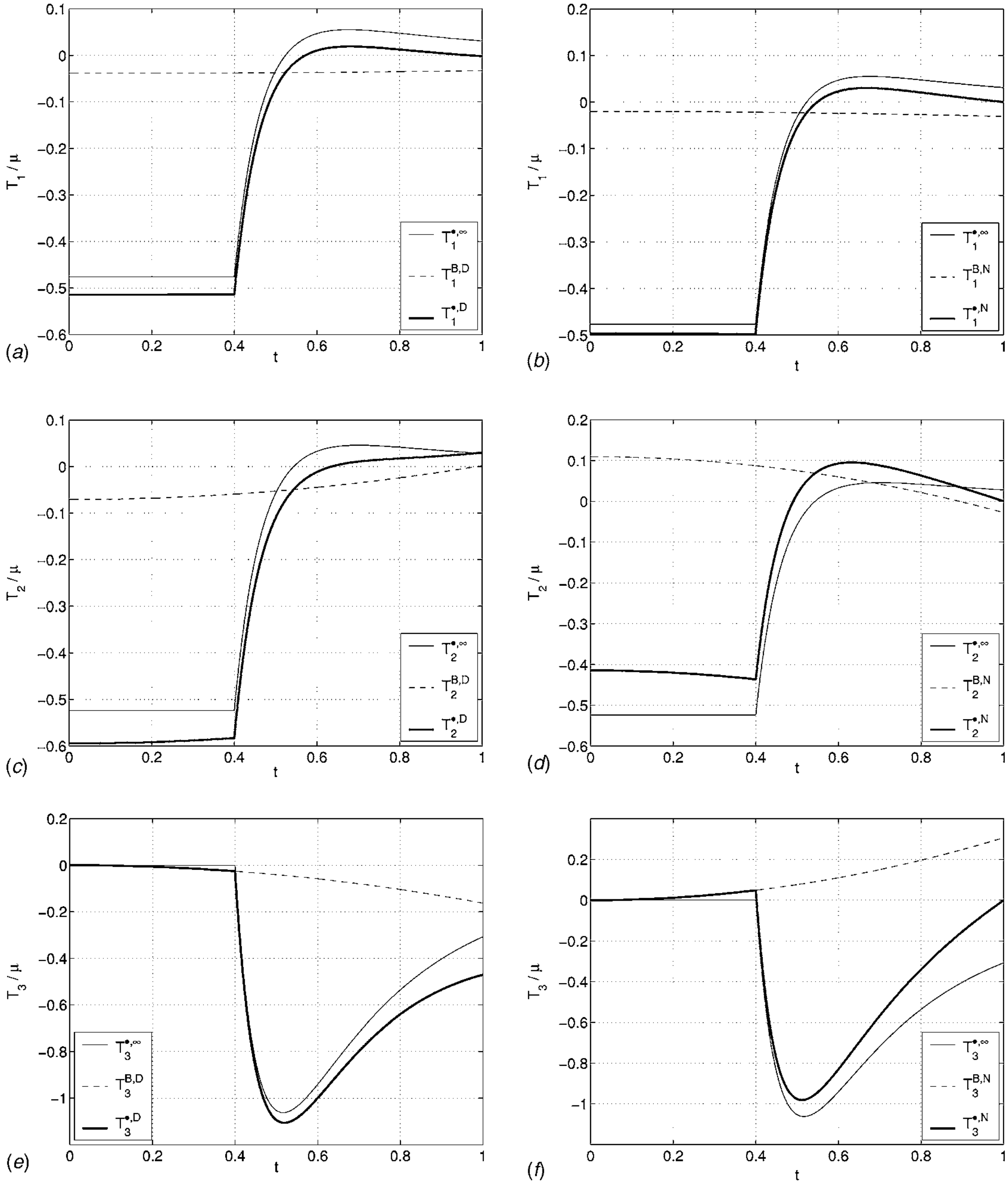


Fig. 5 The components of the radial basis arrays $\mathbf{T}^{\cdot,\infty}$, $\mathbf{T}^{B,D}$, $\mathbf{T}^{\cdot,D}$, $\mathbf{T}^{B,N}$, and $\mathbf{T}^{\cdot,N}$

$$\bar{r}_i = \frac{A}{r}(n_i - t\bar{x}_i) \tag{A1}$$

or

$$n_i = \frac{r}{A}\bar{r}_i + t\bar{x}_i.$$

Recall that $t=|\mathbf{x}|/A$. The relations defined in Eq. (A1) are illustrated in both Figs. 1 and 6(a).

The surface integration over the RVE is performed w.r.t. the surface of a unit sphere, S_2, centered at point $\mathbf{x}$. According to Fig. 6, we define a new basis $\hat{\mathbf{e}}_i$ at $\mathbf{x}$ such that $\hat{\mathbf{e}}_3=\bar{\mathbf{x}}$ Vector $\bar{\mathbf{r}}$ is then described by the spherical coordinates φ and θ, i.e., $\bar{\mathbf{r}}=(\cos\varphi\sin\theta\sin\varphi\sin\theta\cos\theta)^T$.

Denote dS as the surface element of $\partial\Omega$ (the outer surface of the RVE). The projection of dS to the perpendicular direction of $\bar{\mathbf{r}}$ is denoted by $d\hat{S}$, and is given by $d\hat{S}=r^2\sin\theta\, d\theta\, d\varphi$. It is related to dS by

$$dS = \frac{d\hat{S}}{\cos\psi} = \frac{r^2}{\cos\psi}\sin\theta d\theta\, d\varphi = \frac{r^2}{\cos\psi}dS_2 \tag{A2}$$

where $dS_2=\sin\theta\, d\theta\, d\varphi$ is the surface element on the unit sphere S_2. Considering the shaded triangle ($\mathbf{0xy}$) in Fig. 6, we find that

$$\frac{A}{r} = \frac{1}{\sqrt{1-2t\cos\phi+t^2}} \tag{A3}$$

and

$$\cos\psi = \sqrt{1-t^2\sin^2\theta} \tag{A4}$$

Furthermore from $y_i y_i=A^2$, one can derive the relation

$$r = A(-t\cos\theta + \sqrt{1-t^2\sin^2\theta}) \tag{A5}$$

Figure 6(b) shows that for every point P on the surface of the unit sphere there exists a point P^* such that $\bar{\mathbf{r}}(P)=-\bar{\mathbf{r}}(P^*)$. Thus any function, $\mathcal{L}^o(\bar{\mathbf{r}})=\bar{r}_i, \bar{r}_i\bar{r}_j, \bar{r}_i\bar{r}_j\bar{r}_m, \ldots$, which is odd in $\bar{\mathbf{r}}$, satisfies

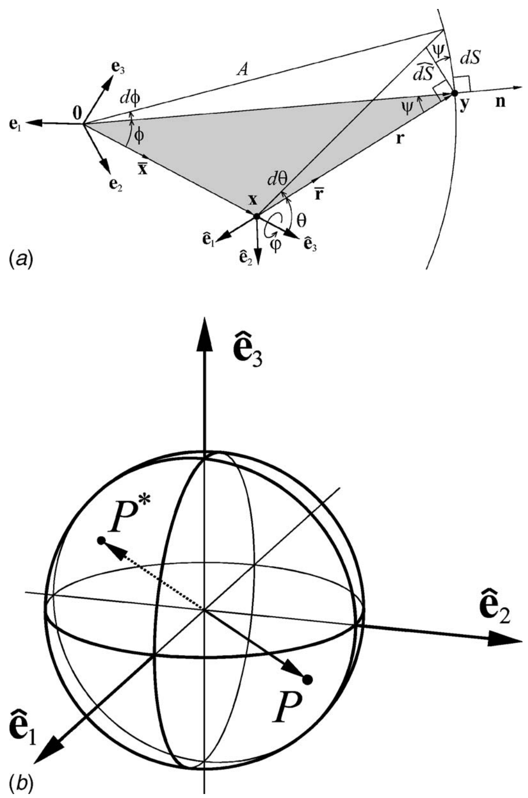


Fig. 6 (*a*) Relation between *dS*, *dϕ*, and *dθ*; and (*b*) unit sphere.

$\mathcal{L}^o[\bar{\mathbf{r}}(P)]=-\mathcal{L}^o[\bar{\mathbf{r}}(P^*)]$, and therefore the integration of an odd function of $\bar{\mathbf{r}}$, i.e., $\mathcal{L}^o(\bar{\mathbf{r}})$, over the surface of the sphere will be zero. In particular, by applying Eqs. (A2) and (A4), we find that

$$\int_{\partial\Omega}\frac{\mathcal{L}^o(\bar{\mathbf{r}})}{r^2}dS=\int_{S_2}A\frac{\mathcal{L}^o(\bar{\mathbf{r}})}{\sqrt{1-t^2\sin^2\theta}}dS_2=0 \qquad \text{(A6)}$$

Note that $\sin^2\theta$ is an even function in $\bar{\mathbf{r}}$, i.e., $\sin^2\theta(\bar{\mathbf{r}})=\sin^2\theta(-\bar{\mathbf{r}})$. Further, we denote an even function of $\bar{\mathbf{r}}$ as $\mathcal{L}^e(\bar{\mathbf{r}})$, if $\mathcal{L}^e[\bar{\mathbf{r}}(P)]=\mathcal{L}^e[\bar{\mathbf{r}}(P^*)]$. Then, by virtue of Eqs. (A3)–(A5), it follows that

$$\int_{\partial\Omega}\frac{\mathcal{L}^e(\bar{\mathbf{r}})}{r}dS=\int_{S_2}A\left(1-\frac{t\cos\theta}{\sqrt{1-t^2\sin^2\theta}}\right)\mathcal{L}^e(\bar{\mathbf{r}})dS_2$$

$$=A\int_{S_2}\mathcal{L}^e(\bar{\mathbf{r}})dS_2 \qquad \text{(A7)}$$

because $\cos\theta$ is an odd function in $\bar{\mathbf{r}}$. Using Eqs. (A7) and (A6) we obtain the following seven elemental integrals

$$(1)\quad \int_{\partial\Omega}\frac{1}{r}dS_y=4\pi A \qquad \text{(A8)}$$

$$(2)\quad \int_{\partial\Omega}\frac{\bar{r}_i}{r^2}dS_y=0 \qquad \text{(A9)}$$

$$(3)\quad \int_{\partial\Omega}\frac{\bar{r}_i\bar{r}_j}{r}dS_y=A\int_{S_2}\bar{r}_i\bar{r}_j\,dS_u=\frac{4\pi}{3}A\delta_{ij} \qquad \text{(A10)}$$

$$(4)\quad \int_{\partial\Omega}\frac{\bar{r}_i\bar{r}_j\bar{r}_m}{r^2}dS_y=0 \qquad \text{(A11)}$$

$$(5)\quad \int_{\partial\Omega}\frac{\bar{r}_i\bar{r}_j\bar{r}_m\bar{r}_n}{r}dS_y=A\int_{S_2}\bar{r}_i\bar{r}_j\bar{r}_m\bar{r}_n\,dS_u=\frac{4\pi}{15}A(\delta_{ij}\delta_{mn}+\delta_{im}\delta_{jn}$$

$$+\delta_{in}\delta_{jm}) \qquad \text{(A12)}$$

$$(6)\quad \int_{\partial\Omega}\frac{\bar{r}_i\bar{r}_j\bar{r}_m\bar{r}_n\bar{r}_r}{r^2}dS_y=0 \qquad \text{(A13)}$$

$$(7)\quad \int_{\partial\Omega}\frac{\bar{r}_i\bar{r}_j\bar{r}_m\bar{r}_n\bar{r}_r\bar{r}_s}{r}dS_y=A\int_{S_2}\bar{r}_i\bar{r}_j\bar{r}_m\bar{r}_n\bar{r}_r\bar{r}_s\,dS_u=\frac{4\pi}{105}A(\delta_{ij}\delta_{mn}\delta_{rs}$$

$$+\delta_{im}\delta_{jn}\delta_{rs}+\delta_{in}\delta_{jm}\delta_{rs}+\delta_{ir}\delta_{mn}\delta_{js}$$
$$+\delta_{is}\delta_{mn}\delta_{jr}\delta_{ij}\delta_{mr}\delta_{ns}+\delta_{im}\delta_{jr}\delta_{ns}$$
$$+\delta_{in}\delta_{jr}\delta_{ms}+\delta_{ir}\delta_{mj}\delta_{ns}$$
$$+\delta_{is}\delta_{mj}\delta_{nr}\delta_{ij}\delta_{ms}\delta_{nr}+\delta_{im}\delta_{js}\delta_{nr}$$
$$+\delta_{in}\delta_{js}\delta_{mr}+\delta_{ir}\delta_{ms}\delta_{nj}+\delta_{is}\delta_{mr}\delta_{nj}) \qquad \text{(A14)}$$

Using these seven elemental integrals and Eq. (A1) we obtain all the integrals listed in Eqs. (55)–(61) and (90)–(96).

References

[1] Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems," Proc. R. Soc. London, Ser. A, **241**, pp. 376–396.
[2] Eshelby, J. D., 1959, "The Elastic Field Outside an Ellipsoidal Inclusion," Proc. R. Soc. London, Ser. A, **252**, pp. 561–569.
[3] Eshelby, J. D., 1961, "Elastic Inclusions and Inhomogeneities," *Progress in Solid Mechanics*, N. I. Sneddon and R. Hill, eds., North Holland, Amsterdam, The Netherlands, Vol. 2, pp. 89–104.
[4] Thostenson, E. T., Ren, Z., and Chou, T. W., 2001, "Advances in Science and Technology of Carbon Nanotubes and their Composites: A Review," Compos. Sci. Technol., **61**, pp. 1899–1912.
[5] Shi, D.-L., Feng, X.-Q., Huang, Y.-Y., Hwang, K.-C., and Gao, H., 2004, "The Effect of Nanotube Waviness and Agglomeration on the Elastic Property of Carbon Nanotube-Reinforced Composites," J. Eng. Mater. Technol., **126**, pp. 250–257.
[6] Kinoshita, N., and Mura, T., 1984, "Eigenstrain Problems in a Finite Elastic Body," SIAM J. Appl. Math., **44**, pp. 524–535.
[7] Kröner, E., 1986, "Statistical Modeling," *Modeling Small Deformation of Polycrystals*, J. Gittus and J. Zarka, eds., Elsevier Applied Science, New York, pp. 229–291.
[8] Kröner, E., 1990, "Modified Green's Function in the Theory of Heterogeneous and/or Anisotropic Linearly Elastic Media," *Micromechanics and Inhomogeneity, The Toshio Mura 65th Anniversary Volume*, G. J. Weng, M. Taya, and H. Abe, eds., Springer, New York, pp. 599–622.
[9] Mazilu, P., 1972, "On the Theory of Linear Elasticity in Statically Homogeneous Media," Rev. Roum. Math. Pures Appl., **17**, pp. 261–273.
[10] Luo, H. A., and Weng, G. J., 1987, "On Eshelby's Inclusion Problem in a Three-Phase Spherically Concentric Solid and a Modification of Mori-Tanaka's Method," Mech. Mater., **6**, pp. 347–361.
[11] Li, S., Wang, G., and Sauer, R. A., 2007, "The Eshelby Tensors in a Finite Spherical Domain—Part II: Applications to Homogenization," ASME J. Appl. Mech., **74**, pp. 784–797.
[12] Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*, 4th ed., Cambridge University Press, Cambridge, UK.
[13] Mura, T., 1987, *Micromechanics of Defects in Solids*, 2nd ed., Martinus Nijhoff, Boston.
[14] Ju, J. W., and Sun, L. Z., 1999, "A Novel Formulation for the Exterior Point Eshelby's Tensor of an Ellipsoidal Inclusion," ASME J. Appl. Mech., **66**, pp. 570–574.
[15] Walpole, L. J., 1981, "Elastic Behavior of Composite Materials: Theoretical Foundations," *Advances in Applied Mechanics*, C. S. Yih, ed., Academic, New York, Vol. 21, pp. 169–242.
[16] Sauer, R. A., Wang, G., and Li, S., 2007, "The Composite Eshelby Tensors and Their Application to Homogenization," submitted.
[17] Li, S., Sauer, R. A., and Wang, G., 2005, "Circular Inclusion in a Finite Domain, I: Direchlet-Eshelby Problem," Acta Mech., **179**, pp. 67–90.
[18] Wang, G., Li, S., and Sauer, R. A., 2005, "Circular Inclusion in a Finite Domain, II: Neumann-Eshelby Problem," Acta Mech., **179**, pp. 91–110.

The Eshelby Tensors in a Finite Spherical Domain—Part II: Applications to Homogenization

Shaofan Li[1]
e-mail: li@ce.berkeley.edu

Gang Wang

Roger A. Sauer

Department of Civil and Environmental Engineering,
University of California,
Berkeley, CA 94720

In this part of the work, the Eshelby tensors of a finite spherical domain are applied to various homogenization procedures estimating the effective material properties of multiphase composites. The Eshelby tensors of a finite domain can capture the boundary effect of a representative volume element as well as the size effect of the different phases. Therefore their application to homogenization does not only improve the accuracy of classical homogenization methods, but also leads to some novel homogenization theories. This paper highlights a few of them: a refined dilute suspension method and a modified Mori–Tanaka method, the exterior eigenstrain method, the dual-eigenstrain method, which is a generalized self-consistency method, a shell model, and new variational bounds depending on the different boundary conditions. To the best of the authors' knowledge, this is the first time that a multishell model is used to evaluate the Hashin–Shtrikman bounds for a multiple phase composite ($n \geqslant 3$), which can distinguish some of the subtleties of different microstructures. [DOI: 10.1115/1.2711228]

1 Introduction

In the first part of this work [1], which is referred to as Part I hereafter, the exact solutions of the elastic fields of a spherical inclusion embedded in a finite spherical representative volume (RVE) are obtained under both the prescribed displacement (Dirichlet) boundary condition and the prescribed traction (Neumann) boundary condition.

For simplicity, we refer to the Dirichlet– and Neumann–Eshelby tensors of a finite domain as the finite Eshelby tensors. A salient feature of the finite Eshelby tensors is their ability to capture both the boundary effect, or image force effect, of an RVE and the size effect, i.e., the dependency on the volume fraction of the different phases of a composite. This offers great advantages and flexibilities in homogenization procedures, which is the focus of this second part of our work. Using the new finite Eshelby tensors we can modify the classical homogenization schemes and obtain some remarkable results. Furthermore several new homogenization schemes can be constructed by the application of the finite Eshelby tensors.

In recent years, nanocomposites have emerged as promising materials for future technologies e.g., Refs. [2,3], because of their high strength, excellent conductivity in both heat transfer and electricity. Considerable attention has been devoted to study the interfacial strength, size effects, and agglomeration effects of nanocomposites (e.g. Fisher et al. [4], Odegard et al. [5], Shi et al. [6], and Sharma and Ganti [7]). The classical homogenization techniques have shown limitations to deal with the above issues. There is a call for a refined micromechanics theory for nanocomposites, e.g, Ref. [8]. One of the objectives of this research is towards establishing a refined micromechanics homogenization theory for nanocomposites.

We proceed, in the following section, by deriving expressions for the average finite Eshelby tensors in a RVE. These are needed to characterize the average disturbance fields, which have some important properties. In Sec. 3 we re-examine two conventional homogenization methods by using the average finite Eshelby tensor. Further, in Sec. 4, we discuss the so-called dual eigenstrain method, which is a combination of an exterior and interior eigenstrain homogenization method. This scheme is a generalized self-consistency method, which leads to a new class of predictor–corrector schemes. In Sec. 5, a shell model is proposed to capture microstructure effects on the homogenization of a multiphase composite. Finally, in Sec. 6, the Hashin–Shtrikman (HS) variational bounds are rederived using the finite Eshelby tensors to incorporate the boundary conditions. A multishell model is used to evaluate the exact HS bounds for a multiphase composite with $n \geqslant 3$ using a multivariable optimization procedure. Conclusions are drawn in Sec. 7.

2 Average Eshelby Tensors and Average Disturbance Fields

In Part I we derived the finite Eshelby tensors, $\mathbb{S}^{\cdot,D}$ and $\mathbb{S}^{\cdot,N}$, which are valid for a spherical inclusion Ω_I embedded at the center of a finite, spherical RVE Ω (see Fig. 1 of Part I). In accordance with Part I we adopt the following nomenclature to describe the problem: The radii of inclusion and RVE are denoted by a and A, their ratio by $\rho_0 = a/A$. Any point $\mathbf{x}$ inside the RVE can be written as $\mathbf{x} = tA\bar{\mathbf{x}}$, where $t = |\mathbf{x}|/A$ and $\bar{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$ denote the normalized radial distance and direction of $\mathbf{x}$. The elasticity tensors of the two domains Ω_I and Ω_E are denoted by $\mathbb{C}^I$ and $\mathbb{C}^E = \mathbb{C}$.

For clarity, we first derive the expression of the average finite Eshelby tensors and discuss their relation with the average disturbance strain field.

2.1 Average Eshelby Tensors.

The spatial averaging operator is defined as

$$\langle \dots \rangle_\Omega = \frac{1}{|\Omega|} \int_\Omega \dots d\Omega \tag{1}$$

where $|\Omega|$ denotes the volume of the spatial domain Ω. Due to the radial isotropic structure of the finite Eshelby tensors, $\mathbb{S}^{\cdot,\star}_{ijmn}(\mathbf{x}) = \mathbf{\Theta}^T_{ijmn}(\bar{\mathbf{x}})\mathbf{S}^{\cdot,\star}(t)$ $(\cdot = I, E;\ \star = D, N)$, their average over the RVE domain Ω can be written as

[1] Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received April 6, 2006; final manuscript received June 14, 2006. Review conducted by Subrata Mukherjee. Paper presented at the 2005 ASME International Mechanical Engineering Congress (IMECE2005), Orlando, FL, USA, November 5–11, 2005.

$$\langle \mathbb{S}^{\cdot,\star}_{ijmn}(\mathbf{x})\rangle_{\Omega} = \frac{1}{|\Omega|}\int_{\Omega} \mathbf{S}^{\cdot,\star}(t)\cdot \mathbf{\Theta}_{ijmn}(\bar{\mathbf{x}})d\Omega = \langle 3t^2\mathbf{S}^{\cdot,\star}\rangle_{[0,1]}\cdot\langle \mathbf{\Theta}_{ijmn}\rangle_{S_2} \tag{2}$$

where S_2 denotes the surface of a sphere with unit radius, and where

$$\langle \mathbf{\Theta}_{ijmn}\rangle_{S_2} := \frac{1}{4\pi}\int_{S_2} \mathbf{\Theta}_{ijmn}\, dS_2 \tag{3}$$

$$\langle 3t^2\mathbf{S}^{\cdot,\star}\rangle_{[a,b]} := \frac{1}{b^3-a^3}\int_a^b 3t^2\mathbf{S}^{\cdot,\star}\, dt \tag{4}$$

The above decomposition is possible since $\mathbf{S}$ is independent of the orientation $\bar{\mathbf{x}}$ and $\mathbf{\Theta}_{ijmn}$ is independent of the radial distance t. Performing the averaging of $\mathbf{\Theta}_{ijmn}$ over the unit sphere yields

$$\langle \mathbf{\Theta}_{ijmn}\rangle_{S_2} = \frac{1}{4\pi}\int_{S_2}\mathbf{\Theta}_{ijmn}\, dS_2 = \begin{bmatrix} 3\mathbb{E}^{(1)}_{ijmn} \\ 2\mathbb{E}^{(1)}_{ijmn} + 2\mathbb{E}^{(2)}_{ijmn} \\ \mathbb{E}^{(1)}_{ijmn} \\ \mathbb{E}^{(1)}_{ijmn} \\ \frac{4}{3}\mathbb{E}^{(1)}_{ijmn} + \frac{4}{3}\mathbb{E}^{(2)}_{ijmn} \\ \frac{1}{3}\mathbb{E}^{(1)}_{ijmn} + \frac{2}{15}\mathbb{E}^{(2)}_{ijmn} \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 2 & 2 \\ 1 & 0 \\ 1 & 0 \\ \frac{4}{3} & \frac{4}{3} \\ \frac{1}{3} & \frac{2}{15} \end{bmatrix} \times \begin{bmatrix} \mathbb{E}^{(1)}_{ijmn} \\ \mathbb{E}^{(2)}_{ijmn} \end{bmatrix} \tag{5}$$

where $\mathbb{E}^{(1)}_{ijmn}$ and $\mathbb{E}^{(2)}_{ijmn}$ are the following isotropic basis tensors

$$\mathbb{E}^{(1)}_{ijmn} = \tfrac{1}{3}\delta_{ij}\delta_{mn}, \quad \mathbb{E}^{(2)}_{ijmn} = \tfrac{1}{2}(\delta_{im}\delta_{jn} + \delta_{in}\delta_{jm}) - \tfrac{1}{3}\delta_{ij}\delta_{mn} \tag{6}$$

For each boundary condition (Dirichlet or Neumann), we have two Eshelby tensors, interior $\mathbb{S}^{I,\star}(\mathbf{x})$ for $\mathbf{x}\in\Omega_I$, or exterior $\mathbb{S}^{E,\star}(\mathbf{x})$ for $\mathbf{x}\in\Omega/\Omega_I := \Omega_E$. Their average over the respective domains follows as

$$\langle \mathbb{S}^{I,\star}_{ijmn}\rangle_{\Omega_I} = \langle 3t^2\mathbf{S}^I\rangle_{[0,\rho_0]}\cdot\langle\mathbf{\Theta}_{ijmn}\rangle_{S_2},$$

$$\langle \mathbb{S}^{E,\star}_{ijmn}\rangle_{\Omega_E} = \langle 3t^2\mathbf{S}^E\rangle_{[\rho_0,1]}\cdot\langle\mathbf{\Theta}_{ijmn}\rangle_{S_2} \tag{7}$$

Since the averaging of $\mathbf{\Theta}_{ijmn}$ over S_2 is an isotropic tensor, we obtain

$$\langle \mathbb{S}^{I,\star}_{ijmn}\rangle_{\Omega_I} = s_1^{I,\star}\mathbb{E}^{(1)}_{ijmn} + s_2^{I,\star}\mathbb{E}^{(2)}_{ijmn} \tag{8}$$

$$\langle \mathbb{S}^{E,\star}_{ijmn}\rangle_{\Omega_E} = s_1^{E,\star}\mathbb{E}^{(1)}_{ijmn} + s_2^{E,\star}\mathbb{E}^{(2)}_{ijmn} \tag{9}$$

The coefficients $s_1^{I,\star}$, $s_2^{I,\star}$ and $s_1^{E,\star}$, $s_2^{E,\star}$ depend on the volume fraction $f := \rho_0^3$ and are given as

$$s_1^{I,D} = \frac{(1+\nu)(1-f)}{3(1-\nu)}, \quad s_2^{I,D} = \frac{2(4-5\nu)(1-f)}{15(1-\nu)} - 21\gamma_u[f](1-f^{2/3}) \tag{10}$$

$$s_1^{E,D} = -\frac{(1+\nu)f}{3(1-\nu)}, \quad s_2^{E,D} = -\frac{2(4-5\nu)f}{15(1-\nu)} + 21\gamma_u[f]f\frac{1-f^{2/3}}{1-f} \tag{11}$$

for the Dirichlet boundary condition (BC) and

$$s_1^{I,N} = \frac{1+\nu+2(1-2\nu)f}{3(1-\nu)},$$

$$s_2^{I,N} = \frac{2(4-5\nu)+(7-5\nu)f}{15(1-\nu)} + 21\gamma_t[f](1-f^{2/3}) \tag{12}$$

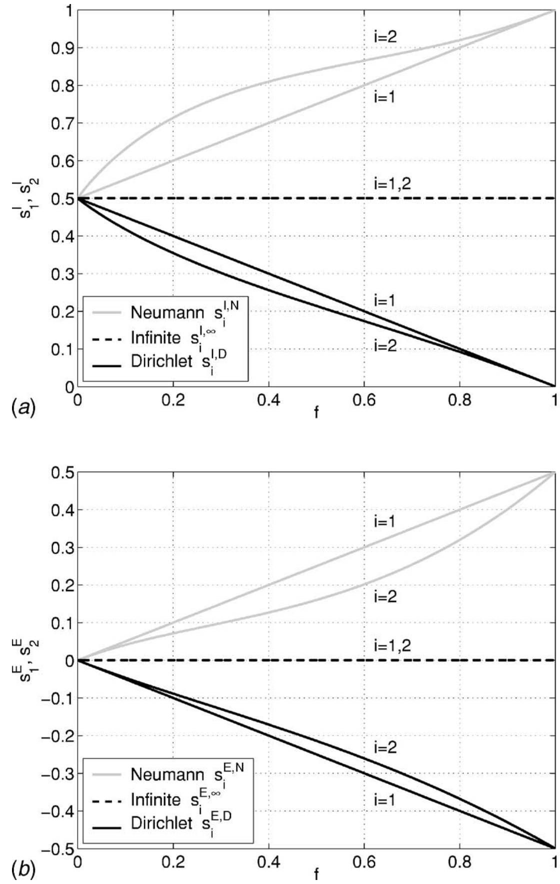


Fig. 1 Average Eshelby tensor coefficients s_1^I, s_1^E ($i=1$) and s_2^I, s_2^E ($i=2$)

$$s_1^{E,N} = \frac{2(1-2\nu)f}{3(1-\nu)}, \quad s_2^{E,N} = \frac{(7-5\nu)f}{15(1-\nu)} - 21\gamma_t[f]f\frac{1-f^{2/3}}{1-f} \tag{13}$$

for the Neumann BC. Here we have denoted

$$\gamma_u[f] := \frac{f(1-f^{2/3})}{10(1-\nu)(7-10\nu)}, \quad \gamma_t[f] := \frac{4f(1-f^{2/3})}{10(1-\nu)(7+5\nu)} \tag{14}$$

In fact, Eqs. (10)–(13) are the precise formulas of the size-effect characterization of of the inclusion problem. One can find that this effect is linear for the bulk modulus, whereas it is nonlinear in the shear modulus. In contrast to the average finite Eshelby tensors we recall the average Eshelby tensor for a spherical inclusion in an unbounded medium

$$\langle \mathbb{S}^{\cdot,\infty}_{ijmn}\rangle_{\Omega_\cdot} = s_1^{\cdot,\infty}\mathbb{E}^{(1)}_{ijmn} + s_2^{\cdot,\infty}\mathbb{E}^{(2)}_{ijmn}, \quad \cdot = I \text{ or } E \tag{15}$$

$$s_1^{I,\infty} = \frac{1+\nu}{3(1-\nu)}, \quad s_2^{I,\infty} = \frac{2(4-5\nu)}{15(1-\nu)} \tag{16}$$

$$s_1^{E,\infty} = 0, \quad s_2^{E,\infty} = 0 \tag{17}$$

Figure 1 displays the behaviors of all the coefficients $s_i^{\cdot,\star}$ in dependence of f. The Poisson's ratio is chosen as $\nu=0.2$. We observe that for the Dirichlet case the coefficients decrease, while for the Neumann case they increase with growing f. The classical Eshelby tensors do not depend on f.

Note that when $f\to 0$ in Eqs. (10)–(13) we recover the expressions for the average of the classical Eshelby tensors. The fact that $s_i^{E,\infty}=0$ implies the well-known Tanaka–Mori Lemma (see below). Let us define the difference $\Delta s_i^\star = s_i^{I,\star} - s_i^{E,\star}$; we have

$$\Delta s_1^D = \Delta s_1^N = \Delta s_1^\infty = \frac{1+\nu}{3(1-\nu)} \tag{18}$$

$$\Delta s_2^D = \Delta s_2^\infty - 21\gamma_u[f]\frac{1-f^{2/3}}{1-f}, \quad \Delta s_2^\infty = \frac{2(4-5\nu)}{15(1-\nu)} \tag{19}$$

$$\Delta s_2^N = \Delta s_2^\infty + 21\gamma_t[f]\frac{1-f^{2/3}}{1-f} \tag{20}$$

2.2 The Average Disturbance Fields. The finite Eshelby tensors can be conveniently used to represent the average disturbance fields. Recall the classical Tanaka–Mori Lemma [9]: the exterior average disturbance strain in the exterior domain is zero (see Eq. (17))

$$\langle \boldsymbol{\epsilon}^d \rangle_{\Omega_E} = \langle \mathbb{S}^{E,\infty} \rangle_{\Omega_E} : \boldsymbol{\epsilon}^* = (s_1^{E,\infty}\mathbb{E}^{(1)} + s_2^{E,\infty}\mathbb{E}^{(2)}) : \boldsymbol{\epsilon}^* = 0 \tag{21}$$

A similar result holds for the disturbance stress field for a linear elastic medium. Using the new finite Eshelby tensors $\mathbb{S}^D$ and $\mathbb{S}^N$ the original Tanaka–Mori Lemma result is modified. The exterior average disturbance strain field is neither zero for the Dirichlet problem

$$\langle \boldsymbol{\epsilon}^d \rangle_{\Omega_E} = \langle \mathbb{S}^{E,D} \rangle_{\Omega_E} : \boldsymbol{\epsilon}^* = (s_1^{E,D}\mathbb{E}^{(1)} + s_2^{E,D}\mathbb{E}^{(2)}) : \boldsymbol{\epsilon}^* \neq 0 \tag{22}$$

nor is it zero for the Neumann problem

$$\langle \boldsymbol{\epsilon}^d \rangle_{\Omega_E} = \langle \mathbb{S}^{E,N} \rangle_{\Omega_E} : \boldsymbol{\epsilon}^* = (s_1^{E,N}\mathbb{E}^{(1)} + s_2^{E,N}\mathbb{E}^{(2)}) : \boldsymbol{\epsilon}^* \neq 0 \tag{23}$$

(unless $f=0$). However, in view of Eqs. (11) and (13), we find that for both problems

$$\langle \boldsymbol{\epsilon}^d \rangle_{\Omega_E} = \mathcal{O}(f) \tag{24}$$

which can be viewed as a modified Tanaka–Mori Lemma. One then recovers the original result as $f \to 0$.

Moreover, consider the Dirichlet problem. We can exactly satisfy a key assumption, the average strain theorem

$$\langle \boldsymbol{\epsilon} \rangle_\Omega = \langle \boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d \rangle_\Omega = \boldsymbol{\epsilon}^0 + \langle \boldsymbol{\epsilon}^d \rangle_\Omega = \boldsymbol{\epsilon}^0 \tag{25}$$

since the average disturbance strain field in Ω is zero

$$\langle \boldsymbol{\epsilon}^d \rangle_\Omega = f\langle \boldsymbol{\epsilon}^d \rangle_{\Omega_I} + (1-f)\langle \boldsymbol{\epsilon}^d \rangle_{\Omega_E} = [f\langle \mathbb{S}^{I,D} \rangle_{\Omega_I} + (1-f)\langle \mathbb{S}^{E,D} \rangle_{\Omega_E}] : \boldsymbol{\epsilon}^* = 0 \tag{26}$$

Likewise, for the Neumann problem, the average stress theorem

$$\langle \boldsymbol{\sigma} \rangle_\Omega = \boldsymbol{\sigma}^0 \tag{27}$$

is exactly satisfied since $\langle \boldsymbol{\sigma}^d \rangle_\Omega = f\langle \boldsymbol{\sigma}^d \rangle_{\Omega_I} + (1-f)\langle \boldsymbol{\sigma}^d \rangle_{\Omega_E} = 0$ due to

$$f\langle \mathbb{T}^{I,N} \rangle_{\Omega_I} + (1-f)\langle \mathbb{T}^{E,N} \rangle_{\Omega_E} = \mathbb{O} \tag{28}$$

where $\mathbb{T}^{I,N}$ and $\mathbb{T}^{E,N}$ are the conjugate Neumann Eshelby tensors related to the Neumann Eshelby tensors by the expressions

$$\langle \mathbb{S}^{I,N} \rangle_{\Omega_I} + \langle \mathbb{T}^{I,N} \rangle_{\Omega_I} = \mathbb{I}^s$$

and

$$\langle \mathbb{S}^{E,N} \rangle_{\Omega_E} + \langle \mathbb{T}^{E,N} \rangle_{\Omega_E} = \mathbb{O} \tag{29}$$

where $\mathbb{I}^s$ is the fourth-order symmetric unit tensor and $\mathbb{O}$ is the fourth-order null tensor.

3 Improvement of the Classical Homogenization Methods

We now use the finite Eshelby tensors in two classical homogenization procedures to estimate effective material properties, namely, the homogenization for composites with dilute suspension and the Mori–Tanaka model.

3.1 Dilute Suspension Model. The dilute suspension method predicts two different effective elastic tensors depending on the different boundary conditions e.g., Ref. [10]. We first consider the prescribed macrostrain BC, i.e., the Dirichlet boundary value problem (BVP) ($\mathbf{u}^d=0$ on $\partial\Omega$), as discussed in Part I. The average stress consistency condition for the considered homogenization scheme (for prescribed eigenstrain within Ω_I as motivated in Part I) is

$$\mathbb{C}^I : (\boldsymbol{\epsilon}^0 + \langle \boldsymbol{\epsilon}^d \rangle_{\Omega_I}) = \mathbb{C} : (\boldsymbol{\epsilon}^0 + \langle \boldsymbol{\epsilon}^d \rangle_{\Omega_I} - \boldsymbol{\epsilon}^*), \quad \forall \in \Omega_I \tag{30}$$

Note that $\mathbb{C}^I$, $\mathbb{C}$, $\boldsymbol{\epsilon}^0$, and $\boldsymbol{\epsilon}^*$ are considered constant. From Eq. (30) we obtain

$$\boldsymbol{\epsilon}^0 + \langle \boldsymbol{\epsilon}^d \rangle_{\Omega_I} = \mathbb{A} : \boldsymbol{\epsilon}^* \tag{31}$$

where $\mathbb{A} := (\mathbb{C} - \mathbb{C}^I)^{-1} : \mathbb{C}$. Consider the interior average of the disturbance strain field

$$\langle \boldsymbol{\epsilon}^d \rangle_{\Omega_I} = \langle \mathbb{S}^{I,D} \rangle_{\Omega_I} : \boldsymbol{\epsilon}^* \tag{32}$$

and substitute Eq. (32) into Eq. (31). This yields

$$\boldsymbol{\epsilon}^* = [\mathbb{A} - \langle \mathbb{S}^{I,D} \rangle_{\Omega_I}]^{-1} : \boldsymbol{\epsilon}^0 \tag{33}$$

and consequently,

$$\langle \boldsymbol{\epsilon} \rangle_{\Omega_I} = \boldsymbol{\epsilon}^0 + \langle \boldsymbol{\epsilon}^d \rangle_{\Omega_I} = \mathbb{A} : [\mathbb{A} - \langle \mathbb{S}^{I,D} \rangle_{\Omega_I}]^{-1} : \boldsymbol{\epsilon}^0 \tag{34}$$

Following the standard procedure, e.g., Ref. [10], we find the estimate of the effective elasticity tensor for the prescribed macrostrain BC

$$\bar{\mathbb{C}} = \mathbb{C} - f\mathbb{C} : (\mathbb{A} - \langle \mathbb{S}^{I,D} \rangle_{\Omega_I})^{-1} \tag{35}$$

The only difference between Eq. (35) and the classical solution for dilute suspension is that a different Eshelby tensor is used. Considering isotropic materials, the effective bulk and shear moduli become

$$\bar{\kappa} = \kappa - f\kappa\left(\frac{1}{1-\kappa^I/\kappa} - s_1^{I,D}\right)^{-1}, \quad \bar{\mu} = \mu - f\mu\left(\frac{1}{1-\mu^I/\mu} - s_2^{I,D}\right)^{-1} \tag{36}$$

For the prescribed macrostress boundary condition, the new dilute suspension estimate is

$$\bar{\mathbb{D}} = \mathbb{D} + f\mathbb{D} : (\mathbb{A} - \langle \mathbb{S}^{I,N} \rangle_{\Omega_I})^{-1} \tag{37}$$

where $\langle \mathbb{S}^{I,N} \rangle_{\Omega_I}$ is the interior average Neumann–Eshelby tensor. For isotropic composites, the corresponding effective bulk and shear moduli are

$$\bar{\kappa}^{-1} = \kappa^{-1} + f\kappa^{-1}\left(\frac{1}{1-\kappa^I/\kappa} - s_1^{I,N}\right)^{-1},$$

$$\bar{\mu}^{-1} = \mu^{-1} + f\mu^{-1}\left(\frac{1}{1-\mu^I/\mu} - s_2^{I,N}\right)^{-1} \tag{38}$$

Figure 2 shows the curves of the normalized bulk modulus, $\bar{\kappa}/\kappa$, and shear modulus, $\bar{\mu}/\mu$, in dependence of the volume fraction f of the inclusion. The material properties of the inclusion are chosen as $\kappa^I/\kappa=10$, $\mu^I/\mu=4$, with $\nu=0.1$. We have plotted the result Eq. (35) using the Dirichlet–Eshelby tensor $\mathbb{S}^{I,D}$ (dark) and (37) using the Neumann–Eshelby tensor $\mathbb{S}^{I,N}$ (light). We compare the new results with the conventional dilute suspension results using the infinite Eshelby tensor $\mathbb{S}^{I,\infty}$ in Eq. (35) (dashed line 2) and in Eq. (37) (dashed line 1).

From this figure, we can observe the well-known result that the classical solution is not self-consistent, i.e $\bar{\mathbb{D}} \neq \bar{\mathbb{C}}^{-1}$. When we use the new finite Eshelby tensors this situation is significantly improved. For the effective bulk modulus, the new scheme is self consistent, i.e., the two $\bar{\kappa}$ in Eqs. (36) and (38) are equal. The estimated effective shear modulus is not self consistent, but it is quite close as shown in Fig. 2.

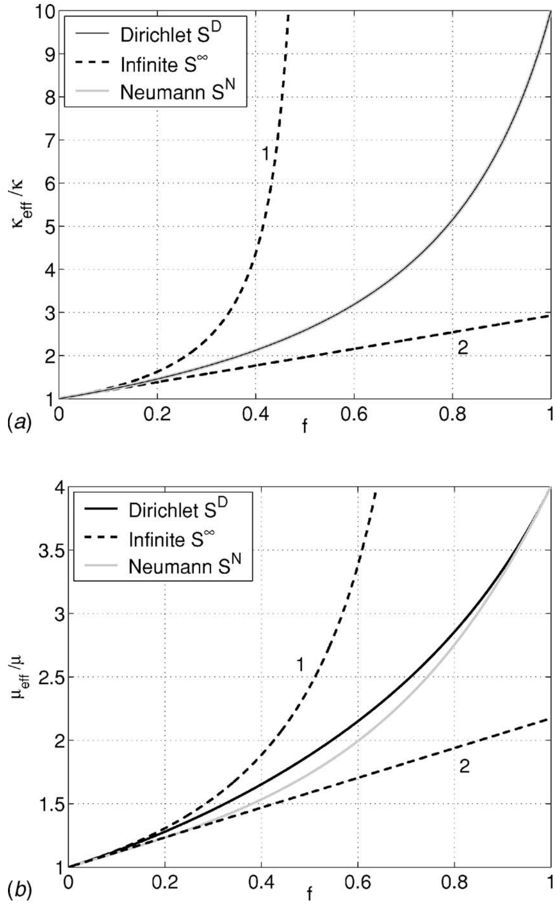


Fig. 2 Effective moduli $\bar{\kappa}$, $\bar{\mu}$ (or κ_{eff} and μ_{eff}) obtained by using the dilute suspension method

3.2 A Refined Mori–Tanaka Model. The original Mori and Tanaka model [11] is derived for an infinite RVE. In the following, we rederive the Mori–Tanaka estimate for a two-phase composite in a finite RVE.

In reality, the boundary condition of an RVE is neither a prescribed displacement boundary condition nor is it a prescribed traction boundary condition. One can thus define a "general finite Eshelby tensor" as the linear combination of the Dirichlet–Eshelby tensor and the Neumann–Eshelby tensor corresponding to general boundary conditions

$$\mathbb{S}^{\cdot,F} = \alpha \mathbb{S}^{\cdot,D} + (1-\alpha)\mathbb{S}^{\cdot,N}, \quad \cdot = I, \text{ or } E \tag{39}$$

For detailed justification, derivation, and discussion of this concept, readers are referred to Ref. [12].

The essence of the Mori–Tanaka procedure is the following incremental homogenization procedure. Let us denote the current background strain of the RVE as $\langle \boldsymbol{\epsilon}^b \rangle_\Omega$, which may or may not be the average strain of the RVE. Adding an inclusion (or a cluster of inclusions represented by a single inclusion) into the RVE, the new average strain $\langle \boldsymbol{\epsilon} \rangle$ in each phase will be the sum of the background strain and the disturbance strain

$$\langle \boldsymbol{\epsilon} \rangle_{\Omega_I} = \langle \boldsymbol{\epsilon}^b \rangle_\Omega + \langle \boldsymbol{\epsilon}^d \rangle_{\Omega_I} \tag{40}$$

$$\langle \boldsymbol{\epsilon} \rangle_{\Omega_E} = \langle \boldsymbol{\epsilon}^b \rangle_\Omega + \langle \boldsymbol{\epsilon}^d \rangle_{\Omega_E} \tag{41}$$

The classical Tanaka–Mori Lemma states that $\langle \boldsymbol{\epsilon}^d \rangle_{\Omega_E} = 0$. This is only true when the RVE is infinite, since $\langle \mathbb{S}^{E,\infty} \rangle_{\Omega_E} = 0$. For a finite RVE, we have to take into account the change of the effective material properties in the matrix

$$\langle \boldsymbol{\epsilon} \rangle_{\Omega_I} = \langle \boldsymbol{\epsilon}^b \rangle_\Omega + \langle \mathbb{S}^{I,F} \rangle_{\Omega_I} : \boldsymbol{\epsilon}^* \tag{42}$$

$$\langle \boldsymbol{\epsilon} \rangle_{\Omega_E} = \langle \boldsymbol{\epsilon}^b \rangle_\Omega + \langle \mathbb{S}^{E,F} \rangle_{\Omega_E} : \boldsymbol{\epsilon}^* \tag{43}$$

Consider the average stress consistency condition (for $\mathbf{x} \in \Omega_I$)

$$\mathbb{C}^I : \langle \boldsymbol{\epsilon} \rangle_{\Omega_I} = \mathbb{C} : (\langle \boldsymbol{\epsilon} \rangle_{\Omega_I} - \boldsymbol{\epsilon}^*) \tag{44}$$

Solving Eqs. (44) for $\langle \boldsymbol{\epsilon} \rangle_{\Omega_I}$ yields

$$\langle \boldsymbol{\epsilon} \rangle_{\Omega_I} = \mathbb{A} : \boldsymbol{\epsilon}^* \tag{45}$$

where $\mathbb{A} = (\mathbb{C} - \mathbb{C}^I)^{-1} : \mathbb{C}$. Considering Eq. (42), we can express the eigenstrain in terms of the background strain as

$$\boldsymbol{\epsilon}^* = [\mathbb{A} - \langle \mathbb{S}^{I,F} \rangle_{\Omega_I}]^{-1} : \langle \boldsymbol{\epsilon}^b \rangle_\Omega \tag{46}$$

Considering the basic average equation of the strain

$$\langle \boldsymbol{\epsilon} \rangle_\Omega = f \langle \boldsymbol{\epsilon} \rangle_{\Omega_I} + (1-f) \langle \boldsymbol{\epsilon} \rangle_{\Omega_E} \tag{47}$$

and substituting Eqs. (42), (43), and (46) into Eq. (47), we can express the average strain $\langle \boldsymbol{\epsilon} \rangle_\Omega$ in terms of the background strain as

$$\langle \boldsymbol{\epsilon} \rangle_\Omega = \mathcal{A}^F : \langle \boldsymbol{\epsilon}^b \rangle_\Omega \tag{48}$$

Here $\mathcal{A}^F$ is the concentration tensor defined as

$$\mathcal{A}^F = [\mathbb{A} - (1-f)(\langle \mathbb{S}^{I,F} \rangle_{\Omega_I} - \langle \mathbb{S}^{E,F} \rangle_{\Omega_E})] : (\mathbb{A} - \langle \mathbb{S}^{I,F} \rangle_{\Omega_I})^{-1} \tag{49}$$

By virtue of Eqs. (45) and (46), the average stress field inside the inclusion can now be written as

$$\langle \boldsymbol{\sigma} \rangle_{\Omega_I} = \mathbb{C} : (\mathbb{A} - \mathbb{I}^{(4s)}) : (\mathbb{A} - \langle \mathbb{S}^{I,F} \rangle_{\Omega_I})^{-1} : \langle \boldsymbol{\epsilon}^b \rangle_\Omega \tag{50}$$

Applying the basic equation for mixture to the stress field,

$$\langle \boldsymbol{\sigma} \rangle_\Omega = f \langle \boldsymbol{\sigma} \rangle_{\Omega_I} + (1-f) \langle \boldsymbol{\sigma} \rangle_{\Omega_E} \tag{51}$$

and substituting Eqs. (42), (43), and (46) into Eq. (51), we can express the average stress $\langle \boldsymbol{\sigma} \rangle_\Omega$ in terms of the background strain as

$$\langle \boldsymbol{\sigma} \rangle_\Omega = \mathcal{B}^F : \langle \boldsymbol{\epsilon}^b \rangle_\Omega \tag{52}$$

with

$$\mathcal{B}^F = \mathbb{C} : [\mathbb{A} - f\mathbb{I}^{(4s)} - (1-f)(\langle \mathbb{S}^{I,F} \rangle_{\Omega_I} - \langle \mathbb{S}^{E,F} \rangle_{\Omega_E})] : (\mathbb{A} - \langle \mathbb{S}^{I,F} \rangle_{\Omega_I})^{-1} \tag{53}$$

Finally from $\langle \boldsymbol{\sigma} \rangle_\Omega = \bar{\mathbb{C}} : \langle \boldsymbol{\epsilon} \rangle_\Omega$, we obtain the effective elastic tensor

$$\bar{\mathbb{C}} = \mathbb{C} - f\mathbb{C} : [\mathbb{A} - (1-f)(\langle \mathbb{S}^{I,F} \rangle_{\Omega_I} - \langle \mathbb{S}^{E,F} \rangle_{\Omega_E})]^{-1}, \quad \bar{\mathbb{D}} = \bar{\mathbb{C}}^{-1} \tag{54}$$

We note in passing that this model is self consistent.

The homogenization procedure with finite Eshelby tensors, $\mathbb{S}^{I,F}$ and $\mathbb{S}^{E,F}$, furnishes a refined Mori–Tanaka model. For isotropic two-phase composites, the corresponding formulas are

$$\bar{\kappa} = \kappa - f\kappa \left[\frac{1}{1 - \kappa^I/\kappa} - (1-f)\Delta s_1^F \right]^{-1},$$

$$\bar{\mu} = \mu - f\mu \left[\frac{1}{1 - \mu^I/\mu} - (1-f)\Delta s_2^F \right]^{-1} \tag{55}$$

Note that the differences $\Delta s_1^F = s_1^{I,F} - s_1^{E,F}$ and $\Delta s_2^F = s_2^{I,F} - s_2^{E,F}$ are given by Eqs. (18)–(20).

Figure 3 displays the profiles of the normalized effective moduli $\bar{\kappa}/\kappa$ and $\bar{\mu}/\mu$ over the volume fraction of the second phase. The same material data is used for the results shown in Fig. 2.

In the case of the bulk modulus the dark, dashed and the light curves match exactly, i.e., they are the same analytically. Indeed, $\bar{\kappa}$ in Eq. (55) is mathematically identical when applying $\mathbb{S}^{\cdot,\infty}$, $\mathbb{S}^{\cdot,D}$,

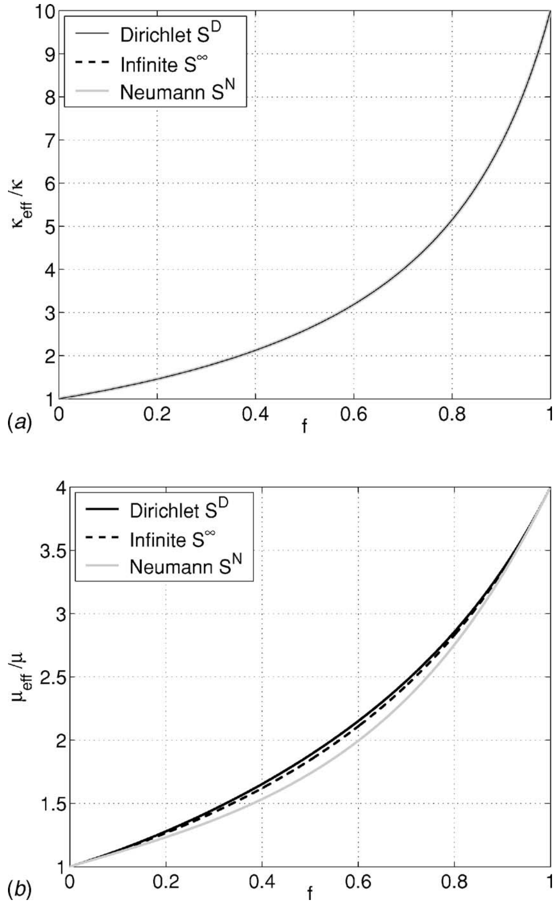


Fig. 3 Effective moduli $\bar{\kappa}$, $\bar{\mu}$ (or κ_{eff} and μ_{eff}) obtained by using the Mori–Tanaka method

or $\mathbb{S}^{\cdot,N}$, since $\Delta s_1^{\infty}=\Delta s_1^{D}=\Delta s_1^{N}$ as noted in Eq. (18). For the shear modulus $\bar{\mu}$ Eq. (55) gives three distinct lines when applying $\mathbb{S}^{\cdot,\infty}$, $\mathbb{S}^{\cdot,D}$, or $\mathbb{S}^{\cdot,N}$.

Remarkably, when comparing Figs. 2 and 3, we find that the dark and light lines match exactly. In other words, it can be shown that by using the finite Eshelby tensors, the dilute suspension method Eqs. (35) and (37) is equivalent to the Mori–Tanaka method Eq. (54), when using the corresponding $\mathbb{S}^{\cdot,D}$ and $\mathbb{S}^{\cdot,N}$. The finite Eshelby tensors $\mathbb{S}^{\cdot,D}$ and $\mathbb{S}^{\cdot,N}$ unify the previously distinct homogenization methods.

4 Exterior/Interior Eigenstrain Method

In the classical eigenstrain homogenization method, since the ambient space (i.e., matrix phase) is assumed to be unbounded, the eigenstrain can only be prescribed inside the inclusion. Therefore the stress or strain consistency condition, i.e., the equivalent eigenstrain principle, is only applicable to the interior. Considering the eigenstrain to be prescribed in the interior domain is the case we have considered so far.

For a finite RVE, the equivalent eigenstrain principle can be equally applied to its interior (inclusion phase) or exterior region (matrix phase). By treating the interior and exterior homogenization scheme with equal footing, one may be able to characterize certain patterns of the phase distribution in an RVE in addition to merely considering the volume fraction of the phases. Such patterns may be the concentration of inhomogeneities towards the center or boundary of the RVE. The exterior eigenstrain method has been studied before by Castles and Mura [13], however, without the knowledge of the finite Eshelby tensor. In this section, we first discuss the exterior eigenstrain method, which relies on the

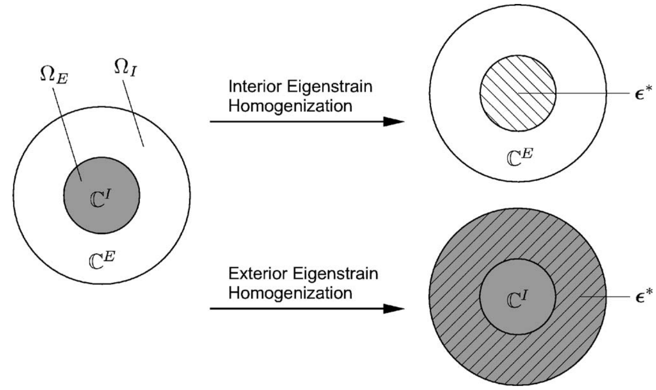


Fig. 4 Illustration of interior and exterior eigenstrain method

interior eigenstrain results previously obtained. Second we introduce a method which considers the simultaneous prescription of interior and exterior eigenstrains.

4.1 Exterior Eigenstrain Method. Analogously to the interior eigenstrain method, the idea of the exterior eigenstrain method is as follows. We choose the interior phase, characterized by elasticity $\mathbb{C}^I$, as the comparison solid of the RVE. To account for the difference of elastic properties, a uniform eigenstrain is prescribed in the *exterior* region of the RVE, i.e.

$$\boldsymbol{\epsilon}^*(\mathbf{x}) = \begin{cases} 0, & \forall \mathbf{x} \in \Omega_I \\ \boldsymbol{\epsilon}^*, & \forall \mathbf{x} \in \Omega_E \end{cases} \tag{56}$$

The concept is illustrated in Fig. 4.

It follows that the constitutive relation between the disturbance stress and strain fields has the form

$$\boldsymbol{\sigma}^d(\mathbf{x}) = \mathbb{C}^I:[\boldsymbol{\epsilon}^d(\mathbf{x}) - \boldsymbol{\epsilon}^*(\mathbf{x})], \quad \forall\, \mathbf{x} \in \Omega \tag{57}$$

Accordingly, Somigliana's identity reads

$$u_m^d(\mathbf{x}) = -\int_{\Omega_E} \mathbb{C}^I_{ijk\ell} G^{\infty}_{im,j}(\mathbf{x}-\mathbf{y}) d\Omega_y \varepsilon^*_{k\ell} + \int_{\partial\Omega} \mathbb{C}^I_{ijk\ell} u_k^d(\mathbf{y}) G^{\infty}_{im,j}(\mathbf{x} - \mathbf{y}) n_\ell(\mathbf{y}) dS_y + \int_{\partial\Omega} \mathbb{C}^I_{ijk\ell} u^d_{k,\ell}(\mathbf{y}) n_j(\mathbf{y}) G^{\infty}_{im}(\mathbf{x}-\mathbf{y}) dS_y \tag{58}$$

By considering either the Dirichlet or Neumann boundary condition prescribed on the RVE boundary, the above equation can be solved to relate the disturbance strain field to the prescribed *exterior* eigenstrain through the so called *exterior Eshelby tensors* denoted by $\bar{\mathbb{S}}^{\cdot,\star}$. They are defined from

$$\boldsymbol{\epsilon}^d(\mathbf{x}) = \begin{cases} \bar{\mathbb{S}}^{E,\star}(\mathbf{x}):\boldsymbol{\epsilon}^*, & \forall \mathbf{x} \in \Omega_E \\ \bar{\mathbb{S}}^{I,\star}(\mathbf{x}):\boldsymbol{\epsilon}^*, & \forall \mathbf{x} \in \Omega_I \end{cases} \tag{59}$$

where the superscripts $\cdot = I$ or E denote the tensors evaluated at the interior or exterior regions, and where $\star = D$ or N stands for either the Dirichlet (prescribed displacement) or Neumann boundary (prescribed traction) condition.

The disturbance fields in Eq. (58) can be solved exactly by means of superposition. Since $\Omega_E = \Omega / \Omega_I$ the resulting exterior eigenstrain Eshelby tensors can be written as a combination of the interior eigenstrain Eshelby tensors, which have been solved in Part I, as

$$\bar{\mathbb{S}}^{E,\star} = \mathbb{S}^{I,\star}(\mathbb{C}^I, f=1) - \mathbb{S}^{E,\star}(\mathbb{C}^I, f_I), \quad \mathbf{x} \in \Omega_E \tag{60}$$

$$\bar{\mathbb{S}}^{I,\star} = \mathbb{S}^{I,\star}(\mathbb{C}^I, f=1) - \mathbb{S}^{I,\star}(\mathbb{C}^I, f_I), \quad \mathbf{x} \in \Omega_I \tag{61}$$

We emphasize that for this case (the exterior eigenstrain method) the Eshelby tensor $\mathbb{S}^{\cdot,\star}$ in the above equations takes the material property $\mathbb{C}^I$ of the inclusion (the comparison solid), and the volume fraction f_I of the inclusion phase. Given the boundary condition $\star = D$ or N, the first term of the equations above can be easily evaluated (see Part I for the expressions of $\mathbb{S}^{I,\star}$)

$$\mathbb{S}^{I,D}(\mathbb{C}^I, f=1) = \mathbb{O}$$

and

$$\mathbb{S}^{I,N}(\mathbb{C}^I, f=1) = \mathbb{I}^s \tag{62}$$

where $\mathbb{O}$ and $\mathbb{I}^s$ are the fourth-order zero tensor and identity tensor, respectively. Therefore the exterior eigenstrain Eshelby tensors can be related explicitly to the interior eigenstrain Eshelby tensors by

$$\bar{\mathbb{S}}^{\cdot,D} = -\mathbb{S}^{\cdot,D}(\mathbb{C}^I, f_I), \quad \mathbf{x} \in \Omega \tag{63}$$

$$\bar{\mathbb{S}}^{\cdot,N} = \mathbb{I}^s - \mathbb{S}^{\cdot,N}(\mathbb{C}^I, f_I), \quad \mathbf{x} \in \Omega \tag{64}$$

To proceed further, the average of the exterior eigenstrain Eshelby tensors $\bar{\mathbb{S}}^{\cdot,\star}$ is needed. Following Sec. 2, we find

$$\langle \bar{\mathbb{S}}^{I,\star}_{ijmn} \rangle_{\Omega_I} = \bar{s}_1^{I,\star} \mathbb{E}^{(1)}_{ijmn} + \bar{s}_2^{I,\star} \mathbb{E}^{(2)}_{ijmn} \tag{65}$$

$$\langle \bar{\mathbb{S}}^{E,\star}_{ijmn} \rangle_{\Omega_E} = \bar{s}_1^{E,\star} \mathbb{E}^{(1)}_{ijmn} + \bar{s}_2^{E,\star} \mathbb{E}^{(2)}_{ijmn} \tag{66}$$

with

$$\bar{s}_1^{I,D} = -\frac{(1+\nu)(1-f)}{3(1-\nu)}, \quad \bar{s}_2^{I,D} = -\frac{2(4-5\nu)(1-f)}{15(1-\nu)} + 21\gamma_u(1-f^{2/3})$$

$$\bar{s}_1^{E,D} = \frac{(1+\nu)f}{3(1-\nu)}, \quad \bar{s}_2^{E,D} = \frac{2(4-5\nu)f}{15(1-\nu)} - 21\gamma_u f \frac{1-f^{2/3}}{1-f} \tag{67}$$

for the Dirichlet BVP and

$$\bar{s}_1^{I,N} = \frac{2(1-2\nu)(1-f)}{3(1-\nu)}, \quad \bar{s}_2^{I,N} = \frac{(7-5\nu)(1-f)}{15(1-\nu)} - 21\gamma_t(1-f^{2/3})$$

$$\bar{s}_1^{E,N} = 1 - \frac{2(1-2\nu)f}{3(1-\nu)}, \quad \bar{s}_2^{E,N} = 1 - \frac{(7-5\nu)f}{15(1-\nu)} + 21\gamma_t f \frac{1-f^{2/3}}{1-f} \tag{68}$$

for the Neumann BVP. Here γ_u and γ_t are as given in Eq. (14). Note that we have omitted all superscripts on the right hand sides, with the understanding that all material properties in the above expressions (and in γ_u and γ_t) are in terms of the inclusion phase (i.e., $\nu = \nu_I$). Further, the volume fraction above is that of the inclusion (i.e., $f = f_I$). Substituting $f_I = 1 - f_E$ above and comparing Eqs. (67) and (68) with Eqs. (10)–(13), the following connections can be established between the exterior and interior eigenstrain Eshelby tensors

$$\bar{s}_1^{I,\star}(\nu, f_E) = s_1^{E,\star}(\nu, f_I), \quad \bar{s}_2^{I,\star}(\nu, f_E) \approx s_2^{E,\star}(\nu, f_I) \tag{69}$$

$$\bar{s}_1^{E,\star}(\nu, f_E) = s_1^{I,\star}(\nu, f_I), \quad \bar{s}_2^{E,\star}(\nu, f_E) \approx s_2^{I,\star}(\nu, f_I) \tag{70}$$

where $\star = D, N$. Note that the equality holds between the bulk coefficients s_1, whereas the deviatoric coefficients s_2 are only approximately equal. Likewise we can substitute $f_I = 1 - f_E$ into Eqs. (10)–(13) and compare these equations to Eqs. (67) and (68). Then we obtain

$$\bar{s}_1^{I,\star}(\nu, f_I) = s_1^{E,\star}(\nu, f_E), \quad \bar{s}_2^{I,\star}(\nu, f_I) \approx s_2^{E,\star}(\nu, f_E) \tag{71}$$

$$\bar{s}_1^{E,\star}(\nu, f_I) = s_1^{I,\star}(\nu, f_E), \quad \bar{s}_2^{E,\star}(\nu, f_I) \approx s_2^{I,\star}(\nu, f_E) \tag{72}$$

Next we consider the Mori–Tanaka model as an example to illustrate the exterior eigenstrain method and its relation to the interior eigenstrain method. Recall the Mori–Tanaka formula for the interior eigenstrain homogenization Eq. (54)

$$\bar{\mathbb{C}} = \mathbb{C}^E - f_I \mathbb{C}^E : \{(\mathbb{C}^E - \mathbb{C}^I)^{-1} : \mathbb{C}^E - f_E[\langle \mathbb{S}^{I,\star}(\mathbb{C}^E, f_I) \rangle_{\Omega_I} - \langle \mathbb{S}^{E,\star}(\mathbb{C}^E, f_I) \rangle_{\Omega_E}]\}^{-1} \tag{73}$$

Let us consider a two-phase composite with elasticities $\mathbb{C}_1$ and $\mathbb{C}_2$. Geometrically, the two phases can be arranged in two ways. We either let phase 1 be the matrix and phase 2 the inclusion ($\mathbb{C}^E = \mathbb{C}_1$, $\mathbb{C}^I = \mathbb{C}_2$) or vice versa ($\mathbb{C}^I = \mathbb{C}_1$, $\mathbb{C}^E = \mathbb{C}_2$). The equation above then takes the two forms

$$\bar{\mathbb{C}} = \mathbb{C}_1 - f_2 \mathbb{C}_1 : \{(\mathbb{C}_1 - \mathbb{C}_2)^{-1} : \mathbb{C}_1 - f_1[\langle \mathbb{S}^{I,\star}(\mathbb{C}_1, f_2) \rangle_{\Omega_I} - \langle \mathbb{S}^{E,\star}(\mathbb{C}_1, f_2) \rangle_{\Omega_E}]\}^{-1} \tag{74}$$

$$\bar{\mathbb{C}} = \mathbb{C}_2 - f_1 \mathbb{C}_2 : \{(\mathbb{C}_2 - \mathbb{C}_1)^{-1} : \mathbb{C}_2 - f_2[\langle \mathbb{S}^{I,\star}(\mathbb{C}_2, f_1) \rangle_{\Omega_I} - \langle \mathbb{S}^{E,\star}(\mathbb{C}_2, f_1) \rangle_{\Omega_E}]\}^{-1} \tag{75}$$

Reexamining the Mori–Tanaka method via exterior eigenstrain homogenization, we obtain

$$\bar{\mathbb{C}} = \mathbb{C}^I - f_E \mathbb{C}^I : \{(\mathbb{C}^I - \mathbb{C}^E)^{-1} : \mathbb{C}^I - f_I[\langle \bar{\mathbb{S}}^{E,\star}(\mathbb{C}^I, f_I) \rangle_{\Omega_E} - \langle \bar{\mathbb{S}}^{I,\star}(\mathbb{C}^I, f_I) \rangle_{\Omega_I}]\}^{-1} \tag{76}$$

which we call *the exterior eigenstrain Mori–Tanaka formula*. For a two-phase composite with elastic stiffnesses $\mathbb{C}_1$ and $\mathbb{C}_2$, we then have

$$\bar{\mathbb{C}} = \mathbb{C}_2 - f_1 \mathbb{C}_2 : \{(\mathbb{C}_2 - \mathbb{C}_1)^{-1} : \mathbb{C}_2 - f_2[\langle \bar{\mathbb{S}}^{E,\star}(\mathbb{C}_2, f_2) \rangle_{\Omega_E} - \langle \bar{\mathbb{S}}^{I,\star}(\mathbb{C}_2, f_2) \rangle_{\Omega_I}]\}^{-1} \tag{77}$$

$$\bar{\mathbb{C}} = \mathbb{C}_1 - f_2 \mathbb{C}_1 : \{(\mathbb{C}_1 - \mathbb{C}_2)^{-1} : \mathbb{C}_1 - f_1[\langle \bar{\mathbb{S}}^{E,\star}(\mathbb{C}_1, f_1) \rangle_{\Omega_E} - \langle \bar{\mathbb{S}}^{I,\star}(\mathbb{C}_1, f_1) \rangle_{\Omega_I}]\}^{-1} \tag{78}$$

Equations (74), (75), (77), and (78) constitute the four flavors of the Mori–Tanaka method. In view of relations (69)–(72) we can see that Eq. (77) is approximately equal to Eq. (75) and that Eq. (78) is approximately equal to Eq. (74). In fact, for the effective bulk modulus $\bar{\kappa}$ this approximation becomes an equality. For the effective shear modulus $\bar{\mu}$, however, there are slight differences. These differences can be seen in Fig. 5, which shows the effective shear modulus $\bar{\mu}$ for the four cases.

The material properties used in the calculation are $\kappa_2 = 4\kappa_1$, $\mu_2 = 10\mu_1$, and $\nu_2 = 0.3$. The traction boundary condition ($\star = N$) is used in the calculation. One can see how close the pairs 1, 4 and 2, 3 are. In the case of the effective bulk modulus these pairs are equal. We can therefore conclude that exchanging the material phases is approximately equal to exchanging the regions where the eigenstrain is prescribed.

This, however, does not mean that the exterior eigenstrain method has no technical merits. In the following sections we shall discuss two new models that are built upon the idea of the exterior eigenstrain method. The first, the dual eigenstrain method, furnishes a method that allows the smooth transition between curves 1 and 3 or between 2 and 4. Second, in Sec. 5, the shell model, a novel multiphase model, is a further generalization of this idea.

4.2 Dual Eigenstrain Method. We have seen in Fig. 5 that, for fixed phase distribution, the interior and exterior eigenstrain methods give very different homogenization results (i.e., the difference between 1 and 3 or between 2 and 4). We therefore want to consider a model that prescribes an eigenstrain field in both the interior and exterior regions of the inclusion simultaneously

$$\boldsymbol{\epsilon}^*(\mathbf{x}) = \begin{cases} \boldsymbol{\epsilon}_I^*, & \forall \mathbf{x} \in \Omega_I \\ \boldsymbol{\epsilon}_E^*, & \forall \mathbf{x} \in \Omega_E \end{cases} \tag{79}$$

The model, termed the *dual eigenstrain method* is discussed in

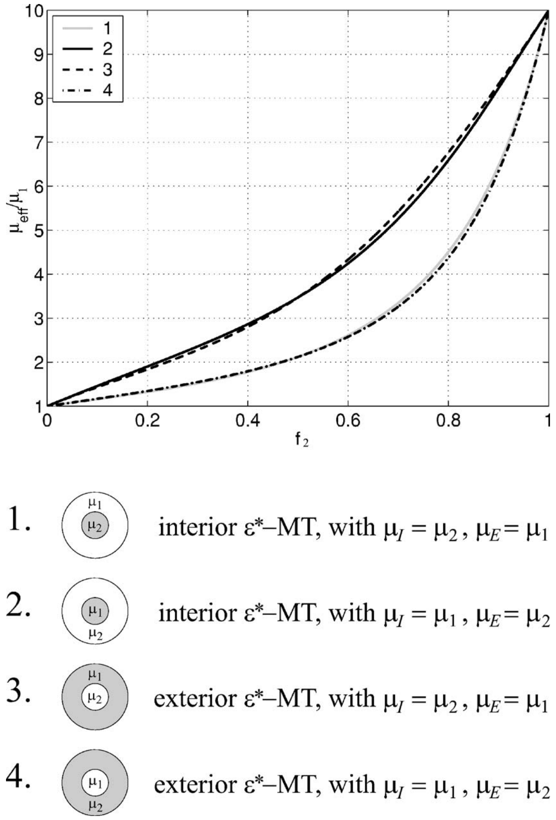


Fig. 5 Mori–Tanaka homogenization for the interior and exterior eigenstrain methods

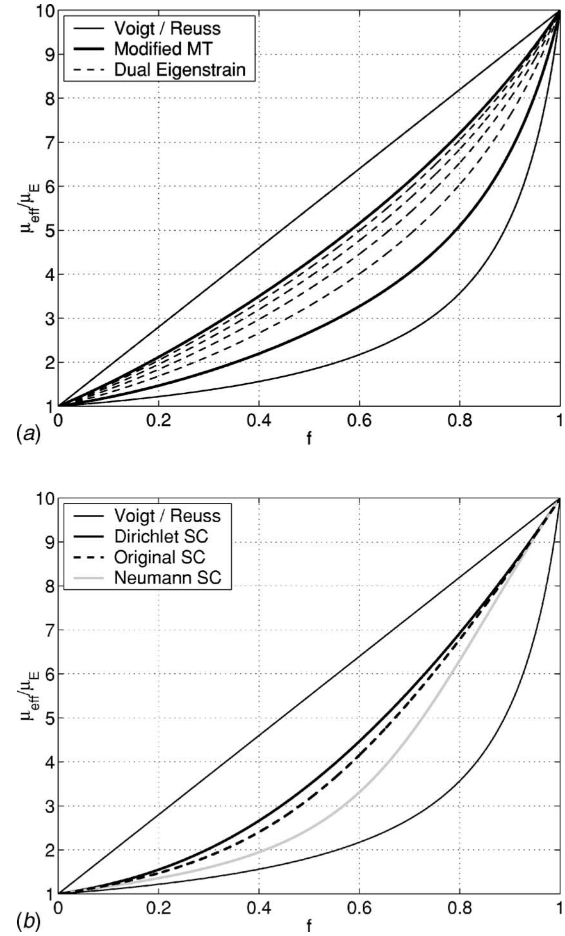


Fig. 6 Effective shear modulus for: (*a*) $\tilde{\mathbb{C}} = a\mathbb{C}^I + (1-a)\mathbb{C}^E$, and (*b*) $\tilde{\mathbb{C}} = \bar{\mathbb{C}}$

detail in Ref. [12] and we only illustrate the main concept of the method here. The central idea of the *dual eigenstrain method* is to treat the homogenization of both the inclusion and matrix phase equally. Such a model has no preference between the two material phases, because neither phase is chosen as the reference state. The comparison solid for the RVE is rather characterized by $\tilde{\mathbb{C}}$, which can be considered as an estimate of the effective modulus $\bar{\mathbb{C}}$. Concrete choices for $\tilde{\mathbb{C}}$ are considered later.

Denoting the average background strain as $\boldsymbol{\epsilon}^b$, the dual stress consistency conditions then become

$$\mathbb{C}^I:[\boldsymbol{\epsilon}^b + \boldsymbol{\epsilon}^d(\mathbf{x})] = \tilde{\mathbb{C}}:[\boldsymbol{\epsilon}^b + \boldsymbol{\epsilon}^d(\mathbf{x}) - \boldsymbol{\epsilon}_I^*], \quad \forall\, \mathbf{x} \in \Omega_I \tag{80}$$

$$\mathbb{C}^E:[\boldsymbol{\epsilon}^b + \boldsymbol{\epsilon}^d(\mathbf{x})) = \tilde{\mathbb{C}}:(\boldsymbol{\epsilon}^b + \boldsymbol{\epsilon}^d(\mathbf{x}) - \boldsymbol{\epsilon}_E^*), \quad \forall\, \mathbf{x} \in \Omega_E \tag{81}$$

The disturbance strain field $\boldsymbol{\epsilon}^d$ is the superposition of the disturbance strain field due to the interior eigenstrain and the disturbance strain field due to the exterior eigenstrain field, i.e.

$$\boldsymbol{\epsilon}^d(\mathbf{x}) = \begin{cases} \mathbb{S}^{I,\star}:\boldsymbol{\epsilon}_I^* + \bar{\mathbb{S}}^{I,\star}:\boldsymbol{\epsilon}_E^*, & \forall \mathbf{x} \in \Omega_I \\ \mathbb{S}^{E,\star}:\boldsymbol{\epsilon}_I^* + \bar{\mathbb{S}}^{E,\star}:\boldsymbol{\epsilon}_E^*, & \forall \mathbf{x} \in \Omega_E \end{cases} \tag{82}$$

Here $\mathbb{S}^{\cdot,\star}$ is the interior eigenstrain finite Eshelby tensor as derived in Part I and $\bar{\mathbb{S}}^{\cdot,\star}$ is the exterior eigenstrain Eshelby tensor as given in the preceding section; (both accepting $\cdot = I$ or E and $\star = D$ or N). Both $\mathbb{S}^{\cdot,\star}$ and $\bar{\mathbb{S}}^{\cdot,\star}$ take $\tilde{\mathbb{C}}$ as the comparison solid. We note that the dual eigenstrain method contains the two special cases $\boldsymbol{\epsilon}_E^* = 0$, with $\tilde{\mathbb{C}} = \mathbb{C}^E$ and $\boldsymbol{\epsilon}_I^* = 0$, with $\tilde{\mathbb{C}} = \mathbb{C}^I$, which are the interior and exterior eigenstrain methods, respectively.

From here on the derivation proceeds in a similar manner as the interior eigenstrain case (see Sec. 3). The effective elasticity modulus is obtained as

$$\bar{\mathbb{C}} = [f\mathbb{C}^I:\mathcal{A}_E + (1-f)\mathbb{C}^E:\mathcal{A}_I]:[f\mathcal{A}_E + (1-f)\mathcal{A}_I]^{-1} \tag{83}$$

with the concentration tensors

$$\mathcal{A}_E = \mathbb{I}^s - \tilde{\mathbb{C}}^{-1}:(\tilde{\mathbb{C}} - \mathbb{C}^E):\Delta\mathbb{S}^\star$$

$$\mathcal{A}_I = \mathbb{I}^s - \tilde{\mathbb{C}}^{-1}:(\tilde{\mathbb{C}} - \mathbb{C}^I):\Delta\mathbb{S}^\star \tag{84}$$

and the difference

$$\Delta\mathbb{S}^\star := \langle \mathbb{S}^{I,\star}(\tilde{\mathbb{C}}, f = f_I)\rangle_{\Omega_E} - \langle \mathbb{S}^{E,\star}(\tilde{\mathbb{C}}, f = f_I)\rangle_{\Omega_I} \tag{85}$$

Here we have indicated that $\mathbb{S}^{\cdot,\star}$ depends on the comparison solid $\tilde{\mathbb{C}}$ and the volume fraction $f = f_I$. We note that the coefficients of $\Delta\mathbb{S}^\star$ follow from Eqs. (18)–(20) given in Sec. 2, with setting $\nu = \tilde{\nu}$ and $f = f_I$. Choosing either $\tilde{\mathbb{C}} = \mathbb{C}^E$ or $\tilde{\mathbb{C}} = \mathbb{C}^I$, the method degenerates to the interior homogenization Eq. (73) or the exterior homogenization Eq. (76), respectively. If we let $\tilde{\mathbb{C}}$ assume a value between $\mathbb{C}^I$ and $\mathbb{C}^E$, the effective modulus $\bar{\mathbb{C}}$ given by the dual eigenstrain method can be expected to lie in between these two special cases. As an example consider the convex combination

$$\tilde{\mathbb{C}} = a\mathbb{C}^I + (1-a)\mathbb{C}^E, \quad 0 \leqslant a \leqslant 1 \tag{86}$$

Figure 6(*a*) shows the effective shear modulus $\bar{\mu}$ obtained from Eq. (83) using Eq. (86) for $a = \{0, 0.2, 0.4, 0.6, 0.8, 1\}$.

The material properties have been chosen as before ($\kappa_I = 4\kappa_E$, $\mu_I = 10\mu_E$, and $\nu_E = 0.3$). The boundary condition in Fig. 6(*a*) is chosen as the Dirichlet BC ($\star = D$). We observe that $\bar{\mu}$, computed

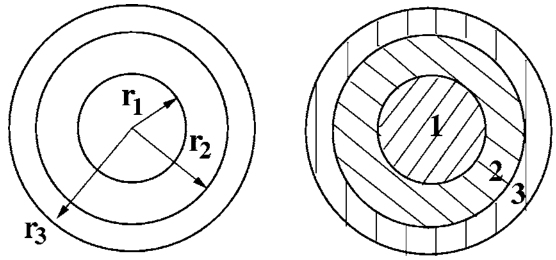


Fig. 7 A three-layer shell model

by the dual eigenstrain method, lies in between the two special cases, the interior and exterior eigenstrain MT method obtained for $a=0$ and $a=1$, respectively. We remark that there are other interesting choices for $\tilde{\mathbb{C}}$ to consider (see Ref. [12]). For instance, we can use $\tilde{\mathbb{C}}$ as a predictor of the effective modulus $\bar{\mathbb{C}}$. One such predictor is the Voigt bound. Equation (83) then becomes a corrected value for $\bar{\mathbb{C}}$. This predictor–corrector scheme can be viewed as a generalization of the classical self-consistent method. An implicit self-consistent method arises when considering $\tilde{\mathbb{C}}=\bar{\mathbb{C}}$ in Eq. (83). The effective shear modulus for this case is shown in Fig. 6(*b*) for both $\star=D$ and N. For comparison the original self-consistent method is also shown. The material data are the same as before.

As a final remark, let us consider the dual eigenstrain method in view of the two possible ways of arranging the material phases. As we have discussed in the preceding section we can either have $\mathbb{C}^E=\mathbb{C}^1$ and $\mathbb{C}^I=\mathbb{C}^2$ or the flipped case $\mathbb{C}^I=\mathbb{C}^1$ and $\mathbb{C}^E=\mathbb{C}^2$. Thus the dual eigenstrain method Eq. (83) results in two distinct formulas.

5 A Shell Model

To utilize the finite Eshelby tensors to represent different microstructures, a so-called spherical shell model is developed, that is a n-phase composite RVE modeled by n concentric spherical shells. To illustrate the model, we present the detailed study of a three-layer shell model (see Fig. 7).

For the three-layer shell model, the RVE consists of three concentric spherical shells, which are labeled as

$$\Omega_1(\mathbf{x})=\{\mathbf{x}||\mathbf{x}|<r_1\},\quad \Omega_2(\mathbf{x})=\{\mathbf{x}|r_1<|\mathbf{x}|<r_2\},$$

$$\Omega_3(\mathbf{x})=\{\mathbf{x}|r_2<|\mathbf{x}|<r_3\}$$

Here the radius of the RVE is r_3, and the volume fraction of the three shells are

$$f_1=\left(\frac{r_1}{r_3}\right)^3,\quad f_2=\frac{r_2^3-r_1^3}{r_3^3},\quad f_3=\frac{r_3^3-r_2^3}{r_3^3} \tag{87}$$

with

$$f_1+f_2+f_3=1$$

To derive the Eshelby tensors for each shell, we consider three partially overlapped concentric spheres

$$\Omega_{\mathrm{I}}(\mathbf{x})=\{\mathbf{x}||\mathbf{x}|<r_1\},\quad \Omega_{\mathrm{II}}(\mathbf{x})=\{\mathbf{x}||\mathbf{x}|<r_2\},\quad \Omega_{\mathrm{III}}(\mathbf{x})=\{\mathbf{x}||\mathbf{x}|<r_3\}$$

The interior and exterior Eshelby tensors for each sphere Ω_J are denoted as

$$\mathbb{S}^{J,F}(\mathbf{x}):=\begin{cases}\mathbb{S}^{I,F}(\mathbf{x}), & \forall\mathbf{x}\in\Omega_J, \quad J=\mathrm{I,II,III}\\ \mathbb{S}^{E,F}(\mathbf{x}), & \forall\mathbf{x}\in\Omega/\Omega_J, \quad J=\mathrm{I,II,III}\end{cases} \tag{88}$$

where the superscript F represents the general boundary conditions, see Eq. (39). Subsequently the average of the Eshelby tensor is required for each shell. We first denote the average of the Eshelby tensor of the overlapping spheres

$$\mathbb{S}^{Jj,F}:=\langle\mathbb{S}^{J,F}\rangle_{\Omega_j},\quad J=\mathrm{I,II,III}\quad\text{and } j=1,2,3 \tag{89}$$

where the first superscript J (Roman numbers) denotes the sphere, Ω_J, in which the eigenstrain is prescribed, and where the second superscript j (Arabic numbers) denotes the shell, Ω_j, over which the average is taken. Similarly we denote the average Eshelby tensor of the shell domains as

$$\mathbb{S}^{ij,F}:=\langle\mathbb{S}^{i,F}\rangle_{\Omega_j},\quad i=1,2,3\quad\text{and } j=1,2,3 \tag{90}$$

Again, the first subscript index, i, refers to the shell region, Ω_i, in which the eigenstrains are prescribed, and the second index, j, denotes the shell region, Ω_j, over which the average is taken. As we have shown in Sec. 2 the average Eshelby tensors can be written as

$$\mathbb{S}^{ij,F}=s_1^{ij,F}\mathbb{E}^{(1)}+s^{ij,F}\mathbb{E}^{(2)},\quad i,j=1,2,3 \tag{91}$$

The idea is to use the Eshelby tensors of three overlapping spheres to represent the Eshelby tensors of the shells via superposition. For the first spherical shell (the inner most shell) we write

$$\mathbb{S}^{11,F}=\mathbb{S}^{I1,F}=s_1^{11,F}\mathbb{E}^{(1)}+s_2^{11,F}\mathbb{E}^{(2)}$$

$$\mathbb{S}^{1j,F}=\mathbb{S}^{Ij,F}=s_1^{1j,F}\mathbb{E}^{(1)}+s_2^{1j,F}\mathbb{E}^{(2)} \tag{92}$$

Here, $s_\alpha^{11,F}$ are the coefficients of the interior Eshelby tensor, whereas $s_\alpha^{1j,F}$, $j=2,3$ are the coefficients of the exterior Eshelby tensors. Using superposition, the Eshelby tensors for the second and third spherical shells can be obtained by using the combination of the average Eshelby tensors of the three overlapping spheres

$$\mathbb{S}^{2i,F}=\mathbb{S}^{\mathrm{II}i,F}-\mathbb{S}^{\mathrm{I}i,F},\quad i=1,2,3 \tag{93}$$

$$\mathbb{S}^{3i,F}=\mathbb{S}^{\mathrm{III}i,F}-\mathbb{S}^{\mathrm{II}i,F},\quad i=1,2,3 \tag{94}$$

Therefore, for $\alpha=1,2$

$$s_\alpha^{2i,F}=s_\alpha^{\mathrm{II}i,F}-s_\alpha^{\mathrm{I}i,F}$$

and

$$s_\alpha^{3i,F}=s_\alpha^{\mathrm{III}i,F}-s_\alpha^{\mathrm{II}i,F},\quad i=1,2,3 \tag{95}$$

To this end, all the coefficients of the Eshelby tensors for each shell layer are expressed in terms the of the Eshelby coefficients for solid spheres, Ω_I, Ω_{II}, and Ω_{III}, which are documented in the Appendix for a three-sphere RVE.

To illustrate the application of the shell model, we consider a simple homogenization example of a two-phase composite material, with elastic modulus $\mathbb{C}_2=\mathbb{C}_e$ in Ω_2 and $\mathbb{C}_1=\mathbb{C}_3=\mathbb{C}$ in Ω_1 and Ω_3. We prescribe the eigenstrain in Ω_2,

$$\boldsymbol{\epsilon}^*(\mathbf{x})=\begin{cases}\boldsymbol{\epsilon}^*, & \forall\mathbf{x}\in\Omega_2\\ 0, & \text{otherwise}\end{cases} \tag{96}$$

We assume the RVE is subjected to the macrostrain boundary condition, i.e., $\mathbf{u}=\boldsymbol{\epsilon}^0\mathbf{x}$, $\forall\mathbf{x}\in\partial\Omega$ and impose the following stress consistency condition

$$\mathbb{C}_e:(\boldsymbol{\epsilon}^0+\boldsymbol{\epsilon}^d)=\mathbb{C}:(\boldsymbol{\epsilon}^0+\boldsymbol{\epsilon}^d-\boldsymbol{\epsilon}^*),\quad\forall\,\mathbf{x}\in\Omega_2 \tag{97}$$

One can then derive the effective elastic stiffness similar to the formula of dilute suspension homogenization Eq. (35)

$$\bar{\mathbb{C}}=\mathbb{C}-f\mathbb{C}:(\mathbb{A}_e-\mathbb{S}^{22,D})^{-1}, \tag{98}$$

where $\mathbb{A}_e=(\mathbb{C}-\mathbb{C}_e)^{-1}:\mathbb{C}$. We note that here $\mathbb{S}^{22,D}$ is a function of both volume fraction of the second phase Ω_2 (i.e., $f=f_2$), and the geometric allocation or separation of the first phase, which can be characterized by a nondimensional parameter $\beta:=f_1/(f_1+f_3)$ $=f_1/(1-f)$. The coefficients of $\mathbb{S}^{22,D}$ are found to be

$$s_1^{22,D}=\frac{(1+\nu)(1-f)}{3(1-\nu)} \tag{99}$$

$$s_2^{22,D} = \frac{8-10\nu}{15(1-\nu)}(1-\beta)(1-f) - 21\gamma_u[\beta(1-f)+f]\cdot\left\{1 - \frac{[\beta(1-f)+f)^{5/3} - [\beta(1-f)]^{5/3}}{f}\right\} + \frac{2\beta(1-f)}{15(1-\nu)} + 21\gamma_u[\beta(1-f)]\frac{f-\beta(1-f)-[\beta(1-f)+f]^{5/3}+[\beta(1-f)]^{5/3}}{f-\beta(1-f)} \tag{100}$$

For the general boundary condition of Eq. (39), the following stress consistency is imposed

$$\mathbb{C}_e:(\boldsymbol{\epsilon}^b + \boldsymbol{\epsilon}^d) = \mathbb{C}:(\boldsymbol{\epsilon}^b + \boldsymbol{\epsilon}^d - \boldsymbol{\epsilon}^*), \quad \forall\, \mathbf{x} \in \Omega_2 \tag{101}$$

This leads to the usual relationships between the average strain and eigenstrain, as well as between the eigenstrain and the background strain, i.e.

$$\langle \boldsymbol{\epsilon} \rangle_{\Omega_2} = \mathbb{A}_e:\boldsymbol{\epsilon}^* \tag{102}$$

and

$$\boldsymbol{\epsilon}^* = (\mathbb{A}_e - \mathbb{S}^{22,F}):\boldsymbol{\epsilon}^b, \quad \text{with } \mathbb{A}_e = (\mathbb{C} - \mathbb{C}_e)^{-1}:\mathbb{C}$$

Then the average strain in the RVE will be

$$\begin{aligned}\langle \boldsymbol{\epsilon} \rangle_\Omega &= f_1\langle \boldsymbol{\epsilon} \rangle_{\Omega_1} + f_2\langle \boldsymbol{\epsilon} \rangle_{\Omega_2} + f_3\langle \boldsymbol{\epsilon} \rangle_{\Omega_3} \\ &= \boldsymbol{\epsilon}^b + f_1\mathbb{S}^{21,F}:\boldsymbol{\epsilon}^* + f_2\mathbb{S}^{22,F}:\boldsymbol{\epsilon}^* + f_3\mathbb{S}^{23,F}:\boldsymbol{\epsilon}^* \\ &= [\mathbb{A}_e + f_1\mathbb{S}^{21,F} - (f_1+f_3)\mathbb{S}^{22,F} + f_3\mathbb{S}^{23,F}]:[\mathbb{A}_e - \mathbb{S}^{22,F}]^{-1}:\boldsymbol{\epsilon}^b\end{aligned} \tag{103}$$

Further, the average stress in the RVE becomes

$$\langle \boldsymbol{\sigma} \rangle_\Omega = f_1\langle \boldsymbol{\sigma} \rangle_{\Omega_1} + f_2\langle \boldsymbol{\sigma} \rangle_{\Omega_2} + f_3\langle \boldsymbol{\sigma} \rangle_{\Omega_3} = \{(f_1\mathbb{C} + f_2\mathbb{C}_e + f_3\mathbb{C}):(\mathbb{A}_e - \mathbb{S}^{22,F}) + [\mathbb{C}(f_1\mathbb{S}^{21,F} + f_3\mathbb{S}^{23,F}) + f_2\mathbb{C}_e\mathbb{S}^{22,F}]\}:(\mathbb{A}_e - \mathbb{S}^{22,F})^{-1}:\boldsymbol{\epsilon}^b \tag{104}$$

Substituting Eq. (103) into Eq. (104) leads to the effective elastic tensor

$$\bar{\mathbb{C}} = \mathbb{C} - f\mathbb{C}:\{\mathbb{A}_e - (1-f)[\mathbb{S}^{22,F} - \beta\mathbb{S}^{21,F} - (1-\beta)\mathbb{S}^{23,F}]\}^{-1}, \tag{105}$$

where $f=f_2$ and

$$\beta = \frac{f_1}{f_1+f_3} = \frac{f_1}{1-f}$$

It is interesting to point out that the above formalism resembles the classical Mori–Tanaka model Eq. (54). For the shell model with the eigenstrain prescribed in Ω_2, the contribution from the exterior Eshelby tensor is represented by a linear combination of $\mathbb{S}^{21,F}$ and $\mathbb{S}^{23,F}$ through parameter $\beta \in [0,1]$, which can be used to characterize the evolution of the microstructure. Figure 8 shows that the microstructure evolution can have some influence on the effective shear modulus. In Fig. 8(*a*), the range of the effective shear modulus for $\beta=0, 0.1, \ldots, 0.9, 1.0$ is plotted for the Dirichlet, Neumann, and averaged ($\alpha=0.5$ in Eq. (39)) boundary conditions, respectively. The differences can be seen more clearly if, for a given volume fraction, the effective modulus is plotted over β. This is shown in Fig. 8(*b*) for the volume fraction $f=0.5$, which demonstrates the dependency of the shell model on the microstructure. Note that this dependency has little influence for the average case but is considerably stronger for both the Dirichlet and Neumann case.

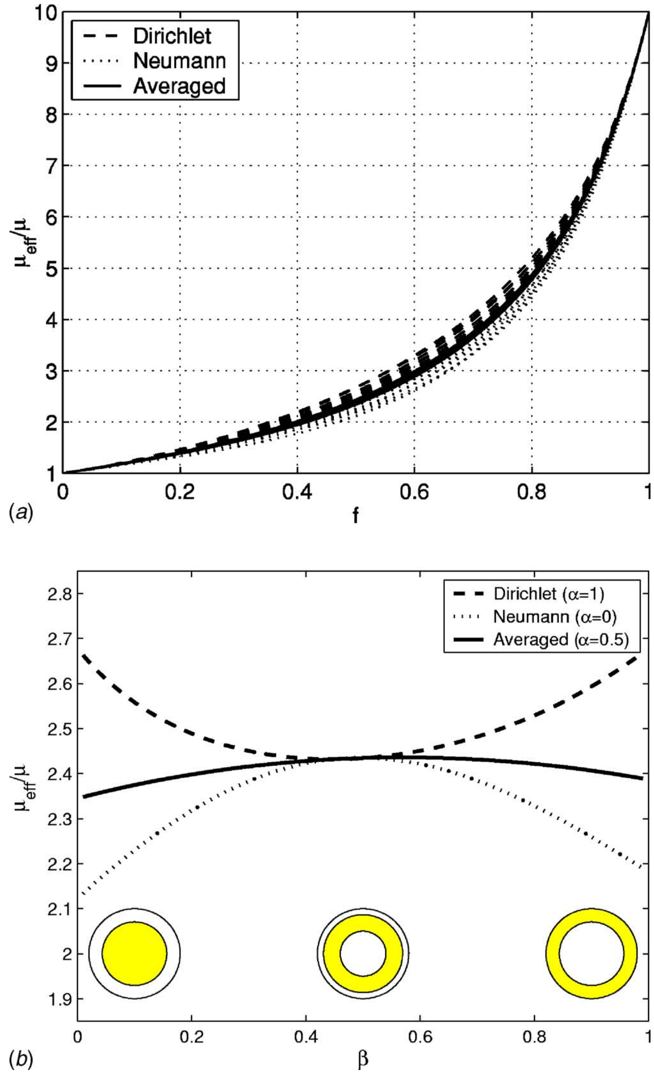


Fig. 8 Influence of β on the effective shear modulus

6 New Variational Bounds

One of the useful homogenization methods for composite materials are the Hashin–Shtrikman variational principles, which have been extensively used in deriving bounds for effective material properties. In the procedure of deriving the variational bounds, the Eshelby tensor is needed in order to estimate the disturbance strain field due to stress polarization or to estimate the disturbance stress field due to the eigenstrain.

Since the classical Eshelby tensor is obtained for an inclusion solution in an unbounded region, in principle, it can not be directly used in the derivation of the variational bounds of a composite with finite volume. In the past, additional probability arguments and approximations based on assumptions of the statistical nature of the inclusion distribution, have been employed to justify the use of the classical Eshelby tensor, e.g., Ref. [14].

In this section, we show that the finite Eshelby tensors are a perfect fit for the Hashin–Shtrikman variational principles [15,16]. They can be directly used in combination with the Hashin–Shtrikman principles to derive variational bounds without resorting to additional statistical arguments. By using the shell model proposed in the previous section, a systematic, multivariable optimization procedure is developed for multiphase composites.

We consider the case that the finite spherical RVE is subjected to a displacement boundary condition, i.e.

$$\mathbf{u}(\mathbf{x}) = \boldsymbol{\epsilon}^0\mathbf{x}, \quad \forall\, \mathbf{x} \in \partial\Omega \tag{106}$$

The standard statement of the Hashin–Shtrikman principles may be expressed in the following form

$$\underline{I}_p(\mathbf{p},\boldsymbol{\epsilon}^d) \leqslant \inf_{\boldsymbol{\epsilon}^d \in E} W(\boldsymbol{\epsilon}^d) \leqslant \bar{I}_p(\mathbf{p},\boldsymbol{\epsilon}^d) \tag{107}$$

where

$$W(\boldsymbol{\epsilon}) = \frac{1}{2|\Omega|}\int_\Omega \boldsymbol{\sigma}:\boldsymbol{\epsilon}\, d\Omega \tag{108}$$

is the strain energy density, and

$$I_p(\mathbf{p},\boldsymbol{\epsilon}^d) = W_0(\boldsymbol{\epsilon}^0) - \frac{1}{2|\Omega|}\int_\Omega \{\mathbf{p}:(\mathbb{C}-\mathbb{C}_0)^{-1}:\mathbf{p} - \mathbf{p}:\boldsymbol{\epsilon}^d - 2\mathbf{p}:\boldsymbol{\epsilon}^0\} d\Omega \tag{109}$$

Here $\mathbb{C}$ is the elastic tensor of the composite and $\mathbb{C}_0$ is the elastic tensor of a comparison solid such that

$$I_p = \begin{cases} \bar{I}_p, & \text{if } \Delta\mathbb{C} = \mathbb{C} - \mathbb{C}_0 < 0 \\ \underline{I}_p, & \text{if } \Delta\mathbb{C} = \mathbb{C} - \mathbb{C}_0 > 0 \end{cases} \tag{110}$$

Here $W_0(\boldsymbol{\epsilon}^0)$ is the strain energy density of the comparison solid; $\mathbf{p}$ is the stress polarization; and $\boldsymbol{\epsilon}^d$ is the disturbance strain field due to the stress polarization. They are related by the following subsidiary boundary value problem

$$\nabla \cdot [\mathbb{C}_0:\nabla\mathbf{u}^d(\mathbf{x}) + \mathbf{p}(\mathbf{x})] = 0, \quad \forall\, \mathbf{x} \in \Omega, \quad \mathbf{u}^d(\mathbf{x}) = 0, \quad \forall\, \mathbf{x} \in \partial\Omega \tag{111}$$

We consider the composite to be made of n distinct phases and assume that each phase may be represented by a hollow spherical shell inside the RVE. The homogenization or statistical model of the composite is that any macropoint of the composite is modeled as a RVE consisting of n distinct concentric spherical shells with domain Ω_i so that $\cup_i^n \Omega_i = \Omega$ and $\cap_{i=1}^n \Omega_i = \varnothing$. The stress polarization is chosen as a piecewise constant tensorial field

$$\mathbf{p}(\mathbf{x}) = \sum_{i=1}^n \mathbf{p}_i \chi(\Omega_i)$$

with

$$\chi(\Omega_i) = \begin{cases} 1, & \forall\, \mathbf{x} \in \Omega_i \\ 0, & \forall\, \mathbf{x} \notin \Omega_i \end{cases} \tag{112}$$

Let us consider

$$\mathbf{p}_i = p_i \mathbb{I}^{(2)} + \tau_i \mathbb{J}^{(2)} \tag{113}$$

where $\mathbb{I}^{(2)}$ is the second-order unit tensor and $\mathbb{J}^{(2)}$ is its counterpart, the so-called deviatoric unit tensor, both defined as

$$\mathbb{I}^{(2)} = \delta_{ij}\mathbf{e}_i \otimes \mathbf{e}_j, \quad \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases},$$

$$\mathbb{J}^{(2)} = \beta_{ij}\mathbf{e}_i \otimes \mathbf{e}_j, \quad \beta_{ij} = \begin{cases} 0, & i = j \\ 1, & i \neq j \end{cases} \tag{114}$$

Based on the finite spherical inclusion model, the average disturbance strain will be the summation of the average disturbance strain in each phase

$$\langle \boldsymbol{\epsilon}^d \rangle_\Omega = -\sum_{i=1}^n \sum_{j=1}^n \mathbb{C}_0^{-1}:\langle \mathbb{S}^{ij,D} \rangle:\mathbf{p}_i \tag{115}$$

As shown in Sec. 2, the average Eshelby tensor can be written as

$$\langle \mathbb{S}^{ij,D} \rangle = s_1^{ij,D}\mathbb{E}^{(1)} + s_2^{ij,D}\mathbb{E}^{(2)} \tag{116}$$

We choose the prescribed boundary field as

$$\boldsymbol{\epsilon}^0 = \bar{\epsilon}\mathbb{I}^{(2)} + \bar{\gamma}\mathbb{J}^{(2)} \tag{117}$$

so that we obtain

$$\begin{aligned} I_p &= W_0(\bar{\boldsymbol{\epsilon}}) - \frac{1}{2|\Omega|}\int_\Omega \{\mathbf{p}:(\mathbb{C}-\mathbb{C}_0)^{-1}:\mathbf{p} - \mathbf{p}:\boldsymbol{\epsilon}^d - 2\mathbf{p}:\bar{\boldsymbol{\epsilon}}\} d\Omega \\ &= \frac{9}{2}\kappa_0\bar{\epsilon}^2 + 6\mu_0\bar{\gamma}^2 - \sum_{i=1}^n \left[\frac{f_i p_i^2}{2(\kappa_i - \kappa_0)} + \frac{3 f_i \tau_i^2}{2(\mu_i - \mu_0)} \right] \\ &\quad - \sum_{i=1}^n \sum_{j=1}^n \left(\frac{f_i s_1^{ji,D} p_i p_j}{2\kappa_0} + \frac{3 f_i s_2^{ji,D} \tau_i \tau_j}{2\mu_0} \right) + \sum_{i=1}^n (3 f_i p_i \bar{\epsilon} + 6 f_i \tau_i \bar{\gamma}) \end{aligned} \tag{118}$$

We first let $\partial I_p / \partial p_i = 0$. One can thus find

$$-\frac{p_i}{\kappa_i - \kappa_0} - \frac{s_1^{ii,D} p_i}{\kappa_0} - \sum_{j \neq i} \frac{s_1^{ji,D} p_j}{2\kappa_0} + 3\bar{\epsilon} = 0, \quad \forall\, i = 1,2,\cdots n \tag{119}$$

Hence the stationary value of each p_i can be obtained through the following system of equations

$$\begin{bmatrix} \ddots & & & & & \\ \cdots & \cdots & \left(\dfrac{s_1^{ii,D}}{\kappa_0} + \dfrac{1}{\kappa_i - \kappa_0} \right) & \cdots & \dfrac{s_1^{ji,D}}{2\kappa_0} & \cdots \\ & & & & \ddots & \\ & & & & & \ddots \\ & & & & & & \ddots \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ p_i \\ \vdots \\ p_j \\ \vdots \\ p_n \end{bmatrix} = 3\bar{\epsilon} \begin{bmatrix} 1 \\ \vdots \\ \vdots \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} \tag{120}$$

We further let $\partial I_p / \partial \tau_i = 0$, which leads to

$$-\frac{3 f_i \tau_i}{\mu_i - \mu_0} - \frac{3 f_i s_2^{ii,D} \tau_i}{\mu_0} - \sum_{j \neq i} \frac{3 f_i s_2^{ji,D} \tau_j}{2\mu_0} + 6 f_i \bar{\gamma} = 0 \tag{121}$$

or in matrix form

$$\begin{bmatrix} \ddots & & & & & \\ \cdots & \cdots & \left(\dfrac{s_2^{ii,D}}{\mu_0} + \dfrac{1}{\mu_i - \mu_0} \right) & \cdots & \dfrac{s_2^{ji,D}}{2\mu_0} & \cdots \\ & & & & \ddots & \\ & & & & & \ddots \\ & & & & & & \ddots \end{bmatrix} \begin{bmatrix} \tau_1 \\ \vdots \\ \tau_i \\ \vdots \\ \tau_j \\ \vdots \\ \tau_n \end{bmatrix} = 2\bar{\gamma} \begin{bmatrix} 1 \\ \vdots \\ \vdots \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} \tag{122}$$

Remark 6.1. In the past, when deriving variational bounds for multiphase composites, the same infinite Eshelby tensor was used for all the phases (except the comparison phase) without discrimination. This procedure excludes the interactions among different phases at the outset. By applying the shell model, proposed in the last section, with the finite Eshelby tensor this interaction can now be taken into account.

6.1 Two-Phase Composites. We now consider an isotropic two-phase composite, with $\kappa_2 > \kappa_1$ and $\mu_2 > \mu_1$. For the effective bulk modulus, we find the following bound under the prescribed displacement boundary condition

$$\kappa_1 + \frac{f_2}{\dfrac{1}{\kappa_2 - \kappa_1} + \dfrac{s_1^{22,D}}{\kappa_1}} \leqslant \bar{\kappa} \leqslant \kappa_2 + \frac{f_1}{\dfrac{1}{\kappa_1 - \kappa_2} + \dfrac{s_1^{11,D}}{\kappa_2}} \tag{123}$$

where

$$s_1^{22,D} = \frac{(1+\nu_1)f_1}{3(1-\nu_1)}$$

and

$$s_1^{11,D} = \frac{(1+\nu_2)f_2}{3(1-\nu_2)} \quad (124)$$

A similar result can be derived for the Neumann boundary condition

$$\kappa_1^{-1} + \frac{f_2}{\dfrac{1}{\kappa_2^{-1}-\kappa_1^{-1}} + \dfrac{1-s_1^{22,N}}{\kappa_1^{-1}}} \geqslant \bar{\kappa}^{-1} \geqslant \kappa_2^{-1} + \frac{f_1}{\dfrac{1}{\kappa_1^{-1}-\kappa_2^{-1}} + \dfrac{1-s_1^{11,N}}{\kappa_2^{-1}}} \quad (125)$$

where

$$s_1^{22,N} = \frac{1+\nu_1+2(1-2\nu_1)f_2}{3(1-\nu_1)}$$

and

$$s_1^{11,N} = \frac{1+\nu_2+2(1-2\nu_2)f_1}{3(1-\nu_2)} \quad (126)$$

It can be shown, by algebraic manipulation, that the bounds (123) and (125) are identical. Furthermore, they are equal to the original Hashin–Strikman bounds, because the coefficients (124) and (126) are equal to those of the original infinite Eshelby tensor.

Similarly, the bounds for the shear modulus can be obtained as

$$\mu_1 + \frac{f_2}{\dfrac{1}{\mu_2-\mu_1} + \dfrac{s_2^{22,D}}{\mu_1}} \leqslant \bar{\mu} \leqslant \mu_2 + \frac{f_1}{\dfrac{1}{\mu_1-\mu_2} + \dfrac{s_2^{11,D}}{\mu_2}} \quad (127)$$

where

$$s_2^{22,D} = \frac{2(4-5\nu_1)f_1}{15(1-\nu_1)} - \frac{21f_2(1-f_2^{2/3})^2}{10(1-\nu_1)(7-10\nu_1)} \quad (128)$$

$$s_2^{11,D} = \frac{2(4-5\nu_2)f_2}{15(1-\nu_2)} - \frac{21f_1(1-f_1^{2/3})^2}{10(1-\nu_2)(7-10\nu_2)} \quad (129)$$

and

$$\mu_1^{-1} + \frac{f_2}{\dfrac{1}{\mu_2^{-1}-\mu_1^{-1}} + \dfrac{1-s_2^{22,N}}{\mu_1^{-1}}} \geqslant \bar{\mu}^{-1} \geqslant \mu_2^{-1} + \frac{f_1}{\dfrac{1}{\mu_1^{-1}-\mu_2^{-1}} + \dfrac{1-s_2^{11,N}}{\mu_2^{-1}}} \quad (130)$$

where

$$s_2^{22,N} = \frac{2(4-5\nu_1)+(7-5\nu_1)f_2}{15(1-\nu_1)} + \frac{84f_2^2(1-f_2^{2/3})^2}{10(1-\nu_1)(7+5\nu_1)} \quad (131)$$

$$s_2^{11,N} = \frac{2(4-5\nu_2)+(7-5\nu_2)f_1}{15(1-\nu_2)} + \frac{84f_1^2(1-f_1^{2/3})^2}{10(1-\nu_2)(7+5\nu_2)} \quad (132)$$

Now the shear modulus bounds (127) and (130), are distinct, and they are different from the original Hashin–Shtrikman bounds based on the classical Eshelby tensor in an unbounded RVE. The new variational bounds for both bulk and shear modulus are displayed in Fig. 9 with respect to f_2. The material data is chosen as $\kappa_2=4\kappa_1$, $\mu_2=10\mu_1$, and $\nu_1=0.3$ (implying $\nu_2=0.083$).

Figure 9(*a*) shows that the boundary conditions have no effect on the bulk modulus, whose bounds coincide with the original HS bounds. On the other hand the boundary conditions do affect the variational bounds of the shear modulus. In Fig. 9(*b*), the three sets of the variational bounds (Dirichlet, Neumann, and the original) for the shear modulus are juxtaposed in comparison. We note that the difference between these three pairs is solely caused by the second term in coefficients $s_2^{ii,D}$ and $s_2^{ii,N}$. Without the second term in Eqs. (128), (129), (131), and (132), the three sets of bounds will coincide.

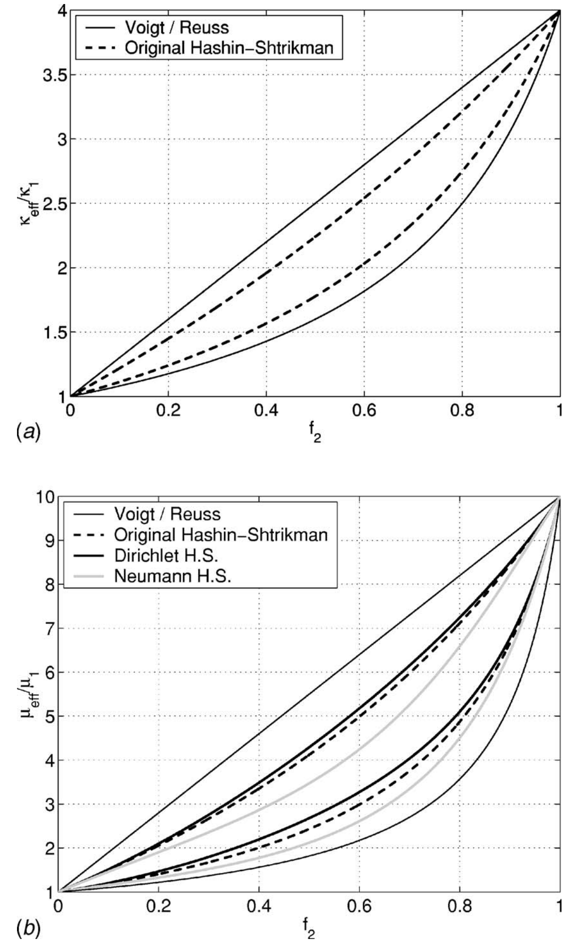


Fig. 9 Improved Hashin–Shtrikman bounds for the effective bulk and shear moduli

Remark 6.2. There is a difference between material ordering, i.e., $\kappa_1 \leqslant \kappa_2 \leqslant \cdots \leqslant \kappa_n$ and geometric ordering, i.e., concentric spherical shells $r_1 \leqslant r_2 \leqslant \cdots \leqslant r_n$. Since one does not necessarily place the phase with the smallest material constants in the inner most region of the RVE, the combination of mappings between material ordering and geometric ordering is multiple. There are differences in the homogenization results due to these different combinations.

For a two-phase composite, there are two ways to place the phase which is not the comparison phase in an RVE: either in the interior of the RVE or in the exterior of the RVE. By alternating the material phase from the interior region of the RVE to the exterior region of the RVE, the interior homogenization becomes the exterior homogenization, and they correspond to different finite Eshelby tensors as seen in Sec. 4. Therefore, in principle, we can obtain for each boundary condition two distinct pairs of the variational bounds, namely one corresponding to the interior eigenstrain and one corresponding to the exterior eigenstrain method (see Fig. 5 illustrating the different combinations possible for each boundary condition). For isotropic composites, alternating the phase position has no effect on the variational bounds for the bulk modulus, because the bulk part of the interior eigenstrain Eshelby tensor equals the bulk part of the exterior eigenstrain Eshelby tensor and thus the two pairs of bounds coincide.

On the other hand, for the shear modulus, alternating the phase position yields new variational bounds. (These are not shown in Fig. 9 since they will only deviate slightly from the bounds shown in the figure.) Altogether we have two pairs of distinct variational bounds for the shear modulus under each boundary condition.

For multiphase composites ($n \geqslant 3$), the dependence on phase position may become more pronounced.

6.2 Three-Phase Composites. Consider a three-phase isotropic composite with $\kappa_3 > \kappa_2 > \kappa_1$ and $\mu_3 > \mu_2 > \mu_1$. To obtain the lower bound, we choose $\kappa_0 = \kappa_1$ and $p_1 = 0$. One can then solve the stationarity condition Eq. (120) for p_2 and p_3

$$p_2 = 3\bar{\epsilon}\underline{p}_2, \quad \underline{p}_2 = \frac{1}{\Delta_{\ell 1}}\left(\frac{s_1^{33,D} - 0.5 s_1^{32,D}}{\kappa_1} + \frac{1}{\kappa_3 - \kappa_1}\right) \tag{133}$$

$$p_3 = 3\bar{\epsilon}\underline{p}_3, \quad \underline{p}_3 = \frac{1}{\Delta_{\ell 1}}\left(\frac{s_1^{22,D} - 0.5 s_1^{23,D}}{\kappa_1} + \frac{1}{\kappa_2 - \kappa_1}\right) \tag{134}$$

where

$$\Delta_{\ell 1} = \left(\frac{s_1^{22,D}}{\kappa_1} + \frac{1}{\kappa_2 - \kappa_1}\right)\left(\frac{s_1^{33,D}}{\kappa_1} + \frac{1}{\kappa_3 - \kappa_1}\right) - \frac{s_1^{32,D} s_1^{23,D}}{4\kappa_1^2} \tag{135}$$

Similarly, one can solve Eq. (120) for the stationary values of p_1 and p_2 for the upper bound by setting $\kappa_0 = \kappa_3$ and $p_3 = 0$, i.e.

$$p_1 = 3\bar{\epsilon}\bar{p}_1, \quad \bar{p}_1 = \frac{1}{\Delta_{u1}}\left(\frac{s_1^{22,D} - 0.5 s_1^{21,D}}{\kappa_3} + \frac{1}{\kappa_2 - \kappa_3}\right) \tag{136}$$

$$p_2 = 3\bar{\epsilon}\bar{p}_2, \quad \bar{p}_2 = \frac{1}{\Delta_{u1}}\left(\frac{s_1^{11,D} - 0.5 s_1^{12,D}}{\kappa_3} + \frac{1}{\kappa_1 - \kappa_3}\right) \tag{137}$$

where

$$\Delta_{u1} = \left(\frac{s_1^{11,D}}{\kappa_3} + \frac{1}{\kappa_1 - \kappa_3}\right)\left(\frac{s_1^{22,D}}{\kappa_3} + \frac{1}{\kappa_2 - \kappa_3}\right) - \frac{s_1^{12,D} s_1^{21,D}}{4\kappa_3^2} \tag{138}$$

Substituting the stationary values Eqs. (133), (134), (136), and (137), into the Hashin–Shtrikman variational principle Eq. (107), we find the explicit variational bounds of the bulk modulus for three-phase composites

$$\begin{aligned}
&\kappa_1 - \left[\frac{f_2 \underline{p}_2^2}{(\kappa_2 - \kappa_1)} + \frac{f_3 \underline{p}_3^2}{(\kappa_3 - \kappa_1)}\right] - \frac{1}{\kappa_1}(f_2 \underline{p}_2^2 s_1^{22,D} + f_2 \underline{p}_2 \underline{p}_3 s_1^{32,D} \\
&\quad + f_3 \underline{p}_2 \underline{p}_3 s_1^{23,D} + f_3 \underline{p}_3^2 s_1^{33,D}) + 2(f_2 \underline{p}_2 + f_3 \underline{p}_3) \leqslant \bar{\kappa} \leqslant \kappa_3 \\
&\quad - \left[\frac{f_1 \bar{p}_1^2}{(\kappa_1 - \kappa_3)} + \frac{f_2 \bar{p}_2^2}{(\kappa_2 - \kappa_3)}\right] - \frac{1}{\kappa_3}(f_1 \bar{p}_1^2 s_1^{11,D} + f_1 \bar{p}_1 \bar{p}_2 s_1^{21,D} \\
&\quad + f_2 \bar{p}_1 \bar{p}_2 s_1^{12,D} + f_2 \bar{p}_2^2 s_1^{22,D}) + 2(f_1 \bar{p}_1 + f_2 \bar{p}_2)
\end{aligned} \tag{139}$$

Similarly, for the bounds of the shear modulus we have

$$\begin{aligned}
&\mu_1 - \left[\frac{f_2 \underline{\tau}_2^2}{(\mu_2 - \mu_1)} + \frac{f_3 \underline{\tau}_3^2}{(\mu_3 - \mu_1)}\right] - \frac{1}{\mu_1}(f_2 \underline{\tau}_2^2 s_2^{22,D} + f_2 \underline{\tau}_2 \underline{\tau}_3 s_2^{32,D} \\
&\quad + f_3 \underline{\tau}_2 \underline{\tau}_3 s_2^{23,D} + f_3 \underline{\tau}_3^2 s_2^{33,D}) + 2(f_2 \underline{\tau}_2 + f_3 \underline{\tau}_3) \leqslant \bar{\mu} \leqslant \mu_3 \\
&\quad - \left[\frac{f_1 \bar{\tau}_1^2}{(\mu_1 - \mu_3)} + \frac{f_2 \bar{\tau}_2^2}{(\mu_2 - \mu_3)}\right] - \frac{1}{\mu_3}(f_1 \bar{\tau}_1^2 s_2^{11,D} + f_1 \bar{\tau}_1 \bar{\tau}_2 s_2^{21,D} \\
&\quad + f_2 \bar{\tau}_1 \bar{\tau}_2 s_2^{12,D} + f_2 \bar{\tau}_2^2 s_2^{22,D}) + 2(f_1 \bar{\tau}_1 + f_2 \bar{\tau}_2)
\end{aligned} \tag{140}$$

where

$$\underline{\tau}_2 = \frac{1}{\Delta_{\ell 2}}\left(\frac{s_2^{33,D} - 0.5 s_2^{32,D}}{\mu_1} + \frac{1}{\mu_3 - \mu_1}\right) \tag{141}$$

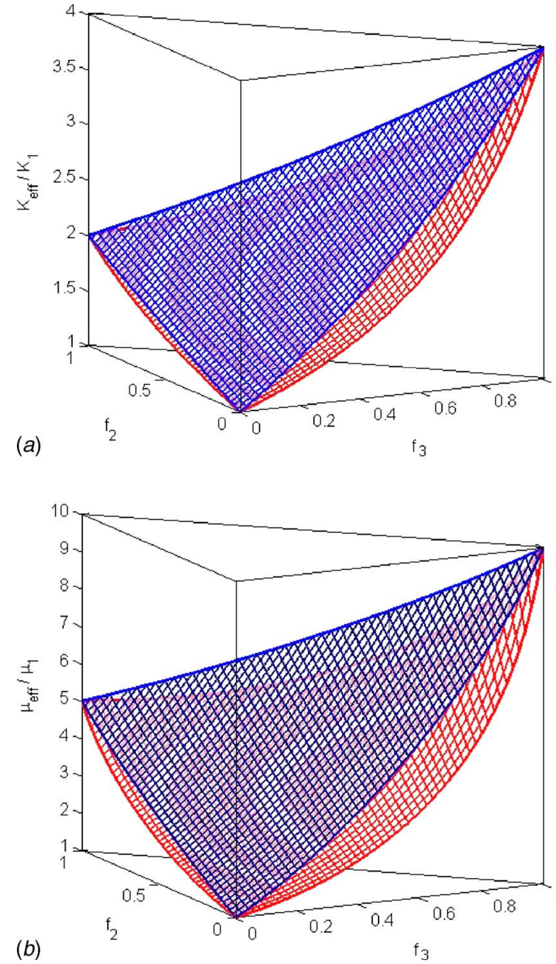


Fig. 10 Variational bounds for a three-phase composite material: (*a*) bounds for bulk modulus; and (*b*) bounds for shear modulus.

$$\underline{\tau}_3 = \frac{1}{\Delta_{\ell 2}}\left(\frac{s_2^{22,D} - 0.5 s_2^{23,D}}{\mu_1} + \frac{1}{\mu_2 - \mu_1}\right) \tag{142}$$

$$\bar{\tau}_1 = \frac{1}{\Delta_{u2}}\left(\frac{s_2^{22,D} - 0.5 s_2^{21,D}}{\mu_3} + \frac{1}{\mu_2 - \mu_3}\right) \tag{143}$$

$$\bar{\tau}_2 = \frac{1}{\Delta_{u2}}\left(\frac{s_2^{11,D} - 0.5 s_2^{12,D}}{\mu_3} + \frac{1}{\mu_1 - \mu_3}\right) \tag{144}$$

and

$$\Delta_{\ell 2} = \left(\frac{s_2^{22,D}}{\mu_1} + \frac{1}{\mu_2 - \mu_1}\right)\left(\frac{s_2^{33,D}}{\mu_1} + \frac{1}{\mu_3 - \mu_1}\right) - \frac{s_2^{23,D} s_2^{32,D}}{4\mu_1^2} \tag{145}$$

$$\Delta_{u2} = \left(\frac{s_2^{11,D}}{\mu_3} + \frac{1}{\mu_1 - \mu_3}\right)\left(\frac{s_2^{22,D}}{\mu_3} + \frac{1}{\mu_2 - \mu_3}\right) - \frac{s_2^{12,D} s_2^{21,D}}{4\mu_3^2} \tag{146}$$

Figure 10 shows the variational bounds for the effective bulk and shear modulus of a three-phase composite using the modulus ratios $\kappa_3{:}\kappa_2{:}\kappa_1 = 4{:}2{:}1$, $\mu_3{:}\mu_2{:}\mu_1 = 10{:}5{:}1$ and Poisson's ratio $\nu_1 = 0.3$. The unique features of variational bounds (139) and (140) are: (1) the boundary conditions are accurately taken into account without resorting to any approximation and ad hoc arguments; (2) interaction among different phases, or in other words, the correlation among different phases are precisely taken into account by

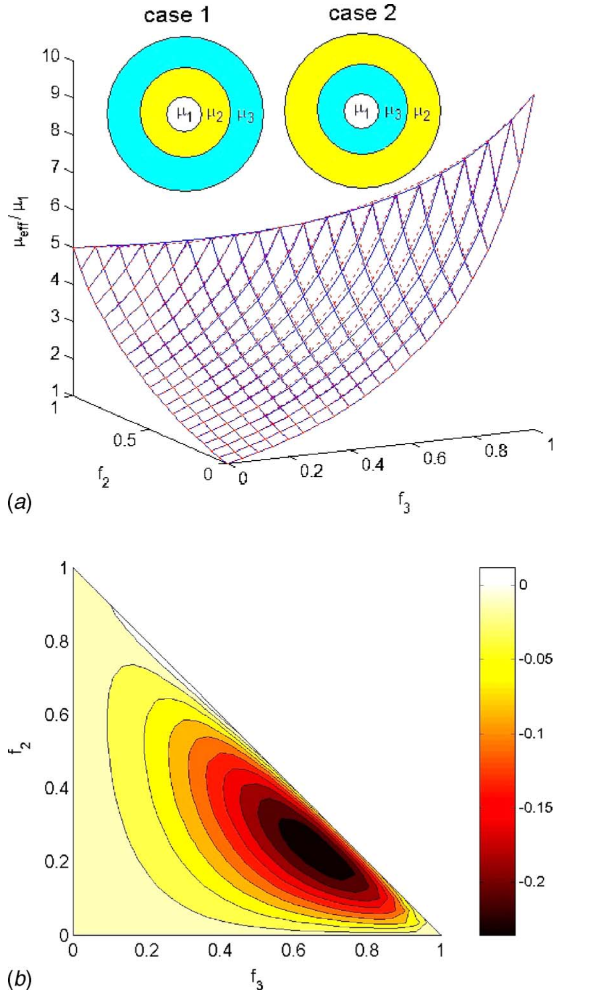


Fig. 11 Influence of phase position on three-phase variational bounds

the cross-term Eshelby tensor $\mathbb{S}^{ij,D}$, $i \neq j$. This feature is absent in the classical HS bounds; (3) Microstructures of the composite are distinguished by mapping different combinations of the geometric ordering to the material ordering. For the bounds shown in Fig. 10, the geometric ordering coincides with the material ordering in ascending order, i.e., $(\kappa_1,\mu_1) \Rightarrow \Omega_1$, $(\kappa_2,\mu_2) \Rightarrow \Omega_2$, and $(\kappa_3,\mu_3) \Rightarrow \Omega_3$.

To examine the effect of the microstructure on the variational bounds, we exchange the material ordering within the domains Ω_2 and Ω_3. Figure 11(*a*) shows a plot of the two lower bound surfaces of the shear modulus. The contour of the difference is shown in Fig. 11(*b*). One can see that the maximum difference is about 0.2, demonstrating the the material ordering has little impact for this case.

7 Closure

In this paper, the finite Eshelby tensors obtained in Part I of our work are applied to develop various homogenization methods. It is shown that the special features of the finite Eshelby tensors can improve the accuracy of conventional homogenization methods and lead to more accurate predictions on effective material properties of composites.

For instance, we have found that for two-phase composites, there are at least two sets of Hashin–Shtrikman variational bounds corresponding to two different boundary conditions. This discovery may be instrumental for numerical homogenization procedures.

Furthermore, we have developed some new homogenization schemes such as the exterior eigenstrain method, dual eigenstrain method, i.e., a generalized self-consistency method, the shell model, and multiphase Hashin–Shtrikman bounds, which will enrich the engineering homogenization repertoire and provide sharper estimates on effective material properties of multi-phase composites.

The applications of the finite Eshelby tensor are multitude, and they are not limited to applications of homogenization theory. As indicated by the multilayer shell model, the finite Eshelby tensors provide the basic module to construct the multi-inclusion model and interface model, which can be used in modeling quantum dots, nano-onions, spinodal decomposition, and functionally graded materials. Some of these studies will be reported in separate papers.

Acknowledgment

This work is supported by a grant from National Science Foundation (Grant No. CMS-0239130), which is greatly appreciated.

Appendix: Table of the Eshelby Coefficients for the Three-Layer Shell Model

In this Appendix, a complete list of the coefficients for the average Eshelby tensors of a three-sphere RVE is documented. The notation of these coefficients is explained and defined in Sec. 5

$$s_1^{I1,D} = \frac{(1+\nu)(1-f_1)}{3(1-\nu)}$$

$$s_1^{I1,N} = \frac{(1+\nu)+2(1-2\nu)f_1}{3(1-\nu)}$$

$$s_2^{I1,D} = \frac{2(4-5\nu)(1-f_1)}{15(1-\nu)} - 21\gamma_u[f_1](1-f_1^{2/3})$$

$$s_2^{I1,N} = \frac{2(4-5\nu)+(7-5\nu)f_1}{15(1-\nu)} + 21\gamma_t[f_1](1-f_1^{2/3})$$

$$s_1^{I2,D} = -\frac{(1+\nu)f_1}{3(1-\nu)}$$

$$s_1^{I2,N} = \frac{2(1-2\nu)f_1}{3(1-\nu)}$$

$$s_2^{I2,D} = -\frac{2(4-5\nu)f_1}{15(1-\nu)} - 21\gamma_u[f_1]\left[1-\frac{(f_1+f_2)^{5/3}-f_1^{5/3}}{f_2}\right]$$

$$s_2^{I2,N} = \frac{(7-5\nu)f_1}{15(1-\nu)} + 21\gamma_t[f_1]\left[1-\frac{(f_1+f_2)^{5/3}-f_1^{5/3}}{f_2}\right]$$

$$s_1^{I3,D} = -\frac{(1+\nu)f_1}{3(1-\nu)}$$

$$s_1^{I3,N} = \frac{2(1-2\nu)f_1}{3(1-\nu)}$$

$$s_2^{I3,D} = -\frac{2(4-5\nu)f_1}{15(1-\nu)} + 21\gamma_u[f_1]\frac{(f_1+f_2)[1-(f_1+f_2)^{2/3}]}{f_3}$$

$$s_2^{I3,N} = \frac{(7-5\nu)f_1}{15(1-\nu)} - 21\gamma_t[f_1]\frac{(f_1+f_2)[1-(f_1+f_2)^{2/3}]}{f_3}$$

$$s_1^{II1,D} = \frac{(1+\nu)f_3}{3(1-\nu)}$$

$$s_1^{II1,N} = \frac{(1+\nu)+2(1-2\nu)(f_1+f_2)}{3(1-\nu)}$$

$$s_2^{II1,D} = \frac{2(4-5\nu)f_3}{15(1-\nu)} - 21\gamma_u(f_1+f_2)(1-f_1^{2/3})$$

$$s_2^{II1,N} = \frac{2(4-5\nu)+(7-5\nu)(f_1+f_2)}{15(1-\nu)} + 21\gamma_t(f_1+f_2)(1-f_1^{2/3})$$

$$s_1^{II2,D} = \frac{(1+\nu)f_3}{3(1-\nu)}$$

$$s_1^{II2,N} = \frac{(1+\nu)+2(1-2\nu)(f_1+f_2)}{3(1-\nu)}$$

$$s_2^{II2,D} = \frac{2(4-5\nu)}{15(1-\nu)}f_3 - 21\gamma_u(f_1+f_2)\left[1-\frac{(f_1+f_2)^{5/3}-f_1^{5/3}}{f_2}\right]$$

$$s_2^{II2,N} = \frac{2(4-5\nu)+(7-5\nu)(f_1+f_2)}{15(1-\nu)} + 21\gamma_t(f_1+f_2)\left[1-\frac{(f_1+f_2)^{5/3}-f_1^{5/3}}{f_2}\right]$$

$$s_1^{II3,D} = -\frac{(1+\nu)(f_1+f_2)}{3(1-\nu)}$$

$$s_1^{II3,N} = \frac{2(1-2\nu)(f_1+f_2)}{3(1-\nu)}$$

$$s_2^{II3,D} = -\frac{2(4-5\nu)}{15(1-\nu)}(f_1+f_2) + 21\gamma_u(f_1+f_2)\frac{(f_1+f_2)[1-(f_1+f_2)^{2/3}]}{f_3}$$

$$s_2^{II3,N} = \frac{7-5\nu}{15(1-\nu)}(f_1+f_2) - 21\gamma_t(f_1+f_2)\frac{(f_1+f_2)[1-(f_1+f_2)^{2/3}]}{f_3}$$

$$s_1^{III3,D} = 0$$

$$s_1^{III3,N} = 1$$

$$s_2^{III3,D} = 0$$

$$s_2^{III3,N} = 1$$

$$s_1^{III1,D} = 0$$

$$s_1^{III1,N} = 1$$

$$s_2^{III1,D} = 0$$

$$s_2^{III1,N} = 1$$

$$s_1^{III2,D} = 0$$

$$s_1^{III2,N} = 1$$

$$s_2^{III2,D} = 0$$

$$s_2^{III2,N} = 1$$

where

$$\gamma_u[x] := \frac{x(1-x^{2/3})}{10(1-\nu)(7-10\nu)}$$

and

$$\gamma_t[x] := \frac{4x(1-x^{2/3})}{10(1-\nu)(7+5\nu)} \tag{A1}$$

References

[1] Li, S., Sauer, R. A., and Wang, G., 2006, "The Eshelby Tensors in a Finite Spherical Domain—Part I: Theoretical Formulations," ASME J. Appl. Mech., **74**, pp. 770–783.

[2] Calvert, P., 1999, "Nanotube Composites: A Recipe for Strength," Nature (London), **399**, pp. 210–211.

[3] Thostenson, E. T., Ren, Z., and Chou, T. W., 2001, "Advances in Science and Technology of Carbon Nanotubes and Their Composites: A Review," Compos. Sci. Technol., **61**, pp. 1899–1912.

[4] Fisher, F. T., Bradshaw, R. D., and Brinson, L. C., 2002, "Effects of Nanotube Waviness on the Modulus of Nanotube-Reinforced Polymers," Appl. Phys. Lett., **80**, pp. 4647–4649.

[5] Odegard, G. M., Gates, T. S., Wise, K. E., Park, C., and Siochi, E. J., 2003, "Constitutive Modeling of Nanotube-Reinforced Polymer Composites," Compos. Sci. Technol., **63**, pp. 1671–1687.

[6] Shi, D.-L., Feng, X.-Q., Huang, Y.-Y., Hwang, K.-C., and Gao, H., 2004, "The Effect of Nanotube Waviness and Agglomeration on the Elastic Property of Carbon Nanotube-Reinforced Composites," J. Eng. Mater. Technol., **126**, pp. 250–257.

[7] Sharma, P., and Ganti, S., 2004, "Size-Dependent Eshelby's Tensor for Embedded Nano-Inclusions Incorporating Surface/Interface Energies," ASME J. Appl. Mech., **71**, pp. 663–671.

[8] Ovid'ko, I. A., and Scheinerman, A. G., 2005, "Elastic Fields of Inclusion in Nanocomposite Solids," Rev. Adv. Mater. Sci., **9**, pp. 17–33.

[9] Tanaka, K., and Mori, T., 1972, "Note on Volume Integrals of the Elastic Field Around an Ellipsoidal Inclusion," J. Elast., **2**, pp. 199–200.

[10] Nemat-Nasser, S., and Hori, M., 1999, *Micromechanics: Overall Properties of Heterogeneous Materials*, Elsevier, Amsterdam, The Netherlands.

[11] Mori, T., and Tanaka, K., 1973, "Average Stress in Matrix and Average Elastic Energy of Materials with Misfiting Inclusion," Acta Metall., **21**, pp. 571–574.

[12] Sauer, R. A., Wang, G., and Li, S., 2007, "The Composite Eshelby Tensors and Their Application to Homogenization," submitted.

[13] Castles, R. R., and Mura, T., 1985, "The Analysis of Eigenstrains Outside of an Ellipsoidal Inclusion," J. Elast., **15**, pp. 27–34.

[14] Willis, J. R., 1981, "Variational and Related Methods for the Overall Properties of Composites," *Advances in Applied Mechanics*, C. S. Vih, ed., Academic, New York, Vol. 21, pp. 1–78.

[15] Hashin, Z., and Shtrikman, S., 1962, "On Some Variational Principles in Anisotropic and Nonhomogeneous Elasticity," J. Mech. Phys. Solids, **10**, pp. 335–342.

[16] Hashin, Z., and Shtrikman, S., 1962, "A Variational Approach to the Theory of the Elastic Behavior of Polycrystals," J. Mech. Phys. Solids, **10**, pp. 343–352.

Stochastic Finite Element Buckling Analysis of Laminated Plates With Circular Cutout Under Uniaxial Compression

A. K. Onkar

C. S. Upadhyay

D. Yadav[1]
e-mail: dy@iitk.ac.in

Department of Aerospace Engineering,
Indian Institute of Technology,
Kanpur 208016, India

A generalized stochastic buckling analysis of laminated composite plates, with and without centrally located circular cutouts having random material properties, is presented under uniaxial compressive loading. In this analysis, the layerwise plate model is used to solve both prebuckling and buckling problems. The stochastic analysis is done based on mean centered first-order perturbation technique. The mean buckling strength of composite plates is validated with results available in the literature. It has been observed that the present analysis can predict buckling load accurately even for plates with large cutouts. Micromechanics based approach is used to study the effect of variation in microlevel constituents on the effective macrolevel properties like elastic moduli. Consequently, the effect of uncertainty in these material properties on the buckling strength of the laminated plates is studied. Parametric studies are carried out to see the effect of hole size, layups, and boundary conditions on the mean and variance of plate buckling strength.
[DOI: 10.1115/1.2711230]

Keywords: stochastic buckling analysis, layerwise plate model, laminated composite plate, centrally located cutout, random material properties

1 Introduction

Buckling behavior of laminated composite plates subjected to in-plane loads is an important consideration in the preliminary design of aircraft and launch vehicle components. The sizing of many structural subcomponents of these vehicles is often determined by stability constraints. Plates with circular holes and other openings are extensively used as structural members in aircraft design. The buckling behavior of such plates has always received much attention by researchers. These holes can be access holes, holes for hardware to pass through, or in the case of fuselage, windows and doors. In some cases holes are used to reduce the weight of the structure. In aerospace and many other applications these structural components are also made up of composite material to further reduce the weight of the structure. The outstanding mechanical properties of composite structures, such as durability and corrosion-resistance characteristics combined with low density, make it more attractive compared to conventional materials. The mechanical properties of composite materials depend on a wide variety of variables at the microlevel, for example, the fiber and matrix material properties; and fabrication variables at all stages of the fabrication process, such as fiber volume ratio, misalignment of ply orientation, fiber waviness or undulation, intralamina voids, incomplete curing of resin, excess resin between plies, and variation in ply thickness. These variables are statistical in nature, therefore, the mechanical properties of a composite material should be quantified probabilistically. The influence of these microlevel variables on the macrolevel effective elastic properties of composites has been studied both experimentally [1] and numerically using various methods [2–5]. It has been reported that the coefficient of variation (COV) in the macrolevel effective elastic moduli of lamina could be up to 15% depending on the number of microlevel random variables taken into the analysis. The variations in macrolevel effective properties ultimately lead to the variation in the response of the structure. Mechanical problems of such composite structures are solved by the use of one of the following computational methodologies: the stochastic finite element methods (SFEM) [6,7], stochastic spectral techniques [8], or the Monte Carlo simulation (MCS) approach [9,10]. The present paper applies the first order perturbation technique based stochastic finite element method.

Any stochastic problem can be split into: (a) one set of mean problem and (b) R set of random problems. These can be solved separately to determine the mean and the variance of the response of our interest. Here R denotes the number of primary random variables chosen for the analysis at the macrolevel.

The first part, i.e., the mean stability analysis of plates with cutouts has been studied by many investigators using various methods under various inplane loadings. When a plate contains a hole, it is known that the mean tensile strength is reduced due to the stress concentration around the hole. However, the mean buckling behavior of a plate with a hole under compressive loading is quite different. For the mean buckling problem the accurate prebuckled stresses must first be found, and then the stability of the plate in the presence of the prebuckled stresses studied. The former problem is an in-plane problem, like the tensile problem, and mainly involves in-plane stiffness. The stability problem, however, involves both in-plane and out-of-plane effects, in particular, bending stiffness. Thus the mechanics of the mean buckling problem is much more complicated than the mechanics of the tensile problem. Numerical investigation of buckling behavior of both isotropic and composite plates with centrally located cutouts have been studied by various methods and plate models using the finite element method [11–16]. The effect of various shape of cutout like, square, circular, elliptical, its location (eccentricity), and the size of cutout on the mean buckling load has also been reported in Refs. [17–20]. There are also few researchers who have investigated the mean buckling behavior of composite plates experimentally [13,14,20,21].

The second part, i.e., the statistics of buckling load of compos-

[1]Corresponding author.
Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received March 6, 2006; final manuscript received September 19, 2006. Review conducted by Edmundo Corona.

ite plates has not been given much attention by researchers. However, there are few studies which talk about the statistics of buckling load of composite plates without hole [22–25]. The material properties, fiber angles, laminate thickness, and different in-plane random loads are treated as basic random variables (BRVs). Both SFEM and MCS have been used to quantify the structural response uncertainty. An overview of different stochastic methods has been presented for solving classical problems of solid mechanics in Ref. [26] with an elementary illustration of the perturbation based stochastic finite element method. It has also been shown that the MCS method is the most general approach and can be applied to solve any stochastic problem. However, a drawback of the method is that it requires solution of a large set of deterministic problems corresponding to generated samples of the differential operator and input. The effect of uncertainty in the initial geometric imperfections (like radius, thickness) becomes important in the stochastic buckling analysis of shells. The stochastic finite element analysis of shell structures without any cutout has been performed for both uncertain material and geometric properties in conjunction with MCS to obtain the response variability [27,28]. These properties were assumed to be described by uncorrelated two-dimensional homogeneous stochastic fields. The buckling behavior of composite plates with cutout having random material properties has not been addressed adequately.

In this paper the variations in the macrolevel material properties are related to the scatter in the microlevel constituent parameters. A stochastic finite element formulation based on mean centered first-order perturbation technique is presented for the buckling analysis of laminated composite plates, with uncertain material properties, having a circular cutout. The mean buckling analysis is done by first solving the linear elastic problem to get accurate prebuckling stresses and then these stresses are used to solve the generalized eigenvalue problem for the lowest eigenvalue. The Green's strain–displacement relation is used in the formulation to include the complete geometric nonlinearity effect. A layerwise plate model is used to predict the state of prebuckled stresses in the plate. This approach can be used to study any layered structure: symmetric, antisymmetric or unsymmetric. The validation of the present mean analysis is performed by comparing the results with those reported in the literature. The statistics of the buckling strength is determined by considering uncertainties in the effective material properties of composite laminates. The validation of the buckling strength statistics for plates without cutout is performed by comparing the results with analytical solutions based on the Kirchhoff–Love plate model. A good agreement between the analytical and SFEM solutions has been observed for the problem studied. Typical results for the covariance of buckling strength of laminated plates with cutout are presented. Effects of ply orientations, layup sequences, and hole size along with change in standard deviation of material properties have also been investigated for different boundary conditions.

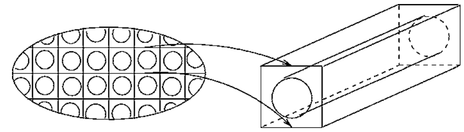


Fig. 1 Periodic fiber reinforced composite

2 Micromechanics Based Approach to Find Scatter in Material Properties

The properties of composites display considerable scatter because of the uncertainties involved at many levels—properties of the constituents, fabrication and manufacturing processes, geometrical parameters of laminates, fiber orientations, volume fraction, etc. It is not possible to control variations in all these parameters completely and thus scatter in geometric and effective material properties is inherent. In the present study only uncertainties in the effective material properties are taken, which are calculated by assuming variations in the properties of the microlevel constituents.

Let us consider the representative volume element (RVE), shown in Fig. 1, as a volume of the material that exhibits statistically homogeneous material properties. Further it is assumed that the composite is periodic in a random sense, i.e., in each RVE the variation in the microlevel material properties remain same. To study the effect of variation in microlevel constituents, i.e., elastic properties of fiber (E_f, ν_f) and matrix phase (E_m, ν_m) and volume concentration of fiber phase (V_f), on the effective properties of fiber reinforced material in a RVE, a composite cylinder model based homogenization approach is used. The analytical expressions for the effective macrolevel properties in terms of the microlevel properties for perfectly bonded thin orthotropic or transversly isotropic layers have been adopted from Ref. [29]. A first-order perturbation technique is used to study the effect of scatter in microlevel constituents on the variations in the macrolevel effective elastic moduli. Table 1 brings out the influence of 5% COV in all five basic microlevel constituents, namely E_f, ν_f, E_m, ν_m, and V_f, changing simultaneously on COV in the effective moduli of various fiber-reinforced composite systems. In the present investigation, four different composite systems: graphite–, boron–, carbon–, and glass–epoxy, are chosen. From the table it is observed that:

1. The COV in E_{ll} is found to be $\approx 7\%$ for all types of composite systems; and
2. The COV in G_{lt} is found to be a maximum of $\approx 12\%$, whereas G_{tt} has minimum variation of $\approx 3\%$. The variation in E_{tt}, ν_{lt}, and ν_{tt} is found to be $\approx 4\%$.

These variations in the effective macrolevel properties will finally lead to variation in the buckling load of the system which is discussed in the next section. Based on the experimental results [23]

Table 1 Effect of COV of 5% in all microlevel properties changing simultaneously on the various macrolevel effective material properties for different composite systems

Different composite system	COV in macrolevel effective material properties (%)					
	E_{ll}	E_{tt}	ν_{lt}	ν_{tt}	G_{lt}	G_{tt}
Graphite–epoxy	6.97	3.88	3.42	3.30	11.68	2.68
Boron epoxy	7.00	3.95	3.38	3.42	11.92	2.69
Carbon–epoxy	6.96	3.92	3.39	3.40	11.59	2.68
Glass–epoxy	6.72	3.71	3.42	3.26	10.22	2.62

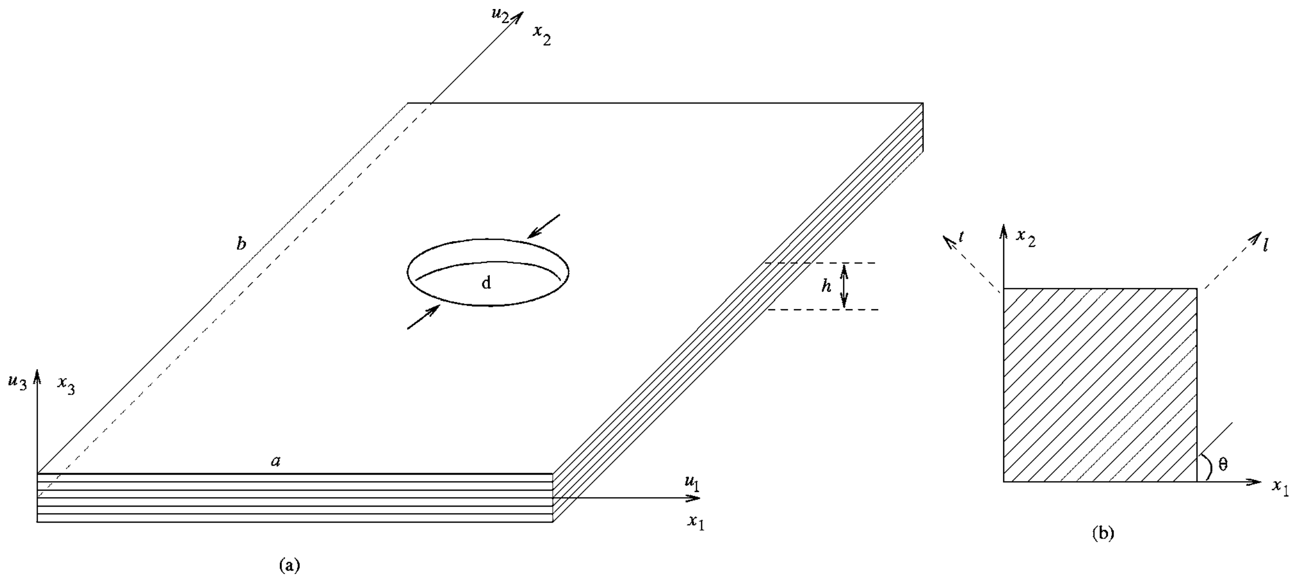


Fig. 2 Geometry of a laminated composite plate with centrally located cutout

it will further be assumed that there are no spatial variations in the effective moduli.

3 Stochastic Finite Element Buckling Formulation

A rectangular composite laminated plate, having a centrally located cutout, with its coordinate definitions and material direction of a typical lamina are shown in Fig. 2. It is assumed that the laminated plate is composed of perfectly bonded thin orthotropic (or transversely isotropic) layers. Due to the applied in-plane loads, the Green's strain produced can be represented as

$$\{\varepsilon\} = \{\varepsilon^L\} + \{\varepsilon^{\mathrm{NL}}\}; \quad \varepsilon_{ij}^L = \frac{1}{2}\left\{\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right\} \tag{1}$$

and

$$\varepsilon_{ij}^{\mathrm{NL}} = \frac{1}{2}\left\{\frac{\partial u_k}{\partial x_i}\frac{\partial u_k}{\partial x_j}\right\} \quad (i,j,k = 1,2,3)$$

where $\{\varepsilon^L\}$ and $\{\varepsilon^{\mathrm{NL}}\}$ are the linear and the geometric nonlinear strain, respectively; and u_i is the components of displacement fields.

Buckling can be modeled as a linearized stability analysis of the geometrically nonlinear elasticity problem by assuming that the prebuckling deformations are small. In the present formulation the plate is assumed to be linearly elastic with a stochastic elasticity tensor field C_{ijkl}. Let the plate with random material properties be subjected to reference in-plane load q_i^{ref}. The total potential energy corresponding to the linear state of the system for uncertain stiffness can be written as [6]

$$\Pi(u^{\mathrm{ref}}) = \frac{1}{2}\int_\Omega C_{ijkl}\varepsilon_{ij}^{\mathrm{ref}}\varepsilon_{kl}^{\mathrm{ref}} d\Omega - \int_{\Gamma_1} q_i^{\mathrm{ref}} u_i^{\mathrm{ref}}\, d\Gamma_1 \quad (i,j,k,l = 1,2,3) \tag{2}$$

where Ω denotes the undeformed configuration of the plate and its boundary is denoted by $\Gamma = \Gamma_0 \cup \Gamma_1$. Γ_0 denotes the Dirichlet part and Γ_1 denotes the Neumann part of the lateral boundary of the plate. $\varepsilon_{ij}^{\mathrm{ref}}$ denotes the linear strain tensor defined by Eq. (1) and q_i^{ref} is the boundary traction. In the present analysis the reference loads $q_1^{\mathrm{ref}} = q^{\mathrm{ref}}$ and $q_2^{\mathrm{ref}} = q_3^{\mathrm{ref}} = 0$ are taken.

The exact solution minimizes Π on the set of all kinematically admissible functions denoted by V, i.e., $u \in V$ such that $V = \{u \in H^1(\Omega) : M(u) = 0 \text{ on } \Gamma_0\}$. This yields

$$\delta\Pi(u^{\mathrm{ref}}) = \int_\Omega C_{ijkl}\varepsilon_{ij}^{\mathrm{ref}}\delta\varepsilon_{kl}^{\mathrm{ref}}\, d\Omega - \int_{\Gamma_1} q_i^{ref}\delta u_i^{\mathrm{ref}}\, d\Gamma_1 = 0 \quad (i,j,k,l = 1,2,3) \tag{3}$$

Using Taylor series expansion based perturbation approach, the zeroth- and first-order equations can be written as follows [6]:

zeroth order

$$\int_\Omega C_{ijkl}^0 \varepsilon_{ij}^{0(\mathrm{ref})}\delta\varepsilon_{kl}^{0(\mathrm{ref})}\, d\Omega = \int_{\Gamma_1} q_i^{0(\mathrm{ref})}\delta u_i^{0(\mathrm{ref})}\, d\Gamma_1 \quad (i,j,k,l = 1,\ldots,3) \tag{4a}$$

first order

$$\int_\Omega C_{ijkl}^0 \varepsilon_{ij}^{,r(\mathrm{ref})}\delta\varepsilon_{kl}^{0(\mathrm{ref})}\, d\Omega = \int_{\Gamma_1} q_i^{,r(\mathrm{ref})}\delta u_i^{0(\mathrm{ref})}\, d\Gamma_1 - \int_\Omega C_{ijkl}^{,r}\varepsilon_{ij}^{0(\mathrm{ref})}\delta\varepsilon_{kl}^{0(\mathrm{ref})}\, d\Omega \quad (i,j,k,l = 1,\ldots,3; r = 1,2,\ldots,R) \tag{4b}$$

where R is the number of BRVs chosen for the analysis.

Having solved Eq. (3), for linear displacements u_i^{ref} due to the reference load q^{ref} (taken equal to unity), we look for buckling load parameter λ such that the system goes to a new equilibrium position. Among all λ the critical or minimum value of the load parameter is denoted by λ_{cr} (with total load $q^{\mathrm{cr}} = \lambda_{\mathrm{cr}}\, q^{\mathrm{ref}}$). From this equilibrium position, we perturb the system by an amount u_i^p (with strain ε_{ij}^p) such that the system retains the new position of equilibrium (neutral stability). The total potential energy of the perturbed system can be written as

$$\Pi(u^p) = \int_\Omega \sigma_{ij}^{\mathrm{cr}}\varepsilon_{ij}^p\, d\Omega + \frac{1}{2}\int_\Omega \sigma_{ij}^p\varepsilon_{ij}^p\, d\Omega - \int_{\Gamma_1} q_i^{\mathrm{cr}} u_i^p\, d\Gamma_1 \quad (i,j = 1,2,3) \tag{5}$$

where $\sigma_{ij}^{\mathrm{cr}}$ is the current linear state of stress (prebuckled) due to

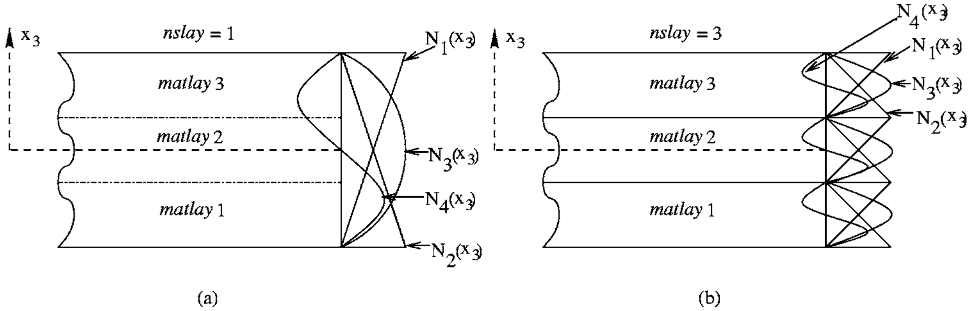


Fig. 3 Representation of transverse function over the thickness of plate: (*a*) equivalent layer; and (*b*) layer by layer

the critical load q^{cr}, and is given by $\sigma_{ij}^{\mathrm{cr}}=\lambda_{\mathrm{cr}}\,\sigma_{ij}^{\mathrm{ref}}$. Upon substituting the perturbational strain ε_{ij}^{p} from Eq. (1) and linearizing the above equation, we arrive at the following expression

$$\Pi(u^p)=\frac{1}{2}\int_\Omega C_{ijkl}\varepsilon_{ij}^{L(p)}\varepsilon_{kl}^{L(p)}\,d\Omega + \lambda_{\mathrm{cr}}\int_\Omega \sigma_{ij}^{\mathrm{ref}}\varepsilon_{ij}^{\mathrm{NL}(p)}\,d\Omega \quad (i,j,k,l=1,2,3) \tag{6}$$

where $\varepsilon_{kl}^{L(p)}$ is the linear part and $\varepsilon_{kl}^{\mathrm{NL}(p)}$ is the nonlinear part of the perturbational strain ε_{kl}^{p}. Now we find λ_{cr} and $u\in V$, $u\neq 0$ such that Π is minimum. This yields

$$\delta\Pi(u^p)=\int_\Omega C_{ijkl}\varepsilon_{ij}^{L(p)}\delta\varepsilon_{kl}^{L(p)}\,d\Omega+\lambda_{\mathrm{cr}}\int_\Omega \sigma_{ij}^{\mathrm{ref}}\delta\varepsilon_{ij}^{\mathrm{NL}(p)}\,d\Omega = 0 \quad (i,j,k,l=1,2,3) \tag{7}$$

The stochastic finite element based buckling analysis of composite plates having random material properties is preformed based on mean-centered first-order perturbation technique. The zeroth- and first-order variational relations are obtained from the above equation using Taylor series expansion based stochastic variational principles as [6]:

zeroth order

$$\int_\Omega C_{ijkl}^{0}\varepsilon_{ij}^{0(L)}\delta\varepsilon_{kl}^{0(L)}\,d\Omega+\lambda_{\mathrm{cr}}^{0}\int_\Omega \sigma_{ij}^{0(\mathrm{ref})}\delta\varepsilon_{ij}^{0(\mathrm{NL})}\,d\Omega = 0 \quad (i,j,k,l=1,2,3) \tag{8a}$$

first order

$$\int_\Omega (C_{ijkl}^{,r}\varepsilon_{ij}^{0(L)}+C_{ijkl}^{0}\varepsilon_{ij}^{,r(L)})\delta\varepsilon_{kl}^{0(L)}\,d\Omega+\lambda_{\mathrm{cr}}^{0}\int_\Omega \sigma_{ij}^{,r(\mathrm{ref})}\delta\varepsilon_{ij}^{0(\mathrm{NL})}\,d\Omega + \lambda_{\mathrm{cr}}^{,r}\int_\Omega \sigma_{ij}^{0(\mathrm{ref})}\delta\varepsilon_{ij}^{0(\mathrm{NL})}\,d\Omega = 0 \quad (i,j,k,l=1,2,3;r=1,2,\ldots,R) \tag{8b}$$

The symbols $(.)^0$ and $(.)^{,r}$ represent the value of the function and its first-order partial derivative with respect to the BRVs, respectively, evaluated at the mean value of the BRVs. The superscript p has been dropped from the above equations for the sake of brevity.

3.1 Discretization. In the present layerwise finite element discretization of the laminated plate, in addition to the in-plane mesh (x_1-x_2 plane), discretization into solution layers along the thickness direction (x_3) is also performed. Hence the displacement field at any point in the laminate using layerwise plate model is written as [30]

$$\begin{aligned}u_1(x_1,x_2,x_3)&=U_{ij}N_i(x_1,x_2)N_j(x_3)\\ u_2(x_1,x_2,x_3)&=V_{ij}N_i(x_1,x_2)N_j(x_3)\\ u_3(x_1,x_2,x_3)&=W_{ij}N_i(x_1,x_2)N_j(x_3)\end{aligned}$$

$$(i=1,\ldots,n_{x_1x_2};j=1,2,\ldots,n_{x_3}) \tag{9}$$

where $n_{x_3}=\mathrm{nslay}\times p_{x_3}+1$; p_{x_3} denotes the order of approximation in the thickness direction. Here nslay is the total number of solution layers in the laminate. A solution layer, herein, means an entity generated by discretizing the plate along thickness direction, and may not necessarily represent a physical (or material) layer. It may represent a subdivision of a physical layer into sublaminae or consist of number of physical layers lumped into a single solution layer termed as equivalent solution layer (see Figs. 3(*a*) and 3(*b*)). $M_i(x_1,x_2)$ and $N_j(x_3)$ are the two-dimensional (2D) and (1D) Legendre shape functions and U_{ij}, V_{ij}, and W_{ij} are the coefficients of the displacement components. $n_{x_1x_2}$ is the total number of degrees of freedoms in the x_1-x_2 plane. Note that $n_{x_1x_2}$ depends on the order of the in-plane approximation, $p_{x_1x_2}$, and the in-plane mesh. If $N_j(x_3)$ is defined over the thickness of the laminate (i.e., $n_{x_3}=p_{x_3}+1$), the model of Eq. (9) reduces to an "equivalent layer" model, i.e., the transverse functions are taken to be smooth polynomials in terms of x_3 over all the laminae, as shown in Fig. 3(*a*). Here we have taken p_{x_3} to be same for u_1, u_2, and u_3. This is because the buckling problem has both the membrane (the pre-

buckled stress due to in-plane loading) and the bending effect (the buckled mode). To resolve both these effects the transverse order of approximation should be same for all the displacement components [31]. Note that generally the membrane effect is ignored and the buckling analysis is done using an approximation suitable for bending effect only.

Upon substituting the above finite element approximation into Eqs. (8*a*) and (8*b*) and by employing nonlinear Green's strain–displacement relationships the discretized finite element system equations are:

zeroth order

$$(K_{ij}^{0}+\lambda_k^{0}K_{ij}^{(G)0})\Delta_j^{k(0)}=0 \quad (i,j,k=1,2,\ldots,n;\ \text{no sum over } k) \tag{10a}$$

first order

$$(K_{ij}^{0}+\lambda_k^{0}K_{ij}^{(G)0})\Delta_j^{k,r}+(K_{ij}^{,r}+\lambda_k^{0}K_{ij}^{(G),r}+\lambda_k^{,r}K_{ij}^{(G)0})\Delta_j^{k(0)} = 0 \quad (i,j,k=1,2,\ldots,n;r=1,2,\ldots,R;\ \text{no sum over} k) \tag{10b}$$

K_{ij}^{0} and $K_{ij}^{(G)0}$are the mean elastic and geometric stiffness matrices of the structure, respectively. Correspondingly $K_{ij}^{,r}$ and $K_{ij}^{(G),r}$ are the first-order partial derivative of elastic and geometric stiffness matrix, respectively, with respect to the rth BRV. $\Delta_j^{k(0)}$ and $\Delta_j^{k,r}$ represent the kth mean eigenvector and its first-order partial derivatives, having components $\{U_{lm}^0, V_{lm}^0, W_{lm}^0\}$ and $\{U_{lm}^{,r}, V_{lm}^{,r}, W_{lm}^{,r}\}$, respectively. It may be noted that λ_{cr}^0 is the minimum among all λ_i^0.

The zeroth-order equation (Eq. (10a)) is a generalized eigenvalue problem which is solved to get the mean critical buckling load of the system. To obtain the statistics of the critical buckling load, we premultiply both sides of Eq. (10b) by the mean eigenvector Δ_i^0 obtained from Eq. (10a) for minimum mean eigenvalue λ_{cr}^0. This gives

$$\Delta_i^0(K_{ij}^0+\lambda_{cr}^0K_{ij}^{(G)0})\Delta_i^{,r}=-\lambda_{cr}^{,r}(\Delta_i^0K_{ij}^{(G)0}\Delta_j^0)-\Delta_i^0(K_{ij}^{,r}+\lambda_{cr}^0K_{ij}^{(G),r})\Delta_j^0 \tag{11}$$

Since both K_{ij}^0 and $K_{ij}^{(G)0}$are symmetric, the left hand side equals zero by the definition of the zeroth-order equation. By employing $K_{ij}^{(G)0}$ orthonormality conditions, the first term on the right hand side equation reduces to $\lambda_{cr}^{,r}$. The expression for the first-order eigenvalue then takes the following form

$$\lambda_{cr}^{,r}=-\Delta_i^0(K_{ij}^{,r}+\lambda_{cr}^0K_{ij}^{(G),r})\Delta_j^0 \quad (i,j=1,2,\ldots,n;r=1,2,\ldots,R) \tag{12}$$

It may be noted again that the eigenvector in the above expression is $K_{ij}^{(G)0}$ orthonormal.

The total minimum eigenvalue or load parameter can be written as

$$\lambda_{cr}(b_r)=\lambda_{cr}^0(b_r^0)+\lambda_{cr}^{,r}(b_r^0)(b_r-b_r^0) \quad (r=1,2,\ldots,R) \tag{13}$$

The second-order statistics of the eigenvalue can be evaluated by first squaring and then taking expectation of the above equation. The statistics of buckling load can be obtained by multiplying the statistics of the critical load parameter with the reference load.

4 Analytical Approach for Specially Orthotropic Plates Without Cutout

As mentioned earlier, the stochastic buckling analysis of plate with cutout is not available in the literature. Hence, an analytical approach is being developed for a simplified case of laminated plates without cutout to validate the present stochastic formulation. However, the mean buckling analysis, being the same as the deterministic analysis, is also validated with analytical results available in the literature.

Suppose a specially orthotropic rectangular laminate, which has a single or multiple specially orthotropic layers that are symmetrically arranged about the middle surface, is subjected to in-plane compressive loading. By assuming the prebuckled stress to be uniform throughout the plate and equal to the applied load, we get only one uncoupled differential equation. Based on Kirchhoff–Love plate theory, the buckling differential equation for such plates can be written as [32]

$$D_{11}\frac{\partial^4u_3}{\partial x_1^4}+2(D_{12}+2D_{66})\frac{\partial^4u_3}{\partial x_1^2\,\partial x_2^2}+D_{22}\frac{\partial^4u_3}{\partial x_2^4}+\bar{N}\frac{\partial^2u_3}{\partial x_1^2}=0 \tag{14}$$

The following admissible function, which satisfies the simply support boundary condition on all edges of the plate (with edges free to move in their respective in-plane normal directions), is used

$$u_3=U_{mn}\sin\frac{m\pi x_1}{a}\sin\frac{n\pi x_2}{b} \quad (m,n=1,2,\ldots,\infty) \tag{15}$$

where U_{mn} are the maximum displacements in x_1 and x_2 directions for a particular value of (m,n). Here m and n are the number of buckle half-wavelengths in the x_1 and x_2 directions, respectively. After substituting the above expression into Eq. (14) we get the expression for buckling load as

Table 2 Three different boundary conditions

Boundary condition	At $x_1=0$ and $x_1=a$	At $x_2=0$ and $x_2=b$
SSSS	$u_2=u_3=0$	$u_1=u_3=0$
SCSC	$u_2=u_3=0$	$u_1=u_2=u_3=0$
SFSF	$u_2=u_3=0$	$u_1=u_2=u_3\neq 0$

$$\bar{N}=\pi^2\left[D_{11}\left(\frac{m}{a}\right)^2+2(D_{12}+2D_{66})\left(\frac{n}{b}\right)^2+D_{22}\left(\frac{n}{b}\right)^4\left(\frac{a}{m}\right)^2\right] \tag{16}$$

The smallest or critical value of $\bar{N}$ obviously occurs when $n=1$.

To quantify the statistics of the critical buckling load, a mean centered first-order perturbation approach is adopted. It is assumed that all effective material properties of each lamina are random and uncorrelated to each other and the dispersion in each component about its mean value is small. Based on Taylor series expansion of the involved random terms in Eq. (16), we arrive at the following equations:

zeroth order

$$\bar{N}_{cr}^0=\pi^2\left[D_{11}^0\left(\frac{m^2}{a^2}\right)+\left(\frac{2}{b^2}\right)(D_{12}^0+2D_{66}^0)+D_{22}^0\left(\frac{a^2}{m^2b^4}\right)\right] \quad (m=1,2,\ldots,\infty) \tag{17a}$$

first order

$$\bar{N}_{cr}^{,r}=\pi^2\left[D_{11}^{,r}\left(\frac{m^2}{a^2}\right)+\left(\frac{2}{b^2}\right)(D_{12}^{,r}+2D_{66}^{,r})+D_{22}^{,r}\left(\frac{a^2}{m^2b^4}\right)\right] \quad (m=1,2,\ldots,\infty;r=1,2,\ldots,R) \tag{17b}$$

The zeroth-order equation (Eq. (17a)) gives the mean critical buckling load. Subsequently, the expression for its variance can be expressed as

$$\mathrm{Var}(\bar{N}_{cr})=E\{[\bar{N}_{cr}^{,r}(b_r^0)(b_r-b_r^0)]^2\} \quad (r=1,2,\ldots,R) \tag{18}$$

5 Results and Discussion

The layerwise based stochastic finite element model described in the previous section is used to illustrate the buckling behavior statistics of symmetric composite laminated plates with cutouts. However, the method has no limitation and can be used for any layup sequences. In this section, first some of the mean prebuckled behavior characteristics of the compression loaded square laminated plates with cutouts are discussed. These brief discussions are intended to provide insight into composite plate behavior. Next, the mean buckling behavior of plates with cutouts is described. The effect of randomness in the material properties on the buckling load of plates with cutout is also investigated. The present solutions for mean and variance of buckling load of plates are verified with available results in the literature. The convergence study of finite element solution is done for plates with all cutout sizes employed. The analysis is performed using equivalent layer model with $p_{x_1x_2}=3$ and $p_{x_3}=3$. Three different boundary conditions SSSS, SCSC, and SFSF, as described in Table 2, are used. It may be noted that the boundary conditions of loaded edges are always kept as simple support while the boundary conditions of unloaded edges are changed.

5.1 Mean Prebuckled Stress. In this section four layered symmetric laminated plates $[\theta/-\theta/-\theta/\theta]$, where θ changes from 0 deg to 90 deg, is taken to demonstrate the mean prebuckled stresses of composite laminated plates with circular hole under uniaxial compression. The effect of boundary conditions on the prebuckled stress intensity $\bar{\sigma}_{ij}=(\sigma_{ij}/q^{\text{ref}})$ is also obtained. Although all the six components of stresses exist in the plate under uniaxial compressive loading, band plots of the prebuckled stress distribution are drawn only for in-plane stress σ_{11} at the bottom face of the plates for some representative cases.

Results have been presented for a graphite/epoxy square laminated plate of width 50 mm and $b/h=50$ with following mean material properties

$$E_{ll}=40E_{tt}, \quad G_{lt}=0.6E_{tt}, \quad G_{tt}=0.5E_{tt},$$

$$\nu_{lt}=\nu_{tt}=0.25, \quad E_{tt}=10 \text{ GPa}$$

Figures 4(*a*)–4(*g*) show the distribution of stress intensity $\bar{\sigma}_{11}$ across the area of plate, with a centrally located hole of $d/b=0.2$, under inplane compression having SSSS boundary condition. Here seven different layups, by taking $\theta=0$ deg, 15 deg, 30 deg, 45 deg, 60 deg, 75 deg, and 90 deg, are chosen to see the effect of ply orientation on the prebuckled stress distribution. It is observed that:

1. The stress distribution for 0 deg laminates is almost uniform through out the plate except near the hole where a very high stress concentration exists compared to the other laminates;
2. For $\theta=15$ deg and 30 deg the stress pattern is symmetric along the fiber axis;
3. Stress is significant near the boundary of plate for $\theta=45$ deg as compared to that for $\theta=0$ deg, 15 deg, and 30 deg. Stress concentration near the hole is also found to be less as compared to $\theta=0$ deg, 15 deg, and 30 deg;
4. The stress near the boundary becomes more significant for plates with $\theta=60$ deg and 75 deg. Stress concentration near the hole is found to be much less, in this case, as compared to other laminates;
5. For $\theta=60$ deg, the tensile stress is found to be much higher as compared to the compressive stress. This increase in tensile stress increases the prebuckling stiffness of the plates; and
6. For plates with $\theta=90$ deg, the stress distribution is found to be uniform through the plates except near the hole where a low stress concentration exists.

The intensity of stress depends on both boundary constraints and ply orientations. For these particular plates, the prebuckling stiffness can be higher for $\theta=45$ deg, 60 deg, and 75 deg as compared to the plates with other ply orientations. It is also observed that plates with SCSC boundary constraint (figure not shown) have generally higher prebuckling stiffness as compared to SSSS plates. However, for plates with $\theta=60$ deg, this stiffening is found to be less prominent as compared to the stiffening effect coming from the tensile stress developed in the laminates due to SSSS boundary condition. In the case of plates with SFSF boundary condition (figure not shown), free edges do not contribute to increase in the prebuckling stiffness of the plates as compared to simply supported and clamped edges. Ply orientations only play a role in increasing the prebuckling stiffness, which is higher for $\theta=0$ deg laminates and decreases monotonically when ply orientation changes from 0 deg to 90 deg. The effect of the prebuckling stiffness on the mean buckling strength will be discussed in Sec. 5.2.3.

5.2 Stochastic Buckling Load. In this section the mean and covariance of buckling loads of laminated composite plates with cutouts is studied using layerwise plate model based stochastic finite element method. The verification of the results for the mean buckling strength of laminated plates without cutouts is compared with available analytical solutions in the literature [33]. The results for the mean buckling strength of laminated plates with cutouts are also compared with the experimental results of Nemeth [13]. Second-order statistics of buckling load, evaluated using the stochastic finite element method, is also validated with the analytical solution presented in Sec. 4. In the present analysis the effective elastic moduli (E_{ll}, E_{tt}, ν_{lt}, ν_{tt}, G_{lt}, G_{tt}) of the lamina are treated as the basic random variables. The variations in these properties are taken based on a micromechanics approach as described earlier. It is assumed that the microlevel variations in the material properties are uncorrelated. A parametric study is conducted to see the effect of hole size, layups, and boundary constraints on the mean and variance of the buckling load for laminated plates with cutouts.

5.2.1 Validation for Mean Buckling Load. First the mean buckling load of plates without cutout, based on layerwise plate model using prebuckled stresses, is compared with Reddy's analytical results [33]. Table 3 shows the normalized mean buckling load of two-layered antisymmetric cross-ply square laminates with $b/h=5$ and 10 and having different boundary conditions. The mean material properties used here is same as described in Sec. 5.1. The buckling load obtained using the equivalent layer (by assuming both material layers as an equivalent solution layer) model and using uniform prebuckled stress assumption is also presented in the table. It is observed that:

1. The mean buckling load obtained, using uniform stress assumption, along with equivalent single layer model, lies between the values obtained using FSDT and HSDT [33]. CPT grossly overpredicts the buckling load;
2. The conventional 2D plate models overpredict buckling loads as compared to those obtained with the layerwise plate model; and
3. The differences in the buckling load obtained from the present layerwise model (with actual prebuckled stress) and the conventional HSDT plate model [33] are found to be more significant for the SCSC boundary condition compared to SSSS and SFSF boundary conditions. This is due to the increased effect of the boundary constraint in the case of SCSC boundary condition compared to the other two boundary conditions.

The difference between the present and the conventional plate solutions is significant because the uniform stress assumption neglects all other in-plane and out-of-plane stresses which are significant for thick plates.

The present mean buckling load for plates with cutout is also compared with experimental results of Nemeth [13]. Following Ref. [13], results are obtained for a 24 layered symmetric angle ply laminate of size 254×254 mm, with a centrally located cutout and the following mean material properties

$$E_{ll}=127.8 \text{ GPa}, \quad E_{tt}=11 \text{ GPa}, \quad G_{lt}=G_{tt}=5.7 \text{ GPa},$$

$$\nu_{lt}=\nu_{tt}=0.35, \quad t_{\text{ply}}=0.127 \text{ mm}$$

Table 4 shows a comparison of the mean buckling load of $[(\pm 30 \text{ deg})_6]_s$ square laminate with centrally located cutout of different sizes. From the table it is observed that:

1. The present mean buckling load for plates with CSSS boundary condition is close to those obtained by Nemeth with hole of size $d/b\leqslant 0.316$. The present analysis gives slightly lower buckling load because CSSS plates are less stiff compared to CSCS plates;
2. The % difference in the buckling load obtained from the present analysis and the experimental results increases for plates with hole of size $d/b\geqslant 0.6$;
3. The present mean buckling load for plates with CCSC boundary condition is also found to be close to those obtained by Nemeth with a hole of $d/b\leqslant 0.316$; and

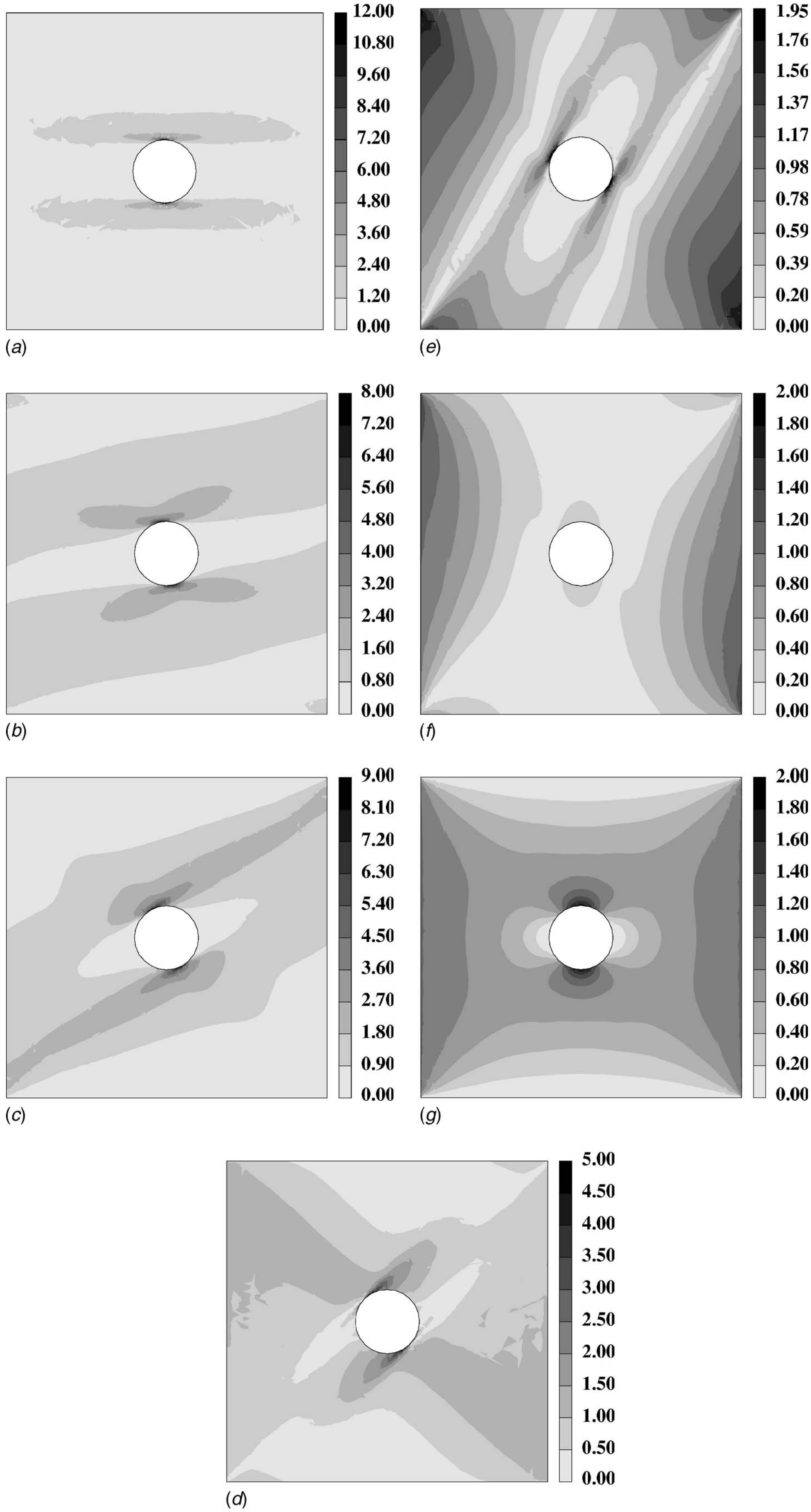


Fig. 4 Distribution of the intensity of stress ($\bar{\sigma}_{11}$) for $[\theta/-\theta/-\theta/\theta]$ square plates with SSSS boundary condition and having different ply orientation: (*a*) 0 deg; (*b*) 15 deg; (*c*) 30 deg; (*d*) 45 deg; (*e*) 60 deg; (*f*) 75 deg; and (*g*) 90 deg

Table 3 Comparison of nondimensioned mean buckling load for [0 deg/90 deg] square laminates without cutout having different thickness ratios and support conditions

Various boundary conditions	Normalized mean buckling load $\hat{N}=\lambda_0^{cr}b^2/(E_{tt}h^3)$					
	Present result				Reddy[a] result (various 2D plate model)	
	Using prebuckled stress (layer by layer)		using uniform stress (equivalent layer)			
	$b/h=5$	$b/h=10$	$b/h=5$	$b/h=10$	$b/h=5$	$b/h=10$
SSSS	7.874	11.070	8.473	11.525	8.769 (HSDT) 8.277 (FSDT) 12.957 (CPT)	11.562 (HSDT) 11.353 (FSDT) 12.957 (CPT)
SCSC	8.945	18.746	10.652	20.825	11.490 (HSDT) 9.757 (FSDT) 31.280 (CPT)	21.464 (HSDT) 20.067 (FSDT) 31.280 (CPT)
SFSF	3.32	4.625	3.813	4.931	3.905 (HSDT) 3.682 (FSDT) 5.425 (CPT)	4.940 (HSDT) 4.851 (FSDT) 5.425 (CPT)

[a]See Ref. [33].

4. For hole size of $d/b \geqslant 0.6$, the present mean buckling load for plates with CCSC boundary condition are close to the experimental one.

Remark. In Ref. [13] the boundary conditions are given as CSCS, with the loaded edges clamped. Physically and numerically this can not be simulated. In the present investigation, the applied boundary conditions are more appropriately modeled using the CCSC for which the difference with the experimental results is $\leqslant 19\%$. However, when a softer boundary condition CSSS is used the discrepancy between the two results goes to $\approx 40\%$ in some cases.

5.2.2 *Validation for Statistics of Buckling Load*. In order to validate the SFEM implementation, the statistics of the buckling load obtained using a layerwise plate model is validated with the analytical solutions described in Sec. 4. As analytical solution of buckling of composite plate with cutout is generally not possible; only plates without cutout are taken to validate the statistics of buckling load with the closed form solutions. Table 5 shows the effect of variation in the material properties, by adopting COV for E_{ll}, E_{tt}, ν_{lt}, ν_{tt}, G_{lt}, and G_{tt} as 7%, 4%, 4%, 4%, 12%, and 3%, respectively, as obtained in Sec. 2, on the buckling load for a thin single layered orthotropic plate with $b/h=100$ and having SSSS boundary conditions. The effect of both prebuckled stress and uniform stress assumption on the statistics of buckling load parameter is shown. From the table it is observed that:

1. The standard deviation (SD) of buckling load obtained using the layerwise model by assuming uniform state of stress is very close to the analytical results;
2. Stochastic analysis based on uniform stress assumption predicts lower SD of buckling load as compared to that based on prebuckled stress analysis; and
3. The difference between the SD of buckling load using both assumptions becomes more significant as the aspect ratio of the plate increases.

The layerwise plate model, using uniform stress assumption, gives slightly lower SD of buckling load compared to the closed form value because this model is less stiff as compared to the Kirchhoff–Love model. This validates the SFEM implementation of the present study.

5.2.3 *Mean Buckling Load*. Figures 5(*a*)–5(*c*) show the effect of hole size (d/b) on the mean buckling load of four layered square symmetric laminated plate $[\theta/-\theta/-\theta/\theta]$ of width 50 mm and $b/h=50$ having SSSS, SCSC, and SFSF boundary conditions, respectively. Seven different layups, with θ ranging from 0 deg to 90 deg; as defined in Sec. 5.1, are chosen to see the effect of ply orientation on the mean buckling load. It is observed that:

1. The mean buckling load of plates with $\theta=0$ deg and 15 deg, under both SSSS and SCSC boundary conditions, decreases monotonically with increase in cutout size. This is because the stress concentration (in the presence of the cutout) near the hole increases with increase in cutout size which results

Table 4 Comparison of mean buckling load for $[(\pm 30\text{ deg})_6]_s$ square laminates with central located circular cutout having different support conditions

Hole size (d/b)	Nemeth[a] experiment result (kN)	Present mean buckling load (kN)			
		CSSS plate	% difference	CCSC plate	% difference
0.00	42.894	38.5695	−10.0818	47.3220	10.3230
0.105	40.510	38.1292	−5.8770	46.8968	13.6188
0.316	37.668	35.5696	−5.5707	45.0000	19.4648
0.60	38.686	25.7490	−33.4410	36.9539	−4.4770
0.66	38.922	23.2350	−40.3036	34.5836	−11.1464

[a]See Ref. [13].

Table 5 Influence of dispersion in material property on the SD of buckling load for simply supported specially orthotropic plates with different aspect ratio and b/h=100.

Aspect ratio (a/b)	SD of buckling load (MPa)		
	Layerwise plate model		
	Using prebuckled stress	Using uniform stress	Analytical result (with uniform stress)
1.0	0.647	0.644	0.645
2.0	0.461	0.287	0.287
3.0	0.530	0.339	0.339

in monotonic reduction in prebuckling stiffness of plates.

2. The mean buckling load of 45 deg, 60 deg, and 75 deg plates, under SSSS boundary conditions, is found to be higher than the 0 deg, 15 deg, 30 deg, and 90 deg plates for all hole sizes considered. For these plates boundary constraints play a significant role and increase the prebuckling stiffness of the plate as discussed in Sec. 5.1. As a result the intensity $\bar{\sigma}_{11}$ is found to be significant near the boundary and less significant toward the middle of the plate. Since edges are constrained, this stress increase near the edges may cause the plate to sustain higher buckling loads.
3. The mean buckling load of plates with θ=60 deg and 75 deg, under SSSS boundary conditions, increases with increase in cutout size. This is because the prebuckling stiffness of plates increases with increase in cutout size.
4. The mean buckling load of 90 deg plates remains uniform with increase in hole size under SSSS boundary conditions. This is because no significant change in the stress concentration near the hole is found with increase in hole size.
5. The buckling load of solid plates with SCSC boundary condition is higher than those of solid plates with SSSS boundary condition for all ply orientations. This is because the prebuckling stiffness of plates increases with SCSC boundary constraints compared to plates with SSSS boundary condition.
6. SCSC boundary condition does not play any significant role in increasing the prebuckling stiffness of plates in the presence of a cutout as opposed to plates with SSSS boundary condition.
7. The buckling loads of SCSC plates with $\theta \geqslant 30$ deg remain almost uniform with increase in cutout size. This is because no substantial change in prebuckling stiffness is found for such plates with increase in cutout size.
8. For plates with SFSF boundary conditions, the mean buckling loads of solid plates are higher than those of the plates with holes and it decreases monotonically with increase in cutout size for all cases studied. This is because the stress concentration near the hole increases with increase in cutout size, which results in monotonic reduction in prebuckling stiffness with increase in cutout size.

From the mean buckling loads calculated above, the smallest value of buckling load is found for plate with θ=90 deg under SFSF boundary condition and with cutout of maximum diameter (i.e., d/b=0.6). It can be seen that as the hole size increases, the buckling load of 0 deg laminates is lower as compared to all other ply orientations for plates with SSSS and SCSC boundary conditions. These results suggests that for SSSS and SCSC plates, 45 deg, 60 deg, or 75 deg ply can be chosen to get the maximum mean buckling load, whereas for SFSF plates 0 deg ply should

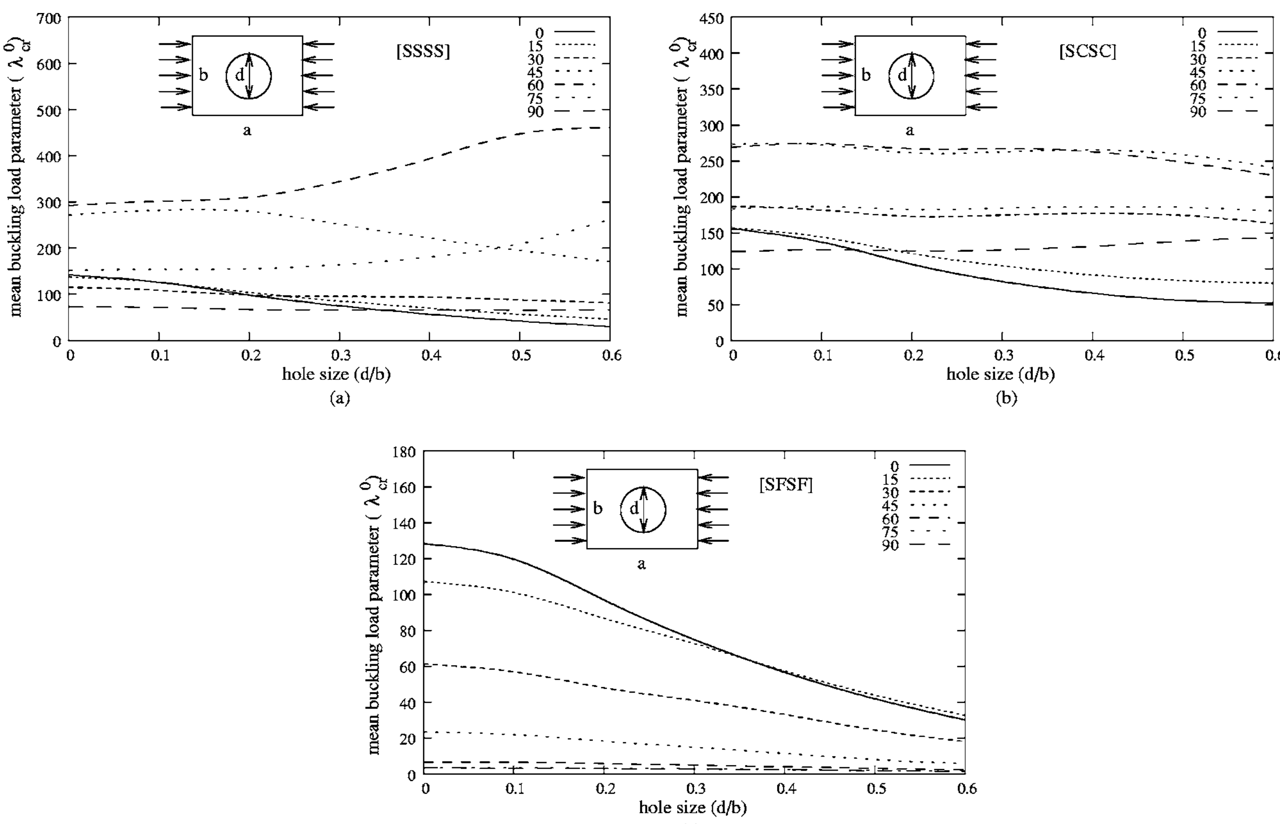


Fig. 5 The effect of hole size and ply orientation on the mean critical load parameter for $[\theta/-\theta/-\theta/\theta]$ square plates with different boundary conditions: (*a*) SSSS; (*b*) SCSC; and (*c*) SFSF

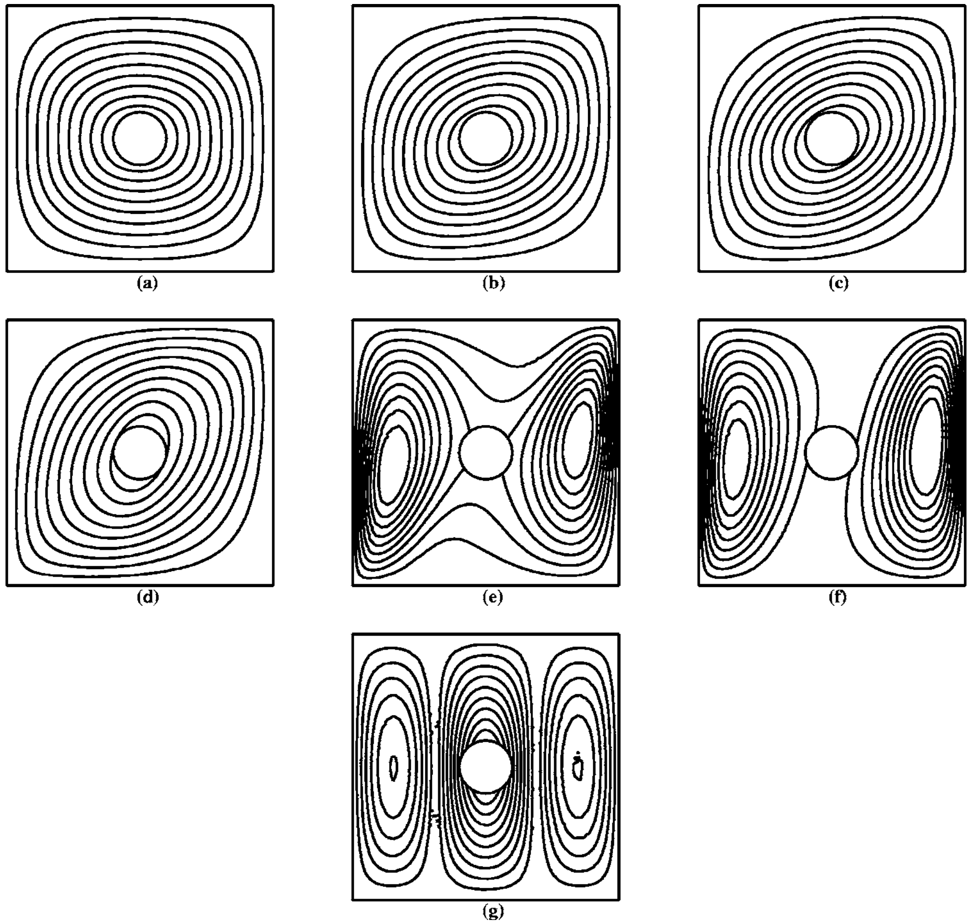


Fig. 6 The first eigen modes for $[\theta/-\theta/-\theta/\theta]$ square plates under SSSS boundary condition with different, (*a*) 0 deg; (*b*) 15 deg; (*c*) 30 deg; (*d*) 45 deg; (*e*) 60 deg; (*f*) 75 deg; and (*g*) 90 deg

always be chosen to get the maximum buckling strength. Note that the prebuckled stress state is significantly influenced by the boundary condition, size of cutout, and ply orientation.

The above mentioned parameters also strongly influence the mean buckling modes of plates. The first eigenmodes for plate with hole size $d/b=0.2$ under SSSS boundary condition is shown in Figs. 6(*a*)–6(*g*) for various ply orientations. It is seen that plates with $\theta \leqslant 45$ deg buckle in the first mode, whereas for $\theta=60$ deg and 75 deg buckling happens in the second mode. The plate with $\theta=90$ deg is found to buckle in the third mode.

5.2.4 Variance of Buckling Load. In this subsection the effect of variation in the effective material properties, with COV the same as mentioned in Sec. 5.2.2, on the second-order statistics of buckling load for composite laminated plates with cutout is presented. First the influence of individual random variables on the COV of critical buckling load characteristics is shown in Table 6 for $[\theta/-\theta/-\theta/\theta]$ square symmetric laminated plates with cutout size $d/b=0.1$. The SSSS boundary condition is applied. From the table it is observed that the buckling load is most sensitive to change in E_{ll} for all layups except for $\theta=90$ deg for which G_{lt} has a dominant effect on critical buckling load. The effect of dispersion in E_{tt}, ν_{lt}, ν_{tt}, and G_{tt} on the buckling load is much less compared to E_{ll} and G_{lt}.

Having seen the effect of variation in individual BRV on the

Table 6 Effect of variation of individual material properties on COV of critical buckling load for $[\theta/-\theta/-\theta/\theta]$ square laminates with hole size $d/b=0.1$ having SSSS boundary condition and $b/h=50$

Angle (θ)	COV of critical buckling load (%)					
	E_{ll}	E_{tt}	ν_{lt}	ν_{tt}	G_{lt}	G_{tt}
0	8.388	0.283	0.032	2.482E-06	2.145	0.004
15	8.012	0.596	0.007	3.045E-04	1.786	0.019
30	6.282	0.847	0.009	0.002	4.134	0.050
45	8.197	0.448	0.016	5.643E-04	1.374	0.132
60	7.161	0.930	0.051	0.013	1.584	0.175
75	6.046	1.514	0.037	0.061	2.407	0.071
90	2.446	1.813	0.109	0.035	8.514	0.022

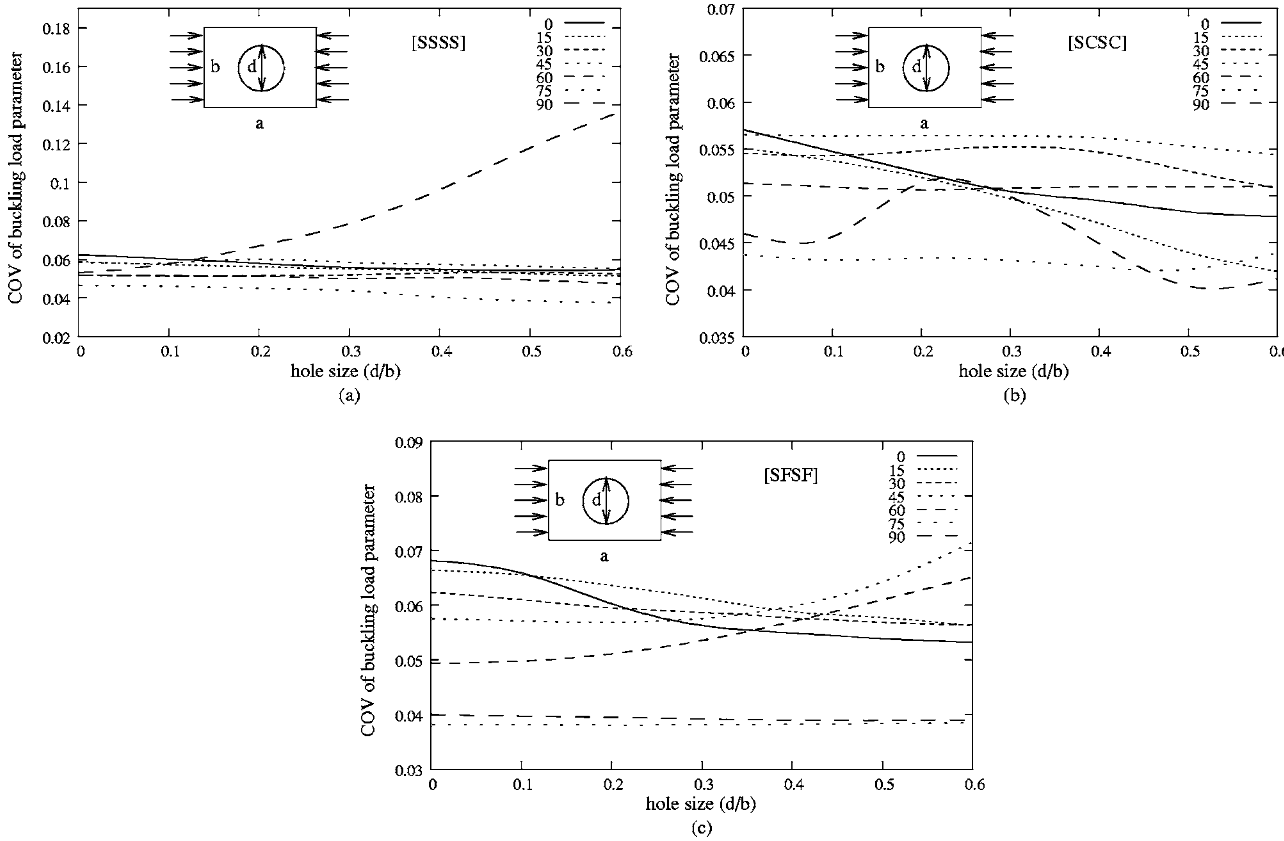


Fig. 7 Influence of dispersion in all material properties changing simultaneously on the COV of critical load parameter for $[\theta/-\theta/-\theta/\theta]$ square laminates with various hole size and different boundary conditions: (*a*) SSSS; (*b*) SCSC; and (*c*) SFSF

critical buckling load, it is desirable to see the effect of simultaneous variation in all BRVs on the buckling load of plates with cutouts. The influence of all BRVs changing simultaneously on the buckling load is depicted in Figs. 7(*a*)–7(*c*) for $[\theta/-\theta/-\theta/\theta]$ square plates with various hole sizes under SSSS, SCSC, and SFSF boundary conditions, respectively. It is observed that:

1. The COV of buckling load lies between 4% and 7% for plates with SSSS boundary conditions for all θ taken, except for plates with $\theta=90$ deg, for which the maximum COV of buckling load is $\approx 14\%$;
2. The COV of buckling load is less affected by increase in hole size and remains constant for plates with different ply orientation except for $\theta=90$ deg and SSSS boundary conditions. For this plate the COV of buckling load increases with increase in hole size;
3. The COV of buckling load for plates with SCSC boundary condition lies between 4% and 6%;
4. For SCSC plates $\theta=0$ deg, 15 deg, the COV of buckling load decreases slightly with increase in hole size, whereas for $\theta=45$ deg, 60 deg, and 75 deg the COV of buckling load is almost insensitive to hole size variation and remains constant. For $\theta=90$ deg the COV first increases and then decreases with increase in hole size;
5. The COV of buckling load for plates with SFSF boundary condition lies between 3% and 7%; and
6. For SFSF plates with $\theta=0$ deg, 15 deg, 30 deg the COV of buckling load decreases slightly with increase in hole size, whereas for $\theta=45$ deg, 60 deg the COV of bucking load increases with increase in hole size. Plates with $\theta=75$ deg, 90 deg the COV of buckling load remains constant.

6 Conclusions

In the present linearized stochastic buckling analysis, an attempt has been made to study the mean and variance of critical buckling load of laminated composite plates with circular cutout. The stochastic analysis has been done to account for the effect of dispersion in the effective material properties, obtained using the micromechanics based approach, on the critical buckling load. From the numerical results the following conclusions are drawn:

1. The mean buckling load strongly depends on the ply orientation and boundary constraints.
2. The mean buckling load generally decreases with the increase of cutout size. However, in some cases (due to interplay of boundary conditions and ply orientation) the mean buckling load may increase with increasing hole size.
3. Due to the cutout the internal stress distribution becomes nonuniform. The non-uniformity is sensitive to the cutout size and ply orientation. In determining buckling load and the corresponding mode shape, the role of orthotropy becomes more significant.
4. In the case of plates with SFSF boundary conditions, the buckling loads of solid plates without a cutout are always higher than those of the plates with holes. The buckling load decreases monotonically with increase in cutout size.
5. The uniform prebuckled state of stress assumption leads to prediction of lower mean and SD for the critical buckling load as compared to that obtained by using the prebuckled stress.
6. In most cases the COV of buckling load of plates with cut-

out for the adopted variation in material properties is found to be $\approx 6\%$.

7. Depending on the ply orientation and applied boundary condition, the variation of the buckling load may be significant in some cases.
8. Assuming an upper and lower limit in the population as ± 3 SD, a COV of 6% would be equivalent to a limiting upper and lower limit of 18% in the response. In the case of laminates with $\theta = 90$ deg and SSSS boundary condition this limit can even go up to 35%.
9. In the design process the effect of dispersion in the material parameters on the buckling load has to be accounted for. Due to material dispersion buckling can happen at load levels that are significantly lower than the mean buckling load.

References

[1] Lin, S. C., and Kam, T. Y., 2000, "Probabilistic Failure Analysis of Transversely Loaded Laminated Composite Plates Using First Order Second Moment Method," J. Eng. Mech., **126**(8), pp. 812–820.

[2] Babuska, I., Andersson, B., Smith, P. J., and Levin, K., 1999, "Damage Analysis of Fiber Composites; Part I: Statistical Analysis on Fiber Scale," Comput. Methods Appl. Mech. Eng., **172**, pp. 27–77.

[3] Kaminski, M., and Kleiber, M., 2000, "Perturbation Based Stochastic Finite Element Method for Homogenization of Two-Phase Elastic Composites," Comput. Struct., **78**, pp. 811–826.

[4] Chamis, C. C., 2004, "Probabilistic Simulation of Multi-Scale Composite Behavior," Theor. Appl. Fract. Mech., **41**(1–3), pp. 51–61.

[5] Buryachenko, B. A., Pagano, N. J., Kim, R. Y., and Spowart, J. E., 2003, "Quantitative Description and Numerical Simulation of Random Microstructures of Composites and their Effective Elastic Moduli," Int. J. Solids Struct., **40**, pp. 47–72.

[6] Kleiber, M., and Hien, T. D., 1992, *The Stochastic Finite Element Method*, Wiley, New York.

[7] Liu, W. M., Belytchko, T., and Mani, A., 1986, "Probabilistic Finite Elements for Nonlinear Structural Dynamics," Comput. Methods Appl. Mech. Eng., **56**, pp. 61–81.

[8] Ghanem, R. G., and Spanos, P. D., 1991, *Stochastic Finite Element: A Spectral Approach*, Springer, New York.

[9] Cruz, M. E., and Patera, A. T., 1995, "A Parallel Monte Carlo Finite Element Procedure for the Analysis of Multicomponent Media," Int. J. Numer. Methods Eng., **38**, pp. 1087–1121.

[10] Hurtado, J. E., and Barbat, A. H., 1998, "Monte Carlo Techniques in Computational Stochastic Mechanics," Arch. Comput. Methods Eng., **5**(1), pp. 3–30.

[11] Yettram, A. L., and Brown, C. J., 1986, "The Elastic Stability of Square Perforated Plates Under Biaxial Loadings," Comput. Struct., **22**(4), pp. 589–594.

[12] Nemeth, M. P., Stein, M., and Johnson, E. R., 1986, "An Approximate Buckling Analysis for Rectangular Orthotropic Plates With Centrally Located Cutouts," NASA Technical Paper No.-2528, NASA, Washington, D.C.

[13] Nemeth, M. P., 1986, "Importance of Anisotropy on Buckling of Compression Loaded Symmetric Composite Plates," AIAA J., **24**, pp. 1831–1835.

[14] Larsson, P. L., 1987, "On Buckling of Orthotropic Compressed Plates With Circular Holes," Compos. Struct., **7**, pp. 103–121.

[15] Lin, C. C., and Kuo, C. S., 1989, "Buckling of Laminated Plates With Holes," J. Compos. Mater., **23**, pp. 536–553.

[16] Srivasta, K. S., and Krishnamurthy, A. V., 1992, "Stability of Laminated Composite Plates With Cut-Outs," Comput. Struct., **43**(2), pp. 273–279.

[17] Sabir, A. B., and Chow, F. Y., 1986, "Elastic Buckling of Plates Containing Eccentrically Located Circular Holes," Thin-Walled Struct., **4**(2), pp. 135–150.

[18] Shakerley, T. M., and Brown, C. J., 1996, "Elastic Buckling of Plates With Eccentrically Positioned Rectangular Perforations," Int. J. Mech. Sci., **38**(8–9), pp. 825–838.

[19] Noor, A. K., and Kim, Y. H., 1996, "Buckling and Postbuckling of Composite Panels With Cutouts Subjected to Combined Edge Shear and Temperature Change," Comput. Struct., **60**(2), pp. 203–222.

[20] Bailey, R., and Wood, J., 1997, "Postbuckling Behaviour of Square Compression Loaded Graphite Epoxy Panels With Square and Elliptical Cutouts," Thin-Walled Struct., **28**(3–4), pp. 373–397.

[21] Nemeth, M. P., 1988, "Buckling Behavior of Compression Loaded Symmetrically Laminated Angle-Ply Plates With Holes," AIAA J., **26**(3), pp. 330–336.

[22] Zhang, J., and Ellingwood, B., 1995, "Effects of Uncertain Material Properties on Structural Stability," J. Struct. Eng., **121**(4), pp. 705–716.

[23] Lin, S. C., 2000, "Buckling Failure Analysis of Random Composite Laminates Subjected to Random Loads," Int. J. Solids Struct., **37**, pp. 7563–7576.

[24] Singh, B. N., Yadav, D., and Iyengar, N. G. R., 2001, "Initial Buckling of Composite Cylindrical Panels With Random Material Properties," Compos. Struct., **53**, pp. 55–64.

[25] Elishakoff, I., 2000, "Uncertain Buckling: Its Past, Present and Future," Int. J. Solids Struct., **37**, pp. 6869–6889.

[26] Grigoriu, M., 2000, "Stochastic Mechanics," Int. J. Solids Struct., **37**, pp. 197–214.

[27] Argyris, J., Papadrakakis, M., and Stefanou, G., 2002, "Stochastic Finite Element Analysis of Shells," Comput. Methods Appl. Mech. Eng., **191**, pp. 4781–4804.

[28] Stefanou, G., and Papadrakakis, M., 2004, "Stochastic Finite Element Analysis of Shell With Combined Random Material and Geometric Properties," Comput. Methods Appl. Mech. Eng., **193**, pp. 139–160.

[29] Christensen, R. M., 1979, *Mechanics of Composite Materials*, Krieger, Malabar, FL.

[30] Ahmad, N. U., and Basu, P. K., 1994, "Higher-Order Finite Element Modeling of Laminated Composite Plates," Int. J. Numer. Methods Eng., **37**, pp. 123–139.

[31] Actis, R. L., Szabo, B. A., and Schwab, C., 1999, "Hierarchic Models for Laminated Plates and Shells," Comput. Methods Appl. Mech. Eng., **172**, pp. 79–107.

[32] Jones, R. M., 1975, *Mechanics of Composite Materials*, Scripta Book Company, Washington, DC.

[33] Reddy, J. N., and Khedir, A. A., 1989, "Buckling and Vibration of Laminated Composite Plates Using Various Plate Theories," AIAA J., **27**(12), pp. 1808–1817.

On the Dynamics of Bodies With Continual Mass Variation

L. Cveticanin

I. Kovacic
e-mail: ivanakov@uns.ns.ac.yu

Faculty of Technical Sciences,
21000 Novi Sad,
Trg D Obradovica 6, Serbia

In this paper the differential equations of the general motion of the rigid body with continual mass variation are considered. The impact force and the impact torque that occur due to addition or separation of the body with velocity and angular velocity which differs from the velocity of mass center and angular velocity of the existing body are introduced. The theoretical consideration is applied for solving a real technical problem when the band winds up on the drum. The plane motion of the drum on which the band winds up is considered. The influence of the velocity of the band on the angular velocity of the drum and the motion of the drum mass center is obtained.
[DOI: 10.1115/1.2711231]

1 Introduction

The problem of the motion of the bodies with continual mass variation has been of interest for many researchers since the 17th century.

Cayley [1,2] was the first to consider the influence of mass variation on motion of the body. The class of dynamical problems he studied were the "continuous impact problems," i.e., the problems when the infinitesimal small mass is continuously added to a system which causes the velocity of the system to be continuously changed for a definite value. Two examples were considered in Ref. [3]: one, when the chain is on the table and is dropping vertically down from the table, and the second, when the chain is moving straightforward on a horizontal plane without friction under the influence of a mass M which is fixed at the end of a chain that is rolling around a drum and changing the length during motion.

In celestial mechanics the problem of continual mass variation was mentioned for the first time in 1886, and was connected with the secular acceleration of the Moon. This characteristic of Moon motion was discovered by Galileo at the end of the 17th century and theoretically considered by Laplace. Dufour [4] explained that the mass of the earth varies continuously to the falling shooting stars and also to combustion or spending in the atmosphere. He found that the dust of shooting stars, which fell on the surface of France in 1 year, can cover a volume of 0.1 m^3. Oppalzer [5] was the first to analyze the reason for secular acceleration of the Moon as a result of the increase of Earth and Moon mass. Namely, during a 100 years a 2.8 mm dust layer was formed on the Earth.

Gylden [6] extended the previous investigations by analyzing the relative motion of two variable mass systems under the influence of Newton force. For planet linear mass increase Gylden concluded that it will fall on the Sun. Meshchersky [7] assumed another mass variation for the same problem. He determined that the body moves along a spiral tending toward zero or increases the distance.

The first systematic assumption in mechanics of variable mass was done by Meshchersky [8]. He formed the general differential equation of motion of a body with variable mass, introducing the impact force which exists when the relative velocity of separating or adding body is not zero

$$M\ddot{\mathbf{r}} = \mathbf{F} + \frac{dM}{dt}(\mathbf{u} - \dot{\mathbf{r}}) \tag{1}$$

where $\mathbf{F}$ is the resultant force acting on the body and $(dM/dt)(\mathbf{u} - \dot{\mathbf{r}})$ the impact force.

The theory was applied for translatory motion of the rigid body with variable mass. In the equation

$$M\ddot{\mathbf{r}}_S = \mathbf{F}_S + \frac{dM}{dt}(\mathbf{u} - \dot{\mathbf{r}}_S) \tag{2}$$

$\mathbf{r}_S$ describes the translatory motion of the mass center of the body and $\mathbf{F}_S$ is the resultant force acting in the mass center. Using Eq. (2), the motion of the mass center of the rotor on which the band is winding up is obtained [9,10]. Based on the mentioned equation the motion of various mechanisms with variable mass is considered (see Refs. [11–13]).

Meshchersky [8] considered the case when the relative velocity of separating or adding body is zero, i.e., the impact force is zero. The differential equation for the translatory motion of the rigid body is

$$M(t)\ddot{\mathbf{r}}_S = \mathbf{F}_S \tag{3}$$

where $M(t)$ is the variable mass of the body, and for rotation around the fixed axle

$$J(t)\ddot{\varphi} = \mathcal{M} \tag{4}$$

where φ is the angle position of the body; $J(t)$ is the variable moment of inertia of the body; and $\mathcal{M}$ is the resultant moment acting on the body.

Based on the theory of Meshchersky the modern rocket theory and theory of cosmic flights are developed. For the general motion of the body with variable mass it is assumed that the absolute velocity and angular velocity of separating mass are zero (see for example, Ref. [14]), i.e., the relative velocity and relative angular velocity of separating mass are the same as the absolute velocity of the body and the differential equations of motion are

$$\frac{d}{dt}(M\mathbf{v}) = \mathbf{F}_r, \quad \frac{d}{dt}(\mathbf{I}\boldsymbol{\Omega}) = \mathbf{M}_S \tag{5}$$

where $\mathbf{v}$ is the velocity of mass center, $\boldsymbol{\Omega}$ is the angular velocity of the body, and $\mathbf{M}_S$ is the resultant moment acting on the body.

Using the principle of solidification Bessonov [15] obtained the differential equations of general motion of the body. Bessonov assumed that the relative velocity of adding or separating the body gives the impact force but the absolute angular velocity of the added or separated body is zero, i.e., the relative angular velocity

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received February 24, 2006; final manuscript received September 29, 2006. Review conducted by Igor Mezic.

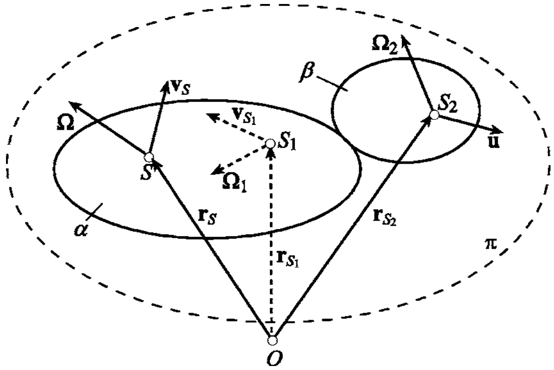


Fig. 1 Model of the mass addition system with position vectors of mass centers $\mathbf{r}_S$ and $\mathbf{r}_{S2}$, velocities of mass centers $\mathbf{v}_S$, and $\mathbf{u}$ and angular velocities Ω and Ω_2 of the bodies α and β before and $\mathbf{r}_{S1}$, $\mathbf{v}_{S1}$, and Ω_1 of the system π after mass variation

of the added or separated body is equal to the angular velocity of the remaining body. The differential equations of motion are the combination of Eqs. (2) and (5)

$$M\ddot{\mathbf{r}}_S = \mathbf{F}_S + \mathbf{\Phi}, \quad \frac{d}{dt}(\mathbf{I\Omega}) = \mathbf{M}_S + \mathbf{M}_S^{\mathbf{\Phi}} \tag{6}$$

where the impact force $\mathbf{\Phi}$ is

$$\mathbf{\Phi} = \frac{dM}{dt}(\mathbf{u} - \mathbf{v}_S) \tag{7}$$

and $\mathbf{M}_S^{\mathbf{\Phi}}$ is the moment of the impact force.

In this paper the theory is expanded to the case when the absolute velocity and the absolute angular velocity of adding or separating the body is not zero. Apart from the impact force the impact torque due to variation of the moment of inertia is also introduced. The differential equations of general motion of the body where the impact force and impact torque act are discussed. The obtained theory is applied for analyzing the plane motion of the rotor on which the band winds up.

2 Dynamics of Body With Continual Mass Variation

Let us consider the general motion of a body with mass M whose mass center is S. The motion of the body is defined with the linear $\mathbf{K}_b$ and angular momentum of the body $\mathbf{L}_{Ob}$ relating to a fixed point O (Fig. 1)

$$\mathbf{K}_b = M\mathbf{v}_S, \quad \mathbf{L}_{Ob} = \mathbf{r}_S \times M\mathbf{v}_S + \mathbf{L}_S \tag{8}$$

and the three kinematic Euler equations. $\mathbf{v}_S$ is the velocity of mass center; $\mathbf{r}_S$ is the position vector of mass center due to the fixed point O; $\mathbf{L}_S=\mathbf{I}_S\mathbf{\Omega}$ is the angular momentum of the body relative to the mass center S; $\mathbf{I}_S$ is the inertia tensor; and $\mathbf{\Omega}$ is the angular velocity of the body.

A body with elementary mass ΔM is added with the velocity of mass center $\mathbf{u}$ and angular velocity $\mathbf{\Omega}_2$ to the existing body. The general case of motion of the free rigid body is described by the following equations

$$\mathbf{K}_{a2} = \Delta M\mathbf{u}, \quad \mathbf{L}_{Oa2} = \mathbf{r}_{S2a} \times \Delta M\mathbf{u} + \mathbf{L}_{S2} \tag{9}$$

where $\mathbf{r}_{S2a}$ is the position vector of the added mass due to the fixed point O; $\mathbf{L}_{S2}=\Delta\mathbf{I}_{S2}\mathbf{\Omega}_2$ is the angular momentum of the added mass; and $\Delta\mathbf{I}_{S2}$ is the moment of inertia according to the axes in mass center S_2.

The two bodies (existing and added) are regarded as one complex system where impact force between these parts is internal within the system, i.e., the existing body and the added body form a unique system π. We assume that the linear momentum before adding mass is equal to the sum of linear momentums of the existing and added body, i.e., $\mathbf{K}_1=\mathbf{K}_b+\mathbf{K}_{a2}$ and the angular momentum before adding mass is equal to the sum of angular momentums of the existing and added body, i.e., $\mathbf{L}_{O1}=\mathbf{L}_{Ob}+\mathbf{L}_{Oa2}$.

After the adding of body the linear and angular momentum of the "new" body are obtained and relating to the fixed point O are

$$\mathbf{K}_2 = (M+\Delta M)\mathbf{v}_{S1}, \quad \mathbf{L}_{O2} = \mathbf{r}_{S1} \times (M+\Delta M)\mathbf{v}_{S1} + \mathbf{L}_{S1} \tag{10}$$

where $\mathbf{v}_{S1}$ is the unknown velocity of mass center S_1 of the new body; $\mathbf{L}_{S1}=\mathbf{I}_{S1}\mathbf{\Omega}_1$ is the angular momentum of the body; and $\mathbf{\Omega}_1$ is the unknown angular velocity of the new body.

Using relations (8)–(10) we obtain the linear and angular momentum differences after and before body addition

$$\mathbf{K}_2 - \mathbf{K}_1 = \Delta\mathbf{K} = M(\mathbf{v}_{S1} - \mathbf{v}_S) + \Delta M(\mathbf{v}_{S1} - \mathbf{u}) \tag{11}$$

$$\mathbf{L}_{O2} - \mathbf{L}_{O1} = \Delta\mathbf{L}_O = \mathbf{r}_{S1} \times (M+\Delta M)\mathbf{v}_{S1} - \mathbf{r}_{S2a} \times \Delta M\mathbf{u} - \mathbf{r}_S \times M\mathbf{v}_S + (\mathbf{L}_{S1} - \mathbf{L}_S) - \mathbf{L}_{S2} \tag{12}$$

As the position vector of gravity center is

$$\mathbf{r}_{S1} = \frac{M}{M+\Delta M}\mathbf{r}_S + \frac{\Delta M}{M+\Delta M}\mathbf{r}_{S2} \tag{13}$$

and

$$\mathbf{r}_{S1} = \mathbf{r}_S + \mathbf{SS}_1, \quad \mathbf{r}_{S2} = \mathbf{r}_S + \mathbf{SS}_2 \tag{14}$$

we obtain

$$\mathbf{SS}_1 = \frac{\Delta M}{M+\Delta M}\mathbf{SS}_2 \tag{15}$$

Substituting Eqs. (13)–(15) into Eq. (12) and using relation (11) we have

$$\Delta\mathbf{L}_O = \mathbf{r}_S \times \Delta\mathbf{K} + \mathbf{SS}_2 \times \Delta M(\mathbf{v}_{S1} - \mathbf{u}) + (\mathbf{L}_{S1} - \mathbf{L}_S) - \Delta\mathbf{I}_{S2}\mathbf{\Omega}_2 \tag{16}$$

Dividing Eqs. (11) and (16) by the interval of time Δt we obtain

$$\frac{\Delta\mathbf{K}}{\Delta t} = M\frac{\Delta\mathbf{v}_{S1}}{\Delta t} + \frac{\Delta M}{\Delta t}(\mathbf{v}_{S1} - \mathbf{u}) \tag{17}$$

$$\frac{\Delta\mathbf{L}_O}{\Delta t} = \mathbf{r}_S \times \frac{\Delta\mathbf{K}}{\Delta t} + \mathbf{SS}_2 \times \frac{\Delta M}{\Delta t}(\mathbf{v}_{S_1} - \mathbf{u}) + \frac{\Delta\mathbf{L}_S}{\Delta t} - \frac{\Delta\mathbf{I}_{S2}}{\Delta t}\mathbf{\Omega}_2 \tag{18}$$

where

$$\Delta\mathbf{v}_{S1} = \mathbf{v}_{S1} - \mathbf{v}_S, \quad \Delta\mathbf{L}_S = \mathbf{L}_{S1} - \mathbf{L}_S$$

If the time interval Δt is infinitesimal dt, and tends to zero, relations (17) and (18) transform into

$$\frac{d\mathbf{K}}{dt} = M\frac{d\mathbf{v}_S}{dt} + \frac{dM}{dt}(\mathbf{v}_S - \mathbf{u}) \tag{19}$$

$$\frac{d\mathbf{L}_O}{dt} = \mathbf{r}_S \times \frac{d\mathbf{K}}{dt} + \mathbf{SS}_2 \times \frac{dM}{dt}(\mathbf{v}_S - \mathbf{u}) + \frac{d\mathbf{L}_S}{dt} - \frac{d\mathbf{I}_S}{dt}\mathbf{\Omega}_2 \tag{20}$$

Relations (19) and (20) represent the time variation of the linear and angular momentum of the body with continual mass and moment of inertia variation.

The total moment and angular moment change of the variable mass system on receiving the body with mass ΔM and moment of inertia $\Delta\mathbf{I}_{S2}$, can be associated with an impulse $\mathbf{F}_r\Delta t$ and corresponding value $\mathbf{M}_O\Delta t$ due to the action of an external resultant force $\mathbf{F}_r$, which is the sum of the external forces acting on the system of bodies, and a resultant moment of the external forces $\mathbf{M}_O$ about the fixed point O. In the limit $\Delta t \to 0$

$$\frac{d\mathbf{K}}{dt} = \mathbf{F}_r, \quad \frac{d\mathbf{L}_O}{dt} = \mathbf{M}_O \tag{21}$$

By introducing the moment of external forces $\mathbf{M}_S$ for the mass center of the body S the connection between the two resultant moments $\mathbf{M}_O$ and $\mathbf{M}_S$ for the two points O and S is

$$\mathbf{M}_O = \mathbf{M}_S + \mathbf{r}_S \times \mathbf{F}_r \tag{22}$$

Introducing Eqs. (21) with Eq. (22) into Eqs. (19) and (20) the differential equations of motion of the body with continual added mass are obtained

$$\frac{d\mathbf{K}}{dt} \equiv \frac{d}{dt}(M\mathbf{v}_S) = \mathbf{F}_r + \frac{dM}{dt}\mathbf{u} \tag{23}$$

$$\frac{d\mathbf{L}_S}{dt} \equiv \frac{d}{dt}(\mathbf{I}_S\boldsymbol{\Omega}) = \mathbf{M}_S + \mathbf{M}_S^{\boldsymbol{\Phi}} + \frac{d\mathbf{I}_S}{dt}\boldsymbol{\Omega}_2 \tag{24}$$

where $\mathbf{I}_S\boldsymbol{\Omega}$ is a vector which is the dot product of the tensor $\mathbf{I}_S$ with the vector $\boldsymbol{\Omega}$, the impact force is

$$\boldsymbol{\Phi} = \frac{dM}{dt}(\mathbf{u} - \mathbf{v}_S) \tag{25}$$

and the moment of the impact force due to point S is

$$\mathbf{M}_S^{\boldsymbol{\Phi}} = \mathbf{SS}_2 \times \boldsymbol{\Phi} \tag{26}$$

For (see Ref. [16])

$$\frac{d}{dt}(\mathbf{I}_S\boldsymbol{\Omega}) = \boldsymbol{\Omega}\frac{d\mathbf{I}_S}{dt} + \mathbf{I}_S\frac{d\boldsymbol{\Omega}}{dt} + \boldsymbol{\Omega} \times \mathbf{I}_S\boldsymbol{\Omega} \tag{27}$$

the differential equations of motion transform to

$$M\frac{d\mathbf{v}_S}{dt} = \mathbf{F}_r + \frac{dM}{dt}(\mathbf{u} - \mathbf{v}_S) \tag{28}$$

$$\mathbf{I}_S\frac{d\boldsymbol{\Omega}}{dt} = \mathbf{M}_S + \mathbf{M}_S^{\boldsymbol{\Phi}} + \frac{d\mathbf{I}}{dt}(\boldsymbol{\Omega}_2 - \boldsymbol{\Omega}) - \boldsymbol{\Omega} \times \mathbf{I}_S\boldsymbol{\Omega} \tag{29}$$

i.e.,

$$M\frac{d\mathbf{v}_S}{dt} = \mathbf{F}_r + \boldsymbol{\Phi} \tag{30}$$

$$\mathbf{I}_S\frac{d\boldsymbol{\Omega}}{dt} = \mathbf{M}_S + \mathbf{M}_S^{\boldsymbol{\Phi}} - \boldsymbol{\Omega} \times \mathbf{I}_S\boldsymbol{\Omega} + \mathbb{R} \tag{31}$$

where the impact torque is

$$\mathbb{R} = \frac{d\mathbf{I}_S}{dt}(\boldsymbol{\Omega}_2 - \boldsymbol{\Omega}) \tag{32}$$

The first equation defines the translational motion and the second the rotation around the mass center S.

Introducing the fixed coordinate system $Oxyz$, the components u, v, w of the velocity $\mathbf{v}$, the components u_2, v_2, w_2 of the velocity $\mathbf{u}$, and F_x, F_y, F_z the components of the resultant $\mathbf{F}_r$, the vector differential equation of translational motion is given with three scalar equations

$$M\frac{du}{dt} = F_x + \frac{dM}{dt}(u_2 - u)$$

$$M\frac{dv}{dt} = F_y + \frac{dM}{dt}(v_2 - v)$$

$$M\frac{dw}{dt} = F_z + \frac{dM}{dt}(w_2 - w) \tag{33}$$

The terms on the right side of Eq. (33)

$$\boldsymbol{\Phi}_x = \frac{dM}{dt}(u_2 - u), \quad \Phi_y = \frac{dM}{dt}(v_2 - v), \quad \Phi_z = \frac{dM}{dt}(w_2 - w) \tag{34}$$

are called the projections of the impact force. The impact force is the consequence of mass variation.

For the reference system $S\xi\eta\zeta$ fixed to the body with the origin in the center of mass of the body the inertial tensor $\mathbf{I}$ has nine components: $I_{\xi\xi}, I_{\eta\eta}, I_{ss}$ are the moments of inertia and $I_{\xi\eta}, I_{\xi s}, I_{\eta s}$ and also $I_{\eta\xi}, I_{s\xi}, I_{s\eta}$ are the products of inertia. If the axes are principal and products of inertia are zero simultaneously the inertial tensor $\mathbf{I}$ has only three principal moments of inertia $I_{\xi\xi}, I_{\eta\eta}, I_{ss}$. The angular velocity $\boldsymbol{\Omega}$ has three components p, q, r in this frame. If p_2, q_2, r_2 are the components of the angular velocity $\boldsymbol{\Omega}_2$, M_ξ, M_η, and M_ζ are the body-axis components of $\mathbf{M}_S$ and $M_\xi^{\boldsymbol{\Phi}}$, $M_\eta^{\boldsymbol{\Phi}}$, and $M_s^{\boldsymbol{\Phi}}$ are the body-axis components of $\mathbf{M}_S^{\boldsymbol{\Phi}}$, the vector equation for rotational motion Eq. (29) in the form of three scalar equations is

$$I_{\xi\xi}\frac{dp}{dt} + (I_{ss} - I_{\eta\eta})qr = M_\xi + M_\xi^{\boldsymbol{\Phi}} + \frac{dI_{\xi\xi}}{dt}(p - p_2)$$

$$I_{\eta\eta}\frac{dq}{dt} + (I_{\xi\xi} - I_{ss})pr = M_\eta + M_\eta^{\boldsymbol{\Phi}} + \frac{dI_{\eta\eta}}{dt}(q - q_2)$$

$$I_{ss}\frac{dr}{dt} + (I_{\eta\eta} - I_{\xi\xi})pq = M_s + M_s^{\boldsymbol{\Phi}} + \frac{dI_{ss}}{dt}(r - r_2) \tag{35}$$

The terms on the right side of Eq. (35)

$$\mathfrak{R}_\xi = \frac{dI_{\xi\xi}}{dt}(p_2 - p), \quad \mathfrak{R}_\eta = \frac{dI_{\eta\eta}}{dt}(q_2 - q), \quad \mathfrak{R}_\zeta = \frac{dI_{ss}}{dt}(r_2 - r) \tag{36}$$

are called the projections of the impact torque. The impact torque is the consequence of variation of moment of inertia of the body.

The system of differential Eqs. (33) and (35) describe the general motion of the body with added mass.

For the case when mass separates the differential equations of motion have the same form as Eqs. (33) and (35) but the signs of separating mass dM/dt and separating moment of inertia dI/dt are negative.

3 Discussion of the Differential Equations of Motion

Comparing the differential equations of general motion of the rigid body with continual body variation Eqs. (33) and (35) with Eq. (6) it can be concluded that in the previous consideration the impact torque is more general. Namely, the impact torque

$$\mathbb{R} = -\frac{d\mathbf{I}}{dt}\boldsymbol{\Omega} \tag{37}$$

represents only a special case of the impact torque Eq. (32)

$$\mathbb{R} = \frac{d\mathbf{I}}{dt}(\boldsymbol{\Omega}_2 - \boldsymbol{\Omega}) \tag{38}$$

where $\boldsymbol{\Omega}_2 = 0$. For the case when the absolute angular velocity of the added or separated body differs from the angular velocity of the existing body the impact torque is defined by Eq. (38) and if the absolute angular velocity is zero the impact torque is defined by Eq. (37).

For the special case when the absolute velocity of the added or separated body is equal to the angular velocity of the existing body $\boldsymbol{\Omega}_2 = \boldsymbol{\Omega}$ the reactive torque Eq. (38) is obtained to be zero

$$\mathbb{R} = 0 \tag{39}$$

In the paper of Meshchersky [8] this special case is considered for the rotating body.

The equations of rigid body motion represent the special case of those of the systems which mass and moment of inertia vary due to add or separating the mass. Namely, for constant mass and moment of inertia the impact force $\boldsymbol{\Phi}$ and the impact torque $\mathbb{R}$ are equal to zero and the equations of motion Eqs. (30) and (31) become

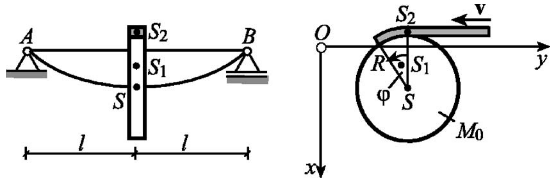


Fig. 2 Model of the drum with winding band

$$M\frac{d\mathbf{v}_S}{dt}=\mathbf{F}_r,\quad \mathbf{I}_S\frac{d\boldsymbol{\Omega}}{dt}+\boldsymbol{\Omega}\times\mathbf{I}_S\boldsymbol{\Omega}=\mathbf{M}_S \tag{40}$$

4 Band is Winding Up on a Drum

Let us analyze the winding up of the band on a drum. The elastic properties of the band at the moment of adding to drum are omitted and are assumed to be inextensible. In general the drum on which the band winds up has plane motion. In the fixed inertial system plotted in Fig. 2 the differential equations of the plane motion of the drum on which the band winding up, according to Eqs. (33) and (35), are

$$\frac{d}{dt}(M\dot{x}_{S1})=F_x+\frac{dM}{dt}v_{xb}$$

$$\frac{d}{dt}(M\dot{y}_{S1})=F_y+\frac{dM}{dt}v_{yb}$$

$$\frac{d}{dt}(I_{S1}\dot{\varphi})=M_{S1}+M_{S1}^{\boldsymbol{\Phi}}+\frac{dI_{S1}}{dt}\Omega_b \tag{41}$$

where

$$M_{S1}^{\boldsymbol{\Phi}}=\frac{dM}{dt}[(\mathbf{S}_1\mathbf{S}_2)_x(v_{yb}-\dot{y}_{S1})-(\mathbf{S}_1\mathbf{S}_2)_y(v_{xb}-\dot{x}_{S1})] \tag{42}$$

$(\mathbf{S}_1\mathbf{S}_2)_x$ and $(\mathbf{S}_1\mathbf{S}_2)_y$ are projections of the position vector of the point of mass addition due to the mass center S_1; Ω_b is the angular velocity of the winding band; v_{xb} and v_{yb} are the projections of the linear velocity of the winding band; M is the mass of the drum with the band; I_{S1} is moment of inertia of the drum with band; $\dot{x}_{S1}$ and $\dot{y}_{S1}$ are the projections of the velocity of the mass center of the drum with band; and $\dot{\varphi}$ is the angular velocity of the drum with band.

The technical requirement for winding up the band is for the absolute velocity of the band v_b to be constant. Only for that condition is rolling up the band on the drum accurate without crumpling the band or its plucking. The band is moving translatory with velocity v horizontally, parallel to the y axle in Fig. 2. The projections of the band velocity are

$$v_{xb}=0,\quad v_{yb}=v \tag{43}$$

Rolling up of one band layer is discussed. The angle of rolling up of the band is in the interval from $\varphi=0$ to $\varphi=2\pi$.

4.1 The Geometric and Physical Properties of the Drum With Band. If the mass of the drum with unrolled band is M_0 and the rolling mass is M_r

$$M_r=\mu\varphi \tag{44}$$

where

$$\mu=Rhb\rho \tag{45}$$

h is the thickness; b is the width; and ρ is the density of band; the mass M variation is a linear function of the angle φ

$$M=M_0+M_r=M_0+\mu\varphi \tag{46}$$

The position of mass center of the drum with unrolled mass is

$$SS_1=\frac{M_r}{M}(SS') \tag{47}$$

where SS' is the distance of the mass center of the unrolled mass on the drum

$$SS'=R\,\frac{\sin(\varphi/2)}{\varphi/2} \tag{48}$$

According to relations (46)–(48) the distance between the mass center of the whole system and the rotation center is obtained

$$SS_1=\frac{2\mu}{M_0+\mu\varphi}\sin\left(\frac{\varphi}{2}\right) \tag{49}$$

The moment of inertia of the drum with the unrolled mass is J_0 and the moment of inertia of the band which is rolling up is

$$J_r=\int_0^{\varphi}R^2\,dM_r=\int_0^{\varphi}R^3hb\rho\,d\varphi=j\varphi \tag{50}$$

where $j=R^3hb\rho=\mu R^2$ is the unit moment of inertia. The total moment of inertia is obtained by superposition of both moments of inertia

$$J_S=J_0+j\varphi \tag{51}$$

Applying the Steiner theorem the moment of inertia for the parallel axis settled in the mass center is obtained

$$J_{S1}=J_S-M(SS_1)^2 \tag{52}$$

4.2 Forces Acting on the System. During winding up of the band the following forces act: the elastic force of the shaft, the damping torque, the impact force, and the impact torque.

The elastic force of the shaft is projected in the fixed coordinate system

$$F_x=-cx_S=-c\left(x_{S1}-SS_1\cos\frac{\varphi}{2}\right),$$

$$F_y=-cy_S=-c\left(y_{S1}-SS_1\sin\frac{\varphi}{2}\right) \tag{53}$$

where c is the rigidity of the shaft.

According to Eqs. (41) and (43) the projections of the impact force Φ and the impact torque $\Im$ are obtained

$$\Phi_x=\frac{dM}{dt}(-\dot{x}_{S1}),\quad \Phi_y=\frac{dM}{dt}(v-\dot{y}_{S1}),\quad \Re=\frac{dI_{S1}}{dt}(-\dot{\varphi})$$

If the rotational damping torque acts

$$M_D=-D\dot{\varphi} \tag{54}$$

where D is the damping coefficient, and the moment of the impact force according to S_1 is considered, the differential equations of the plane motion is obtained

$$M\ddot{x}_{S1}+cx_{S1}=c(SS_1)\cos\frac{\varphi}{2}+\frac{dM}{dt}(-\dot{x}_{S1})$$

$$M\ddot{y}_{S1}+cy_{S1}=c(SS_1)\sin\frac{\varphi}{2}+\frac{dM}{dt}(v-\dot{y}_{S1})$$

$$J_{S1}\ddot{\varphi}+D\dot{\varphi}=\frac{dJ_{S1}}{dt}(-\dot{\varphi})+x_{S1}c(SS_1)\sin\frac{\varphi}{2}-y_{S1}c(SS_1)\cos\frac{\varphi}{2}$$
$$+\frac{dM}{dt}(v-\dot{y}_{S1})(\mathbf{S}_1\mathbf{S}_2)_x-\frac{dM}{dt}(-\dot{x}_{S1})(\mathbf{S}_1\mathbf{S}_2)_y \tag{55}$$

Analyzing relation (49) it can be concluded that $SS_1\ll 1$ and in the first approximation the terms with SS_1 in Eq. (55) can be omitted as the small values. Using relations (46), (51), and (52) the differential Eq. (55) simplifies to

$$M\ddot{x}_S + cx_S = -\mu\dot{\varphi}\dot{x}_S, \quad M\ddot{y}_S + cy_S = \mu\dot{\varphi}(v - \dot{y}_S) \tag{56}$$

$$J_S\ddot{\varphi} + D\dot{\varphi} = -j\dot{\varphi}^2 + \mu\dot{\varphi}(v - \dot{y}_S)(\mathbf{SS}_2)_x + \mu\dot{\varphi}\dot{x}_S(\mathbf{SS}_2)_y \tag{57}$$

The system of differential Eqs. (56) and (57) is nonlinear.

4.3 The Shaft is Rigid. If the shaft of the drum with the winding up band is rigid the motion of the system transforms to a rotation around the rigid axle ($x_S = y_S = 0$)

$$(J_0 + j\varphi)\ddot{\varphi} + D\dot{\varphi} = -j\dot{\varphi}^2 + \mu\dot{\varphi}vR \tag{58}$$

Introducing the new variable $u(\varphi) = \dot{\varphi}$ the differential Eq. (58) is transformed to the Bernoulli equation

$$(J_0 + j\varphi)\frac{du}{d\varphi} + ju = (\mu vR - D) \tag{59}$$

whose solution for the initial condition $\dot{\varphi}(0) = \Omega_b$ has the form

$$\dot{\varphi} = \frac{J_0\Omega_b + (\mu vR - D)\varphi}{J_0 + j\varphi} \tag{60}$$

Relation (60) describes the variation of the angular velocity of the drum when the absolute velocity v of the winding band is constant. The angular velocity of the drum decreases during winding up of a layer.

Integrating the differential Eq. (60) for the initial angle $\varphi(0) = 0$ the time history of angle variation is obtained

$$\varphi + \left(\frac{J_0}{j} - \frac{J_0\Omega_b}{\mu vR - D}\right)\ln\left|1 + \varphi\frac{\mu vR - D}{J_0\Omega_b}\right| = \frac{\mu vR - D}{j}t \tag{61}$$

This form of solution is not convenient for discussion. Introducing the new variable

$$r = 1 + \frac{\mu vR - D}{J_0\Omega_b}\varphi \tag{62}$$

Eq. (61) is

$$r + \frac{D}{j\Omega_b}\ln r = 1 + \frac{(\mu vR - D)^2}{J_0 j\Omega_b}t \tag{63}$$

Let us introduce the new function

$$w = -\ln r - \frac{f_1}{k_1} \tag{64}$$

where

$$f_1 = 1 + \frac{(\mu vR - D)^2}{J_0 j\Omega_b}t, \quad k_1 = -\frac{\mu vR - D - j\Omega_b}{j\Omega_b} \tag{65}$$

After substituting Eq. (64) into Eq. (61) and some transformation the obtained result is

$$w\exp(w) = x \tag{66}$$

where

$$x = \frac{1}{k_1}\exp\left(-\frac{f_1}{k_1}\right) \tag{67}$$

The solution $w(x)$ of Eq. (66) is the Lambert's w function [17]

$$w(x) \equiv \text{lambert}\, w\left[\frac{1}{k_1}\exp\left(-\frac{f_1}{k_1}\right)\right] \tag{68}$$

Substituting into Eq. (64) the solution for r is obtained

$$r = -k_1\left\{\text{lambert}\, w\left[-\frac{1}{k_1}\exp\left(-\frac{f_1}{k_1}\right)\right]\right\} \equiv -k_1 w(x) \tag{69}$$

which gives the implicit solution for Eq. (61)

$$\varphi = J_0\left(\frac{1}{j} - \frac{\Omega_b}{\mu vR - D}\right)w - \frac{J_0\Omega_b}{\mu vR - D} \tag{70}$$

For the case when damping is neglected and assuming that $v = \Omega_b R$ relation (61) is simplified and the angle time function is linear

$$\varphi = \Omega_b t \tag{71}$$

4.4 The Shaft is Elastic. Let us transform the differential Eqs. (56) introducing the variables

$$x_S = x(\varphi), \quad \dot{x}_S = \frac{dx}{d\varphi}\dot{\varphi}, \quad \ddot{x}_S = \frac{d^2x}{d\varphi^2}\dot{\varphi}^2 + \frac{dx}{d\varphi}\ddot{\varphi} \tag{72}$$

$$y_S = y(\varphi), \quad \dot{y}_S = \frac{dy}{d\varphi}\dot{\varphi}, \quad \ddot{y}_S = \frac{d^2y}{d\varphi^2}\dot{\varphi}^2 + \frac{dy}{d\varphi}\ddot{\varphi} \tag{73}$$

The obtained system of differential equations of plane motion is

$$(M_0 + \mu\varphi)\dot{\varphi}^2\frac{d^2x}{d\varphi^2} + \frac{dx}{d\varphi}[\mu\dot{\varphi}^2 + (M_0 + \mu\varphi)\ddot{\varphi}] + cx = 0 \tag{74}$$

$$(M_0 + \mu\varphi)\dot{\varphi}^2\frac{d^2y}{d\varphi^2} + \frac{dy}{d\varphi}[\mu\dot{\varphi}^2 + (M_0 + \mu\varphi)\ddot{\varphi}] + cy = \mu\dot{\varphi}v \tag{75}$$

$$J_S\ddot{\varphi} + (D - \mu vR)\dot{\varphi} + j\dot{\varphi}^2 = -\mu\dot{\varphi}^2 R\frac{dy}{d\varphi} \tag{76}$$

Substituting Eq. (60) into Eqs. (74) and (75), assuming that μ/M_0, j/J_0, μ/J_0, and $D/J_0\Omega_b$ are small parameters, the simplified differential equations are formed

$$\frac{d^2x}{d\varphi^2} + 2\delta\frac{dx}{d\varphi} + \omega^2(\varphi)x = 0 \tag{77}$$

$$\frac{d^2y}{d\varphi^2} + 2\delta\frac{dy}{d\varphi} + \omega^2(\varphi)y = \frac{\mu}{M_0}R \tag{78}$$

$$\ddot{\varphi} + \frac{D - \mu Rv}{J_S}\dot{\varphi} + \frac{j}{J_S}\dot{\varphi}^2 = -\frac{\mu}{J_0}R\dot{\varphi}^2\frac{dy}{d\varphi} \tag{79}$$

where

$$2\delta = 2\frac{\mu}{M_0} - \frac{j}{J_0} - \frac{D}{J_0\Omega_b}, \quad \omega^2(\varphi) \equiv \omega^2 = k^2(1 - A\varphi),$$

$$k^2 = \frac{\omega_1^2}{\Omega_b^2}, \quad \omega_1^2 = \frac{c}{M_0}, \quad A = 3\frac{\mu}{M_0} - 2\frac{j}{J_0} - \frac{D}{J_0\Omega_b}$$

To obtain the approximate analytic solutions of Eqs. (77)–(79) the Bogolubov–Mitropolski method is modified for the nonhomogenous rheo-linear differential equations.

Omitting the terms on the right side of Eq. (79) as small values the approximate solution of Eq. (79) corresponds to the case of rigid shaft Eq. (60). Substituting Eq. (60) into Eq. (78) the solution of the differential Eq. (78) is assumed as

$$y = a(\varphi)\exp(-\delta\varphi)\cos\Psi(\varphi) + \frac{1}{\omega^2(\varphi)}\frac{\mu}{M_0}R \equiv a\exp(-\delta\varphi)\cos\Psi + \frac{1}{\omega^2(\varphi)}\frac{\mu}{M_0}R \tag{80}$$

where

$$\Psi(\varphi) = \int \omega(\varphi)d\varphi + \alpha(\varphi) \tag{81}$$

and the first derivative of the function y

$$\frac{dy}{d\varphi} = [-\delta a \cos\Psi - a\omega(\varphi)\sin\Psi]\exp(-\delta\varphi) - \left(\frac{\mu}{M_0}\right)^2 \frac{R}{k^2(1-A\varphi)^2} \tag{82}$$

with

$$\frac{da}{d\varphi}\cos\Psi - a\frac{d\alpha}{d\varphi}\sin\Psi = 0 \tag{83}$$

Eliminating the second-order small term in Eq. (82) the relation transforms to

$$\frac{dy}{d\varphi} \approx [-\delta a \cos\Psi - a\omega(\varphi)\sin\Psi]\exp(-\delta\varphi) \tag{84}$$

Using relations (80)–(84) the differential Eq. (78) is transformed into a system of two first-order differential equations

$$\frac{da}{d\varphi} = -\frac{a}{\omega}\frac{d\omega}{d\varphi}\sin^2\Psi, \quad \frac{d\alpha}{d\varphi} = \frac{1}{2\omega}\frac{d\omega}{d\varphi}\sin 2\Psi \tag{85}$$

It is at this point where the averaging procedure $1/2\pi\int_0^{2\pi}(\bullet)d\Psi$ is introduced and relations (85) are simplified into

$$\frac{da}{d\varphi} = -\frac{a}{2\omega}\frac{d\omega}{d\varphi}, \quad \frac{d\alpha}{d\varphi} = 0 \tag{86}$$

For the initial conditions

$$\varphi = 0, \quad a = a_0, \quad \alpha = \alpha_0 \tag{87}$$

the solution of Eq. (78) in the first approximation is

$$y_S = y(\varphi) = \frac{a_0}{\sqrt[4]{1-A\varphi}}\exp(-\delta\varphi)\cos[k\sqrt{(1-A\varphi)} + \alpha_0] + \frac{R}{k^2}\frac{\mu}{M_0} \tag{88}$$

According to the suggested procedure the solution of Eq. (77) is

$$x_S = x(\varphi) = \frac{b_0}{\sqrt[4]{1-A\varphi}}\exp(-\delta\varphi)\cos[k\sqrt{(1-A\varphi)} + \beta_0] \tag{89}$$

where b_0 and β_0 are initial amplitude and phase. The parameter values have to satisfy the relation

$$3\frac{\mu}{M_0} - 2\frac{j}{J_0} - \frac{D}{J_0\Omega_b} < \frac{1}{2\pi} \tag{90}$$

The motion of the rotor center depends on the ratio between the small parameters μ/M_0, j/J_0, and $D/J_0\Omega_b$. For small value of the rotational damping and higher velocity of the rolling band the vibrations decrease.

Using the obtained solution Eq. (88) the correction for the angle velocity Eq. (60) can be denoted. Due to the fact that $\dot{y}_S$ tends to zero for technical reasons relation (60) is assumed to be accurate enough.

5 Conclusion

During the process of continual mass variation, the mass and moment of inertia of the rigid body vary due to adding or separating of mass in the short infinitesimal time interval: mass but also the form and the volume of the body are continually varying in time. This causes the body mass center position variation and also the change of the moment of inertia and the products of moment of inertia. Due to mass and moment of inertia variation the impact force and impact torque act. Namely, the absolute velocity and angular velocity of addition or separation differs in general from the velocity of mass center and angular velocity of the existing body and it causes the impact to occur. As the mass variation is continual the impact is substituted with an "impact force" and "impact torque" which continually act on the body. The force and torque depend on the absolute velocity of mass center and angular velocity of the separated or added body.

For rolling up the band on the drum mass and moment of inertia of the drum with band varies. Mass and moment of inertia depend on the angle position of the wound band. Due to geometry variation of the drum with band the mass center position inside the system also varies. This variation seems to be small and is neglected in our consideration. During winding up of the band on the drum the impact occurs due to the difference in velocity and angular velocity of the band and drum. It causes the vibrations of the mass center of the drum. The vibrations of drum mass center depend on the amount of the band winding up on the drum: the higher the amount of band on the drum the smaller the vibrations. The damping property of the drum also has an influence on the vibrations: the higher the damping the smaller the vibrations.

The band is wound up with constant velocity. This requires the angle velocity of the drum to vary. The angle velocity variation is the function of the moment of inertia of the band, which winds up and also of the damping properties of the system: for higher damping the angle velocity decreases faster than for smaller damping; the larger the moment of inertia of the wound up band the slower the decrease of the angular velocity. This result is of technical importance for regulating the rotation of the drum.

References

[1] Cayley, A., 1857, "On a Class of Dynamical Problems," Proc. R. Soc. London, **8**, pp. 506–511.
[2] Cayley, A., 1858, "On a Class of Dynamical Problems," Philos. Mag., **XV**, pp. 306–310.
[3] Cayley, A., 1859, The Collected Mathematical Papers, **IV**(225), pp. 7–11.
[4] Dufour, Ch., 1866, "Sur l'Acceleration Seculaire du Mouvement de la Lune," C. R. Seances Acad. Sci. III, **LXII**, pp. 840–842.
[5] Oppalzer, A., 1884, "Ueber die Ursache, Welche den Unterschied Zwischen der Theoretischen Berechneten Secularacceleration in der Laenge des Mondes und der Thatsaechlichen Bedingen Kann," Astron. Nachr., **108**(2573), pp. 67–72.
[6] Gylden, H., 1884, "Die Bahnbewegungen in Einem Systeme von Zwei Koerpern in dem Falle das Massen Veraenderungen Unterworfen Sind," Astron. Nachr., **109**(2593), pp. 1–6.
[7] Meshcherski, I. V., 1893, "Odin Chasnij Sluchaj Zadachi Gouldena," Astron. Nachr., **132**(3153), p. 9.
[8] Meshcherski, I. V., 1897, "Dinamika Tochki Peremennoj Massji," Magistarskaja dissertacija, Petersburgski Universitet, Petersburg, USSR.
[9] Cveticanin, L., 1992, "The Influence of the Reactive Force on a Nonlinear Oscillator with Variable Parameter," Trans. ASME, J. Vib. Acoust., **114**(4), pp. 578–580.
[10] Cveticanin, L., 1993, "An Approximative Solution of a Coupled Differential Equation with Variable Parameter," Trans. ASME, J. Appl. Mech., **60**(1), pp. 214–217.
[11] Cveticanin, L., 1993, "The Influence of the Reactive Force on the Motion of the Rotor on Which the Band is Winding Up," J. Sound Vib., **167**(2), pp. 382–384.
[12] Cveticanin, L., 1996, "The Influence of the Reactive Force on the Stability of Motion for One-Degree-Freedom Mechanisms with Variable Mass," *Machine Vibration*, **5**, Springer-Verlag, London, pp. 224–228.
[13] Cveticanin, L., 1998, *Dynamics of Machines with Variable Mass*, Gordon and Breach Science Publishers, London, UK.
[14] Cornelisse, J. W., Schoyer, H. F. R., and Wakker, K. F., 1979, *Rocket Propulsion and Spaceflight Dynamics*, Pitman, London, UK.
[15] Bessonov, A. P., 1967, *Osnovji Dinamiki Mehanizmov s Peremennoj Massoj Zvenjev*, Nauka, Moscow, USSR.
[16] Goldstein, H., 1980, *Classical Mechanics*, Addison-Wesley, Reading, MA.
[17] Weisstein, E. W., 2007, "Lambert W-Function," MathWorld—A Wolfram Web Resource, http://mathworld.wolfram.com/LambertW-Function.html

Journal of Applied Mechanics

Technical Briefs

Use of Near-Field Acoustic Levitation in Experimental Sliding Contact

T. A. Stolarski
e-mail: mesttas@brunel.ac.uk

C. I. Woolliscroft

Department of Mechanical Engineering,
School of Engineering and Design,
Brunel University,
Uxbridge, Middlesex UB8 3PH,
United Kingdom

The paper presents an investigation into producing the self-levitation effect using piezo-electric actuators (PZTs). Self-levitation has been demonstrated and results are presented and discussed. A relationship between the levitation distance and weight of the levitating sample has been found. In addition, the orientation and position of the PZTs has been found to affect the levitation distance. Modal shapes of the vibration plates used have been produced through modeling and found to accurately correlate with the experimental results found. Additional evidence suggests that the type of vibration plate material affects the separation distance, possibly due to the material's properties of acoustic reflection. [DOI: 10.1115/1.2424472]

1 Introduction

Increasingly in modern technology the demand for engineering components is to become smaller and smaller. This opens up a completely new area of microtechnology and precision engineering in the nanometer scale. One of the main problems with manufacturing components at such a small scale with equipment such as laser cutters, optical scanners, and silicon chip production tools is the accuracy of any movement at this scale. Static friction has a major effect on controlled precise movements at the nanometer scale. Near frictionless operation of machinery for nanometre accuracy is a definite prerequisite.

Producing frictionless movement can be achieved in a number of ways. However, all currently used techniques are not without certain disadvantages of a practical nature. Near-field acoustic levitation (NFAL) [1–3] is an interesting alternative for near frictionless operation of high precision machinery. NFAL is produced when rapid vibrations between two surfaces close together in a compressible fluid medium such as air create a load carrying gas film. This gas film has a pressure greater than the ambient pressure and so can support a load. The production of rapid vibrations can be achieved by some form of piezo-electric actuator (PZT) as the frequency required is particularly high. The separation between the two surfaces can be very small indeed, in some cases down to a distance of just 10 μm.

Acoustic levitation of solids has been suggested for use in noncontact transportation of glass substrate of liquid crystal display (LCD) in its manufacture [1]. Since the acoustic levitation effect can be used to move objects as well as cause them to levitate, many other applications would be applicable to acoustic levitation where relatively light objects need transporting accurately.

This paper reports on a sliding contact operating on NFAL principle. The ultrasonic frequency range used is inaudible so operation of any NFAL contact should be kept within this region if NFAL is to be recognized as a plausible method for use in industry. The main objective of the study presented was to find proof of evidence demonstrating that the NFAL contact is practically feasible and operates due to the emission of an ultrasonic acoustic wave from a parallel surface. Once the acoustic levitation concept was confirmed it was then possible to investigate relationships between the levitation effect and various parameters affecting it [4].

2 Experimental Setup

In order to investigate the possibility of self-lifting generation due to NFAL a plain rectangular piece of material was rigidly clamped at each end as shown in Fig. 1 [5]. The initial dimensions of the rectangular plate were $200\times100\times5$ mm and it was made from aluminum. The plate had two foil-type piezoelectric actuators on the underside of it, which produce Poisson's ratio contraction effect when operated. In addition, these two PZTs generate an ultrasonic acoustic wave from the surface of the plate to allow small mass objects to levitate utilizing the NFAL effect.

The test rig main structure is made of mild steel and holds the vibrating plate clamped in the horizontal plane. The vibrating plate is clamped at both ends to the steel supporting structure. The three feet to the base are adjustable so that the testing rig can be balanced with the use of a spirit level when set up on a bench.

Vibrating plates used in experiments were: the 1 mm aluminum plate, 1.9 mm aluminum plate, 1.55 mm titanium plate, and the 1.1 mm steel plate. The 1.9 mm aluminum plate was tested first and good results were obtained. The 1 mm aluminum plate was then tested, as thinner plates should lead to higher amplitude of vibrations due to the lower mass of the plate being accelerated and the increase in deformation. The high carbon steel plate was chosen to obtain ground steel gauge plate, as it was thought that the smooth ground finish would increase the NFAL performance. The titanium plate was tested because titanium has certain elastic properties that were thought to be favorable to the NFAL effect.

The floating samples used for the experimental testing had to have five very important features.

1. The sample needed to be as flat and smooth as possible on the side facing the vibrating plate.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received December 20, 2005; final manuscript received May 22, 2006. Review conducted by Thomas W. Shield.

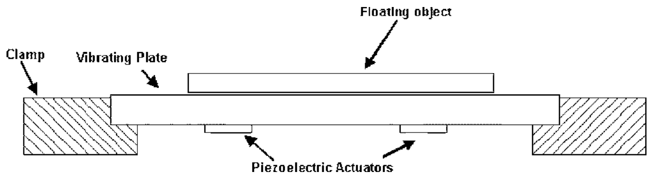


Fig. 1 Piezoelectric driven plate and levitating sample

2. The samples needed to be relatively light. The amount of weight per unit area that can be supported by the NFAL effect may be very small with the small size of the PZTs driving the plate.
3. The floating samples needed to be rigid. If the samples were not rigid then the acoustic waves emitted from the vibrating plate would be partially absorbed by the sample. The levitating sample is a reflector of ultrasonic acoustic waves produced by the vibrating plate. This is a fundamental part of the NFAL principle, having a rigid sample will greatly improve the performance of the NFAL effect observed.
4. The sample had to be at least 1.5 wavelengths of the flexural wavelength of the vibrating plate to levitate stably. Through modeling, it was discovered that the floating sample would need to be at least 30 mm in most modes of vibration for all of the plates.
5. The top surface of the sample had to be metallic. The proximity probe, used to measure separation, can sense only conductive surfaces.

The most readily available samples that were found to cover all of these criteria were glass discs of 49.5 mm diameter and 3 mm thickness and acrylic discs of 49.5 mm diameter and 6 mm thickness. Both of these samples agree with all of the above conditions, although neither of the samples were metallic.

To use the glass and acrylic samples with the proximity probe the upper surface of the sample needed to be metallic. This was overcome by having both of the samples gold plated. The gold plating technique produces a very thin layer of less than 0.1 μm.

However, the layer is sufficient to be detected by the capacitance proximity sensor and thin enough not to affect the sample's smooth surface.

Foil type PZTs were used and they are known to change length when a voltage is put across them in order to vibrate the plate on the rig. These PZTs are capable of operating at ultrasonic frequencies and were suitable for this application. The PZTs were bonded to the plates with epoxy resin in the manner shown in Fig, 2. Initially, two PZTs were used. Their positions were decided as the best position over the plate with regard to the most flexural part of the plate: one element in the center of each half of the plate less the clamped ends of the plate. Also the orientation of the PZTs was chosen to be lengthways as in this direction the plate has more material and so is more flexural.

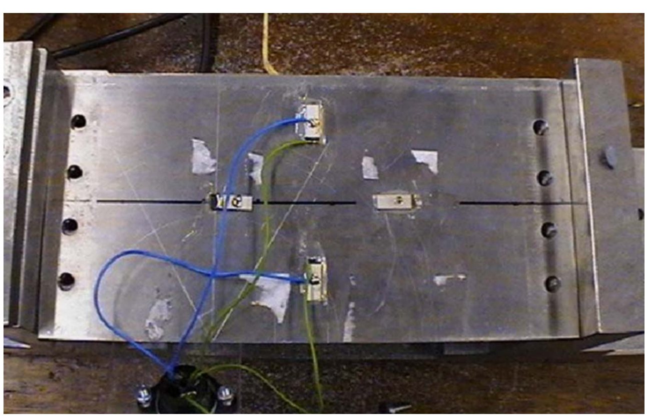


Fig. 2 Vibration plate with four PZTs bonded to the underside of the plate

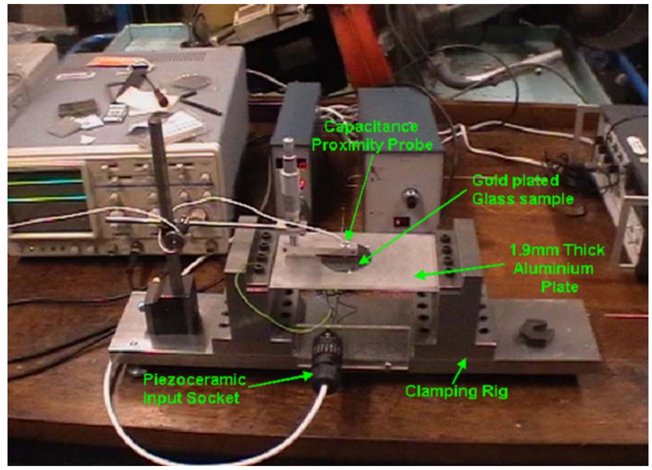


Fig. 3 Apparatus setup

3 Experimental Apparatus and Procedures

The apparatus, as used during experimental investigations, is shown in Fig. 3. The input equipment consists of a 110 V transformer and a sine wave signal generator. The 110 V transformer has to be used to drive the ultrasonic PZT amplifier and the PZT monitor. The sine wave signal generator is needed to create the sinusoidal frequency required to drive the PZTs attached to the plate. This sine wave generator is operated between the frequency range of 10 and 60 KHz and it goes well into the ultrasonic region. Two experiments of fundamental importance for the concept of NAFL were carried out, namely; (1) voltage amplitude versus levitation distance and (2) supported load per unit area (surface density) versus levitation distance.

4 Results and Discussion

The testing began with a 1.9-mm-thick aluminum plate and the lengthways orientation of the PZTs. A preliminary experiment was carried out according to the experimental procedure described earlier. A gold plated glass disc was used as the levitation sample and the plate was operated at a resonant frequency of 25.6 kHz, which was found to be the most powerful ultrasonic resonant frequency.

4.1 Effect of PZT Orientation. The experimental findings showed that the levitation distance of the floating sample was less than 5 μm at full peak to peak voltage. In order to increase the levitation distance of the floating sample two additional PZTs were added to the vibration plate laterally across the centre of the plate now giving the plate four PZTs. The levitation distances obtained with the additional elements were much higher. The results of the different orientations of the PZTs are shown in Fig. 4. It is clearly shown that the levitation distance is almost entirely dependent upon the laterally aligned PZTs and the longitudinally aligned elements do not contribute much to the NFAL effect. As the graph shows, the levitation distance versus input voltage amplitude relationship seems to be a reasonably linear for the laterally applied PZTs when just the values above the 5 μm separation distance are considered. This can also be said for the PZTs orientated longitudinally for voltage amplitudes greater then 120 V peak to peak. However, the separation distances achieved with this orientation of the PZTs is far less than the laterally orientated ones. In fact, the levitation distances attained do not even go above the critical separation distance of 5 μm and so these results could be disregarded altogether as true levitation of the sample

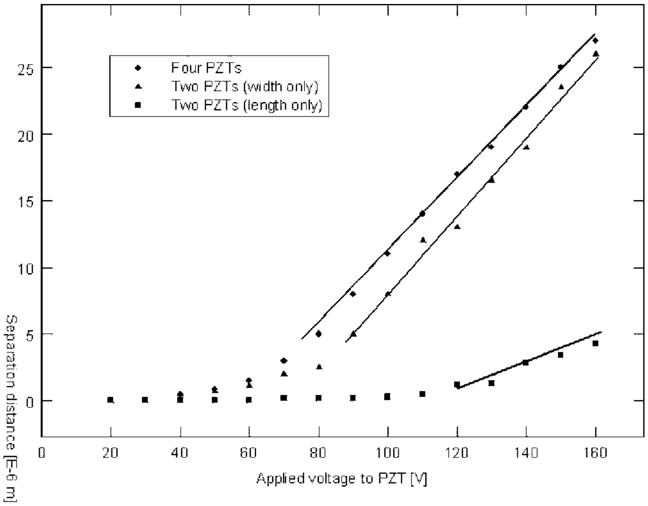


Fig. 4 Graph of separation distance versus voltage across PZTs, for three orientations of the PZTs, at 25.6 kHz

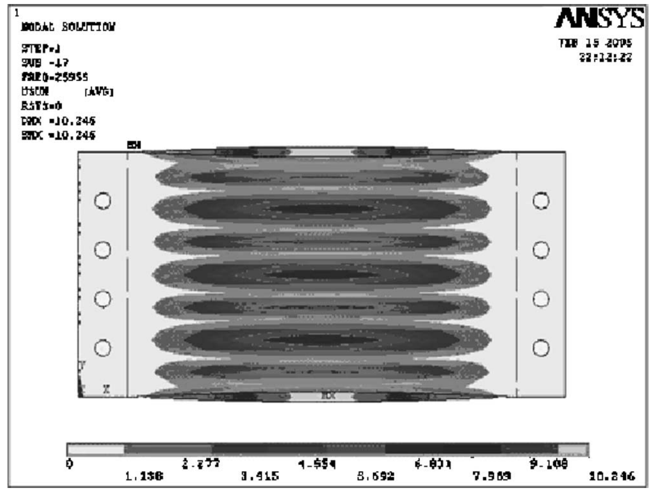


Fig. 6 Modal shape of the plate at 25.955 kHz

may not be fully achieved here and there may still be contact with the vibration plate below. The reason for the orientation of the PZTs affecting the performance of the NFAL so dramatically is thought to be that the lateral PZTs excite the plate in this mode of vibration much more. This can be clarified by looking at the modal shape of the plate at the frequency of 25.6 kHz, shown in Fig. 5. This figure demonstrates the modal shape of the plate when at its resonant frequency of 25.6 kHz. To achieve a pattern depicting the modal shape of the vibrating plate, caster sugar was simply poured onto the plate and the sugar particles gathered around nodes—the points between the oscillating crests and troughs of a standing wave of the flexural wave of the plate. The sugar particles gathered here because the nodes are stationary positions of the plate and so permit the sugar particles to reside. This mode appeared to displace the most perpendicular to the surface of the plate in comparison with other resonant frequencies found between the frequencies of 20 and 50 kHz. This was found both with the maximum levitation distance achieved at this frequency and the sugar particles being most vigorous at this frequency, as they are forced off the oscillatory parts of the plate. The experimental modal shape at this frequency agrees well with the modal shape found within finite element analysis (FEA) modeling at a frequency of 25.955 kHz (see Fig. 6). This is the only modal shape found that has eight longitudinal nodes of flexural vibration and the frequency is very close (within 1.4%) to the frequency found during testing.

The deformed shape of the model (see Fig. 7) shows, in an exaggerated scale, the modal shape of the plate when vibrating at the frequency of the 25.96 kHz model. It can be seen from Fig. 7 that the reason that the PZTs excited the plate in this mode much better when mounted laterally, is because the force of the elements was applied across the peak of a flexural wave. Unlike in the longitudinal direction, this orientation of the PZTs resulted in plate deformation similar to the plate's modal shape at this frequency. This explains why the lateral PZTs produced a levitation distance that is over six times higher than that obtained with the longitudinal PZTs. The best position for the elements to be in would be centered on a crest of one of the peaks/troughs as the shape oscillates between the two. In this position the maximum displacement of the PZT can be utilized as this position of the flexural wave is where the maximum lateral displacement occurs (see Fig. 8). The crests of the flexural wave of the plate are shown to be the most laterally displaced positions locally over the length of a PZT (10 mm) in Fig. 8.

Positioning the lateral PZTs centrally over the crest of the plate flexural waves could increase the levitation distance of the NFAL effect even further, as the current position of the elements is not centrally over the crest of the waves. If the NFAL effect were to be increased even further, then several PZTs of the same dimen-

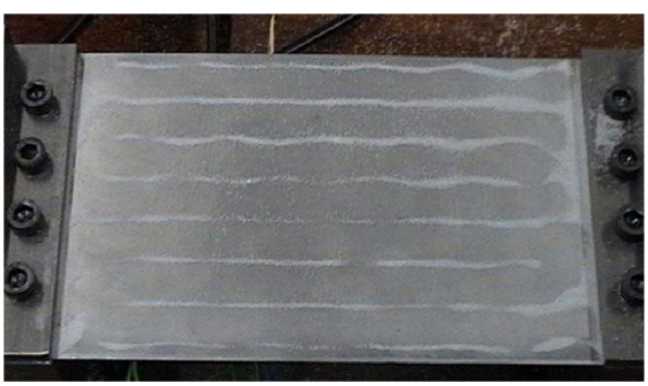


Fig. 5 The pattern of sugar upon the plate when resonating at 25.6 kHz

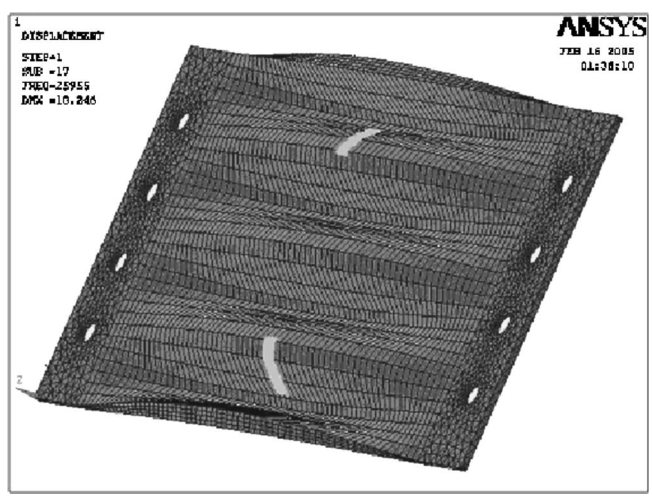


Fig. 7 Exaggerated model of the modal shape of the plate when excited to resonant frequency of 25.955 kHz

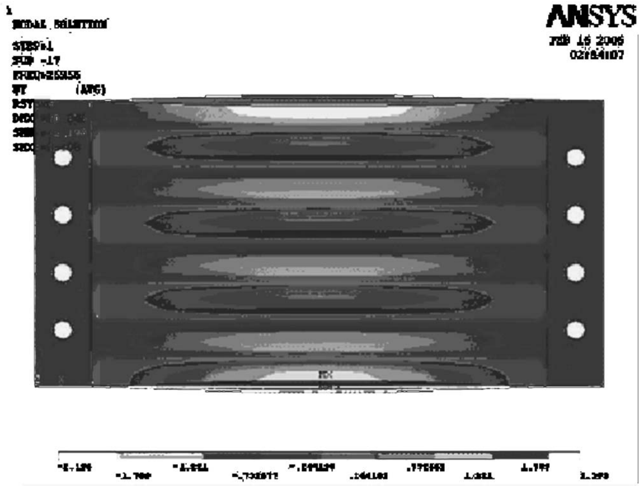


Fig. 8 Contour plot of lateral displacement when in modal shape at frequency of 25.955 kHz

sions used in this experiment could be positioned on the underside of the plate placing each PZT over the crest of one of the flexural waves of the plate.

4.2 Load Carrying Capacity. Aluminum plate with 1.9 mm thickness was used and operated at the frequency of 25.6 kHz. The levitating specimen was the gold plated glass disc. The results obtained from the testing are shown in Fig. 9. This figure shows the combined mass of the levitating sample versus the separation distance achieved for the fixed voltage input amplitude. It appears that the levitation distance is inversely proportional to the square root of the weight per unit area of the floating specimen in all flexural modes, that is

$$L = k\frac{1}{\sqrt{m}} \tag{1}$$

where L is the separation distance; m is the mass of floating specimen; and k is a constant. In order to prove the correlation with the above equation the results shown in Fig. 9 were replotted but this time in a graph of levitation distance versus the inverse of the root of the total mass on the levitating object. It is seen from Fig. 10 that a linear relationship was obtained.

The levitation distance was found to be proportional to the inverse root of the total mass of floating object with the proportionality constant k equal to $1^{*}10^{-4}$.

4.3 Effect of Vibrating Plate Material. In order to ascertain the effect of plate material on levitation two additional plates, 1.6-mm-thick titanium and 1.1-mm-thick steel were made and tested. In Fig. 11, the results obtained are compared with the set of results found for the aluminum plate of 1.9 mm thickness.

As the results show, the 1.9 mm aluminum plate produces the largest separation distances. The 1.1 mm steel plate produces the next highest separation distances and then the 1.6 mm titanium plate produces separation distances that have been recorded but are all below the 5 μm level.

In an attempt to understand the differences between the NFAL effect when using these materials, the mass of each plate must be taken into consideration. The masses involved were: 172.9 g, 144.8 g, and 103.5 g for steel, titanium, and aluminum plates, respectively. Lighter plates should give better NFAL effects as they can be accelerated greater with the same force and so the flexural wave should have higher amplitudes of displacement. This appears to be true at first as the aluminum plate did produce the greatest vibrations; however the steel plate, which is the heaviest plate of the three, is not the worst plate with regard to levitation distance. The distances produced with the steel plate were higher than the ones produced with the titanium plate.

The reason for this noncorrelation of mass of plate versus levitation distance may be because the modal shape for each plate is dependent on the dimensions of the plate and the material properties, namely the Young's modulus, Poisson's ratio, and the density of the material. In order for any material to produce a large NFAL effect the plate must have a modal shape at an achievable frequency with a given equipment, and the PZTs must by positioned and orientated in an optimal way so that maximum displacement of the plate is possible perpendicular to the surface of the plate. The 1.9 mm aluminum plate achieved these goals better than the other plates and so outperformed them by producing a larger separation distance.

5 Conclusions

Based on obtained results the following conclusions can be drawn.

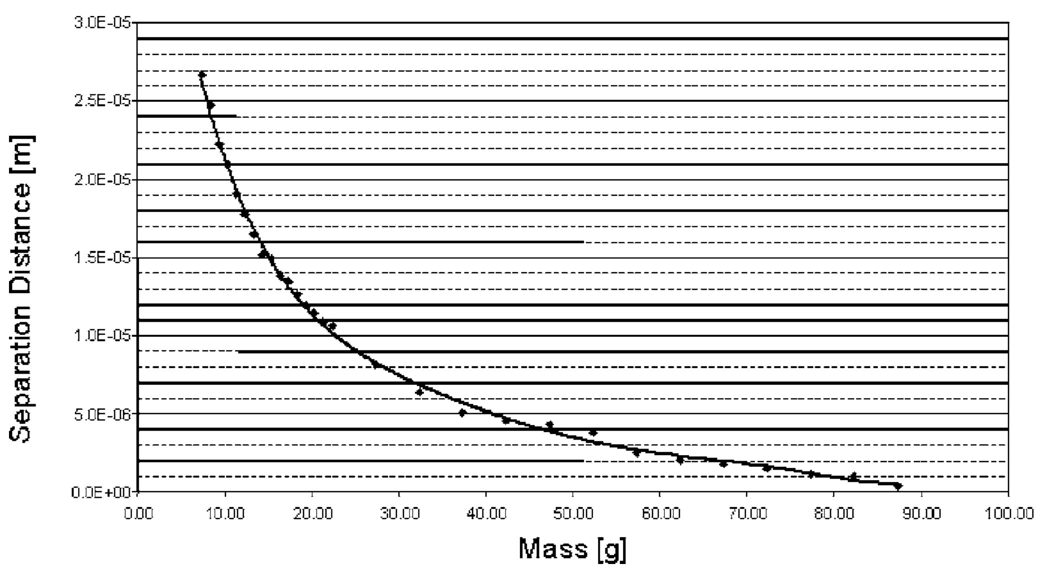


Fig. 9 Separation distance versus total mass of levitating sample

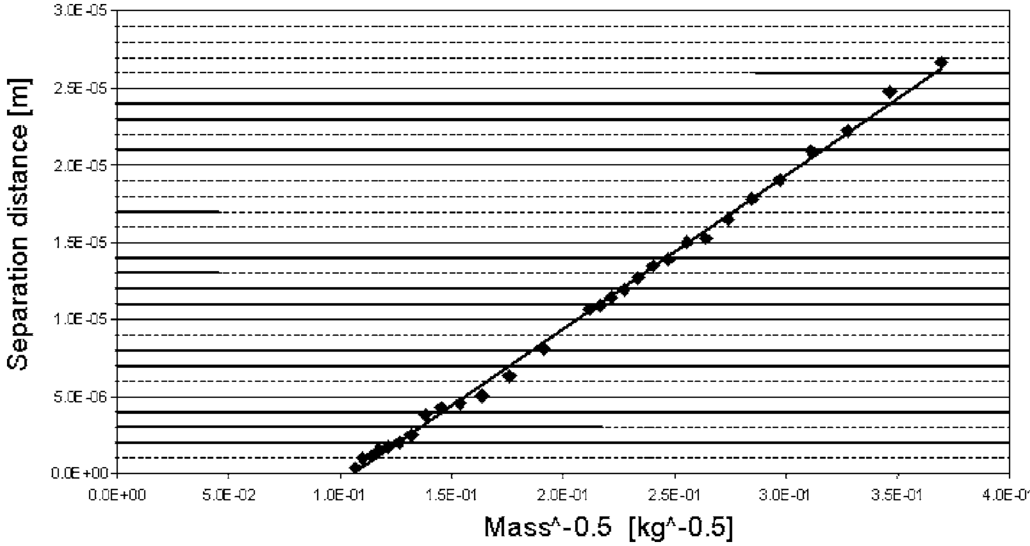


Fig. 10 Separation distance versus inverse root of the total mass of the levitation object

1. Existence of NFLA effect has been experimentally confirmed.
2. For the 1.9 mm aluminum plate, a reasonably linear relationship was found to exist between the PZT input voltage and the levitation distance at the frequency of 25.6 kHz. This can also be said for the titanium, and steel plates at their respective frequencies.

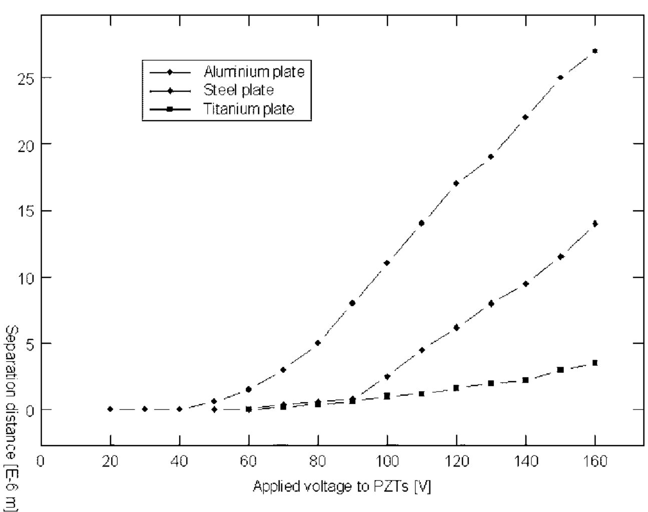


Fig. 11 Separation distance versus applied voltage input for aluminium (1.9 mm thickness), titanium (1.55 mm thickness), and steel (1.1 mm thickness) plates

3. A relationship between surface density and separation distance for the 1.9 mm aluminum plate at a frequency of 25.6 kHz was found to exist, as given by Eq. (1).
4. The orientation and position of the PZTs was found to be one of the most important factors controlling the effectiveness of the NFAL phenomenon. It has been found that attaching the elements in a suitable position on the plate is dependent upon the modal shape produced at the operating frequency of the vibrating plate and the dimensions of the PZTs.
5. Computer modeling results of modal shape agree well with experimental findings. Both the modal shape and the frequency at which it is produced show excellent correlation.
6. Vibration plate material greatly affects the separation distance created with the NFAL effect. This is because the modal shape and resonant frequencies, closely governing the NFAL effect, were also altered by the material.

References

[1] Matsuo, E., Koike, Y., Nakamura, K., Ueha, S., and Hashimoto Y., 2000, "Holding Characteristics of Planar Objects Suspended by Near-Field Acoustic Levitation," Ultrasonics, **38**, pp. 60–63.

[2] Ueha, S., Hashimoto, Y., and Koike Y., 2000, "Non-contact Transportation Using Near-Field Acoustic Levitation," Ultrasonics, **38**, pp. 26–32.

[3] Hashimoto, Y. and Ueha, S., 1996, "Near Field Acoustic Levitation of Planar Specimens Using Flexural Vibration," J. Acoust. Soc. Am., **100**, pp. 2057–2061.

[4] Woolliscroft, C., I., 2005, "Ultrasonic Acoustic Levitation," Mechanical Engineering, BEng dissertation, School of Engineering and Design, Brunel University, May.

[5] Menon, A., 2000, "Essential requirements for 100 Gbits/in.^{2} Head/Media Interface," *Proceedings International Tribology Conference*, October 29–November 2, Nagasaki, Japan.

Effect of Surface Elasticity on the Interaction Between Steps

Gan-Yun Huang

Shou-Wen Yu[1]
e-mail: yusw@mail.tsinghua.edu.cn

Department of Engineering Mechanics,
Tsinghua University,
Beijing 100084,
China

By taking into account the effect of surface elasticity, the problem of a half plane under concentrated normal or shear loads is first considered. The solutions for the displacements or alternatively named surface Green's functions can be obtained by using the Fourier integral transform technique. Based on such solutions, the elastic interaction between two surface steps that are modeled as force dipoles is further investigated. The results show that the effect of surface elasticity on the interaction energy is significant when the distance between the two steps is in the range of several times the intrinsic length scale of the system. Further, surface elasticity seems to influence the interaction between steps with force components parallel to the surface more strongly than that when the steps exhibit force components only normal to the surface. [DOI: 10.1115/1.2424473]

1 Introduction

Step-step interaction plays an important role in many different areas such as epitaxial growth and surface chemistry and therefore has been the subject of numerous fundamental studies. From the perspective of continuum analysis, Marchenko and Parshin (MP) [1] supposed that steps could be modeled as force dipole lines on flat surfaces (generally referred to as the MP model) and thus obtained that the interaction energy between two steps is in the form of d^{-2} with d being the distance between the two steps. Some experiments also verified the results [2]. However, there are also results in contradiction with the MP model. For example, Kouris et al. [3] found that the interaction energy between two steps sometimes does not follow the form of d^{-2}; Prevot et al.'s [4] and Silkrot and Srolovitz's [5] molecular dynamic simulations show the dipole strength is inconsistent with that in the MP model. To explain those results, Kukta et al. [6] have proposed a model by taking into account the geometry of the steps while Prevot and Croset [7] have employed embedded dipoles to model steps. Though much improvement has been achieved in certain circumstances, those models unanimously neglect the effect of surface elasticity, i.e., the elastic property of flat surfaces that has been recognized fairly long ago by many authors. See, for example, Shuttleworth's work [8]. It is worth mentioning that based on that work, Gurtin and Murdoch [9] have developed a generic constitutive relation for the surfaces from the perspective of continuum mechanics. Their work recently has gained some attraction in studying the mechanical behavior of nanosystems that yielded results in good agreement with those by molecular simulations [10,11].

With those in mind, in the present work we attempt to consider the effect of surface elasticity on the interaction energy between surface steps by using Gurtin and Murdoch's theory [9] and the MP model [1]. To this end, the surface Green's functions or the problem of a half plane under concentrated normal or shear loads with the effect of surface elasticity will be solved first.

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received December 28, 2005; final manuscript received April 16, 2006. Review conducted by Yonggang Huang.

[1]Corresponding author.

2 Surface Green's Functions for Surface Elasticity

As will be shown in the next section, it is important to obtain the surface Green's functions for one to study the interaction between two surface steps. To this end, let us consider a semi-infinite system loaded by tractions $F_\alpha(\alpha=1,2)$ under plane strain deformation.. According to Gurtin and Murdoch [9], the basic equations with the effect of surface elasticity can be written

$$\sigma_{\alpha\beta,\beta}=0, \quad \alpha,\beta=1,2 \tag{1}$$

in the bulk and

$$\Sigma_{11,1}+F_1=\sigma_{12}, \quad \sigma_{22}=F_2 \tag{2}$$

on the surface. $\sigma_{\alpha\beta}$ and Σ_{11} are, respectively, the conventional Cauchy stresses and surface stress. It is worth pointing out that Σ_{11} is the only nontrivial component of surface stresses in the present situation. Equation (2) in essence is the equilibrium between surface stress and bulk stress. The stresses and surface stress can be related to strain as follows

$$\sigma_{\alpha\beta}=\lambda\varepsilon_{\gamma\gamma}\delta_{\alpha\beta}+2\mu\varepsilon_{\alpha\beta}; \quad \Sigma_{11}=\tau^0+(\lambda^s+2\mu^s-\tau^0)\varepsilon_{11}(x_1,0) \tag{3}$$

in which λ and μ are Lame constants for the bulk, λ^s and μ^s Lame constants for the surface that can be obtained from atomic calculation, and τ^0 the residual surface stress that will be assumed to be zero for convenience. The strain can be expressed in terms of displacement components u_k

$$\varepsilon_{\alpha\beta}=\tfrac{1}{2}(u_{\alpha,\beta}+u_{\beta,\alpha}) \tag{4}$$

The displacements induced by $F_1=\delta(x_1), F_2=0$ (i.e., a concentrated shear force) or $F_1=0, F_2=\delta(x_1)$ (i.e., a concentrated normal force) and satisfying Eqs. (1) and (2) are defined as surface Green's functions for the present problem and can be obtained by the Fourier integral transform technique. For this purpose, substitute Eq. (4) into Eq. (3) and then into Eq. (1), and we obtain

$$(\lambda+2\mu)u_{1,11}+\mu u_{1,22}+(\lambda+\mu)u_{2,12}=0 \tag{5a}$$

$$(\lambda+\mu)u_{1,12}+\mu u_{2,11}+(\lambda+2\mu)u_{2,22}=0 \tag{5b}$$

Applying Fourier integral transform to the above equations with respect to x_1, one obtains the general solutions of displacements in the transformed domain

$$\tilde{u}_1=\frac{1}{2\mu is}\{s^2A_1+[s^2x_2+2(1-\nu)|s|x_2]A_2\}\exp(|s|x_2) \tag{6}$$

$$\tilde{u}_2=\frac{1}{2\mu}[-|s|A_1+(1-2\nu-|s|x_2)A_2]\exp(|s|x_2) \tag{7}$$

where "$\sim$" stands for Fourier integral transform, s for the transformed variable, and $i=\sqrt{-1}$, A_1, and A_2 are constants to be determined, and $\nu=\lambda/2(\lambda+\mu)$ the Poisson's ratio. Substitution of Eq. (3) into Eq. (2) and using Eqs. (6) and (7) gives

$$A_1=0, \quad A_2=\frac{i}{s(1+l|s|)} \tag{8}$$

for concentrated shear loading, and for concentrated normal loading

$$A_1 = -\frac{1}{s^2}, \quad A_2 = \frac{|s| + ls^2/2(1-\nu)}{1+l|s|}\frac{1}{s^2} \tag{9}$$

with $l=(\lambda^s+2\mu^s)/\mu(1-\nu)$ being an intrinsic length scale for the system under consideration. Consequently, the displacements induced by the concentrated loads described above can be obtained by inserting Eqs. (8) and (9), respectively, into Eqs. (6) and (7) and by further applying inverse Fourier integral transform

$$u_{11}(x_1,x_2) = -\frac{1-\nu}{2\pi\mu}\ln R^2 + \frac{1}{2\pi\mu}[x_2 - 2(1-\nu)l]g_1 \tag{10}$$

$$u_{21}(x_1,x_2) = -\frac{1-2\nu}{2\pi\mu}\tan^{-1}(x_2/x_1) + \frac{1}{2\pi\mu}[l(1-2\nu)+x_2]g_2 \tag{11}$$

$$u_{12}(x_1,x_2) = \frac{1}{2\pi\mu}\left\{(1-2\nu)\tan^{-1}\left(\frac{x_2}{x_1}\right) + \frac{1}{2(1-\nu)}\frac{x_1x_2}{R^2} + (1-2\nu)\times\left[\frac{x_2}{2(1-\nu)}+l\right]g_2\right\} \tag{12}$$

$$u_{22}(x_1,x_2) = -\frac{1}{2\pi\mu}\left\{(1-\nu)\ln R^2 - \frac{1}{2(1-\nu)}\frac{x_2^2}{R^2} + \frac{1-2\nu}{2(1-\nu)}[x_2 + (1-2\nu)l]g_1\right\} \tag{13}$$

in which $u_{\alpha\beta}$ is none other than the surface Green's function that represents the αth component of displacement induced by the unit concentrated surface load along x_β, and

$$R^2 = x_1^2 + x_2^2$$

$$g_1 = -\frac{\exp[(ix_1-x_2)/l]\mathrm{Ei}[(-ix_1+x_2)/l] + \exp[-(ix_1+x_2)/l]\mathrm{Ei}[(ix_1+x_2)/l]}{2l}$$

$$g_2 = \frac{\exp[(ix_1-x_2)/l]\mathrm{Ei}[(-ix_1+x_2)/l] - \exp[-(ix_1+x_2)/l]\mathrm{Ei}[(ix_1+x_2)/l]}{2il}$$

with Ei( ) being the exponential integral function. With expressions (10)–(13) one may easily obtain the fields induced by an arbitrarily distributed surface traction.

3 Elastic Interaction Between Two Surface Steps

According to Marchenko and Parshin [1], surface steps can be modeled by force dipoles acting on a planar surface and the force components assume the form of

$$F_1^{(1)} = p_1^{(1)}\frac{\partial\delta(x_1)}{\partial x_1}, \quad F_2^{(1)} = p_2^{(1)}\frac{\partial\delta(x_1)}{\partial x_1} \tag{14a}$$

$$F_1^{(2)} = p_1^{(2)}\frac{\partial\delta(x_1-d)}{\partial x_1}, \quad F_2^{(2)} = p_2^{(2)}\frac{\partial\delta(x_1-d)}{\partial x_1} \tag{14b}$$

in which $p_j^{(k)}$ is the force dipole strength of step k along x_j. The schematic illustration is plotted in Fig. 1. It can be easily proven that the interaction energy between the two steps can be obtained from

$$\Pi_{\mathrm{int}} = -\frac{1}{2}\int_S [(F_1^{(1)}u_1^{(2)} + F_1^{(2)}u_1^{(1)}) + (F_2^{(1)}u_2^{(2)} + F_2^{(2)}u_2^{(1)})]dS \tag{15}$$

where $u_\beta^{(\alpha)}$ represents the displacement along x_β induced by step α. By combining Eqs. (10)–(15), the interaction energy between two steps can be written explicitly as

$$\Pi_{\mathrm{int}} = p_\alpha^{(1)}p_\beta^{(2)}\frac{\partial^2 u_{\alpha\beta}(d,0)}{\partial x_1^2} \tag{16}$$

4 Numerical Results and Implications

As an example, we consider the two steps are identical and their dipole components are such that $p_1^{(1)}=p_1^{(2)}=p$ and $p_2^{(1)}=p_2^{(2)}=0$ (denoted as Case I) or $p_1^{(1)}=p_1^{(2)}=0$ and $p_2^{(1)}=p_2^{(2)}=p$ (denoted as Case II). The interaction energy between two such steps has been plotted as a function of d/l for $\nu=0.3$ in Fig. 2 where Π_{int} has been normalized by $\Pi_c=p_\alpha^{(1)}p_\beta^{(2)}\partial^2 G_{\alpha\beta}^c(d,0)/\partial x_1^2$ (i.e., the interac-

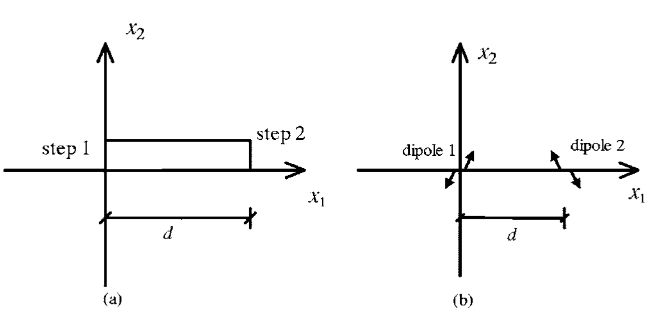


Fig. 1 Schematic illustration of two steps (*a*) and their force dipoles (*b*)

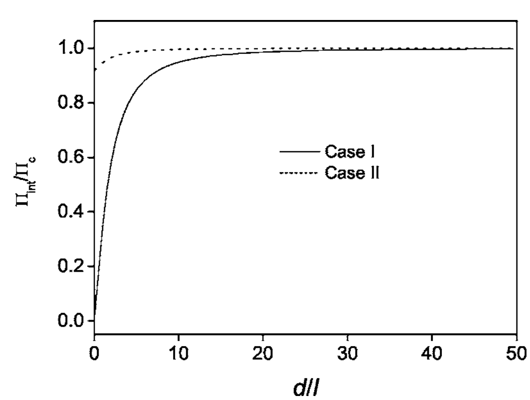


Fig. 2 The normalized interaction energy between two steps as a function of *d*/*l*

tion energy without the effect of surface elasticity). Here $G^c_{\alpha\beta}$ is the conventional surface Green's functions and can be found in many textbooks or papers, e.g., Müller and Saúl [12]. From Fig. 2, it can be seen that Π_{int}/Π_c both in Case I and in Case II increases with d/l and eventually approach 1, which means that in a certain range of the steps' separation, especially when it becomes small, the interaction energy no longer exhibits the form of d^{-2} as the conventional elasticity predicts. Such behavior has been found in some molecular simulations of interaction between surface steps [3]. Meanwhile, deviation of the elastic interaction between other surface defects such as adatoms and vacancies from the result of conventional elasticity theory for small spacing that has been found in Ref. [13] may also be attributed partly to the effect of surface elasticity. It is also noteworthy that for small value of d/l the effect of surface elasticity on the interaction energy in Case II is rather insignificant as compared to that in Case I, which may be considered qualitatively consistent with the results obtained by Hecquet [14]. Therein, it is found that for dipoles with force components only normal to the surface, the surface stress has no effect on their interaction while it does for dipoles with force components parallel to the surface. Consequently, taking into account the effect of surface elasticity is reasonable in characterizing some phenomena in surface science, especially when the geometric dimension like the distance d in the present work becomes close to the intrinsic length like l defined by the ratio between the surface elastic modulus and the bulk elastic modulus. Since according to Shenoy's atomic calculation of elastic properties of some surfaces [15], l may be well in the range of several angstroms, from the above results it can be concluded that the effect of surface elasticity may become significant when the distance between the steps is on the nanometer scale. Such effect has been generally neglected in the investigation of surface evolution and may be worthwhile for future consideration.

Acknowledgment

The present work is partially supported by the National Natural Science Foundation of China under Grant Nos. 10572067, No. 10121202, and SFDP 20030003064.

References

[1] Marchenko, V. I. and Parshin, A., Y., 1980, "Elastic Properties of Crystal Surfaces," Sov. Phys. JETP, **52**, pp. 129–131.
[2] Stewart, J., Pohland, O., and Gibbson, J. M., 1994, "Elastic-Displacement Field of an Isolated Surface Step," Phys. Rev. B, **49**, pp. 13848–13858.
[3] Kouris, D., Peralta, A., and Sieradzki, K., 1999, "Elastic Interaction of Defects on Crystal Surfaces," J. Eng. Mater. Technol., **121**, pp. 129–135.
[4] Prevot, G., Cohen, C., Schmaus, D., Hecquet, P., and Salanon, B., 2002, "Temperature Dependence of the Atomic Relaxations and Vibrations on a Stepped Surface: A Molecular Dynamics Study of Cu(1,1,19)," Surf. Sci., **506**, pp. 272–286.
[5] Shilkrot, L. E. and Srolovitz, D. J., 1996, "Elastic Field of a Surface Step: Atomistic Simulations and Anisotropic Elastic Theory," Phys. Rev. B **53**, pp. 11120–11127.
[6] Kukta, R. V. and Bhattacharya, K., 2002, "A Micromechanical Model of Surface Steps," J. Mech. Phys. Solids, **50**, pp. 615–649.
[7] Prevot, G. and Croset, B., 2004, "Revisiting Elastic Interactions Between Steps on Vicinal Surfaces: The Buried Dipole Model," Phys. Rev. Lett., **92**, p. 256104.
[8] Shuttleworth, R., 1950, "The Surface Tension of Solids," Proc. Phys. Soc., London, Sect. A, **63**, pp. 444–457.
[9] Gurtin, M. E. and Murdoch, A. I., 1975, "A Continuum Theory of Elastic Material Surfaces," Arch. Ration. Mech. Anal., **57**, pp. 291–323.
[10] Miller, R. E., and Shenoy, V. B., 2000, "Size-Dependent Elastic Properties of Nanosized Structural Elements," Nanotechnology, **11**, pp. 139–47.
[11] Sharma, P., Ganti, S., and Bhate, N., 2003, "Effect of Surfaces on the Size-Dependent Elastic State of Nano-Inhomogeneities," Appl. Phys. Lett., **82**, pp. 535–537.
[12] Müller, P., and Saúl, A., 2004, "Elastic Effects on Surface Physics," Surf. Sci. Rep., **54**, pp. 157–258.
[13] Ciobanu, C. V., Bardu, A., and Briggs, R. M., 2005, "Interactions of Carbon Atoms and Dimmer Vacancies on the Si(001) Surface," J. Eng. Mater. Technol., **127**, pp. 462–467.
[14] Hecquet, P., 2004, "Surface Stress Stabilizes Vicinal Surfaces," Surf. Sci., **561**, pp. 127–146.
[15] Shenoy, V. B., 2005, "Atomistic Calculations of Elastic Properties of Metallic Fcc Crystal Surfaces," Phys. Rev. B, **71**, p. 094104.

First-Order Numerical Analysis of Linear Thin Layers

F. Lebon[1]
Laboratoire de Mécanique et d'Acoustique,
Université Aix-Marseille 1,
31 Chemin Joseph-Aiguier,
13402 Marseille Cedex 20, France
e-mail: lebon@lma.cnrs-mrs.fr

S. Ronel
Laboratoire de Biomécanique et de Modélisation Humaine,
Université Lyon 1,
IUT B,
17 rue de France,
69100 Villeurbanne, France
e-mail: sylvie.ronel@iutb.univ-lyon1.fr

This paper deals with the first-order numerical analysis of thin layers. Theoretical results are recalled and compared with numerical data obtained on two classical examples. The effects of concentrated forces are discussed. [DOI: 10.1115/1.2424716]

1 Introduction

The aim of this study was to perform a first-order asymptotic and numerical analysis on linear thin layers. The rigidity of the thin layers is assumed here *not to depend on its thickness*. During the last few decades several authors have developed asymptotic theories applied to thin layers (see Ref. [1]). These problems have mostly been studied at order zero, and only a few authors have performed **first-order** studies [2]. We assume that only the geometrical parameter of the layer (the thickness) tends towards zero, and, analyze the limit problem using matched asympotic expansions [3]. In this limit problem, the layer vanishes geometrically and is replaced by an interface. The aim of this study is to check *quantitatively* the validity of the theoretical approach. This point is a crucial one for mechanical engineers because in practice, the thickness of the thin layer is generally very small.

The paper is organized as follows: in Sec. 2, the mechanical problem and the theoretical results are presented. Section 3 deals with two numerical examples. In Sec. 4, we analyze the influence of the concentrated forces obtained in the theoretical limit problem. In Sec. 5, we draw some conclusions and discuss the perspectives.

2 The Mechanical Problem and Theoretical Results

Let us consider two elastic bodies which are perfectly bonded with a third one which is very thin. For sake of simplicity, we work only in two dimensions. The structure is denoted Ω with boundary $\partial\Omega$ and is referred to the local frame (O,x_1,x_2). A surface load is applied to the part of the structure Γ_1. The structure is embedded in part Γ_0. We take Ω_ε to denote the part of Ω such that $|x_2|>\epsilon/2$ (the adherent) and B_ε to denote the complementarity part $\Omega/\Omega_\varepsilon$ (the adhesive). Segment S is the intersection between Ω and the line $\{x_2=0\}$. We adopt the small perturbations hypothesis and the adhesion between Ω_ε and B_ε is assumed to be perfect. Note that S is the surface to which the adhesive tends geometrically. The material composing is assumed to be elastic. We take λ and μ to denote the Lamé coefficients of the adhesive. Contrary to more classical studies, the order of magnitude of the stiffness is assumed to be the same in the adherent and in the adhesive (see Fig. 1).

2.1 Presentation of the Problem. In what follows, a jump across S is denoted by $[.]$ and a jump across the interface S_ε between Ω_ε and B_ε is denoted by $[.]_\varepsilon$. The equations of the problem are written as follows ($\boldsymbol{\sigma}$ denotes the stress tensor and $\boldsymbol{u}$ the displacement vector):

$$\begin{gathered} \operatorname{div}\boldsymbol{\sigma}=0 \text{ in } \Omega \\ \boldsymbol{\sigma}\mathbf{n}=\mathbf{F} \text{ on } \Gamma_1 \\ \mathbf{u}=0 \text{ on } \Gamma_0 \\ [\mathbf{u}]_\varepsilon=0 \text{ on } S_\varepsilon \\ [\boldsymbol{\sigma}\mathbf{n}]_\varepsilon=0 \text{ on } S_\varepsilon \end{gathered} \tag{1}$$

The constitutive equations are written as follows ($\mathbf{e}=\mathbf{e}(\mathbf{u})$ denotes the strain tensor)

$$\begin{gathered} \boldsymbol{\sigma}=\boldsymbol{A}\boldsymbol{e}(\boldsymbol{u}) \text{ in } \Omega_\varepsilon \\ \boldsymbol{\sigma}=\lambda\,\mathrm{tr}(\boldsymbol{e}(\boldsymbol{u}))\boldsymbol{I}+2\mu\boldsymbol{e}(\boldsymbol{u}) \text{ in } B_\varepsilon \end{gathered} \tag{2}$$

In the previous equation, $\boldsymbol{A}$ denotes a given elasticity tensor.

2.2 Matched Asymptotic Expansions. We assume that the solution of the above problem can be expanded into power series of ε. Using the matched asymptotic expansions method [4], we introduce internal (Eq. (4)) and external (Eq. (3)) expansions of the displacement vector $\mathbf{u}^\varepsilon$ and the stress tensor $\boldsymbol{\sigma}^\varepsilon$ which are valid sufficiently far from the edges. The two expansions are assumed to be coincident in a set of intermediate points (Eq. (5)). We write

$$\boldsymbol{u}^\varepsilon(x_1,x_2)=\sum_{m=0}^{\infty}\varepsilon^m\boldsymbol{u}^m(x_1,x_2),\quad \boldsymbol{\sigma}^\varepsilon(x_1,x_2)=\sum_{m=0}^{\infty}\varepsilon^m\boldsymbol{\sigma}^m(x_1,x_2) \tag{3}$$

$$\boldsymbol{u}^\varepsilon(x_1,x_2)=\sum_{m=0}^{\infty}\varepsilon^m\boldsymbol{v}^m\left(x_1,\frac{x_2}{\varepsilon}\right),\quad \boldsymbol{\sigma}^\varepsilon(x_1,x_2)=\sum_{m=0}^{\infty}\varepsilon^m\boldsymbol{\tau}^m\left(x_1,\frac{x_2}{\varepsilon}\right) \tag{4}$$

$$\begin{gathered} \boldsymbol{v}^0(x_1,\pm\infty)=\boldsymbol{u}^0(x_1,0^\pm) \\ \boldsymbol{\tau}^0(x_1,\pm\infty)=\boldsymbol{\sigma}^0(x_1,0^\pm) \\ \boldsymbol{v}^1(x_1,\pm\infty)=\boldsymbol{u}^1(x_1,0^\pm)+\lim_{y\to\pm\infty} y\frac{\partial\boldsymbol{u}^0}{\partial x_2}(x_1,0^\pm) \\ \boldsymbol{\tau}^1(x_1,\pm\infty)=\boldsymbol{\sigma}^1(x_1,0^\pm)+\lim_{y\to\pm\infty} y\frac{\partial\boldsymbol{\sigma}^0}{\partial x_2}(x_1,0^\pm) \end{gathered} \tag{5}$$

Let $y_2=x_2/\varepsilon$.

Introducing these expansions into Eqs. (1) and (2), we obtain

$$\tau_{ij}^n=\lambda e_{kk}^n\delta_{ij}+2\mu e_{ij}^n,\ n=0,1$$

$$\frac{\partial v_j^0}{\partial y_2}=0,\ j=1,2$$

$$e_{11}^0=\frac{\partial v_1^0}{\partial x_1}$$

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received January 18, 2006; final manuscript received June 27, 2006. Review conducted by Yonggang Huang.

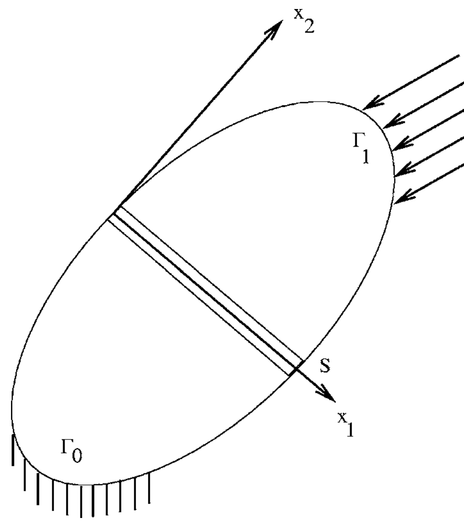


Fig. 1 The mechanical problem

$$e^0_{22} = \frac{\partial v^1_2}{\partial y_2}$$

$$e^0_{12} = \frac{1}{2}\left(\frac{\partial v^0_2}{\partial x_1} + \frac{\partial v^1_1}{\partial y_2}\right)$$

$$\frac{\partial \tau^0_{i2}}{\partial y_2} = 0$$

$$\frac{\partial \tau^0_{i1}}{\partial x_1} + \frac{\partial \tau^1_{i2}}{\partial y_2} = 0 \tag{6}$$

2.3 Asymptotic Results. By integration, Eqs. (5) and (6) (ii) mean that $[u^0_i]=0$, $i=1,2$. Likewise, Eqs. (5) and (6) (vi) mean that $[\sigma^0_{i2}]=0$, $i=1,2$. In conclusion, we have the following order zero system

$$[u^0_1](x_1) = 0 \tag{7a}$$

$$[u^0_2](x_1) = 0 \tag{7b}$$

$$[\sigma^0_{22}](x_1) = 0 \tag{7c}$$

$$[\sigma^0_{12}](x_1) = 0 \tag{7d}$$

In the same way, by integration, Eqs. (5) and (6) (i) with $n=0$ and (6) (iii-v) mean that the jump in the displacements $[u^1_i]$ $i=1,2$ is not equal to zero and depends on σ^0_{i2} and $\partial u^0_i/\partial x_1$. The corresponding results are given in Eq. (8).

The jump in the stresses is obtained using Eqs. (5) and (6) (vii), and (6) (i) (by derivation). In conclusion, we have the following order one system

$$[u^1_1](x_1) = \frac{\sigma^0_{12}}{\mu} - \frac{\partial u^0_2}{\partial x_1} \tag{8a}$$

$$[u^1_2](x_1) = \frac{\sigma^0_{22}}{\lambda + 2\mu} - \frac{\lambda}{\lambda + 2\mu}\frac{\partial u^0_1}{\partial x_1} \tag{8b}$$

$$[\sigma^1_{22}](x_1) = -\frac{\partial \sigma^0_{12}}{\partial x_1} \tag{8c}$$

$$[\sigma^1_{12}](x_1) = -\frac{4\mu(\lambda+\mu)}{\lambda+2\mu}\frac{\partial^2 u^0_1}{\partial x_1^2} - \frac{\lambda}{\lambda+2\mu}\frac{\partial \sigma^0_{22}}{\partial x_1} \tag{8d}$$

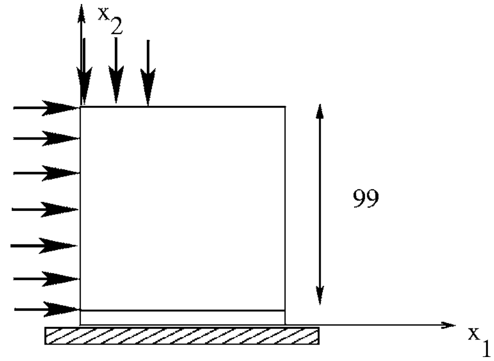


Fig. 2 First example: square bar bonded with a rigid body (dimensions in mm)

Note that the model obtained is **nonlocal**.

3 Numerical Tests

The aim of this section is to check quantitatively the validity of the theory. It is crucial to obtain values of the thickness for which it is possible to substitute the real problem by the limit one. The computations were performed using the ANSYS (Multiphysics solver, Plan82 element, plane stress) software program [5]. The discretization of the thin layer is done by two or four elements in the thickness and 200 elements in the width. This numerical section contains two parts, corresponding to two examples. In the first part, we observe the jumps in the displacements $[u_1]$ and $[u_2]$ and the jumps in the stress vector $[\sigma_{22}]$ and $[\sigma_{12}]$ along the interface zone (based on Eq. (7)). In the second one, we check the validity of Eq. (8).

3.1 First Numerical Example

3.1.1 Geometry of the Problem. In this section, we describe numerical tests performed on a long square bar bonded with a rigid obstacle (Fig. 2). The width of the bar was equal to 99 mm and the thickness of the thin layer was equal to 1 mm. A horizontal load was applied to the whole left part of the structure and a vertical one was applied to only the upper left nodes of the square bar. The mechanical data are given in Table 1.

3.1.2 Numerical Synthesis

3.1.2.1 Jump in the displacements at order 0 (Eqs. (7a) and (7b)). The first step in the numerical test consists of checking that the values of the jump in the displacements at order zero tend toward zero. In this example, the jump in the displacements is equal to the displacement of the upper nodes of the thin layer. It is confirmed that the displacement is small (see Figs. 4(*a*) and 4(*b*)). The values of the displacements tend toward zero: these values are in the $[2.10^{-4}, 5.10^{-4}]$ mm range with u_1 and in the

Table 1 Mechanical data

Example thickness (mm)	1 1	2 1
Substrate: Young's modulus (GPa)	200	200
Substrate: Poisson ratio	0.3	0.3
Thin layer: Young's modulus (GPa)	160	160
Thin layer: Poisson ratio	0.3	0.3
Total x_1 force (N)	1800 (18 nodes)	1800 (18 nodes)
Total x_2 force (N)	−1200 (60 nodes)	−1000 (50 nodes)
Finite element	8-node quadrangle	8-node quadrangle

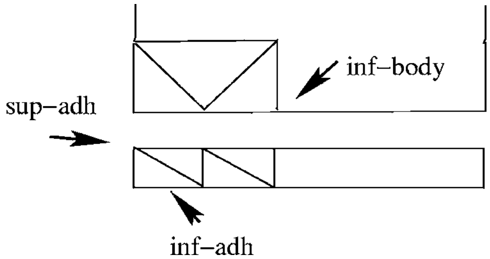


Fig. 3 First example: three lines of nodes

$[-3.10^{-4}, 3.10^{-4}]$ mm range with u_2. These values are approximately 100 times smaller than those in the adherent. It is noticed that in this modeling, contrary to other theories for which the rigidity of the adhesive is small, the jump in the displacements tends toward zero.

3.1.2.2 Jump in the stress vector at order 0 (Eqs. (7c) and (7d)). In the case of the stress vector, we computed three sets of values (Fig. 3): the lower and the upper nodes of the layer and the lower nodes of the body (Figs. 4(*c*) and 4(*d*)). The three curves (denoted inf-adh, sup-adh, and inf-body in Fig. 4) are very similar, the jump tends towards zero, and it is possible to choose any of these three values for the following computations. In particular, the computation of the gradient can be done indifferently on one of the three surfaces.

3.1.2.3 Jump in the displacements at order 1 (Eqs. (8a) and (8b)). In this case, the thin layer is bonded directly with a rigid body and some of the terms in Eqs. (8*a*) and (8*b*) can be simplified. In particular, the jump in the displacement at order zero is equal to zero as are the displacement and its x_1 derivatives. Since the lower nodes are clamped, the jump in the displacements at order one is computed as the displacements of the upper nodes in the thin layer divided by ε. In Eq. (8*a*) the jump in the displacement of u_1 at order one is approximately σ_{12} at order zero divided by μ (Fig. 5(*a*)). In Eq. (8*b*) the jump in the displacements of u_2 at order one is approximately σ_{22} at order zero divided by $\lambda + 2\mu$ (Fig. 5(*b*)).

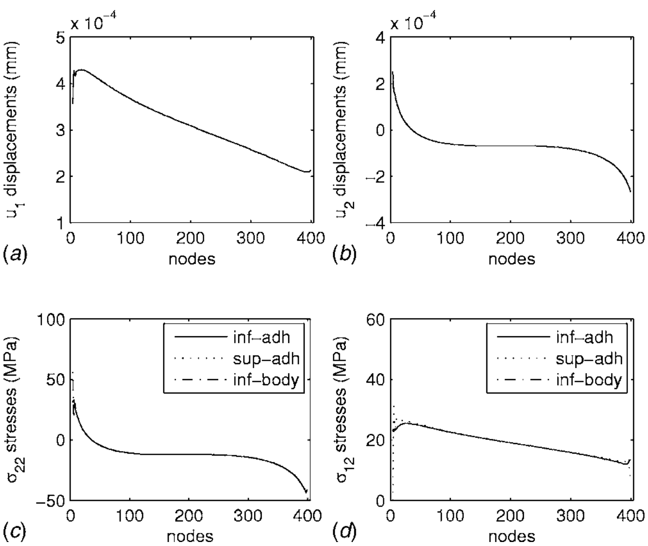


Fig. 4 Square bar: numerical results on the contact zone: (*a*) u_1 displacements; (*b*) u_2 displacements; (*c*) σ_{22} stresses; (*d*) σ_{12} stresses (-) lower zone on the adhesive, (...) upper zone on the adhesive, (-.) lower zone on the elastic body)

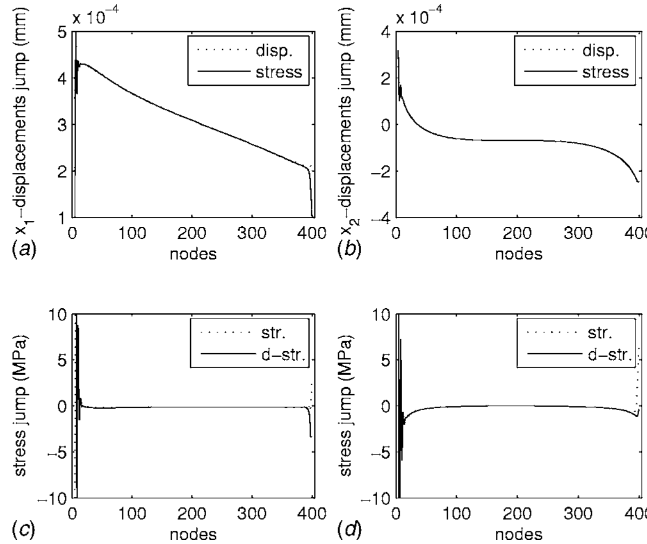


Fig. 5 Square bar: numerical results on the contact zone: (*a*) $[u_1]$ displacements; (*b*) $[u_2]$ displacements; (*c*) $[\sigma_{12}]$ stresses; (*d*) $[\sigma_{22}]$ stresses ((...) jump in the stress, (-) derivative)

3.1.2.4 Jump in the stress vector at order 1 (Eqs. (8c) and (8d)). At order one, it can be seen from Fig. 4 that the stresses are very small. This is computed by dividing the numerical displacement obtained in the upper nodes of the layer by ε. Stresses at order zero are computed in the upper nodes of the layers. For Eqs. (8*c*) and (8*d*), the jump in the stress vector at order one is taken to be equal to the difference between the values obtained in the upper and lower nodes of the layer. Figure 5(*c*) shows that stress σ_{12} (denoted str.) is similar to the derivative of σ_{22} (denoted d-str.) at order zero (upper nodes of the layer) divided by $\lambda/(\lambda+2\mu)$. The derivatives of the stresses at order zero are computed in the upper nodes of the layer. In Fig. 5(*d*), we can see that the jump in the stress σ_{22} at order one is approximately zero as predicted in Eq. (8*c*).

3.1.2.5 Conclusion. In Fig. 5, we have checked the results by comparing each term of Eqs. (8). The agreement found to exist between the curves confirms the validity of the theory presented in Ref. [2] and developed in this paper. One notes a divergence of the theory on the edges. From a mathematical point of view, this theory is valid only in the open set and not in the closed set. This problem will be taken into account thereafter.

3.2 Second Numerical Example

3.2.1 Geometry of the Problem. In this section, we describe numerical tests performed on two bars connected with a thin layer (Fig. 6). The width of the bars is equal to 99.5 mm and the thickness of the thin layer is equal to 1 mm. The lower bar is clamped underneath. A horizontal load is applied to the whole left side of the top bar and a vertical one is applied to the upper left part of the top bar.

3.2.2 Numerical Synthesis

3.2.2.1 Jump in the displacements at order 0 (Eqs. (7a) and (7b)). The first step in the numerical validation procedure consists of checking that the values of the jump in the displacements at order zero tend toward zero. In this example, the jump in the displacements is equal to the difference between the displacements of the upper nodes and the lower nodes of the thin layer. The jump was found to be small (Figs. 7(*a*) and 7(*b*)). The values of the jumps tend to zero. These values are in the $[2.10^{-4}, 5.10^{-4}]$ mm range with u_1 and in the $[-2.10^{-4}, 1.10^{-4}]$ mm range with u_2.

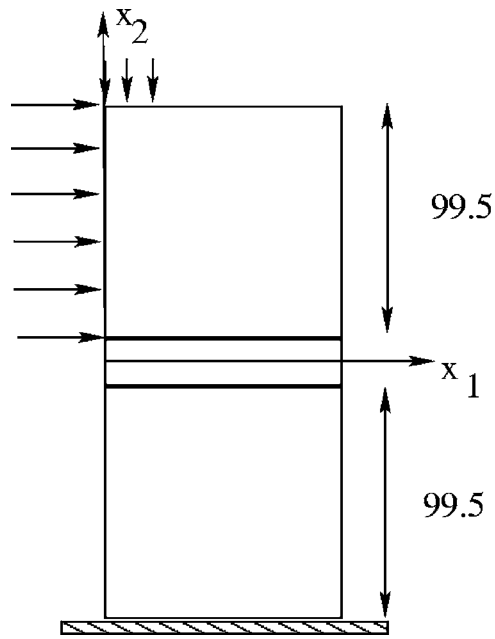


Fig. 6 Second example: two bonded bars (dimensions in mm)

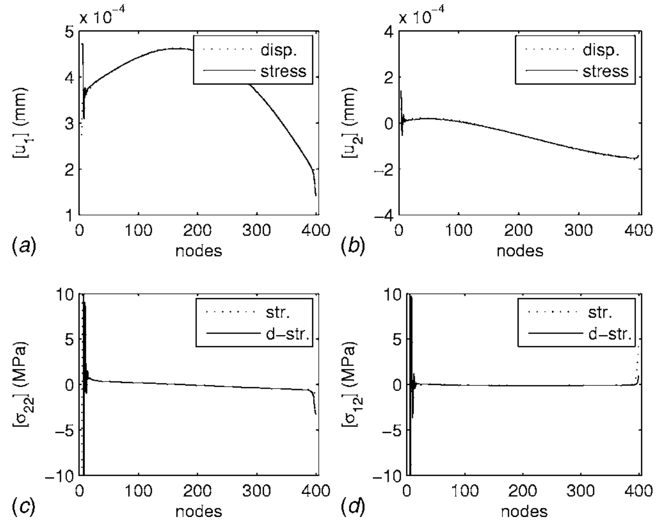


Fig. 8 Two bars: numerical data on the contact zone (*a*) $[u_1]$ displacements; (*b*) $[u_2]$ displacements (… displacements, - stress); (*c*) $[\sigma_{22}]$ stresses; (*d*) $[\sigma_{12}]$ stresses ((…) jump in the stress, (-) derivative)

3.2.2.2 Jump in the stress vector at order 0 (Eqs. (7c) and (7d)). In the case of the stress vector, we computed the stress difference between the lower and upper nodes of the layer (Figs. 7(*c*) and 7(*d*)). The two values were found to be similar and the jump tends toward zero.

3.2.2.3 Jump in the displacements at order 1 (Eqs. (8a) and (8b)). In this case, the thin layer is not bonded directly with a rigid body and the terms in Eqs. (8*a*) and (8*b*) are not simplified. The jump in the displacements at order one is taken to be the computed jump in the displacements divided by ε. The derivative of the displacement (Eqs. (8*a*) and (8*b*)) is approximated for the upper nodes of the thin layer. We compared Eqs. (8*a*) and (8*b*) with the numerical jump in the displacement at order one (denoted disp. in Fig. 8). It can be seen from Figs. 8(*a*) and 8(*b*) that these values are very similar.

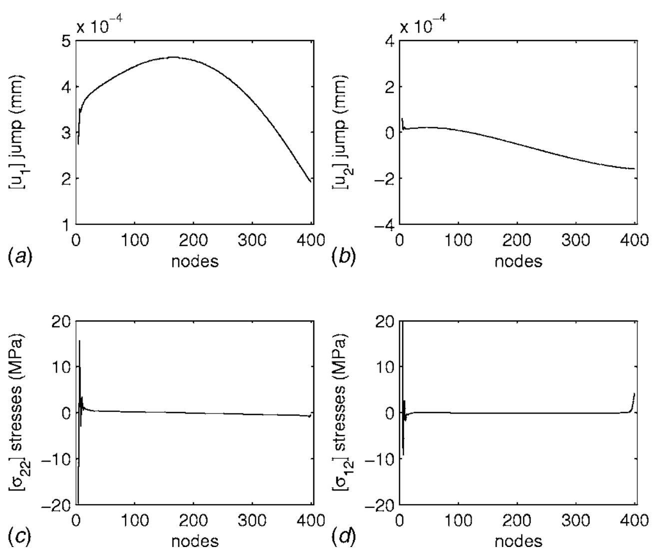


Fig. 7 Two bars: numerical results on the contact zone: (*a*) $[u_1]$ displacements; (*b*) $[u_2]$ displacements; (*c*) $[\sigma_{22}]$ stresses; and (*d*) $[\sigma_{12}]$ stresses

3.2.2.4 Jump in the stress vector at order 1 (Eqs. (8c) and (8d)). In Eqs. (8*c*) and (8*d*), the jump in the stress vector at order one is taken to be equal to the difference between the values in the upper and lower nodes of the layer. The derivatives of the stresses at order zero are computed in the upper nodes of the layer. As Fig. 8(*c*) shows, stress σ_{12} (denoted str.) is similar to the derivative of σ_{22} (denoted d-str.) at order zero (upper nodes of the layer) divided by $\lambda/(\lambda+2\mu)$ (Eq. (8*d*)). In Fig. 8(*d*), we can see that the jump in stress σ_{22} at order one is similar to the values obtained upon computing the right hand side of Eq. (8*c*).

3.2.2.5 Conclusion. The numerical data given in Fig. 8 are compared with the theoretical results obtained in each term of Eq. (8). The good agreement obtained confirms the validity of the theory proposed in Ref. [2].

4 Comments on Concentrated Forces at the Edges

As seen in the previous sections (Figs. 4, 5, 7, and 8), the results presented in Eq. (8) are no longer valid near the edges. In this case, it is necessary to include concentrated forces in the model. In particular, the limit model does not take into account the singularities of stresses on the edges. In this paragraph, one presents a way of taking into account partial effects of the singularities near the edges. Let us consider a small circle centered at the edge of the thin layer (see Fig. 9). Using the divergence formula and Eqs. (1), (4), and (6) (i)(vii), we obtain

$$\int_{C_\varepsilon} \sigma n + \int_{D_\varepsilon} \sigma n = 0$$

$$\int_{C_\varepsilon} \sigma^1 n + \int_{-1/2}^{-1/2} \tau^0 n = 0$$

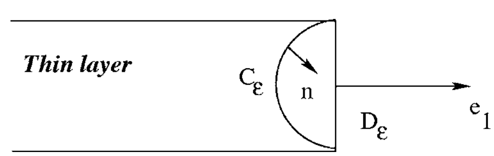


Fig. 9 Concentrated forces

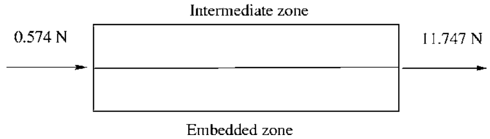


Fig. 10 Application of the concentrated forces

$$\int_{C_\varepsilon} \sigma^1 n = -\left(\frac{4\mu(\lambda+\mu)}{\lambda+2\mu} \frac{\partial u_1^0}{\partial x_1} - \frac{\lambda}{\lambda+2\mu} \sigma_{22}^0 \right) e_1 \tag{9}$$

Therefore the concentrated forces at the edges $P^{\pm}$, denoted F, are

$$F = -\left(\frac{4\mu(\lambda+\mu)}{\lambda+2\mu} \frac{\partial u_1^0}{\partial x_1} - \frac{\lambda}{\lambda+2\mu} \sigma_{22}^0 \right) (P^{\pm}) e_1 \tag{10}$$

The validity of these theoretical results can be seen by comparing Eq. (10) with the computed numerical data. We applied to the limit problem (first example) a concentrated force exerted on the first line of elements as shown in Eq. (10). The value of this force is obtained by performing computations on the real data (see Fig. 10). In Fig. 11, the limit problem with and without concentrated forces and the initial problem with the thin layer are compared. The results show the considerable improvement obtained in the case involving concentrated forces (Fig. 11).

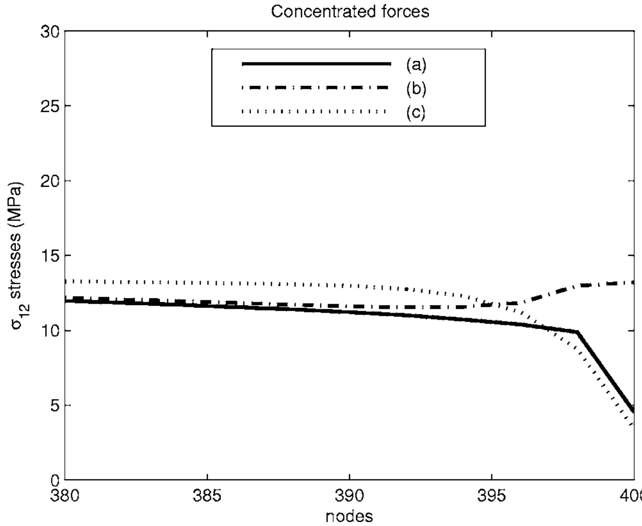


Fig. 11 Numerical solution on the edge (*a*) limit problem with concentrated forces; (*b*) limit problem without concentrated forces; and (*c*) initial problem with two layers

5 Conclusion

In this study, we have both developed and numerically validated an asymptotic model for the interface described in Ref. [2]. This interface model is nonlocal. Good agreement was obtained between theoretical and numerical data. One obtains a law that could be modeled numerically at order 1 by particular cohesive elements. We now intend to develop a theory on similar lines dealing with the nonlinear constitutive laws pertaining to the thin layer, in particular, taking into account damage or heterogeneities (as in adhesives reinforced by particles).

References

[1] Lebon, F., and Ronel-Idrissi, S., 2004, "Asymptotic Analysis of Mohr–Coulomb and Drucker–Prager Soft Thin Layers," Steel Compos. Struct., **4**, pp. 133–147.

[2] Abdelmoula, R., Coutris, M., and Marigo, J. J., 1998, "Comportement Asymptotique d'une Interface Élastique Mince," C. R. Acad. Sci. III, **326**, pp. 237–242.

[3] Sanchez-Palencia, E., 1980, "Non Homogeneous Materials and Vibration Theory," *Lecture Notes in Physics*, Springer, Berlin, Vol. No. 127.

[4] Eckhaus, W., 1979, *Asymptotic Analysis of Singular Perturbations*, North-Holland, Dordredht, The Netherlands.

[5] ANSYS, 2002, ANSYS 6.1, SASIP.

The Contact Problem in a Compressible Hyperelastic Material

G. F. Wang
e-mail: wang_gf@sohu.com

T. J. Wang

School of Aerospace,
Xian Jiaotong University,
Xian 710049, China

P. Schiavone[1]
Department of Mechanical Engineering,
University of Alberta,
Edmonton, AB T6G 2G8, Canada
e-mail: P.Schiavone@ualberta.ca

We consider the contact problem for a particular class of compressible hyperelastic materials of harmonic type undergoing finite plane deformations. Using complex variable techniques, we derive subsidiary results concerning a half-plane problem corresponding to this class of materials. Using these results, we solve the contact problem for a harmonic material in the case of a uniform load acting on a finite area. Finally, we show how we can then deduce the corresponding results for the case of a point load. [DOI: 10.1115/1.2711229]

Keywords: contact problem, finite elastic deformations, harmonic materials

1 Introduction

The model of a compressible hyperelastic harmonic material was first proposed by John [1] and subsequently applied by a number of authors to various problems in finite elasticity. Recently, the analysis of this class of materials has attracted considerable attention in the literature. For example, Knowles and Sternberg [2] have considered the singularity induced by certain mixed boundary conditions in nonlinear elastostatics; Varley and Cumberbatch [3] solved the finite deformation problem of an elliptic hole in a harmonic material (their results showed good agreement with experimental data for some rubber-like materials); Abeyaratne and Horgan [4] investigated the pressurized hollow sphere problem in finite elastostatics; Li and Steigmann [5] studied the finite deformation of an annular membrane induced by the rotation of a rigid hub; Horgan [6] obtained the axisymmetric solutions for compressible nonlinearly elastic solids and Aguiar and Fosdick [7] analyzed the finite deformation field near a corner using asymptotic analysis. In these papers, it was found that theoretical results based on this model were in good agreement with experimental data.

Complex variable techniques [8] have been shown to be extremely powerful and effective tools for the study of problems in linear plane elasticity. Analogously, complex variable techniques have been used successfully in the case of finite deformations. In particular, in Ref. [3], Varley and Cumberbatch developed a complex variable formulation of a wide class of compressible hyperelastic materials of harmonic type. More recently, Ru [9] developed a simple, yet powerful complex-variable formulation of a class of problems involving the plane-strain deformations of a set of compressible hyperelastic materials of harmonic type including, in particular, problems involving an interface crack. Based on Ru's formulation, Ru et al. [10] investigated the uniformity of stresses inside an elliptic inclusion in finite plane elastostatics; Wang et al. [11,12] considered the design of harmonic shapes in finite elasticity and considered the surface instability of compliant materials under van der Waals forces, respectively.

The contact problem is one of the most important in solid mechanics. It has wide engineering applications, for example, in friction, lubrication, adhesion etc. In Johnson's distinguished book, ***Contact Mechanics*** [13], many typical contact problems have been included. However, the analogous results for the case of *finite deformation* are very rare, particularly in analytical form.

In the present paper, we investigate the contact problem for a hyperelastic harmonic material. Our paper is organized as follows. The basic equations describing plane strain deformations of a harmonic material are summarized in Sec. 2. In Sec. 3, using the continuation properties of analytical functions, the complex potential for a half-plane problem is formulated. Finally, in Sec. 4, we solve contact problems for harmonic materials subjected to a uniform traction on a finite region and a point force, respectively.

2 Basic Equations for a Harmonic Material

In this section, we present only a brief review of the equations governing finite plane (strain) deformations of a harmonic material. Further details can be found in Ru [9] and Knowles and Sternberg [2].

Let the complex variable $z=x_1+ix_2$ represent the initial coordinates of a material particle in the undeformed configuration, and $w(z)=y_1(z)+iy_2(z)$ the corresponding spatial coordinates in the deformed configuration. Define the deformation gradient tensor as

$$F_{ij}=\frac{\partial y_i}{\partial x_j} \tag{1}$$

For a particular class of harmonic materials, the strain energy density defined with respect to the undeformed unit area can be expressed by

$$W=2\mu[F(I)-J], \quad F'(I)=\frac{1}{4\alpha}[I+\sqrt{I^2-16\alpha\beta}] \tag{2}$$

where I and J are the scalar invariants of FF^T given by

$$I=\sqrt{F_{ij}F_{ij}+2J}, \quad J=(\det[F])^2 \tag{3}$$

μ is the shear modulus and $1/2\leqslant\alpha<1$, $\beta>0$ are two material constants.

According to the formulation given by Ru [9], the deformation $w(z)$ can be written in terms of two analytic functions φ and ψ as

$$iw(z)=\alpha\varphi(z)+i\overline{\psi(z)}+\frac{\beta z}{\overline{\varphi'(z)}} \tag{4}$$

and the complex Piola stress function $\chi(z)$ is given by

$$\chi(z)=2\mu i\left[(\alpha-1)\varphi(z)+i\overline{\psi(z)}+\frac{\beta z}{\overline{\varphi'(z)}}\right] \tag{5}$$

The stress components can be obtained by the Piola stress function as

$$-i\sigma_{12}+\sigma_{22}=\chi_{,1}, \quad -\sigma_{21}+i\sigma_{11}=\chi_{,2} \tag{6}$$

Using Eqs. (5) and (6), we can obtain the stress components on x_2 plane as

$$-i\sigma_{12}+\sigma_{22}=2\mu i\left[(\alpha-1)\varphi'(z)+i\overline{\psi'(z)}+\frac{\beta}{\overline{\varphi'(z)}}-\frac{\beta z}{\overline{\varphi'(z)}^2}\overline{\varphi''(z)}\right] \tag{7}$$

---

[1]Corresponding author.

Contributed by the Applied Mechanics Division of ASME for publication in the JOURNAL OF APPLIED MECHANICS. Manuscript received April 11, 2006; final manuscript received August 14, 2006. Review conducted by Matthew R. Begley.

3 Complex Potential for a Half-Plane Problem

To solve the contact problem, we use the continuation properties of analytic functions [8] to formulate the complex potential for a half-plane harmonic solid. This is the basis on which we analyze the surface stability, contact mechanics, etc, of the harmonic solid.

Suppose that a region L of the surface $x_2=0$ of the half plane S^+ ($x_2>0$) is unstressed so that

$$\lim_{x_2\to 0+}\left[(\alpha-1)\varphi'(z)+i\overline{\psi'(z)}+\frac{\beta}{\overline{\varphi'(z)}}-\frac{\beta z}{\overline{\varphi'(z)}^2}\overline{\varphi''(z)}\right]=0 \quad (8)$$

that is

$$(\alpha-1)\varphi'^+(x_1)+i\overline{\psi'^+(x_1)}+\frac{\beta}{\overline{\varphi'^+(x_1)}}-\frac{\beta x_1}{\overline{\varphi'^+(x_1)}^2}\overline{\varphi''^+(x_1)}=0 \quad (9)$$

Here we use the notation $\lim_{x_2\to 0+}\varphi'(z)=\varphi'^+(x_1)$.

Noting that $\lim_{x_2\to 0-}\overline{\varphi(\bar{z})}=\lim_{x_2\to 0+}\overline{\varphi(z)}=\overline{\varphi^+(x_1)}$. In terms of the associated functions $\overline{\varphi(\bar{z})}$ and $\overline{\psi(\bar{z})}$, which are analytic in S^- ($x_2<0$), the boundary condition can be written in alternative form as

$$\varphi'^+(x_1)=-\frac{1}{(\alpha-1)}\lim_{x_2\to 0-}\left[i\overline{\psi'(\bar{z})}+\frac{\beta}{\overline{\varphi'(\bar{z})}}-\frac{\beta z}{\overline{\varphi'(\bar{z})}^2}\overline{\varphi''(\bar{z})}\right] \quad (10)$$

It is natural to extend the definition of $\varphi'(z)$ from S^+ to S^- by setting

$$\varphi'(z)=\varphi'(z), \quad (z\in S^+)$$

$$\varphi'(z)=-\frac{1}{(\alpha-1)}\left[i\overline{\psi'(\bar{z})}+\frac{\beta}{\overline{\varphi'(\bar{z})}}-\frac{\beta z}{\overline{\varphi'(\bar{z})}^2}\overline{\varphi''(\bar{z})}\right], \quad (z\in S^-) \quad (11)$$

Since $\varphi'(z)$ is continued analytically from S^+ into S^- across the boundary $x_2=0$, these equations can be integrated to give

$$\varphi(z)=\varphi(z), \quad (z\in S^+)$$

$$\varphi(z)=-\frac{1}{(\alpha-1)}\left[i\overline{\psi(\bar{z})}+\frac{\beta z}{\overline{\varphi'(\bar{z})}}\right], \quad (z\in S^-) \quad (12)$$

Next, we express $\psi(z)$ ($z\in S^+$) in terms of $\varphi(z)$ defined in both S^+ into S^-, and obtain

$$i\overline{\psi(z)}=-(\alpha-1)\varphi(\bar{z})-\frac{\beta\bar{z}}{\overline{\varphi'(z)}}, \quad (z\in S^+) \quad (13)$$

The displacement can then be expressed as

$$iw(z)=\alpha\varphi(z)-(\alpha-1)\varphi(\bar{z})+\frac{\beta(z-\bar{z})}{\overline{\varphi'(z)}} \quad (14)$$

and the complex Piola stress function $\chi(z)$ is given by

$$\chi(z)=2\mu i\left\{(\alpha-1)[\varphi(z)-\varphi(\bar{z})]+\frac{\beta(z-\bar{z})}{\overline{\varphi'(z)}}\right\} \quad (15)$$

The stress components are finally evaluated by

$$-i\sigma_{12}+\sigma_{22}=2\mu i\left\{(\alpha-1)[\varphi'(z)-\varphi'(\bar{z})]-\frac{\beta(z-\bar{z})}{\overline{\varphi'(z)}^2}\overline{\varphi''(z)}\right\}$$

$$-\sigma_{21}+i\sigma_{11}=-2\mu\left\{(\alpha-1)[\varphi'(z)+\varphi'(\bar{z})]+\frac{2\beta}{\overline{\varphi'(z)}}+\frac{\beta(z-\bar{z})}{\overline{\varphi'(z)}^2}\overline{\varphi''(z)}\right\} \quad (16)$$

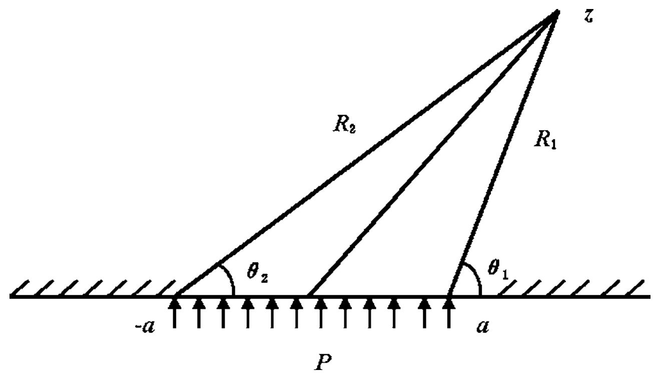


Fig. 1 The contact problem in a hyperelastic material subjected to uniform loading

These formulae hold for ($z\in S^+$), where $\varphi(z)$ is analytic in both S^+ and S^-. Though we formulate the previous results using the homogeneous boundary condition, it loses no generality when we treat the above continuation as the stress continuation.

4 Contact Problem

Based on the formulation in Sec. 3, we consider the contact problem for a harmonic material. For a material occupying the half plane S^+, the stress boundary condition is given by

$$-i\sigma_{12}+\sigma_{22}=-[p(x_1)+is(x_1)], \quad (x_2=0) \quad (17)$$

where $p(x_1)$ is the normal pressure and $s(x_1)$ is the shear stress applied to the boundary. By using the stress expression in Eq. (16), we have

$$\varphi'^+(x_1)-\varphi'^-(x_1)=\frac{i}{2\mu(\alpha-1)}[p(x_1)+is(x_1)], \quad (x_2=0) \quad (18)$$

Suppose that the stresses are zero at infinity. This requires,

$$\varphi'(z)=\eta i, \quad (|z|\to\infty) \quad (19)$$

where $\eta=\sqrt{\beta/(1-\alpha)}$.

If we suppose further that a finite resultant force acts on the surface, we can obtain the unique solution for Eq. (18) as

$$\varphi'(z)=\frac{1}{2\mu(\alpha-1)}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{p(t)+is(t)}{t-z}dt+\eta i \quad (20)$$

As an example, we consider a uniform pressure p and shear s acting over the region $|x_1|\leqslant a$ with the remainder of the boundary free (as shown in Fig. 1). We obtain

$$\varphi'(z)=\frac{p+is}{2\mu(\alpha-1)}\frac{1}{2\pi}\ln\left(\frac{z-a}{z+a}\right)+\eta i \quad (21)$$

For this case, the stress field is given by

$$\frac{i\sigma_{22}+\sigma_{12}}{2\mu}=-2i(\alpha-1)(p'+is')(\theta_1-\theta_2)+\frac{\beta(p'-is')(e^{i2\theta_1}-e^{i2\theta_2})}{\left[(p'-is')\ln\left(\frac{R_1}{R_2}e^{i(\theta_2-\theta_1)}\right)-\eta i\right]^2}$$

$$\frac{\sigma_{21}-i\sigma_{11}}{2\mu}=2(\alpha-1)\left[(p'+is')\ln\left(\frac{R_1}{R_2}\right)+\eta i\right]+\frac{2\beta}{(p'-is')\ln\left(\frac{R_1}{R_2}e^{i(\theta_2-\theta_1)}\right)-\eta i}+\frac{\beta(p'-is')(e^{i2\theta_1}-e^{i2\theta_2})}{\left[(p'-is')\ln\left(\frac{R_1}{R_2}e^{i(\theta_2-\theta_1)}\right)-\eta i\right]^2} \tag{22}$$

with

$$p'+is'=\frac{p+is}{2\mu(\alpha-1)}\frac{1}{2\pi},\quad z-a=R_1e^{i\theta_1},\quad z+a=R_2e^{i\theta_2} \tag{23}$$

These results can be used to obtain the corresponding results for the case of a point force. In fact, as $a\to 0$, let p and s tend to infinity in such a way that

$$2ap\to Y,\quad 2as\to X \tag{24}$$

where X and Y are constants. Taking the limit in Eq. (21) as $a\to 0$, we have

$$\varphi'(z)=-\frac{Y'+iX'}{z}+\eta i \tag{25}$$

where

$$Y'+iX'=\frac{Y+iX}{2\mu(\alpha-1)}\frac{1}{2\pi}$$

The corresponding stresses are evaluated by

$$\frac{i\sigma_{22}+\sigma_{12}}{2\mu}=(\alpha-1)(Y'+iX')\left(\frac{1}{z}-\frac{1}{\bar{z}}\right)+\frac{\beta(Y'-iX')(z-\bar{z})}{(Y'-iX'+\eta i\bar{z})^2}$$

$$\frac{\sigma_{21}-i\sigma_{11}}{2\mu}=(\alpha-1)\left[2\eta i-(Y'+iX')\left(\frac{1}{z}+\frac{1}{\bar{z}}\right)\right]-\frac{2\beta\bar{z}}{Y'-iX'+\bar{z}\eta i}+\frac{\beta(Y'-iX')(z-\bar{z})}{(Y'-iX'+\bar{z}\eta i)^2} \tag{26}$$

Though the potentials $\varphi'(z)$ in Eqs. (21) and (25) are quite similar to those in a linearly elastic material [8], the stress distributions are quite different, as seen in Eqs. (22) and (26).

5 Conclusions

Using the complex variable method, we formulate the general equations for the half-plane problem of a particular class of compressible hyperelastic materials of harmonic type. Subsequently, we investigate the corresponding contact problems in the case of finite plane deformations. Stress distributions are obtained in analytical form for the two cases of a uniform loading on a finite area and then for a point force. It is clear from Eqs. (22) and (26) that the stress distributions are significantly different from the corresponding distributions in a linearly elastic solid, illustrating the influence of the nonlinear behavior of this particular class of materials.

Acknowledgment

This work was supported by the Natural Sciences and Engineering Research Council of Canada through grant No. NSERC OGP 115112.

References

[1] John, F., 1960, "Plane Strain Problems for a Perfectly Elastic Material of Harmonic Shape," Commun. Pure Appl. Math. **XIII**, pp. 239–290.
[2] Knowles, J. K., and Sternberg, E., 1975, "On the Singularity Induced by Certain Mixed Boundary Conditions in Linearized and Nonlinear Elastostatics," Int. J. Solids Struct., **11**, pp. 1173–1201.
[3] Varley, E., and Cumberbatch, E., 1980, "Finite Deformation of Elastic Material Surrounding Cylindrical Holes," J. Elast., **10**, pp. 341–405.
[4] Abeyaratne, R., and Horgan, C. O., 1984, "The Pressurized Hollow Sphere Problem in Finite Elastostatics for a Class of Compressible Materials," Int. J. Solids Struct., **20**, pp. 715–723.
[5] Li, X., and Steigmann, D. J., 1993, "Finite Plane Twist of an Annular Membrane," Q. J. Mech. Appl. Math., **46**, pp. 601–625.
[6] Horgan, C. O., 1995, "On Axisymmetric Solutions for Compressible Nonlinearly Elastic Solids," ZAMP, **46**, pp. 107–125.
[7] Aguiar, A., and Fosdick, R., 2001, "Self-Interaction in Elasticity," Int. J. Solids Struct., **38**, pp. 4797–4823.
[8] England, A. E., 1971, *Complex Variable Methods in Elasticity*, Wiley Interscience, London, UK.
[9] Ru, C. Q., 2002, "On Complex-Variable Formulation for Finite Plane Elastostatics of Harmonic Materials," Acta Mech., **156**, pp. 219–234.
[10] Ru, C. Q., Schiavone, P., Sudak, L. J., and Mioduchowski, A., 2005, "Uniformity of Stresses Inside an Elliptic Inclusion in Finite Plane Elastostatics," Int. J. Non-Linear Mech., **40**(2–3), pp. 281–287.
[11] Wang, G. F., Schiavone, P., and Ru, C.-Q., 2005, "Harmonic Shapes in Finite Elasticity Under Non-Uniform Loading," ASME J. Appl. Mech., **72**(5), pp. 691–694.
[12] Wang, G. F., Schiavone, P., and Ru, C.-Q., 2005, "Surface Instability of a Semi-Infinite Harmonic Solid Under van der Waals Attraction," Acta Mech., **180**(1–4), pp. 1–10.
[13] Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, London, UK.

Journal of Applied Mechanics

Erratum

Erratum: "Plane-Strain Propagation of a Fluid-Driven Fracture: Small Toughness Solution" [Journal of Applied Mechanics, 2005, 72(6), pp. 916–928]

D. I. Garagash and E. Detournay

Equations (28), (35*b*), (47), and (C16), and expression for Π_1 in Appendix D in the published paper contain typographical errors. The corrected equations are written as follows:

$$\hat{\mathcal{K}} = \left(\tfrac{2}{3}\right)^{-1/3}\gamma^{-1/2}\mathcal{K} \tag{28}$$

$$\hat{\varepsilon} = \frac{E'\mu' v_{\mathrm{tip}}}{K'^2} \tag{35b}$$

$$\mathcal{E}(\mathcal{K}) = B_1\mathcal{K}^b = \left(\tfrac{2}{3}\right)^{1/3}\beta_1\hat{\mathcal{K}}^b \tag{47}$$

$$\hat{\Omega}_0(\hat{\xi}) = \hat{\xi}^{1/2} + F_{00}(\hat{\xi}) + \frac{4}{\pi}\sum_{j=1}^{n-1}\left[a_i F_{\ln}(\hat{\xi},\hat{\eta}) + b_i\,\delta F\left(\tfrac{2}{3},\hat{\xi},\hat{\eta}\right)\right]_{\hat{\eta}=\hat{\xi}_j}^{\hat{\eta}=\hat{\xi}_{j+1}} + F_{0\infty}(\hat{\xi}) \tag{C16}$$

$$\Pi_1 = b_{11}\,{}_2\mathrm{F}_1\left(c_{11},1;\tfrac{1}{2};\xi^2\right) + b_{12}\,{}_2\mathrm{F}_1\left(c_{12},1;\tfrac{1}{2};\xi^2\right) + b_{13}\,\xi^2\,{}_2\mathrm{F}_1\left(c_{13},2;\tfrac{3}{2};\xi^2\right) + b_{14}\,(2-\pi|\xi|) \qquad \text{(Appendix D, p. 927)}$$

We are grateful to Dr. Andrew Bunger for bringing to our attention the second error.

