

Journal of
Mechanics of
Materials and Structures

Volume 2, N° 3

March 2007

 mathematical sciences publishers

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES

<http://www.jomms.org>

EDITOR-IN-CHIEF Charles R. Steele
ASSOCIATE EDITOR Marie-Louise Steele
Division of Mechanics and Computation
Stanford University
Stanford, CA 94305
USA

BOARD OF EDITORS

D. BIGONI University of Trento, Italy
H. D. BUI École Polytechnique, France
J. P. CARTER University of Sydney, Australia
R. M. CHRISTENSEN Stanford University, U.S.A.
G. M. L. GLADWELL University of Waterloo, Canada
D. H. HODGES Georgia Institute of Technology, U.S.A.
J. HUTCHINSON Harvard University, U.S.A.
C. HWU National Cheng Kung University, R.O. China
IWONA JASIUK University of Illinois at Urbana-Champaign
B. L. KARIHALOO University of Wales, U.K.
Y. Y. KIM Seoul National University, Republic of Korea
Z. MROZ Academy of Science, Poland
D. PAMPLONA Universidade Católica do Rio de Janeiro, Brazil
M. B. RUBIN Technion, Haifa, Israel
Y. SHINDO Tohoku University, Japan
A. N. SHUPIKOV Ukrainian Academy of Sciences, Ukraine
T. TARNAI University Budapest, Hungary
F. Y. M. WAN University of California, Irvine, U.S.A.
P. WRIGGERS Universität Hannover, Germany
W. YANG Tsinghua University, P.R. China
F. ZIEGLER Technische Universität Wien, Austria

PRODUCTION


PAULO NEY DE SOUZA Production Manager
SHEILA NEWBERY Senior Production Editor
SILVIO LEVY Scientific Editor

See inside back cover or <http://www.jomms.org> for submission guidelines.

Regular subscription rate: \$500 a year.

Subscriptions, requests for back issues, and changes of address should be sent to Mathematical Sciences Publishers, 798 Evans Hall, Department of Mathematics, University of California, Berkeley, CA 94720-3840.

©Copyright 2008. Journal of Mechanics of Materials and Structures. All rights reserved.

 mathematical sciences publishers

A PROPOSED METHOD FOR FATIGUE CRACK DETECTION AND MONITORING USING THE BREATHING CRACK PHENOMENON AND WAVELET ANALYSIS

VIET KHOA NGUYEN AND OLUREMI A. OLATUNBOSUN

In this paper the dynamic behavior of a fatigue cracked beam is investigated. The purpose is to reveal the nonlinear behavior of the structure with fatigue damage by using wavelet transform. A cracked cantilever beam is modeled by the finite element (FE) method using ALGOR™ software. A breathing crack is described in the FE method as a surface to surface contact of the two edges of the crack during vibration. Strain time history in the area adjacent to the crack has been analyzed using data processing techniques. Nonlinear effects in signals are usually difficult to detect by conventional data processing methods such as fast Fourier transform. However wavelet transform has recently been shown to be an effective method of detecting such nonlinear effects in signals. Modulus maxima, an important property of wavelet transform, have been used as an indicator of the crack size. Numerical results obtained from the FE analysis are presented in this paper, as well as some experimental results. It is shown that detection of fatigue cracks using breathing behavior and wavelet transform can be used to develop a vibration-based crack detection technique.

1. Introduction

Fatigue cracks of a structural member often occur after a period of loading cycles. This leads to structural failure. For this reason, methods of detecting and localizing cracks have been the subject of intensive investigation in the last two decades. In these methods, nondestructive methods for crack detection that are based on the changes in the dynamic properties of the structure caused by damage have been developed as an important tool for early detection of imminent failure in mechanical and civil engineering structures.

There are two main categories of crack models used in the methods: open crack models and breathing crack models. Most researchers have assumed that the crack in a structural member is open and remains open during vibration. This assumption was made to avoid the complexity resulting from nonlinear behavior when a breathing crack is presented. Nevertheless, during vibration, a crack will open and close due to an externally applied loading. During the vibration of a structure, edges of the crack come into and out of contact, leading to sudden changes in the dynamic response of the structure. This phenomenon is known as the breathing process of the crack. By introducing the concept of a breathing crack, an intensive investigation can reveal small changes in the dynamic response of the cracked element. These changes in dynamic response can be useful for detection of cracks.

Keywords: crack detection, crack monitoring, fatigue crack, breathing crack, closing crack, wavelet transform, wavelet analysis.

Nash [1969] investigated the dynamic response of a cracked beam under impact load. In this research, the governing equation of a cracked beam was derived with the aid of a variational principle developed by Gurtin [1964]. The governing equation was then solved by the approximate method known as the *small increment method* presented by Timoshenko [1913]. Carlson [1974] and Gudmundson [1983] studied the influences of closing cracks on the dynamical characteristics of a cracked cantilever beam. The relative increase in natural frequencies due to the closing crack phenomenon has been found to be much smaller than the decrease due to an open crack. Closing cracks in beams have been investigated by Chen and Chen [1988], Actis and Dimarogonas [1989], and Collin et al. [1992] in terms of nonlinear behavior of the longitudinal free and forced vibration using direct numerical integration. Kisa and Brandon [2000] studied the effects of closure of cracks on the dynamics of a cracked cantilever beam using successive modal transformations. Xastrau [1985] used the finite element method to study the steady state responses of a simply supported beam with multiple closing cracks. Matveev and Bovsunovsky [2002] investigated nonlinear distortions of the vibration characteristics due to fatigue damage. In their studies, the level of nonlinear distortions of the displacement, acceleration and strain waves of the cracked beam were studied and the comparative evaluation of their sensitivity was carried out using a successive algorithm (cycle by cycle). The Euler–Bernoulli beam model is used widely for breathing crack studies, but not the Timoshenko beam model. This is because the Timoshenko model is much more complicated when the effects of shearing deflection and rotational inertia are taken into account [Timoshenko and Young 1955]. However analytical results of the two models are the same for general slender beams [Kikidis and Papadopoulos 1992]. The existence of nonlinear dynamic behavior due to a breathing crack was proved by these authors. They related the breathing crack to changes in natural frequency and mode shapes. However they were not able to establish practical methods of detecting a crack using the changes in frequency and mode shapes because of the difficulty in detecting minute changes in frequency and mode shapes.

On-line methods for crack detection are preferred to other methods because they do not require the test to be stopped for inspection of the test specimen for damage during testing, thus providing considerable saving in time and cost. Most of these on-line methods are based on changes in dynamic characteristics of the object, for example, frequencies, mode shapes, transfer functions, and so on. In the field of signal processing, Fourier transform has been very useful and has been widely applied for a long time. However, when transforming a signal from the time domain (or space domain) into the frequency domain using Fourier transform, the time (or space) information is lost so that it is impossible to examine simultaneously the time and frequency characteristics of an event. Moreover, Fourier transform mainly works with stationary signals, while in practice many signals appear in the form of nonstationary signals. To overcome this shortcoming of the original Fourier transform, some signal processing methods based on Fourier transform have been developed. For example, short-time Fourier transform (STFT) was proposed by Gabor [1946]. By this method, the transformed signal can retain some information in the time domain. Nevertheless, this information is less precise when only one constant window is applied to the whole data set.

The wavelet transform method, like STFT, analyzes the signal in two dimensions: time (or space) and frequency. Instead of using only a constant width window as in STFT, wavelet transform uses a variable parameter called scale. The scale in wavelet transform can play a role similar to frequency in STFT. Because of this, wavelet transform is able to analyze signals locally. From its wavelet transform,

hidden details or irregular changes in a signal could be revealed efficiently. The vibration signal of a damaged structure might contain irregular events, thus, damaged structures could be examined by wavelet transform.

Ovanosova and Suarez [2004] used wavelet transform to analyze the deflection of an open cracked beam. The position of the crack was found using bior6.8 wavelet. These authors have also developed the method for application to frame structures and obtained satisfactory results. Douka et al. [2004] presented a method for crack identification in plates based on wavelet analysis. The crack is considered to be open. The position of the crack is determined by the sudden change in the spatial variation of the transformed displacement response. To estimate the depth of the crack, an intensity factor is defined which relates the depth of the crack to the coefficients of the wavelet transform. However, there are practical difficulties in applying these methods since they are based on minute changes in mode shapes or deflections which are very difficult to measure in practice. Also, a large number of transducers are required for such measurement. Furthermore, the assumption of open cracks may be inaccurate as many fatigue cracks are breathing cracks in practice.

In brief, the existence of breathing cracks has been proved by a series of papers. Some authors have tried applying the breathing crack phenomenon in order to detect cracks but it is still complicated or difficult to deploy in practice. Crack detection using wavelet transform is mainly based on the change of mode shapes of the structure after it is damaged but the practical measurement of mode shapes of structures is not easy.

In this paper a new approach, using a combination of the nonlinear behavior of a cracked beam during vibration and wavelet transform to detect cracks is presented. For this approach, the signal that needs to be measured to detect the crack is simple: the strain time history at one point close to the crack position. When a crack appears under load, based on the breathing crack phenomenon, strain time history will include distortion during subsequent dynamic loading. Therefore application of wavelet transform to the strain time history signal provides a means of monitoring the appearance of a crack and its depth.

2. Breathing crack

Consider a cracked Euler–Bernoulli cantilever beam as in Figure 1. Free bending vibration of a beam is described by the differential equation

$$\frac{\partial^4 z(x, t)}{\partial x^4} + \frac{\rho A}{EI} \frac{\partial^2 z(x, t)}{\partial t^2} = 0, \quad (1)$$

where E is the Young's modulus, ρ is the material density of the beam, A and $I = bh^3/12$ are the cross section area and inertia moment, respectively, and b and h are the width and the height of the cross section, respectively.

The solution of Equation (1) is written in the form

$$z(x, t) = \sum_{i=1}^{\infty} w_i(x)(P_i \sin \omega_i t + R_i \cos \omega_i t), \quad (2)$$

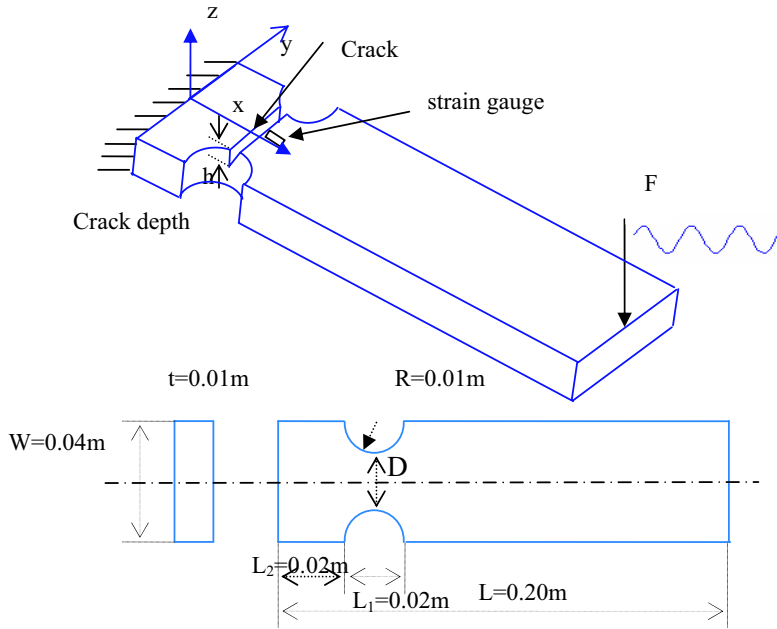


Figure 1. A cantilever beam with a crack.

where $w_i(x)$ is the i -th mode shape and ω_i is the i -th natural frequency. Mode shapes are expressed in the form [Matveev and Bovsunovsky 2002]

$$w_i(x) = A_i S(k_i x) + B_i T(k_i x) + C_i U(k_i x) + D_i V(k_i x), \tag{3}$$

where $k_i^4 = \frac{\omega_i^2 \rho A}{EI}$. S , T , U , and D are Krylov functions such that

$$\begin{aligned} S(x) &= \sin(x) - \sinh(x), & T(x) &= \cos(x) - \cosh(x), \\ U(x) &= \sin(x) + \sinh(x), & D(x) &= \cos(x) + \cosh(x). \end{aligned}$$

The coefficients A_i , B_i , C_i , D_i can be calculated from the boundary conditions.

A cantilever with a crack can be modelled in sections. The crack is modeled by a short section with reduced cross section moment of inertia, while the sections on either side of the crack have the full cross section moment of inertia. For this model, the general solution of Equation (1) for the section j will be

$$z_{oj}(x, t) = \sum_{i=1}^{\infty} w_{ij}(x) (P_{oi} \sin \omega_{oi} t + R_{oi} \cos \omega_{oi} t), \tag{4}$$

and

$$w_{ij}(x) = A_{ij} S(k_{ij} x) + B_{ij} T(k_{ij} x) + C_{ij} U(k_{ij} x) + D_{ij} V(k_{ij} x), \tag{5}$$

where $w_{ij}(x)$ is the i -th mode shape of j -th section, and $k_{ij}^4 = \frac{\omega_{oi}^2 \rho A}{EI_j}$; $j = 1, 2$. The subscript o means the open crack.

A beam with a breathing crack is modeled as follows: at the half-cycles while the crack is closed, the vibrations of the beam are described by Equation (2) (with the assumption that there is no reduction of cross section at the crack position, this leads to the assumption that the stiffness of the intact beam and the stiffness of the beam with a closed crack are the same); at the half cycles while the crack is open, the vibration of the beam is described by Equation (4). Thus, the vibration of a cracked beam is described by a combination of Equations (2) and (4).

Matveev and Bovsunovsky [2002] presented the solutions of the vibration of the beam in two halves of one complete cycle with certain initial conditions as

$$z(x, t) = w_s(x) \sin \omega_s t, \tag{6}$$

$$z_{oj}(x, t) = P_{os} w_{sj}(x) \sin \omega_{os} t. \tag{7}$$

Solution (6) is the vibration of the beam in the first half-cycle with amplitude $w_s(x)$ and solution (7) is the vibration of the beam in the second half-cycle with amplitude $P_{os} w_{sj}(x)$, where s denotes the initial mode shape, and $P_{os} < 1$ and P_{os} is dependent on crack depth and material of the beam. Therefore, the amplitudes of the vibration of the beam are different in two halves of one complete cycle. This means that a breathing crack causes a distortion or singularity of the harmonic time functions, describing vibration characteristics such as displacement, strain, or acceleration. It is shown in [Matveev and Bovsunovsky 2002] that, for the different half cycles of vibration while the crack is closed or open, the normalized distribution function of strain along the beam length can be expressed as

$$\bar{\varepsilon}(x, t) = \bar{M}_s(x) \sin \omega_s t, \tag{8}$$

$$\bar{\varepsilon}_o(x, t) = f_\varepsilon(x, \gamma) P_{cs} \bar{M}_{sj} \sin \omega_{os} t, \tag{9}$$

where s signifies the initial mode shape, $f_\varepsilon(x, \gamma)$ is distortion function due to the crack's effect on strain distribution, and $\bar{M}_s(x)$ and $\bar{M}_{ij}(x)$ are normalized distribution functions of bending moment along the length of intact and cracked beams, respectively.

3. The wavelet transforms

Wavelet transform analysis uses a little wavelike function known as a wavelet. A more accurate description is that a wavelet is a function which has local wavelike properties. In mathematical terms, the wavelet transform is a convolution of the wavelet function with the signal. A wavelet transform is defined as [Daubechies 1992]

$$Wf(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt, \tag{10}$$

where $f(t)$ is the input signal, a is a real number called scale or dilation, and b is a real number called position. $Wf(a, b)$ are wavelet coefficients of the function f , $\psi \left(\frac{t-b}{a} \right)$ is the wavelet function, and $*$ denotes complex conjugation.

Let $\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi^* \left(\frac{t-b}{a} \right)$. Equation (10) can be rewritten as

$$Wf(a, b) = \int_{-\infty}^{+\infty} f(t) \psi_{a,b} dt, \tag{11}$$

which describes the continuous wavelet transform (CWT). The CWT has an inverse transform:

$$f(t) = C_\varphi^{-1} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Wf(a, b) \psi_{a,b} db \frac{da}{a^2}, \quad (12)$$

$$C_\varphi = 2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty. \quad (13)$$

Modulus maxima. An important factor for detection of singularity are the local maxima of wavelet coefficients $Wf_{a,b}$, which are the local maxima of the derivative of $f(t)$ smoothed by $\phi(t)$. Mallat and Hwang [1992] presented a method using local maxima to detect singularity, and gave definitions as follows:

- (i) Modulus maxima are points (a_0, t_0) such that $|C(a_0, t)| < |C(a_0, t_0)|$, when t belongs to either the right or the left neighborhood of t_0 , and $|C(a_0, t)| \leq |C(a_0, t_0)|$ when t belongs to the other side of the neighborhood of t_0 .
- (ii) The maxima line is any connected curve in the scale space (a, b) along which all points are modulus maxima.

There is a connection of the regularity of a function at a point $t = t_0$ with the decay of the local maxima of the wavelet modulus across scales. To detect singularities, the asymptotic decay of the wavelet modulus maxima must be examined as the scale tends to zero. If the coefficients decay at the same rate as the scale decreases to zero, then t_0 is a singular point of $f(t)$. According to Mallat,

$$\log_2(|Wf(a, b)|) \leq \log_2 A + h \log_2(a), \quad (14)$$

where A is a constant and h is the Lipschitz regularity of function or Lipschitz exponent [Mallat and Hwang 1992]. The Lipschitz exponent gives more precise information about the differentiability of a function. For example, a function is not differentiable at $t = u$ if $0 < h < 1$. Therefore, the Lipschitz exponent h characterizes the nature of the singularity of the function $f(t)$ at $t = u$. A and h can be easily calculated from the intercept and gradient of the straight line (14) that is the asymptotic line of the curve $\log_2(|Wf(a, b)|)$ versus $\log_2(a)$. It is important to note that the Lipschitz exponent h describes the type of a singularity, as will be shown in more detail in the next sections.

4. Phase shift distribution

4.1. Definition. The definition of phase shift distribution is as follows:

- (i) During vibration, the difference phase between two strain time history signals at two points on a structure is called the phase shift.
- (ii) For a set of points along the structure, a set of phase shifts between these points and one reference point is called the phase shift distribution.

Methods for calculating phase shift distribution differ depending on types of excitation. Here, we consider two types: harmonic and random excitation.

4.2. Phase shift under harmonic excitation. It is known that under harmonic excitation at one frequency, the response of a structure will be a harmonic function of the same frequency, generally with a phase shift. The calculation of phase shift between output and harmonic input signal is presented as follows [Elsden and Ley 1969]. Assuming that the input signal is $f(t) = A \sin \omega t$ and the output signal is $\sigma(t) = R \sin(\omega t + \phi)$, the integrator outputs are

$$\frac{1}{NT} \int_0^N R \sin(\omega t + \phi) \sin \omega t \cdot dt = \frac{1}{2} R \cos \phi = \frac{a}{2}, \tag{15}$$

$$\frac{1}{NT} \int_0^N R \sin(\omega t + \phi) \cos \omega t \cdot dt = \frac{1}{2} R \sin \phi = \frac{b}{2}, \tag{16}$$

where $R = \sqrt{a^2 + b^2}$, and NT is the integration time of N cycles of the wave form of period T . Dividing (16) by (15) we have

$$\phi = \tan^{-1} \left(\frac{b}{a} \right). \tag{17}$$

For M signals $\{\sigma_i(t)\}$, and $i = 1 \div M$ corresponding to measurement points from 1 to M , we have M phase shifts with the reference as the input signal. If one signal, for example $\sigma_1(t)$, in $\{\sigma_i(t)\}$, is selected to be the reference signal, the phase shift distribution is defined as

$$\Phi = \{0, \phi_2 - \phi_1, \phi_3 - \phi_1, \dots, \phi_M - \phi_1\}. \tag{18}$$

4.3. Phase shift under random excitation. For random excitation, the phase shift distribution can be obtained from the concept of frequency response function [Wirsching et al. 1995]. A frequency response function is a mathematical representation of the relation between the input and output of a linear time-invariant system. If $x(t)$ and $y(t)$ are the input and output signals, the transfer function for every frequency ω is written in the form

$$|H(j\omega)| = \frac{Y(j\omega)}{X(j\omega)}, \tag{19}$$

where $Y(j\omega)$ and $X(j\omega)$ are Fourier transforms of $x(t)$ and $y(t)$.

The phase shift for any frequency ω can be calculated as

$$\begin{aligned} \phi(\omega) &= \text{angular} [H(j\omega)] \\ &= \text{angular} [Y(j\omega)] - \text{angular} [X(j\omega)]. \end{aligned} \tag{20}$$

If $y^*(t)$ is another output signal, the phase shift between $y^*(t)$ and $x(t)$ is

$$\begin{aligned} \phi^*(\omega) &= \text{angular} [H^*(j\omega)] \\ &= \text{angular} [Y^*(j\omega)] - \text{angular} [X(j\omega)], \end{aligned} \tag{21}$$

where Y^* is Fourier transform of $y^*(t)$, and H^* is frequency response function between $x(t)$ and $y^*(t)$. From (20) and (21), the phase shift between two output signals $y(t)$ and $y^*(t)$ is calculated as

$$\begin{aligned} \phi^*(\omega) - \phi(\omega) &= \text{angular} [H^*(\omega)] - \text{angular} [H(\omega)] \\ &= \text{angular} [Y^*(j\omega)] - \text{angular} [Y(j\omega)]. \end{aligned} \tag{22}$$

Case	Crack depth (%)
1	0
2	10
3	20
4	30
5	40
6	50
7	60

Table 1. Seven cases with cracks of varying depths at crack position $x = 30$ mm.

Similar to harmonic excitation situation, if we have M signals $\{\sigma_i(t)\}$ corresponding to measurement points from 1 to M , the phase shift distribution with reference to $\sigma_1(t)$ is defined as

$$\Phi = \{0, \phi_2 - \phi_1, \phi_3 - \phi_1, \dots, \phi_N - \phi_1\}. \quad (23)$$

5. Numerical studies

5.1. Strain analysis. To analyze the dynamic response of a cracked beam, the ALGOR - finite element software was used. The crack is described as shown in Figure 1. In this model the crack includes two close edges. In ALGOR, a surface to surface contact tool is used in a mechanical event simulation (MES) to model the breathing crack. It is expected that when the load is a sinusoidal function, the strain function should not be completely sinusoidal, but should be distorted when the crack closes.

From the definition of a wavelet transform, Equation (10), it is obvious that wavelet coefficients are proportional to vibration amplitude. On the other hand, vibration amplitude of the beam is proportional to load amplitude. This leads to the fact that the distortion of strain is proportional to force level. For this reason, a normalization of strain time history is needed before data processing, to detect the level of the crack.

Seven levels of the crack from zero to 60% were examined. These seven cases are numbered as in Table 1. The two types of applied forces are harmonic and random functions.

5.2. Detection of crack existence and crack depth using wavelet transforms.

5.2.1. Detection of crack existence. The wavelet transforms were applied for analyzing strain-time signals at the point adjacent to the crack under dynamic load for seven levels of crack depth. The wavelet function used was db2, which is the Doubachives wavelet family with two vanishing moments. In this analysis, the discrete wavelet transform was applied. The original signal is strain time history.

Figures 2–5 describe strain-time history and its wavelet transform for specific levels of the crack. In each figure, the upper graph shows strain time history and the lower graph shows its wavelet transform. Obviously, the nonlinear phenomena of strain signals cannot be detected visually in all strain-time history graphs. To reveal the nonlinear phenomena caused by the appearance of the crack, wavelet transform is applied. As can be seen in the lower graphs, when there is no crack, or the crack depth is smaller than 30%, no discontinuity in the wavelet transform is shown. However, when the crack depth is equal or

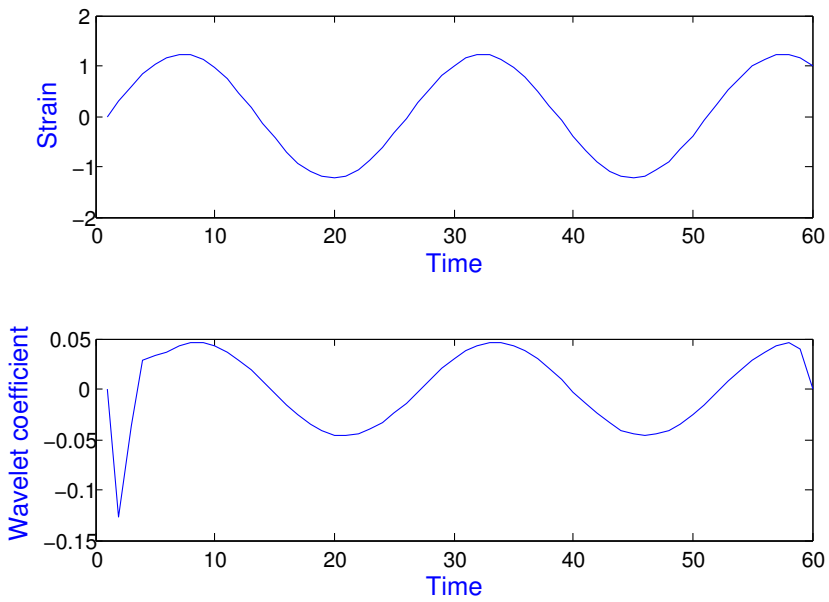


Figure 2. Intact specimen.

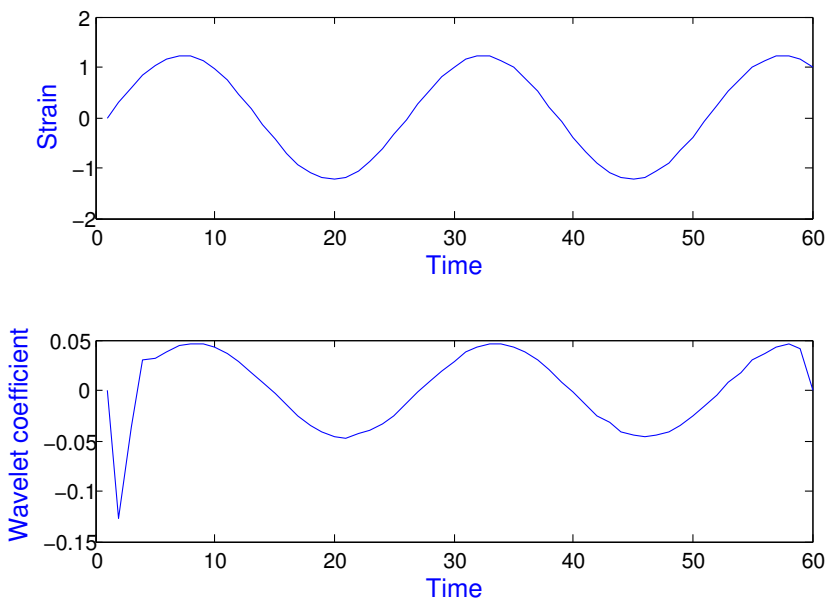


Figure 3. Specimen with 20% crack depth.

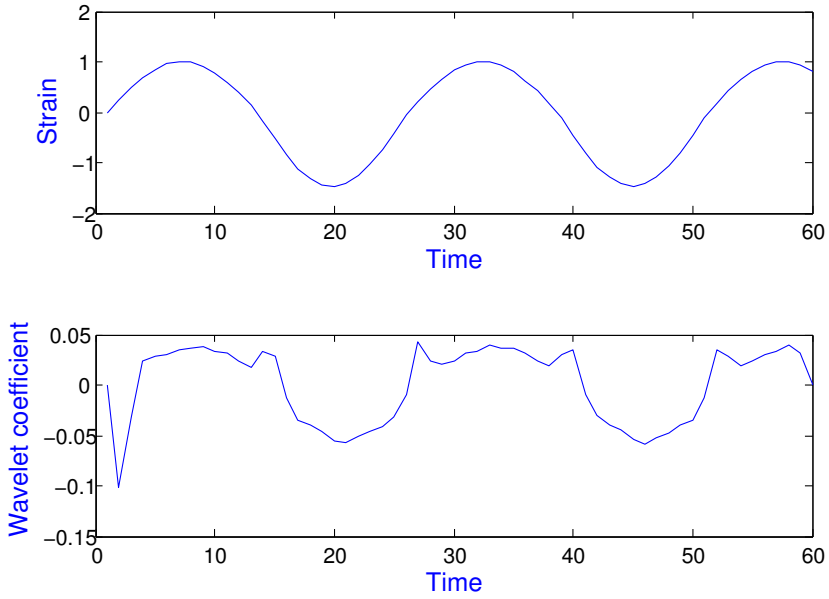


Figure 4. Specimen with 40% crack depth.

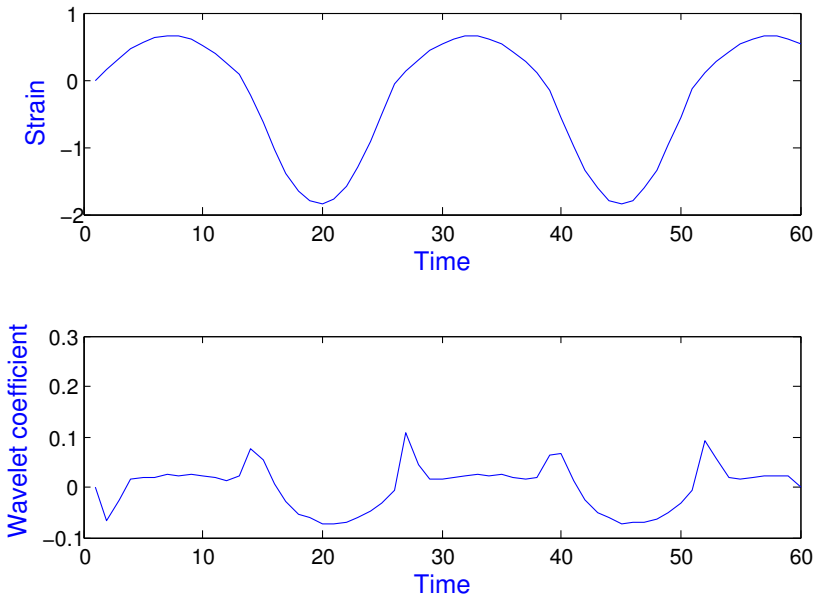


Figure 5. Specimen with 60% crack depth.

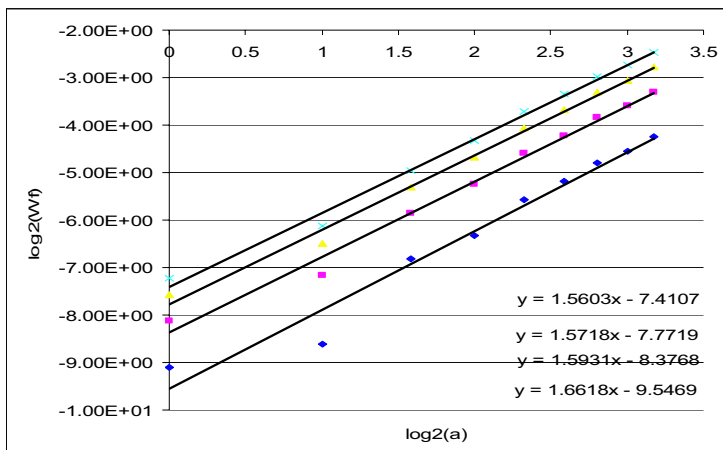


Figure 6. Maxima lines versus scale for five levels of crack: the lowest line is for the crack of 30% and the highest line is for the crack of 60%.

greater than 30%, the wavelet transform clearly shows discontinuities at moments when breathing of the crack occurs. These discontinuities of the wavelet transform indicate that the test specimen contains a crack.

5.2.2. Detection of crack depth. To determine crack depth, the modulus maxima lines of the wavelet transform have to be investigated. For this purpose, Equation (14) is used to describe the relationship between crack depth and the coefficients of the wavelet transform.

As can be seen in Figure 6 the wavelet maxima lines versus scale are shown for crack depths varying from 30 % to 60% of the wavelet transform. These four lines are parallel lines of the same slope, or in other words, of the same Lipschitz exponent. The Lipschitz exponent for all cases has a constant value of 1.6, which indicates that singularities of the same type are caused by the same physical process — in this case, the existence of a crack. From Figure 6, when the Lipschitz exponent h is fixed, only the intercept A changes when the crack depth changes. Thus, each parallel line can be distinguished by its intercept A . This intercept increases when crack depth increases. The intercept A therefore can be considered as an intensity factor which relates the extent of the fatigue crack to the wavelet coefficients.

Establishing a graph of intercept A versus crack depth from Figure 6, a relationship between intercept A (or intensity factor) and crack depth is obtained as shown in Figure 7. The values of A plotted are the linear values derived from the log-log plot in Figure 6. It can be seen that this relationship is the straight line in semilog plot. Because of this, Figure 7 can be used to predict the crack depth if the intensity factor is known.

5.3. Detection of crack position. A method of detection of crack position using phase shift distribution of strain signals along the structure is proposed. The idea is that the strain wave traveling along the beam might be influenced by a sudden change in its transmission direction, such as a crack. Therefore, a phase shift of the strain signal at points along the beam could be influenced as well and could be used to detect

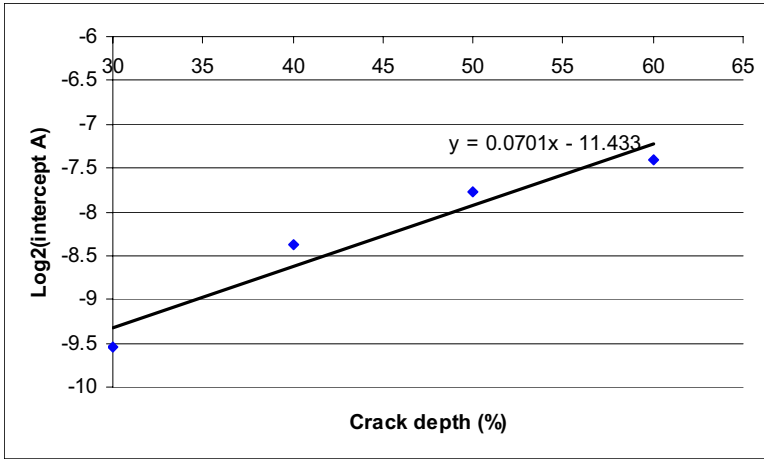


Figure 7. Intensity factor versus crack depth.

the position of the crack. Due to the existence of damage, the phase shift distribution of a damaged structure is expected to differ from that of an intact structure, at the position of the damage.

To investigate the phase shift distribution, six strain time history signals are obtained at six points, from finite element (FE) analysis as shown in Figure 8. The relative positions of the six points labeled from 1 to 6 are 0, 5.5, 10.6, 16.4, 22.7, and 28.7 mm, respectively. The crack position is between points 3 and 4, at a position of 13.5 mm.

5.3.1. Detection of crack position using phase shift under harmonic excitation. Figures 9–12 present phase shift distributions calculated from FE analysis results for the beam with increasing crack depth

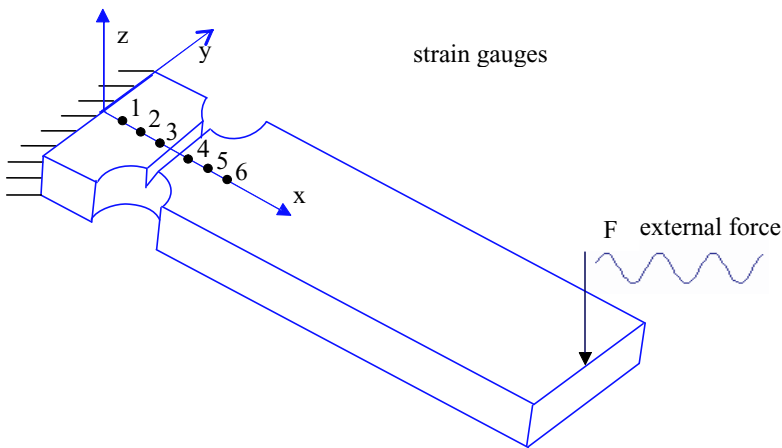


Figure 8. Positions of the six points along the beam.

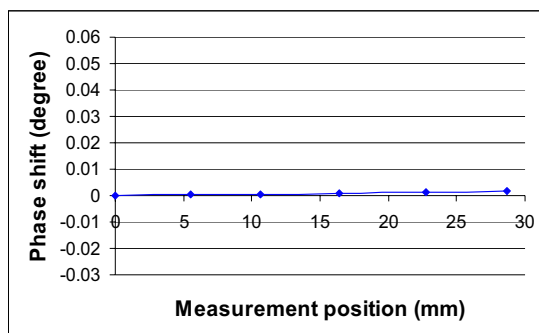


Figure 9. Phase shift distribution of the intact beam.

in the case of harmonic excitation. As can be seen in Figure 9, when the beam has no crack, the phase shift distribution is smooth. In Figure 10 when the crack depth increases to 20%, there is a distortion in the phase shift distribution at a position between point 3 and point 4. It is interesting that the shape of the distortion remains the same, and only the amplitude increases when the crack depth increases (see Figures 11 and 12).

From this result, it is obvious that the crack position can be detected as the position of the distortion in the phase shift distribution.

5.3.2. Detection of crack position using phase shift under random excitation. Figures 13–16 show phase shift distribution of the beam under random excitation for specific cases when crack depth increases from zero to 60%. The phase shift distribution is calculated at the resonant frequency.

Clearly, the shape and position of the distortion in phase shift distribution changes when the excitation and crack depth vary. This means that under random excitation, there is no rule to relate the position of the crack to the distortion of phase shift distribution. In other words, phase shift distribution cannot be used to detect the crack position under random excitation.

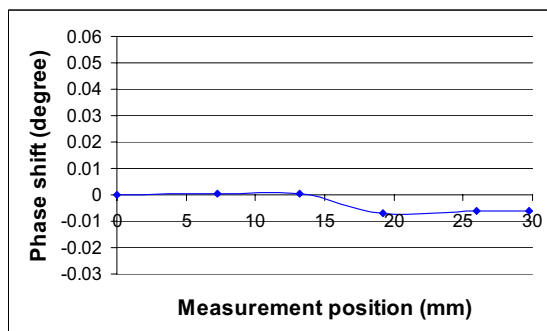


Figure 10. Phase shift distribution of the beam with crack of 20% depth.

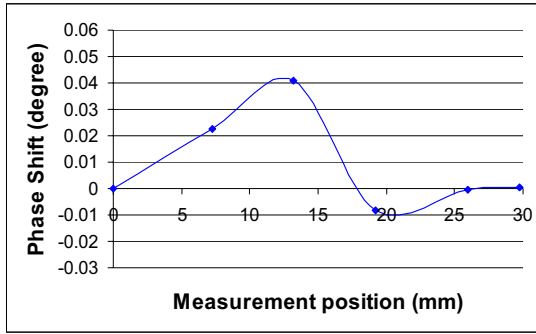


Figure 11. Phase shift distribution of the beam with crack of 40% depth.

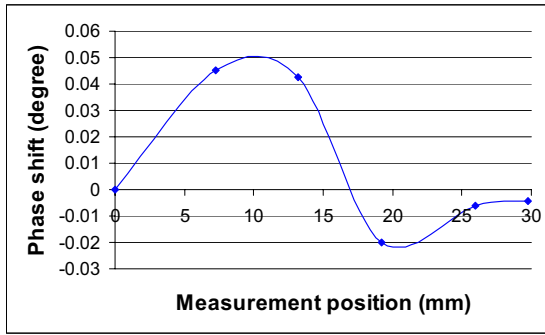


Figure 12. Phase shift distribution of the beam with crack of 60% depth.

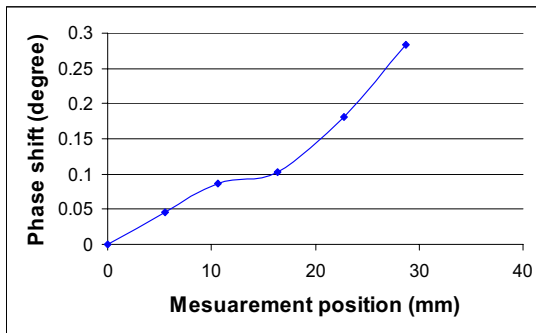


Figure 13. Phase shift distribution of the intact beam.

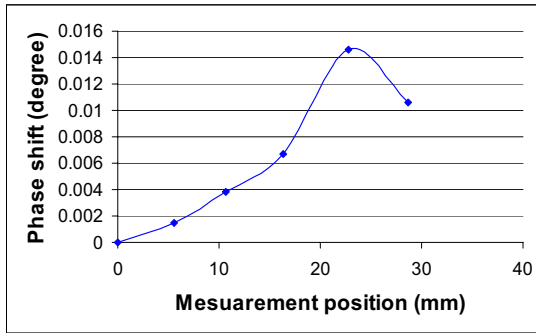


Figure 14. Phase shift distribution of the beam with crack of 20% depth.

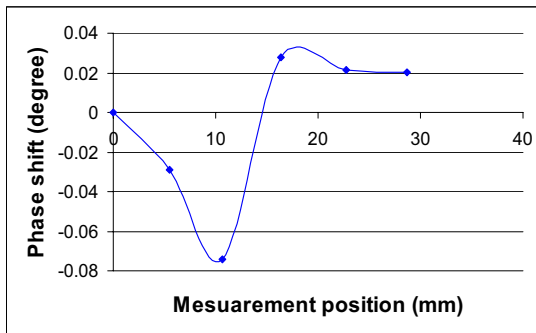


Figure 15. Phase shift distribution of the beam with crack of 40% depth.

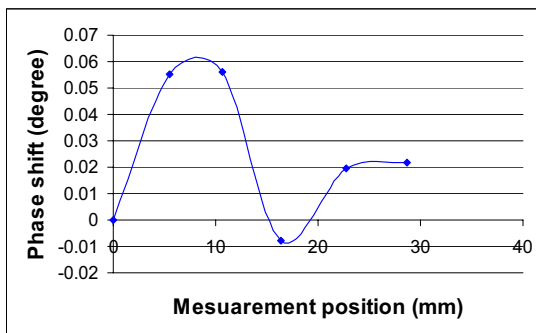


Figure 16. Phase shift distribution of the beam with crack of 60% depth.

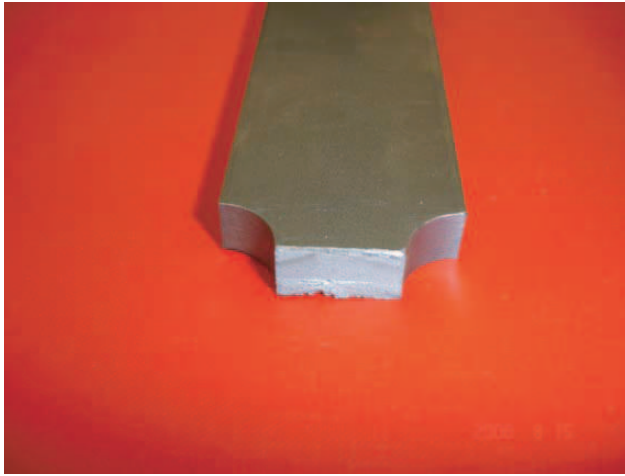


Figure 17. Crack shape of the specimen.

6. Experimental results

To validate the method for crack detection using the breathing crack model and wavelet analysis, fatigue tests on a cantilever beam have been carried out. A specimen made from mild steel BS970 was used, with a size similar to the beam in the FE model (see Figure 1). The excitation force was generated using an electro-hydraulic actuator and was controlled by LabView using a National Instruments device and Kelsey Instruments controller. Exciting frequency was 10 Hz. The crack shape obtained from experiment is quite similar to the numerical case: the crack line is fairly straight and edges of the crack are reasonably flat (see Figure 17).

6.1. Detection of crack existence. Figures 18–21 describe strain time history and its wavelet transform for six cases taken at different stages during the fatigue test on the specimen. In these figures, wavelet transform at level 2 and wavelet function db2 are applied.

Visually, the distortions in strain time histories caused by the crack are not apparent (see upper graphs in Figures 18–21). However, these distortions are highlighted from the wavelet transform (lower graphs in Figures 19–21). Figure 18 shows the measured strain and its wavelet transform for the specimen at the beginning of the test when no crack has yet developed. Obviously, there is no peak in the wavelet transform for this case. It means that when there is no crack, the wavelet transform does not give any information about the crack. As can be seen in Figures 18–21, when the crack is from 19% to 58%, the wavelet transforms show clear peaks. This indicates the presence of distortion in the strain time history and, hence a crack in the specimen. However, the peaks in wavelet transform of the measured signals are not as clean as in the FEM results. This can be explained by the background noise in the measured signals during testing, resulting from sources such as the electro-hydraulic pump and the exciter mechanism.

6.2. Detection of crack depth. Figure 22 shows a comparison of the relationship between crack depth and the intercept A in two cases: experimental and FE analysis. As can be seen in this figure, the relationship in semilog plot between crack depth and value of intercept A obtained from experimental

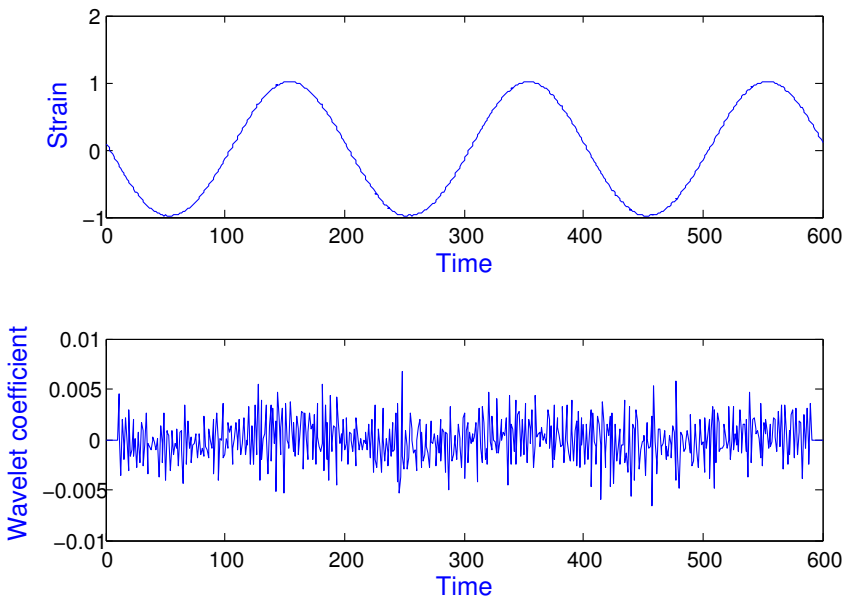


Figure 18. Strain and its wavelet transform; crack is 0%.

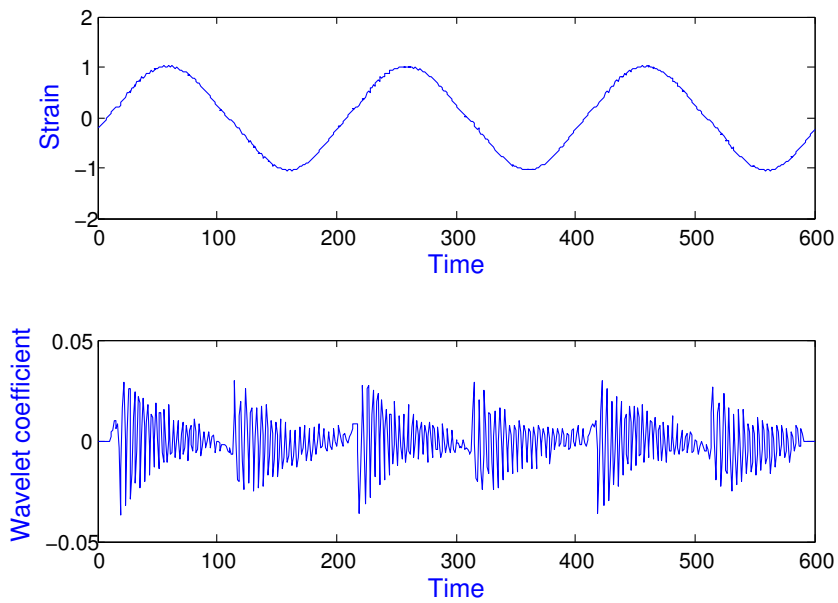


Figure 19. Strain and its wavelet transform; crack is 19%.

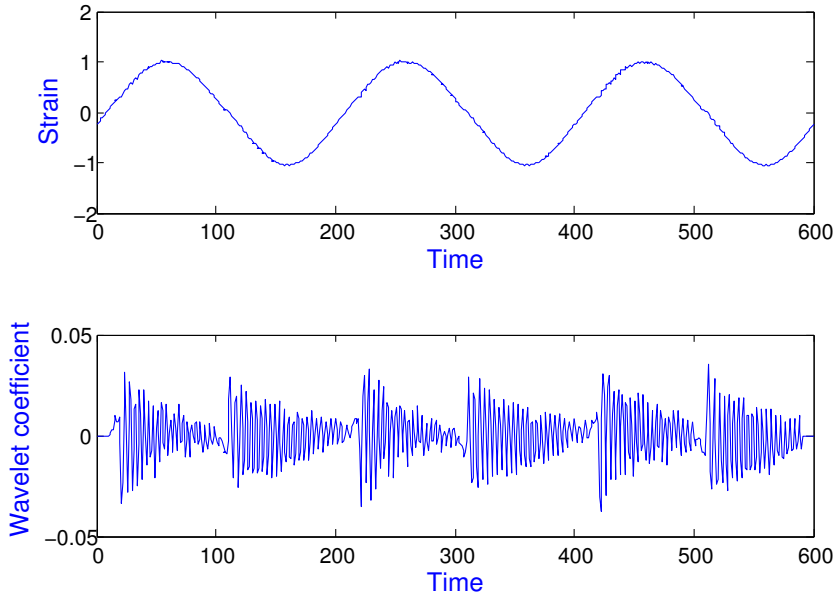


Figure 20. Strain and its wavelet transform; crack is 41%.

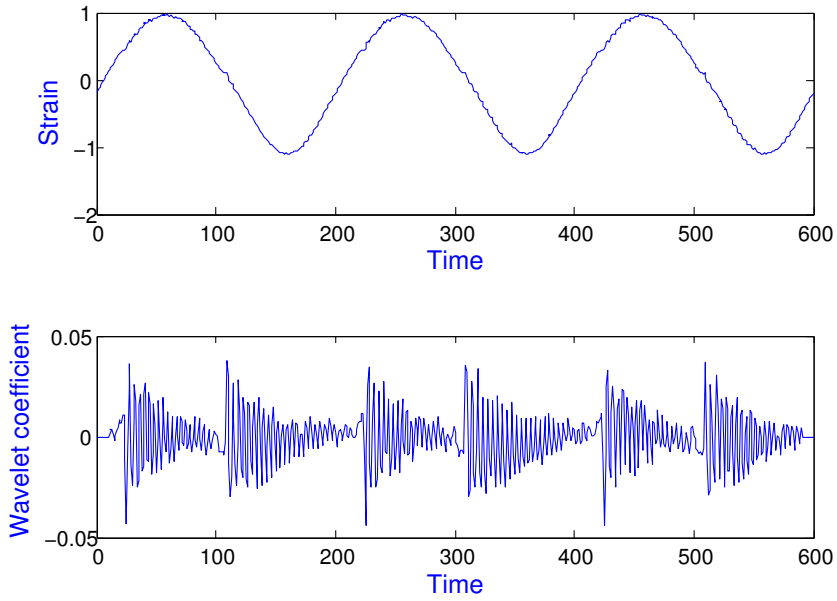


Figure 21. Strain and its wavelet transform; crack is 58%.

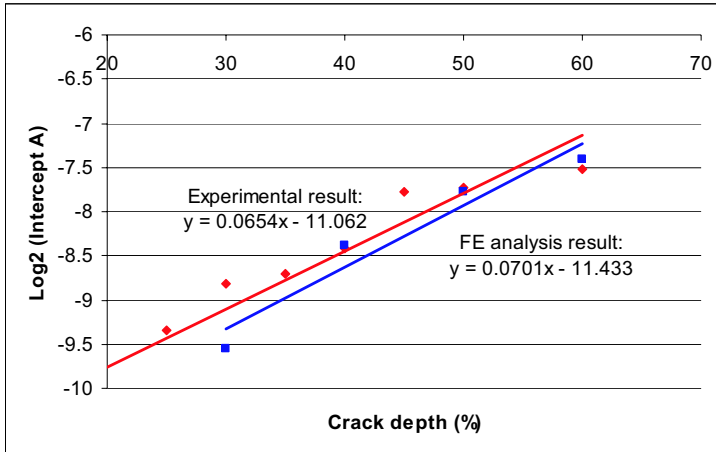


Figure 22. Relationship between crack depth and intercept. The red line is the experimental result, and the blue line is the FE analysis result.

results is linear. The intercept A increases when the crack depth increases. Note that there are still differences between numerical and experimental results. The Lipschitz exponent from the experiment is 1.0 and differs from FE analysis, where the value of the Lipschitz exponent is 1.6. This can be explained by the fact that the two edges of the crack do not come into contact purely, but rather the contact of these two edges may contain sticking, frictionless sliding, and friction sliding phenomena. On the other hand, the surfaces of two edges of a crack are not as smooth as when modeled in FE analysis. Background noise, as mentioned in Section 6.1, may also be a contributory factor to the difference between numerical and experimental results.

6.3. Detection of crack position. To measure phase shift distribution, four strain gauges are used. The relative positions of these points from 1–4 are 0, 17, 25, and 46 mm, respectively. Three tests have been carried out and gave the same results. The actual crack position is found at 20 mm which means the crack is between point 2 and point 3. Figures 23–26 show the phase shift distribution for specific cases taken at different stages during the fatigue test on the specimen.

Figure 23 shows the phase shift distribution for the specimen at the beginning of the test when no crack has yet developed. As can be seen in this figure, the phase shift distribution when there is no crack is above the horizontal axis and quite smooth as expected. From Figures 24–26, when the crack for these cases is from 19% to 58%, the shapes of phase shift distributions remain the same while the distortion is found clearly at a position between point 2 and point 3. Thus, it can be said that the crack position is well detected by the phase shift distribution.

7. Conclusions

A method for crack detection in a beam based on the breathing crack phenomenon and wavelet transform has been presented. The appearance of the crack is detected by the singularities in strain time history caused by the breathing phenomenon of the crack. Such singularities are usually difficult to discriminate

visually. However, applying the wavelet transform has given clear pictures of these singularities. The relationship between crack depth and intensity factor was also determined. The linear relationship between crack depth and intensity factor in a semilog plot has been obtained in both numerical and experimental data.

A fatigue test with a simple specimen has been carried out in the laboratory. The experimental results are in agreement with the results of FE analysis in terms of detecting the appearance of the crack. When there is no crack or the crack depth is small, there is no discontinuity in the wavelet transform. Although the experimental signals were noisy, the appearance of the crack was detected quite clearly by peaks in the wavelet transform. The estimation of crack depth based on experimental data has been carried out. The relationship between crack depth and the intensity factor A is established. It is shown that the intensity factor (intercept A) increases when the crack depth increases. However, there is a significant difference in experimental results when compared with numerical analysis in terms of detecting crack depth, due to the influences of real conditions during the test.

A new approach based on phase shift distribution for the detection of crack position is also proposed. The position of the crack is accurately detected when the specimen is excited under harmonic load. The crack position is determined by the position of the distortion in phase shift distribution. The phase distribution has to be determined under harmonic excitation, as it has been shown that this technique does not work under random excitation.

To conclude, the method of combining the breathing crack phenomenon and wavelet transform is adequate to apply for monitoring the appearance of a crack in the test specimen. The nonlinearity resulting from the breathing crack phenomenon is clearly shown by wavelet transform of the strain time history. The advantages of this technique are that it is capable of being implemented in a practical manner, and the strain time history signal can be easily measured in practice. Also, the amount of measured data required for crack detection is easily manageable for the implementation of on-line crack detection and monitoring.

This paper has established the proposed technique for structural health monitoring using a simple structure. Further studies on real components under real test conditions are now in progress.

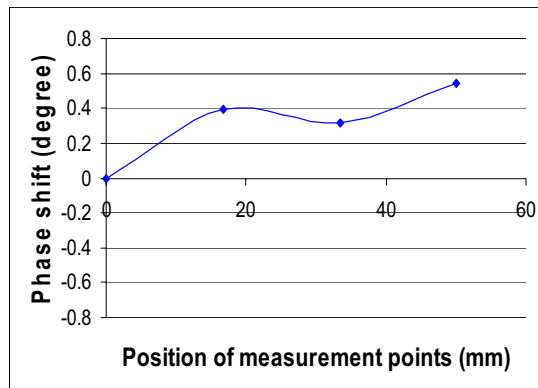


Figure 23. Phase shift distribution of the intact beam.

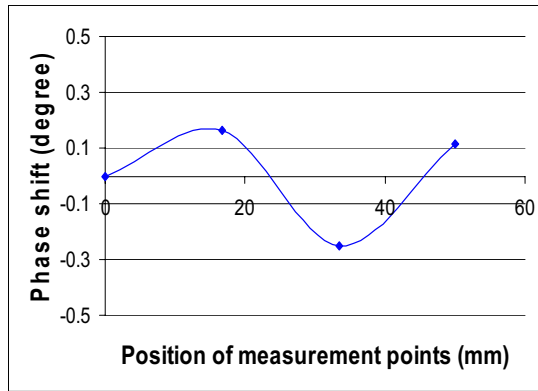


Figure 24. Phase shift distribution of the beam with crack of 19% depth.

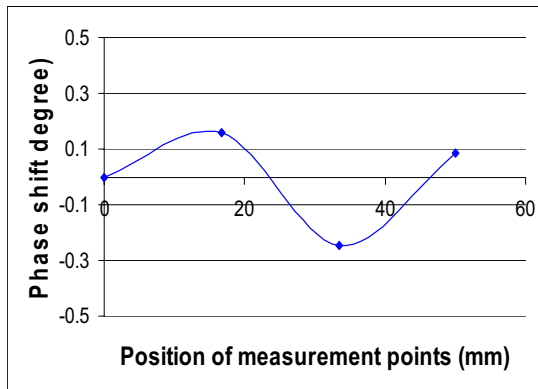


Figure 25. Phase shift distribution of the beam with crack of 41% depth.

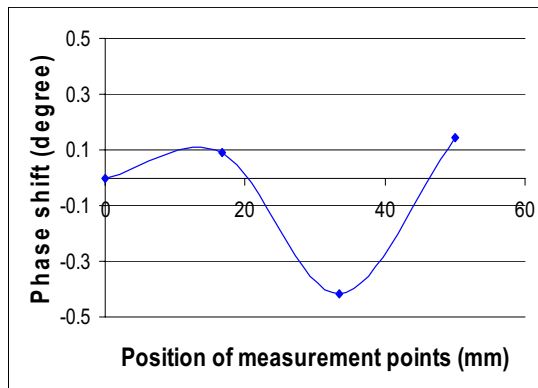


Figure 26. Phase shift distribution of the beam with crack of 58% depth.

References

- [Actis and Dimarogonas 1989] R. L. Actis and A. D. Dimarogonas, “Non-linear effects due to closing cracks in vibrating beams”, pp. 99–104 in *Structural vibration and acoustics: 12th Biennial Conference on Mechanical Vibration and Noise* (Montréal), edited by T. S. Sankar et al., ASME Design Engineering **18**(3), ASME, New York, 1989.
- [Carlson 1974] R. L. Carlson, “An experimental study of the parametric excitation of a tensioned sheet with a crack-like opening”, *Exp. Mech.* **14**:11 (1974), 452–458.
- [Chen and Chen 1988] L. W. Chen and C. L. Chen, “Vibration and stability of cracked thick rotating blades”, *Comput. Struct.* **28**:1 (1988), 67–74.
- [Collin et al. 1992] K. R. Collin, P. H. Plaut, and P. H. Wauer, “Free and forced longitudinal vibrations of a cantilevered bar with crack”, *J. Vib. Acoust. (ASME)* **114** (1992), 171–177.
- [Daubechies 1992] I. Daubechies, *Ten lectures on wavelets*, CBMS Regional Conference Series in Applied Mathematics **61**, SIAM, Philadelphia, 1992.
- [Douka et al. 2004] E. Douka, S. Loutridis, and A. Trochidis, “Crack identification in plates using wavelet analysis”, *J. Sound Vib.* **270**:1–2 (2004), 279–295.
- [Elsden and Ley 1969] C. S. Elsdén and A. J. Ley, “A digital transfer function analyser based on pulse rate techniques”, *Automatica* **5**:1 (1969), 51–58.
- [Gabor 1946] D. Gabor, “Theory of communication”, *J. Inst. Elect. Eng.* **93** (1946), 429–457.
- [Gudmundson 1983] P. Gudmundson, “The dynamic behaviour of slender structures with cross-sectional cracks”, *J. Mech. Phys. Solids* **31**:4 (1983), 329–345.
- [Gurtin 1964] M. E. Gurtin, “Variational principles for linear elastodynamics”, *Arch. Ration. Mech. An.* **16**:1 (1964), 34–50. MR 35 #5173 Zbl 0124.40001
- [Kikidis and Papadopoulos 1992] M. L. Kikidis and C. A. Papadopoulos, “Slenderness ratio effect on cracked beam”, *J. Sound Vib.* **155**:1 (1992), 1–11.
- [Kisa and Brandon 2000] M. Kisa and J. Brandon, “The effects of closure of cracks on the dynamics of a cracked cantilever beam”, *J. Sound Vib.* **238**:1 (2000), 1–18.
- [Mallat and Hwang 1992] S. Mallat and W. Hwang, “Singularity detection and processing with wavelets”, *IEEE Trans. Inform. Theory* **38**:2 (1992), 617–643.
- [Matveev and Bovsunovsky 2002] V. V. Matveev and A. P. Bovsunovsky, “Vibration-based diagnostics of fatigue damage of beam-like structures”, *J. Sound Vib.* **249**:1 (2002), 23–40.
- [Nash 1969] G. E. Nash, “Mechanical aspect of the dynamic tear test”, *J. Basic Eng. (ASME)* **91** (1969), 535–543.
- [Ovanesova and Suarez 2004] A. V. Ovanesova and L. E. Suarez, “Applications of wavelet transforms to damage detection in frame structures”, *Eng. Struct.* **26**:1 (2004), 39–49.
- [Timoshenko 1913] S. P. Timoshenko, “Zur Frage nach der Wirkung eines Stosses auf einen Balken”, *Z. Math. Phys.* **62** (1913), 198–209.
- [Timoshenko and Young 1955] S. P. Timoshenko and D. H. Young, *Vibration problems in engineering*, 3rd ed., Van Nostrand, New York, 1955.
- [Wirsching et al. 1995] P. H. Wirsching, T. L. Paez, and K. Ortiz, *Random vibrations: theory and practice*, Wiley, New York, 1995.
- [Xastrau 1985] B. Xastrau, “Vibration of cracked structures”, *Arch. Mech.* **37** (1985), 731–743.

Received 20 Jun 2006. Accepted 23 Sep 2006.

VIET KHOA NGUYEN: nguyenvietkhoa_vc@yahoo.com

School of Mechanical Engineering, The University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

OLUREMI A. OLATUNBOSUN: O.A.Olatunbosun@bham.ac.uk

School of Mechanical Engineering, The University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

MIXED PIEZOELECTRIC PLATE ELEMENTS WITH CONTINUOUS TRANSVERSE ELECTRIC DISPLACEMENTS

ERASMO CARRERA AND CHRISTIAN FAGIANO

This paper proposes mixed finite elements, FEs, with an *a priori* continuous transverse electric displacement component \mathcal{D}_z . The Reissner Mixed Variational Theorem (RMVT) and the Unified Formulation (UF) are applied to the analysis of multilayered anisotropic plates with embedded piezoelectric layers. Two forms of RMVT are compared. In a first, *partial*, form (P-RMVT), the field variables are displacements \mathbf{u} , electric potential Φ and transverse stresses σ_n . The second, *full*, form (F-RMVT) adds \mathcal{D}_z as an independent variable. F-RMVT allows the *a priori* and complete fulfillment of interlaminar continuity of both mechanical and electrical variables.

We treat both equivalent single-layer models (ESLM), where the number of variables is kept independent of the number of layers, an layerwise models (LWM), in which the number of variables depends in each layer. According to the UF the order N of the expansions assumed for the \mathbf{u} , ϕ , σ_n and \mathcal{D}_z fields in the plate thickness direction z as well as the number of the element nodes N_n have been taken as free parameters.

In most cases the results of the classical formulation which are based on Principle of Virtual Displacements (PVD) are given for comparison purposes. The superiority of the F-RMVT results, with respect to the P-RMVT and to PVD ones, is shown by few examples for which three-dimensional solution is available. In particular, the F-RMVT results to be very effective for the evaluation of interlaminar continuous \mathcal{D}_z fields.

1. Introduction

In recent years piezoelectric materials have been integrated with structural systems to build smart structures which are the candidates for next generation structures of aerospace vehicles as well as for some advanced products in the automotive and ship industries. Piezoelectric materials are, in fact, capable of altering the response of the structures through sensing and actuation [Tiersten 1969]. By integrating the surface bonded and embedded actuators in structural systems, the desired localized strains may be induced in the structures thanks to the application of an appropriate voltage to the actuators. Such an electromechanical coupling allows closed-loop control systems to be built up, in which piezomaterials play the role of both the actuators and the sensors. An intelligent structure can therefore be built in which, for instance, thermomechanical deformations or vibrations can be reduced by using appropriate control laws. For details see [Chopra 1996; 2002] and the related literature.

In order to successfully incorporate actuator/sensors in a structures, the mechanical interaction between the piezoelectric layers and the hosting structure must be completely understood, that is, an

Keywords: piezoelectric plates, finite elements, mixed method, transverse continuity, unified formulation.

This work has been carried out in the framework of STREP EU project CASSEM under contract NMP-CT-2005-013517.

appropriate use of piezoelectric materials, requires an accurate description of both the electrical and mechanical fields in the constitutive layers. Early mechanical models were developed by Crawley and de Luis [1987], Lee [1990] and Mitchell and Reddy [1995], among others. More recent works are [Yang and Batra 1995; Wang et al. 1997; Vidoli and Batra 2000; Batra and Vidoli 2002]. A recent assessment of classical and refined theories with displacements and electrical variables for plates can be found in [Ballhause et al. 2005]. Equivalent single-layer (ESL) and layerwise (LW) theories have been compared in the framework of the application of the Principle of Virtual Displacement (PVD) applications (it is intended that the number of independent variables is kept independent by the number of the layers in the ESL models). Numerous benchmark, exact solution analyses have been conducted for piezoelectric plates; some are given in [Heyliger and Saravanos 1995]. However, these benchmark solutions are restricted to simple geometries and special boundary conditions. The treatment of more realistic problems would require the use of efficient computational tools such as the finite element method (FEM).

The present paper focuses on FEM electromechanical two-dimensional modelings of smart structures with embedded piezo layers. Finite element studies were conducted by Robbins and Reddy [1991]. A finite element that accounts for a first order shear deformation theory (FSDT) description of displacement and layerwise form of the electric potential was developed in [Sheikh et al. 2001]. The numerical, membrane and bending behavior of the FEs based on FSDT was analyzed in [Auricchio et al. 2001] in the framework of a suitable variational formulation. The third-order theory was applied by Thornbuegh and Chattopadhyay [2002] to derived finite elements that account for electromechanical coupling. Similar elements have more recently been considered in [Shu 2005]. Extension of the third-order Ambartsumian zigzag multilayered theory [Carrera 2003a] to the finite analysis of electromechanical problems has been proposed by Oh and Cho [2004]. An extension of numerically efficient plate/shell elements based on mixed interpolation of tensorial components (MITC) to piezoelectric plates has recently been provided by Kögl and Bucalem [2005a; 2005b]. We also mention the review papers [Saravanos and Heyliger 1999; Benjeddou 2000; Wang and Yang 2000].

Our contributions to the application of the Reissner Mixed Variational Theorem (RMVT) to multilayered made structures started with [Carrera 1995; 1996; 2001], and have included closed-form solution analyses [Carrera 1999a; 1999b] and FE applications [Carrera and DeMasi 2002a; 2002b], showing the RMVT is a very suitable tool to provide quasi-3D description of stress and strain fields in anisotropic laminated structures. The RMVT was also employed in the framework of Unified Formulation (UF), dealt with in detail in [Carrera 2001]. The main feature of UF is that it allows one to formulate both ESLM and LW models in terms of a few fundamental nuclei whose forms do not depend on either the order of the expansion N used for the various variables (in the thickness direction) or on by the number of nodes of the element N_n . The Murakami zigzag Function (MZZF) [Carrera 2001] was used to reproduce the zigzag form of displacement field in the ESLM case. A classical formulation, based on PVD, was developed for comparison purposes.

A first application of RMVT to piezoelectric plates was provided in [Carrera 1997], where an MITC-type plate element was extended to nonlinear dynamic analysis of piezoelectric, composite plate. The UF formulation was applied, in the PVD framework, to piezoelectric plates in [Ballhause et al. 2005]; attention was restricted to analytical closed form solutions. RMVT closed form solutions were presented in [D'Ottavio and Kröplin 2006], while extension to shell has been provided in [Carrera et al. 2005]. Finite element applications have also been provided recently [Carrera and Boscolo 2006].

All these RMVT works have been restricted to the *a priori* fulfillment of the interlaminar continuity of the mechanical variables (transverse normal and shear stress fields), that is, the continuity of transverse normal component \mathcal{D}_z of the electric displacement vector was not *a priori* guaranteed. This form of RMVT is herein referred to as the *partial* form, or P-RMVT. RMVT has also been applied in [Garcia Lage et al. 2004a] to develop LW piezoelectric plate elements in the static case. The transverse component of electric displacement \mathcal{D}_z was considered as an assumed variable. We refer to such an extension as *full* RMVT applications, namely F-RMVT. Garcia Lage and his coauthors restricted their attention to the quadratic distribution of displacements (mechanical and electrical) and transverse stress unknowns, and treated only a layerwise model. These restrictions have not allowed us to analyze the features of the *a priori* assumption of interlaminar continuous transverse electric displacement.

Here we compares P-RMVT and F-RMVT in the framework of UF, extending the analysis of [Carrera and Boscolo 2006] to include the normal electrical displacement \mathcal{D}_z as an assumed *a priori* variable. A number of new finite elements are derived and systematically compared to those based on P-RMVT and PVD. ESLM and LW variable description analyses are compared to available 3D solutions. Up to forth-order expansions in the thickness plate/layers have been implemented.

The paper is organized as follows. Section 2 gives the necessary preliminaries. Section 3 introduces the two RMVT forms for piezoelectric continua along with variationally consistent constitutive equations. The UF for finite element applications are derived in Section 3, and the FE matrices themselves in Section 5. Section 6 contains numerical results and discussion.

2. Preliminaries

Figure 1 shows the geometry and the coordinate system of a laminated plate with N_l layers, including piezoelectric layers. The reference system is denoted by x, y, z ; the correspondent plate dimensions are denoted by a, b, h , the last of which is the thickness.

The material properties of a piezoelectric continuum can be expressed in different forms; we use the so-called *e*-form [Ikeda 1996]. The relevant energy is then the electric Gibbs energy G_2 , which takes the form

$$G_2 = \frac{1}{2} \boldsymbol{\epsilon}^T \mathbf{C}^{\mathcal{E}} \boldsymbol{\epsilon} - \boldsymbol{\mathcal{E}}^T \boldsymbol{e} \boldsymbol{\epsilon} - \frac{1}{2} \boldsymbol{\mathcal{E}}^T \boldsymbol{\epsilon}^{\mathcal{E}} \boldsymbol{\mathcal{E}}, \tag{1}$$

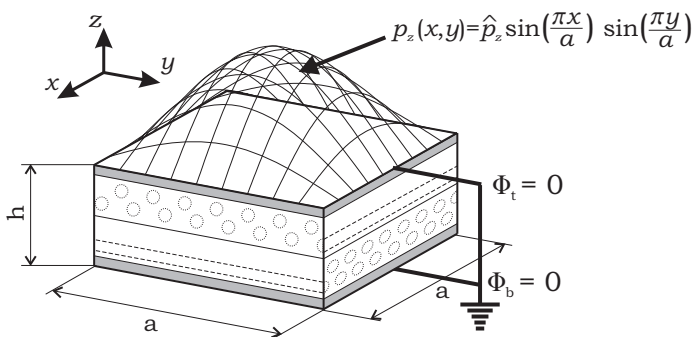


Figure 1. Geometry of Piezoelectric Plate

where $\boldsymbol{\epsilon}^T = \{\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}, \epsilon_{xz}, \epsilon_{yz}, \epsilon_{xy}\}$ is the strain tensor (we use bold letters for arrays and T to denote transposition), $\mathcal{E}^T = \{\mathcal{E}_x, \mathcal{E}_y, \mathcal{E}_z\}$ is the electric field vector, $\mathbf{C}^{\mathcal{E}}$ is the stiffness matrix calculated at constant \mathcal{E} , \mathbf{e} is the piezoelectric matrix that couples electrical and mechanical fields, and $\boldsymbol{\epsilon}^{\epsilon} = \{\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}\}$ is the permittivity matrix calculated at ϵ -constant.

The constitutive equations will be written out in Section 3 in a form suitable for the F-RMVT application.

Geometrical relations. The strain-displacement geometrical (subscript G) relations in the linear case are

$$\boldsymbol{\epsilon}_{pG}^k = \mathbf{D}_p \mathbf{u}^k, \quad \boldsymbol{\epsilon}_{nG}^k = (\mathbf{D}_{np} + \mathbf{D}_{nz}) \mathbf{u}^k. \quad (2)$$

The superscript k is the layer index. Strains have been split into in-plane (subscript p) and out-of-plane (subscript n, for “normal”) components:

$$\boldsymbol{\epsilon}_p^k = \{\epsilon_{xx}, \epsilon_{yy}, \epsilon_{xy}\}, \quad \boldsymbol{\epsilon}_n^k = \{\epsilon_{xz}, \epsilon_{yz}, \epsilon_{zz}\},$$

while $\mathbf{u}^k = \{u_x, u_y, u_z\}$ is the vector of the displacement components. The differential matrices are given explicitly by

$$\mathbf{D}_p = \begin{bmatrix} \partial_x & 0 & 0 \\ 0 & \partial_y & 0 \\ \partial_y & \partial_x & 0 \end{bmatrix}, \quad \mathbf{D}_{np} = \begin{bmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_y \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{D}_{nz} = \begin{bmatrix} \partial_z & 0 & 0 \\ 0 & \partial_z & 0 \\ 0 & 0 & \partial_z \end{bmatrix}. \quad (3)$$

The electric field \mathcal{E} is related to the electric potential by the gradient relation

$$\mathcal{E}^{kT} = [-\partial_x \quad -\partial_y \quad -\partial_z] \Phi^k. \quad (4)$$

The electric potential Φ being a scalar, one obtains by separating in-plane and normal components the equality

$$\mathcal{E}^k = (\mathbf{D}_{ep} + \mathbf{D}_{ez}) \Phi^k, \quad (5)$$

where

$$\mathbf{D}_{ep}^T = [-\partial_x \quad -\partial_y \quad 0], \quad \mathbf{D}_{ez}^T = [0 \quad 0 \quad -\partial_z]. \quad (6)$$

3. Variational statements for piezoelectric continua

The classical variational tool most often used to develop FEs, is the principle of virtual displacements (PVD), which, for a piezoelectric continuum, can be written

$$\sum_{k=1}^{N_l} \int_{\Omega_k} \int_{A_k} (\delta \boldsymbol{\epsilon}_{pG}^{kT} \boldsymbol{\sigma}_{pC}^k + \delta \boldsymbol{\epsilon}_{nG}^{kT} \boldsymbol{\sigma}_{nC}^k - \delta \mathcal{E}_G^{kT} \mathcal{D}_C^k) d\Omega_k dz = \delta L_e. \quad (7)$$

Here δ denotes virtual variations, A_k is the layer domain in the thickness direction, Ω_k denotes the reference surface of the layer, and δL_e denotes the virtual variation of the work made by applied loadings. The in-plane and out-of-plane stress components are

$$\boldsymbol{\sigma}_p^T = \{\sigma_{xx}, \sigma_{yy}, \sigma_{xy}\}, \quad \boldsymbol{\sigma}_n^T = \{\sigma_{xz}, \sigma_{yz}, \sigma_{zz}\}.$$

The electrical work is obtained via the electrical displacement vector:

$$\mathcal{D} = \{\mathcal{D}_x, \mathcal{D}_y, \mathcal{D}_z\}.$$

A subscript C will denote stress and electrical displacements from the constitutive law, and a subscript G strains and electrical fields from the geometrical relation. The PVD allows one to assume two independent fields for \mathbf{u} and Φ . The remaining variables are obtained from the constitutive law of the piezoelectric layers.

The RMVT was proposed in [Reissner 1984] for purely mechanical problems. A critical review on its use was given in [Carrera 2001]. A main feature of the RMVT is that it allows one to assume two independent fields for displacements \mathbf{u} and transverse stresses σ_n . This allows the *a priori* fulfillment of the necessary continuity (equilibrium) conditions of transverse normal and shear stresses at each layer interfaces. In the static case, for pure mechanical problems RMVT states that

$$\sum_{k=1}^{N_l} \int_{\Omega_k} \int_{A_k} (\delta \epsilon_{pG}^{kT} \sigma_{pC}^k + \delta \epsilon_{nG}^{kT} \sigma_{nM}^k - \delta \sigma_{nM}^{kT} (\epsilon_{nG}^k - \epsilon_{nC}^k)) d\Omega_k dz = \delta L_e. \quad (8)$$

The second term in the integrand forces the compatibility of transverse strain obtained by the material's constitutive law (which are different from those related to PVD applications) and by the geometric relation. The subscript M denotes those variables which are assumed in a given model.

By introducing the electrical work, we can write the RMVT for piezoelectric continua as

$$\sum_{k=1}^{N_l} \int_{\Omega_k} \int_{A_k} (\delta \epsilon_{pG}^{kT} \sigma_{pC}^k + \delta \epsilon_{nG}^{kT} \sigma_{nM}^k - \delta \mathcal{E}_G^{kT} \mathcal{D}_C^k - \delta \sigma_{nM}^{kT} (\epsilon_{nG}^k - \epsilon_{nC}^k)) d\Omega_k dz = \delta L_e. \quad (9)$$

This form of the RMVT will be called the *partial* extension of RMVT to piezoelectric continua, or P-RMVT.

A full extension of the RMVT can be obtained by introducing the transverse components of electric displacement \mathcal{D}_z as additional variables. The RMVT then assumes the following *full* form, or F-RMVT:

$$\sum_{k=1}^{N_l} \int_{A_k} \int_{h_k} (\delta \epsilon_{pG}^{kT} \sigma_{pC}^k + \delta \epsilon_{nG}^{kT} \sigma_{nM}^k - \delta \mathcal{E}_{pG}^{kT} \mathcal{D}_{pC}^k - \delta \mathcal{E}_{nG}^{kT} \mathcal{D}_{nM}^k + \delta \sigma_{nM}^{kT} (\epsilon_{nG}^k - \epsilon_{nC}^k) - \mathcal{D}_{nM}^T (\mathcal{E}_{nG}^k - \mathcal{E}_{nC}^k)) dA_k dz = \delta L_e. \quad (10)$$

The electrical displacement and electrical field vectors have been split into in-plane and normal components (as for the stresses σ and strains ϵ):

$$\mathcal{D}_p = \{\mathcal{D}_x, \mathcal{D}_y\}, \quad \mathcal{D}_n = \{\mathcal{D}_z\}, \quad \mathcal{E}_p = \{E_x, E_y\}, \quad \mathcal{E}_n = \{E_z\}.$$

The constitutive equations of the k -layer are conveniently written as

$$\begin{aligned} \sigma_{pC}^k &= \mathbf{C}_{pp}^k \epsilon_{pG}^k + \mathbf{C}_{pn}^k \epsilon_{nC}^k - \mathbf{e}_{pp}^{kT} \mathcal{E}_{pG}^k - \mathbf{e}_{np}^{kT} \mathcal{E}_{nC}^k, & \mathcal{D}_{pC}^k &= \mathbf{e}_{pp}^k \epsilon_{pG}^k + \mathbf{e}_{pn}^k \epsilon_{nC}^k + \mathbf{e}_{pp}^{kT} \mathcal{E}_{pG}^k + \mathbf{e}_{pn}^{kT} \mathcal{E}_{nC}^k, \\ \sigma_{nM}^k &= \mathbf{C}_{pn}^{kT} \epsilon_{pG}^k + \mathbf{C}_{nn}^k \epsilon_{nC}^k - \mathbf{e}_{pn}^{kT} \mathcal{E}_{pG}^k - \mathbf{e}_{nn}^{kT} \mathcal{E}_{nC}^k, & \mathcal{D}_{nC}^k &= \mathbf{e}_{np}^k \epsilon_{pG}^k + \mathbf{e}_{nn}^k \epsilon_{nC}^k + \mathbf{e}_{pn}^{kT} \mathcal{E}_{pG}^k + \mathbf{e}_{nn}^{kT} \mathcal{E}_{nC}^k, \end{aligned} \quad (11)$$

where we have introduced the following arrays:

- Stiffness matrices:

$$\mathbf{C}_{pp}^k = \begin{bmatrix} C_{11} & C_{12} & C_{16} \\ C_{12} & C_{22} & C_{26} \\ C_{16} & C_{26} & C_{66} \end{bmatrix}^k, \quad \mathbf{C}_{pn}^k = \begin{bmatrix} 0 & 0 & C_{13} \\ 0 & 0 & C_{23} \\ 0 & 0 & C_{36} \end{bmatrix}^k, \quad \mathbf{C}_{nn}^k = \begin{bmatrix} C_{55} & C_{45} & 0 \\ C_{45} & C_{44} & 0 \\ 0 & 0 & C_{33} \end{bmatrix}^k. \quad (12)$$

- Piezoelectric matrices:

$$\mathbf{e}_{pp}^k = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}^k, \quad \mathbf{e}_{pn}^k = \begin{bmatrix} e_{15} & e_{14} & 0 \\ e_{25} & e_{24} & 0 \end{bmatrix}^k, \quad \mathbf{e}_{np}^k = [e_{31} \ e_{32} \ e_{33}]^k, \quad \mathbf{e}_{nn}^k = [0 \ 0 \ e_{33}]^k. \quad (13)$$

- Permittivity matrices:

$$\boldsymbol{\epsilon}_{pp}^k = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} \\ \epsilon_{12} & \epsilon_{22} \end{bmatrix}^k, \quad \boldsymbol{\epsilon}_{pn}^k = \begin{bmatrix} 0 \\ 0 \end{bmatrix}^k, \quad \boldsymbol{\epsilon}_{nn}^k = [\epsilon_{33}]^k. \quad (14)$$

Application of the F-RMVT requires one to express the in-plane stresses $\boldsymbol{\sigma}_{pC}$, the normal strains $\boldsymbol{\epsilon}_{nC}$, the normal electric field \mathcal{E}_{nC} and the in-plane electric displacements \mathcal{D}_{pC} in terms of the remaining variables. Thus the constitutive equations (11) can be solved as follows:

$$\begin{aligned} \boldsymbol{\sigma}_{pC} &= \hat{\mathbf{C}}_{spm} \boldsymbol{\epsilon}_{pG} + \hat{\mathbf{C}}_{snm} \boldsymbol{\sigma}_{nM} + \hat{\mathbf{C}}_{spe} \mathcal{E}_{pG} + \hat{\mathbf{C}}_{sne} \mathcal{D}_{nM}, \\ \boldsymbol{\epsilon}_{nC} &= \hat{\mathbf{C}}_{dpm} \boldsymbol{\epsilon}_{pG} + \hat{\mathbf{C}}_{dnm} \boldsymbol{\sigma}_{nM} + \hat{\mathbf{C}}_{dpe} \mathcal{E}_{pG} + \hat{\mathbf{C}}_{dne} \mathcal{D}_{nM}, \\ \mathcal{D}_{pC} &= \hat{\mathbf{C}}_{fpm} \boldsymbol{\epsilon}_{pG} + \hat{\mathbf{C}}_{fnm} \boldsymbol{\sigma}_{nM} + \hat{\mathbf{C}}_{fpe} \mathcal{E}_{pG} + \hat{\mathbf{C}}_{fne} \mathcal{D}_{nM}, \\ \mathbf{E}_{nC} &= \hat{\mathbf{C}}_{epm} \boldsymbol{\epsilon}_{pG} + \hat{\mathbf{C}}_{enm} \boldsymbol{\sigma}_{nM} + \hat{\mathbf{C}}_{epe} \mathcal{E}_{pG} + \hat{\mathbf{C}}_{ene} \mathcal{D}_{nM}. \end{aligned} \quad (15)$$

The matrices above are obtained from by those in (11) by means of the relations

$$\begin{aligned} \bar{\mathbf{C}}_{dpm}^k &= -\mathbf{C}_{nn}^{k-1} \mathbf{C}_{pn}^{kT} - (\mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT}) (\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT} + \boldsymbol{\epsilon}_{nn}^k)^{-1} (\mathbf{e}_{np}^k - \mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{C}_{pn}^{kT}), \\ \bar{\mathbf{C}}_{dnm}^k &= \mathbf{C}_{nn}^{k-1} - (\mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT}) (\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT} + \boldsymbol{\epsilon}_{nn}^k)^{-1} (\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1}), \\ \bar{\mathbf{C}}_{dpe}^k &= \mathbf{C}_{nn}^{k-1} \mathbf{e}_{pn}^{kT} - (\mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT}) (\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT} + \boldsymbol{\epsilon}_{nn}^k)^{-1} (\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{e}_{pn}^{kT} + \boldsymbol{\epsilon}_{pn}^k), \\ \bar{\mathbf{C}}_{dne}^k &= (\mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT}) (\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT} + \boldsymbol{\epsilon}_{nn}^k)^{-1}, \\ \bar{\mathbf{C}}_{epm}^k &= -(\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT} + \boldsymbol{\epsilon}_{nn}^k)^{-1} (\mathbf{e}_{np}^k - \mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{C}_{pn}^{kT}), \\ \bar{\mathbf{C}}_{enm}^k &= -(\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT} + \boldsymbol{\epsilon}_{nn}^k)^{-1} (\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1}), \\ \bar{\mathbf{C}}_{epe}^k &= -(\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT} + \boldsymbol{\epsilon}_{nn}^k)^{-1} (\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{e}_{pn}^{kT} + \boldsymbol{\epsilon}_{pn}^k), \\ \bar{\mathbf{C}}_{ene}^k &= (\mathbf{e}_{nn}^k \mathbf{C}_{nn}^{k-1} \mathbf{e}_{nn}^{kT} + \boldsymbol{\epsilon}_{nn}^k)^{-1}, \\ \bar{\mathbf{C}}_{spm}^k &= \mathbf{C}_{pp}^k + \mathbf{C}_{pn}^k \bar{\mathbf{C}}_{dpm}^k - \mathbf{e}_{np}^{kT} \bar{\mathbf{C}}_{epm}^k, & \bar{\mathbf{C}}_{snm}^k &= \mathbf{C}_{pn}^k \bar{\mathbf{C}}_{dnm}^k - \mathbf{e}_{np}^{kT} \bar{\mathbf{C}}_{enm}^k, \\ \bar{\mathbf{C}}_{spe}^k &= \mathbf{C}_{pn}^k \bar{\mathbf{C}}_{dpe}^k - \mathbf{e}_{pp}^{kT} - \mathbf{e}_{np}^{kT} \bar{\mathbf{C}}_{epe}^k, & \bar{\mathbf{C}}_{sne}^k &= \mathbf{C}_{pn}^k \bar{\mathbf{C}}_{dne}^k - \mathbf{e}_{np}^{kT} \bar{\mathbf{C}}_{ene}^k, \\ \bar{\mathbf{C}}_{fpm}^k &= \mathbf{e}_{pp}^k + \mathbf{e}_{pn}^k \bar{\mathbf{C}}_{dpm}^k + \boldsymbol{\epsilon}_{np}^k \bar{\mathbf{C}}_{epm}^k, & \bar{\mathbf{C}}_{fnm}^k &= \mathbf{e}_{pn}^k \bar{\mathbf{C}}_{dnm}^k + \boldsymbol{\epsilon}_{pn}^k \bar{\mathbf{C}}_{enm}^k, \\ \bar{\mathbf{C}}_{fpe}^k &= \mathbf{e}_{pn}^k \bar{\mathbf{C}}_{dpe}^k + \boldsymbol{\epsilon}_{pp}^k + \boldsymbol{\epsilon}_{pn}^k \bar{\mathbf{C}}_{epe}^k, & \bar{\mathbf{C}}_{fne}^k &= \mathbf{e}_{pn}^k \bar{\mathbf{C}}_{dne}^k + \boldsymbol{\epsilon}_{pn}^k \bar{\mathbf{C}}_{ene}^k. \end{aligned}$$

It must be noted that

$$\begin{aligned} \hat{\mathbf{C}}_{dnm}^k &= \hat{\mathbf{C}}_{dnm}^{kT}, & \hat{\mathbf{C}}_{spm}^k &= \hat{\mathbf{C}}_{spm}^{kT}, & \hat{\mathbf{C}}_{fpe}^k &= \hat{\mathbf{C}}_{fpe}^{kT}, & \hat{\mathbf{C}}_{ene}^k &= \hat{\mathbf{C}}_{ene}^{kT}, \\ \hat{\mathbf{C}}_{dpm}^k &= -\hat{\mathbf{C}}_{snm}^{kT}, & \hat{\mathbf{C}}_{dne}^k &= \hat{\mathbf{C}}_{enm}^{kT}, & \hat{\mathbf{C}}_{dpe}^k &= \hat{\mathbf{C}}_{fnm}^{kT}, & \hat{\mathbf{C}}_{epm}^k &= \hat{\mathbf{C}}_{sne}^{kT}. \end{aligned}$$

4. Unified formulation for plate elements

The unified formulation is a technique that allows one to handle in a unified manner a large variety of plate modelings and finite elements. In this formulation, the finite element matrices are written in terms of a few *fundamental nuclei*, which do not formally depend on: the expansion N used in the z -direction, the number of the node N_n of the element, or the variables description (LW or ESL).

The unknown variables \mathbf{u} , $\boldsymbol{\sigma}_n$, Φ and \mathcal{D}_z are expressed in terms of the layer thickness coordinate:

$$\begin{aligned} (\mathbf{u}^k(x, y, z), \phi^k(x, y, z), \boldsymbol{\sigma}_n^k(x, y, z), \mathcal{D}_n^k(x, y, z)) &= F_b(z) (\mathbf{u}_b^k(x, y), \phi_b^k(x, y), \boldsymbol{\sigma}_{nb}^k(x, y), \mathcal{D}_{nb}^k(x, y)) \\ &+ F_r(z) (\mathbf{u}_r^k(x, y), \phi_r^k(x, y), \boldsymbol{\sigma}_{nr}^k(x, y), \mathcal{D}_{nr}^k(x, y)) + F_t(z) (\mathbf{u}_t^k(x, y), \phi_t^k(x, y), \boldsymbol{\sigma}_{nt}^k(x, y), \mathcal{D}_{nt}^k(x, y)). \end{aligned} \quad (16)$$

The subscript t and b denote the linear part of the thickness expansion (t and b will be used to denote top- and bottom-layer variable values in layerwise cases), while subscript r refers to higher-order terms: $r = 2, \dots, N-1$. In compact form,

$$(\mathbf{u}^k(x, y, z), \phi^k(x, y, z), \boldsymbol{\sigma}_n^k(x, y, z), \mathcal{D}_n^k(x, y, z)) = F_\tau(z) (\mathbf{u}^k(x, y), \phi^k(x, y), \boldsymbol{\sigma}_n^k(x, y), \mathcal{D}_n^k(x, y))_\tau. \quad (17)$$

Here $(\mathbf{u}^k(x, y), \phi^k(x, y), \boldsymbol{\sigma}_n^k(x, y), \mathcal{D}_n^k(x, y))_\tau$ are two-dimensional unknowns, the $F_\tau(z)$ are the base functions of the expansion, and the summation convention over repeated index has been adopted. The base functions could be, in general, different for each variable. Different choices for $F_\tau(z)$ will lead to different plate/shell theories. The choices made in our study are briefly discussed below; detailed descriptions can be found in the works cited.

. Layer-wise elements.

The thickness functions are given by combinations of Legendre polynomials P_j as

$$F_t = \frac{P_0(\zeta_k) + P_1(\zeta_k)}{2}, \quad F_b = \frac{P_0(\zeta_k) - P_1(\zeta_k)}{2}, \quad F_r = P_r(\zeta_k) - P_{r-2}(\zeta_k), \quad r = 2, 3, \dots, N, \quad (18)$$

for $\zeta = z_k/2h_k$, where z_k is the local layer thickness coordinate and h_k is the layer thickness, so $-1 \leq \zeta_k \leq 1$. As mentioned, t and b denote top and bottom; that is, the chosen functions have the properties

$$\zeta_k = \begin{cases} 1 & : F_t = 1, F_b = 0, F_r = 0, \\ -1 & : F_t = 0, F_b = 1, F_r = 0, \end{cases} \quad (19)$$

Thanks to these properties the interlaminar continuity of the assumed variables can be easily linked in the assembly procedure from layer-level matrices to multilayer-level matrices.

The resulting elements will be denoted by the acronyms LFM1 to LFM4, in which L means layerwise, FM states that F-RMVT has been employed, and the digit is the order of the expansion. Particular cases of P-RMVT and PVD will also be used in the numerical analysis; these applications will be denoted by LPM1 to LPM4 and LD1 to LD4, respectively.

Equivalent single-layer model. In this case the layerwise expansion is preserved for the transverse stresses, electric potential and electric displacements, while a Taylor-type expansion is used for the displacement components:

$$\mathbf{u}(x, y, z) = \mathbf{u}_\tau(x, y) z^\tau, \quad \tau = 0, N$$

The base functions related to displacements can be chosen as

$$F_b(z) = 1, \quad F_r(z) = z^r, \quad r = 1, N-1, \quad F_t(z) = z^N.$$

These theories will be denoted with the acronyms EFMC1 to EFMC3, in which E means equivalent single-layer, FM means full mixed, and C that interlaminar continuity conditions are fulfilled for transverse stresses, electric potential and transverse electric displacement. The digit, as before, denotes the expansion order. Results related to P-RMVT application will be denoted by EPMC1 to EPMC3. When the Murakami zigzag function is used (which allows the introduction of piecewise continuous displacement fields in the thickness plate direction; see [Carrera 2001]), the resulting elements are referred to as EFMZC1 to EFMZC3 and EPMZC1 to EPMZC3 for the full and partial cases.

Finite element approximations. Finite element approximations to the plate reference surface domain are introduced by means of isoparametric descriptions for the various field variables:

$$\left(\mathbf{u}_\tau^k, \Phi_\tau^k, \boldsymbol{\sigma}_{n\tau}^k, \mathcal{D}_{n\tau}^k \right) (x, y) = N_i(x, y) \left(\mathbf{q}_{\tau i}^k, g_{\tau i}^k, \mathbf{f}_{\tau i}^k, d_{\tau i}^k \right), \quad i = 1, 2, \dots, N_n, \quad (20)$$

where the $N_i(x, y)$ are the shape functions, $\mathbf{q}_{\tau i}^k$ the nodal unknown displacements, $g_{\tau i}^k$ the nodal unknown electric potentials, $\mathbf{f}_{\tau i}^k$ the nodal unknowns normal stresses and $d_{\tau i}^k$ the nodal unknown normal electrical displacements. The cases of 9, 8 and 4 nodes are considered in the numerical implementation referred to as Q9, Q8 and Q4 finite elements [Carrera and DeMasi 2002b].

5. Derivation of finite element matrices

This section is devoted to the fundamental nuclei of the F-RMVT finite element matrices. The RMVT and PVD matrices can be found in [Carrera and DeMasi 2002a; Carrera and Boscolo 2006].

By starting from Equation (10), the fundamental nuclei are derived in several steps:

1. The constitutive relations (15) are introduced in the F-RMVT statement at (10).
2. The geometric relations are used to express strain in terms of displacements and electric field in terms of electric potential.
3. The through-the-thickness assumptions by means of the Unified Formulation are introduced.
4. The FE shape functions are used to eliminate the in-plane plate coordinates by numerical integration.
5. Matrix products are made, yielding the explicit forms of the fundamental nuclei.

We omit the details for brevity. The final form of the governing equations is

$$\begin{aligned}
 \delta \mathbf{q}_{\tau i}^k{}^T : \quad & \mathbf{K}_{uu}^{k\tau sij} \mathbf{q}_{sj}^k + \mathbf{K}_{u\sigma}^{k\tau sij} \mathbf{f}_{sj}^k + \mathbf{K}_{ue}^{k\tau sij} \mathbf{g}_{sj}^k + \mathbf{K}_{ud}^{k\tau sij} \mathbf{d}_{sj}^k = \mathbf{P}_{u\tau}^k, \\
 \delta \mathbf{f}_{\tau i}^{kT} : \quad & \mathbf{K}_{\sigma u}^{k\tau sij} \mathbf{q}_{sj}^k + \mathbf{K}_{\sigma\sigma}^{k\tau sij} \mathbf{f}_{sj}^k + \mathbf{K}_{\sigma e}^{k\tau sij} \mathbf{g}_{sj}^k + \mathbf{K}_{\sigma d}^{k\tau sij} \mathbf{d}_{sj}^k = 0, \\
 \delta \mathbf{g}_{\tau i}^k{}^T : \quad & \mathbf{K}_{eu}^{k\tau sij} \mathbf{q}_{sj}^k + \mathbf{K}_{e\sigma}^{k\tau sij} \mathbf{f}_{sj}^k + \mathbf{K}_{ee}^{k\tau sij} \mathbf{g}_{sj}^k + \mathbf{K}_{ed}^{k\tau sij} \mathbf{d}_{sj}^k = \mathbf{P}_{e\tau}^k, \\
 \delta \mathbf{d}_{\tau i}^k{}^T : \quad & \mathbf{K}_{du}^{k\tau sij} \mathbf{q}_{sj}^k + \mathbf{K}_{d\sigma}^{k\tau sij} \mathbf{f}_{sj}^k + \mathbf{K}_{de}^{k\tau sij} \mathbf{g}_{sj}^k + \mathbf{K}_{dd}^{k\tau sij} \mathbf{d}_{sj}^k = 0.
 \end{aligned} \tag{21}$$

The mechanical and electrical loading terms on the right-hand side are

$$\mathbf{P}_{u\tau}^k = \mathbf{K}_{up}^{k\tau sij} \mathbf{p}_{sj}, \quad \mathbf{P}_{e\tau}^k = -\mathbf{K}_{ef}^{k\tau sij} \Psi_{sj}. \tag{22}$$

The explicit forms of the fundamental nuclei thus obtained are

$$\begin{aligned}
 \mathbf{K}_{uu}^{k\tau sij} &= \int_{\Omega} (\mathbf{D}_p^T N_i \hat{\mathbf{C}}_{spm}^k E_{\tau s}^{uu} \mathbf{D}_p N_j) d\Omega, \\
 \mathbf{K}_{u\sigma}^{k\tau sij} &= \int_{\Omega} (\mathbf{D}_p^T N_i \hat{\mathbf{C}}_{snm}^k E_{\tau s}^{u\sigma} N_j + N_i \mathbf{D}_{np}^T E_{\tau s}^{u\sigma} N_j + N_i \mathbf{I}^T E_{\tau,zs}^{u\sigma} N_j) d\Omega, \\
 \mathbf{K}_{ue}^{k\tau sij} &= \int_{\Omega} (\mathbf{D}_p^T N_i \hat{\mathbf{C}}_{spe}^k E_{\tau s}^{ue} \mathbf{D}_{ep} N_j) d\Omega, \\
 \mathbf{K}_{ud}^{k\tau sij} &= \int_{\Omega} (\mathbf{D}_p^T N_i \hat{\mathbf{C}}_{sne}^k E_{\tau s}^{ue} N_j) d\Omega, \\
 \mathbf{K}_{\sigma u}^{k\tau sij} &= \int_{\Omega} (N_i E_{\tau s}^{\sigma u} \mathbf{D}_{np} N_j + N_i E_{\tau s,z}^{\sigma u} \mathbf{I} N_j - N_i \hat{\mathbf{C}}_{dpm}^k E_{\tau s}^{\sigma u} \mathbf{D}_p N_j) d\Omega, \\
 \mathbf{K}_{\sigma\sigma}^{k\tau sij} &= - \int_{\Omega} (N_i \hat{\mathbf{C}}_{dnm}^k E_{\tau s}^{\sigma\sigma} N_j) d\Omega, \\
 \mathbf{K}_{\sigma e}^{k\tau sij} &= - \int_{\Omega} (N_i \hat{\mathbf{C}}_{dpe}^k E_{\tau s}^{\sigma e} \mathbf{D}_{ep} N_j) d\Omega, \\
 \mathbf{K}_{\sigma d}^{k\tau sij} &= - \int_{\Omega} (N_i \hat{\mathbf{C}}_{dne}^k E_{\tau s}^{\sigma e} N_j) d\Omega, \\
 \mathbf{K}_{eu}^{k\tau sij} &= - \int_{\Omega} (\mathbf{D}_{ep}^T N_i \hat{\mathbf{C}}_{fpm}^k E_{\tau s}^{eu} \mathbf{D}_p N_j) d\Omega, \\
 \mathbf{K}_{e\sigma}^{k\tau sij} &= - \int_{\Omega} (N_i \mathbf{D}_{ep}^T \hat{\mathbf{C}}_{fnm}^k E_{\tau s}^{e\sigma} N_j) d\Omega, \\
 \mathbf{K}_{ee}^{k\tau sij} &= - \int_{\Omega} (N_i \mathbf{D}_{ep}^T \hat{\mathbf{C}}_{fpe}^k E_{\tau s}^{ee} \mathbf{D}_{ep} N_j) d\Omega, \\
 \mathbf{K}_{ed}^{k\tau sij} &= - \int_{\Omega} (N_i \mathbf{D}_{ep}^T \hat{\mathbf{C}}_{fne}^k E_{\tau s}^{ee} \mathbf{D}_{ep} N_j + \mathbf{I}^* N_i E_{\tau,zs}^{e\sigma} N_j) d\Omega, \\
 \mathbf{K}_{du}^{k\tau sij} &= \int_{\Omega} (N_i \hat{\mathbf{C}}_{epm}^k E_{\tau s}^{du} \mathbf{D}_p N_j) d\Omega, \\
 \mathbf{K}_{d\sigma}^{k\tau sij} &= \int_{\Omega} (N_i \hat{\mathbf{C}}_{enm}^k E_{\tau s}^{d\sigma} N_j) d\Omega,
 \end{aligned}$$

$$\begin{aligned}\mathbf{K}_{de}^{k\tau sij} &= \int_{\Omega} (-N_i E_{\tau s, z}^{d\phi} \mathbf{I}^* N_j + N_i \hat{\mathbf{C}}_{epe}^k E_{\tau s}^{d\phi} \mathbf{D}_{ep} N_j) d\Omega, \\ \mathbf{K}_{dd}^{k\tau sij} &= \int_{\Omega} (N_i \hat{\mathbf{C}}_{ene}^k E_{\tau s}^{dd} N_j) d\Omega, \\ \mathbf{K}_{up}^{k\tau sij} &= \int_{\Omega} F_{\tau}^1 (N_i N_j \mathbf{m}_s^k) F_s^1 d\Omega, \\ \mathbf{K}_{ef}^{k\tau sij} &= \int_{\Omega} F_{\tau}^1 (N_i N_j n_s^k) F_s^1 d\Omega.\end{aligned}$$

\mathbf{I} is the unit matrix and $\mathbf{I}^{*T} = \{0, 0, -1\}$. The following integrals have been defined:

$$E_{\tau s}^{\alpha\beta} = \int_{A_k} F_{\tau}^{\alpha} F_s^{\beta} dz, \quad E_{\tau, z s}^{\alpha\beta} = \int_{A_k} F_{\tau, z}^{\alpha} F_s^{\beta} dz, \quad E_{\tau s, z}^{\alpha\beta} = \int_{A_k} F_{\tau}^{\alpha} F_{s, z}^{\beta} dz,$$

where α and β can assume any of the values $u, \sigma, \Phi, \mathcal{D}$ to denote thickness function used for the related variables.

Table 1 summarizes the dimensions of the nuclei. By varying the subscripts τ, s, k, i, j over their ranges one obtains the element matrices; see [Carrera 2003b].

$\mathbf{K}_{uu}^{k\tau sij}$	[3 × 3]	$\mathbf{K}_{eu}^{k\tau sij}$	[1 × 3]	$\mathbf{K}_{\sigma u}^{k\tau sij}$	[3 × 3]	$\mathbf{K}_{du}^{k\tau sij}$	[1 × 3]
$\mathbf{K}_{u\sigma}^{k\tau sij}$	[3 × 3]	$\mathbf{K}_{e\sigma}^{k\tau sij}$	[1 × 3]	$\mathbf{K}_{\sigma\sigma}^{k\tau sij}$	[3 × 3]	$\mathbf{K}_{d\sigma}^{k\tau sij}$	[1 × 3]
$\mathbf{K}_{ue}^{k\tau sij}$	[3 × 1]	$\mathbf{K}_{ee}^{k\tau sij}$	[1 × 1]	$\mathbf{K}_{\sigma e}^{k\tau sij}$	[3 × 1]	$\mathbf{K}_{de}^{k\tau sij}$	[1 × 1]
$\mathbf{K}_{ud}^{k\tau sij}$	[3 × 1]	$\mathbf{K}_{ed}^{k\tau sij}$	[1 × 1]	$\mathbf{K}_{\sigma d}^{k\tau sij}$	[3 × 1]	$\mathbf{K}_{dd}^{k\tau sij}$	[1 × 1]
$\mathbf{K}_{ef}^{k\tau sij}$	[1 × 1]	$\mathbf{K}_{up}^{k\tau sij}$	[3 × 3]			$\mathbf{M}_{\ddot{u}\ddot{u}}^{k\tau sij}$	[3 × 3]

Table 1. Dimensions of the fundamental nuclei.

6. Numerical results

This section shows the performance of the mixed FEs developed on the basis of interlaminar *a priori* continuous transversal electric displacements \mathcal{D}_z , comparing it with a mixed elements approach that does not incorporate such continuity, as with one based of PVD applications. Further comparisons are given with the results in [Garcia Lage et al. 2004b] and with three dimensional solutions in [Heyliger 1994]. To compare the analysis with closed-form exact solutions, attention has been restricted to simply supported square plates. We retain the reduced integration technique that was successfully applied in [Carrera and DeMasi 2002b]. LW as well as ESL analyses have been performed for Q4, Q8 and Q9 elements.

We consider four-layer plates, with the two inner layers consisting of cross-ply $[0^\circ/90^\circ]$ carbon fiber and the external skins made of piezoceramic material PZT-4. The material properties are shown in Table 2 on page 432. The two composite layers have thickness $h_2 = h_3 = 0.4h$ and the skins have $h_1 = h_4 = 0.1h$. The unit value is assigned to the plate thickness. A bisinusoidal distribution of transversal pressure with amplitude $\hat{p}_z = 1$ is applied to the top surface (this coincides with a sensor configuration case).

Figure 2 shows the in-plane displacement u_y distribution in the thickness direction for the selected plate elements (z is the horizontal axis). Better results are obtained for the LFM and EFMZC analyses with respect to ones based on P-RMVT. The number of elements for the plate side N_e has been placed to the right of the acronym. Layerwise analysis leads to much better results than ESL.

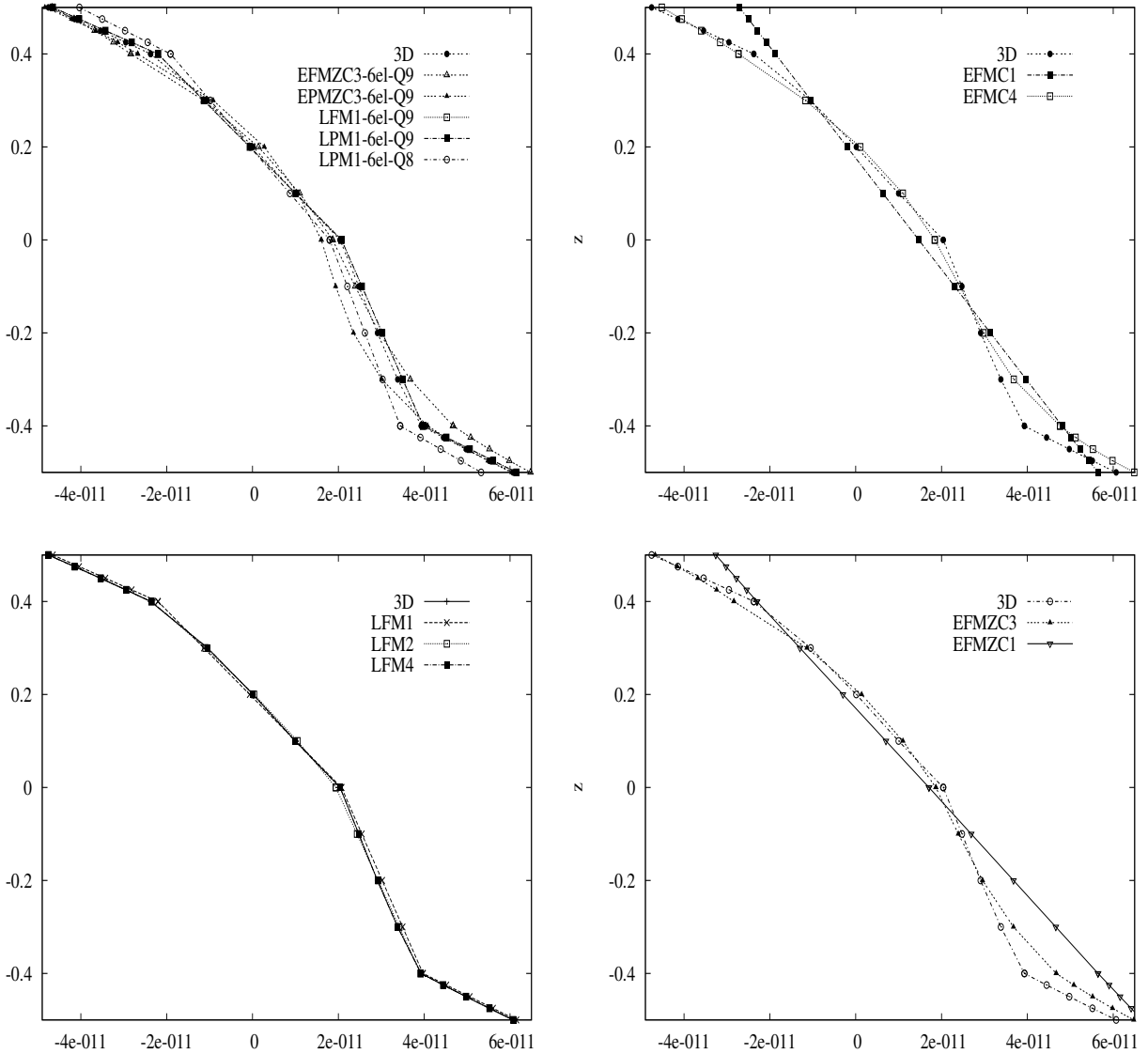


Figure 2. Performance of various FEs in predicting the displacement $u_y(a/2, 0)$ versus z . The a/h ratio equals 4. Curves labeled “3D” show the exact solution reported in [Heyliger 1994]; the remaining curves show the results obtained from FE approaches based on F-RMVT and P-RMVT (upper left), ESL theory (upper right); LW theory (lower left), and ESL theory incorporating Murakami’s zigzag function (lower right).

Property	PZT-4	Gr/Ep	PVDF	Property	PZT-4	Gr/Ep	PVDF
E_1 (GPa)	81.3	132.38	236.99	e_{15} (C/m ²)	12.72	0	-0.01
E_2 (GPa)	81.3	10.756	23.19	e_{24} (C/m ²)	12.72	0	-0.01
E_3 (GPa)	64.5	10.756	10.43	e_{31} (C/m ²)	-5.20	0	-0.13
ν_{12}	0.329	0.24	0.1541	e_{32} (C/m ²)	-5.20	0	-0.14
ν_{13}	0.432	0.24	0.1787	e_{33} (C/m ²)	15.08	0	-0.28
ν_{23}	0.432	0.49	0.1780	ϵ_{11}/ϵ_0	1475	3.5	12.50
G_{23} (GPa)	25.6	3.606	2.15	ϵ_{22}/ϵ_0	1475	3.0	11.98
G_{13} (GPa)	25.6	5.6537	4.4	ϵ_{33}/ϵ_0	1300	3.0	11.98
G_{12} (GPa)	30.6	5.6537	6.43	ρ	1	1	1

Table 2. Mechanical and electrical material properties.

The same conclusions can be drawn for the transversal normal stress evaluation in Figure 3. The use of LW elements with at least a parabolic distribution ($N = 2$) in each layer is required. Remarkable improvements are obtained when the Murakami zigzag function is used.

Data related to the transversal electrical displacement \mathcal{D}_z , shown in Figure 4, are of particular interest. Various numbers of nodes for elements and FE meshes are compared (top left). There are difficulties when certain FEs are used to predict \mathcal{D}_z in the piezoelectric layers (top right pane of figure); the results' accuracy is very much dependent on the choice of a model, and the use of elements of type LM2 (at least) appears to be necessary for correct predictions. This suggests that the use of F-RMVT may be mandatory for the accurate computation of *interlaminar continuous* \mathcal{D}_z at a reasonable computational cost, and that

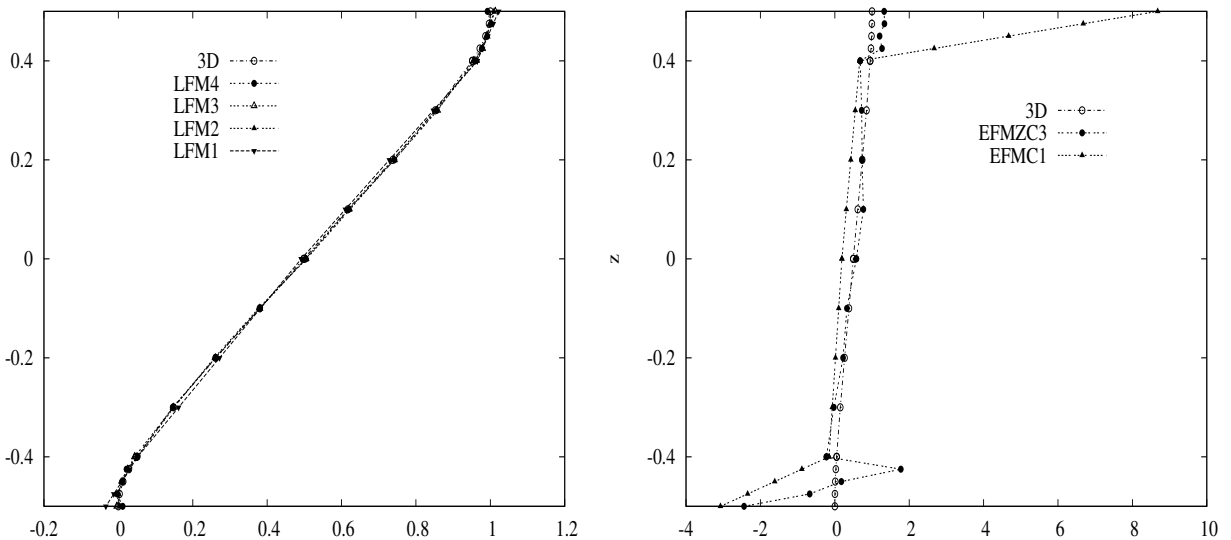


Figure 3. Performance of various FEs in predicting the transverse normal stress $\sigma_{zz}(a/2, b/2)$ versus z . The ratio a/h equals 4.

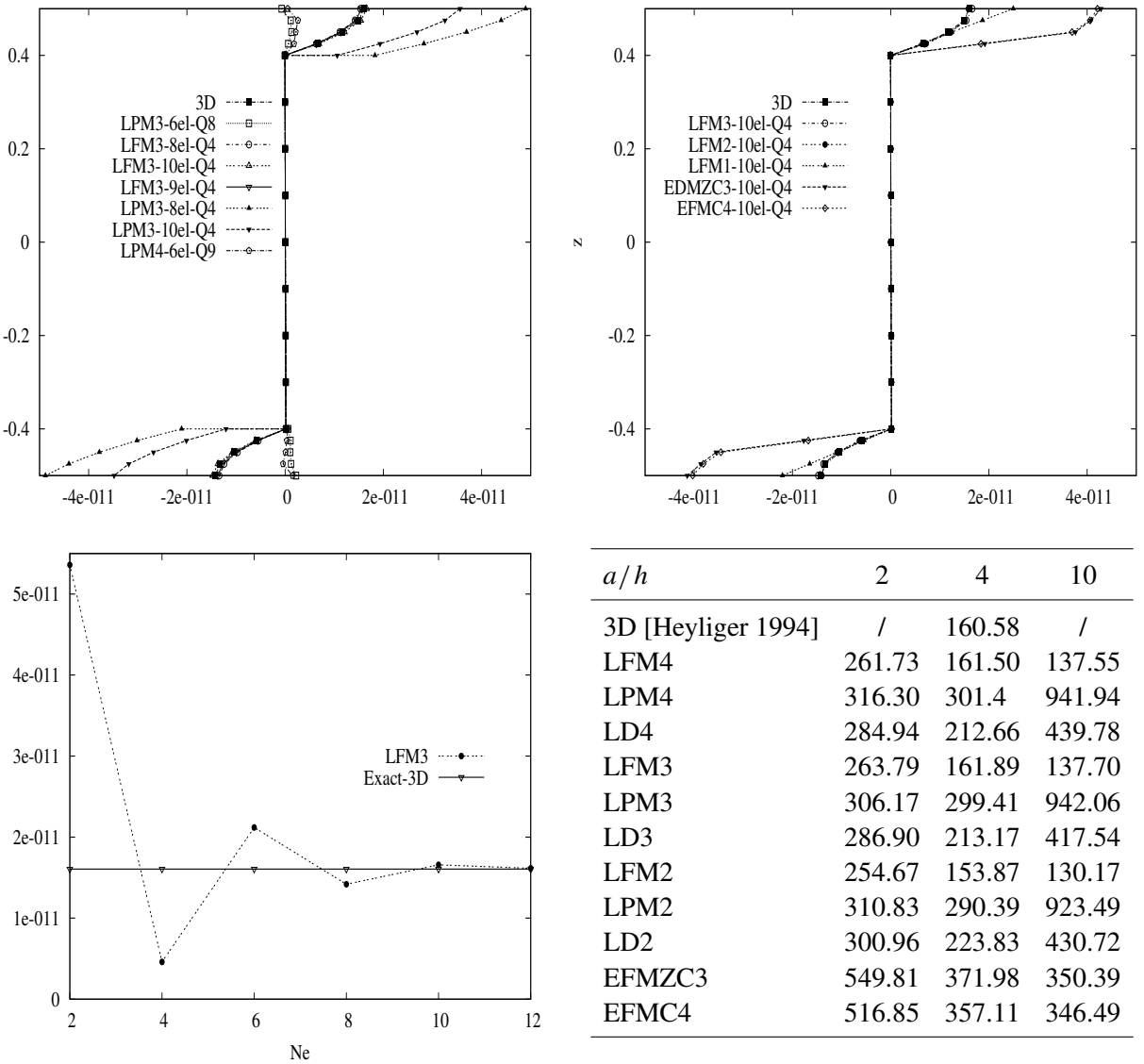


Figure 4. Top: performance of various FEs in predicting the transverse electric displacement $\mathcal{D}_z(a/2, b/2)$ versus z , with $a/h = 4$. (Left: LW elements with various number of nodes per element; right: LW and ESL elements for the Q4 case.) Bottom left: Convergence analysis for Q4 elements. Bottom right: Dependence of $\mathcal{D}_z(a/2, b/2, h) \times 10^{13}$ on the ratio a/h , for a $[12 \times 12]$ mesh and Q4 element.

P-RMVT and ESL results may be unacceptable. (Since the electric charge Q over a piezoelectric patch is obtained by integrating the \mathcal{D}_z distribution over the patch’s surface, wrong \mathcal{D}_z values lead to wrong Q values, potentially rendering the closed-loop control completely meaningless.)

Note that the accuracy obtainable with LFM2 is comparable with what we get with LPM4, confirming that the use of P-RMVT is advantageous as far as computational effort is concerned. For the sake of

completeness, Figure 4 shows the convergence rate of the Q4 elements; they are consistent with those found for of pure mechanical problems in our earlier work. Various plate thickness ratio values are considered in the table at the bottom right of Figure 4, showing the importance of UF as a tool to establish an assessment of simplified, classical and advanced FEs for piezoelectric plate analysis.

These results are confirmed in the evaluation of the electrical voltage distribution versus, shown in Figure 5. The largest discrepancies among the theories are experienced in the evaluation of electrical displacements.

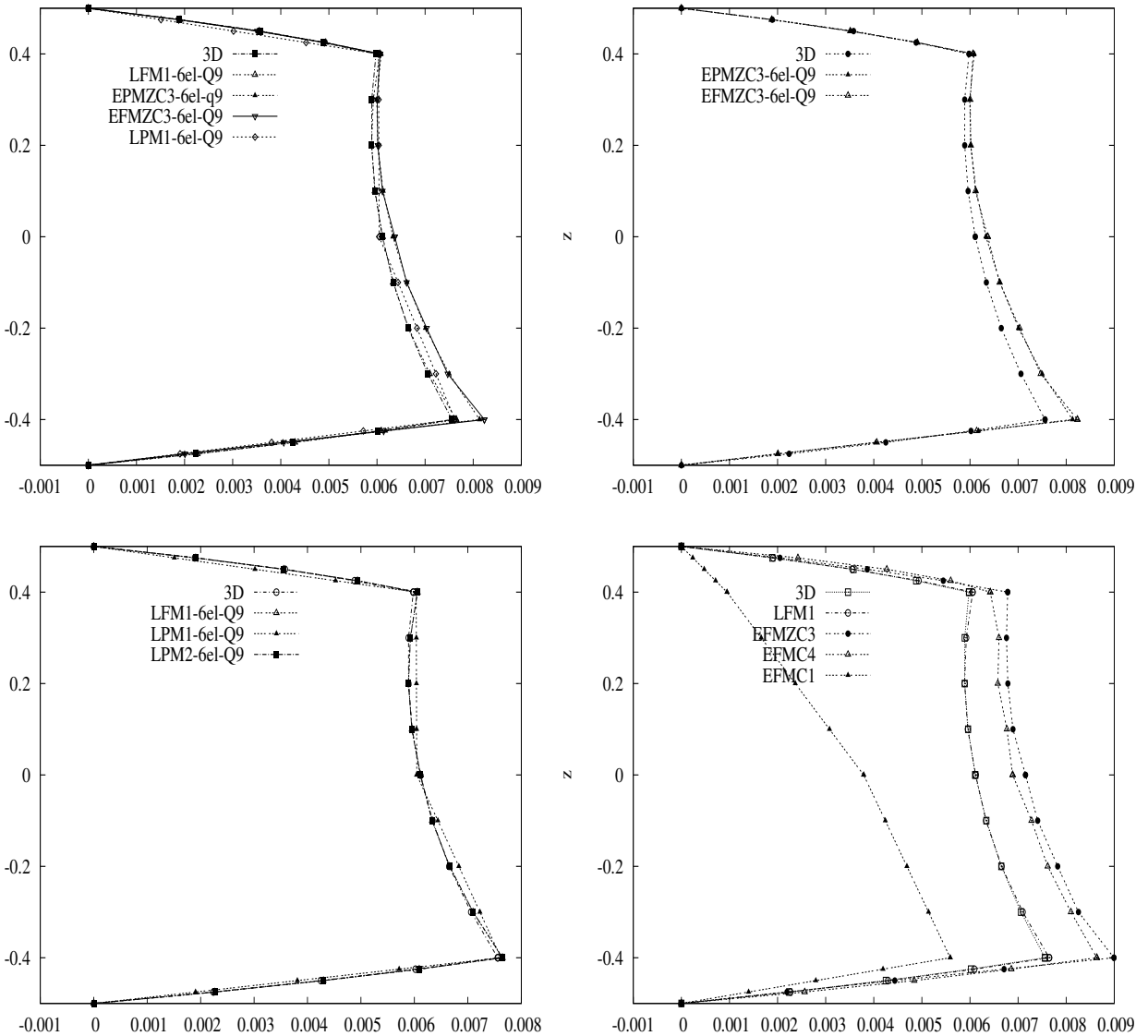


Figure 5. Top: performance of various FEs in predicting the transverse electric potential $\Phi(a/2, b/2)$ versus z , with $a/h = 4$, a $[6 \times 6]$ mesh and a Q9 element.

a/h	2	4	10	a/h	2	4	10
exact 3D	/	30,03	/	LFM1	4.751	30.13	587.4
LFM4	4.949	30.27	587.1	LPM1	4.761	30.16	587.8
LPM4	4.947	30.27	587.1	EFMZC3	4.731	31.11	623.4
LD4	4.909	30.03	582.2	EPMZC3	4.487	28.91	579.3
LFM3	4.953	30.27	587.1	EFMZC2	4.719	24.55	566.2
LPM3	4.952	30.27	587.1	EPMZC2	2.881	21.26	529.5
LD3	4.909	30.03	582.2	EFMC4	5.224	31.66	623.6
LFM2	4.928	30.23	586.9	EPMC4	5.564	28.97	579.4
LPM2	4.928	30.23	587.0	EFMC3	4.956	30.95	621.2
LD2	4.894	29.98	581.9	EPMC3	4.713	28.96	578.2

Table 3. Evaluation of $u_z(a/2, b/2, 0) \times 10^{11}$; mesh $[6 \times 6]$ and Q9 elements. The exact 3D data are taken from [Heyliger 1994].

Table 3 compares our results, for both mechanical and electrical variables, with the three-dimensional exact solution and the results of Garcia Lage et al. [2004b]. A square plate is considered with a lay-up $[0^\circ/90^\circ/0]$ for the internal layers; two piezoelectric layers of PVDF materials (see Table 2) are used as external skins. As in this last reference, the peak value of the applied pressure is 3 Pa. The relative errors are displayed in Table 4. The superiority of the full implementation of RMVT is still remarkable.

7. Concluding remarks

The paper extends the Unified Formulation and the Reissner Mixed Variational Theorem to the development of finite elements for the static analysis of piezoelectric plates with *a priori* continuous transverse electrical displacement components \mathcal{D}_z . The following main conclusions can be drawn.

- (1) It has been confirmed that UF is a valuable tool in the hierarchical analysis of piezoelectric plates using the finite element method. The implemented FEs, in fact, can provide very accurate descriptions of both mechanical and electrical fields.
- (2) FEs with interlaminar continuous \mathcal{D}_z appear to be very suitable for piezoelectric plate analysis. Better results are obtained with respect to the other FEs herein compared.
- (3) In order to preserve computational efforts, the use of the proposed elements would seem to be mandatory if accurate evaluations of \mathcal{D}_z and the related electric charge are required.

Future developments should be directed towards considering the analysis of piezoelectric plate with localized patches as sensors and/or actuators. Other plate lay-ups and the effect of additional boundary conditions and geometries should be examined. The case of imposed \mathcal{D}_z at the interface should in particular be analyzed.

References

- [Auricchio et al. 2001] F. Auricchio, P. Bisegna, and C. Lovadina, “Finite element approximation of piezoelectric plates”, *Int. J. Numer. Methods Eng.* **50**:6 (2001), 1469–1499.

	$\sigma_{zz}(a/2, b/2, 0.5)$ [Pa]	$\sigma_{yz}(a/2, 0, 0)$ [Pa]	$\sigma_{xx}(a/2, b/2, 0.5)$ [Pa]
Exact 3D	3.000	2.614	33.71
LFM4	3.0043 (0.14%)	2.6239 (0.37%)	34.54 (2.46%)
LPM4	3.0063 (0.21%)	2.6486 (1.32%)	34.54 (2.46%)
LD4	3.0081 (0.27%)	2.7302 (4.44%)	34.50 (2.34%)
Literature	3.022 (0.73%)	2.610(−0.16%)	33.08 (−1.87%)
	$\mathcal{D}_z(a/2, b/2, 0.5) \times 10^{11}$ [C/m ²]	$\Phi(a/2, b/2, 0) \times 10^3$ [V]	$\sigma_{yy}(a/2, b/2, 0.166^+)$ [Pa]
Exact 3D	−4.970	1.28	−19.63
LFM4	−5.0234 (1.07%)	1.2835 (0.27%)	−19.41 (−1.12%)
LPM4	−5.0456 (1.52%)	1.2834 (0.26%)	−19.39 (−1.22%)
LD4	−5.0509 (1.62%)	1.2837 (0.28%)	−19.41 (−1.12%)
Literature	−4.998 (0.56%)	1.279(−0.07%)	−19.46 (−0.87%)
	$u_z(a/2, b/2, 0.5) \times 10^{11}$ [m]	$u_x(0, b/2, 0.5) \times 10^{12}$ [m]	$u_y(a/2, 0, 0.5) \times 10^{12}$ [m]
Exact 3D	1.529	−1.719	−3.507
LFM4	1.5394 (0.68%)	−1.719 (0.0%)	−3.5045 (−0.07%)
LPM4	1.5395 (0.68%)	−1.7203 (0.075%)	−3.5053 (−0.048%)
LD4	1.5276 (−0.091%)	−1.7224 (0.075%)	−3.5046 (−0.068%)
Literature	1.532 (0.19%)	−1.675 (−2.56%)	−3.475 (−0.90%)

Table 4. Comparison of present analysis with respect to available results: mesh [6 × 6], $a/h = 4$, Q8 element (+ means top value). The literature values are from [Garcia Lage et al. 2004b]

- [Ballhause et al. 2005] D. Ballhause, M. D’Ottavio, B. Kröplin, and E. Carrera, “A unified formulation to assess multilayered theories for piezoelectric plates”, *Comput. Struct.* **83**:15-16 (2005), 1217–1235.
- [Batra and Vidoli 2002] R. C. Batra and S. Vidoli, “Higher-order piezoelectric plate theory derived from a three-dimensional variational principle”, *AIAA J.* **40**:1 (2002), 91–104.
- [Benjeddou 2000] A. Benjeddou, “Advances in piezoelectric finite element modeling of adaptive structural elements: a survey”, *Comput. Struct.* **76**:1-3 (2000), 347–363.
- [Carrera 1995] E. Carrera, “A class of two dimensional theories for multilayered plates analysis”, *Mem. Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur.* **19-20** (1995), 49–87.
- [Carrera 1996] E. Carrera, “ C^0 Reissner–Mindlin multilayered plate elements including zig-zag and interlaminar stresses continuity”, *Int. J. Numer. Methods Eng.* **39**:11 (1996), 1797–1820.
- [Carrera 1997] E. Carrera, “An improved Reissner–Mindlin type element for the electromechanical analysis of multilayered plates including piezo layers”, *J. Intell. Mater. Syst. Struct.* **8**:3 (1997), 232–248.
- [Carrera 1999a] E. Carrera, “A Reissner’s mixed variational theorem applied to vibrational analysis of multilayered shell”, *J. Appl. Mech. (ASME)* **66**:1 (1999), 69–78.
- [Carrera 1999b] E. Carrera, “A study of transverse normal stress effects on vibration of multilayered plates and shells”, *J. Sound Vib.* **225**:5 (1999), 803–829.
- [Carrera 2001] E. Carrera, “Developments, ideas, and evaluations based upon Reissner’s mixed variational theorem in the modeling of multilayered plates and shells”, *Appl. Mech. Rev.* **54**:4 (2001), 301–329.

- [Carrera 2003a] E. Carrera, “Historical review of zig-zag theories for multilayered plates and shell”, *Appl. Mech. Rev.* **56**:3 (2003), 287–308.
- [Carrera 2003b] E. Carrera, “Theories and finite elements for multilayered plates and shells: a unified compact formulation with numerical assessment and benchmarking”, *Arch. Comput. Methods Eng.* **10**:3 (2003), 215–296. Zbl 02097772
- [Carrera and Boscolo 2006] E. Carrera and M. Boscolo, “Classical and mixed finite elements for static and dynamics analysis of piezoelectric plates”, *Int. J. Numer. Methods Eng.* (2006). Published online 6 Nov 2006.
- [Carrera and DeMasi 2002a] E. Carrera and L. DeMasi, “Classical and advanced multilayered plate elements based upon PVD and RMVT, 1: Derivation of finite element matrices”, *Int. J. Numer. Methods Eng.* **55**:2 (2002), 191–231.
- [Carrera and DeMasi 2002b] E. Carrera and L. DeMasi, “Classical and advanced multilayered plate elements based upon PVD and RMVT, 2: Numerical implementations”, *Int. J. Numer. Methods Eng.* **55**:3 (2002), 253–291.
- [Carrera et al. 2005] E. Carrera, S. Brischetto, and M. D’Ottavio, “Vibrations of piezoelectric shells by unified formulations in the Reissner’s mixed theorem”, in *SMART ’05: Second ECCOMAS Thematic Conference on Smart Structures and Materials* (Lisbon), edited by C. A. Mota Soares et al., 2005.
- [Chopra 1996] I. Chopra, “Review of current status of smart structures and integrated systems”, pp. 20–62 in *Smart structures and materials 1996: Smart structures and integrated systems* (San Diego, CA), edited by I. Chopra, Proceedings of SPIE **2717**, SPIE, Bellingham, WA, 1996.
- [Chopra 2002] I. Chopra, “Review of state of art of smart structures and integrated systems”, *AIAA J.* **40**:11 (2002), 2145–2187.
- [Crawley and de Luis 1987] E. F. Crawley and J. de Luis, “Use of piezoelectric actuators as elements of intelligent structures”, *AIAA J.* **25**:10 (1987), 1373–1385.
- [D’Ottavio and Kröplin 2006] M. D’Ottavio and B. Kröplin, “An extension of Reissner mixed variational theorem to piezoelectric laminates”, *Mech. Adv. Mater. Struct.* **13**:2 (2006), 139–150.
- [Garcia Lage et al. 2004a] R. Garcia Lage, C. M. Mota Soares, C. A. Mota Soares, and J. N. Reddy, “Layerwise partial mixed finite element analysis of magneto-electro-elastic plates”, *Comput. Struct.* **82**:17-19 (2004), 1293–1301.
- [Garcia Lage et al. 2004b] R. Garcia Lage, C. M. Mota Soares, C. A. Mota Soares, and J. N. Reddy, “Modeling of piezo-laminated plates using layer-wise mixed finite elements”, *Comput. Struct.* **82**:23-26 (2004), 1849–1863.
- [Heyliger 1994] P. Heyliger, “Static behavior of laminated elastic/piezoelectric plates”, *AIAA J.* **32**:12 (1994), 2481–2484.
- [Heyliger and Saravanos 1995] P. Heyliger and D. A. Saravanos, “Exact free-vibration analysis of laminated plates with embedded piezoelectric layers”, *J. Acoust. Soc. Am.* **98**:3 (1995), 1547–1557.
- [Ikeda 1996] T. Ikeda, *Fundamentals of piezoelectricity*, corrected ed., Oxford University Press, New York, 1996.
- [Kögl and Bucalem 2005a] M. Kögl and M. L. Bucalem, “Analysis of smart laminates using piezoelectric MITC plate and shell elements”, *Comput. Struct.* **83**:15-16 (2005), 1153–1163.
- [Kögl and Bucalem 2005b] M. Kögl and M. L. Bucalem, “A family of piezoelectric MITC plate elements”, *Comput. Struct.* **83**:15-16 (2005), 1277–1297.
- [Lee 1990] C. K. Lee, “Theory of laminated piezoelectric plates for the design of distributed sensors/actuators, I: Governing equations and reciprocal relationships”, *J. Acoust. Soc. Am.* **87**:3 (1990), 1144–1158.
- [Mitchell and Reddy 1995] J. A. Mitchell and J. N. Reddy, “High frequency vibrations of piezoelectric crystal plates”, *Int. J. Solids Struct.* **32**:16 (1995), 2345–2367.
- [Oh and Cho 2004] J. Oh and M. Cho, “A finite element based on cubic zig-zag plate theory for the prediction of thermo-electric-mechanical behaviors”, *Int. J. Solids Struct.* **41**:5-6 (2004), 1357–1375. In press.
- [Reissner 1984] E. Reissner, “On a certain mixed variational theory and a proposed application”, *Int. J. Numer. Methods Eng.* **20**:7 (1984), 1366–1368.
- [Robbins and Reddy 1991] D. H. Robbins and J. N. Reddy, “Analysis of piezoelectric actuated beams using a layer-wise displacements theory”, *Comput. Struct.* **41**:2 (1991), 265–279.
- [Saravanos and Heyliger 1999] D. A. Saravanos and P. R. Heyliger, “Mechanics and computational models for laminated piezoelectric beams, plates, and shells”, *Appl. Mech. Rev.* **52**:10 (1999), 305–320.
- [Sheikh et al. 2001] A. H. Sheikh, P. Topdar, and S. Halder, “An appropriate fe model for through-thickness variation of displacement and potential in thin moderately thick smart laminates”, *Compos. Struct.* **51**:4 (2001), 401–409.

- [Shu 2005] X. Shu, “Free-vibration of laminated piezoelectric composite plates based on an accurate theory”, *Compos. Struct.* **67**:4 (2005), 375–382.
- [Thornbuegh and Chattopadhyay 2002] R. P. Thornbuegh and A. Chattopadhyay, “Simultaneous modeling of mechanical and electrical response of smart composite structures”, *AIAA J.* **40**:8 (2002), 1603–1610.
- [Tiersten 1969] H. F. Tiersten, *Linear piezoelectric plate vibrations*, Plenum Press, New York, 1969.
- [Vidoli and Batra 2000] S. Vidoli and R. C. Batra, “Derivation of plate and rod equations for a piezoelectric body from a mixed three-dimensional variational principle”, *J. Elasticity* **59**:1-3 (2000), 23–50.
- [Wang and Yang 2000] J. Wang and J. Yang, “High-order theories of piezoelectric plates and applications”, *Appl. Mech. Rev.* **53**:4 (2000), 87–96.
- [Wang et al. 1997] J. Wang, Y.-K. Yong, and T. Imai, “Finite element analysis of the piezoelectric vibrations of quartz plate resonators with higher-order plate theory”, pp. 650–658 in *Proceedings of the 1997 IEEE International Frequency Control Symposium* (Orlando, FL), IEEE, New York, 1997.
- [Yang and Batra 1995] J. S. Yang and R. C. Batra, “Mixed variational principles in non-linear electroelasticity”, *Int. J. Non-Linear Mech.* **30**:5 (1995), 719–725.

Received 9 Oct 2006. Accepted 29 Nov 2006.

ERASMO CARRERA: erasmo.carrera@polito.it

Dept. of Aeronautics and Aerospace Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy

CHRISTIAN FAGIANO: C.Fagiano@tudelft.nl

Dept. of Aeronautics and Aerospace Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy

Current address: Department of Mechanics, Aerospace Structures and Materials, Delft University of Technology, P.O. Box 5058, 2600 GB Delft, Netherlands

A SEMI-INFINITE HIGHER-ORDER DISPLACEMENT DISCONTINUITY METHOD AND ITS APPLICATION TO THE QUASISTATIC ANALYSIS OF RADIAL CRACKS PRODUCED BY BLASTING

HASAN HOSSEINI_NASAB AND MOHAMMAD FATEHI MARJI

We introduce a higher-order indirect boundary element method in a traction-free half-plane known as semi-infinite displacement discontinuity method. The method is modified to use the linear elastic fracture mechanics principles for radial crack analysis in brittle materials like rocks. In this numerical method there is no need to discretize the traction-free boundary of the half-plane into higher-order elements thus decreasing the number of elements without affecting the accuracy of the solution to the desired problems. The use of higher-order elements increases the accuracy so that it is possible to discretize both the boundary of the body and radial cracks by the same higher-order elements, therefore there may be no need to use the more complicated hybrid methods. A special crack tip element is added for each crack tip to increase the accuracy of displacement discontinuities near the crack ends due to their singularities. Based on the brittle behavior of most rocks, linear elastic fracture mechanics principles have been used to find the fracture mechanics parameters (mode-I and mode-II mixed mode stress intensity factors) of radial cracks occurring in common blasting operations. Arbitrary fracture criteria can be implemented in this code, but here a simple maximum tangential stress criterion is used to predict the angle of deviation (initiation) of radial cracks. Although this code is specially designed to include the traction-free half-plane problems, it is somewhat comprehensive so that any number of radial cracks with any length in the finite, infinite and semi-infinite planes can be treated easily. The validity of the method is proved by solving simple examples and some previously solved problems in the literature.

1. Introduction

In this work we formulate a new higher-order semi-infinite displacement discontinuity method and use it to analyze a number of crack problems. This method assumes linear or quadratic variation of displacement discontinuity in a semi-infinite body with a traction-free surface. It is based on the use of two or three collocation points (for linear or quadratic displacement discontinuity variation, respectively) over a two-element or three-element “patch” centered at the source element. This method is suitable for solving some fracture mechanics problems, because the special crack tip elements can easily be incorporated in this algorithm. We adopt the hybrid element formulation: higher-order elements are used for the discretization of all boundaries excluding the crack tips, and a special crack tip element is used for discretization of the crack ends.

Fracture mechanics has been suggested as a possible tool for solving a variety of rock engineering problems, such as rock cutting, hydrofracturing, explosive fracturing, rock stability, etc. In 1957, Irwin modified the basic theory of fracture of Griffith [1925] and introduced the important parameters called

Keywords: DDM, half-plane problems, higher-order elements, radial cracks, LEFM.

stress intensity factors (SIFs), to express the stress and displacement fields near the crack tip. Three SIFs, denoted by K_I , K_{II} and K_{III} , were introduced, corresponding to three basic fracture modes: opening or tensile (mode I), sliding or shearing (mode II) and tearing (mode III).

Recently linear elastic fracture mechanics (LEFM) principles have been widely used in rock fracture mechanics (RFM) [Rossmann 1983; Whittaker et al. 1992; Aliabadi 1998]. Based on LEFM principles, a superposition of the three fracture modes describes the general case of loading called mixed mode loading. For a given cracked body under a certain type of loading, the SIFs are known and the displacements and stresses near the crack tip are accordingly determined. Therefore, the problem of LEFM reduces to the determination of the crack tip SIFs. Hybrid element formulations have been used in the literature of fracture mechanics [Guo et al. 1992; Scavia 1992; Scavia 1995; Tan et al. 1996; Carpinteri and Yang 1997; Bobet 2001; Shen et al. 2004]. Because of their complexity, fracture mechanics problems are usually solved numerically by using the complicated hybrid methods [Scavia 1995; Stephansson 2002]. Due to brittle behavior of most rocks, the linear elastic fracture mechanics principles have been used to find the fracture mechanics parameters; that is, the (I and II) mixed-mode stress intensity factors (SIFs) of radial cracks occur in the common blasting operations. For the prediction of crack initiation and its angle of deviation the maximum tangential stress criterion introduced in [Erdogan and Sih 1963], is used which compares the computed SIFs with the fracture toughness (material properties like yield strength) that should be obtained experimentally [Huang and Wang 1985; Ouchterlony 1988; Stephansson et al. 2001; Backers et al. 2004; Shen et al. 2004]. A general numerical modeling for quasistatic crack analysis in semi-infinite plane is given and as a practical problem, the radial cracks around the blast holes are numerically analyzed. Any number of blast holes with any gas pressurization ratios along the emanating cracks can be studied by this model. Suitable normal gas pressurization ratios along the radial cracks are used, to solve the problem. As it was expected, the radial crack propagation takes place under tension (Mode I or opening mode of fracture), which is mainly responsible for rock blasting [Ouchterlony 1983].

2. Higher order (linear and quadratic elements) displacement discontinuity in a half-plane

General solutions and higher-order elements. Many boundary-value problems are set in traction-free half-spaces. Here we consider a two-dimensional traction-free half-plane. To implement the higher-order displacement discontinuity elements numerically, we need the analytical solution to the problem of a constant displacement discontinuity D_i integrated over a line element along the x -axis in an infinite elastic solid Crouch [1976] showed that the general solution to this problem, over a line element of length $2a$, can be expressed in terms of two harmonic functions $f(x, y)$ and $g(x, y)$ of x and y , in which the displacements are

$$\begin{aligned} u_x &= (2(1 - \nu)f_{,y} - yf_{,xx}) + (-(1 - 2\nu)g_{,x} - yg_{,xy}), \\ u_y &= ((1 - 2\nu)f_{,x} - yf_{,xy}) + (2(1 - \nu)g_{,y} - yg_{,yy}) \end{aligned} \quad (1)$$

and the stresses are

$$\begin{aligned} \sigma_{xx} &= 2G_s(2f_{,xy} + yf_{,xyy}) + 2G_s(g_{,yy} + yg_{,yyy}), \\ \sigma_{yy} &= 2G_s(-yf_{,xyy}) + 2G_s(g_{,yy} - yg_{,yyy}), \\ \sigma_{xy} &= 2G_s(2f_{,yy} + yf_{,yyy}) + 2G_s(-yg_{,xyy}), \end{aligned} \quad (2)$$

where $f_{,x} = \partial f / \partial x$ and so on.

Shou and Crouch [1995] proposed a new higher-order displacement discontinuity method for solving plane elasticity problems, exploiting the use of quadratic elements for analysis of crack problems in infinite bodies. In this paper the same kind of the higher-order displacement discontinuity elements is used for the analysis of crack problems in semi-infinite bodies, but both linear and quadratic elements are considered. The general higher-order expression of harmonic functions $f(x, y)$ and $g(x, y)$ can be rearranged as

$$f(x, y) = \frac{-1}{4\pi(1-\nu)} \sum_{j=1}^k D_x^j F_j(I_{j-1}), \quad g(x, y) = \frac{-1}{4\pi(1-\nu)} \sum_{j=1}^k D_y^j F_j(I_{j-1}). \quad (3)$$

Here k ranges over 1, 2 for linear elements and over 1, 2, 3 for quadratic elements. Similarly, the higher-order displacement discontinuity $D_i(\varepsilon)$ can be expressed as

$$D_i(\varepsilon) = \sum_{j=1}^k N_j(\varepsilon) D_i^j \quad \text{with } k \text{ as above and } i = x, y. \quad (4)$$

The displacement discontinuity using linear elements is based on analytical integration of linear collocation shape functions over collinear, straight-line displacement discontinuity elements. Figure 1a shows the linear displacement discontinuity distribution, which can be written as

$$D_i(\varepsilon) = N_1(\varepsilon) D_i^1 + N_2(\varepsilon) D_i^2, \quad i = x, y, \quad (5)$$

where D_i^1 and D_i^2 are the linear nodal displacement discontinuities and

$$N_1(\varepsilon) = -(\varepsilon - a_1)/2a_1, \quad N_2(\varepsilon) = (\varepsilon + a_1)/2a_1 \quad (6)$$

are the linear collocation shape functions using $a_1 = a_2$. A linear element has two nodes, which are at the centers of its two equal subelements (Figure 1a).

Similarly, the quadratic element displacement discontinuity is based on the analytic integration of quadratic collocation shape functions over collinear, straight-line displacement discontinuity elements. Figure 1b shows the quadratic displacement discontinuity distribution, which can be written as

$$D_i(\varepsilon) = N_1(\varepsilon) D_i^1 + N_2(\varepsilon) D_i^2 + N_3(\varepsilon) D_i^3, \quad i = x, y, \quad (7)$$

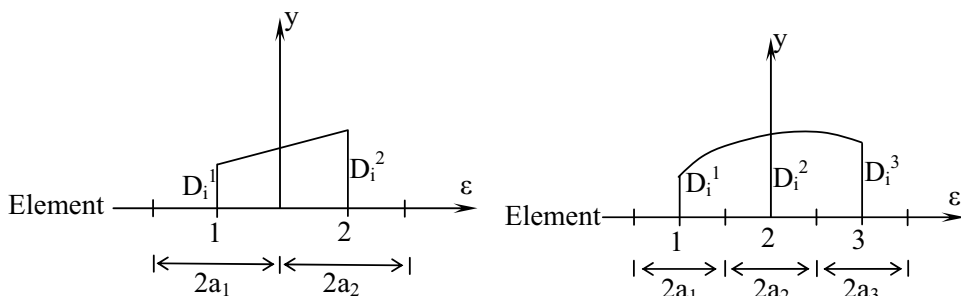


Figure 1. Linear (left) and quadratic (right) collocations for the higher-order displacement discontinuity elements.

where D_i^1 , D_i^2 , and D_i^3 are the quadratic nodal displacement discontinuities and

$$N_1(\varepsilon) = \varepsilon(\varepsilon - 2a_1)/8a_1^2, \quad N_2(\varepsilon) = -(\varepsilon^2 - 4a_1^2)/4a_1^2, \quad N_3(\varepsilon) = \varepsilon(\varepsilon + 2a_1)/8a_1^2 \quad (8)$$

are the quadratic collocation shape functions using $a_1 = a_2 = a_3$. A quadratic element has three nodes, at the centers of its three equal subelements (Figure 1b).

Considering a linear variation for $D_i(\varepsilon)$ as given in (5), the common function F_j in (3) is defined as

$$F_j(I_0, I_1) = \int N_j(\varepsilon) \ln((x - \varepsilon) + y^2)^{1/2} d\varepsilon, \quad j = 1, 2. \quad (9)$$

Here the integrals I_0 and I_1 are expressed as

$$I_0(x, y) = \int_{-a}^a \ln((x - \varepsilon)^2 + y^2)^{1/2} d\varepsilon = y(\theta_1 - \theta_2) - (x - a) \ln(r_1) + (x + a) \ln r_2 - 2a \quad (10)$$

$$I_1(x, y) = \int_{-a}^a \varepsilon \ln((x - \varepsilon)^2 + y^2)^{1/2} d\varepsilon = xy(\theta_1 - \theta_2) + \frac{1}{2}(y^2 - x^2 + a^2) \ln \frac{r_1}{r_2} - ax, \quad (11)$$

where we have defined

$$\theta_1 = \arctan \frac{y}{x - a}, \quad \theta_2 = \arctan \frac{y}{x + a}, \quad r_1 = ((x - a)^2 + y^2)^{1/2}, \quad r_2 = ((x + a)^2 + y^2)^{1/2}. \quad (12)$$

Similarly, considering a quadratic variation for $D_i(\varepsilon)$ as given in Equation (7), the common function F_j in (3) is defined as

$$F_j(I_0, I_1, I_2) = \int N_j(\varepsilon) \ln((x - \varepsilon) + y^2)^{1/2} d\varepsilon, \quad j = 1, 2, 3, \quad (13)$$

where I_0 and I_1 are as in (5) and (13) and

$$\begin{aligned} I_2(x, y) &= \int_{-a}^a \varepsilon^2 \ln((x - \varepsilon)^2 + y^2)^{1/2} d\varepsilon \quad (14) \\ &= \frac{y}{3}(3x^2 - y^2)(\theta_1 - \theta_2) + \frac{1}{3}(3xy^2 - x^3 + a^3) \ln(r_1) - \frac{1}{3}(3xy^2 - x^3 - a^3) \ln(r_2) - \frac{2a}{3} \left(x^2 - y^2 + \frac{a^2}{3}\right). \end{aligned}$$

A routine computation yields the partial derivatives of the integrals I_0 , I_1 and I_2 with respect to x and y . These derivatives are needed in the calculation of displacements and stresses in semi-infinite plane problems. As an example,

$$\begin{aligned} I_{2,yyyy} &= -2 \left(\frac{(x - a)}{r_1^2} - \frac{(x + a)}{r_2^2} \right) + 2a \left(\frac{(x - a)^2 - y^2}{r_1^4} + \frac{(x + a)^2 - y^2}{r_2^4} \right) \\ &\quad - 2a^2 \left(\frac{(x - a)^2(r_1^2 - 4y^2)}{r_1^6} - \frac{(x + a)^2(r_2^2 - 4y^2)}{r_2^6} \right), \end{aligned}$$

where r_1 , r_2 , θ_1 , and θ_2 are defined in (12).

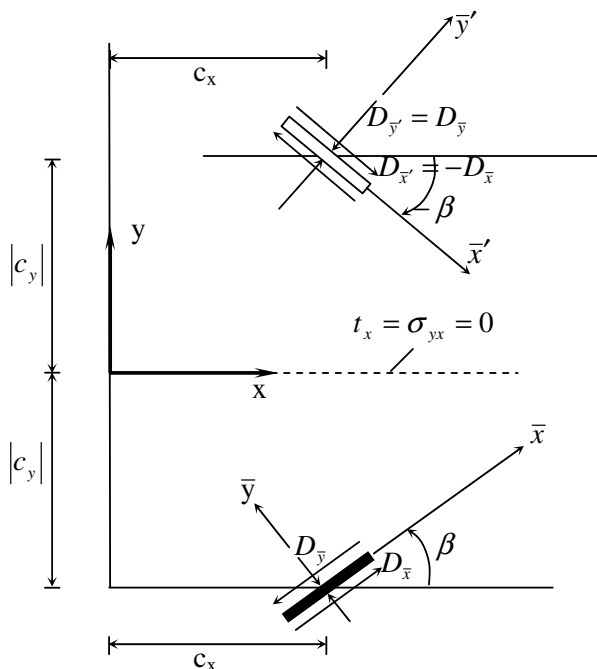


Figure 2. The actual and image displacement discontinuities in half-plane $y \leq 0$.

Half-plane solution. The analytical solution to a constant element displacement discontinuity over the line segment $|x| \leq a, y = 0$ in the semi-infinite region $y \leq 0$ is found in [Crouch and Starfield 1983] using the method of images to be

$$u_i = u_i^A + u_i^I + u_i^S, \quad \sigma_{ij} = \sigma_{ij}^A + \sigma_{ij}^I + \sigma_{ij}^S, \tag{15}$$

where the displacements and stresses due to the actual displacement discontinuity are denoted by a superscript A , those due to its image by I and those resulting from the supplementary solution by S .

Based on the local \bar{x}, \bar{y} coordinates and the image local \bar{x}', \bar{y}' coordinates shown in Figure 2, the displacements and stresses given in (15) can be obtained in global x, y coordinates by using the coordinate, displacement and stress transformation rules explained in [Crouch and Starfield 1983]. Considering the geometry shown in Figure 2, the displacements and stresses due to actual displacement discontinuities can be written directly from (1) and (2). The local \bar{x}, \bar{y} coordinates are related to the global x, y coordinates by the transformation formulas

$$\begin{aligned} \bar{x} &= (x - c_x) \cos \beta + (y - c_y) \sin \beta, \\ \bar{y} &= -(x - c_x) \sin \beta + (y - c_y) \cos \beta. \end{aligned} \tag{16}$$

Denoting the common potential function $F_j(x, y)$ by $F_j^A(\bar{x}, \bar{y}) = F_{j1}^A$ and its derivatives by $F_{j,\bar{x}}^A = F_{j2}^A, F_{j,\bar{y}}^A = F_{j3}^A, F_{j,\bar{x}\bar{y}}^A = F_{j4}^A, F_{j,\bar{x}\bar{x}}^A = -F_{j,\bar{y}\bar{y}}^A = F_{j5}^A, F_{j,\bar{x}\bar{y}\bar{y}}^A = F_{j6}^A, F_{j,\bar{y}\bar{y}\bar{y}}^A = F_{j7}^A$, for the actual displacement

discontinuities, the actual displacements in terms of the global x, y coordinates are given by

$$\begin{aligned}
 u_x^A &= \frac{-1}{4\pi(1-\nu)} \sum_{j=1}^3 \left((-1-2\nu) \sin \beta F_{j2}^A + 2(1-\nu) \cos \beta F_{j3}^A + \bar{y}(\sin \beta F_{j4}^A - \cos \beta F_{j5}^A) D_{\bar{x}}^j \right. \\
 &\quad \left. + (-1-2\nu) \cos \beta F_{j2}^A - 2(1-\nu) \sin \beta F_{j3}^A - \bar{y}(\cos \beta F_{j4}^A + \sin \beta F_{j5}^A) D_{\bar{y}}^j \right), \\
 u_y^A &= \frac{-1}{4\pi(1-\nu)} \sum_{j=1}^3 \left(((1-2\nu) \cos \beta F_{j2}^A + 2(1-\nu) \sin \beta F_{j3}^A - \bar{y}(\cos \beta F_{j4}^A + \sin \beta F_{j5}^A)) D_{\bar{x}}^j \right. \\
 &\quad \left. + (-1-2\nu) \sin \beta F_{j2}^A + 2(1-\nu) \cos \beta F_{j3}^A - \bar{y}(\sin \beta F_{j4}^A - \cos \beta F_{j5}^A) D_{\bar{y}}^j \right), \quad (17)
 \end{aligned}$$

and the actual stresses by

$$\begin{aligned}
 \sigma_{xx}^A &= \frac{-2G}{4\pi(1-\nu)} \sum_{j=1}^3 \left(2 \cos^2 \beta F_{j4}^A + \sin 2\beta F_{j5}^A + \bar{y}(\cos 2\beta F_{j6}^A - \sin 2\beta F_{j7}^A) D_{\bar{x}}^j \right. \\
 &\quad \left. + (-F_{j5}^A + \bar{y}(\sin 2\beta F_{j6}^A + \cos 2\beta F_{j7}^A)) D_{\bar{y}}^j \right), \\
 \sigma_{yy}^A &= \frac{-2G}{4\pi(1-\nu)} \sum_{j=1}^3 \left((2 \sin^2 \beta F_{j4}^A - \sin 2\beta F_{j5}^A - \bar{y}(\cos 2\beta F_{j6}^A + \sin 2\beta F_{j7}^A)) D_{\bar{x}}^j \right. \\
 &\quad \left. + (-F_{j5}^A - \bar{y}(\sin 2\beta F_{j6}^A + \cos 2\beta F_{j7}^A)) D_{\bar{y}}^j \right), \\
 \sigma_{xy}^A &= \frac{-2G}{4\pi(1-\nu)} \sum_{j=1}^3 \left((\sin 2\beta F_{j4}^A - \cos 2\beta F_{j5}^A + \bar{y}(\sin 2\beta F_{j6}^A + \cos 2\beta F_{j7}^A)) D_{\bar{x}}^j \right. \\
 &\quad \left. - \bar{y}(\cos 2\beta F_{j6}^A - \sin 2\beta F_{j7}^A) D_{\bar{y}}^j \right). \quad (18)
 \end{aligned}$$

The displacements and stresses due to the image displacement discontinuity can be expressed in term of a single function $F_j^I(\bar{x}', \bar{y}')$ in which the image local \bar{x}', \bar{y}' coordinates (as shown in Figure 2) are related to the x, y coordinates by the transformation formula

$$\begin{aligned}
 \bar{x}' &= (x - c_x) \cos \beta - (y + c_y) \sin \beta, \\
 \bar{y}' &= (x - c_x) \sin \beta + (y + c_y) \cos \beta,
 \end{aligned} \quad (19)$$

which is obtained by replacing c_y and β in (16) by c_y and $-\beta$.

It can be shown that the supplementary solution for the displacements and stresses can be expressed in term of the function $F_j^I(\bar{x}', \bar{y}')$ and its derivatives. The final expressions for the combined displacements $u_i^I + u_i^S$ and stresses $\sigma_{ij}^I + \sigma_{ij}^S$ as given in the Appendix.

The displacement discontinuity functions $D_i(\varepsilon)$ in (4) can be used either in a constant element form or in a higher-order element form as follows, to solve the displacements and stresses of (1), (2) and (15). Two degrees of freedom are used for each node at the center of each element. Crawford and Curran [1982] have developed a higher-order displacement discontinuity for linear and quadratic elements using four and six degrees of freedom respectively. Shou and Crouch [1995] have introduced a new higher-order displacement discontinuity for two-dimensional infinite plane problems using only two degrees of freedom for each element while still preserving the advantages of the approach in [Crawford and Curran 1982].

In this study, Shou and Crouch's approach is extended and modified for both linear and quadratic displacement discontinuity variations (the original formulation covers only quadratic displacement variations in infinite plane problems) in order to solve the half-plane boundary-value problems with traction-free surfaces. For linear and quadratic displacement variations, the interpolation of displacement values over two- and three-element patches, respectively, is the basis of our formulation (recall Figure 1).

The general solution to plane elasticity problems involves two types of boundary conditions: the stress boundary conditions $\sigma_s^i = (\sigma_s^i)_0$, $\sigma_n^i = (\sigma_n^i)_0$ and the displacement boundary conditions $u_s^i = (u_s^i)_0$, $u_n^i = (u_n^i)_0$. In these equations the right-hand sides stand for the given boundary values of the stresses and displacements for the local s and n coordinates (that is, the same as the local \bar{x} , \bar{y} coordinates shown in Figure 2) defining at the center of each two-element patch (linear variation) or three-element patch (quadratic variation). Finally, then, we obtain a system of $2 \times 2N$ or $2 \times 3N$ algebraic equations in as many unknown displacement discontinuity components:

$$b_s^i = \sum_{j=1}^N C_{ss}(i, j) D_s^j + \sum_{j=1}^N C_{sn}(i, j) D_n^j, \quad b_n^i = \sum_{j=1}^N C_{ns}(i, j) D_s^j + \sum_{j=1}^N C_{nn}(i, j) D_n^j, \quad i = 1, N. \quad (20)$$

The quantities b_s^i and b_n^i stand for the known boundary values of stress and displacement, and $C_{ss}(i, j)$, etc., are the corresponding influence coefficients [Crouch and Starfield 1983]. For the solution of cracked body problems in half-planes with traction-free surfaces, we developed two computer programs: SIDDLCR for semi-infinite displacement discontinuity method using linear displacement discontinuity elements for crack analysis and SIDDQCR for semi-infinite displacement discontinuity method using quadratic displacement discontinuity elements for crack analysis. Since the solution of the infinite plane case is part of the solution of the half-plane problems (i.e., the actual solution part given in (17) and (18)), these two computer codes can actually solve general elasticity problems in finite, infinite and semi-infinite planes.

3. Crack tip element formulation and stress intensity factor computation

Consider a body of arbitrary shape with a crack of arbitrary size, subjected to arbitrary tensile and shear loadings (mode-I and mode-II loading). The stresses and displacements near the crack tip are given in [Rossmannith 1983; Whittaker et al. 1992] and other textbooks; but since we use the displacement discontinuity method here we need the formulations given for the SIFs K_I and K_{II} in terms of the normal and shear displacement discontinuities [Whittaker et al. 1992; Scavia 1995]:

$$K_I = \frac{G}{4(1-\nu)} \left(\frac{2\pi}{a} \right)^{1/2} D_y(a), \quad K_{II} = \frac{G}{4(1-\nu)} \left(\frac{2\pi}{a} \right)^{1/2} D_x(a). \quad (21)$$

Due to the singularity variations $1/\sqrt{r}$ and \sqrt{r} for the stresses and displacements in the vicinity of the crack tip the accuracy of the displacement discontinuity method decreases, and usually a special treatment of the crack at the tip is necessary. A special crack tip element which already has been introduced in literature (see [Crouch and Starfield 1983], for example) is used here to represent the singularity. Using a special crack tip element of length $2a$ as shown in Figure 3, we obtain the parabolic displacement

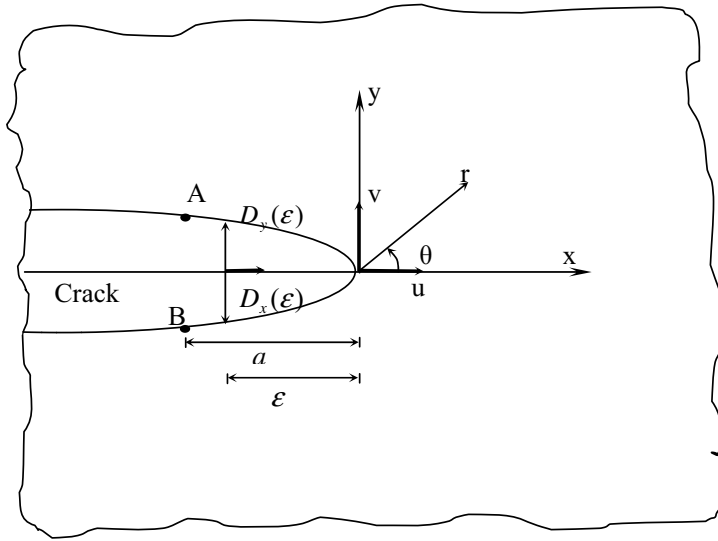


Figure 3. Displacement correlation technique for the special crack tip element.

discontinuity variations along this element as

$$D_i(\varepsilon) = D_i(a)(\varepsilon/a)^{1/2}, \quad i = x, y, \tag{22}$$

where ε is the distance from crack tip and $D_y(a)$ and $D_x(a)$ are the opening (normal) and sliding (shear) displacement discontinuities at the center of special crack tip element.

The potential functions $f_C(x, y)$ and $g_C(x, y)$ for the crack tip element can be expressed as

$$\begin{aligned} f_C(x, y) &= \frac{-1}{4\pi(1-\nu)} \int_{=a}^a \frac{D_x(a)}{a^{1/2}} \varepsilon^{1/2} \ln((x-\varepsilon)^2 + y^2)^{1/2} d\varepsilon, \\ g_C(x, y) &= \frac{-1}{4\pi(1-\nu)} \int_{=a}^a \frac{D_y(a)}{a^{1/2}} \varepsilon^{1/2} \ln((x-\varepsilon)^2 + y^2)^{1/2} d\varepsilon. \end{aligned} \tag{23}$$

These functions have a common integral of the form

$$I_C = \int_0^{2a} \varepsilon^{1/2} \ln((x-\varepsilon)^2 + y^2)^{1/2} d\varepsilon. \tag{24}$$

The derivatives of this integral, which are used in calculation of the crack tip displacement discontinuities in semi-infinite plane problems, are easily computed:

$$I_{C,x} = \int_0^{2a} \frac{\varepsilon^{1/2}(x-\varepsilon)}{(x-\varepsilon)^2 + y^2} d\varepsilon = xA_1 - A_2, \quad I_{C,y} = \int_0^{2a} \frac{\varepsilon^{1/2}y}{(x-\varepsilon)^2 + y^2} d\varepsilon = yA_1,$$

where we have introduced

$$A_1 = \frac{1}{\rho} \left(\frac{1}{2} \left(\cos \varphi - \frac{x}{y} \sin \varphi \right) \ln \frac{2a - 2\sqrt{2a\rho} \cos \varphi + \rho^2}{2a + 2\sqrt{2a\rho} \cos \varphi + \rho^2} + \left(\sin \varphi + \frac{x}{y} \cos \varphi \right) \arctan \left(\frac{2\sqrt{2a\rho} \sin \varphi}{\rho^2 - 2a} \right) \right),$$

$$A_2 = \rho \left(\frac{1}{2} \left(\cos \varphi + \frac{x}{y} \sin \varphi \right) \ln \frac{2a - 2\sqrt{2a\rho} \cos \varphi + \rho^2}{2a + 2\sqrt{2a\rho} \cos \varphi + \rho^2} + \left(\sin \varphi + \frac{x}{y} \cos \varphi \right) \arctan \left(\frac{2\sqrt{2a\rho} \sin \varphi}{\rho^2 - 2a} \right) \right),$$

with $\rho = (x^2 + y^2)^{1/4}$ and $\varphi = \frac{1}{2} \arctan(y/x)$. Moreover,

$$\begin{aligned} I_{c,xy} &= yA_{1,x}, & I_{c,xyy} &= A_{1,x} + yA_{1,xy}, & I_{c,xyyy} &= 2A_{1,xy} + yA_{1,xyy}, \\ I_{c,yy} &= A_1 + yA_{1,y} = -I_{c,xx}, & I_{c,yyy} &= 2A_{1,y} + yA_{1,yy}, & I_{c,yyyy} &= 3A_{1,yy} + yA_{1,yyy}. \end{aligned}$$

4. Crack initiation and direction of its propagation

Several mixed mode fracture criteria are well known from the literature [Ingraffea 1981; 1987; Huang and Wang 1985; Zipf and Bieniawski 1989; Ouchterlony 1988; Stephansson et al. 2001; Rao et al. 2003; Backers et al. 2004; Shen et al. 2004], any of them can be applied to crack analysis problems using this model. In this study as the blast hole radial cracks are mostly in opening mode case, the simple maximum tangential stress criterion or σ -criterion is used here to predict the angle of crack initiation. This criterion is a mixed mode fracture criterion which is widely used and well fitted with the experimental results [Ingraffea 1983; Broek 1989; Guo et al. 1992; Carpinteri and Yang 1997].

Based on this criterion the crack tip will start propagating when

$$\cos \frac{\theta_0}{2} \left(K_I \cos^2 \frac{\theta_0}{2} - \frac{3}{2} K_{II} \sin \frac{\theta_0}{2} \right) = K_{IC}, \quad (25)$$

where K_{IC} is the mode-I fracture toughness of the material and θ_0 is the propagation angle. The latter value corresponding to the crack tip should satisfy the condition

$$K_I \sin \theta_0 + K_{II} (3 \cos \theta_0 - 1) = 0. \quad (26)$$

5. Verification of higher-order semi-infinite displacement discontinuity

Verification of this method is made through the solution of simple example problems. We first take a center slant crack in an infinite body, as shown in Figure 4. The slant angle, β , changes counterclockwise from the x -axis, and the tensile stress $\sigma = 10$ MPa acts parallel to the x -axis. A half crack length $b = 1$ meter, modulus of elasticity $E = 10$ GPa, Poisson's ratio $\nu = 0.2$, fracture toughness $K_{IC} = 2$ MPa \sqrt{m} are assumed. The analytical solution of the first and second mode stress intensity factors K_I and K_{II} for the infinite body problem are given as (see [Guo et al. 1990; Whittaker et al. 1992])

$$\begin{aligned} K_I &= \sigma (\pi b)^{1/2} \sin^2 \beta \quad \implies \frac{K_I}{\sigma \sqrt{\pi b}} = \sin^2 \beta, \\ K_{II} &= \sigma (\pi b)^{1/2} \sin \beta \cos \beta \quad \implies \frac{K_{II}}{\sigma \sqrt{\pi b}} = \sin \beta \cos \beta. \end{aligned} \quad (27)$$

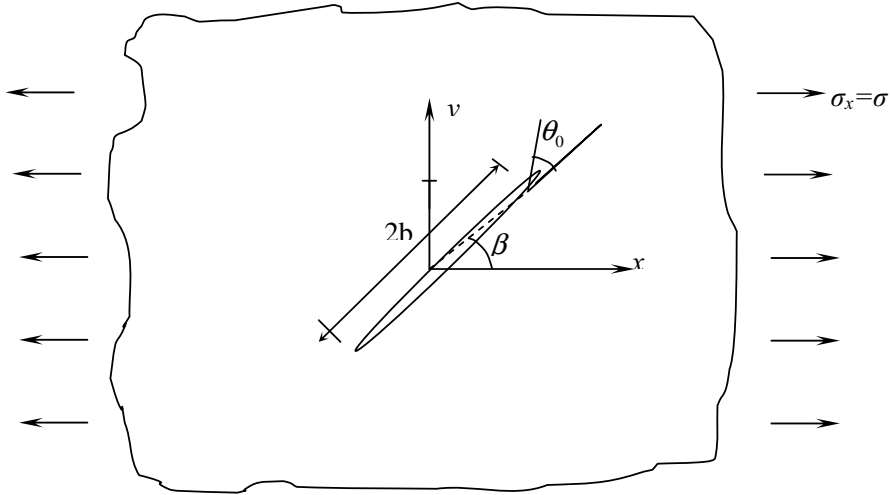


Figure 4. Center slant cracks in an infinite body under far-field tension.

The normalized mixed mode stress intensity factors, $K_I/(\sigma\sqrt{\pi b})$ and $K_{II}/(\sigma\sqrt{\pi b})$, are obtained analytically for different crack inclination from Equations (27), and numerically by means of the two software programs SDDLRCR and SIDDQCR (page 445), using a total of 98 nodes and a crack tip length equal to one-tenth the half crack length—that is, an L/b ratio of 0.1. Some of the results obtained are tabulated in Table 1, and give an idea of the accuracy and usefulness of the programs.

To investigate the effect of the number of elements on accuracy, we solved numerically two problems, respectively with 45° and 30° slanted cracks, choosing $L/b = 0.1$ and a varying number of nodes. The normalized numerical results are given in Table 2. They show that using any number of nodes above 24 gives very accurate values for both SDDLRCR and SIDDQCR.

Angle β	$K_I/(\sigma\sqrt{\pi b})$			$K_{II}/(\sigma\sqrt{\pi b})$		
	Analytic	SIDDQCR	SDDLRCR	Analytic	SIDDQCR	SDDLRCR
10°	0.0302	0.0302	0.0309	0.1711	0.1711	0.1752
20°	0.1170	0.1171	0.1198	0.3214	0.3216	0.3292
30°	0.2500	0.2502	0.2561	0.4330	0.4334	0.4435
40°	0.4132	0.4135	0.4176	0.4924	0.4920	0.4977
50°	0.5868	0.5864	0.5932	0.4924	0.4921	0.4977
60°	0.7500	0.7495	0.7581	0.4330	0.4327	0.4369
70°	0.8830	0.8824	0.8926	0.3214	0.3212	0.3249
80°	0.9696	0.9692	0.9803	0.1711	0.1709	0.1713
90°	1.0000	0.9996	1.011	0.0000	0.0000	0.0000

Table 1. Analytical and numerical values of the normalized stress intensity factors for the slant center crack at different orientation from the loaded axis (x -axis), for $L/b = 0.1$ and 98 nodes.

Nodes	$K_I/(\sigma\sqrt{\pi b})$ (45°)		$K_I/(\sigma\sqrt{\pi b})$ (30°)		$K_{II}/(\sigma\sqrt{\pi b})$ (30°)	
	SIDDQCR	SIDDLRCR	SIDDQCR	SIDDLRCR	SIDDQCR	SIDDLRCR
12	0.5116	0.5396	0.2558	0.2698	0.4431	0.4673
24	0.5014	0.5160	0.2507	0.2580	0.4342	0.4469
36	0.4999	0.5097	0.2500	0.2549	0.4330	0.4414
48	0.4997	0.5069	0.2499	0.2535	0.4327	0.4390
60	0.4997	0.5054	0.2498	0.2527	0.4327	0.4377
72	0.4997	0.5044	0.2498	0.2522	0.4327	0.4368
84	0.4998	0.5038	0.2499	0.2517	0.4328	0.4363
96	0.4998	0.5033	0.2499	0.2516	0.4328	0.4358
108	0.4999	0.5029	0.2499	0.2514	0.4328	0.4355
120	0.4999	0.5026	0.2499	0.2513	0.4329	0.4353

Table 2. Numerical values of the stress intensity factors for the 45° and 30° slant center cracks using varying number of nodes and $L/b = 0.1$.

Finally, to show the effect of the length of the crack tip element on the accuracy of the results, the same two problems are solved numerically using 98 nodes and different L/b ratios. The results are given in Table 3. We see that for any L/b ratio above 0.025, both programs give very accurate values.

Because of its simplicity, the center slant crack problem has been solved by various investigators such as Guo et al. [1990], who used constant element displacement discontinuity with a special crack tip element for angles 30°, 40°, 50°, 60°, 70° and 80°. These authors used a different fracture criterion for evaluating the crack initiation angle θ_0 and compared their results with the results obtained by other

L/b	$K_I/(\sigma\sqrt{\pi b})$ (45°)		$K_I/(\sigma\sqrt{\pi b})$ (30°)		$K_{II}/(\sigma\sqrt{\pi b})$ (30°)	
	SIDDQCR	SIDDLRCR	SIDDQCR	SIDDLRCR	SIDDQCR	SIDDLRCR
0.025	0.5028	0.5166	0.2514	0.2614	0.4354	0.4467
0.050	0.5000	0.5072	0.2500	0.2536	0.4330	0.4393
0.075	0.4998	0.5045	0.2498	0.2523	0.4328	0.4369
0.100	0.4998	0.5033	0.2499	0.2516	0.4328	0.4358
0.125	0.4999	0.5026	0.2499	0.2513	0.4329	0.4352
0.150	0.5001	0.5021	0.2500	0.2511	0.4331	0.4349
0.175	0.5003	0.5019	0.2501	0.2510	0.4333	0.4347
0.200	0.5005	0.5018	0.2503	0.2509	0.4335	0.4346
0.225	0.5008	0.5018	0.2504	0.2509	0.4337	0.4346
0.250	0.5011	0.5019	0.2506	0.2510	0.4340	0.4347
0.275	0.5015	0.5021	0.2507	0.2510	0.4343	0.4348
0.300	0.5019	0.5023	0.2509	0.2512	0.4346	0.4350

Table 3. Numerical values of the stress intensity factors for the 45° and 30° slant center cracks using varying L/b ratios and 98 nodes.

Angle β	θ_0 , present work		θ_0 as reported in [Guo et al. 1990]			
	SIDDQCR	SIDDLQR	σ -criterion	S-criterion	Exper.	Numer.
30	60.00	60.00	60.2	63.5	62.4	67.0
40	55.65	55.65	55.7	56.7	55.1	59.0
50	50.29	50.29	50.2	49.5	51.1	51.0
60	43.22	43.22	43.2	41.9	43.1	41.0
70	33.26	33.26	33.2	31.8	30.7	29.0
80	18.91	18.91	19.3	18.5	17.3	15.0

Table 4. Crack initiation angle θ_0 obtained by different methods for the center slant crack problem.

researchers using different fracture theories. Table 4 compares the results obtained for crack initiation angle θ_0 with SIDDQCR and SIDDLQR, using the maximum tangential stress theory proposed by Erdogan and Sih [1963], and the results obtained by other methods as given by Guo et al. The numerical results obtained here are very close to those predicted by the σ -criterion.

For the verification of the semi-infinite higher-order displacement discontinuity method, the problem of a 45° slant crack with different depths (C/b ratio) from the free surface of the half-plane is considered. This is the problem shown in Figure 5, where $C = C_y$ is the depth at the center of the slant crack from the free surface of the half-plane as shown in Figure 3.

The normalized stress intensity factors $K_I/(\sigma\sqrt{\pi b})$ and $K_{II}/(\sigma\sqrt{\pi b})$ of the upper and lower crack tips were obtained using SIDDQCR. The numerical results for the 45° crack are given in Table 5, where

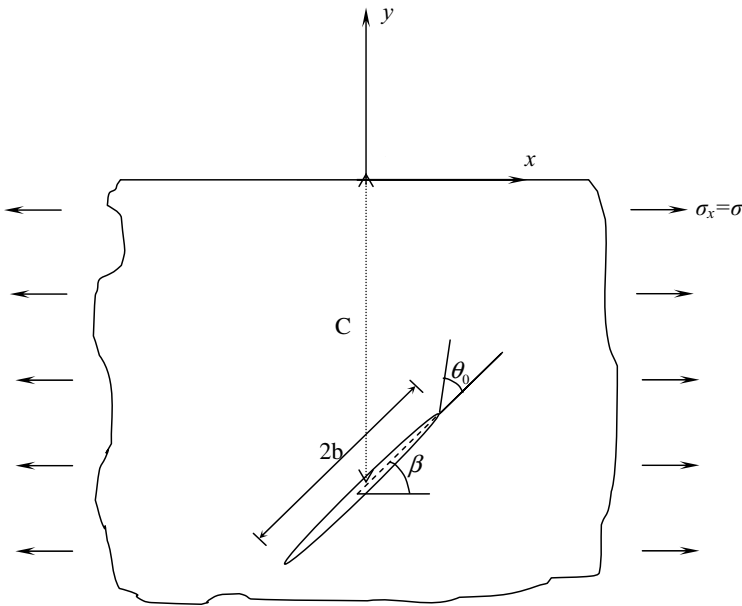


Figure 5. Slant cracks in a semi-infinite body under far field tension.

C/b ratio	$K_I/(\sigma\sqrt{\pi b})$		$K_{II}/(\sigma\sqrt{\pi b})$	
	Upper Tip	Lower Tip	Upper Tip	Lower Tip
1	0.6147	0.5278	0.4603	0.5530
2	0.5115	0.5067	0.4669	0.4957
3	0.4989	0.4996	0.4785	0.4907
4	0.4959	0.4971	0.4841	0.4905
5	0.4951	0.4961	0.4874	0.4911
6	0.4950	0.4958	0.4895	0.4919
7	0.4951	0.4958	0.4909	0.4927
8	0.4953	0.4959	0.4920	0.4930
9	0.4956	0.4960	0.4929	0.4938
10	0.4958	0.4962	0.4935	0.4943

Table 5. Normalized stress intensity factors $K_I/(\sigma\sqrt{\pi b})$ and $K_{II}/(\sigma\sqrt{\pi b})$ of the upper and lower crack tips for a 45 degrees inclined crack in a semi-infinite body with different C/b ratios, using 98 nodes and $L/b = 0.1$.

we have used a total of 98 nodes (including the two crack tip elements) and an L/b ratio of 0.1. The table shows that as the crack becomes very close to the free surface of the half-plane, the mode-I stress intensity factor K_I of the upper crack tip (the one nearer the free surface) increases more rapidly compared to that of the lower tip, but the mode-II stress intensity factor K_{II} decreases. The analytical solution for the 45° center slant crack in an infinite plate gives $K_I/(\sigma\sqrt{\pi b}) = K_{II}/(\sigma\sqrt{\pi b}) = 0.5$, and as can be seen from the table, the numerical values tend to this analytical value as the depth C increases to infinity.

6. Numerical analysis of radial cracks in blasting

Radial crack propagation in blasting operations is a complicated and interesting phenomenon. The initiation and propagation mechanisms have been investigated in [Ingraffea 1983; Mortazavi and Katsabanis 2001; Cho et al. 2004]. Generally, two forms of radial crack analysis have been used: dynamic crack analysis (considering stress wave and/or gas pressurization theories) and quasistatic crack analysis (considering only gas pressurization theory); see [Kutter and Fairhurst 1971; Courtesen 1979; Ash 1985; Donzé et al. 1997; Cho et al. 2004]. Dynamic crack analysis is far beyond our scope; instead we briefly consider quasistatic radial crack analysis due to gas pressurization, to show the effectiveness of the present model for solving some crack problems occurring in rock fracture mechanics. Analytical solutions of the radial crack propagation in an infinite elastic rock have been discussed in the literature; see [Ouchterlony 1983; Whittaker et al. 1992] and references therein. Ouchterlony has extensively analyzed various load configurations in relation to rock blasting and determined the stress intensity factors for them by using the conformal mapping method. This has provided valuable information on crack-growth behavior due to internal pressure, and the effect of gas penetration in the radial cracks has also been investigated.

In this section we discuss bench blasting problems based on Ouchterlony's setup and the present approach. We solve the problems shown in Figure 6, considering two limiting cases: empty cracks, meaning that no gas pressure penetrates through the radial cracks during blasting; and fully pressurized

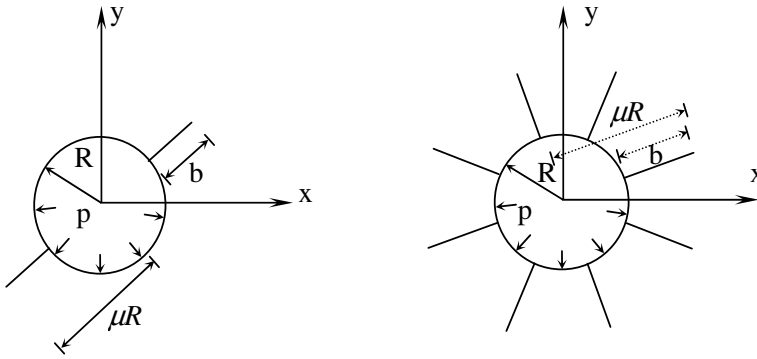


Figure 6. Pressurized circular holes (blast holes) with radial cracks in infinite planes. The examples shown have two (left) and eight (right) radial cracks.

cracks, meaning that the gas pressure fully penetrates the radial cracks, so the gas pressure in the radial cracks equals the total blast hole pressure.

SIDDLR and SIDDQCR have been used for the numerical solution. The following assumptions and data have used throughout: hole radius, $R = 1$ m; inside pressure, $p = 10$ MPa; modulus of elasticity, $E = 10$ GPa; Poisson's ratio, $\nu = 0.2$; rock fracture toughness, $K_{IC} = 2$ MPa $m^{1/2}$; ratio of crack tip element to crack length, $L/b = 0.1$. Exploiting symmetry, 30 quadratic elements or 40 linear elements are used to discretize the boundary of a circular hole, and 10 quadratic elements or 15 linear elements are used for the discretization of each radial crack (excluding the crack tip elements). We compute numerically the normalized mode-I stress intensity factor $K_I/(p\sqrt{\pi\mu R})$ for different crack length ratios $\mu = (b + R)/R$ of a pressurized blast hole under uniform inside pressure p with 2 and 8 symmetric radial cracks emanating from the hole. These numerical results, for empty and fully pressurized radial cracks, are compared in Figures 7 and 8 with the corresponding analytical values given in [Ouchterlony 1983]. We see that, particularly for long radial cracks, the analytical and numerical values of normalized stress intensity factors (mode-I) are very close to each other. The blasting pressures p are around 0.56 to 1.0 GPa, but the results plotted here are normalized in the form of $K_I/(p\sqrt{\pi\mu R})$.

In a similar manner we analyzed numerically the problem shown in Figure 9 using SIDDQCR. The calculated fracture parameters are shown in Table 6 for the two extreme cases of empty cracks and fully pressurized cracks against different ratios B/R of burden radius to blast hole (B/R can be viewed as a normalized hole depth relative to the free surface of the half-plane). All these results were obtained through SIDDQCR, using a constant value of $\mu = 2.5$ for the crack length ratio.

The analytical results for the problem of a pressurized circular hole with four symmetric empty radial cracks in an infinite plane are: $K_I/(p\sqrt{\pi\mu R}) = 0.1966$ and $K_{II}/(p\sqrt{\pi\mu R}) = 0.0$, and for the fully pressurized radial cracks are: $K_I/(p\sqrt{\pi\mu R}) = 0.9085$ and $K_{II}/(p\sqrt{\pi\mu R}) = 0.0$ respectively [Ouchterlony 1983].

Table 6 compares the different results obtained for the upper crack (the crack near to the free surface of the half-plane), and the lower crack. The results given in this table show that as the burden (B/R ratio) increases the mixed mode stress intensity factors K_I and K_{II} tend to their corresponding analytical values for the infinite plane case.

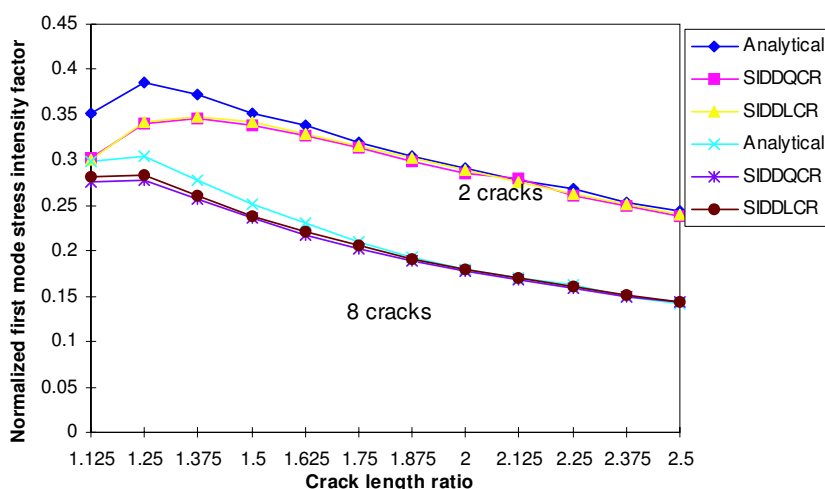


Figure 7. Blast hole with 2 and 8 empty radial cracks: comparison of analytical and numerical values of $K_I/(p\sqrt{\pi\mu R})$ for varying crack length ratio μ .

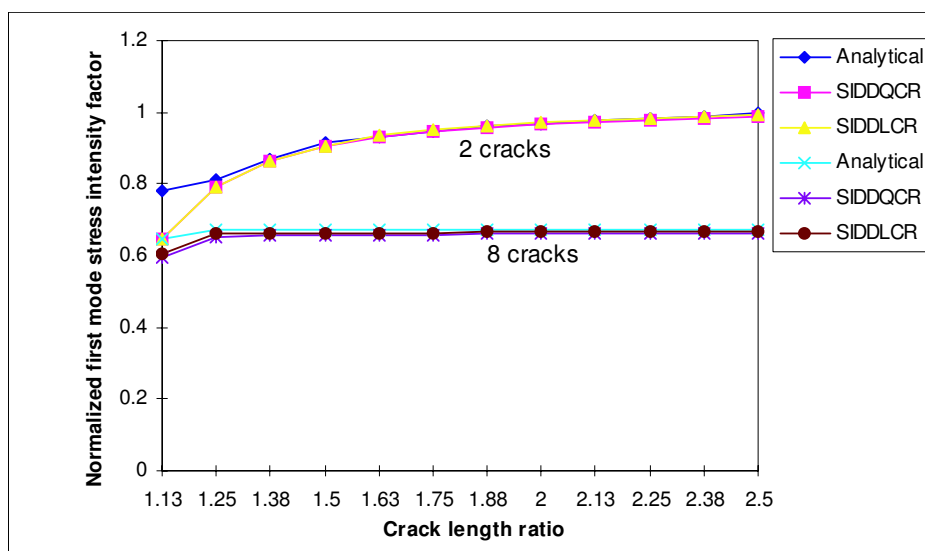


Figure 8. Blast hole with 2 and 8 fully pressurized radial cracks: comparison of analytical and numerical values of $K_I/(p\sqrt{\pi\mu R})$ for varying crack length ratio μ .

The results obtained in this paper show that, as the Mode II stress intensity factor (K_{II}) is very small compared to the Mode I (K_I), and for the practical problems; the crack initiation angle θ_0 is also very small, for both upper and lower cracks, then it may be concluded that the radial cracks produced in blasting operations propagates nearly in their own planes due to the high influence of K_I . However, when the radial cracks become very close to the free surface (that is, for small B/R ratios), the upper cracks divert away from the free surface in the direction of the crack deviation angle θ_0 .

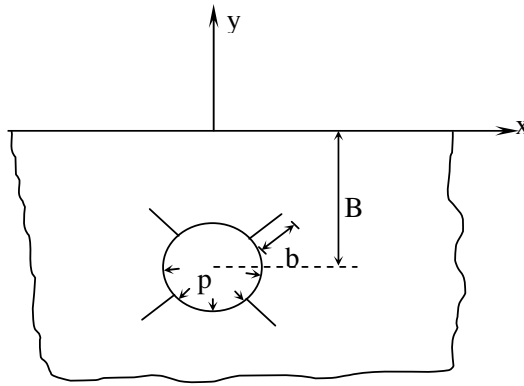


Figure 9. A blast hole in a semi-infinite plane with 4 symmetric radial cracks.

7. Conclusion

The higher-order displacement discontinuity method is used to solve boundary value problems in finite, infinite and semi-infinite plane problems. The special crack tip formulation is also modified and used for solving the crack problems in semi-infinite planes. Then based on the formulation derived in Sections 2 and 3, the two computer programs SIDDLCR (semi-infinite displacement discontinuity method with linear elements for crack analysis) and SIDDQCR (semi-infinite displacement discontinuity method with quadratic elements for crack analysis) were developed. These programs use a special crack tip element

B/R	$K_I/(p\sqrt{\pi\mu R})$		$K_{II}/(p\sqrt{\pi\mu R})$		θ_0 (degrees)	
	Up. crack	Lo. crack	Up. crack	Lo. crack	Up. crack	Lo. crack
2.00	2.2036	1.3534	0.1954	0.0194	-10.0	-1.6
2.25	1.8940	1.2366	0.1230	-0.0307	-7.4	2.8
2.50	1.6993	1.1766	0.1285	-0.0465	-8.6	4.5
2.75	1.5298	1.1352	0.1229	-0.0565	-9.1	5.7
3.00	1.3937	1.1022	0.1109	-0.0626	-9.0	6.5
3.25	1.2690	1.0578	0.0949	-0.0552	-8.5	5.9
3.50	1.1927	1.0392	0.0862	-0.0573	-8.2	6.3
3.75	1.1116	1.0230	0.0609	-0.0587	-6.2	6.5
4.00	1.0344	1.0085	0.0470	-0.0580	-5.5	6.5
4.25	1.0114	0.9960	0.0362	-0.0573	-4.1	6.5
4.50	0.9863	0.9903	0.0223	-0.0523	-2.5	6.0
4.75	0.9582	0.9816	0.0091	-0.0510	-1.1	5.9
5.00	0.9390	0.9745	0.0044	-0.0487	-0.5	5.7

Table 6. Normalized stress intensity factors and crack propagation angle θ_0 for a pressurized hole under uniform inside pressure p , with four fully pressurized radial cracks, for varying values of B/R (corresponding to different depths).

with a simple mixed mode fracture criterion — the maximum tangential stress criterion, based on linear elastic fracture mechanics principles — to quasistatically analyze the radial cracks. For the verification of this method some example problems of cracked bodies in infinite and semi-infinite planes are solved and the results are compared with their existing analytical results and/or with the results obtained by some previous researchers. The computed results obtained by these codes are very accurate as compared to the previous results (given in the literature), which proves the validity and accuracy of the proposed method. For completeness, some of the necessary formulations which are derived and used in these computer codes are also given in the appendices of this paper.

Appendix: Supplementary and image solutions for displacements and stresses

We maintain the notation of Section 2 and Figure 2 and define

$$F_{j8}^I(\bar{x}', \bar{y}') = \frac{\partial^4 F_j^I(\bar{x}', \bar{y}')}{\partial \bar{x}' \partial \bar{y}'^3}, \quad F_{j9}^I(\bar{x}', \bar{y}') = \frac{\partial^4 F_j^I(\bar{x}', \bar{y}')}{\partial \bar{y}'^4}.$$

The combined displacements $u_i^I + u_i^S$ are

$$\begin{aligned} u_x^I + u_x^S &= \frac{-1}{4\pi(1-\nu)} \\ &\times \sum_{j=1}^3 \left(\left((1-2\nu) \sin \beta F_{j2}^I - 2(1-\nu) \cos \beta F_{j3}^I + ((3-4\nu)(y \sin 2\beta - \bar{y} \sin \beta) + 2y \sin 2\beta) F_{j4}^I \right. \right. \\ &\quad \left. \left. + ((3-4\nu)(y \cos 2\beta - \bar{y} \cos \beta) - y(1-2 \cos 2\beta)) F_{j5}^I \right. \right. \\ &\quad \left. \left. + 2y(y \sin 3\beta - \bar{y} \sin 2\beta) F_{j6}^I - 2y(y \cos 3\beta - \bar{y} \cos 2\beta) F_{j7}^I \right) D_{\bar{x}}^j \right. \\ &\quad \left. + \left((1-2\nu) \cos \beta F_{j2}^I + 2(1-\nu) \sin \beta F_{j3}^I - ((3-4\nu)(y \cos 2\beta - \bar{y} \cos \beta) - y) F_{j4}^I \right. \right. \\ &\quad \left. \left. + (3-4\nu)(y \sin 2\beta - \bar{y} \sin \beta) F_{j5}^I \right. \right. \\ &\quad \left. \left. - 2y(y \cos 3\beta - \bar{y} \cos 2\beta) F_{j6}^I - 2y(y \sin 3\beta - \bar{y} \sin 2\beta) F_{j7}^I \right) D_{\bar{y}}^j \right), \end{aligned}$$

$$\begin{aligned} u_y^I + u_y^S &= \frac{-1}{4\pi(1-\nu)} \\ &\times \sum_{j=1}^3 \left(\left((1-2\nu) \cos \beta F_{j2}^I - 2(1-\nu) \sin \beta F_{j3}^I - ((3-4\nu)(y \cos 2\beta - \bar{y} \cos \beta) + y(1-2 \cos 2\beta)) F_{j4}^I \right. \right. \\ &\quad \left. \left. + ((3-4\nu)(y \sin 2\beta - \bar{y} \sin \beta) - 2y \sin 2\beta) F_{j5}^I \right. \right. \\ &\quad \left. \left. + 2y(y \cos 3\beta - \bar{y} \cos 2\beta) F_{j6}^I + 2y(y \sin 3\beta - \bar{y} \sin 2\beta) F_{j7}^I \right) D_{\bar{x}}^j \right. \\ &\quad \left. + \left((1-2\nu) \sin \beta F_{j2}^I - 2(1-\nu) \cos \beta F_{j3}^I - (3-4\nu)(y \sin 2\beta - \bar{y} \sin \beta) F_{j4}^I \right. \right. \\ &\quad \left. \left. - ((3-4\nu)(y \cos 2\beta - \bar{y} \cos \beta) + y) F_{j5}^I \right. \right. \\ &\quad \left. \left. + 2y(y \sin 3\beta - \bar{y} \sin 2\beta) F_{j6}^I - 2y(y \cos 3\beta - \bar{y} \cos 2\beta) F_{j7}^I \right) D_{\bar{y}}^j \right), \end{aligned}$$

The stresses $\sigma_{ij}^I + \sigma_{ij}^S$ associated with these displacements are

$$\sigma_{xx}^I + \sigma_{xx}^S = \frac{-2G}{4\pi(1-\nu)} \left(\left(F_{j4}^I - 3(\cos 2\beta F_{j4}^I - \sin 2\beta) F_{j5}^I + (2y(\cos \beta - 3 \cos 3\beta) + 3\bar{y} \cos 2\beta) F_{j6}^I \right. \right. \\ \left. \left. + (2y(\sin \beta - 3 \sin 3\beta) + 3\bar{y} \sin 2\beta) F_{j7}^I \right. \right. \\ \left. \left. - 2y(y \cos 4\beta - \bar{y} \cos 3\beta) F_{j8}^I - 2y(y \sin 4\beta - \bar{y} \sin 3\beta) F_{j9}^I \right) D_{\bar{x}}^j \right. \\ \left. + \left(F_{j5}^I + (2y(\sin \beta - 2 \sin 3\beta) + 3\bar{y} \sin 2\beta) F_{j6}^I - (2y(\cos \beta - 2 \cos 3\beta) + 3\bar{y} \cos 2\beta) F_{j7}^I \right. \right. \\ \left. \left. - 2y(y \sin 4\beta - \bar{y} \sin 3\beta) F_{j8}^I + 2y(y \cos 4\beta - \bar{y} \cos 3\beta) F_{j9}^I \right) D_{\bar{y}}^j \right),$$

$$\sigma_{yy}^I + \sigma_{yy}^S = \frac{-2G}{4\pi(1-\nu)} \left(\left(F_{j4}^I - (\cos 2\beta F_{j4}^I - \sin 2\beta) F_{j5}^I - (4y \sin 2\beta - \bar{y} \cos 2\beta) F_{j6}^I \right. \right. \\ \left. \left. + (4y \sin \beta \cos 2\beta + \bar{y} \sin 2\beta) F_{j7}^I + 2y(y \sin 4\beta - \bar{y} \sin 3\beta) F_{j9}^I \right) D_{\bar{x}}^j \right. \\ \left. + \left(F_{j5}^I - (2y \sin \beta - \bar{y} \sin 2\beta) F_{j6}^I + (2y \cos \beta - \bar{y} \cos 2\beta) F_{j7}^I \right. \right. \\ \left. \left. + 2y(y \sin 4\beta - \bar{y} \sin 3\beta) F_{j8}^I - 2y(y \cos 4\beta - \bar{y} \cos 3\beta) F_{j9}^I \right) D_{\bar{y}}^j \right),$$

$$\sigma_{xy}^I + \sigma_{xy}^S = \frac{-2G}{4\pi(1-\nu)} \left(\left((\sin 2\beta + \cos 2\beta) F_{j4}^I + (2y \sin \beta (1 + 4 \cos 2\beta) - \bar{y} \sin 2\beta) F_{j6}^I \right. \right. \\ \left. \left. + (2y \cos \beta (3 - 4 \cos 2\beta) + \bar{y} \sin 2\beta) F_{j7}^I \right. \right. \\ \left. \left. + 2y(y \sin 4\beta - \bar{y} \sin 3\beta) F_{j8}^I - 2y(y \cos 4\beta - \bar{y} \cos 3\beta) F_{j9}^I \right) D_{\bar{x}}^j \right. \\ \left. + \left((4y \sin \beta \sin 2\beta + \bar{y} \cos 2\beta) F_{j6}^I - (4y \sin \beta \cos 2\beta - \bar{y} \sin 2\beta) F_{j7}^I \right. \right. \\ \left. \left. - 2y(y \cos 4\beta - \bar{y} \cos 3\beta) F_{j8}^I - 2y(y \sin 4\beta - \bar{y} \sin 3\beta) F_{j9}^I \right) D_{\bar{y}}^j \right),$$

References

- [Aliabadi 1998] M. H. Aliabadi, *Fracture of rocks*, Computational Mechanics Publications, Southampton, 1998.
- [Ash 1985] R. L. Ash, "Flexural rupture as a rock breakage mechanism in blasting", pp. 371–378 in *Fragmentation by blasting*, edited by W. L. Fournery et al., Bethel, CT, 1985.
- [Backers et al. 2004] T. Backers, G. Dresen, E. Rybacki, and S. O., "New data on mode II fracture toughness of rock from the punch-through shear test", *Int. J. Rock Mech. Min.* **41**:S1 (2004), 2–7. SINOROCK 2004 Symposium, Paper 1A 01.
- [Bobet 2001] A. Bobet, "A hybridized displacement discontinuity method for mixed mode I-II-III loading", *Int. J. Rock Mech. Min.* **38**:8 (2001), 1121–1134.
- [Broek 1989] D. Broek, *The practical use of fracture mechanics*, 4th ed., Kluwer, Dordrecht, 1989.
- [Carpinteri and Yang 1997] A. Carpinteri and G. Yang, "Size effects in brittle specimen with microcrack interaction", *Comput. Struct.* **63**:3 (1997), 429–437.

- [Cho et al. 2004] S. H. Cho, Y. Nakamura, and K. Kaneko, "Dynamic fracture process analysis of rock subjected to a stress wave and gas pressurization", *Int. J. Rock Mech. Min.* **41**:S1 (2004), 433–440. Paper 2A 20, SINOROCK2004 Symposium.
- [Courtesen 1979] D. L. Courtesen, "Cavities and gas penetration from blasts in stressed rock with flooded joints", *Acta Astronaut.* **6**:3–4 (1979), 341–363.
- [Crawford and Curran 1982] A. M. Crawford and J. H. Curran, "Higher order functional variation displacement discontinuity elements", *Int. J. Rock Mech. Min.* **19**:3 (1982), 143–148.
- [Crouch 1976] S. L. Crouch, "Solution of plane elasticity problems by the displacement discontinuity method, I: Infinite body solution", *Int. J. Numer. Methods Eng.* **10**:2 (1976), 301–343.
- [Crouch and Starfield 1983] S. L. Crouch and A. M. Starfield, *Boundary element methods in solid mechanics*, Allen and Unwin, London, 1983. MR 84i:73060 Zbl 0528.73083
- [Donzé et al. 1997] F. V. Donzé, J. Bouchez, and S. A. Magnier, "Modeling fractures in rock blasting", *Int. J. Rock Mech. Min.* **34**:8 (1997), 1153–1163.
- [Erdogan and Sih 1963] F. Erdogan and G. C. Sih, "On the crack extension in plates under plane loading and transverse shear", *J. Basic Eng. (ASME)* **85** (1963), 519–527.
- [Griffith 1925] A. A. Griffith, "The theory of rupture", pp. 56–63 in *Proceedings of the First International Congress for Applied Mechanics* (Delft, 1924), edited by C. B. Biezeno and J. M. Burgers, Waltman, Delft, 1925.
- [Guo et al. 1990] H. Guo, N. I. Aziz, and R. A. Schmidt, "Linear elastic crack tip modeling by the displacement discontinuity method", *Eng. Fract. Mech.* **36**:6 (1990), 933–943.
- [Guo et al. 1992] H. Guo, N. I. Aziz, and R. A. Schmidt, "Rock cutting study using linear elastic fracture mechanics", *Eng. Fract. Mech.* **41**:5 (1992), 771–778.
- [Huang and Wang 1985] J.-A. Huang and S. Wang, "An experimental investigation concerning the comprehensive fracture toughness of some brittle rocks", *Int. J. Rock Mech. Min.* **22**:2 (1985), 99–104.
- [Ingraffea 1981] A. R. Ingraffea, "Mixed-mode fracture initiation in Indiana sandstone and Westerly granite", pp. 199–204 in *Rock mechanics from research to application: proceedings of the 22nd US Symposium on Rock Mechanics* (MIT), edited by H. H. Einstein, MIT, Cambridge, MA, 1981.
- [Ingraffea 1983] A. R. Ingraffea, "Numerical modeling of fracture propagation", pp. 151–208 in *Rock fracture mechanics*, edited by H. P. Rossmannith, Courses and Lectures / International Centre for Mechanical Sciences **275**, Springer, Vienna, 1983.
- [Ingraffea 1987] A. R. Ingraffea, "Theory of crack initiation and propagation in rock", pp. 71–110 in *Fracture mechanics of rock*, edited by B. K. Atkinson, Academy Press, London, 1987.
- [Kutter and Fairhurst 1971] H. K. Kutter and C. Fairhurst, "On the fracture process in blasting", *Int. J. Rock Mech. Min.* **8**:3 (1971), 181–188.
- [Mortazavi and Katsabanis 2001] A. Mortazavi and P. D. Katsabanis, "Modeling burden size and strata dip effects on surface blasting process", *Int. J. Rock Mech. Min.* **38** (2001), 481–498.
- [Ouchterlony 1983] F. Ouchterlony, "Analysis of cracks", pp. 31–67 in *Rock fracture mechanics*, edited by H. P. Rossmannith, Springer, Vienna, 1983.
- [Ouchterlony 1988] F. Ouchterlony, "Suggested methods for determining the fracture toughness of rock: ISRM commission on testing methods", *Int. J. Rock Mech. Min.* **25**:2 (1988), 71–96.
- [Rao et al. 2003] Q. Rao, Z. Sun, O. Stephansson, C. Li, and B. Stillborg, "Shear fracture (mode II) of brittle rock", *Int. J. Rock Mech. Min.* **40**:3 (2003), 355–375.
- [Rossmannith 1983] H. P. Rossmannith, *Rock fracture mechanics*, Springer, New York, 1983.
- [Scavia 1992] C. Scavia, "A numerical technique for the analysis of cracks subjected to normal compressive stresses", *Int. J. Numer. Methods Eng.* **33**:5 (1992), 929–942.
- [Scavia 1995] C. Scavia, "A method for the study of crack propagation in rock structures", *Géotechnique* **45** (1995), 447–463.
- [Shen et al. 2004] B. Shen, O. Stephansson, M. Rinne, H.-S. Lee, L. Jing, and K. Roshoff, "A fracture propagation code and its application to nuclear waste disposal", *Int. J. Rock Mech. Min.* **41**:S1 (2004), 472–477.
- [Shou and Crouch 1995] K. J. Shou and S. L. Crouch, "A higher order displacement discontinuity method for analysis of crack problems", *Int. J. Rock Mech. Min.* **32**:1 (1995), 49–55.

- [Stephansson 2002] O. Stephansson, “Recent rock fracture mechanics developments”, pp. 37–60 in *Proceedings of the First Iranian Rock Mechanics Conference* (Tehran), 2002.
- [Stephansson et al. 2001] O. Stephansson, T. Backers, G. Dresen, and E. Rybacki, “Shear fracture mechanics of rocks and a new testing method for K_{IIC} ”, pp. 163–168 in *Rock mechanics: a challenge for society: proceedings of the ISRM Regional Symposium: EUROCK 2001* (Espoo, Finland), edited by P. Sarkka and P. Eloranta, Taylor and Francis, London, 2001.
- [Tan et al. 1996] X. C. Tan, S. Q. Kou, and P. A. Lindqvist, *Simulation of rock fragmentation by indenters using DDM and fracture mechanics*, edited by M. Aubertin et al., A. A. Balkema, Rotterdam, 1996.
- [Whittaker et al. 1992] B. N. Whittaker, R. N. Singh, and G. Sun, *Rock fracture mechanics: principles, design and applications*, Developments in Geotechnical Engineering **71**, Elsevier, Amsterdam, 1992.
- [Zipf and Bieniawski 1989] R. K. Zipf and Z. T. Bieniawski, “Development of the mixed mode testing system for geological materials”, pp. 338–352 in *Fracture of concrete and rock: SEM-RILEM International Conference* (Houston, TX, 1987), edited by S. P. Shah and S. E. Swartz, Springer, New York, 1989.

Received 29 Nov 2006. Accepted 29 Nov 2006.

HASAN HOSSEINI_NASAB: hhn@yazduni.ac.ir

Industrial Engineering Department, Engineering Faculty, Yazd University, Safaiyeh, Yazd, Iran

MOHAMMAD FATEHI MARJI: mfatehi@yazduni.ac.ir

Mining Engineering Department, Engineering Faculty, Yazd University, Safaiyeh, Yazd, Iran

IN-SITU OPTIMIZED PWAS PHASED ARRAYS FOR LAMB WAVE STRUCTURAL HEALTH MONITORING

LINGYU YU AND VICTOR GIURGIUTIU

The use of piezoelectric wafer active sensors (PWAS) phased arrays for Lamb wave damage detection in thin-wall structures is presented. The PWAS capability to tune into specific Lamb-wave modes (which is an enabling factor for our approach) is first reviewed. Then, a generic beamforming formulation that does not require the conventional parallel-ray approximation is developed for PWAS phased arrays in connection with the delay-and-sum beamforming principles. This generic formulation is applied to a 1-D linear PWAS phased array. Particularly, 1-D PWAS array beamforming reduces to the simplified parallel ray algorithm when the parallel ray approximation is invoked. The embedded ultrasonic structural radar (EUSR) algorithm is presented. A couple of simple experiments are used to show that the linear EUSR PWAS phased array system can successfully detect cracks in large aluminum thin plates.

To improve the EUSR image quality, advanced signal processing is studied for possible integration into the EUSR system. The approaches include Hilbert transform for envelope detection, thresholding techniques for removing background noise, discrete wavelet transform for denoising, continuous wavelet transform for single frequency component extraction, and cross-correlation for time-of-flight detection.

The optimization of linear PWAS arrays is studied next. First we consider the effect of several parameters affecting the phased-array beamforming: (1) number of elements M ; (2) elementary spacing d ; (3) steering angle ϕ_0 ; (4) location of the target r . Second, we examine the so-called nonuniform PWAS arrays which are generated by assigning different excitation weights to each of the array elements. The design of two nonuniform linear PWAS arrays, the binomial array and the Dolph–Chebyshev array, is presented. Significant improvement of the EUSR image is observed when using these nonuniform arrays.

1. Introduction

The current use of nondestructive evaluation (NDE) technologies is limited by the fact that the NDE inspection can be carried out only if the area to be inspected is accessible. In most cases, this can be achieved only during maintenance checks. For economic reasons, more frequent inspections or even continuous monitoring are not feasible with the existing NDE techniques [Beral and Speckmann 2003]. One method of conducting in-service monitoring of structural hot spots is through structural health monitoring (SHM) using guided waves and in-situ sensors. Guided waves can travel large distance with very little amplitude loss and inspect large area from a single transducer position [Rose 1995; 1999; 2001]. In-situ sensors, if sufficiently light and reliable, would offer on-demand structural interrogation capabilities.

Keywords: phased array, piezoelectric wafer, active sensors, embedded ultrasonics structural radar, beamforming, Lamb waves, damage detection, structural health monitoring, EUSR, PWAS, SHM, NDE.

Financial support from the National Science Foundation, grants NSF CMS-0408578 and NSF CMS-0528873 and from Air Force Office of Scientific Research grant FA9550-04-0085 are thankfully acknowledged.

1.1. Motivation. Traditionally, guided waves are generated by ultrasonic transducers that obliquely impinge onto a thin-wall structure. However, such ultrasonic transducers are unsuitable for in-situ SHM due to their cost, weight, and size. They cannot be attached in large numbers onto a structure without incurring important cost and weight penalties. The advent of commercially-available low-cost piezoelectric wafers has opened new opportunities for ultrasonic testing. Piezoelectric wafer active sensors (PWAS) are inexpensive lightweight unobtrusive transducers, which are well suited for SHM applications. These devices can be permanently attached to the structures for the generation and detection of guided waves [Giurgiutiu and Zagrai 2000]. Typical PWAS weigh 68 mg, are 0.2 mm thick, and cost less than US\$10.

The advantages of using a phased array of transducers for ultrasonic testing are multiple [Moles et al. 2005]. Ultrasonic phased arrays use ultrasonic elements and electronic time delays to create wave beams by constructive wave interference. Rather than using a single transducer, the phased array utilizes a group of transducers located at distinct spatial locations. By sequentially firing the individual elements of an array at slightly different times, the ultrasonic wave front can be focused or steered in specific directions [Krautkramer and Krautkramer 1990]. The relative amplitudes of the signals radiated by the individual array elements determine the effective radiation pattern of the array. If the amplitude at a certain distance is the same for all directions, the wave field of the ultrasonic element is called omnidirectional. An array constructed of such elements, will have an effective radiation pattern that is uniquely determined by the beam steering algorithm alone. Inspection of a wide area can be achieved from a single location by electronically sweeping and/or refocusing the wave beam without physically manipulating the transducer. The backscattered ultrasonic signals received and stored in the computer can be numerically analyzed and mapped into an inspection image.

Current ultrasonic phased array technology employs pressure waves generated by ultrasonic transducers through normal impingement on the structural surface. Such phased arrays have shown clear advantages in the inspection of very thick specimens and in the sidewise inspection of thick slabs, where electronic beam scanning and focusing have produced significant improvements in the inspection efficiency. However, they cannot be efficiently used in thin-wall structures because of the small relative thickness of such structures. New transducers are therefore needed for efficient Lamb wave SHM using phased array technology.

PWAS phased arrays are capable of in-situ scanning of a large structural area using guided waves but without the need for actual physical scanning being performed. This advantage stems from the capability of PWAS phased array to see large areas from a single location [Giurgiutiu and Bao 2002]. A permanently mounted array of unobtrusive PWAS transducers was shown to map an entire half-plate and detect a small crack using the *embedded ultrasonics structural radar* (EUSR) methodology [Giurgiutiu et al. 2006]. The EUSR image resembles the C-scan of conventional ultrasonic surface scanning but without the need for actual physical motion of the transducer over the structural surface. Building onto these earlier results, this paper presents the results of a sustained effort to improve the EUSR-PWAS phased array concept through a generic phased-array formulation and the use of advanced signal processing methods. Phased-array design optimization is also discussed for better wave beam steering. This effort has resulted in increased detection capability, refined detection resolution, and extended detection range [Giurgiutiu and Yu 2006].

1.2. Background. The PWAS phased array technique is based on two important concepts: the use of PWAS for Lamb wave generation and detection, and the tuning between Lamb waves and PWAS that permits the transmission and reception of selective Lamb-wave modes, as appropriate for phased-array implementation. These concepts are briefly reviewed in this section.

PWAS coupled Lamb waves. When guided waves travel between two parallel surfaces which are sufficiently close, for example, a plate having a thickness of the order of a wavelength or so, they are called Lamb waves [Viktorov 1967]. In our work, Lamb waves were excited and detected using PWAS. PWAS operates on the piezoelectric principle that couples the electrical and mechanical variables in the material

$$S_{ij} = s_{ijkl}^E T_{kl} + d_{kij} E_k, \quad D_j = d_{jkl} T_{kl} + \epsilon_{jk}^T E_k,$$

where s_{ijkl}^E is the mechanical compliance of the material measured at zero electric field ($E = 0$), ϵ_{jk}^T is the dielectric permittivity measured at zero mechanical stress ($T = 0$), and d_{kij} represents the piezoelectric coupling effect. In Lamb wave application, PWAS couple their in-plane motion with the Lamb wave particle motion on the material surface. As an active sensor, PWAS can be used as transmitters and receivers to excite and detect Lamb waves. A surface mounted PWAS will couple with both the symmetric Lamb wave modes (S_0, S_1, S_2, \dots) and the antisymmetric Lamb wave modes (A_0, A_1, A_2, \dots). Figure 1 illustrates PWAS interaction with the S_0 and A_0 Lamb wave modes. Similarly to conventional

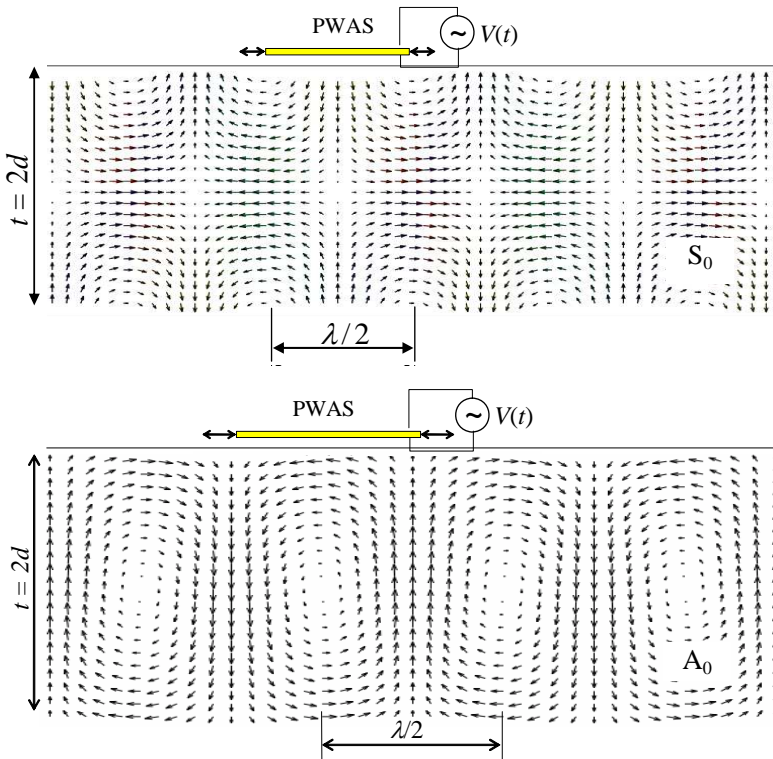


Figure 1. PWAS interaction with Lamb modes: top, symmetric mode S_0 ; bottom, anti-symmetric mode A_0 [Giurgiutiu 2005].

ultrasonic transducers, PWAS can be used in pitch-catch, pulse-echo, phased array, etc. An extensive description of the use of PWAS for SHM can be found in [Giurgiutiu et al. 2006].

Lamb wave mode selection through PWAS tuning. Lamb waves are dispersive, that is, their phase velocity varies with the frequency, f , and the thickness of the material, $2d$. In addition, for any given frequency, at least two modes are present. The simultaneous presence of two or more Lamb wave modes complicates the use of Lamb waves for detecting structural damage. Besides, certain damage is more detectable with certain Lamb wave modes. In particular, the S0 wave mode is better for the detection of through-the-thickness cracks and A0 mode is better for corrosion and disbonding. For PWAS phased array application, it is better to selectively excite low-dispersion Lamb wave modes, such as S0 at low fd value. Given the PWAS length $2a$, wall thickness $2d$, and material properties μ and λ , it is possible to find tuning frequencies at which only one mode is excited; see [Giurgiutiu 2005]. A plane-strain analysis of the PWAS-structure interaction using the space-domain Fourier analysis was developed in the same paper to illustrate the principle of PWAS Lamb-wave mode tuning and open a path for a more comprehensive analysis. The analysis yielded the forward wave solution in the form

$$\varepsilon_x(x, t) = -j \frac{a\tau_0}{\mu} \sum_{\xi^S} \sin \xi^S a \frac{N_S(\xi^S)}{D'_S(\xi^S)} e^{j(\xi^S x - \omega t)} d\xi - j \frac{a\tau_0}{\mu} \sum_{\xi^A} \sin \xi^A a \frac{N_A(\xi^A)}{D'_A(\xi^A)} e^{j(\xi^A x - \omega t)} d\xi, \quad (1)$$

where τ_0 is the interfacial shear stress and ξ^S, ξ^A are the symmetric and antisymmetric wave numbers. The expressions for N_S, D_S, N_A, D_A are detailed in [Giurgiutiu 2005] and will not be repeated here. The summations in (1) cover all the symmetric (ξ^S) and antisymmetric (ξ^A) Lamb wave modes that exist for a given value of ω in a given structure. The function $\sin(\xi a)$ in (1) displays maxima when the PWAS length $2a$ equals an odd multiple of the half wavelength. Several such maxima and minima exist, each associated with a certain Lamb wave mode and its wavelength. These maxima and minima allow us to achieve Lamb mode tuning. A selected Lamb mode can be tuned by choosing the appropriate frequency for given PWAS dimensions. A plot of the strain solution in the $0 \sim 700$ kHz bandwidth for 7-mm square PWAS installed on a 1-mm thick aluminum plate is presented in Figure 2, left, while the right half of the figure gives the dispersion curves. An S0 tuning frequency can be found around 210 kHz, where the amplitude of the A0 mode is minimized while that of the S0 is still strong. In this way, we achieve tuning of the S0 mode and reject the A0 mode. Lamb wave tuning offers considerable advantages by allowing us to select the Lamb waves that are most appropriate for the particular application.

2. Lamb waves PWAS phased array design and implementation

One can employ various methods to steer and focus ultrasonic waves in certain directions. Among these methods, delay-and-sum beamforming is the oldest and simplest [Johnson and Dudgeon 1993]. The delay-and-sum beamforming was implemented in PWAS phased arrays as the embedded ultrasonic structural radar (EUSR) algorithm [Giurgiutiu et al. 2006; Yu and Giurgiutiu 2005b]. EUSR assumes that data from an M -element PWAS array is collected in a round-robin fashion by using one element at a time as transmitter and all the elements as receivers. With a total of M^2 data signals, EUSR conducts the beam scanning in virtual time as a signal post-processing operation. The EUSR beamforming and scanning procedure does not require complex devices, nor multichannel electronic circuitry as needed by the conventional ultrasonic phased array equipment. EUSR implementation of PWAS phased arrays

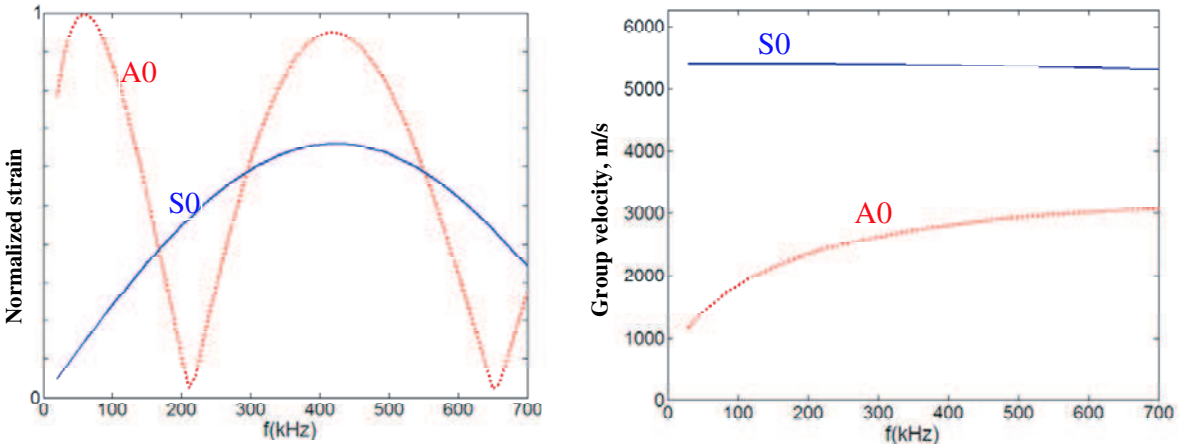


Figure 2. Lamb wave S0 mode tuning. Left: predicted Lamb wave strain curve of a 1 mm aluminum plate installed with a 7 mm square PWAS. Right: dispersion curves within 0 ~ 700 kHz range.

requires only a function generator, an oscilloscope, a switching device, and a computer [Giurgiutiu et al. 2006].

The original EUSR algorithm uses the parallel rays approximation to simplify the beamforming calculation. This simplifying assumption is only valid if the target is far away from the phased array. If the scanning field is not sufficiently far away, the parallel ray approximation error is not negligible and the method breaks down. This can be explained as follows. When the target is not sufficiently far away from the array, that is, it is in the near field (Figure 3a), the propagating wave front is curved (circular wave front) and the wave propagating directions vary with the PWAS location. In this situation, wave propagation direction varies from PWAS element to PWAS element and individual direction vectors need to be assigned to each PWAS. Using antenna theory [Johnson and Dudgeon 1993], we define the near

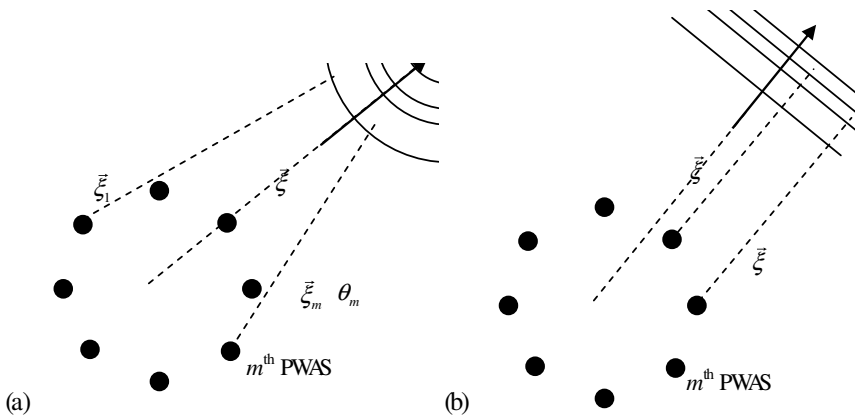


Figure 3. Wave fronts of a PWAS array: (a) near field; (b) far field.

field and far field, respectively, as the regions in which

$$0.62\sqrt{\frac{D^3}{\lambda}} < R_{\text{near}} \leq \frac{2D^2}{\lambda}, \quad R_{\text{far}} > \frac{2D^2}{\lambda}, \quad (2)$$

where D is the array aperture and λ is the excitation wavelength. Only when the target meets the far field criterion, the wave fronts are plane waves and wave propagation directions are the same for all of the array elements.

In the region $R \leq 0.62\sqrt{D^3/\lambda}$, phased array theory is no longer valid; however, other methods for damage detection such as PWAS electromechanical impedance technique [Zagrai and Giurgiutiu 2001] can be employed.

2.1. Generic beamforming formulation for PWAS Lamb waves phased arrays. We developed generic PWAS beamforming formulas for wave beam steering at any point either in the near field or in the far field. The following assumptions are made:

- All the PWAS elements in the array lie in the same plane and behave as point-wise sources and receivers (omnidirectional);
- Monochromatic excitation and reception, without considering the dispersion of guided waves and the waves are propagating at a constant speed c in an isotropic material;
- Simultaneous excitation along the array elements.

The beamforming onto a target $P(\vec{r}, \phi)$ using an M -PWAS phased array with elements $\{\vec{s}_m\}$ is considered. The origin of coordinate system is defined in the phase center, that is, $\sum \vec{s}_m/M = 0$, for $m = 0, 1, \dots, M-1$. For all waves arriving at the target P , the delay-and-sum beamforming consists of two steps:

- (1) applying a delay Δ_m and a weighting factor w_m to the propagating wave from the m -th PWAS, $f(\vec{r}_m, t)$;
- (2) summing up the output signals of the total of M PWAS.

This procedure can be expressed as

$$z(\vec{r}, t) = \sum_{m=0}^{M-1} w_m f(\vec{r}_m, t - \Delta_m).$$

For a single-tone radial wave, the wave front at a point \vec{r} away from the source can be expressed as

$$f(\vec{r}, t) = \frac{A}{\sqrt{|\vec{r}|}} e^{j(\omega t - \vec{k} \cdot \vec{r})} \quad (3)$$

with \vec{k} is the wave number, $\vec{k} = \vec{\xi} \cdot \omega/c$, and ω is wave frequency of the wave. For an M -element array, we apply (3) to each array element m , where $m = 0, 1, \dots, M-1$ (see Figure 4). The m -th element is assumed at location \vec{s}_m , whereas the direction vector from m -th element to the target is defined as $\vec{\xi}_m$. The following notations apply

$$\vec{\xi} = \frac{\vec{r}}{r}, \quad r = |\vec{r}|, \quad \vec{k} = \vec{\xi} \cdot \frac{\omega}{c}, \quad \vec{r}_m = \vec{r} - \vec{s}_m, \quad r_m = |\vec{r}_m|, \quad \vec{\xi}_m = \frac{\vec{r}_m}{|\vec{r}_m|}, \quad \vec{k}_m = \vec{\xi}_m \cdot \frac{\omega}{c},$$

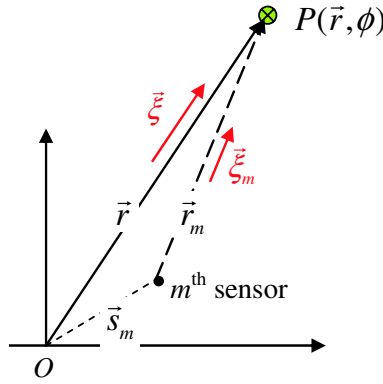


Figure 4. Schematic of the geometry of the m -th PWAS and the target at $P(\vec{r}, \phi)$.

where \vec{k}_m is the wave number of the wave propagating in the direction of $\vec{\xi}_m$.

The synthetic wave front arriving at target $P(\vec{r}, \phi)$ resulting from the superposition of the waves generated by the M sources can be written as

$$z(\vec{r}, t) = \sum_{m=0}^{M-1} w_m f(\vec{r}_m, t) = \sum_{m=0}^{M-1} w_m \cdot \frac{A}{\sqrt{r_m}} e^{j[\omega t - \vec{k}_m \cdot \vec{r}_m]}. \tag{4}$$

Generic beamforming: the triangular algorithm. Equation (4) is rewritten in the form

$$z(\vec{r}, t) = f\left(t - \frac{r}{c}\right) \cdot \sum_{m=0}^{M-1} w_m \frac{1}{\sqrt{r_m/r}} e^{j\omega \frac{r-r_m}{c}}, \tag{5}$$

where the first multiplier represents the wave front emitting from a single source at the origin and is independent of the array elements. This wave is to be used as a reference for calculating the needed time delay for each elementary wave. The second multiplier is a summation which is the total effects caused by individual sources, various weighting factors and locations. Equation (5) shows that the synthetic signal $z(\vec{r}, t)$ is controlled by the multiplier

$$\sum_{m=0}^{M-1} w_m \frac{1}{\sqrt{r_m/r}} e^{j\omega \frac{r-r_m}{c}},$$

which depends on the weighting factors w_m , the location r of the target, and the locations \vec{s}_m of the PWAS sources. We identify this multiplier as the *beamforming factor*, BF , given by

$$BF(w_m, r, s_m) = \sum_{m=0}^{M-1} w_m \frac{1}{\sqrt{r_m/r}} e^{j\omega \frac{r-r_m}{c}}. \tag{6}$$

Equation (6) is the direct beamforming of a group of M -PWAS fired simultaneously. To achieve beamforming in a certain direction ϕ_0 , we apply certain delays to the signals from each element. The delays

are defined as $\Delta_m(\phi_0)$, $m = 0, 1, \dots, M - 1$. Then, the beamforming formula in (6) gives

$$BF(w_m, r, s_m, \phi_0, \Delta_m(\phi_0)) = \sum_{m=0}^{M-1} w_m \frac{1}{\sqrt{r_m/r}} e^{j\omega\left(\frac{r-r_m}{c} - \Delta_m(\phi_0)\right)},$$

where r_m is a function of s_m and ϕ_0 . Beamforming is achieved when the delays $\Delta_m(\phi_0)$ are chosen in such a way as to maximize BF . We observe that a maximum of $BF(w_m, r, s_m, \phi_0, \Delta_m(\phi_0))$ is achieved if the exponential in summation equal to one, that is,

$$\frac{r - r_m}{c} - \Delta_m(\phi_0) = 0 \quad (\text{beamforming condition}), \tag{7}$$

which yields the delays to be applied to each element:

$$\Delta_m(s_m, \phi_0) = \frac{r(\phi_0) - r_m(s_m, \phi_0)}{c}, \quad m = 0, 1, \dots, M - 1. \tag{8}$$

Notice that the delays not only depend on the desired maximum beamforming direction, but also depend on the location of the PWAS sources. Using the delays $\Delta_m(s_m, \phi_0)$, the beamforming factor at the particular direction ϕ_0 reaches a maximum, that is,

$$BF_{\max}(w_m, r, s_m, \phi_0) = \sum_{m=0}^{M-1} w_m \frac{1}{\sqrt{r_m(s_m, \phi_0)/r}}.$$

The maximum beamforming depends on the beamforming direction ϕ_0 , besides the weighting w_m , target $P(r, \phi)$, and PWAS source location s_m .

Further manipulation of the beamforming can be obtained by adjusting the weighting factors w_m . Thus far, we conclude that by applying proper time delays and weightings, the phased array beamforming at a desired direction can be achieved. When the beam steering direction ϕ_0 changes, the phased array beamforming will accomplish scanning.

Far field beamforming: the parallel algorithm. If the target meets the far field condition, the parallel ray assumption applies. The generic formulas developed in the previous section can be simplified to become independent of the target radial position. Since the propagation directions of the waves are now parallel to each other, we have

$$\vec{r}_m \parallel \vec{r}, \quad \vec{\xi}_m \approx \vec{\xi}, \quad \sqrt{r_m} \approx \sqrt{r}, \quad \vec{k}_m \approx \vec{\xi} \cdot \frac{\omega}{c} = \vec{k}, \quad m = 0, 1, \dots, M - 1.$$

After applying delay and weighting, the beamforming factor becomes

$$BF(w_m, \vec{s}_m, \phi_0) = \sum_{m=0}^{M-1} w_m e^{j\omega\left(\frac{\vec{\xi} \cdot \vec{s}_m}{c} - \Delta_m(\phi_0)\right)}. \tag{9}$$

For a particular direction ϕ_0 , the beamforming factor can be achieved by setting $\Delta_m(\phi_0) = \vec{\xi} \cdot \vec{s}_m/c$. Substituting that into (9) gives a beamforming factor that depends only on weight w_m , that is,

$$BF \max(w_m) = \sum_{m=0}^{M-1} w_m.$$

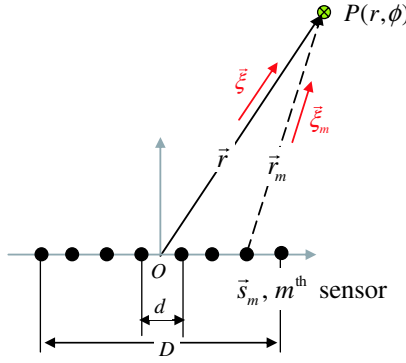


Figure 5. Schematic of an M -PWAS phased array. The coordinate origin is located in the middle of the array.

2.2. Directional beamforming of 1-D linear PWAS array. The generic PWAS phased array beamforming formulas derived in the previous section will be first applied to the simplest and also one of the most practical arrays, the 1-D linear array. This is achieved by placing the PWAS along a straight line. Here, a M -PWAS linear array uniformly spaced at d is considered, as illustrated in Figure 5. The span (aperture) D of the array is $D = (M - 1) \cdot d$. With the coordinate system origin located in the middle of the array, the location vector of the m -th element and the vector r_m are

$$\vec{s}_m = \left[d \left(m - \frac{M-1}{2} \right), 0 \right], \quad \vec{r}_m = \vec{r} - \vec{s}_m.$$

Recall from Equation (5) that the synthetic wave front received at $P(r, \phi)$ from the M -PWAS array is

$$z(\vec{r}, t) = f\left(t - \frac{r}{c}\right) \cdot \sum_{m=0}^{M-1} w_m \frac{1}{\sqrt{r_m/r}} e^{j\omega \frac{r-r_m}{c}}.$$

The first multiplier represents a wave emitting from the origin and it is independent of the array elements. This wave is to be used as a reference for calculating the needed time delay for each elementary wave. The second multiplier, which controls the array beamforming, can be simplified by normalizing r_m by the quantity r , resulting in the beamforming factor

$$BF(w_m, M) = \frac{1}{M} \cdot \sum_{m=0}^{M-1} w_m \frac{\exp\left(j \frac{2\pi}{\lambda} (1 - r_m)\right)}{\sqrt{r_m}}.$$

The scale factor $1/M$ is used to normalize the beamforming factor. By further introducing two new parameters, d/λ and r/d , the beamforming is rewritten as

$$BF(w_m, M, \frac{d}{\lambda}, \frac{r}{d}) = \frac{1}{M} \cdot \sum_{m=0}^{M-1} w_m \frac{\exp\left(j 2\pi \frac{d}{\lambda} \frac{r}{d} (1 - r_m)\right)}{\sqrt{r_m}}. \quad (10)$$

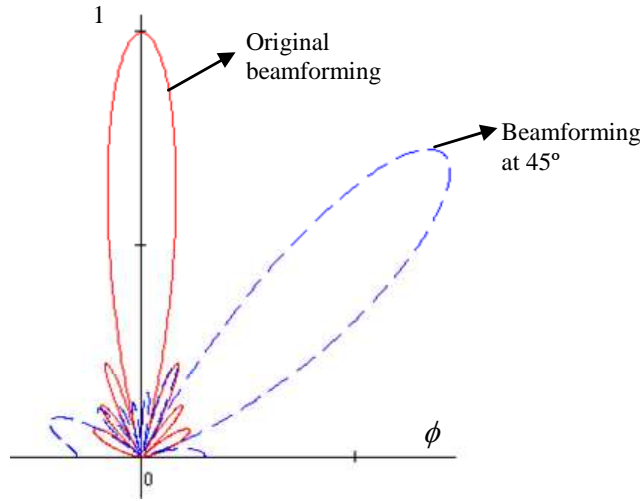


Figure 6. The original beamforming and directional beamforming at 45° of an 8 PWAS phased array with $d/\lambda = 0.5$, $r/d = 10$, $\{w_m\} = 1$.

For the far field situation, the simplified beamforming is independent of r/d , that is,

$$BF\left(w_m, M, \frac{d}{\lambda}\right) = \frac{1}{M} \cdot \sum_{m=0}^{M-1} w_m \exp\left(j2\pi \frac{d}{\lambda} \left(m - \frac{M-1}{2}\right) \cos \phi\right). \quad (11)$$

The beamforming factor in (10) and (11) has a maximum value for $\phi_0 = 90^\circ$. This is the inherent beamforming of the linear array. The inherent beamforming for an 8-PWAS array with $\{w_m\} = 1$, $d/\lambda = 0.5$, and $r/d = 10$ is shown in Figure 6 (solid line). Notice that, indeed, the maximum beam is obtained at 90° .

Now applying “delays” $\delta_m(\phi_0)$ to steer the beam towards a preferred direction ϕ_0 the beamforming is

$$BF\left(w_m, M, \frac{d}{\lambda}, \frac{r}{d}, \phi_0\right) = \frac{1}{M} \cdot \sum_{m=0}^{M-1} w_m \frac{\exp\left(j2\pi \frac{d}{\lambda} \frac{r}{d} (1 - r_m - \delta_m(\phi_0))\right)}{\sqrt{r_m}}. \quad (12)$$

The beamforming factor of (12) reaches its maximum in direction ϕ_0 when the delay $\delta_m(\phi_0)$ is

$$\delta_m(\phi_0) = 1 - r_m(\phi_0) = 1 - \frac{|\vec{r}(\phi_0) - \vec{s}_m|}{|\vec{r}(\phi_0)|}.$$

By changing the value of ϕ_0 from 0° to 180° , we can generate a scanning beam. Simulation results of the directional beamforming at $\phi_0 = 45^\circ$ are also shown in Figure 6.

2.3. PWAS array implementation: embedded ultrasonic structural radar (EUSR). The embedded ultrasonic structural radar (EUSR) methodology [Giurgiutiu and Bao 2002] was first developed under the parallel rays approximation and the assumption of a 1-D linear PWAS array permanently attached onto the structure. We generalize the original EUSR algorithm by implementing the beamforming of a generic array configuration using the exact wave propagation paths theory presented in previous sections.

For an M -PWAS array, with all elements fired simultaneously with the same excitation $s_T(t)$, the total signal arriving at target $P(r, \phi_0)$ is

$$s_P(t) = \sum_{m=0}^{M-1} w_m \frac{1}{\sqrt{r_m}} s_T\left(t - \frac{r}{c} + \frac{r - r_m}{c}\right), \quad (13)$$

where $1/\sqrt{r_m}$ represents the decay due to the omnidirectional 2-D wave radiation, r/c is the time traveling to the target from the origin O , and $(r - r_m)/c$ is the time to the target from the m -th element. Wave-energy conservation, that is, no dissipation, is assumed. If the elements are fired with certain delays rather than simultaneously, that is, $\Delta_m(\phi) = (r - r_m)/c$, Equation (13) becomes

$$s_P(t) = s_T\left(t - \frac{r}{c}\right) \sum_{m=0}^{M-1} w_m \frac{1}{\sqrt{r_m}}.$$

This shows a factor of the sum over the signal magnitude of the individual excitation $s_T(t)$.

If the delay $\Delta_m(\phi)$ is taken at $\phi = \pm\phi_0$, a maximum transmitting wave directed to the target $P(r, \phi_0)$ is thus obtained through implementing delays in the firing of the elements in the array.

After the transmission signals arrive at target P , they will be scattered and picked up by the array. The PWAS transducers serve now as receivers. The signal received at the m -th PWAS will arrive quicker by $\Delta_m(\phi) = (r - r_m)/c$. To synchronize all the received signals, we simply need to delay them with $\Delta_m(\phi) = (r - r_m)/c$. Assume that at $P(r, \phi_0)$ the incoming signal is backscattered with a backscatter coefficient A ; thus, the signal received at each PWAS will be

$$\frac{1}{\sqrt{r_m}} s_T\left(t - \frac{r}{c} + \frac{r - r_m}{c}\right) A \sum_{m=0}^{M-1} w_m \frac{1}{\sqrt{r_m}}.$$

The receiver beamforming is obtained by assembling all the signals arriving at the same time, that is,

$$s_R(t) = \left(\sum_{m=0}^{M-1} w'_m \frac{1}{\sqrt{r_m}} \right) s_T\left(t - \frac{r}{c}\right) A \sum_{m=0}^{M-1} w_m \frac{1}{\sqrt{r_m}}, \quad (14)$$

where appropriate delays $\Delta_m(\phi) = (r - r_m)/c$ were used, and w'_m as the weighting for reception beamforming. Thus, the assembled received signal is further scaled by the factor $\sum_{m=0}^{M-1} w'_m / \sqrt{r_m}$.

If the location of target $P(r, \phi_0)$ is indicated by the angle ϕ_0 , the coarse estimation of ϕ_0 can be implemented by using the ϕ_0 sweeping method. For linear PWAS array, the EUSR algorithm scans through 0° to 180° by incrementing ϕ_0 at 1° each time, and then finding the direction where the maximum received energy, $\max E_R(\phi_0)$, is obtained. The received energy, $E_R(\phi_0)$ is defined as

$$E_R(\phi_0) = \int_{t_p}^{t_p+t_a} |s_R(t)|^2 dt.$$

2.4. Proof-of-concept laboratory experiments. A smoothed toneburst signal of carrying frequency f_c was used to excite S0 Lamb wave mode in the PWAS array experiments. The PWAS Lamb-wave tuning principles [Giurgiutiu 2005; Bottai and Giurgiutiu 2005] were used to achieve the tuning frequency. The general form of the tone-burst excitation was $s_T(t) = s_0(t) \cos 2\pi f_c t$, where $s_0(t)$ is a short-duration

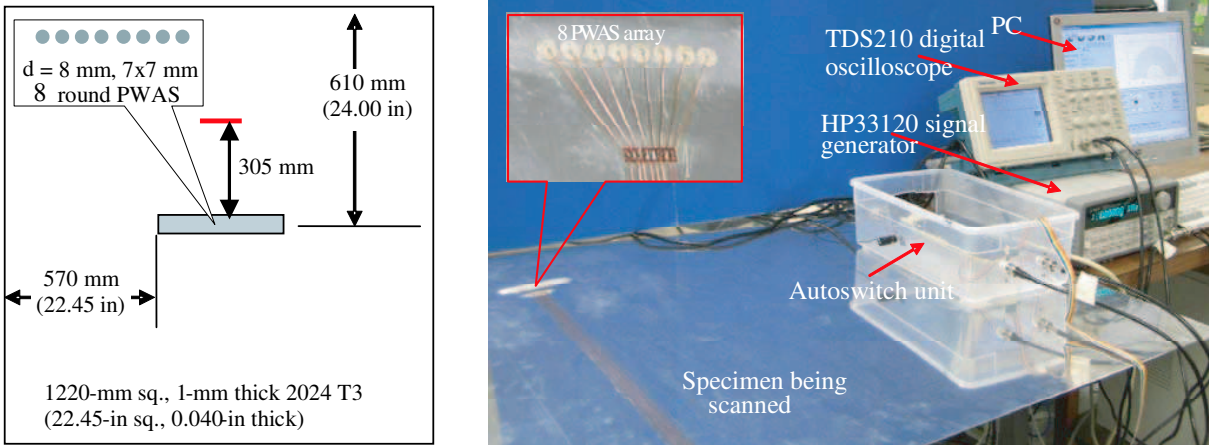


Figure 7. EUSR experiment using an 8-PWAS array: specimen layout and experiment setup.

smoothing Hanning window between 0 and t_p that is applied to the carrier signal of frequency f_c . The values t_p is calculated by $t_p = N_c/f_c$ where N_c is the count number. Since S0 is quasi-nondispersive for our fd range, we assume a constant wave speed c , and the wavelength λ was determined as $\lambda = c/f_c$.

The laboratory proof-of-concept experiments use a set of 1220 mm square 2024 T3 aluminum plates of 1 mm thickness with different damage types. A PWAS array consisting of eight 7-mm round PWAS spaced at 8 mm pitch was installed in the center of the plate, as illustrated in Figure 7, left. The damage was simulated by fully penetrated 1-mm wide and 20-mm long slits representing cracks and pin holes of various diameters. A 3-count toneburst signal was used as the excitation. The carrying frequency f_c was defined in the range 100 ~ 600 kHz until the S0 modes was tuned at 300 kHz.

The experimental setup (Figure 7, right) includes a HP33120 function generator to send the excitation and a TDS210 digital oscilloscope to collect the signals. The collected signals were stored in a computer through the GPIB interface. A round-robin pattern was used for signal excitation and collection. At each time, one PWAS acted as transmitter to send out the excitation signal and all the PWAS served as receivers to pick up the reflection signals.

Then all the elements in the array serve as transmitters in turn. By this means, when the round-robin data collection is done, a total of M^2 signals are recorded. The switch between transmitters and receivers is implemented through the PWAS-ASCU unit described in [Liu and Giurgiutiu 2005]. A typical raw signal recorded in the experiments is shown in Figure 8. The signal-to-noise ratio for the raw signal is found to be 7.6.

The delay-and-sum beamforming procedure was processed as a post processing routine to generate a virtual scanning beam using the improved EUSR algorithm. This is advantageous and different from the traditional phased arrays instruments which have high device requirements for sending out all the phased signals simultaneously through parallel channels.

The virtual scanning result from the EUSR is visualized as a 2-D image by mapping the wave magnitude. In addition, a display of an A-scan signal at a user selected direction is also provided. Figure 9 gives an example of using the PWAS array EUSR to scan a specimen (plate #1) with a broadside crack located

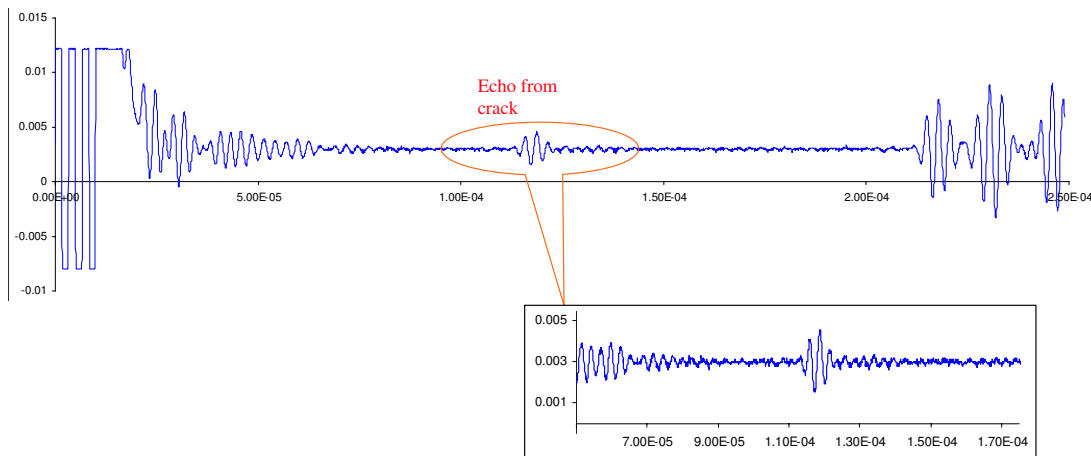


Figure 8. A typical raw signal collected in the proof-of-concept experiment with close-up plot of echo from crack.

about 305 mm in front of the array. The excitation frequency was tuned to 300 kHz where the S0 wave was dominant (Figure 9c). At this frequency, the S0 wave propagates at speed 5440 m/s with $\lambda = 18$ mm. The dark stripe seen in Figure 9b at position of 90° and about 305 mm away from the array indicated the presence of the broadside crack. Figure 9d gives the A-scan signal at 90° and 30°. Note here that an observation window is used to maximize the display by removing the initial bang and the reflections

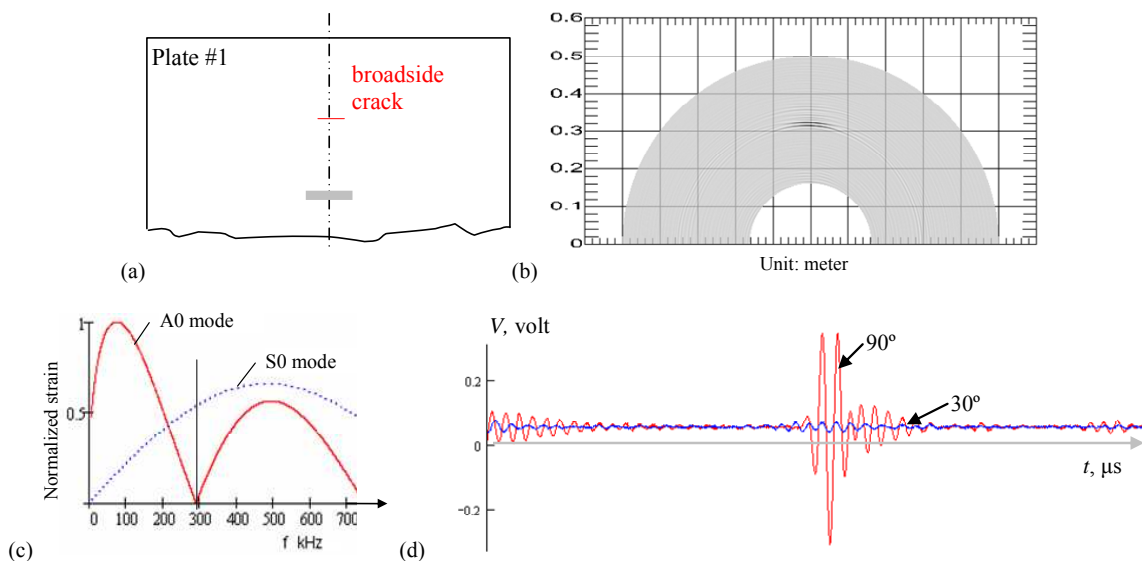


Figure 9. Proof-of-concept PWAS EUSR scanning experiment for detecting a broadside crack: (a) schematic of half a plate with a broadside crack; (b) mapped EUSR scanning image; (c) S0 mode tuning curve for using 8 mm round PWAS in a 1 mm thick aluminum plate; (d) selected A-scan at 90°.

from the boundaries. The estimated signal-to-noise ratio for the crack echo was signal-to-noise ratio of approximately 38, a great enhancement when compared with the raw signal.

The scanning image obtained from the experiment shows that EUSR can correctly show the presence of the crack and indicates that the crack is centered at 90° and around 305 mm in front of the array. These results match quite well the actual situation. However, there are a couple of dark stripes on the picture in Figure 9b and the estimation of the crack size is rather coarse. Further work was found needed to improve the image quality, as shown in the next section.

3. Advanced signal processing for EUSR

Observing the A-scan signals in Figure 9d, it is realized that the multiple lines shown in scanned image is actually caused by the multiple local maxima in the rectified signal. Such an influence can be removed by rectifying the signal and replacing it with its envelope. Moreover, background noise is present at all the time, resulting the circular ripples in the scanned image. To remove these problems, we added the Hilbert transform and thresholding to the post-processing of the EUSR system. These modules we inserted immediately after the beamforming algorithm and before the visualization section. In addition, we applied denoising and filtering at a particular frequency component using wavelet analysis and time-of-flight detection through cross correlation methods.

3.1. Essential signal processing. Two methods, the Hilbert transform and the thresholding process, are used as essential signal processing to enhance the EUSR image quality.

Envelope extraction using Hilbert transform. Envelope represents the absolute amplitude of a signal. It can be easily extracted through the Hilbert transform. Hilbert transform is widely used to construct an analytical signal which has the envelope of the original signal. The Hilbert transform of a signal $x(t)$ is defined as (see [Poularikas 2006, Chapter 10])

$$H(x(t)) = -\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} dt.$$

To build an analytical signal $\tilde{x}(t)$, the original signal $x(t)$ is employed to construct the real part, the Hilbert transform $H(x(t))$ is employed to construct the imaginary part, that is, $\tilde{x}(t) = x(t) + jH(x(t))$. The analytical signal $\tilde{x}(t)$ has the property that it has the same envelope as the original signal, that is, $|\tilde{x}(t)| = |x(t)|$. Hence we can find the envelop of the original signal $x(t)$ by taking the magnitude of the analytical signal $\tilde{x}(t)$. Figure 10 demonstrates the envelope of an A-scan signal recorded at 90° extracted with the Hilbert transform [Yu and Giurgiutiu 2005a]. The original echo of the crack shows many local maxima, which introduces difficulties in automatically detecting the location of the crack (as the case shown in Figure 9b). After extracting the envelope, there is only one peak to consider, and peak detection method can be easily applied.

Thresholding. As shown in Figure 9b, some background noise shows up in the image, resulting in the ripples in the image. To remove these ripples, a thresholding process is included after the Hilbert transform. By this means, noise signal below the threshold value will be discarded. The original and improved EUSR scanning image of the broadside crack specimen are shown in Figure 11.

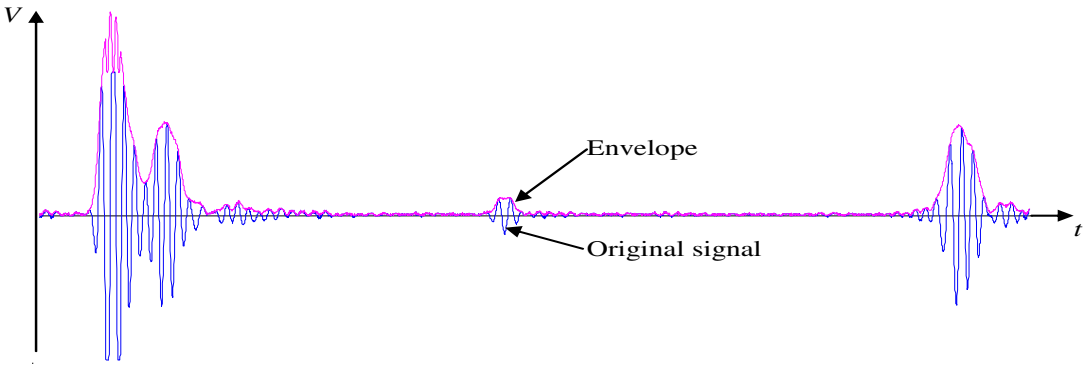


Figure 10. Envelope extraction by the Hilbert transform.

3.2. Additional advanced signal processing. As options, more modules have been developed for more signal processing objective as well. These modules include denoising by discrete wavelet transform, single frequency component filtering by continuous wavelet transform, and precise measurement of TOF by cross correlation method. Details of this development can be found in Yu and Giurgiutiu [2005a]. An overview is given next.

Denoising by discrete wavelet transform (DWT). Based on multiresolution analysis (MRA), the discrete Wavelet transform (DWT) provides a tool for decomposing signals into a set of elementary mutually-orthogonal building blocks, called wavelets. The DWT is defined as

$$c_{m,n} = \int_{-\infty}^{+\infty} x(t)\Psi_{m,n}(t)dt, \quad \Psi_{m,n}(t) = 2^{-m/2}\Psi(2^{-m}t - n),$$

where $\Psi_{m,n}(t)$ are orthonormal wavelets obtained by shifting and dilating a mother wavelet $\Psi(t)$. The coefficients $c_{m,n}$ are usually considered as the result of a filter being applied to the signal; the filter is working as a highpass filter and a lowpass filter to extract the signal’s approximation and detail information, respectively. After each filtering process, half of the frequencies of the original signal are removed. At the same time, half the signal samples can be discarded by a downsampling process of a factor of 2

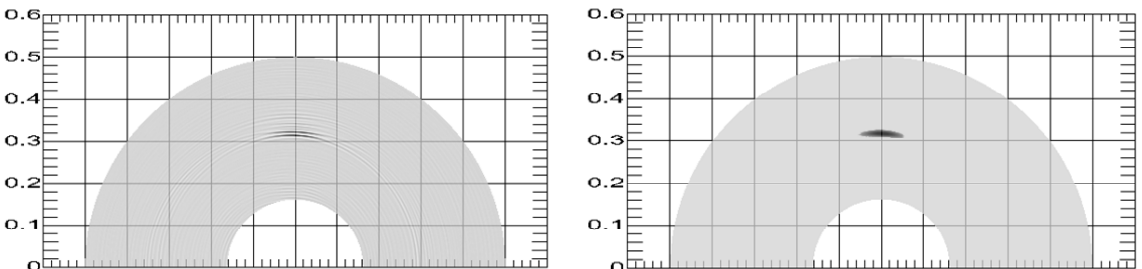


Figure 11. EUSR inspection results: original image (left) and improved image after Hilbert transform and thresholding process (right).

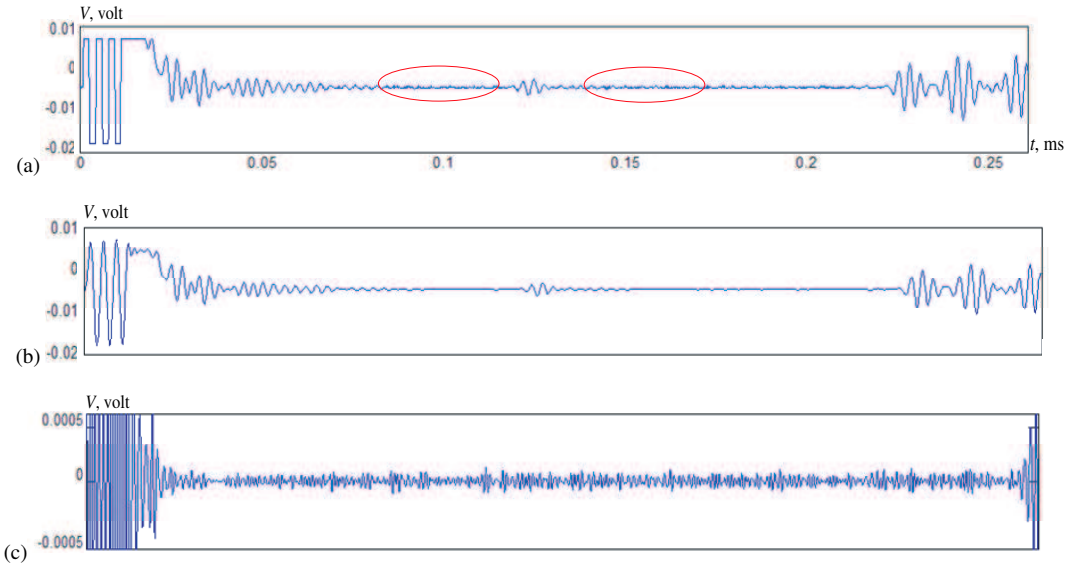


Figure 12. A 3-level DWT denoising using Mayer wavelet: (a) original signal from broadside experiment; (b) denoised clean signal; (c) removed noise.

consistent with the Nyquist rule. Such a decomposition process will continue for l times (called DWT level) until data remaining is within the user-defined allowable error range, defined in the filter bank theory [Boggess and Narcowich 2001]. Therefore, the original signal is separated into two parts, the approximation part, that is, the denoised signal, and the details part, that is, the removed noise. As an example, a raw signal recorded in the broadside experiment was processed by a 3 level DWT using the Mayer wavelet. Figure 12a shows the processed original data. Note that, in this proof-of-concept experiment, the noise level is not severe. After DWT denoising, certain noise (shown in Figure 12c at a zoomed scale) was removed and a much cleaner signal was obtained (Figure 12b). The small ripples were significantly reduced (see the circled segments between wave packets). The large components at the beginning of the noise caused by truncation in initial bang, which were not removed, but they will be later discarded through time windowing.

Filtering by continuous wavelet transform (CWT). Another way to remove disturbance is offered by the continuous wavelet transform (CWT). The CWT of signal $x(t)$ using the mother wavelet $\psi(t)$ is

$$CWT(a, \tau) = \frac{1}{\sqrt{|a|}} \int x(t) \psi^* \left(\frac{t-\tau}{a} \right) dt, \quad (15)$$

where a is the scale (or dilation) factor, τ is the translation (or time shift) of the wavelet with respect to the signal, and the factor $1/\sqrt{|a|}$ is introduced for energy normalization at different scales. The resulting scale-frequency representation of the magnitude squared, $|CWT(a, \tau)|^2$, is the scalogram

$$|CWT(a, \tau)|^2 = \frac{1}{|a|} \left| \int x(t) \psi^* \left(\frac{t-\tau}{a} \right) dt \right|^2.$$

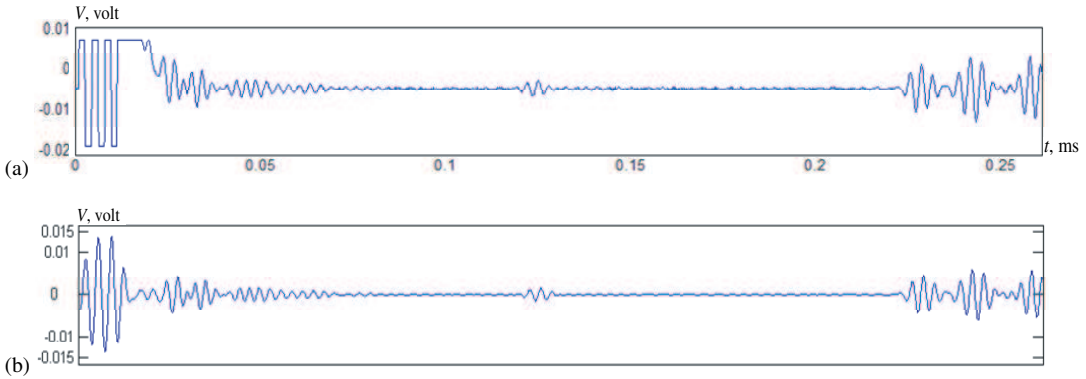


Figure 13. CWT filtering at 300 kHz using Mayer wavelet: (a) original signal; (b) CWT filtered result.

The relation between the scale and the filtering frequency f is

$$f = \frac{\text{dimensionless center frequency of the wavelet}}{(\text{scale}) \times (\text{sampling interval})}. \quad (16)$$

With Equations (15) and (16), we see that a certain frequency component of interest can be extracted by performing the CWT at the corresponding single scale. We named this process CWT filtering. In PWAS array application, though excitation at tuned frequency is employed, the signal contains multifrequency components due to tone-burst characteristics (having certain bandwidth). With CWT filtering, only the frequency component at the wavelet level corresponding to the excitation frequency is used in the reception signals, reducing the influence of spectrum spread and dispersion. By this means, disturbances from other frequencies can be removed. Figure 13 gives the filtering result of the same raw signal as used in the DWT denoising. The CWT filtering used the Mayer wavelet to extract the frequency component at 300 kHz, which is the excitation frequency. Filtered result is shown in Figure 13b. In addition to the reduced background noise, it is interesting to observe that the 300 kHz component in the initial bang was also extracted with more focused energy. This illustrates the advantage of CWT over DWT filtering.

TOF detection by cross correlation. Cross correlation is widely used to detect similarities in two signals. The cross correlation $R_{xy}(m)$ of two discrete signals $x(n)$ and $y(m)$ of length N is defined by

$$R_{xy}(m) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)y(n-m).$$

When the input signal $x(n)$ slides along the time axis of the correlated signal $y(m)$ with a small step, the similarity of the overlapped part of two signals is compared. If the two signals are completely not similar (unrelated), the correlation coefficient is 0, while if they are identical, the correlation coefficient is 1. For real signals, the cross correlation method can reduce the noise since the noise is not related to the signals and not auto-related either.

In our application, since the reflection caused by the damage resembles the original excitation, its arrival time (TOF) can be extracted using the cross correlation method. Figure 14 illustrates this idea. A signal received at PWAS (Figure 14b) is correlated with the 3-count toneburst excitation (Figure 14a),

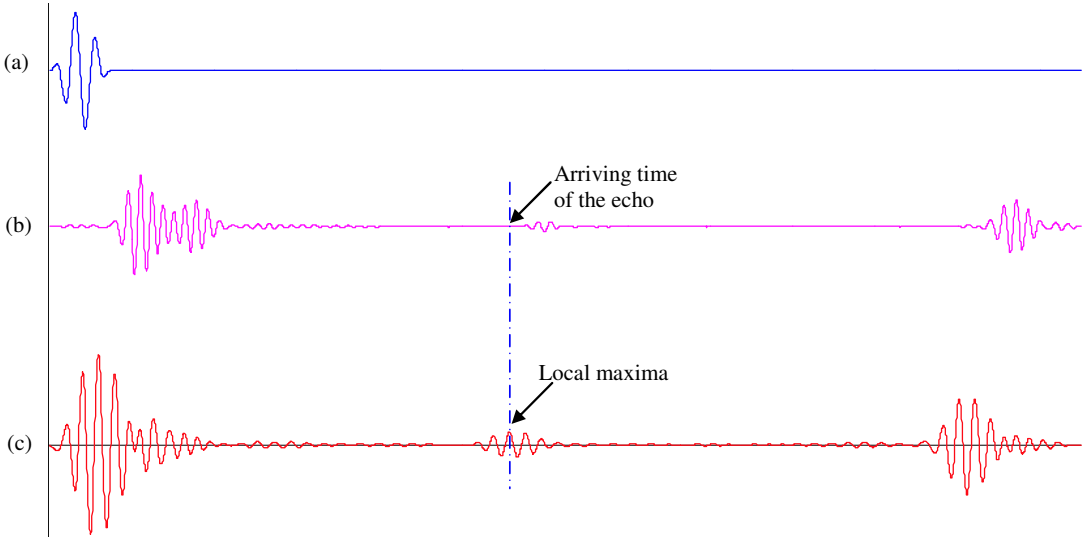


Figure 14. Cross-correlation for detecting TOF of the echo from the crack: (a) window-smoothed tone-burst excitation; (b) received signal; (c) cross-correlation coefficients of the signals.

generating a sequence of cross correlation coefficients (Figure 14c). The local maxima represent the time instants when the wave packets have largest similarity with the toneburst, that is, at the arrival time of the wave packet. To locate the TOF of echo from the crack, a time window can be used to cut off influences from the initial bang and boundary reflections. After the TOF is obtained, the location of the crack l can be estimated using the wave propagation speed c , that is, $l = c \cdot \text{TOF}$.

3.3. Experiments with improved EUSR. To further confirm EUSR detection ability, more experiments were conducted on different specimens. These specimens include: (1) plate #2 with a broadside crack inclined at 30° ; (2) plate #3 with a horizontally positioned offside crack located at 137° ; (3) plate #4 with a broadside pin hole of increasing diameter (0.5 mm, 1 mm, 1.57 mm, 2 mm); (4) plate #5 with two aligned horizontally positioned offside cracks at 67° and 117° , respectively. Specimen schematics are shown in Figure 15. All plates have the same dimension as the one used in the proof-of-concept experiment and the perpendicular distances from the damage to the array are 301 mm for all the plates.

To compare the imaging of the horizontal broadside crack specimen and the 30° -inclined crack specimen, these EUSR images are presented side-by-side in Figure 16. It is noticed that the image of the inclined crack is a little further away from the array. This is because some incipient waves were reflected away due to the slope. The inclined-crack footprint in the EUSR image is also smaller than the parallel crack. This leads to the idea of an effective crack length, $L_e = L \cos \alpha$, where $\alpha = 30^\circ$. The slope of the crack was not clearly indicated in Figure 16, right. We believe that this is due to the lack of resolution inherent in the 300 kHz S0 Lamb wave used in the detection, which is with the crack length ($\lambda_{S0} \approx 16$ mm at 300 kHz).

The pin hole damage detection in plate #4 was started from diameter 0.5 mm and then gradually enlarged to 1 mm, 1.57 mm, and 2 mm (in accordance with available drill bit sizes). EUSR failed to

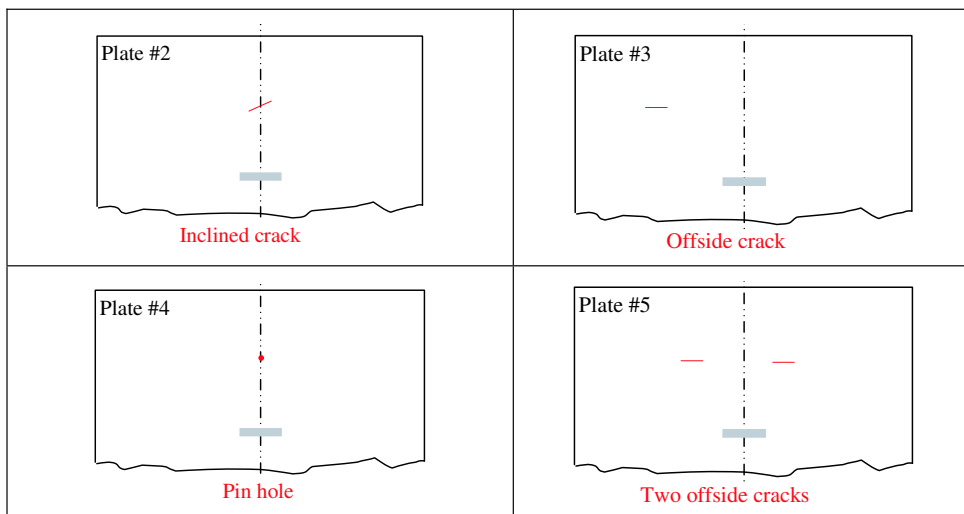


Figure 15. EUSR proof-of-concept experimental specimen schematics (all cracks are 20 mm long, 1 mm wide).

detect the pin hole of 0.5 mm diameter. Neither did it detect the pin hole of 1 mm diameter. When the hole was enlarged to 1.57 mm, EUSR gave a precise indication for the pin hole location, as shown in Figure 17, left. We considered this dimension as the minimum detectable damage dimension at this inspection frequency. The EUSR image of the 2 mm pin hole was even stronger, as shown by Figure 17, right, which is more intense than the left counterpart for the 1.57-mm pin hole.

Figure 18, left, shows the EUSR scanning image of the plate #3, which has a single offside crack at 137° . Since the crack orientation is parallel to the array alignment, most of the incipient waves are reflected away and no specular reflection is present. However, diffraction at crack tips creates sufficient back scatter waves to permit detection. The resulting EUSR image shows the presence of the crack, verifying that EUSR PWAS phased array can detect cracks which are located in nonoptimal positions. The scanning of plate #5, which has two offside cracks at 67° and 117° , is shown in Figure 18, right.

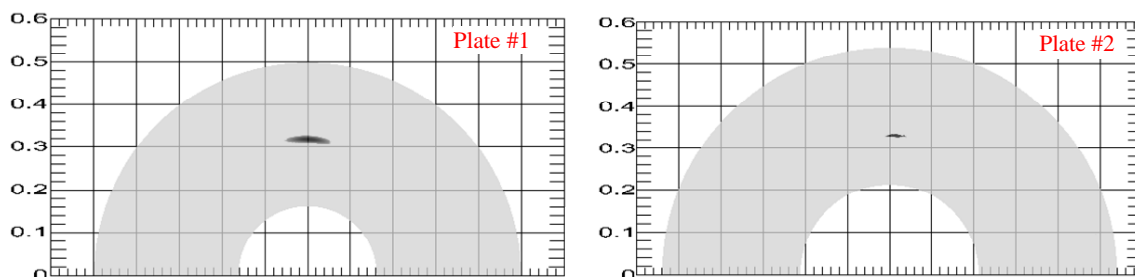


Figure 16. EUSR scanning images of the broadside crack specimens. Left: EUSR image of plate #1 with a broadside crack. Right: EUSR image of plate #2 with an inclined broadside crack.

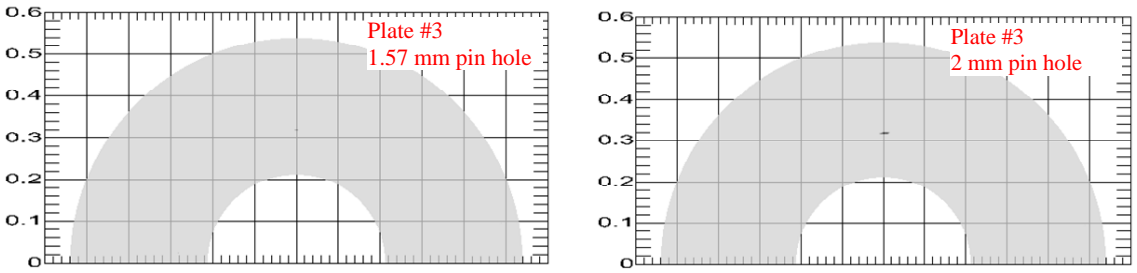


Figure 17. EUSR scanning images of the pin hole specimen. Left: EUSR image of the 1.57 mm pin hole. Right: EUSR image of the 2 mm pin hole.

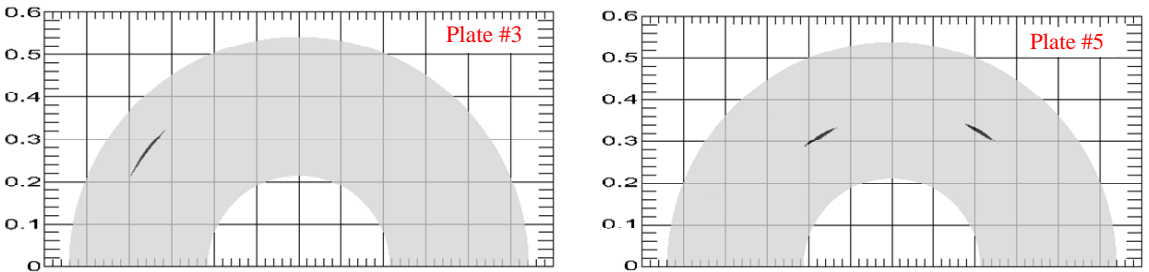


Figure 18. EUSR scanning images of the offside crack specimens. Left: EUSR image of plate #3 with an offside crack at 137° . Right; EUSR image of plate #5 with two offside cracks at 67° and 137° , respectively.

Both cracks are clearly detected in the image with correct center positions. This proves the multicrack detection capabilities of the EUSR PWAS method.

4. PWAS phased array optimization

The PWAS phased array beamforming at certain direction ϕ_0 is affected by several parameters:

- spacing between neighboring PWAS, d ;
- number of PWAS, M ;
- steering angle, ϕ_0 ;
- weighting factors, $\{w_m\}$.

Among these parameters, the effect of spacing d is always measured by the wavelength λ . Since $\lambda = c/f$, when the wavelength λ changes with the frequency, the ratio of d/λ will also change. For the triangular beamforming, there is an extra parameter needed to be considered, the ratio of r/d . Therefore, the effect of parameter d can be represented by the ratio d/λ and the ratio r/d . If all the weighting factors are the same ($\{w_m\} = \text{constant}$), that is, all PWAS in the array are excited uniformly, this type of phased array is called uniform array. Otherwise, it is called nonuniform array. We will use the

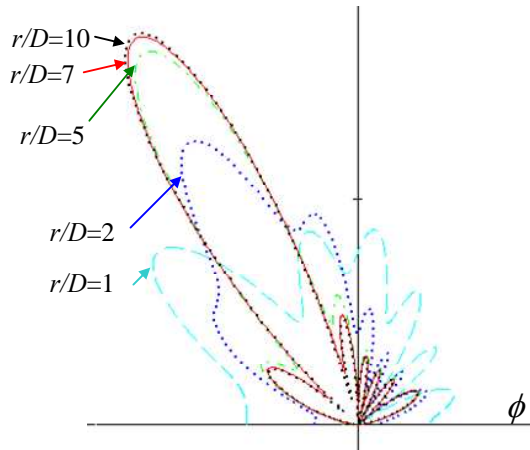


Figure 19. Beamforming at 120° of an eight PWAS phased array spacing at 0.5λ with various r/D value.

triangular beamforming algorithm to explore how these parameters affect the uniform PWAS phased array beamforming and then find out how the beamforming is modified by the weighting factors.

4.1. Uniform PWAS phased array. First we consider an M -element uniform PWAS array equally spaced at pitch d with equal weights, $w_m = 1$.

Effect of the r/d ratio. The ratio r/d determines whether a target is located in the far field or the near field of the phased array. If the target is in the near field, we have to use the exact triangular algorithm. Otherwise, we use the parallel ray approximation. To better quantify this ratio, the array span D is used instead of d using $D = (M - 1)d$. The field definitions of (2) can be expressed in terms of r/D

$$0.62\sqrt{(M-1)\frac{d}{\lambda}} < \left(\frac{r}{D}\right)_{\text{near}} \leq 2(M-1)\frac{d}{\lambda}, \quad \left(\frac{r}{D}\right)_{\text{far}} > 2(M-1)\frac{d}{\lambda}.$$

These definitions show that, for a particular application of an eight PWAS array spaced at half wavelength ($d = \lambda/2$), the near field is contained in the interval $r \in (1.16D, 7D]$, while the far field is the outside area ($r > 7D$). Beamforming simulations with r/D value evolving from 1, to 2, 5, 7, and 10 are shown in Figure 19 for a PWAS array directed at 120° .

The simulation used parameters from the laboratory experiments, that is, aluminum plate of 1 mm thickness and 3-count toneburst at $f_c = 300$ kHz. Figure 19 shows that very close to the array field ($r/D = 1$), directional beamforming does no longer exist (dashed line). However, as the target moves away from the array (entering near field, $r/D = 2$), directional beamforming starts to take shape. The beamforming is getting better when far-field conditions are approached ($r/D = 5$). In far field, where parallel rays approximation applies, the effect of r/D vanishes. No significant difference can be noticed between $r/D = 7$ and $r/D = 10$.

Effect of the d/λ ratio. The ratio d/λ shows the influence of spacing on array beamforming. Simulation results of an eight PWAS array directed at 120° with various d/λ values are shown in Figure 20. By

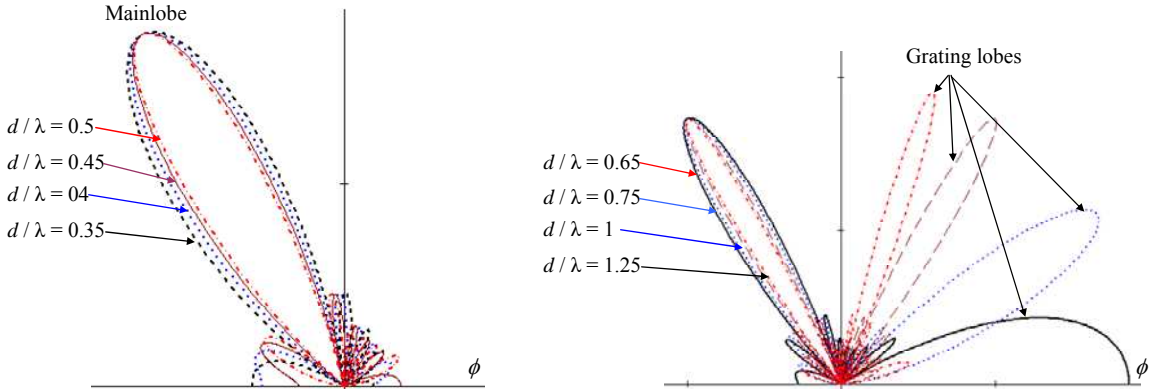


Figure 20. Beamforming at 120° of an eight-PWAS array with $r/D = 10$ using various d/λ : (a) $d/\lambda \leq 0.5$; (b) $d/\lambda > 0.5$.

comparing the beams, it can be seen that the beam width becomes smaller and smaller as d/λ increases. However, the number of sidelobes also increases. For larger d/λ value, narrow beam width (better resolution/directivity) is achieved, but more sidelobe disturbing are also presented.

It is also noticed that, for beamforming at $d/\lambda = 0.75$ and beyond (Figure 20b), besides the mainlobe at the desired angle of 120° , there are other strong and disturbing lobes showing up at other undesired directions. Such lobes are called grating lobes; they are caused by spatial aliasing [Johnson and Dudgeon 1993]. Grating lobes can have magnitudes as big as the mainlobe magnitude. The grating lobes are not desired and should be avoided because they give misleading scanning results. According to the spatial sampling theorem [Johnson and Dudgeon 1993], in order to avoid spatial aliasing, the spacing d between elements should be smaller or equal to the half wavelength ($d/\lambda \leq 0.5$). Otherwise grating lobes may appear. In practical implementation, this rule should be verified after frequency tuning. For the 300 kHz tuning frequency used in the proof-of-concept experiments, the ratio d/λ has the value $d/\lambda = 0.44$ ($d = 8$ mm; $c = 5440$ m/s). Therefore, no grating lobes are were yet presented.

Effect of the number M of PWAS in the array. The discussion about the influence of the spacing pitch on the array beamforming indicates that larger spacing may give better directional beams despite of their byproduct, larger sidelobes. However, the spacing cannot be unlimitedly increased because of the spatial sampling theorem. The number of elements in the array is another factor that affects the beamforming. Figure 21 demonstrates how beamforming is modified by different M values, for example, $M = 8$ and $M = 16$. The comparison shows the beamforming of a 16-PWAS array yields a much narrower mainlobe and slightly stronger sidelobes than that of an 8-PWAS array. Increasing the number of elements is a simple way to enhance beamforming with the small penalty of larger sidelobes. However, in practice, more elements will result in wiring issue and will be limited by the available installation space.

Effect of the steering angle ϕ_0 . Steering angle ϕ_0 (beamforming direction) is another factor that affects the beamforming. Figure 22a and b shows beamforming at 0° , 30° , 60° , 90° , 120° , 150° directions using eight-PWAS and sixteen-PWAS are arrays, respectively.

An overall examination of Figure 22 indicates that the best beamforming is achieved at $\phi_0 = 90^\circ$, with a slender and focused beam. As the angle moves to either side of the 90° position, the beamforming

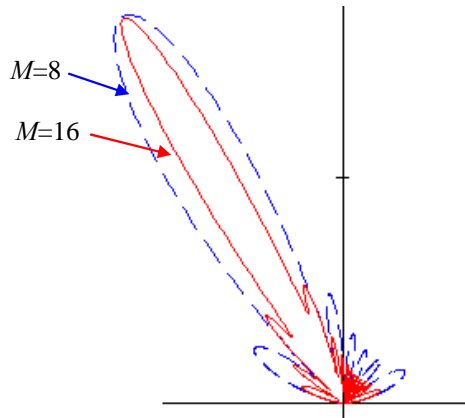


Figure 21. Beamforming at 120° of PWAS array with $r/D = 10$ and $d = 0.5\lambda$ using different M . The beam in dash is obtained with $M = 8$ and the beam in solid is obtained with $M = 16$.

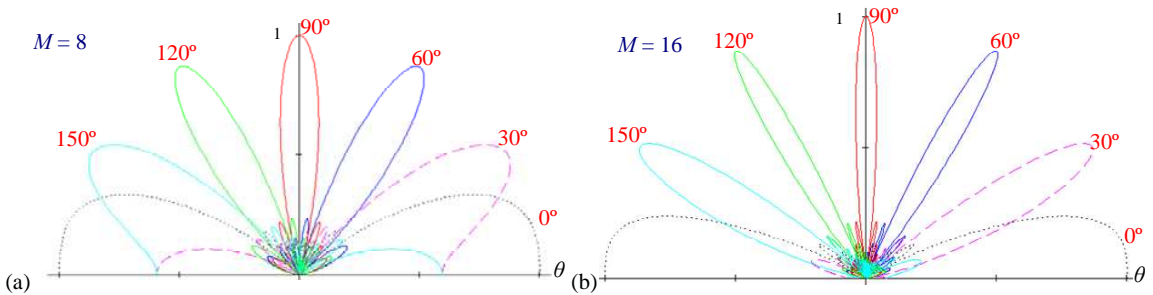


Figure 22. Beamforming at various steering angles with $D/r = 10$ and $d = 0.5\lambda$: (a) an eight PWAS array; (b) a sixteen PWAS array.

worsens. At 0° and 180° , the beamforming breaks down. In fact, at 0° , the beamforming split into two identical halves, one at 0° , the other at 180° . Such a lobe at 180° is called backlobe. As the direction increases to 30° and 60° , the back lobe shrinks and the main lobe becomes more directional. Beamforming at 120° and 150° are symmetrical about the vertical center to that at 60° and 30° . When the mainlobe gets closer to 180° (such as 150°), the back lobe starts again increasing. The directional mainlobe is completely lost again at 180° . This observations show that a linear PWAS array does not have a complete 180° view but a smaller one. The result in Figure 22b confirms the conclusion, but certain improvements due to the increased number of elements are apparent. Comparing the results for $M = 8$ to those for $M = 16$, we see that an array with larger M gives larger view area. No directional beam was obtained at 30° by the eight-PWAS array, but it was achieved by the sixteen-PWAS array.

4.2. Nonuniform PWAS phased array. Now we consider the case of having different excitations for the array's elements, that is, beamforming of a nonuniform PWAS array. The various excitations, if known,

can be processed as weighting factors and implemented in the EUSR algorithm, that is,

$$BF\left(w_m, M, \frac{d}{\lambda}, \frac{r}{d}, \phi_0\right) = \frac{1}{M} \cdot \sum_{m=0}^{M-1} w_m \frac{\exp\left(j2\pi \frac{d}{\lambda} \frac{r}{d} (1 - r_m - \delta_m(\phi_0))\right)}{\sqrt{r_m}}.$$

Two widely used distributions, the binomial distribution and the Dolph–Chebyshev distribution, will be used to determine the effect of relative excitation amplitudes in the array.

Binomial array. The coefficients w_m for a binomial array can be obtained by using the binomial expansion of the expression $(1 + x)^M$ [Johnson and Dudgeon 1993], that is,

$$(1 + x)^{M-1} = 1 + (M - 1)x + \frac{(M - 1)(M - 2)}{2!}x^2 + \frac{(M - 1)(M - 2)(M - 3)}{3!}x^3 + \dots .$$

The positive coefficients of the series at different values of M serve as the relative amplitude distribution w_m for an M -PWAS array. Such a nonuniform array is thereby named a binomial array. For $M = 8$, the amplitude weights are $\{1, 7, 21, 35, 35, 21, 7, 1\}$. Beamforming simulation results for a binomial array at different d/λ values are shown in Figure 23a. Larger d/λ yields thinner mainlobe, that is, better resolution/directivity. Figure 23a also shows that the beamforming of the binomial array has no sidelobes. Actually, this is the most significant characteristic of binomial arrays. Comparison with the beamforming of an equivalent uniform array is given in Figure 23b (at $d/\lambda = 0.5$). Although the binomial array has a wider mainlobe, it has no sidelobe at all which gives stronger signal suppression in undesired directions.

The disadvantage of using binomial array becomes, however, apparent if we look at directional beamforming away from 90° , for example, $\phi_0 = 45^\circ$. Figure 23c shows that the binomial array had already a deteriorated beamforming at 45° whereas the equivalent uniform array still has a directional beam. In addition, the zero sidelobe advantage is lost as we move away from vertical. The binomial arrays have

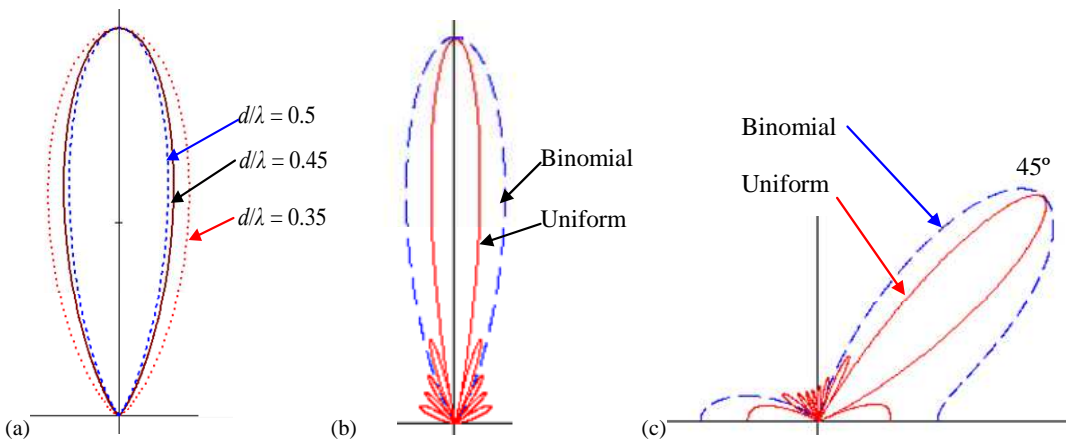


Figure 23. Beamforming of eight PWAS binomial array with $r/D = 10$: (a) beamforming at 90° of a binomial array with various d/λ values; (b) beamforming at 90° of a binomial array and an equivalent uniform array with $d/\lambda = 0.5$; (c) beamforming at 45° of a binomial and equivalent uniform array with $d/\lambda = 0.5$.

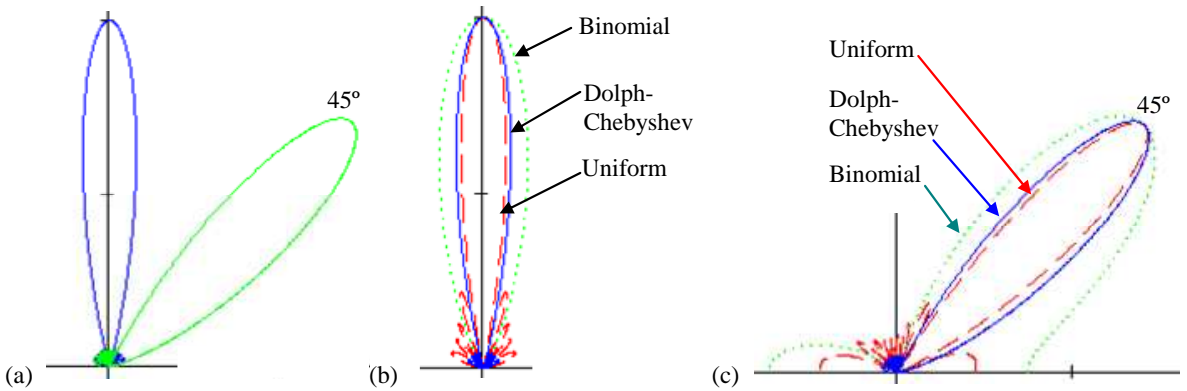


Figure 24. Beamforming of an eight PWAS Dolph-Chebyshev array with $r/D = 10$ and $d/\lambda = 0.5$: (a) beamforming at 90° and 45° ; (b) beamforming at 90° of all three arrays; (c) beamforming at 45° of all three arrays.

smaller view area than the equivalent uniform arrays and sidelobes may show up when departing from front focus.

Dolph-Chebyshev array. Though binomial arrays have the unique property of zero sidelobe at 90° , they give larger beam width and smaller view area. Another feasible nonuniform array is the Dolph-Chebyshev array utilizing the Chebyshev distribution. By assigning a sidelobe level, that is, the ratio of mainlobe magnitude to the first sidelobe magnitude, the Dolph-Chebyshev array coefficients can be derived [Balanis 2005]. To build an 8-PWAS Dolph-Chebyshev array, for a desired sidelobe level of 20, the normalized coefficients w_m are given by $\{0.357, 0.485, 0.706, 0.89, 1, 1, 0.89, 0.706, 0.485, 0.357\}$. Beamforming at 90° and 45° of this Dolph-Chebyshev array with $d/\lambda = 0.5$ and $r/D = 10$ is shown in Figure 24a. It can be seen that, with the designed sidelobe level, the sidelobes are significantly suppressed. If beamforming at 90° of all the three arrays are compared (Figure 24b), we see

- (1) considering the sidelobe level, the binomial array has the smallest sidelobes (zero sidelobe), followed by the Dolph-Chebyshev array while the uniform array has the highest sidelobe level;
- (2) considering the mainlobe width (directionality), the uniform array has the thinnest mainlobe and the Dolph-Chebyshev array is slightly larger, with the binomial array of largest width.

Beamforming at 45° is shown in Figure 24c. At this direction, the Dolph-Chebyshev array still has low sidelobes and directional beamforming. In summary, the Dolph-Chebyshev array is a good compromise between sidelobe level and mainlobe width. While sidelobes are suppressed, its directivity at off angles is still well maintained.

4.3. Experiments with weighted EUSR for nonuniform arrays. The simulation results presented in the previous section have shown that the Dolph-Chebyshev array is expected to have much smaller sidelobes when compared to the equivalent uniform array, while having almost the same mainlobe width. For the binomial array, it is expected to have much smaller sidelobe level but with enlarged mainlobe width. Using the experimental data from the broadside crack specimen collected with the 8-PWAS uniform

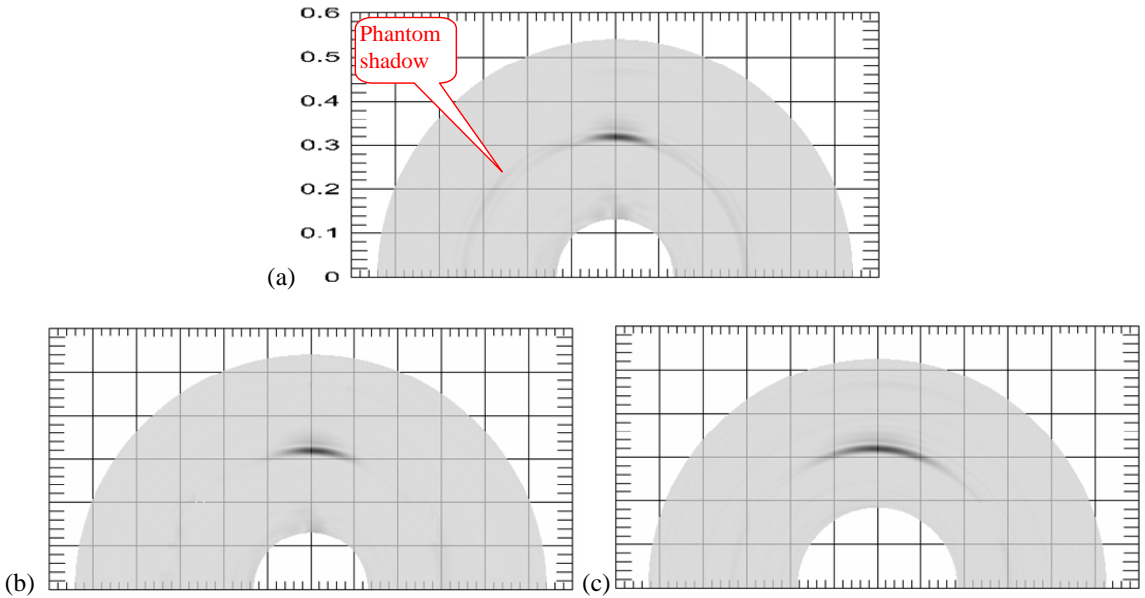


Figure 25. Crack detection using weighted EUSR algorithms on eight PWAS nonuniform arrays: (a) image of the original uniform array; (b) image of the Dolph–Chebyshev array; (c) image of the binomial array.

array, we implemented the nonuniform arrays with the weighted EUSR algorithm. Figure 25a gives the uniform array’s EUSR image before thresholding process. A phantom shadow is present as a ring in which the crack shows up. Recalling the beamforming simulation results in Figure 24b, we know that this is caused by the sidelobe effect. With either the Dolph–Chebyshev array or the binomial array, the corresponding weighted EUSR algorithm should give better images with the phantom removed or reduced. The scanning results of the two nonuniform arrays are shown in Figure 25b and Figure 25c, respectively.

Immediately we observe that the sidelobe phantom has been sufficiently suppressed in both images, as expected. But the crack size in the binomial array image is much wider than that of the uniform array. The wider crack image further verifies that the binomial array produces the largest mainlobe width among the three array considered in this study.

5. Conclusions

In this paper, we presented new developments of a novel SHM method, the embedded ultrasonic structural radar (EUSR) [Giurgiutiu et al. 2006]. The EUSR method uses piezoelectric wafer active sensors (PWAS) phased array to create directional Lamb wave for systematically interrogating large areas of a thin-wall structure. After a brief presentation of using PWAS transducers to send and receive Lamb waves, we introduced the frequency tuning technique that enables the application of phased array theory to PWAS coupled Lamb waves.

Original EUSR algorithm was developed on the parallel rays approximation in order to simplify the beamforming calculation. However, the simplifying assumption is only valid when the target is far away from the phased array and if the scanning field is not sufficiently far away, the method will deteriorate. To overcome this problem, a generic beamforming formula, called triangular algorithm, was developed by using exact wave traveling path for the calculation. Research results show that by applying proper time delays and weightings, the phased array beamforming at a desired direction can be achieved through this generic algorithm.

To verify the damage detection ability of PWAS phased array algorithm, a set of laboratory experiments using a linear 8-PWAS array were conducted to detect various damages in aluminum plates. The beamforming is implemented as the signal post-processing algorithm, known as the EUSR method. This post processing approach used in PWAS phase array implementation is advantageous over the conventional ultrasonic phased arrays methods since sophisticated multichannel electronic devices are not needed to precisely delay the individual signals going to the phased array elements. EUSR system collects data in a round-robin fashion, then performs the scanning in virtual time, and presents the result as a 2-D image. A proof-of-concept experiment was conducted and successfully verified the crack detection ability of the EUSR approach.

To improve the display image quality, several advanced signal processing methods were discussed for current or potential applications in EUSR routine. Signal preprocessing methods include the DWT denoising to remove the noise and CWT f_c -component filtering to eliminate the influence resulted from spectrum spreading and Lamb waves dispersion. The signal post-processing after EUSR algorithm includes two steps. First Hilbert transform is used to extract the A-scan signal envelope thus avoiding the ambiguity caused by multiple local peaks. Then the thresholding process is applied to remove the background noise. Thus, a clear image with sharp contrast can be obtained to indicate the presence of damage in the structure with better precision. To locate the position of the crack, TOF detection module using the cross-correlation method was also implemented. Several experiments were conducted to estimate the detection ability of the EUSR algorithm.

Beamforming properties of linear PWAS array were investigated. We found that several parameters can be used to manipulate the output beamforming. They are: ratio d/λ , ratio r/d , number of elements M , steering angle ϕ_0 , and weighting factors w_m . The effects are concluded as following:

- Larger d/λ results in thinner mainlobe width. However, the requirement of $d/\lambda \leq 0.5$ should be met to observe the spatial sampling theorem. Otherwise, grating lobes are present.
- Larger r/d ratio results in better beamforming for the triangular algorithm.
- Larger M results in thinner mainlobe though the sidelobe level gets larger which wiring may become an issue.
- The effective viewable area of linear arrays is less than 180° . The actual range is affected by number of elements and weighting.
- Weighting can further modify the beamforming, affecting, the mainlobe width, the sidelobes, and viewable area.

Experimentally, two types of nonuniform PWAS arrays were implemented, the Dolph–Chebyshev array and the binomial array. This was done by using a weighted EUSR algorithm. The results coincide

with the simulation results: (1) in terms of mainlobe width, uniform array is the best, followed by the Dolph–Chebyshev array while the binomial array yields the largest mainlobe; (2) in terms of sidelobes, the binomial array can achieve the “no sidelobe” effect at 90° when spaced at half wavelength, while the Dolph–Chebyshev array has manipulated sidelobe levels.

In a summary, the original contributions of this paper can be summarized as follows:

1. The paper presents an exact beamforming formulation using the exact wave traveling path rather than using the conventional parallel ray approximation.
2. The paper explores how the PWAS phased array beamforming is affected by several phased array parameters.
3. The paper presents how advanced signal processing methods can be used to improve the PWAS phased array performance.
4. The paper verifies the potential application of nonuniform PWAS phased array in the pattern of binomial array and Dolph–Chebyshev array.

Using proof-of-concept experiments, we have shown that the PWAS phased arrays implemented with the EUSR algorithm has great potential for in-situ structural health monitoring. To apply this technique to structures with complicated geometries or to detect small targets (weak reflectors) etc., more research needs to be done. The use of the generic PWAS array beamforming formulas can be applied to 2-D array configurations to obtain full range 360° scanning. Sustained theoretical research and technological development will continue to be conducted to more this technique toward practical SHM implementation.

References

- [Balanis 2005] C. A. Balanis, *Antenna theory analysis and design*, Wiley, Hoboken, NJ, 2005.
- [Beral and Speckmann 2003] B. Beral and H. Speckmann, “Structural health monitoring (SHM) for aircraft structures: a challenge for system developers and aircraft manufacturers”, pp. 12–29 in *Proceedings of Structural Health Monitoring 2003: From diagnostics and prognostics to structural health monitoring* (Stanford, CA), edited by F.-K. Chang, Destech, Lancaster, PA, 2003.
- [Boggess and Narcowich 2001] A. Boggess and F. Narcowich, *A first course in wavelets with Fourier analysis*, Prentice-Hall, Upper Saddle River, NJ, 2001.
- [Bottai and Giurgiutiu 2005] G. Bottai and V. Giurgiutiu, “Simulation of the lamb wave interaction between piezoelectric wafer active sensors and host structure”, pp. 259–270 in *Smart structures and materials 2005: sensors and smart structures technologies for civil, mechanical, and aerospace systems* (San Diego), edited by M. Tomizuka, Proceedings of SPIE **5765**, SPIE, Bellingham, WA, 2005.
- [Giurgiutiu 2005] V. Giurgiutiu, “Tuned lamb wave excitation and detection with piezoelectric wafer active sensors for structural health monitoring”, *J. Intell. Mater. Syst. Struct.* **16**:4 (2005), 291–306.
- [Giurgiutiu and Bao 2002] V. Giurgiutiu and J. Bao, “Embedded-ultrasonics structural radar for the nondestructive evaluation of thin-wall structures”, pp. paper # 2002–39017 in *Proceedings of the 2002 ASME International Mechanical Engineering Congress and Exposition* (New Orleans), ASME, New York, 2002.
- [Giurgiutiu and Yu 2006] V. Giurgiutiu and L. Yu, “Optimized embedded ultrasonics structural radar (EUSR) with piezoelectric wafer active sensors (PWAS) phased arrays for in-situ wide-area damage detection during SHM and NDE”, invention disclosure submitted to USC-USCRF, June 2006.
- [Giurgiutiu and Zagrai 2000] V. Giurgiutiu and A. Zagrai, “Characterization of piezoelectric wafer active sensors”, *J. Intell. Mater. Syst. Struct.* **11** (2000), 959–976.

- [Giurgiutiu et al. 2006] V. Giurgiutiu, J. Bao, and A. N. Zagrai, "Structural health monitoring system utilizing guided Lamb waves embedded ultrasonic structural radar", US Patent No. US6996480B2, Feb. 7 2006.
- [Johnson and Dudgeon 1993] D. H. Johnson and D. E. Dudgeon, *Array signal processing: concepts and techniques*, Prentice-Hall, Upper Saddle River, NJ, 1993.
- [Krautkramer and Krautkramer 1990] J. Krautkramer and H. Krautkramer, *Ultrasonic testing of materials*, vol. Berlin, 4th rev. ed., Springer, 1990.
- [Liu and Giurgiutiu 2005] W. Liu and V. Giurgiutiu, "Automation of data collection for PWAS-based structural health monitoring", pp. 1139–1147 in *Smart structures and materials 2005: sensors and smart structures technologies for civil, mechanical, and aerospace systems*, edited by M. Tomizuka, Proceedings of SPIE **5765**, SPIE, Bellingham, WA, 2005.
- [Moles et al. 2005] M. Moles, N. Dube, S. Labbe, and E. Ginzl, "Review of ultrasonic phased arrays for pressure vessel and pipeline", *J. Pressure Vessel Technol.* **127**:3 (2005), 351–356.
- [Poularikas 2006] A. D. Poularikas (ed.), *The transforms and applications handbook*, CRC Press, Boca Raton, FL, 2006.
- [Rose 1995] J. L. Rose, "Recent advances in guided wave NDE", pp. 761–770 in *1995 IEEE ultrasonics symposium*, 1995.
- [Rose 1999] J. L. Rose, *Ultrasonic waves in solid media*, Cambridge University Press, 1999.
- [Rose 2001] J. L. Rose, "A vision of ultrasonic guided wave inspection potential", pp. 1–5 in *Proceeding of the 7th ASME NDE topical conference*, NDE Topical Conference **20**, 2001.
- [Viktorov 1967] I. A. Viktorov, *Rayleigh and Lamb waves: physical theory and applications*, Plenum Press, New York, 1967.
- [Yu and Giurgiutiu 2005a] L. Yu and V. Giurgiutiu, "Advanced signal processing for enhanced damage detection with embedded ultrasonics structural radar using piezoelectric wafer active sensors", *Smart Struct. Syst.* **1**:2 (2005), 185–215.
- [Yu and Giurgiutiu 2005b] L. Yu and V. Giurgiutiu, "Improvement of damage detection with the embedded ultrasonics structural radar for structural health monitoring", pp. 1081–1090 in *Proceedings of Structural Health Monitoring 2005: Advances and challenges for implementation* (Stanford, CA), edited by F.-K. Chang, Destech, Lancaster, PA, 2005.
- [Zagrai and Giurgiutiu 2001] A. Zagrai and V. Giurgiutiu, "Electro-mechanical impedance method for crack detection in thin plates", *J. Intell. Mater. Syst. Struct.* **12**:10 (2001), 709–718.

Received 20 Jul 2006. Accepted 12 Jan 2007.

LINGYU YU: yu3@engr.sc.edu

Department of Mechanical Engineering, University of South Carolina, Columbia, SC 29208, United States

VICTOR GIURGIUTIU: giurgiut@engr.sc.edu

Department of Mechanical Engineering, University of South Carolina, Columbia, SC 29208, United States

ONE-DIMENSIONAL THERMOELASTIC WAVES IN ELASTIC HALF-SPACE WITH DUAL PHASE-LAG EFFECTS

SNEHANSHU KR. ROYCHOUDHURI

The theory of thermoelasticity with dual phase-lag effects is employed to study the problem of one-dimensional disturbances in an elastic half-space with its plane boundary subjected to (i) a constant step input of temperature and zero stress, and (ii) a constant step input of stress and zero temperature. The Laplace transform method is used to solve the problem. Expressions for displacement, temperature and stress fields are obtained for small values of time. It is found that the solutions consist of two coupled waves both of which propagate with finite speeds and attenuation, influenced by the two delay times and thermoelastic coupling constant. The discontinuities that occur at the wave fronts are obtained. The characteristic features of the underlying theory are analyzed by comparing the results of the present analysis with their counterparts in coupled thermoelasticity theory (CTE) and in other generalized thermoelasticity theories ETE, TRDTE and TEWOED.

1. Introduction

Thermoelasticity theories which involve finite speed of thermal signals (second sound) have created much interest during the last three decades. The conventional coupled dynamic thermoelasticity theory (CTE), based on the mixed parabolic-hyperbolic governing equations of [Biot 1956; Chadwick 1960], predicts an infinite speed of propagation of thermoelastic disturbances. To remove the paradox of infinite speed for propagation of thermoelastic disturbances, several generalized thermoelasticity theories have been developed, which involve hyperbolic governing equations. Among these generalized theories, the extended thermoelasticity theory (ETE) proposed by Lord and Shulman [1967] involving one relaxation time (called single-phase-lag model) and the temperature-rate-dependent theory of thermoelasticity (TRDTE) proposed by Green and Lindsay [1972] involving two relaxation times are two important models of generalized theory of thermoelasticity. Experimental studies [Kaminski 1990; Mitra et al. 1995; Tzou 1995a; 1995b] indicate that the relaxation times can be of relevance in the cases involving a rapidly propagating crack tip, a localized moving heat source with high intensity, shock wave propagation, laser technique etc. Because of the experimental evidence in support of finiteness of heat propagation speed, the generalized thermoelasticity theories are considered to be more realistic than the conventional theory in dealing with practical problems involving very large heat fluxes at short intervals like those occurring in laser units and energy channels. For a review of the relevant literature, see [Chandrasekharaiah 1986; Ignaczak 1989].

Green and Naghdi [1977; 1992; 1993] formulated three different models of thermoelasticity among which, in one of these models, there is no dissipation of thermoelastic energy. This model is referred to as

Keywords: thermoelastic waves, dual phase-lag effects, thermoelastic half-space.

the G–N model of thermoelasticity without energy dissipation (TEWOED). Problems concerning generalized thermoelasticity theories and G–N theory have been studied by many authors [RoyChoudhuri and Debnath 1983; RoyChoudhuri 1984; 1985; 1987; Dhaliwal and Rokne 1988; 1989; RoyChoudhuri 1990; Chandrasekharaiah and Murthy 1993; Chandrasekharaiah and Srinath 1996; RoyChoudhuri and Banerjee 2004; RoyChoudhuri and Bandyopadhyay 2005; RoyChoudhuri and Dutta 2005; 2005]. Tzou [1995a; 1995b] and Ozisik and Tzou [1994] have developed a new model called dual phase-lag model for heat transport mechanism in which Fourier’s law is replaced by an approximation to a modification of Fourier’s law with two different time translations for the heat flux and the temperature gradient. According to this model, classical Fourier’s law $\vec{q} = -k\vec{\nabla}T$ has been generalized as $\vec{q}(P, t + \tau_q) = -k\vec{\nabla}T(P, t + \tau_T)$ where the temperature gradient $\vec{\nabla}T$ at a point P of the material at time $t + \tau_T$ corresponds to the heat flux vector \vec{q} at the same point at time $t + \tau_q$. Here k is the thermal conductivity of the material. The delay time τ_T is interpreted as that caused by the microstructural interactions (small-scale heat transport mechanisms occurring in microscale) and is called the phase-lag of the temperature gradient. The other delay time is τ_q interpreted as the relaxation time due to the fast transient effects of thermal inertia (small-scale effects of heat transport in time) and is called the phase-lag of the heat flux. If $\tau_q = \tau_T$, $\tau_T = 0$, Tzou [1995a; 1995b] refers to the model as the single phase-lag model. The case $\tau_q \neq \tau_T (\neq 0)$ corresponds to the dual phase-lag model of the constitutive equation connecting the heat flux vector and the temperature gradient. The case $\tau_q = \tau_T (\neq 0)$ becomes identical with the classical Fourier’s law. Further for materials with $\tau_q > \tau_T$, the heat flux vector is the result of a temperature gradient and for materials with $\tau_T > \tau_q$, the temperature gradient is the result of a heat flux vector. For a review of the relevant literature, see [Chandrasekharaiah 1998]. A hyperbolic thermoelastic model was developed in this same reference, taking into account the phase-lag of both temperature gradient and heat flux vector and also the second order term in τ_q in Taylor’s expansion of heat flux vector and the first order term in τ_T in Taylor’s expansion of the temperature gradient in the generalization of classical Fourier’s law. It may be pointed out that ETE was formulated by taking into account the thermal relaxation time, which is in fact the phase-lag of the heat flux vector (single phase-lag model).

The purpose of the present paper is to consider thermoelastic interaction in an elastic half-space in the context of the thermoelasticity theory based on the Tzou model [1995a; 1995b] of heat transport mechanism with dual phase-lag effects. The plane boundary is subjected to (i) a constant step input of temperature and zero stress and (ii) a constant step input of stress and zero temperature. Laplace transform is used as a mathematical tool. The expressions for displacement, temperature and stress in the half-space are derived for small times. The solution for displacement, temperature and stress consist of two waves – one, the predominantly elastic wave (E-wave) and the other, the predominantly thermal wave (T-wave) in nature, both propagating with finite speeds modified by the nondimensional delay times τ_q^* and τ_T^* . It is observed that the displacement is continuous at both the wave fronts while both the temperature and stress fields suffer finite jumps at these locations. Further the waves suffer exponential attenuation at both the wave fronts as in ETE and TRDTE. Similar problems have been studied in [Dhaliwal and Rokne 1988; 1989] in the context of ETE and TRDTE, and in [Chandrasekharaiah and Srinath 1996] in the context of TEWOED. The results of the present analysis are compared with those derived in the context of ETE, TRDTE, TEWOED and CTE. The present investigation has brought to light some similarities and differences for the theories ETE, TRDTE, TEWOED and CTE.

2. Formulation of the problem: basic equations

An isotropic elastic homogeneous half-space is considered. The plane boundary is subjected to a constant step input of temperature and zero stress. We study the disturbances produced in the half-space. The solid is subjected to one-dimensional deformation so that all the field variables are functions of the spatial coordinate x and time t . If θ is the temperature increase at time t above the uniform reference temperature θ_0 and \vec{u} the displacement vector, the heat transport equation which includes dual phase-lag effects (see [Ozisik and Tzou 1994; Tzou 1995a; 1995b]) is

$$k\left(1 + \tau_T \frac{\partial}{\partial t}\right) \nabla^2 \theta = \left(1 + \tau_q \frac{\partial}{\partial t} + \frac{1}{2} \tau_q^2 \frac{\partial^2}{\partial t^2}\right) (\rho c_v \dot{\theta} + \beta \theta_0 \dot{\Delta} - \rho R), \tag{1}$$

where τ_T and τ_q are the phase-lag of the temperature gradient and of the heat flux respectively, often referred to as the delay times, k is the thermal conductivity of the solid, and R is the heat source term. In addition, $\Delta = \text{div } \vec{u}$ and $\beta = (3\lambda + 2\mu)\alpha_t$. The displacement equation of motion is

$$\mu \nabla^2 \vec{u} + (\lambda + \mu) \text{grad } \Delta - \beta \text{grad } \theta + \rho \vec{F} = \rho \ddot{\vec{u}}. \tag{2}$$

Here λ and μ are Lamé constants, ρ is the constant mass density of the solid, α_t is the coefficient of linear thermal expansion of the material, k is the thermal conductivity, c_v is the specific of the solid.

For one-dimensional deformation $\vec{u} = (u(x, t), 0, 0)$ and $\theta = \theta(x, t)$. In absence of heat source and body forces, the Equations (1) and (2), in case of one-dimensional disturbances, reduce to

$$k\left(1 + \tau_T \frac{\partial}{\partial t}\right) \frac{\partial^2 \theta}{\partial x^2} = \left(1 + \tau_q \frac{\partial}{\partial t} + \frac{1}{2} \tau_q^2 \frac{\partial^2}{\partial t^2}\right) \left(\rho c_v \dot{\theta} + \beta \theta_0 \frac{\partial^2 u}{\partial x \partial t}\right) \tag{3}$$

and

$$(\lambda + 2\mu) \frac{\partial^2 u}{\partial x^2} - \beta \frac{\partial \theta}{\partial x} = \rho \frac{\partial^2 u}{\partial t^2}. \tag{4}$$

We introduce the following nondimensional variables

$$\xi = \frac{c_1 x}{\kappa}, \quad \eta = \frac{c_1^2 t}{\kappa}, \quad \Theta = \frac{\theta}{\theta_0}, \quad U = \frac{c_1(\lambda + 2\mu)u}{\kappa \beta \theta_0},$$

where $\kappa = k/\rho c_v$ is the thermal diffusivity. The Equations (3)–(4) reduce to the following nondimensional forms

$$\left(1 + \tau_q^* \frac{\partial}{\partial \eta}\right) \frac{\partial^2 \Theta}{\partial \xi^2} = \left(1 + \tau_q^* \frac{\partial}{\partial \eta} + \frac{1}{2} \tau_q^{*2} \frac{\partial^2}{\partial \eta^2}\right) \left(\frac{\partial \Theta}{\partial \eta} + \varepsilon \frac{\partial^2 U}{\partial \xi \partial \eta}\right) \tag{5}$$

and

$$\frac{\partial^2 U}{\partial \xi^2} - \frac{\partial \Theta}{\partial \xi} = \frac{\partial^2 U}{\partial \eta^2}, \tag{6}$$

where

$$\tau_q^* = \frac{\tau_q c_1^2}{\kappa}, \quad \tau_T^* = \frac{\tau_T c_1^2}{\kappa}, \quad \varepsilon = \frac{\beta^2 \theta_0}{\rho^2 c_v c_1^2}, \quad \tau = \frac{\sigma_{xx}}{\beta \theta_0} = \frac{\partial U}{\partial \xi} - \Theta$$

are respectively the nondimensional delay times, the thermoelasticity coupling, and the nondimensional stress.

If τ_q^{*2} is neglected and $\tau_T^* = 0$, on setting $\tau_q^* = \tau$ = thermal relaxation parameter, the equation (5) and (6) reduce to L-S theory.

Further if τ_q^{*2} is neglected and $\tau_T^* \neq 0$, Equation (5) reduces to

$$\left(1 + \tau_T^* \frac{\partial}{\partial \eta}\right) \frac{\partial^2 \Theta}{\partial \xi^2} = \left(1 + \tau_q^* \frac{\partial}{\partial \eta}\right) \left(\frac{\partial \Theta}{\partial \eta} + \varepsilon \frac{\partial^2 U}{\partial \xi \partial \eta}\right). \tag{7}$$

This equation with the equation of motion (6) then constitutes a coupled system of field equations of a thermoelasticity theory with non-Fourier heat transport equation (7).

3. Solution of the problem in the Laplace transform domain

We now proceed to study one-dimensional thermoelastic disturbances in the half-space $\xi \geq 0$ on the basis of Equations (5)–(6). We define the Laplace transforms of the functions $U(\xi, \eta)$ and $\Theta(\xi, \eta)$ by

$$\{\bar{U}(\xi, s), \bar{\Theta}(\xi, s)\} = \int_0^\infty \{U(\xi, \eta), \Theta(\xi, \eta)\} e^{-s\eta} d\eta,$$

where $\text{Re}(s) > 0$, s is the Laplace transform parameter. We assume that the medium is at rest at $\eta = 0$ and has its temperature, temperature-velocity and temperature acceleration equal to zero at $\eta = 0$. This means that

$$U = \frac{\partial U}{\partial \eta} = \frac{\partial^2 U}{\partial \eta^2} = 0 \quad \text{and} \quad \Theta = \frac{\partial \Theta}{\partial \eta} = \frac{\partial^2 \Theta}{\partial \eta^2} = 0, \quad \text{for } \eta = 0, \xi \geq 0. \tag{8}$$

If the disturbances are caused by the sudden application of a constant step in temperature on the boundary which is stress-free at time $\eta > 0$ (Danilovskaya’s problem [1950]), then this leads to the boundary conditions

$$\Theta(0, \eta) = \Theta_0 H(\eta), \quad \tau(0, \eta) = 0, \quad \eta > 0,$$

where Θ_0 is a positive constant, and $H(\eta)$ is the Heaviside unit step function, taking the value 1 if $\eta > 0$ and 0 if $\eta \leq 0$.

On using stress-strain-temperature relations, the conditions become

$$\theta(0, \eta) = \Theta_0 H(\eta), \quad \frac{\partial U}{\partial \xi}(0, \eta) = \Theta_0 H(\eta), \quad \text{for } \eta > 0. \tag{9}$$

Alternatively, if the thermoelastic interactions are caused by a uniform step in the stress applied to the boundary of the half-space, which is held at reference temperature θ_0 , then the following boundary conditions hold:

$$\tau(0, \eta) = -\tau_0 H(\eta), \quad \Theta(0, \eta) = 0, \quad \eta > 0,$$

where τ_0 is a positive constant, or

$$\frac{\partial U}{\partial \xi}(0, \eta) = -\tau_0 H(\eta), \quad \Theta(0, \eta) = 0, \quad \text{for } \eta > 0. \tag{10}$$

Now the equations (5) and (6), on taking Laplace transform, reduce to

$$(ND^2 - M)\bar{\Theta} = \varepsilon MD\bar{U}, \tag{11}$$

$$(D^2 - s^2)\bar{U} = D\bar{\Theta}, \tag{12}$$

where

$$D = \frac{d}{d\xi}, \quad M = s \left(1 + \tau_q^* s + \frac{1}{2} \tau_q^{*2} s^2 \right), \quad N = 1 + \tau_T^* s.$$

This leads to the following equation satisfied by \bar{U} and $\bar{\Theta}$

$$[ND^4 - (Ns^2 + M + M\varepsilon)D^2 + Ms^2](\bar{U}, \bar{\Theta}) = 0. \tag{13}$$

The solutions of equation (13), vanishing as $\xi \rightarrow \infty$, are assumed to take the form

$$\bar{U} = c_1 e^{-m_1 \xi} + c_2 e^{-m_2 \xi}, \quad \bar{\Theta} = c_1^1 e^{-m_1 \xi} + c_2^1 e^{-m_2 \xi}, \tag{14}$$

where $m_{1,2}$ are the roots with positive real part of the equation

$$Nm^4 - (Ns^2 + M + M\varepsilon)m^2 + Ms^2 = 0. \tag{15}$$

Again, on taking the Laplace transform of the boundary conditions (9) we have

$$\bar{\Theta} = \frac{\Theta_0}{s} \quad \text{and} \quad \frac{d\bar{U}}{d\xi} = \frac{\Theta_0}{s}, \quad \text{on } \xi = 0. \tag{16}$$

Substituting the solutions (14) into (12) and equating the coefficients of like exponentials, we obtain

$$c_1^1 = \frac{c_1(s^2 - m_1^2)}{m_1}, \quad c_2^1 = \frac{c_2(s^2 - m_2^2)}{m_2}$$

Using the conditions (16) and solving for c_1, c_2 , we arrive at the following solutions in the Laplace transform domain:

Case (i):

$$\begin{aligned} \bar{U} &= \frac{\Theta_0}{s} \frac{1}{(m_2^2 - m_1^2)} [m_1 e^{-m_1 \xi} - m_2 e^{-m_2 \xi}], \\ \bar{\Theta} &= \frac{\Theta_0}{s} \frac{1}{(m_2^2 - m_1^2)} [(s^2 - m_1^2) e^{-m_1 \xi} - (s^2 - m_2^2) e^{-m_2 \xi}], \\ \bar{\tau} &= \frac{\Theta_0 s}{(m_2^2 - m_1^2)} [e^{-m_2 \xi} - e^{-m_1 \xi}], \quad \text{for } \xi > 0. \end{aligned} \tag{17}$$

Case (ii):

$$\begin{aligned} \bar{U} &= \frac{\tau_0}{s^3(m_1^2 - m_2^2)} [m_1(s^2 - m_2^2) e^{-m_1 \xi} - m_2(s^2 - m_1^2) e^{-m_2 \xi}], \\ \bar{\Theta} &= \frac{\tau_0(s^2 - m_1^2)(s^2 - m_2^2)}{s^3(m_1^2 - m_2^2)} [e^{-m_1 \xi} - e^{-m_2 \xi}], \\ \bar{\tau} &= \frac{\tau_0}{s(m_1^2 - m_2^2)} [-(s^2 - m_2^2) e^{-m_1 \xi} + (s^2 - m_1^2) e^{-m_2 \xi}], \quad \text{for } \xi > 0. \end{aligned} \tag{18}$$

The roots of the biquadratic equation (15) are given by

$$m_{1,2}^2 = \frac{1}{2N} \left[Ns^2 + M(1 + \varepsilon) \pm \left\{ (Ns^2 + M(1 + \varepsilon))^2 - 4MNs^2 \right\}^{1/2} \right]. \tag{19}$$

Clearly the roots given by (19) are real if s is real, since

$$(Ns^2 + M(1 + \varepsilon))^2 - 4MNs^2 = (M - Ns^2)^2 + M^2\varepsilon^2 + 2M\varepsilon(M + Ns^2) > 0.$$

The inverse Laplace transforms of (17)–(18) then determine U, Θ, τ . Since $m_{1,2}$ involve the Laplace parameter s , determination of U, Θ, τ is difficult. Since the second sound effects are short-lived, it is sufficient to derive and analyze the solutions for small η . This is done by taking Laplace parameter s to be large.

Taking the sign $+$ in (19), we have for large s ,

$$m_1 \cong \frac{s}{v_1} + \frac{1}{2} \frac{\lambda_2}{\lambda_1} \frac{1}{v_1} + \frac{1}{2v_1} \left(\frac{\lambda_3}{\lambda_1} - \frac{1}{4} \frac{\lambda_3^2}{\lambda_1^2} \right) \frac{1}{s}. \tag{20}$$

Taking the sign $-$ in (19), we have for large s

$$m_2 \cong \frac{s}{v_2} + \frac{1}{2} \frac{\mu_2}{\mu_1} \frac{1}{v_2} + \frac{1}{2v_2} \left(\frac{\mu_3}{\mu_1} - \frac{1}{4} \frac{\mu_3^2}{\mu_1^2} \right) \frac{1}{s}, \tag{21}$$

where $\lambda_1 = A + \sqrt{A^2 - 4F}$, $\lambda_2 = B + \frac{L_1}{2} \sqrt{A^2 - 4F} - \frac{A + \sqrt{A^2 - 4F}}{\tau_T^*}$,

$$v_1 = \left(\frac{2\tau_T^*}{\lambda_1} \right)^{1/2}, \quad v_2 = \left(\frac{2\tau_T^*}{\mu_1} \right)^{1/2},$$

$$\lambda_3 = C + \frac{\sqrt{A^2 - 4F}}{8} (4L_2^2 - L_1^2) + \frac{1}{\tau_T^{*2}} (A + \sqrt{A^2 - 4F}) - \frac{1}{\tau_T^*} \left(B + \frac{L_1}{2} \sqrt{A^2 - 4F} \right), \tag{22}$$

$$L_1 = \frac{2AB - 4D}{A^2 - 4F}, \quad L_2 = \frac{B^2 + 2CA - 4E}{A^2 - 4F}, \quad A = \tau_T^* + \frac{1}{2} (1 + \varepsilon) \tau_q^{*2},$$

$$B = 1 + (1 + \varepsilon) \tau_q^*, \quad C = 1 + \varepsilon, \quad D = \frac{1}{2} \tau_q^{*2} + \tau_T^* \tau_q^*,$$

$$E = \tau_T^* + \tau_q^*, \quad F = \frac{1}{2} \tau_T^* \tau_q^{*2},$$

$$\mu_1 = A - \sqrt{A^2 - 4F},$$

$$\mu_2 = B - \frac{L_1}{2} \sqrt{A^2 - 4F} - \frac{1}{\tau_T^*} (A - \sqrt{A^2 - 4F}),$$

$$\mu_3 = C - \frac{\sqrt{A^2 - 4F}}{8} (4L_2^2 - L_1^2) + \frac{1}{\tau_T^{*2}} (A - \sqrt{A^2 - 4F}) - \frac{1}{\tau_T^*} \left(B - \frac{L_1}{2} \sqrt{A^2 - 4F} \right).$$

We note that

$$A^2 - 4F = \frac{1}{4} [(2\tau_T^* - \tau_q^*)^2 + \varepsilon^2 \tau_q^{*4} + 4\tau_T^* \tau_q^{*2} \varepsilon] > 0$$

and $A > 0$.

This indicates that $v_{1,2}$ are both real.

Clearly $\lambda_1, \mu_1 > 0$ since $A > \sqrt{A^2 - 4F}$ and $F > 0$. Further $\lambda_1 > \mu_1$ implies $v_2 > v_1$.

Now we are to prove that, under suitable restrictions on material constants, λ_2 and μ_2 are positive. We have

$$\lambda_2 = B + \frac{L_1}{2} \sqrt{A^2 - 4F} - \frac{A + \sqrt{A^2 - 4F}}{\tau_T^*}, \quad \mu_2 = B - \frac{L_1}{2} \sqrt{A^2 - 4F} - \frac{A - \sqrt{A^2 - 4F}}{\tau_T^*}.$$

Now

$$\lambda_2 > 0, \text{ if } B + \frac{L_1}{2} \sqrt{A^2 - 4F} > \frac{A + \sqrt{A^2 - 4F}}{\tau_T^*}.$$

That is, if

$$B > \frac{A}{\tau_T^*} + \frac{\sqrt{A^2 - 4F}}{\tau_T^*} \left(1 - \frac{L_1 \tau_T^*}{2}\right).$$

Similarly,

$$\mu_2 > 0, \text{ if } B > \frac{A}{\tau_T^*} - \frac{\sqrt{A^2 - 4F}}{\tau_T^*} \left(1 - \frac{L_1 \tau_T^*}{2}\right).$$

We impose the restriction on material parameters such that $1 > \frac{L_1}{2} \tau_T^*$. Then since $A, \tau_T^* > 0, B$ must be positive under the restriction $1 > \frac{L_1}{2} \tau_T^*$. The required restriction, on substitution for L_1 , yields $A^2 - 4F > \tau_T^*(AB - 2D)$. This leads to the inequality

$$(\tau_q^* - 2\tau_T^*)(1 + \varepsilon)\tau_q^{*2} - 2\tau_T^*(1 - \varepsilon) > 0.$$

This is satisfied if $\tau_q^* > 2\tau_T^*$ and

$$(1 + \varepsilon)\tau_q^{*2} - 2\tau_T^*(1 - \varepsilon) > (1 + \varepsilon)4\tau_T^{*2} - 2\tau_T^*(1 - \varepsilon) = 2\tau_T^*\{2(1 + \varepsilon)\tau_T^* + \varepsilon - 1\} > 0 \quad \text{or} \quad 2(1 + \varepsilon)\tau_T^* > 1 - \varepsilon.$$

Thus the two conditions are $2(1 + \varepsilon)\tau_T^* > 1 - \varepsilon$, that is,

$$\tau_T^* > \frac{1 - \varepsilon}{2(1 + \varepsilon)} \quad \text{and} \quad \tau_q^* > \sqrt{\frac{2\tau_T^*(1 - \varepsilon)}{(1 + \varepsilon)}}.$$

Further since $\tau_q^* > 2\tau_T^*$, we must have

$$\tau_q^* > \frac{1 - \varepsilon}{1 + \varepsilon} =: \varepsilon_0.$$

The required restrictions on material constants for $\lambda_2 > 0$ then reduce to

$$\tau_q > \frac{k\varepsilon_0}{\rho c_v c_1^2} \quad \text{and} \quad \tau_T > \frac{k}{2\rho c_v c_1^2} \varepsilon_0.$$

Since $\lambda_2 > 0$ implies $\mu_2 > 0$, the inequalities

$$\tau_q > \frac{k\varepsilon_0}{\rho c_v c_1^2} \quad \text{and} \quad \tau_T > \frac{k\varepsilon_0}{2\rho c_v c_1^2}$$

imply that $\mu_2 > 0$. Using the results

$$m_1 \cong \frac{s}{v_1} + \frac{1}{2} \frac{\lambda_2}{\lambda_1} \frac{1}{v_1} \quad \text{and} \quad m_2 \cong \frac{s}{v_2} + \frac{1}{2} \frac{\mu_2}{\mu_1} \frac{1}{v_2}$$

for large s , we obtain the following results after simplification:

$$\frac{m_1}{s(m_2^2 - m_1^2)} \cong \frac{1}{L_0 v_1} \frac{1}{s^2} - \frac{1}{L_0} \left(\frac{M_0}{L_0} + \frac{\lambda_2}{2\lambda_1} \right) \frac{1}{v_1} \frac{1}{s^3},$$

$$\frac{m_2}{s(m_2^2 - m_1^2)} \cong \frac{1}{L_0 v_2} \frac{1}{s^2} - \frac{1}{L_0} \left(\frac{M_0}{L_0} + \frac{\mu_2}{2\mu_1} \right) \frac{1}{v_2} \frac{1}{s^3},$$

$$\frac{s^2 - m_2^2}{s(m_2^2 - m_1^2)} \cong \frac{1}{L_0} \left(\frac{v_2^2 - 1}{v_2^2} \frac{1}{s} - \left(\frac{M_0}{L_0} \frac{v_2^2 - 1}{v_2^2} + \frac{\mu_2}{\mu_1 v_2^2} \right) \frac{1}{s^2} + \left(\frac{\lambda_2}{\lambda_1 v_1^2} \frac{M_0}{L_0} + \frac{v_1^2 - 1}{v_1^2} \frac{(M_0^2 - N_0 L_0)}{L_0^2} \right) \frac{1}{s^3} \right),$$

$$\frac{s^2 - m_1^2}{s(m_2^2 - m_1^2)} \cong \frac{1}{L_0} \left(\frac{v_1^2 - 1}{v_1^2} \frac{1}{s} - \left(\frac{M_0}{L_0} \frac{v_1^2 - 1}{v_1^2} + \frac{\lambda_2}{\lambda_1 v_1^2} \right) \frac{1}{s^2} + \left(\frac{\mu_2}{\mu_1 v_2^2} \frac{M_0}{L_0} + \frac{v_2^2 - 1}{v_2^2} \frac{(M_0^2 - N_0 L_0)}{L_0^2} \right) \frac{1}{s^3} \right),$$

$$\begin{aligned} \frac{(s^2 - m_1^2)(s^2 - m_2^2)}{s^3(m_1^2 - m_2^2)} \cong & -\frac{1}{L_0} \left(\frac{(v_1^2 - 1)(v_2^2 - 1)}{v_1^2 v_2^2} \frac{1}{s} - \left(\frac{(v_1^2 - 1)\mu_2}{v_1^2 \cdot \mu_1 v_2^2} + \frac{(v_2^2 - 1)}{v_2^2} \alpha_1 \right) \frac{1}{s^2} \right. \\ & \left. + \left(\frac{\mu_2}{\mu_1 v_2^2} \alpha_1 + \frac{(v_2^2 - 1)}{v_2^2} \alpha_2 \right) \frac{1}{s^3} \right), \end{aligned}$$

where

$$L_0 = \frac{1}{v_2^2} - \frac{1}{v_1^2}, \quad M_0 = \frac{\mu_2}{\mu_1 v_2^2} - \frac{\lambda_2}{\lambda_1 v_1^2}, \quad N_0 = \frac{1}{4} \left(\frac{\mu_2^2}{\mu_1^2 v_2^2} - \frac{\lambda_2^2}{\lambda_1^2 v_1^2} \right),$$

$$\alpha_1 = \frac{M_0 (v_1^2 - 1)}{L_0 v_1^2} + \frac{\lambda_2}{\lambda_1 v_1^2}, \quad \alpha_2 = \frac{\lambda_2}{\lambda_1 v_1^2} \frac{M_0}{L_0} + \frac{(v_1^2 - 1)}{v_1^2} \frac{(M_0^2 - N_0 L_0)}{L_0^2}.$$

Finally we obtain the following solutions for displacement, temperature and stress fields in the Laplace transform domain for large s :

Case (i):

$$\begin{aligned} \bar{U}(\xi, s) \cong & \Theta_0 \left[\left\{ \frac{1}{L_0 v_1} \frac{1}{s^2} - \frac{1}{L_0} \left(\frac{M_0}{L_0} + \frac{\lambda_2}{2\lambda_1} \right) \frac{1}{v_1} \frac{1}{s^3} \right\} e^{-\left(\frac{s}{v_1} + \frac{\lambda_2}{2\lambda_1 v_1} \right) \xi} \right. \\ & \left. - \left\{ \frac{1}{L_0 v_2} \frac{1}{s^2} - \frac{1}{L_0} \left(\frac{M_0}{L_0} + \frac{\mu_2}{2\mu_1} \right) \frac{1}{v_2} \frac{1}{s^3} \right\} e^{-\left(\frac{s}{v_2} + \frac{\mu_2}{2\mu_1 v_2} \right) \xi} \right], \end{aligned}$$

$$\begin{aligned} \bar{\Theta}(\xi, s) \cong & \frac{\Theta_0}{L_0} \left[\left\{ \frac{(v_1^2 - 1)}{v_1^2} \frac{1}{s} - \left(\frac{M_0}{L_0} \frac{v_1^2 - 1}{v_1^2} + \frac{\lambda_2}{\lambda_1 v_1^2} \right) \frac{1}{s^2} \right. \right. \\ & \left. \left. + \left(\frac{\lambda_2}{\lambda_1 v_1^2} \frac{M_0}{L_0} + \frac{(v_1^2 - 1)}{v_1^2} \frac{(M_0^2 - N_0 L_0)}{L_0^2} \right) \frac{1}{s^3} \right\} e^{-\left(\frac{s}{v_1} + \frac{\lambda_2}{2\lambda_1 v_1} \right) \xi} - \left\{ \frac{(v_2^2 - 1)}{v_2^2} \frac{1}{s} - \left(\frac{M_0}{L_0} \frac{(v_2^2 - 1)}{v_2^2} + \frac{\mu_2}{\mu_1 v_2^2} \right) \frac{1}{s^2} \right. \right. \\ & \left. \left. + \left(\frac{\mu_2}{\mu_1 v_2^2} \frac{M_0}{L_0} + \frac{(v_2^2 - 1)}{v_2^2} \frac{(M_0^2 - N_0 L_0)}{L_0^2} \right) \frac{1}{s^3} \right\} e^{-\left(\frac{s}{v_2} + \frac{\mu_2}{2\mu_1 v_2} \right) \xi} \right], \end{aligned}$$

$$\bar{\tau}(\xi, s) \cong \frac{\Theta_0}{L_0} \left[- \left\{ \frac{1}{s} - \frac{M_0}{L_0} \frac{1}{s^2} + \frac{M_0^2 - L_0 N_0}{L_0^2} \frac{1}{s^3} \right\} e^{-\left(\frac{s}{v_1} + \frac{\lambda_2}{2\lambda_1 v_1}\right)\xi} + \left\{ \frac{1}{s} - \frac{M_0}{L_0} \frac{1}{s^2} + \frac{M_0^2 - L_0 N_0}{L_0^2} \frac{1}{s^3} \right\} e^{-\left(\frac{s}{v_2} + \frac{\mu_2}{2\mu_1 v_2}\right)\xi} \right].$$

Case (ii):

$$\begin{aligned} \bar{U}(\xi, s) \cong \frac{\tau_0}{L_0} & \left[- \frac{1}{v_1} \left\{ \frac{(v_2^2 - 1)}{v_2^2} \frac{1}{s^2} - \left(\frac{M_0}{L_0} \frac{v_2^2 - 1}{v_2^2} + \frac{\mu_2}{\mu_1 v_2^2} \right) \frac{1}{s^3} \right. \right. \\ & + \left. \left. \left(\frac{\mu_2}{\mu_1 v_2^2} \frac{M_0}{L_0} + \frac{(v_2^2 - 1)(M_0^2 - N_0 L_0)}{v_2^2 L_0^2} \right) \frac{1}{s^4} \right\} e^{-\left(\frac{s}{v_1} + \frac{\lambda_2}{2\lambda_1 v_1}\right)\xi} \right. \\ & + \frac{1}{v_2} \left\{ \frac{(v_1^2 - 1)}{v_1^2} \frac{1}{s^2} - \left(\frac{M_0}{L_0} \frac{(v_1^2 - 1)}{v_1^2} + \frac{\lambda_2}{\lambda_1 v_1^2} \right) \frac{1}{s^3} \right. \\ & \left. \left. + \left(\frac{\lambda_2}{\lambda_1 v_1^2} \frac{M_0}{L_0} + \frac{(v_1^2 - 1)(M_0^2 - N_0 L_0)}{v_1^2 L_0^2} \right) \frac{1}{s^4} \right\} e^{-\left(\frac{s}{v_2} + \frac{\mu_2}{2\mu_1 v_2}\right)\xi} \right], \end{aligned}$$

$$\begin{aligned} \bar{\Theta}(\xi, s) \cong - \frac{\tau_0}{L_0} & \left[\frac{(v_1^2 - 1)(v_2^2 - 1)}{v_1^2 v_2^2} \frac{1}{s} - \left\{ \frac{(v_1^2 - 1)\mu_2}{v_1^2 v_2^2 \mu_1} + \frac{(v_2^2 - 1)}{v_2^2} \left(\frac{M_0}{L_0} \frac{(v_1^2 - 1)}{v_1^2} + \frac{\lambda_2}{\lambda_1 v_1^2} \right) \right\} \cdot \frac{1}{s^2} \right. \\ & \left. + \left\{ \frac{\mu_2}{\mu_1 v_2^2} \alpha_1 + \frac{v_2^2 - 1}{v_2^2} \alpha_2 \right\} \frac{1}{s^3} \right] \times \left\{ e^{-\left(\frac{s}{v_1} + \frac{\lambda_2}{2\lambda_1 v_1}\right)\xi} - e^{-\left(\frac{s}{v_2} + \frac{\mu_2}{2\mu_1 v_2}\right)\xi} \right\}, \end{aligned}$$

$$\begin{aligned} \bar{\tau}(\xi, s) \cong \frac{\tau_0}{L_0} & \left[\left\{ \frac{(v_2^2 - 1)}{v_2^2} \frac{1}{s} - \left(\frac{M_0}{L_0} \frac{(v_2^2 - 1)}{v_2^2} + \frac{\mu_2}{\mu_1 v_2^2} \right) \frac{1}{s^2} \right. \right. \\ & + \left. \left. \left(\frac{\mu_2}{\mu_1 v_2^2} \frac{M_0}{L_0} + \frac{(v_2^2 - 1)(M_0^2 - N_0 L_0)}{v_2^2 L_0^2} \right) \frac{1}{s^3} \right\} e^{-\left(\frac{s}{v_1} + \frac{\lambda_2}{2\lambda_1 v_1}\right)\xi} \right. \\ & - \left\{ \frac{(v_1^2 - 1)}{v_1^2} \frac{1}{s} - \left(\frac{M_0}{L_0} \frac{(v_1^2 - 1)}{v_1^2} + \frac{\lambda_2}{\lambda_1 v_1^2} \right) \frac{1}{s^2} \right. \\ & \left. \left. + \left(\frac{\lambda_2}{\lambda_1 v_1^2} \frac{M_0}{L_0} + \frac{(v_1^2 - 1)(M_0^2 - L_0 N_0)}{v_1^2 L_0^2} \right) \frac{1}{s^3} \right\} e^{-\left(\frac{s}{v_2} + \frac{\mu_2}{2\mu_1 v_2}\right)\xi} \right]. \end{aligned}$$

4. Derivation of small-time solutions

Inverse Laplace transforms of the expressions yield the following small-time solutions for displacement, temperature and stress fields.

Case (i):

$$\begin{aligned}
 U(\xi, \eta) \cong \Theta_0 \left[e^{-\frac{\lambda_2}{2\lambda_1 v_1} \xi} \left\{ \frac{1}{L_0 v_1} \left(\eta - \frac{\xi}{v_1} \right) \right. \right. \\
 \left. \left. - \frac{1}{L_0} \left(\frac{M_0}{L_0} + \frac{\lambda_2}{2\lambda_1} \right) \frac{1}{v_1} \frac{1}{2} \left(\eta - \frac{\xi}{v_1} \right)^2 \right\} H \left(\eta - \frac{\xi}{v_1} \right) - e^{-\frac{\mu_2}{2\mu_1 v_2} \xi} \left\{ \frac{1}{L_0 v_2} \left(\eta - \frac{\xi}{v_2} \right), \right. \right. \\
 \left. \left. - \frac{1}{L_0} \left(\frac{M_0}{L_0} + \frac{\mu_2}{2\mu_1} \right) \frac{1}{v_1} \frac{1}{2} \left(\eta - \frac{\xi}{v_2} \right)^2 \right\} H \left(\eta - \frac{\xi}{v_2} \right) \right], \quad (23)
 \end{aligned}$$

$$\begin{aligned}
 \Theta(\xi, \eta) \cong \frac{\Theta_0}{L_0} \left[e^{-\frac{\lambda_2}{2\lambda_1 v_1} \xi} \left\{ \frac{(v_1^2 - 1)}{v_1^2} - \left(\frac{M_0 (v_1^2 - 1)^2}{L_0 v_1^2} + \frac{\lambda_2}{\lambda_1 v_1^2} \right) \left(\eta - \frac{\xi}{v_1} \right) \right. \right. \\
 \left. \left. + \left(\frac{\lambda_2 M_0}{\lambda_1 v_1^2 L_0} + \frac{(v_1^2 - 1) (M_0^2 - N_0 L_0)}{v_1^2 L_0^2} \right) \frac{1}{2} \left(\eta - \frac{\xi}{v_1} \right)^2 \right\} H \left(\eta - \frac{\xi}{v_1} \right) \right. \\
 \left. - e^{-\frac{\mu_2}{2\mu_1 v_2} \xi} \left\{ \frac{(v_2^2 - 1)}{v_2^2} - \left(\frac{M_0 (v_2^2 - 1)}{L_0 v_2^2} + \frac{\mu_2}{\mu_1 v_2^2} \right) \left(\eta - \frac{\xi}{v_2} \right) \right. \right. \\
 \left. \left. + \left(\frac{\mu_2 M_0}{\mu_1 v_2^2 L_0} + \frac{(v_2^2 - 1) (M_0^2 - N_0 L_0)}{v_2^2 L_0^2} \right) \frac{1}{2} \left(\eta - \frac{\xi}{v_2} \right)^2 \right\} H \left(\eta - \frac{\xi}{v_2} \right) \right], \quad (24)
 \end{aligned}$$

$$\begin{aligned}
 \tau(\xi, \eta) \cong \frac{\Theta_0}{L_0} \left[-e^{-\frac{\lambda_2}{2\lambda_1 v_1} \xi} \left\{ 1 - \frac{M_0}{L_0} \left(\eta - \frac{\xi}{v_1} \right) + \frac{M_0^2 - L_0 N_0}{L_0^2} \frac{1}{2} \left(\eta - \frac{\xi}{v_2} \right)^2 \right\} H \left(\eta - \frac{\xi}{v_1} \right) \right. \\
 \left. + e^{-\frac{\mu_2}{2\mu_1 v_2} \xi} \left\{ 1 - \frac{M_0}{L_0} \left(\eta - \frac{\xi}{v_2} \right) + \frac{(M_0^2 - L_0 N_0)}{L_0^2} \frac{1}{2} \left(\eta - \frac{\xi}{v_2} \right)^2 \right\} H \left(\eta - \frac{\xi}{v_2} \right) \right]. \quad (25)
 \end{aligned}$$

Case (ii):

$$\begin{aligned}
 U(\xi, \eta) \cong \frac{\tau_0}{L_0} \left[-\frac{e^{-\frac{\lambda_2}{2\lambda_1 v_1} \xi}}{v_1} \left\{ \frac{(v_2^2 - 1)}{v_2^2} \left(\eta - \frac{\xi}{v_1} \right) - \left(\frac{M_0 v_2^2 - 1}{L_0 v_2^2} + \frac{\mu_2}{\mu_1 v_2^2} \right) \frac{1}{2} \left(\eta - \frac{\xi}{v_1} \right)^2 \right. \right. \\
 \left. \left. + \left(\frac{\mu_2 M_0}{\mu_1 v_2^2 L_0} + \frac{(v_2^2 - 1) (M_0^2 - N_0 L_0)}{v_2^2 L_0^2} \right) \frac{1}{6} \left(\eta - \frac{\xi}{v_1} \right)^3 \right\} H \left(\eta - \frac{\xi}{v_1} \right) \right. \\
 \left. + \frac{e^{-\frac{\mu_2}{2\mu_1 v_2} \xi}}{v_2} \left\{ \frac{(v_1^2 - 1)}{v_1^2} \left(\eta - \frac{\xi}{v_2} \right) - \left(\frac{M_0 v_1^2 - 1}{L_0 v_1^2} + \frac{\lambda_2}{\lambda_1 v_1^2} \right) \frac{1}{2} \left(\eta - \frac{\xi}{v_2} \right)^2 \right. \right. \\
 \left. \left. + \left(\frac{\lambda_2 M_0}{\lambda_1 v_1^2 L_0} + \frac{(v_1^2 - 1) (M_0^2 - N_0 L_0)}{v_1^2 L_0^2} \right) \frac{1}{6} \left(\eta - \frac{\xi}{v_2} \right)^3 \right\} H \left(\eta - \frac{\xi}{v_2} \right) \right], \quad (26)
 \end{aligned}$$

$$\Theta(\xi, \eta) \cong -\frac{\tau_0}{L_0} \left[\phi\left(\eta - \frac{\xi}{v_1}\right) H\left(\eta - \frac{\xi}{v_1}\right) e^{-\frac{\lambda_2}{2\lambda_1 v_1} \xi} - \phi\left(\eta - \frac{\xi}{v_2}\right) H\left(\eta - \frac{\xi}{v_2}\right) e^{-\frac{\mu_2}{2\mu_1 v_2} \xi} \right], \tag{27}$$

$$\begin{aligned} \tau(\xi, \eta) \cong & \frac{\tau_0}{L_0} \left[e^{-\frac{\lambda_2}{2\lambda_1 v_1} \xi} \left\{ \frac{v_2^2 - 1}{v_2^2} H\left(\eta - \frac{\xi}{v_1}\right) - \left(\frac{M_0 v_2^2 - 1}{L_0 v_2^2} + \frac{\mu_2}{\mu_1 v_2^2} \right) \left(\eta - \frac{\xi}{v_1}\right) H\left(\eta - \frac{\xi}{v_1}\right) \right. \right. \\ & + \left. \left. \left(\frac{\mu_2 M_0}{\mu_1 v_2^2 L_0} + \frac{(v_2^2 - 1)(M_0^2 - N_0 L_0)}{v_2^2 L_0^2} \right) \frac{1}{2} \left(\eta - \frac{\xi}{v_1}\right)^2 H\left(\eta - \frac{\xi}{v_1}\right) \right\} \right. \\ & - e^{-\frac{\mu_2}{2\mu_1 v_2} \xi} \left\{ \frac{v_1^2 - 1}{v_1^2} H\left(\eta - \frac{\xi}{v_2}\right) - \left(\frac{M_0 (v_1^2 - 1)}{L_0 v_1^2} + \frac{\lambda_2}{\lambda_1 v_1^2} \right) \left(\eta - \frac{\xi}{v_2}\right) H\left(\eta - \frac{\xi}{v_2}\right) \right. \\ & \left. \left. + \left(\frac{\lambda_2 M_0}{\lambda_1 v_1^2 L_0} + \frac{(v_1^2 - 1)(M_0^2 - N_0 L_0)}{v_1^2 L_0^2} \right) \frac{1}{2} \left(\eta - \frac{\xi}{v_2}\right)^2 H\left(\eta - \frac{\xi}{v_2}\right) \right\} \right], \tag{28} \end{aligned}$$

where

$$\begin{aligned} \varphi(\eta) = & \frac{(v_1^2 - 1)(v_2^2 - 1)}{v_1^2 v_2^2} - \left\{ \frac{(v_1^2 - 1)\mu_2}{v_1^2 v_2^2 \mu_1} + \frac{(v_2^2 - 1)}{v_2^2} \left(\frac{M_0 (v_1^2 - 1)}{L_0 v_1^2} + \frac{\lambda_2}{\lambda_1 v_1^2} \right) \right\} \eta + \\ & \left\{ \frac{\mu_2}{\mu_1 v_2^2} \alpha_1 + \frac{v_2^2 - 1}{v_2^2} \alpha_2 \right\} \frac{1}{2} \eta^2. \end{aligned}$$

5. Analysis of the solutions

The short-time solutions (23)–(28) for displacement, temperature, and stress fields reveal the existence of two waves. Each of U , Θ and τ is composed of two parts and that each part corresponds to a wave propagating with a finite speed. The speed of the wave corresponding to the first part is v_1 and that corresponding to the second part is v_2 . The faster wave has its speed equal to v_2 and the slower wave has its speed equal to v_1 .

One cannot classify (5) independently of (6) (as hyperbolic or parabolic), since a type of the coupled system (5)–(6) must be determined.

For $\tau_q^{*2} = 0$, we obtain from (22),

$$\begin{aligned} A &= \tau_T^*, & B &= 1 + (1 + \varepsilon)\tau_q^*, \\ C &= 1 + \varepsilon, & D &= \tau_T^* \tau_q^*, \\ E &= \tau_T^* + \tau_q^*, & F &= 0, \\ \lambda_1 &= 2\tau_T^*, & \lambda_2 &= 2\varepsilon\tau_q^*, \\ \mu_1 &= 0, & \mu_2 &= 2\tau_q^*, \end{aligned}$$

$$L_1 = \frac{2 + 2\tau_q^*(\varepsilon - 1)}{\tau_T^*},$$

$$L_2 = \{(1 - \tau_q^*)^2 + \varepsilon^2 \tau_q^{*2} + 2\varepsilon \tau_q^*(1 + \tau_q^*) + 2\tau_T^*(\varepsilon - 1)\}(\tau_T^*)^2.$$

These give

$$v_2 = \infty, \quad v_1 \longrightarrow 1, \quad L_0 \longrightarrow -1, \quad M_0 = -\varepsilon \tau_q^*/\tau_T^*, \quad N_0 = -\frac{\varepsilon^2 \tau_q^{*2}}{4\tau_T^{*2}}.$$

Therefore, for $\tau_q^{*2} = 0$, the system of equations (5)–(6) is of the mixed parabolic-hyperbolic type.

For $\tau_q^{*2} \neq 0$, and $\tau_T^* \neq 0$, v_2 corresponds to the modified speed of thermal signals and v_1 corresponds to the modified elastic dilatational wave speed, both modified by delay times in the thermoelastic solid with dual phase-lag effects. Since $v_1 < v_2$, the faster wave is the predominantly modified Tzou wave (T-wave) and the slower is a predominantly modified elastic wave (E-wave). The first term of the solutions (23)–(28) represents the contribution of the E-wave near its wave front $\xi = v_1\eta$ and the second term represents the contribution of the T-wave near its wave front $\xi = v_2\eta$. We observe also that both the waves experience decay exponentially with the distance (attenuation), which is also the case in CTE, ETE and TRDTE, but not the case in TEWOED where the waves do not experience attenuation (see [Dhaliwal and Rokne 1988;1989; Chandrasekhariah and Srinath 1996;1997]. From (23)–(28), we further note that all of U , Θ , τ are identically zero for $\xi > v_2\eta$. This implies that at a given instant of time $\eta*$, the points of the solid $\xi > 0$ that are beyond the faster wave front $\xi = v_2\eta*$ do not experience any disturbances. This observation confirms that, like ETE, TRDTE, TEWOED the thermoelasticity theory with dual phase-lag effects is also a wave thermoelasticity theory. Moreover it is interesting to record that at a given instant, the region $0 < \xi < v_2\eta*$ is the domain of influence of the disturbance, contrary to the result that this domain extends and the effects occur instantaneously everywhere in the solid as predicted by CTE, see [Boley and Tolins 1962].

By direct inspection of the solutions (23)–(28), we find that in both cases U is continuous whereas both Θ and τ are discontinuous at both the wave fronts. The finite jumps experienced by temperature and stress at the wave fronts are as follows:

Case (i):

$$[\Theta]_{\xi=v_1\eta} = -\frac{\Theta_0}{L_0} \frac{v_1^2 - 1}{v_1^2} e^{-\frac{\lambda_2}{2\lambda_1}\eta},$$

$$[\Theta]_{\xi=v_2\eta} = \frac{\Theta_0}{L_0} \frac{v_2^2 - 1}{v_2^2} e^{-\frac{\mu_2}{2\mu_1}\eta},$$

$$[\tau]_{\xi=v_1\eta} = \frac{\Theta_0}{L_0} e^{-\frac{\lambda_2}{2\lambda_1}\eta},$$

$$[\tau]_{\xi=v_2\eta} = -\frac{\Theta_0}{L_0} e^{-\frac{\mu_2}{2\mu_1}\eta}.$$

Case (ii):

$$\begin{aligned}
 [\Theta]_{\xi=v_1\eta} &= \frac{\tau_0}{L_0} \frac{(v_1^2 - 1)(v_2^2 - 1)}{v_1^2 v_2^2} e^{-\frac{\lambda_2}{2\lambda_1} \cdot \eta}, \\
 [\Theta]_{\xi=v_2\eta} &= -\frac{\tau_0}{L_0} \frac{(v_1^2 - 1)(v_2^2 - 1)}{v_1^2 v_2^2} e^{-\frac{\mu_2}{2\mu_1} \cdot \eta}, \\
 [\tau]_{\xi=v_1\eta} &= -\frac{\tau_0}{L_0} \frac{v_2^2 - 1}{v_2^2} e^{-\frac{\lambda_2}{2\lambda_1} \cdot \eta}, \\
 [\tau]_{\xi=v_2\eta} &= +\frac{\tau_0}{L_0} \frac{v_1^2 - 1}{v_1^2} e^{-\frac{\mu_2}{2\mu_1} \cdot \eta}.
 \end{aligned}
 \tag{29}$$

Here $[f]$ denotes the jump of the function f across a wave front. The finite jumps are not constants but they decay exponentially with distance from the boundary. The same results are observed to occur in ETE, TRDTE [Dhaliwal and Rokne 1988;1989], but not in TEWOED where the jumps are all constants [Chandrasekhariah and Srinath 1996;1997]. However the discontinuity in temperature and stress at both the wave fronts is a situation common in the context of ETE, TRDTE and TEWOED [Norwood and Warren 1969; Sherief and Dhaliwal 1981; Dhaliwal and Rokne 1988;1989; Chandrasekhariah and Srinath 1996]. Clearly the finite jumps of Θ and τ in the present analysis depend on the delay times τ_T^* , τ_q^* and the thermo elastic coupling ε . The magnitudes of these jumps are exact, valid for short times and hold for all possible values of ε , τ_T^* and τ_q^* . The expressions for jumps for case (ii) reveal another interesting phenomenon. The temperature is discontinuous at both the wave fronts in spite of the fact that the boundary load is purely mechanical in nature. This implies that the application of a discontinuous mechanical load on the boundary does generate discontinuities in temperature. This phenomenon is present in ETE as well, see [Sharma 1987] but absent in TRDTE [Chandrasekharaiah and Keshavan 1992]. According to TRDTE, the temperature is continuous when the applied load is purely mechanical in nature [Chandrasekharaiah and Keshavan 1992]. A similar observation has been made in the half-space problem in the context of TEWOED [Chandrasekharaiah 1996; Chandrasekharaiah and Srinath 1997].

If the effect of τ_q^{*2} is neglected with $\tau_T^* \neq 0$, $v_1 \rightarrow 1$, and $v_2 \rightarrow \infty$, then there exists only one wave front $\xi = \eta$ (E-wave front) and the T-wave propagates with infinite speed as expected from Equations (5)–(6) with $\tau_q^{*2} = 0$. In this case the jumps of Θ and τ at the elastic wave front $\xi = \eta$ are given as follows:

Case (i):

$$[\Theta]_{\xi=\eta} = 0, \quad [\tau]_{\xi=\eta} = -\Theta_0 e^{-\frac{1}{2}\varepsilon \cdot \frac{\tau_q^*}{\tau_T^*} \xi}.$$

Case (ii):

$$[\Theta]_{\xi=\eta} = 0, \quad [\tau]_{\xi=\eta} = \tau_0 \cdot e^{-\frac{1}{2}\varepsilon \cdot \frac{\tau_q^*}{\tau_T^*} \xi}.
 \tag{30}$$

The jumps of temperature disappear at the E-wave front whereas that of stress τ exists and depends on ε , τ_T^* , τ_q^* . These results are in complete agreement (for $\tau_T^* = \tau_q^*$) with the corresponding results obtained in the context of CTE with classical Fourier’s law, see [Boley and Tolins 1962]. Moreover, it is interesting to record from solutions (25) and (28) that because of delay times, the position of jumps of stress shifts

from $\xi = \eta$ in CTE to $\xi = v_1\eta$ in the present analysis. Also the dual phase-lag model introduced by [Tzou 1995a;1995b] brings to light the jumps of temperature Θ at the E-wave front too, which are not encountered in CTE.

Acknowledgement

The author expresses deep gratitude and thanks to the reviewer for valuable advice and suggestions for the improvement of the paper.

References

- [Biot 1956] M. A. Biot, "Thermoelasticity and irreversible thermodynamics", *J. Appl. Phys.* **27**:3 (1956), 240–253.
- [Boley and Tolins 1962] B. A. Boley and I. S. Tolins, "Transient coupled thermo-elastic boundary value problems in the half-space", *J. Appl. Mech. (ASME)* **29** (1962), 637–646.
- [Chadwick 1960] P. Chadwick, *In progress in solid mechanics*, vol. I, edited by R. Hill and I. N. Sneddon, North Holland, Amsterdam, 1960.
- [Chandrasekharaiah 1986] D. S. Chandrasekharaiah, "Thermo-elasticity with second sound", *Appl. Mech. Rev.* **39**:3 (1986), 355–375. A review.
- [Chandrasekharaiah 1996] D. S. Chandrasekharaiah, "One-dimensional wave propagation in the linear theory of thermo-elasticity without energy dissipation", *J. Therm. Stresses* **19** (1996), 695–710.
- [Chandrasekharaiah 1998] D. S. Chandrasekharaiah, "Hyperbolic thermo-elasticity: a review of recent literature", *Appl. Mech. Rev.* **51**:12 (1998), 705–729.
- [Chandrasekharaiah and Keshavan 1992] D. S. Chandrasekharaiah and H. R. Keshavan, "Axisymmetric thermoelastic interactions in an unbounded body with cylindrical cavity", *Acta Mech.* **92**:1–4 (1992), 61–76.
- [Chandrasekharaiah and Murthy 1993] D. S. Chandrasekharaiah and H. N. Murthy, "Thermoelastic interactions in an unbounded body with a spherical cavity", *J. Therm. Stresses* **16** (1993), 55–70.
- [Chandrasekharaiah and Srinath 1997] D. S. Chandrasekharaiah and K. S. Srinath, "Axisymmetric thermoelastic interactions without energy dissipation in an unbounded body with cylindrical cavity", *J. Elasticity* **46**:1 (1997), 19–31.
- [Chandrasekharaiah and Srinath 1996] D. S. Chandrasekharaiah and K. S. Srinath, "One-dimensional waves in a thermoelastic half-space without energy dissipation", *Int. J. Eng. Sci.* **34**:13 (1996), 1447–1455.
- [Danilovskaya 1950] V. I. Danilovskaya, "Thermal stresses in an elastic half-space due to sudden heating on the surface", *J. Appl. Math. Mech.* **14** (1950), 316–318.
- [Dhaliwal and Rokne 1988] R. S. Dhaliwal and J. G. Rokne, "One-dimensional generalized thermo-elastic problem for a half-space", *J. Therm. Stresses* **11** (1988), 257–271.
- [Dhaliwal and Rokne 1989] R. S. Dhaliwal and J. G. Rokne, "One dimensional thermal shock problem with two relaxation times", *J. Therm. Stresses* **12** (1989), 259–279.
- [Green and Lindsay 1972] A. E. Green and K. A. Lindsay, "Thermoelasticity", *J. Elasticity* **2**:1 (1972), 1–7.
- [Green and Naghdi 1977] A. E. Green and P. M. Naghdi, "On thermodynamics and the nature of the second law", *Proc. R. Soc. Lond. A* **357**:1690 (1977), 253–270.
- [Green and Naghdi 1992] A. E. Green and P. M. Naghdi, "On undamped heat waves in an elastic solid", *J. Therm. Stresses* **15** (1992), 253–264.
- [Green and Naghdi 1993] A. E. Green and P. M. Naghdi, "Thermoelasticity without energy dissipation", *J. Elasticity* **31**:3 (1993), 189–208.
- [Ignaczak 1989] J. Ignaczak, *In thermal stresses*, vol. III, chap. 4, edited by R. B. Hetnarski, Elsevier, Oxford, 1989.
- [Kaminski 1990] W. Kaminski, "Hyperbolic heat conduction equation for materials with a non-homogenous inner structure", *J. Heat Transf. (ASME)* **112** (1990), 555–560.

- [Lord and Shulman 1967] H. W. Lord and Y. Shulman, "A generalized dynamical theory of thermoelasticity", *J. Mech. Phys. Solids* **15**:5 (1967), 299–309.
- [Mitra et al. 1995] K. Mitra, S. Kumar, and A. Vedaverz, "Experimental evidence of hyperbolic heat conduction in processed meat", *J. Heat Transf. (ASME)* **117** (1995), 568–573.
- [Norwood and Warren 1969] F. R. Norwood and W. E. Warren, "Wave propagation in the generalized dynamical theory of thermoelasticity", *Q. J. Mech. Appl. Math.* **22**:3 (1969), 283–290.
- [Ozizik and Tzou 1994] M. N. Ozizik and D. Y. Tzou, "On the wave theory of heat conduction", *J. Heat Transf. (ASME)* **116** (1994), 526–535.
- [RoyChoudhuri 1984] S. K. Roy Choudhuri, "Electro-magneto-thermo-elastic plane waves in rotating media with thermal relaxation", *Int. J. Eng. Sci.* **22**:5 (1984), 519–530.
- [RoyChoudhuri 1985] S. K. Roy Choudhuri, "Effect of rotation and relaxation times on plane waves in generalised thermo-elasticity", *J. Elasticity* **15**:1 (1985), 59–68.
- [RoyChoudhuri 1987] S. K. Roy Choudhuri, "On magneto thermo-elastic plane waves in infinite rotating media with thermal relaxation", pp. 361–366 in *Proceedings of the IUTAM Symposium on the Electromagnetomechanical Interactions in Deformable Solids and Structures* (Tokyo, 1986), edited by Y. Yamamoto and K. Miya, North-Holland, Amsterdam, 1987.
- [RoyChoudhuri 1990] S. K. Roy Choudhuri, "Magneto-thermo-micro-elastic plane waves in finitely conducting solids with thermal relaxation", pp. 461–468 in *Proceedings of the IUTAM Symposium on Mechanical Modeling of New Electromagnetic Materials* (Stockholm), edited by R. K. T. Hsieh, Elsevier, Amsterdam, 1990.
- [RoyChoudhuri and Bandyopadhyay 2005] S. K. Roychoudhuri and N. Bandyopadhyay, "Thermoelastic wave propagation in a rotating elastic medium without energy dissipation", *Int. J. Math. Math. Sci.* **2005**:1 (2005), 99–107.
- [RoyChoudhuri and Banerjee 2004] S. K. Roychoudhuri and M. Banerjee (Chattopadhyay), "Magnetoelastic plane waves in rotating media in thermoelasticity of type II (G-N model)", *Int. J. Math. Math. Sci.* **2004**:71 (2004), 3917–3929.
- [RoyChoudhuri and Banerjee 2005] S. K. Roy Choudhuri and M. Banerjee (Chattopadhyay), "Magneto-viscoelastic plane waves in rotating media in the generalized thermoelasticity II", *Int. J. Math. Math. Sci.* **2005**:11 (2005), 1819–1834.
- [RoyChoudhuri and Debnath 1983] S. K. Roy Choudhuri and L. Debnath, "Magneto-thermo-elastic plane waves in rotating media", *Int. J. Eng. Sci.* **21**:2 (1983), 155–163.
- [RoyChoudhuri and Dutta 2005] S. K. Roychoudhuri and P. S. Dutta, "Thermo-elastic interaction without energy dissipation in an infinite solid with distributed periodically varying heat sources", *Int. J. Solids Struct.* **42**:14 (2005), 4192–4203.
- [Sharma 1987] J. N. Sharma, "Transient generalized thermoelastic waves in a transversely isotropic medium with a cylindrical hole", *Int. J. Eng. Sci.* **25**:4 (1987), 463–471.
- [Sherief and Dhaliwal 1981] H. H. Sherief and R. S. Dhaliwal, "Generalized one-dimensional thermal shock problem for small times", *J. Therm. Stresses* **4** (1981), 407–420.
- [Tzou 1995a] D. Y. Tzou, "Experimental support for the lagging behavior in heat propagation", *J. Thermophys. Heat Transf.* **9**:4 (1995), 686–693.
- [Tzou 1995b] D. Y. Tzou, "A unified approach for heat conduction from macro to micro scales", *J. Heat Transf. (ASME)* **117** (1995), 8–16.

Received 12 Jan 2006. Accepted 28 Jun 2006.

SNEHANSHU KR. ROYCHOU DHURI: skrc_bu_math@yahoo.com

Department of Mathematics, Burdwan University, Burdwan-713104, West Bengal, India

FERROMAGNETIC SHAPE MEMORY EFFECTS IN AN IRON PALLADIUM ALLOY

JUN CUI AND TOM SHIELD

This paper presents the results of an extensive series of experiments conducted on $\text{Fe}_{70}\text{Pd}_{30}$ using a recently developed apparatus, the Magneto-Mechanical Testing Machine. These experiments were designed to investigate the ferromagnetic shape-memory behavior of $\text{Fe}_{70}\text{Pd}_{30}$ and test the predictions of a theory that assumes the magnetizations of the material are constrained to lie in the easy directions and the material strains are constrained to be the shape-memory transformation strains. It was found that a specimen made of $\text{Fe}_{70}\text{Pd}_{30}$ single crystal lengthens when a magnetic field is applied along its c axis (short axis of FCT lattice) while the specimen is under uniaxial compression in the c direction. This behavior agrees with the predictions of the constrained theory and magnetic anisotropy measurements. The maximum field-induced strain change measured in this material is about 0.009 at 5500 G and -1 MPa, which is one fifth of the theoretical prediction. This is attributed to the magnetization rotation away from the easy directions caused by insufficient magnetic anisotropy. Under -12 MPa of compression the field-induced strain change is considerably smaller reaching only about 0.0008, but this change gives the largest work output observed of $9.6 \times 10^3 \text{ J/m}^3$. This work output is very close to the work output of Terfenol-D under this amount of compressive stress.

1. Introduction

Shape memory alloys (SMAs) are materials that undergo a reversible, diffusionless structural transformation. At temperatures above the transformation temperature, these materials typically have a cubic crystal structure which is called the austenite phase. Below this transformation temperature the material structure has less symmetry in what is called the martensite phase. This reduction in symmetry gives rise to the formation of variants of the martensitic phase. The variants are related by the symmetry of the parent austenitic phase and in the absence of biasing stress all have the same free energy. If the SMA is also ferromagnetic (an FSMA) then there exists the possibility of rearranging the martensite variants using a magnetic field. Such a rearrangement can give changes in shape that involve strains one order of magnitude larger than those possible in “giant” magnetostrictive materials such as Terfenol-D.

Rearranging martensite variants by a magnetic field is a novel idea and has received considerable attention recently. The concept of combining martensitic phase transformation with magnetic field was first introduced by James and Kinderlehrer [1993]. Vasil'ev et al. [1996] measured the magnetostriction in $\langle 110 \rangle$ directions of a specimen made of Ni_2MnGa single crystal, and suggested the existence of field-induced shape memory effect. Ullakko et al. [1996] demonstrated field-induced variant rearrangement involving a 0.002 strain in unstressed Ni_2MnGa single crystal. In 1998, a larger reversible field-induced strain of 0.006 under cyclic fields of 10 kG has been achieved in $\text{Fe}_{70}\text{Pd}_{30}$ [James and Wuttig 1998].

Keywords: ferromagnetic shape memory, FePd, iron-palladium alloy.

Tickle et al. [1999] reported a field-induced strain of 0.045 in Ni₂MnGa. In their experiments, the movement of twin interfaces was directly observed verifying the basic FSMA mechanism. More recently, a field-induced strain of 0.094 was demonstrated by Sozinov et al. [2002]. Theories of field induced variant rearrangement have been developed in [James and Wuttig 1998; DeSimone and James 2002; O'Handley 1998]. According to the theory of James, Desimone, and Wuttig, if the magneto-crystalline anisotropy and elastic moduli of the material are large relative to the applied stress and magnetic fields, the strain-magnetization pair, $(\mathbf{E}(\mathbf{x}), \mathbf{m}(\mathbf{x}))$, can be reasonably assumed to only take the values at local energy minima (bottoms of the energy wells). In this case the free energy is reduced to only involving loading device and demagnetization terms and the minimization process becomes a process of determining the compatible microstructure and domain patterns that minimize this energy. From the solution to this constrained energy minimization problem, the microstructures and shape change of the material can be predicted as a function of the applied stress and magnetic fields. This constrained theory was applied to Fe₇₀Pd₃₀ [James and Wuttig 1998] and Ni₂MnGa [Tickle et al. 1999], and qualitative agreement between the theory and the experimental results was achieved. O'Handley [1998] approached the problem phenomenologically. He proposed a model for the magnetization process and field-induced twin boundary motion. His model includes an energy contribution associated with moving a twin interface and provides a simple explanation for the nearly linear field dependence of strain on magnetic field below saturation.

To date, Ni₂MnGa and Fe₇₀Pd₃₀ are the only two alloys reported exhibiting giant (>0.005) field-induced strain. The alloy Ni₂MnGa has received considerable attention recently, while Fe₇₀Pd₃₀ has attracted less attention. While the martensitic behavior of Fe₇₀Pd₃₀ had been previously studied extensively [Somura and Oshima and Fujita 1980; Oshima 1981; Oshima and Sugiyama 1982; Sugiyama et al. 1984; 1985; 1986; Muto et al. 1988; Oshima et al. 1988; 1990; Muto et al. 1990a; 1990b; 1990c; Tanaka and Oshima 1991; Oshima et al. 1992; Tanaka et al. 1992], only recently has the FSM behavior of this alloy been investigated [Cui and James 2001; Cui et al. 2004]. It was found that the FCC-FCT transformation in Fe₇₀Pd₃₀ is a weak first order thermoelastic transition. The average lattice parameters are $a = 3.822 \pm 0.001 \text{ \AA}$ and $c = 3.629 \pm 0.001 \text{ \AA}$ for the FCT martensite, and $a_0 = 3.755 \pm 0.001 \text{ \AA}$ for the cubic austenite. The latent heat is $10.79 \pm 0.01 \text{ J/cm}^3$. The Curie temperature is 450°C . The saturation magnetization is $1220 \pm 10 \text{ emu/cm}^3$ for the martensite and $1080 \pm 10 \text{ emu/cm}^3$ for the austenite; the easy axis is in [100] or [010] direction (a axes of FCT lattice). The magnetic anisotropy is $-5 \pm 2 \times 10^3 \text{ erg/cm}^3$ for the austenite at 60°C , and $3.46 \pm 0.02 \times 10^5 \text{ erg/cm}^3$ for the martensite at -20°C .

In this paper, a brief summary of the constrained theory is given, followed by the results of a series of magneto-mechanical tests. A comparison of the predictions of constrained theory with the results of these experiments is made and finally some conclusions are drawn.

2. Constrained theory

Constrained theory [DeSimone and James 2002] restricts the range of possible strain-magnetization pairs to those values that give absolute minima in the strain and anisotropy energies. This considerably reduces the complexity of the problem of predicting the microstructures and domain patterns that will form. In this section this theory and the process of solving problems using it will be summarized.

2.1. Constrained free energy. The free energy for a ferromagnetic shape memory material can be written as

$$F = \int_{\Omega} \{\phi(\mathbf{F}, \mathbf{m}, \theta) - \mathbf{h}_0 \cdot \mathbf{m} - \sigma_0 \cdot \mathbf{E}\} d\mathbf{x} + \frac{1}{8\pi} \int_{\mathbb{R}^3} |\nabla \zeta_m|^2 d\mathbf{x},$$

where Ω is the region occupied by the material in the reference configuration. The first term, $\phi(\mathbf{F}, \mathbf{m}, \theta)$, combines the strain and magnetic anisotropy energies and depends on the deformation gradient \mathbf{F} , magnetization \mathbf{m} , and temperature θ . The applied magnetic field is \mathbf{h}_0 , the applied stress is σ_0 , and the strain is \mathbf{E} . The second integral is the demagnetization energy, where the magnetostatic potential ζ_m is obtained by solving the magnetostatic equation

$$\operatorname{div}(-\nabla \zeta_m + 4\pi \mathbf{m}) = 0,$$

which itself follows from Maxwell's equations,

$$\begin{aligned} \operatorname{curl} \mathbf{h} &= 0 \quad (\mathbf{h} = -\nabla \zeta_m), \\ \operatorname{div}(\mathbf{h} + 4\pi \mathbf{m}) &= 0. \end{aligned}$$

According to the constrained theory of DeSimone and James [2002], if the magnetic anisotropy and elastic moduli of the material are large compared to the applied stress and magnetic field, the strain-magnetization pair $(\mathbf{E}(\mathbf{x}), \mathbf{m}(\mathbf{x}))$ will be close to the energy wells, which are given by

$$\phi(\mathbf{F}, \mathbf{m}, \theta) = 0. \tag{1}$$

Assuming the solution is constrained to be exactly on bottom of these energy wells, the constrained free energy is then given by

$$\int_{\Omega} \{-\mathbf{h}_0 \cdot \mathbf{m} - \sigma_0 \cdot \mathbf{E}\} d\mathbf{x} + \frac{1}{8\pi} \int_{\mathbb{R}^3} |\nabla \zeta_m|^2 d\mathbf{x}. \tag{2}$$

2.2. Average microstructural behavior. If the magnetization-strain states are combined to form a compatible microstructure and it is assumed that this microstructure consists of very fine twins, then the problem can be further relaxed (in the terminology of [DeSimone and James 2002]) and only the average values of the strain and magnetization need be considered. These average values lie in the convex hull of the energy wells, \mathcal{C} , which is defined as follows. The alloy $\text{Fe}_{70}\text{Pd}_{30}$ has three tetragonal variants and each variant has a strain and two easy directions that satisfy (1) (at a fixed temperature). Thus the constrained theory strain-magnetization pairs for this alloy are

$$\begin{aligned} \mathbf{E}_1 &= \begin{pmatrix} \epsilon_2 & 0 & 0 \\ 0 & \epsilon_1 & 0 \\ 0 & 0 & \epsilon_1 \end{pmatrix} \text{ and any of } \begin{aligned} &\pm[\mathbf{m}_1^2 = m_s \mathbf{e}_2], \\ &\pm[\mathbf{m}_1^3 = m_s \mathbf{e}_3], \end{aligned} \\ \mathbf{E}_2 &= \begin{pmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_2 & 0 \\ 0 & 0 & \epsilon_1 \end{pmatrix} \text{ and any of } \begin{aligned} &\pm[\mathbf{m}_2^1 = m_s \mathbf{e}_1], \\ &\pm[\mathbf{m}_2^3 = m_s \mathbf{e}_3], \end{aligned} \\ \mathbf{E}_3 &= \begin{pmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_1 & 0 \\ 0 & 0 & \epsilon_2 \end{pmatrix} \text{ and any of } \begin{aligned} &\pm[\mathbf{m}_3^1 = m_s \mathbf{e}_1], \\ &\pm[\mathbf{m}_3^2 = m_s \mathbf{e}_2], \end{aligned} \end{aligned} \tag{3}$$

where $\epsilon_1 = a/a_0 - 1$ and $\epsilon_2 = c/a_0 - 1$ are the transformation strains and m_s is the saturation magnetization of the martensite. The strain tensor and magnetization vector components are in the cubic crystallographic coordinates of the parent austenite phase, which has basis vectors $\mathbf{e}_k, k = 1, 2, 3$.

The set \mathcal{C} , which is comprised of average strain, $\langle \mathbf{E} \rangle$, and average magnetization, $\langle \mathbf{m} \rangle$, pairs, is given by a sum of the combinations in (3); thus $(\langle \mathbf{E} \rangle, \langle \mathbf{m} \rangle)$ lies in \mathcal{C} if and only if two conditions are satisfied:

$$\begin{aligned} \langle \mathbf{m} \rangle &= \lambda_1 \mathbf{m}_1^2 + \lambda_2 (-\mathbf{m}_1^2) + \lambda_3 \mathbf{m}_1^3 + \lambda_4 (-\mathbf{m}_1^3) + \lambda_5 \mathbf{m}_2^1 + \lambda_6 (-\mathbf{m}_2^1) \\ &\quad + \lambda_7 \mathbf{m}_2^3 + \lambda_8 (-\mathbf{m}_2^3) + \lambda_9 \mathbf{m}_3^1 + \lambda_{10} (-\mathbf{m}_3^1) + \lambda_{11} \mathbf{m}_3^2 + \lambda_{12} (-\mathbf{m}_3^2), \\ \langle \mathbf{E} \rangle &= (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) \mathbf{E}_1 + (\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8) \mathbf{E}_2 + (\lambda_9 + \lambda_{10} + \lambda_{11} + \lambda_{12}) \mathbf{E}_3, \end{aligned} \tag{4}$$

where $\lambda_i, i = 1 \dots 12$ are the volume fractions of the domains inside the variants that have magnetizations $\pm \mathbf{m}_j^k$ and strains \mathbf{E}_i in the combinations given in (3). The volume fractions are subject to the constraints

$$\sum_{i=1}^{12} \lambda_i = 1, \quad 0 \leq \lambda_i \leq 1, \quad i = 1 \dots 12. \tag{5}$$

Figure 1 presents a geometrical interpretation of the convex hull.

Equation (4) can be simplified as follows. Let $\xi_j = \lambda_{2j-1} + \lambda_{2j}$ and $\eta_j = \lambda_{2j-1} - \lambda_{2j}$, where $j = 1 \dots 6$. Then note that $\xi_1 + \xi_2, \xi_3 + \xi_4$ and $\xi_5 + \xi_6$ are the volume fractions of variants $\mathbf{E}_1, \mathbf{E}_2$ and \mathbf{E}_3 and $\eta_1 + \eta_2, \eta_3 + \eta_4$ and $\eta_5 + \eta_6$ are the net magnetization associated with these strains, respectively. The average magnetization in (4) then reduces to

$$\langle \mathbf{m} \rangle = \eta_1 \mathbf{m}_1^2 + \eta_2 \mathbf{m}_1^3 + \eta_3 \mathbf{m}_2^1 + \eta_4 \mathbf{m}_2^3 + \eta_5 \mathbf{m}_3^1 + \eta_6 \mathbf{m}_3^2,$$

which for the specific forms of \mathbf{m}_j^k in (3) has components

$$\langle \mathbf{m} \rangle = m_s \begin{pmatrix} \eta_3 + \eta_5 \\ \eta_1 + \eta_6 \\ \eta_2 + \eta_4 \end{pmatrix}.$$

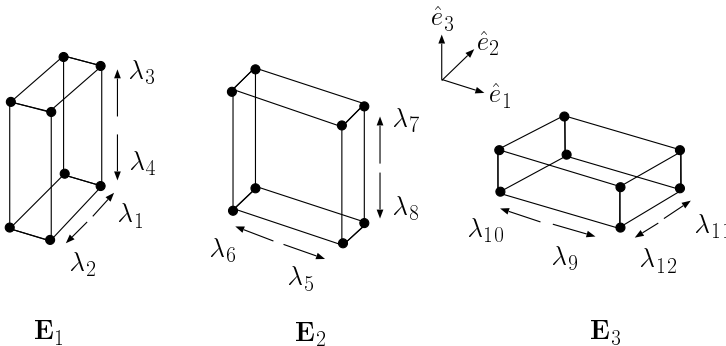


Figure 1. Schematic drawing of an FCT lattice and its three possible variants is shown. The tetragonality of the lattice is exaggerated. The arrows labeled with $\lambda_{1\dots 12}$ are the possible magnetization directions (easy axes) for each variant in the constrained theory.

The average strain defined in (4) also reduces to

$$\langle \mathbf{E} \rangle = (\xi_1 + \xi_2)\mathbf{E}_1 + (\xi_3 + \xi_4)\mathbf{E}_2 + (\xi_5 + \xi_6)\mathbf{E}_3, \tag{6}$$

where the nonzero components of $\langle \mathbf{E} \rangle$ are

$$\begin{aligned} \langle \mathbf{E} \rangle_{11} &= (\xi_1 + \xi_2)\epsilon_2 + (\xi_3 + \xi_4 + \xi_5 + \xi_6)\epsilon_1, \\ \langle \mathbf{E} \rangle_{22} &= (\xi_3 + \xi_4)\epsilon_2 + (\xi_1 + \xi_2 + \xi_5 + \xi_6)\epsilon_1, \\ \langle \mathbf{E} \rangle_{33} &= (\xi_5 + \xi_6)\epsilon_2 + (\xi_1 + \xi_2 + \xi_3 + \xi_4)\epsilon_1, \end{aligned}$$

using the forms of \mathbf{E}_i in (3). The constraints (5) in terms of ξ_j and η_j are

$$\sum_{j=1}^6 \xi_j = 1, \quad 1 \geq \xi_j \geq |\eta_j|, \quad j = 1 \dots 6.$$

In the limit as the microstructure becomes infinitely fine the demagnetization energy can be calculated using the results for a uniformly magnetized body [DeSimone and James 2002], which is $\frac{1}{2}\langle \mathbf{m} \rangle \cdot \mathbf{D}\langle \mathbf{m} \rangle$, where \mathbf{D} is the magnetometric demagnetization matrix for the specimen geometry. Then the constrained free energy (2) reduces to

$$-\mathbf{h}_0 \cdot \langle \mathbf{m} \rangle - \sigma_0 \cdot \langle \mathbf{E} \rangle + \frac{1}{2}\langle \mathbf{m} \rangle \cdot \mathbf{D}\langle \mathbf{m} \rangle,$$

which is to be minimized over $(\langle \mathbf{E} \rangle, \langle \mathbf{m} \rangle) \in \mathcal{C}$. This is an exact result if the specimen has an ellipsoidal shape and an approximation for any other geometry.

For the case of a rectangular specimen with a square base and a height in the x_1 direction which is much longer than the two other dimensions, the demagnetization matrix can be approximated by the demagnetization matrix of an infinitely long cylinder, $\mathbf{D} = \text{diag}(0, 2\pi, 2\pi)$. If the applied field is restricted to the (x_1, x_2) plane and the applied stress is uniaxial in the x_1 direction with magnitude σ , then the constrained free energy, \mathcal{E} , becomes

$$\mathcal{E} = -h_1 m_s (\eta_3 + \eta_5) - h_2 m_s (\eta_1 + \eta_6) + \pi m_s^2 ((\eta_1 + \eta_6)^2 + (\eta_2 + \eta_4)^2) - \sigma (\xi_3 + \xi_4 + \xi_5 + \xi_6)(\epsilon_1 - \epsilon_2) - \sigma \epsilon_2. \tag{7}$$

2.3. Energy minimization. Finding the minimizers of the constrained free energy given by Equation (7) can be broken into two steps: First, a minimization over ξ_j , $j = 1 \dots 6$, is performed using linear programming. In the second step the energy is minimized with respect to the remaining variables, η_j , $j = 1 \dots 6$.

Figure 2 and Table 1 summarize results of this process for a magnetic field with $h_1 \geq 0$ and $h_2 \geq 0$ combined with a uniaxial stress, $\sigma \leq 0$. As shown on this figure, the solution can be divided into four distinct regions, which involve three possible microstructures. In region A, only martensite variant \mathbf{E}_1 is present, and its net magnetization is given by $\eta_1 m_s \mathbf{e}_2$ which depends on the magnitude and direction of magnetic field h_2 . In region D, martensite variants \mathbf{E}_2 and \mathbf{E}_3 may coexist. Their net magnetizations are $\eta_3 m_s \mathbf{e}_1$ and $\eta_5 m_s \mathbf{e}_1$, respectively. In the remaining regions, B and C, all three variants may be present. The admissible microstructures that these variants may form in each of these solution regions are discussed next.

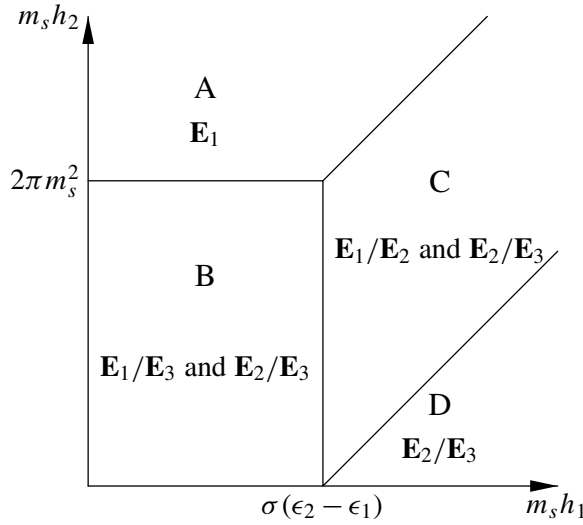


Figure 2. Summary of the minimum energy solution for $h_1 \geq 0$, $h_2 \geq 0$ with constant applied stress $\sigma < 0$, is shown. The volume fraction values for the areas labeled A–D are given in Table 1. The pairs of strains listed in each region are those that minimize the energy and are magnetically compatible across one or both of their possible twin interfaces.

Table 1 shows the results of the energy minimization for the variables η_j only. This is because the variables ξ_j cannot be always fully determined. For example, in region A energy minimization gives $\eta_1 (= \lambda_1 - \lambda_2) = 1$, and because of the constraint $\lambda_i \leq 1$, $i = 1 \dots 12$, the solution is $\lambda_1 = 1$ and $\lambda_i = 0$, $i = 2 \dots 12$ and the ξ_j are fully determined. In region C, energy minimization gives 5 equations and 3 inequalities, plus the constraints. Many combinations of the λ_i satisfy these conditions, so ξ_j cannot be determined. However, without knowing the values of ξ_j , the overall strain of the specimen can still

Area	η_1	η_3	η_5	\mathcal{E}
A	1	0	0	$-h_2 m_s + \pi m_s^2 - \sigma \epsilon_2$
B	$\frac{h_2}{2\pi m_s}$	0	0	$-\frac{h_2^2}{4\pi} - \sigma \epsilon_2$
C	$\frac{(h_2 - h_1)m_s - \sigma (\epsilon_1 - \epsilon_2)}{2\pi m_s^2}$ $0 <, < 1$	$1 - \eta_1 - \eta_5$ ≥ 0	≥ 0	$-\frac{((h_2 - h_1)m_s - \sigma (\epsilon_1 - \epsilon_2))^2}{4\pi m_s^2}$ $-h_1 m_s - \sigma \epsilon_1$
D	0	$1 - \eta_5$ ≥ 0	≥ 0	$-h_1 m_s + \sigma \epsilon_1$

Table 1. Values for the areas A–D of Figure 2. For all areas shown, $\eta_2 = \eta_4 = \eta_6 = 0$.

	$\mathbf{E}_1/\mathbf{E}_2$	$\mathbf{E}_1/\mathbf{E}_3$	$\mathbf{E}_2/\mathbf{E}_3$
$\sqrt{2}\mathbf{a}/\gamma$	$[\bar{1}10]$ $[110]$	$[101]$ $[\bar{1}01]$	$[0\bar{1}1]$ $[011]$
$\sqrt{2}\hat{\mathbf{n}}$	$[\bar{1}\bar{1}0]$ $[1\bar{1}0]$	$[10\bar{1}]$ $[\bar{1}0\bar{1}]$	$[0\bar{1}\bar{1}]$ $[01\bar{1}]$

Table 2. Twin solutions from linear geometric martensite theory. The normal to the twin plane is $\hat{\mathbf{n}}$ and the shape strain is \mathbf{a} . Each pair of variants can form two different twins. The magnitude of the shape strain is $\gamma = |\epsilon_1 - \epsilon_2| = 0.0515$.

be evaluated. For example, in region D, both variants \mathbf{E}_2 and \mathbf{E}_3 can coexist, and it is not possible to fully determine the values λ_i . Nonetheless, the specimen can be determined to be at its maximum length because both variant \mathbf{E}_2 and \mathbf{E}_3 have their long axes aligned with the long axis of the specimen.

2.4. Magnetoelastic compatibility and energy minimizing microstructures. Without considering kinematic and magnetic compatibility, free energy minimization predicts three possible combinations of variants. However, some of these combinations may not be compatible and must be excluded as possible energy minimizing microstructures. The microstructures that are magnetoelastically compatible for $h_1 \geq 0$, and $h_2 \geq 0$ are discussed below. Microstructures in the other three quadrants can be found by symmetry.

If martensite variants \mathbf{E}_i and \mathbf{E}_j meet at a planar interface, linearized kinematic compatibility requires

$$\mathbf{E}_i - \mathbf{E}_j = \frac{1}{2}(\mathbf{a} \otimes \hat{\mathbf{n}} + \hat{\mathbf{n}} \otimes \mathbf{a}), \quad (8)$$

where the vector \mathbf{a} is the shape strain and $\hat{\mathbf{n}}$ is the normal to the interface. Given a pair of variant strains, equation (8) can be solved for \mathbf{a} and $\hat{\mathbf{n}}$. In the tetragonal case considered here, there are two solutions of (8) for each pair of variants, which gives a total of six possible twin interfaces in this material. These are listed in Table 2.

Once the twin interface is determined from kinematics, magnetic compatibility across this interface must also be satisfied. This requires that the jump in the normal component of the magnetization across the interface must be zero, that is,

$$(\mathbf{m}_i - \mathbf{m}_j) \cdot \hat{\mathbf{n}} = 0, \quad (9)$$

where \mathbf{m}_i and \mathbf{m}_j are net magnetization of variants \mathbf{E}_i and \mathbf{E}_j , respectively. The minima presented in Figure 2 and Table 1 will next be checked that they satisfy (9) for interfaces between the predicted variants.

The solution in area A of Figure 2 involves only variant \mathbf{E}_1 , thus there are no twin interfaces across which compatibility must be satisfied.

The microstructure in area D consists of variants \mathbf{E}_2 with a net magnetization of $\eta_3 m_s \mathbf{e}_1$, and \mathbf{E}_3 with a net magnetization of $\eta_5 m_s \mathbf{e}_1$. The two possible twin interfaces between these two variants have normals $[0\bar{1}\bar{1}]$ and $[01\bar{1}]$. The difference between these two magnetizations is $(\eta_3 - \eta_5)m_s \mathbf{e}_1$, which is perpendicular to both of these twin plane normals for any values of η_3 and η_5 . Thus this microstructure is magnetically compatible.

Areas B and C may contain microstructures that involve all three variants. Possible combinations are \mathbf{E}_1 with \mathbf{E}_2 , \mathbf{E}_1 with \mathbf{E}_3 , and \mathbf{E}_2 with \mathbf{E}_3 . The last case is the same as that for area D, and because it was found to be always magnetically compatible, both areas B and C may contain twins with variants \mathbf{E}_2 and \mathbf{E}_3 . The other two possible twins will be considered in turn.

For the combination of \mathbf{E}_1 and \mathbf{E}_2 , the possible interface normals between these two variants are $[\bar{1}\bar{1}0]$ and $[1\bar{1}0]$. From the results of energy minimization, the net magnetization within these two variants are $\eta_1 m_s \mathbf{e}_2$ and $\eta_3 m_s \mathbf{e}_1$, respectively. Their difference is $(-\eta_3, \eta_1, 0)$, which is perpendicular to the interface normal $[\bar{1}\bar{1}0]$ if $\eta_1 = \eta_3$ and perpendicular to $[1\bar{1}0]$ if $\eta_1 = -\eta_3$. Because both η_1 and η_3 are required to be positive by the minimization, the second interface is not possible in either area B or C. Area B minimization requires $\eta_3 = 0$; thus unless $h_2 = 0$, the remaining twin interface is not possible in area B. In area C, the requirement that $\eta_1 = \eta_3$ restricts $\eta_5 = 1 - 2\eta_1$, which is possible and thus so is this twin in area C.

For the combination of \mathbf{E}_1 and \mathbf{E}_3 , the normals to the possible interfaces between these two variants are $[10\bar{1}]$ and $[\bar{1}0\bar{1}]$. Energy minimization determined the net magnetization within these two variants to be $\eta_1 m_s \mathbf{e}_2$ and $\eta_5 m_s \mathbf{e}_1$, respectively. Their difference is $(-\eta_5, \eta_1, 0)$, which is perpendicular to these interface normals only if $\eta_5 = 0$. This is already required by minimization in area B, thus this twin is allowed in this area. In area C, η_5 is nonzero in general and thus this twin is not allowed in this area.

Figure 2 summarizes the results of both energy minimization and the requirements of compatibility and it shows the valid twin combinations in each of the four regions. The applied loading, σ , is in the x_1 direction and the change in length of the material in this direction is of interest because it is the deformation that does work against the applied stress. In area A the normal strain in the x_1 direction is ϵ_2 which is less than one for this material. The microstructure in area D is a combination of variants that both have ϵ_1 as their normal strain in the x_1 direction. Thus selecting an applied field path that goes from area A to area D will result in a change in the x_1 direction normal strain of $\epsilon_1 - \epsilon_2 = (a - c)/a_0 = 0.0515$. The experiments discussed below were designed to test this prediction.

3. Equipment and sample preparation

Combined mechanical and magnetic tests were performed with a facility called the Magneto-Mechanical Test Machine (MMTM). Details of this machine can be found in [Shield 2003]. Two $\text{Fe}_{70}\text{Pd}_{30}$ single crystal specimens, MM6 and MM10, were used in the magneto-mechanical tests that are reported here. They were cut in close proximity to each other from the single crystal boule FePd4 (see [Cui et al. 2004] for more information regarding this boule). The composition of both specimens MM6 and MM10 is near 29.6 at.% Pd. Their martensitic/austenitic transformation temperatures are $M_s \approx 20^\circ\text{C}$, $M_f \approx 16^\circ\text{C}$, $A_s \approx 18^\circ\text{C}$, and $A_f \approx 25^\circ\text{C}$. Specimen MM6 has dimensions of $2.12 \times 2.10 \times 6.52 \text{ mm}^3$, and specimen MM10 has dimensions of $2.11 \times 2.10 \times 8.65 \text{ mm}^3$. Both specimens are rectangular bars with $\{100\}$ surfaces. The long dimension of each specimen is used as the loading direction and will be called the longitudinal specimen direction. The transverse direction is perpendicular to the longitudinal direction and it is along a short specimen dimension. Their orientation was checked by x-ray diffractometry at 60°C . The surface normals vary by about 0.7° from $\langle 100 \rangle$ for MM6, and 0.5° for MM10. Both of them were polished at 70°C with $0.1 \mu\text{m}$ diamond suspension.

4. Test design

A series of tests with different temperatures, amounts of compressive stress and magnetic field paths are summarized in Table 3. This table lists twenty-five tests divided into four sets in the order they will be discussed below. For each test there are three groups of columns. Initially all tests start with the specimen at 35°C and at the conditions (stress and field) listed in the first group of columns. The specimen is then

Test	Start at 35°C		Ferro. Strain Measurements			End to 35°C	
	Field (G)	Stress (MPa)	Temp. (°C)	Stress (MPa)	Field Path(s)	Field (G)	Stress (MPa)
----- Set 1 -----							
T5,T10	0	-1.0	11	-1.0	Rot-9radii	0	-1.0
T6	0	-2.0	11	-2.0	Rot-9radii	0	-2.0
T7	0	-4.0	11	-4.0	Rot-9radii	0	-4.0
T11	0	-6.0	11	-6.0	Rot-9radii	0	-6.0
T8	0	-8.0	11	-8.0	Rot-9radii	0	-8.0
T9	0	-12.0	11	-12.0	Rot-9radii	0	-12.0
----- Set 2 -----							
T13b	0	-1.0	20	-1.0	Rot2-5515	0	-1.0
T13c	0	-1.0	15	-1.0	Rot2-5515	0	-1.0
T13d	0	-1.0	10	-1.0	Rot2-5515	0	-1.0
T14a	0	-1.0	5	-1.0	Rot2-5515	0	-1.0
T14b	0	-1.0	0	-1.0	Rot2-5515	0	-1.0
T14c	0	-1.0	-5	-1.0	Rot2-5515	0	-1.0
----- Set 3 -----							
T12,T15	0	-12.0	11	-1.0	Rot20-5515	0	-1.0
T16	0	-1.0	11	-1.0	Rot20-5515	0	-1.0
T17	0	-1.0	11	-1.0	Rot20-3818	0	-1.0
----- Set 4 -----							
T2	0	0	11	-2.0	Lin-4242, Arc-4242	0	-1.0
			11	-1.0	Lin-4242, Arc-4242		
T3	0	-2.0	11	-2.0	Lin-4242, Arc-4242	0	-1.0
			11	-2.0	Rot1-4242		
			11	-1.0	Lin-4242, Arc-4242		
T4	T4242	-2.0	11	-1.0	Rot1-4242	0	-1.0
			11	-2.0	Lin-4242, Arc-4242		
			11	-2.0	Rot1-4242		
			11	-1.0	Lin-4242, Arc-4242		
			11	-1.0	Rot1-4242	0	-1.0

Table 3. Tests conducted on specimens MM6 and MM10. Terms such as T4242, Lin-4242, Arc-4242, and so on are explained in Section 4.

cooled, at approximately 1°C/min to the temperature listed in the second group of columns. The second group of columns includes three columns. The first two are the temperature and the bias stress at which the measurements conducted and the third column describes the path(s) of the magnetic field applied during the test. After the applied magnetic field path(s) have been completed, the specimen temperature

is raised back up to 35°C under the conditions listed in the last group of columns. Tests in Set 1 have two different sets of field paths that were performed under two different amounts of compressive stress. By starting all tests at 35°C, which is above the austenite finish temperature, all martensitic microstructure is removed from the specimen. This makes sure each test starts from identical and repeatable conditions.

In Table 3 the magnetic field conditions and paths are specified with labels. For the only field applied during cooling, the label T4242 indicates a field applied along the transverse direction of the specimen [100] at a magnitude of 4242 G. There are two basic field paths applied to the specimen to measure its FSM response. Linear paths follow radial lines that start or end at zero field. Rotation or arc paths follow circles or parts of circles, which keep the magnitude of the applied field constant and change its direction. Field path Lin-4242 consists of two segments: first, the magnetic field is applied on the transverse direction of the specimen [100] from 0.0 G to 4242 G, then back to 0.0 G along the same direction; second, it increases on the longitudinal [001] direction from 0.0 G to 4242 G, and returns back to 0.0 G on [001] axis. For the Arc-4242 path, the field first increases along the transverse [100] direction from 0.0 G to 4242 G, then rotates 90° counterclockwise (CCW) toward the longitudinal [001] direction with constant magnitude. Once the rotation is complete, the field decreases from 4242 G to 0.0 G along longitudinal direction. The field rate is 60 G/sec on all segments of both Lin-4242 and Arc-4242 paths. These two paths were designed to investigate the path dependence of the FSM effect, as they both include the points at 4242 G in both the transverse and then the longitudinal directions.

The path Rot-9radii has nine circular segments. It starts by increasing the field in transverse [100] direction from zero to 424 G at a rate of 60 G/sec, the field is then rotated 360° CCW at a rate of 6 G/sec, the first segment finishes when the field returns to 424 G in the transverse direction. The second segment starts at the end of the first segment, the field increases in the transverse direction from 424 G to a new magnitude, 848 G, at a rate of 12 G/sec. It then rotates 360° CCW at a rate of 12 G/sec, and finishes by returning to 848 G in the transverse direction. The next seven segments are similar to the two segments described above with larger field magnitudes. Finally the field is decreased back to 0.0 G while pointing in the transverse direction. The radii of the nine circles are 424, 848, 1273, 1697, 2121, 2969, 3818, 4666, and 5515 G, which are completed at field rates of 6, 12, 18, 24, 30, 42, 54, 66, and 78 G/sec, respectively. These rates result in each 360° circle being completed in 444 seconds.

The path Rot1-4242 consists of a single circle. The field first increases in the transverse [001] direction of the specimen from 0.0 G to 4242 G. Then the field rotates 360° CCW with constant magnitude. After the rotation is complete the field decreases from 4242 G to 0.0 G while pointing in the transverse direction. The field rate is 60 G/sec for the entire path. Paths Rot2-5515 and Rot20-5515 are similar paths that repeat the same circular path twice and twenty times, respectively. In both paths the magnetic field is first applied in the transverse [100] direction of the specimen and increased from 0.0 G to 5515 G. Then the field rotates 720° and 7200° CCW, respectively. After the rotations are complete, the field then decreases from 5515 G to 0.0 G while pointing in the transverse direction. The field rates are 60 G/sec for the linear segments, and 78 G/sec for the circles. The field path Rot20-3818 is similar to the path Rot20-5515, except that the magnitude of the field is 3818 G with a rate of 54 G/sec.

Tests T5 to T10 were designed to study work output, but they also provide information on the effects of stress on phase transformation temperatures. However, due to the limitation of the space, phase transformation effects will be discussed in another paper, along with some other tests which were designed to study the effect of magnetic field on phase transformation temperature [Cui et al. ≥ 2008].

5. Experimental results

We now give the results of the experiments and compare them to the constrained theory presented above.

5.1. Ferromagnetic strain, blocking stress and work output. The key property of interest for FSMAs is the field-induced (or ferromagnetic) strain amplitude, which is defined as the amplitude of the strain change in the specimen caused by an application of a cyclic magnetic field holding stress and temperature constant. In many materials, when the stress exceeds a critical level there will be no ferromagnetic strain no matter how large the applied magnetic field is, and this critical stress is called the blocking stress. Another important property is the work output, which is defined as the product of the ferromagnetic strain amplitude and the constant stress at which this ferromagnetic strain is achieved.

The tests listed in set 1 of Table 3 were designed to give a complete picture of the ferromagnetic strain amplitudes achievable in $\text{Fe}_{70}\text{Pd}_{30}$ over a wide range of applied field magnitudes and compressive stresses. In Figure 3, the applied magnetic field and the strain are plotted versus time for tests T6 to T11.

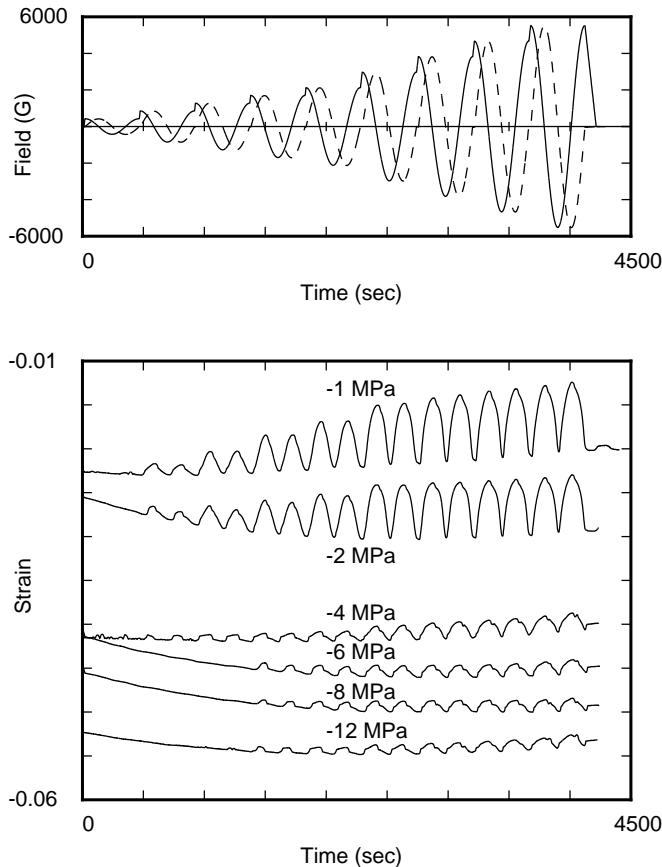


Figure 3. Top: A time history of the applied magnetic fields for field path Rot-9rad. Solid curve is the transverse component of field, and dashed curve is the longitudinal component. Bottom: Time history of the strains measured in tests T5-T9 on specimen MM6 at 11°C .

Because a 180° change in field direction results in a complete ferromagnetic strain cycle, the measured strain goes through two cycles for every 360° of field rotation. The figure shows that the amplitude of the ferromagnetic strain decreases with increasing compressive stress and increases with increasing magnetic field amplitude. The strain at -1 MPa started to oscillate in response to the field after the amplitude of the field exceeded about 800 G; the amplitude of the ferromagnetic strain then increased rapidly with increasing field amplitude until the field reached approximately 3000 G. After this point, the strain amplitude only increased slightly with further increases in the field. The -2 MPa curve is similar to the -1 MPa curve with slightly smaller ferromagnetic strain. The -4 MPa curve has a much smaller strain amplitude than the lower stress curves and the specimen starts to respond to the field at smaller value, 424 G. The -6 MPa curve was significantly different from the -1 , -2 , and -4 MPa curves because it does not oscillate until the field exceeds 1700 G. The -8 and -12 MPa curves are similar to the -6 MPa curve except that the field-induced strains are even smaller.

Table 4 and Figure 4 summarize the field-induced strain amplitude and work output at various stress levels for both specimens. Figure 4 also includes data on Terfenol-D [Clark 1980] and Ni_2MnGa [Tickle and James 1999] for comparison. The ferromagnetic strain amplitude was calculated using the maximum and minimum strain within a single strain cycle due to the largest magnetic field. The ferromagnetic strain amplitude decreases quickly with increasing compressive stress. After -4 MPa, the rate of decrease slows and it is not clear at what stress level the amplitude will decrease to zero. This unusual behavior seems to indicate that there is no blocking stress for this material or that there is a combination of effects active at the same time. In fact, both specimens still responded to field even at -12 MPa. At this level, the field-induced strain and work output of MM10 is similar to that of Terfenol-D, and is larger than that of Ni_2MnGa . Figure 4 shows that Ni_2MnGa has a higher strain at smaller stresses, but stops responding to magnetic field after -8 MPa. Terfenol-D on the other hand stays active beyond -50 MPa, and its largest work output occurs near -20 MPa. The strain available in Terfenol-D is far less than Ni_2MnGa at low stress levels. The behavior of $\text{Fe}_{70}\text{Pd}_{30}$ is in between these two materials. Compared to Ni_2MnGa , $\text{Fe}_{70}\text{Pd}_{30}$ has less strain at small stresses and more strain at higher stress levels. When $\text{Fe}_{70}\text{Pd}_{30}$ is compared to Terfenol-D, the situation is reversed. The behavior of $\text{Fe}_{70}\text{Pd}_{30}$ beyond -12 MPa could not

Stress (MPa)	Sample MM6		Sample MM10	
	Strain	Work output (J/m^3)	Strain	Work output (J/m^3)
-1	0.0084	8.4×10^3	0.0059	5.9×10^3
-2	0.0070	1.4×10^4	0.0042	8.4×10^3
-4	0.0025	1.0×10^4	0.0025	1.0×10^4
-6	0.0019	1.1×10^4	0.0016	9.8×10^3
-8	0.0016	1.3×10^4	0.0011	8.8×10^3
-12	0.0017	2.1×10^4	0.0008	9.6×10^3

Table 4. Ferromagnetic strain amplitude at various bias stresses and the corresponding work output. The MM6 strains are from the crosshead displacement, while the MM10 strains are from the capacitive sensor.

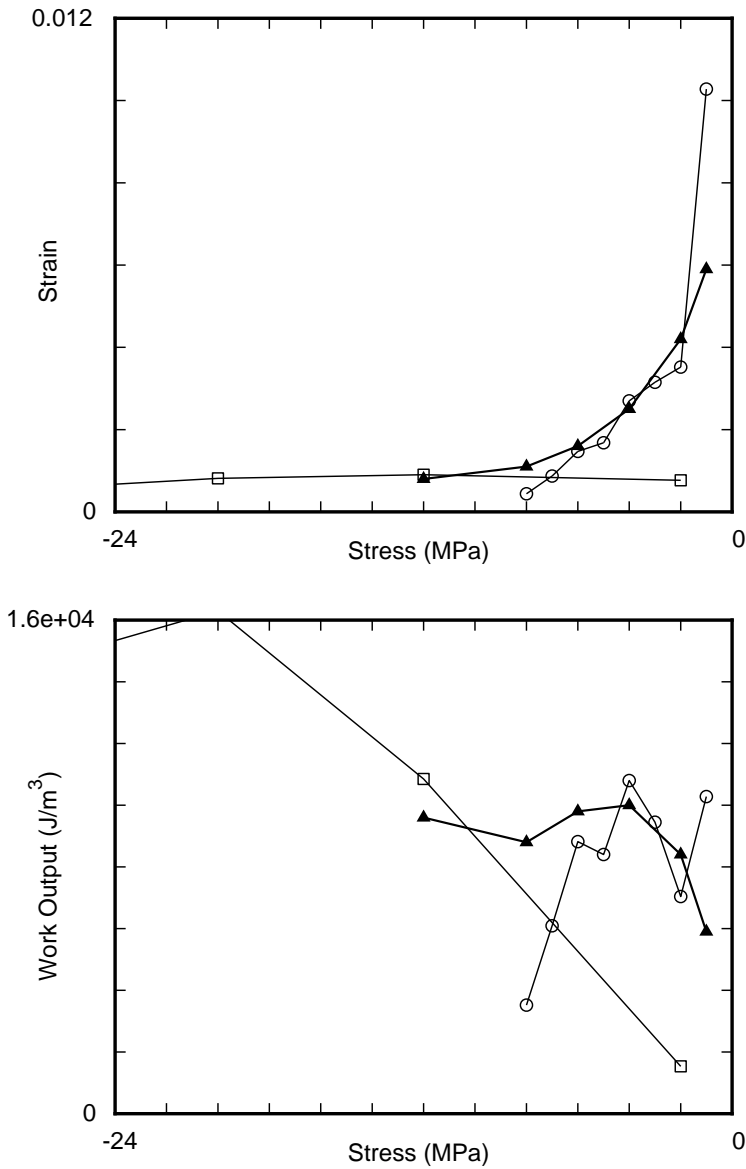


Figure 4. Strains induced by magnetic field at various load conditions and corresponding work output for Fe₇₀Pd₃₀ specimen MM10 (solid triangle), Ni₂MnGa (circle) [Tickle and James 1999], and Terfenol-D (square) [Clark 1980].

be determined because of the presence of the FCT to BCT transformation, which is not reversible. There is some evidence that the yield stress of Fe₇₀Pd₃₀ is higher than -20 MPa (possible slip bands appeared on one specimen that was overloaded).

The details of the strain induced by magnetic field at different levels of stress are provided in Figure 5, which gives the strain versus both the transverse and longitudinal components of applied field. Each

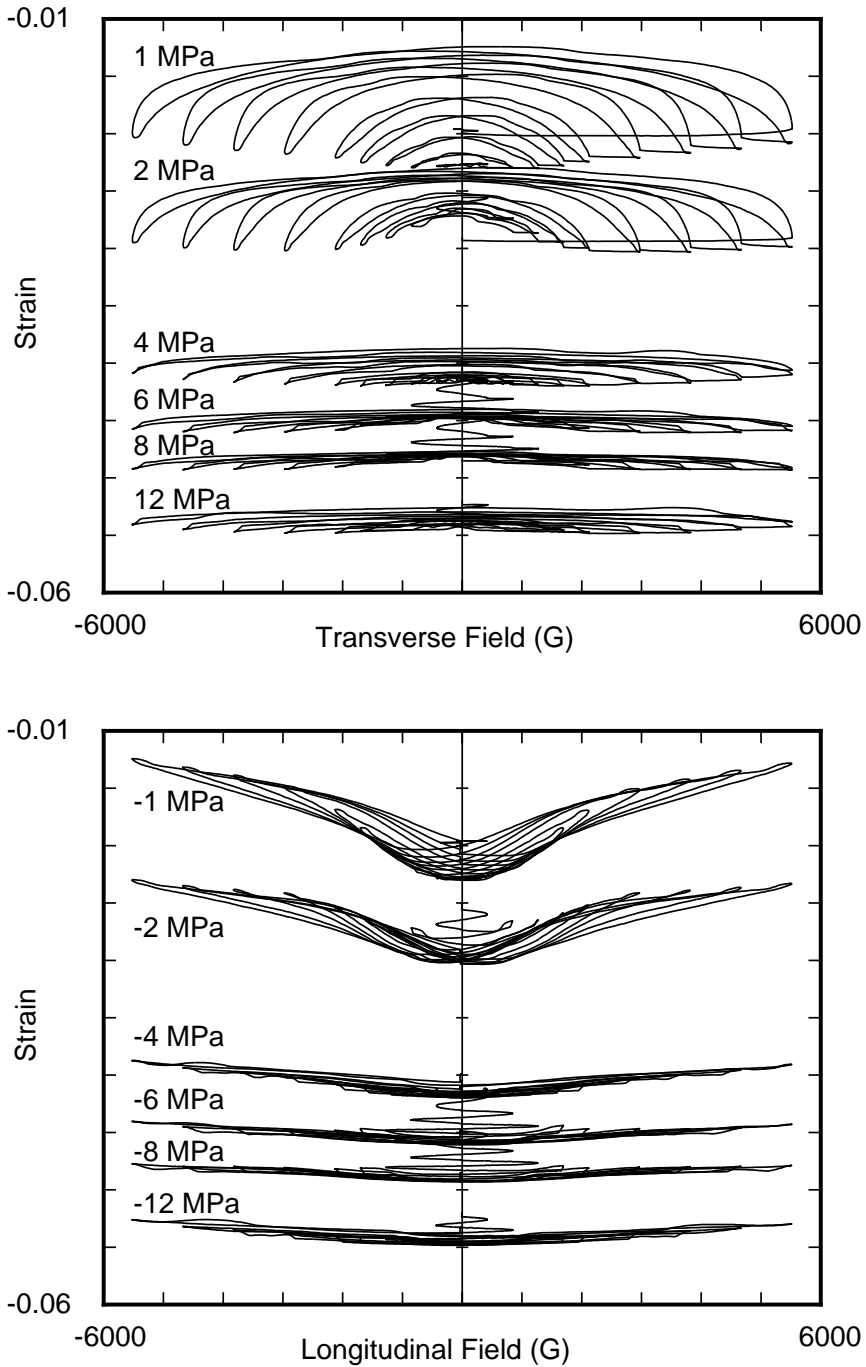


Figure 5. Ferromagnetic strains vs. transverse (top) and longitudinal (bottom) component of field at various load conditions for specimen MM6 are given. Each curve has the same magnetic field path (Rot-9radii) and temperature of 11°C.

set of curves at one load level includes data for all nine field levels in the path Rot-9radii. The curves at -1 MPa show that the changes in strain were small when the transverse component of the field was between -4500 to 4500 G, and were substantial between 4500 to 5515 G and -4500 to -5515 G. If the curves are examined from the perspective of the longitudinal component of field, the changes of the strain were substantial between -1000 to 1000 G, and they were smaller over the rest of the range. This point is better illustrated by Figure 6, which shows the strain versus the angle of the applied magnetic field. In that figure, the maximum and minimum strain were not located at the 90° , 180° , and 270° : there was an approximately 10° shift toward larger angles, which is also the direction of the field rotation (CCW). This indicates the existence of hysteresis. Examining the third quadrant (190° to 280° of the field angle), the specimen was at its shortest length when the field angle was approximately 190° , because of the combined effects of the compressive stress and the transverse field favored the variant with its short axis parallel to the loading direction. When the field rotated away from 190° , the component of the field in the transverse direction decreased and the component of the field in the longitudinal direction increased. As result, the specimen started to lengthen. When the field reached 225° , approximately $2/3$ of the ferromagnetic strain has been achieved, and the remaining $1/3$ of ferromagnetic strain was achieved during the rest of the 45° of this quadrant. In other words, the changes in strain occurred faster in the first 45° of rotation than it did in the second 45° of the field rotation.

The surfaces of the specimens were observed optically during the experiments using a microscope equipped with DIC that allows the surface relief to be observed in false color. The images in Figure 7 were taken when the field was 5500 G in $[001]$ and $[\bar{1}00]$ directions. The strains were -0.013 and

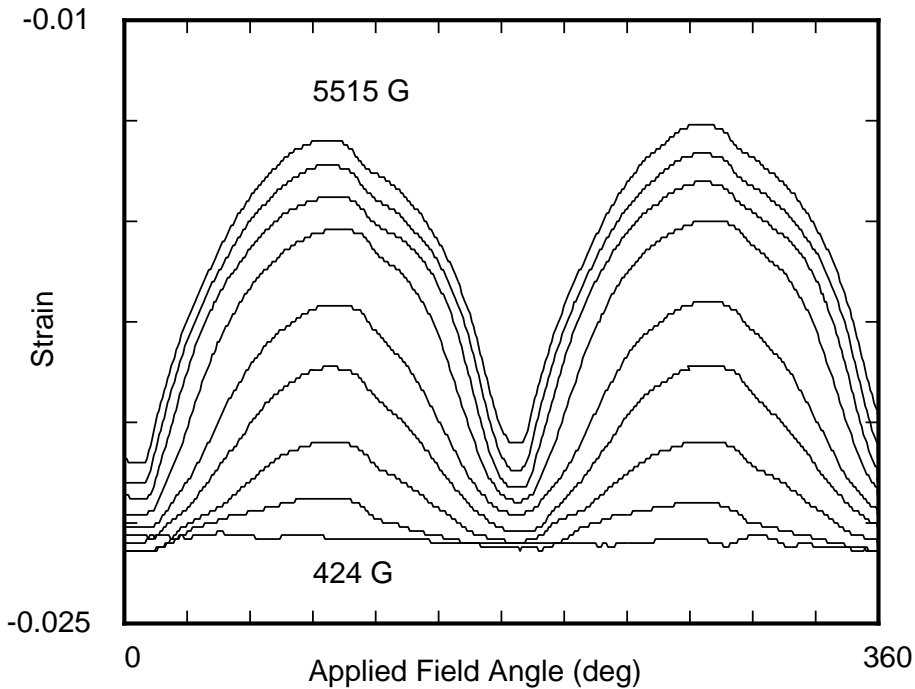


Figure 6. Strain vs. magnetic field rotation angle for test T5 at -1 MPa on specimen MM6.

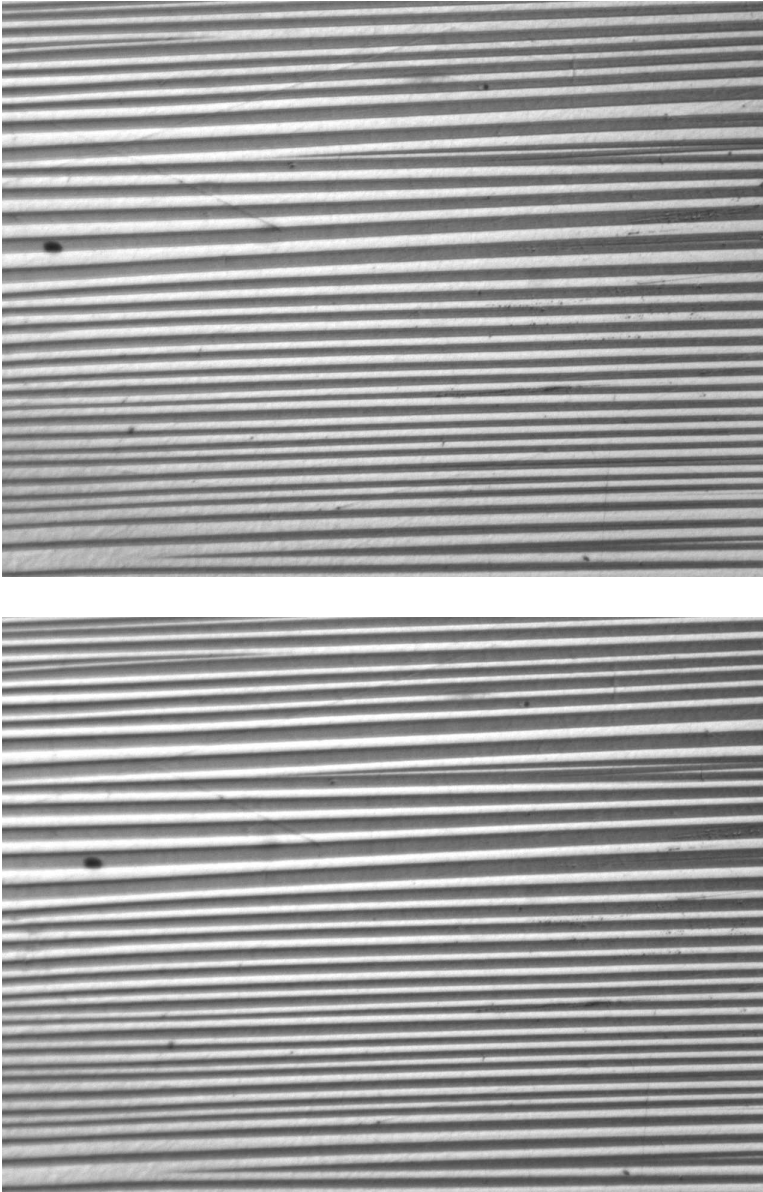


Figure 7. Microstructures of specimen MM6 when magnetic field is 5515 G in [001] (top) and in $[\bar{1}00]$ (bottom). The images are from test T5 at -1 MPa and the field of view is approximately 0.3 mm wide. In these images the vertical [001] direction is up and the transverse [100] direction is horizontal.

-0.020 . The strain difference between these two states is 0.007, which is about 14% of the theoretical ferromagnetic strain of 0.0515. However, comparison of these two images shows no clear differences. When the images are alternated on a computer screen a few small differences are visible in the thickness

Test	Temperature (°C)	Ferromagnetic Strain
T13b	20	0.0048
T13c	15	0.0071
T13d	10	0.0065
T14a	5	0.0069
T14b	0	0.0071
T14c	−5	0.0076

Table 5. Ferromagnetic strain amplitudes measured at various temperatures and under -1 MPa of stress for specimen MM6.

and configuration of the E_1/E_3 twin bands. The fact that the measured strain is so much smaller than that predicted from lattice parameters is the primary unresolved issue about the behavior of $\text{Fe}_{70}\text{Pd}_{30}$.

To assess the variation in the ferromagnetic strain amplitude with temperature, a series of measurements of the ferromagnetic strain at the various temperatures was conducted. The results of tests T13b-d and T14a-c (see Table 3 for test descriptions) are given in Table 5. These results show that the ferromagnetic strain does not depend on temperature, once the phase transformation is complete. Test T13b at 20°C is the temperature at which the martensite transformation begins, so at this temperature there is still a significant amount of austenite in the specimen. This accounts for the decreased strain amplitude measured in this test. The phase transformation is complete by 16°C and all experiments below this temperature have essentially the same strain amplitude.

5.2. Effects of initial microstructures and loading paths. The ferromagnetic shape-memory effect relies on the ability of the magnetic field to rearrange the martensitic variants. Thus the microstructure in the material plays a crucial role in the FSM behavior of the material. If the microstructure is arranged in such a way as to make twin boundary motion easy then the FSM behavior will be enhanced. On the other hand, if a complicated microstructure forms on cooling that does not easily lend itself to field-induced boundary motion, the FSM behavior may be adversely affected. Thus experiments were designed to attempt to generate different microstructures on cooling and compare their responses to applied fields. In addition, different field paths between identical field values were also used to assess the dependence of the FSM effect on the actuation path.

The amount of compressive stress applied to the specimen can be visually observed to affect the microstructure that forms on cooling. Small amounts of compressive stress allow microstructures with vertical twin bands (E_2/E_3 twins) to form, and in some cases remain, in the specimen during transformation. Larger values of compressive stress cause the transformation to form only horizontal twin bands (E_1/E_3 twins, such as those in Figure 7). Test T12 involved cooling the specimen under -12 MPa of stress to 11°C and then reducing the applied stress to -1 MPa before the magnetic field path Rot20-5515 was performed. Test T16 was cooled under -1 MPa of stress and then the same field path was performed. Figure 8 compares the strain responses during the first cycle of each of these tests. The results from T12 are at consistently larger (compressive) strain values, which is as expected due to the microstructural bias

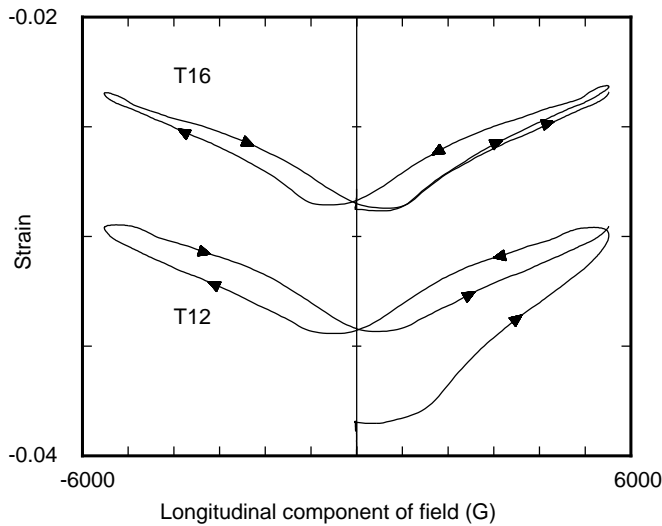


Figure 8. Strain versus field from tests T12 and T16 at -1 MPa and 11°C for specimen MM6 are shown. Test T12 included a bias stress of -12 MPa during the phase transformation, while test T16 had a bias stress of -1 MPa during cooling.

applied by the -12 MPa under which T12 cooled. A large fraction of this extra compressive strain is recovered on the application of the longitudinal field, which favors lengthening of the specimen, however, this recovery does not reverse on the removal of the field and the amplitude of the cyclic strain change is very similar in both tests. Ignoring the first 90° of field change when the irreversible recovery occurs, the average strain amplitude over the rest of the data shown in the figure is 0.0049 in test T12, and 0.0057 in test T16. It should be noted that the specimen is not fully detwinned by the -12 MPa stress, a small number of thin twin bands are still visible on the specimen surface after cooling is complete. When the stress is reduced to -1 MPa the bands become thicker and more numerous.

To study the effects of the field path, the paths Lin-4242 and Arc-4242 were designed. The results from two of the tests using these paths are shown in Figure 9. The strain versus longitudinal field component behavior is almost identical for the two tests. The large difference in the strain versus transverse field component plots is due to the fact that the arc path also has a longitudinal component to the field that grows as the transverse component is reduced. The solid circle and square symbols mark corresponding points on the arc path curves in both plots. This growing longitudinal component causes the specimen to lengthen on the arc path while the linear path causes no lengthening with its purely transverse field. The numerical values of the strain amplitudes for tests T2, T3, and T4 are given in Table 6. These data show that there is essentially no difference in the amplitudes measured for these two different field paths at the two different stress levels considered.

5.3. Comparison with constrained theory. Test T5 with a compressive stress of -1 MPa was selected for comparison with the results of the micromagnetic theory. As shown in Figure 10, there is a large discrepancy in the magnitude of the ferromagnetic strain predicted by theory and the experimental measurements. The theory predicts a ferromagnetic strain amplitude of 0.0515 compared to a measured

Test	Stress (MPa)	Ferromagnetic Strain	
		Lin-4242	Arc-4242
T2	-2	0.0076	0.0074
T2	-1	0.0098	0.0090
T3	-2	0.0074	0.0068
T3	-1	0.0099	0.0090
T4	-2	0.0078	0.0068
T4	-1	0.0095	0.0085

Table 6. Ferromagnetic strain amplitudes measured in the tests T2, T3, and T4 at 11°C on specimen MM6.

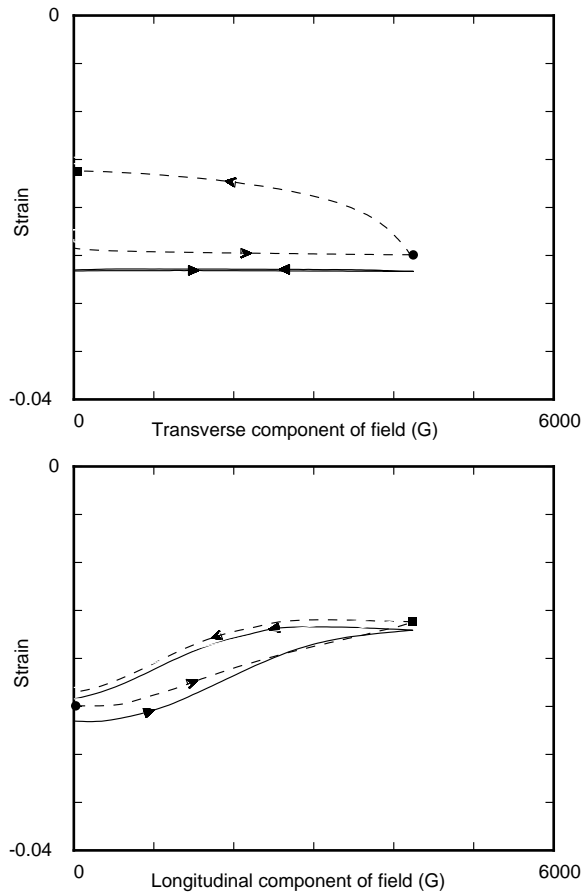


Figure 9. Strain versus field of test T3 at -1 MPa for specimen MM6. Dashed lines represent the strain response of the specimen when the field path is the arc rotation, and solid lines represent those with linear path. The dot and square indicate corresponding points on the two curves from the arc path.

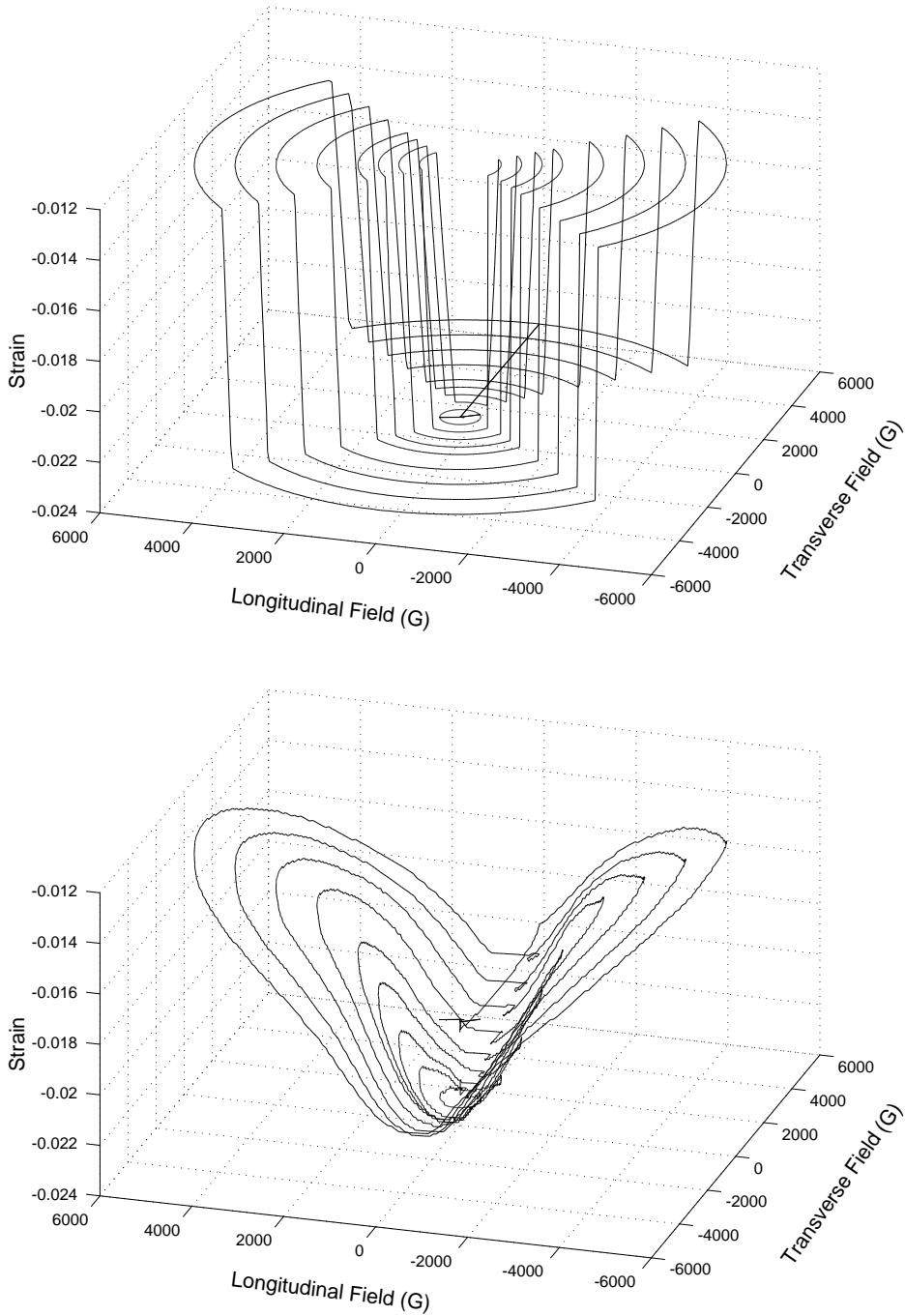


Figure 10. Comparison of the theoretical (top) and the experimental (bottom) results for field-induced strain of specimen MM6 at -1 MPa and 10°C . The total field-induced strain in top figure is scaled to $1/5$ of the theoretical prediction.

value of 0.007. The basic shapes of the two sets of curves are similar and the constrained theory cannot be expected to capture smooth transitions between states that must occur due to the presence of microstructure and demagnetization. Because the theory assumed an infinite specimen, the finite nonellipsoidal shape of the specimen and its effects on the M-H curves were ignored. These effects give the M-H curves its gradual transition to saturation and will similarly affect the strain-field behavior of the specimen. Nonetheless, the curves do show a relative flat minimum when the magnitude of the longitudinal field is less than 440 G, and a tendency to flatten at higher fields. The sign of the strain changes agrees with the theory and this confirms that the long axis of the FCT martensite is the easy axis, otherwise a longitudinal field would make the specimen shorten. In addition, deviations from constrained theory are quite possible due to the finite values of the elastic moduli and magnetic anisotropy of the material. The magnetic anisotropy of $\text{Fe}_{70}\text{Pd}_{30}$ was found to be an order of magnitude smaller than that of Ni_2MnGa [Cui et al. 2004]. This small anisotropy means that it may be energetically less costly for the magnetization to rotate away from the easy directions in the material compared to moving a twin boundary that would result in a change in strain. This may be the reason for the small amount of ferromagnetic strain measured in this material.

6. Conclusions

The results presented above confirm that $\text{Fe}_{70}\text{Pd}_{30}$ is a FSM material and that its behavior agrees qualitatively with the predictions of the constrained theory for FSM materials. However, there are two results that are unexplained. The first is the small amount of ferromagnetic strain amplitude measured in this material. As just discussed in the previous paragraph, this may be due to the small magnetic anisotropy of this material. However, there are also a couple of unusual aspects to the shape-memory behavior of this material that may also contribute to the small strain amplitude. Observations of the microstructure that forms when this material is cooled through the transformation temperature are quite different from what is typically observed in shape memory materials (compared say to CuAlNi). Instead of a sharp planar austenite-martensite (A-M) interface with finely twinned martensite, the martensite in $\text{Fe}_{70}\text{Pd}_{30}$ grows into the austenite in a disorderly fashion. The predominate microstructure with a small volume fraction of martensite is a horizontal needle that appears to consist of a pair of martensite variants with a twin boundary down its middle. These needles grow into the austenite and then merge to give the final twinned martensite structure when the transformation is complete. Often the tips of these needles are arranged along what appears to be a roughly 45° line (the crystallographic theory of martensite predicts an A-M interface at nearly 45° with horizontal martensite twin interfaces) and they sometimes appear to move in unison, but this is far from the distinct A-M interface typically observed in shape memory materials. This type of behavior may indicate the $\text{Fe}_{70}\text{Pd}_{30}$ deviates substantially from the constrained theory. Because this transformation behavior is solely shape-memory related, this would indicate that the deviation is related to the elastic behavior of the material. Thus, elastic deformations may play an important role in the microstructure formation. Without accurate predictions of the microstructures in FSM materials, it will be very difficult to predict their behavior.

The second unusual aspect of the behavior of $\text{Fe}_{70}\text{Pd}_{30}$ involves its behavior in the martensitic state. When cooled under a stress of -12 MPa there are still some horizontal twin bands visible on the specimen surface when cooling is complete. Typical shape-memory materials would be completely detwinned by

this stress once the transformation was complete. Even more interesting, is that when the stress is reduced to -1 MPa from the -12 MPa applied during cooling (see the results for test T12) some twin boundary motion occurs. The twin bands thicken and the specimen lengthens. It is not known what force could be driving twin boundary motion that would grow the variants with their a axes in the compression direction at the expense of the variant with its shorter c axis in this direction when the amount of compressive stress is reduced. A typical shape memory material would not retwin until the stress became tensile. If there is some unknown microstructure related driving force in this material, the applied magnetic field would have to overcome this force to cause the FSM effect thus reducing the amount of strain achieved. Because there are only three tetragonal variants, it may be that there is a fair amount of elastic strain near the ends of specimen because of the lack of enough microstructural degrees of freedom to form a stress free microstructure there. These elastic strains could provide the energy needed for this type of behavior.

The second unexplained aspect of the behavior of $\text{Fe}_{70}\text{Pd}_{30}$ is the apparent lack of a blocking stress. Figure 4 shows that while the strain amplitude decays with increasing applied compressive stress, it is not decaying fast enough for the work output to decay as well. The results at -12 MPa are particularly notable. Aside from this data point, the work output does appear to have reached a maximum at approximately -4 MPa and is beginning to decrease. Because the strains measured at -12 MPa are the smallest, the relative error in this measurement will be the largest. If the strain was only 25% less than the reported value, the decreasing trend in the work output would continue to this stress level as well. The results reported in Figure 4 are from specimen MM10 using the capacitive sensor to measure the displacements. A constant bias or calibration error in this measurement would not effect the trends in the data. Additionally, the measurements are made under a constant stress so any elastic displacements of the compression fixture should cancel out in the amplitude calculation. Strain gauges are difficult to use on these materials, because they tend to debond during the phase transformation due to the large strains involved.

Future work on this material will focus on answering these questions. Measurement of the average magnetization of the material during testing will allow the possibility of magnetization rotation away from the easy axes to be assessed. In addition the shape-memory behavior will be studied further to provide a better understanding of the types of behaviors that have just been discussed. Finally, improvements to the constrained theory, such as including the specimen geometry, and the possibility of including deviations from constrained theory will be considered.

References

- [Clark 1980] A. Clark, *Ferromagnetic materials*, North-Holland, Amsterdam, 1980.
- [Cui and James 2001] J. Cui and R. D. James, "Study of Fe_3Pd and related alloys for ferromagnetic shape memory", *IEEE Trans. Magn.* **37**:4 (Part 1) (2001), 2675–2677.
- [Cui et al. 2004] J. Cui, T. W. Shield, and R. D. James, "Phase transformation and magnetic anisotropy of an iron-palladium ferromagnetic shape-memory alloy", *Acta Mater.* **52**:1 (2004), 35–47.
- [Cui et al. \geq 2008] J. Cui, T. W. Shield, and R. D. James, "Effects of stress and magnetic field on the phase transformation in iron-palladium ferromagnetic shape memory alloys". In preparation.
- [DeSimone and James 2002] A. DeSimone and R. D. James, "A constrained theory of magnetoelasticity", *J. Mech. Phys. Solids* **50**:2 (2002), 283–320.
- [James and Kinderlehrer 1993] R. D. James and D. Kinderlehrer, "Theory of magnetostriction with applications to $\text{Tb}_x\text{Dy}_{1-x}\text{Fe}_2$ ", *Philos. Mag. B* **68** (1993), 237–274.

- [James and Wuttig 1998] R. D. James and M. Wuttig, "Magnetostriction of martensite", *Philos. Mag. A* **77**:5 (1998), 1273–1299.
- [Muto et al. 1988] S. Muto, S. Takeda, R. Oshima, and F. E. Fujita, "High resolution electron microscopy of the tweed microstructure in an Iron-Palladium alloy", *Jpn. J. Appl. Phys.* **27**:8 (1988), 1387–1389.
- [Muto et al. 1990a] S. Muto, R. Oshima, and F. E. Fujita, "Consideration of the tweed structure of Fe-Pd alloys by continuum elasticity theory", pp. 65–70 in *Martensitic transformations: proceedings of the 6th International Conference* (Sydney, 1989), vol. 1, edited by B. C. Muddle, Materials Science Forum **56**, Trans Tech Publications, Aedermannsdorf, Switzerland, and Brookfield, VT, 1990.
- [Muto et al. 1990b] S. Muto, S. Takeda, and R. Oshima, "Analysis of lattice modulations in the tweed structure of an Fe-Pd alloy by image processing of a high resolution electron micrograph", *Jpn. J. Appl. Phys.* **1** **29**:10 (1990), 2066–2071.
- [Muto et al. 1990c] S. Muto, S. Takeda, R. Oshima, and F. E. Fujita, "High resolution electron microscopy of tweed microstructure in Fe-Pd alloys", pp. 45–50 in *Martensitic transformations: proceedings of the 6th International Conference* (Sydney, 1989), vol. 1, edited by B. C. Muddle, Materials Science Forum **56**, Trans Tech Publications, Aedermannsdorf, Switzerland, and Brookfield, VT, 1990.
- [O'Handley 1998] R. C. O'Handley, "Model for strain and magnetization in magnetic shape-memory alloys", *J. Appl. Phys.* **83**:6 (1998), 3263–3270.
- [Oshima 1981] R. Oshima, "Successive martensitic transformations in Fe-Pd alloys", *Scr. Metall.* **15**:8 (1981), 829–833.
- [Oshima and Sugiyama 1982] R. Oshima and M. Sugiyama, "Martensite transformations in iron-palladium alloys", *J. Phys. (Paris) Colloq.* **43**:C4 (1982), 383.
- [Oshima et al. 1988] R. Oshima, M. Sugiyama, and F. E. Fujita, "Tweed structures associated with FCC-FCT transformations in Fe-Pd alloys", *Metall. Trans. A* **19**:4 (1988), 803–810.
- [Oshima et al. 1990] R. Oshima, K. Tanaka, A. Taniyama, and F. E. Fujita, "Study of BCT martensite of Fe-Pd alloys", pp. 175–180 in *Martensitic transformations: proceedings of the 6th International Conference* (Sydney, 1989), vol. 1, edited by B. C. Muddle, Materials Science Forum **56**, Trans Tech Publications, Aedermannsdorf, Switzerland, and Brookfield, VT, 1990.
- [Oshima et al. 1992] R. Oshima, S. Muto, and F. E. Fujita, "Initiation of FCC-FCT thermoelastic martensite transformation from premartensitic state of Fe-30 at Pd alloys", *Mater. Trans. (JIM)* **33**:3 (1992), 197–202.
- [Shield 2003] T. W. Shield, "A magnetomechanical testing machine for ferromagnetic shape-memory alloys", *Rev. Sci. Instrum.* **74**:9 (2003), 4077–4088.
- [Somura and Oshima and Fujita 1980] T. Sohmura, R. Oshima, and F. E. Fujita, "Thermoelastic FCC-FCT martensitic transformation in Fe-Pd alloy", *Scr. Metall.* **14**:8 (1980), 855–856.
- [Sozinov et al. 2002] A. Sozinov, A. A. Likhachev, and K. Ullakko, "Crystal structures and magnetic anisotropy properties of Ni-Mn-Ga martensitic phases with giant magnetic-field-induced strain", *IEEE Trans. Magn.* **38**:5 (Part 1) (2002), 2814–2816.
- [Sugiyama et al. 1984] M. Sugiyama, R. Oshima, and F. E. Fujita, "Martensitic transformation in the Fe-Pd alloy system", *Trans. Jpn. Inst. Met.* **25**:9 (1984), 585–592.
- [Sugiyama et al. 1985] M. Sugiyama, S. Harada, and R. Oshima, "Change in Young's modulus of thermoelastic martensite Fe-Pd alloys", *Scr. Metall.* **19**:3 (1985), 315–317.
- [Sugiyama et al. 1986] M. Sugiyama, R. Oshima, and F. E. Fujita, "Mechanism of FCC-FCT thermoelastic martensite transformation in Fe-Pd alloys", *Trans. Jpn. Inst. Met.* **27**:10 (1986), 719–730.
- [Tanaka and Oshima 1991] K. Tanaka and R. Oshima, "Role of annealing twin in the formation of variant structure of Bct martensite in Fe-Pd alloy", *Mater. Trans. (JIM)* **32**:4 (1991), 325–330.
- [Tanaka et al. 1992] K. Tanaka, K. Hiraga, and R. Oshima, "Origin of tetragonality of BCT martensite in substitutional Fe-Pd(-Ni) disordered alloys", *Mater. Trans. (JIM)* **33**:3 (1992), 215–219.
- [Tickle and James 1999] R. Tickle and R. D. James, "Magnetic and magnetomechanical properties of Ni₂MnGa", *J. Magn. Mater.* **195**:3 (1999), 627–638.
- [Tickle et al. 1999] R. Tickle, R. D. James, T. W. Shield, M. Wuttig, and V. V. Kokorin, "Ferromagnetic shape memory in the NiMnGa system", *IEEE Trans. Magn.* **35**:5 (Part 3) (1999), 4301–4310.

[Ullakko et al. 1996] K. Ullakko, J. K. Huang, C. Kantner, R. C. O'Handley, and V. V. Kokorin, "Large magnetic-field-induced strains in Ni₂MnGa single crystals", *Appl. Phys. Lett.* **69**:13 (1996), 1966–1968.

[Vasil'ev et al. 1996] A. N. Vasil'ev, S. A. Klestov, R. Z. Levitin, and V. V. Snegirev, "Magnetoelastic interaction in the martensitic transformation in an Ni₂MnGa single crystal", *J. Exp. Theor. Phys.* **82**:3 (1996), 524–526.

Received 29 Dec 2005. Accepted 18 Sep 2006.

JUN CUI: cui@research.ge.com

GE Global Research Center, K1-2D47, One Research Circle, Niskayuna, NY 12302, United States

TOM SHIELD: shield@aem.umn.edu

Department of Aerospace Engineering and Mechanics, University of Minnesota, 107 Akerman Hall, 110 Union St. SE, Minneapolis, MN 55455, United States

MACROSCOPIC ELASTIC PROPERTIES OF RANDOMLY PACKED BALLOONS

ISAO TAGUCHI AND MICHIO KURASHIGE

Macroscopic elastic properties of sintered, randomly packed balloons are estimated for various degrees of sintering and for a wide range of balloon wall thickness. Macroscopic elastic moduli are little affected by a balloon's inner/outer diameter ratio for thicker balloons, while those of thinner balloons are very sensitive to the ratio. The elastic moduli rapidly decline with decreasing wall thickness. They are larger in the direction of gravity than in the horizontal one, corresponding to gravity-affected packing structures. Specific elastic moduli have a peak against porosity. Poisson's ratios are negative in some cases of very thin balloon walls and very low sintering degree. Comparison of the present random packing with the regular simple cubic packing reveals that there is a significant difference between them in their elastic properties and their structures of anisotropy, although both have almost the same coordination number.

1. Introduction

Cellular solids have many diverse applications such as hot gas and ion exchange filters, thermal protection systems, energy and sound absorption systems, heat exchangers, catalyst supports and porous implants for surgical treatments [Gibson and Ashby 1997; Ashby et al. 2000]. They may be regarded as porous materials with a very high porosity or very low solid volume fraction [Torquato 2001]. Among the cellular solids, foams are most commonly and widely used. The foams are usually classified by their pore type of either open or closed cells.

The open-cell foams are neither as stiff nor as strong as the closed-cell foams, and they allow fluids to flow through whereas the closed-cell foams do not. Because of this, they can be exploited in multi-functional applications of load supporting and heat dissipation [Wadley 2002; Queheillalt et al. 2002].

Open-cell metallic foams typically achieve mechanical behaviors close to theoretical predictions. Although the mechanical properties of closed-cell metallic foams theoretically exceed those of open-cell foams, defects reduce their measured properties to values similar to those for the open-cell foams [Andrews et al. 1999; Sanders and Gibson 2003b]. Bonded compacts of balloons may be a good alternative in terms of mechanical properties as well as in their ability to allow fluids to flow through connected pores.

In recent years, bonded compacts of balloons (or hollow spheres) made of ceramic, metal, inorganic glass, carbon, etc. have started to become commercially available and have been applied to various engineering practices [Torobin 1986; Norris and Gojny 1990; Andersen et al. 2000]. Because the materials made by bonding balloons contain both types of connected and closed pores, these structures have significant potential for those multifunctional applications which require a combination of impact energy absorption, acoustic attenuation [Gasser et al. 2004a] and/or thermal insulation, in addition to

Keywords: balloon, elastic modulus, random packing, transverse isotropy, negative Poisson's ratio, optimum design.

structural load support; the bonded balloons also have potential for improved mechanical properties and high-porosity structures, that is, for high specific stiffness and strength.

Another important usage of the microballoons is as syntactic foams, which are mixtures of microballoons and polymeric (or metallic) matrix material [Bardella and Genna 2001; Marur 2005]. Because the microballoons are dispersed in the matrix material, the syntactic foams are closed-cell foams and do not allow fluids to flow through connected pores. We will exclude the syntactic forms from the present argument.

Using the finite element method, Sanders and Gibson [2003b] analyzed the mechanical properties of simple cubic (SC) packed hollow-sphere foams, and compared them with those for the open-cell and closed-cell foams. The results indicated that the theoretical values of moduli and strength stand between those of the open- and closed-cell foams. Furthermore, they also obtained those of body-centered cubic (BCC) and face-centered cubic (FCC) packings [Sanders and Gibson 2003a]. It was found that the FCC packing gives the highest values of moduli and strength.

Like Sanders and Gibson [2003a], Gasser et al. [2003] investigated the uniaxial tensile elastic properties of regular FCC hollow sphere packings, and expressed three independent elastic constants of the materials for FCC stacking in terms of polynomial expressions [Gasser et al. 2004b]. They compared the results with the estimations from the formulae proposed by Sanders and Gibson [2003a]; it was shown that the polynomial expressions are valid for the case where the size of necks is much smaller than 0.2 times the balloon radius.

All of the above studies have dealt with regular stackings. However, bonded balloon aggregates are usually fabricated by dumping balloons into a container and in some cases by tapping it to obtain denser packings, leading to random structures. Thus, we need to take into account the randomness of balloon structures. It should be noted that the randomness hinders us from using the finite element analysis of such unit cells as those in SC, BCC and FCC stackings; we have to resort to another method.

For randomly packed solid spherical particles, Kurashige et al. [1999a] proposed a three-step simulation method. This method constructs a random packing of spheres in a computer by sequential deposition of them, followed by replacing the packing, after *sintering* it, by a three-dimensional random network of springs of six degrees of freedom, and then by estimating effective elasticity coefficients by conventional structural analysis. The results by this simulation were in good agreement with experimental ones. Moreover, Kato et al. [2002] extracted a statistical characteristic of the random packing structure from data on the random packing; although this kind of statistic is important to understand some relations between the packing structures and their macroscopic characteristics, it cannot be obtained in other models, such as the self-consistent models. They also estimated the mechanical properties (effective elastic moduli and elastic wave speeds) by the same evaluation method as that in Kurashige et al. [1999a]. From this evaluation, it was found that the packing structure of spherical particle random packings and the elastic property of sintered compacts are of transverse anisotropy, due to gravity. Furthermore, Kurashige et al. [1999b] applied the method to the thermal conductivity estimation of sintered solid particles; they confirmed the usefulness of their method.

The present paper treats elastic properties of the bonded balloon random packing. We deal here with the *sintered* balloons, which imply that the bonding material is identical to that of the balloons, but the present method is applicable to blazed balloon random packings, or those bonded by other methods. We carry out a thorough simulation using the above method [Kurashige et al. 1999a; Kato et al. 2002] to

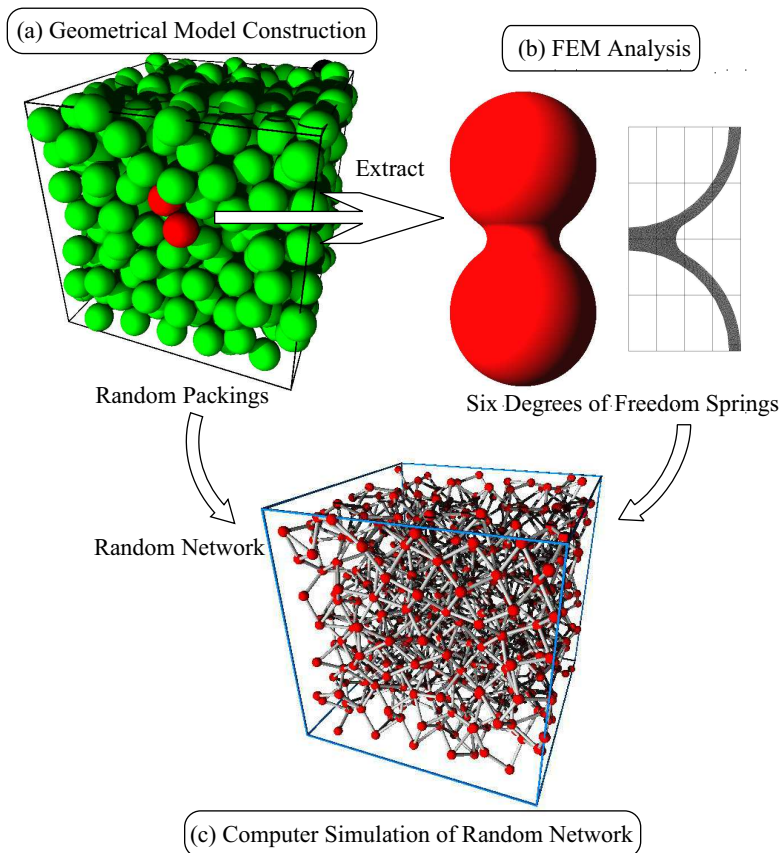


Figure 1. Three step simulation method: (a) geometrical model construction of randomly packed balloon aggregate by the method of a rigid sphere free fall into a virtual box; (b) evaluation of the characteristics of a microscopic structure by FEM analysis of spring constants for a sintered balloon pair; (c) simulation of macroscopic material properties by 3D structural analysis of a random network of springs, yielding desired macroscopic elastic moduli.

examine statistical characteristics of the balloon packings and their macroscopic elastic properties and to propose simple formulae to estimate the elasticities, by going through the following steps. (We will use the term *balloon packings*, although there is no difference in packing structures and their statistical characteristics between solid and hollow spheres.)

First, to generate a structural model of sintered balloon aggregates, we carry out a random packing simulation of the sequential accumulation method of equal-sized balloons into a virtual box (see Figure 1a) combined with the use of periodic boundary conditions and random ups and downs of the box floor

level. Some statistical characteristics of aggregate microstructures are examined by using such parameters as porosity, coordination number, cumulative frequency of diameters of the circles appearing on a cross-section created by cutting the aggregate and cumulative angular distribution of all contact points within the aggregate.

Next, we extract a pair of connected spheres from the packing and regard it as a spring with six degrees of freedom: one for elongation and torsion, and two for bending and shearing. Properties of the microstructure element (that is, spring constants) are estimated by the finite element method; see Figure 1b.

Finally, we carry out tensile and shear tests of a random network of springs, with rigid hinges, of six degrees of freedom with the spring constants already estimated by the finite element method (Figure 1c). From this simulation, we obtain the average tensile (or shear) stress and the average strain resulting from the applied load. Thus, the desired macroscopic Young's and shear moduli and Poisson's ratios are determined.

We also calculate Young's modulus for the SC packing using the present three-step method to confirm its validity. To consider an optimum design, macroscopic specific (that is, per weight) elastic moduli are precisely estimated for a wide range of balloon thickness and sintering degree. Furthermore, we compare the results for the present random packing with those for the regular packing obtained by Sanders and Gibson [2003b].

2. Geometrical model construction

2.1. Equation of motion. To generate a spherical particle random packing structure in a computer, we carry out the simulation by the method of rigid sphere free fall into a box. More details can be found in [Taguchi et al. 2006].

We sequentially release a spherical particle with diameter D from random positions into a virtual box in the computer. The motion equation of the sphere in a vectorial form can be expressed as:

$$c \frac{d\mathbf{x}}{dt} = k\Delta\mathbf{d} - mg\mathbf{k}, \quad (1)$$

where \mathbf{x} is the position vector of a falling sphere's center, \mathbf{k} is the unit vector along the z axis pointing upward and $\Delta\mathbf{d}$ is the overlapping depth vector of the sphere in collision with another still one. Constants c , k , m and g are the viscosity, the spring constant between two or more particles in collision, mass of each particle, and the gravitational acceleration, respectively. We solve equation Equation (1) by the Runge–Kutta–Gill method.

2.2. Simulation method. Consider the virtual box of $0 \leq x \leq B$, $0 \leq y \leq B$, $0 \leq z \leq H$ ($B < H$). We do not stop introducing balloons until the cubic region $0 \leq x \leq B$, $0 \leq y \leq B$, $0 \leq z \leq H$ is completely filled with balloons. For the box size $B/D = 16$ adopted, 5 runs were made with 5 different sets of pseudo-random numbers, giving 5 samples. Five lists of final coordinates of all balloons' centers will be used to analyze the statistical characteristics and to estimate the elastic moduli of sintered balloons later on.

The effects of the boundaries of the virtual box in which spheres accumulate are eliminated by introducing the cyclic boundary condition. The level of the box floor is moved up and down for each sphere

	$B/D = 16$
Sample 1	0.417439
Sample 2	0.417035
Sample 3	0.417106
Sample 4	0.416517
Sample 5	0.417406
Mean	0.417101
Standard deviation	0.000372

Table 1. Porosities of five samples; their mean and standard deviation.

by a magnitude given by pseudo-random numbers in order to avoid packing regularity on the floor. The algorithm by Wichmann and Hill [1982] is used for the generation of the random numbers.

3. Statistical characteristics of geometrical model

3.1. Porosity. Porosity associated only with interstitial void space between the hollow spheres is one of the most important parameters characterizing a random packing of particles. It should be cautioned that the porosity excludes hollow space in the hollow spheres. The porosity of the random packing constructed by the above method is evaluated from the center coordinate lists and shown in Table 1. As seen from this table, the mean porosity is about 41.7%. This mean value is in good agreement with that obtained by Kato et al. [2002].

In general, random packings are classified into three categories [Haughey and Beveridge 1969; Tory et al. 1973]: very loose random packing ($0.44 \leq \phi \leq 0.47$), random loose packing ($0.41 \leq \phi \leq 0.44$), and close random packing ($\phi \approx 0.36$).

The mean porosity shown in Table 1 reveals that our packings fall in the category of random loose packing. The porosity $\phi = 0.417$ obtained here is rather small in the range of the loose random packing, because the present method allows no bridging.

3.2. Coordination numbers. Next, we examine the coordination number for our random packings for the box size $B/D = 16$. The coordination number is defined as the number of contact points of a particle with neighboring ones in a regular or irregular packing of particles. With a few regular packings, it is 6 for the SC, 8 for BCC and 12 for FCC; in contrast, there is a possibility that a sphere is in contact with four to twelve other spheres for the random packing [Nolan and Kavanagh 1992].

Table 2 shows the mean coordination number for all particles for each sample and the mean value and standard deviation for the five samples. The mean value 6.019 is in good agreement with that obtained by Kato et al. [2002]. It is found that the average coordination number for the random packing constructed is very near to that for the regular SC packings.

Figure 2 shows the distribution of the coordination numbers for the five samples. From this figure, it can be seen that the coordination distribution has its peak at 6 and ranges from 4 — 8 for all samples;

	Coordination Number
Sample 1	6.016345
Sample 2	6.020791
Sample 3	6.020047
Sample 4	6.019235
Sample 5	6.020346
Mean	6.018755
Standard deviation	0.001921

Table 2. Coordination number for five samples; their mean and standard deviation.

it has negligibly small percentage for three and nine contact points and no sphere has more than then contact points. Little difference in the distribution can be seen among the samples.

3.3. Cumulative frequency of diameters of circles appearing on a cross-section. If a randomly packed aggregate of balloons is sectioned along a plane, a large number of double circles will appear on its resultant cross-section; the outer circles have various values of diameter D' , which is naturally smaller than or equal to D . Cumulative distribution of the diameter $f(D'/D)$ is theoretically given by

$$f(D'/D) = 1 - \sqrt{1 - (D'/D)^2}, \quad (2)$$

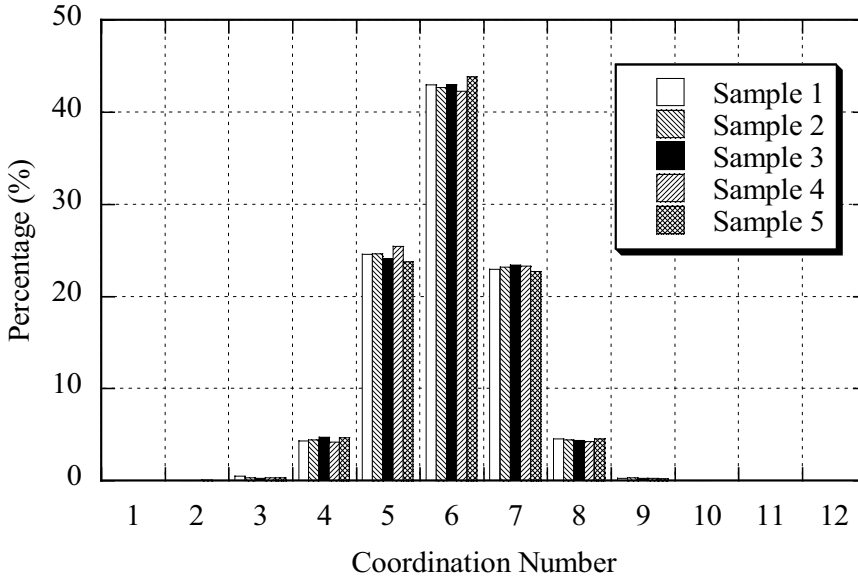


Figure 2. Coordination number distribution for all five samples.

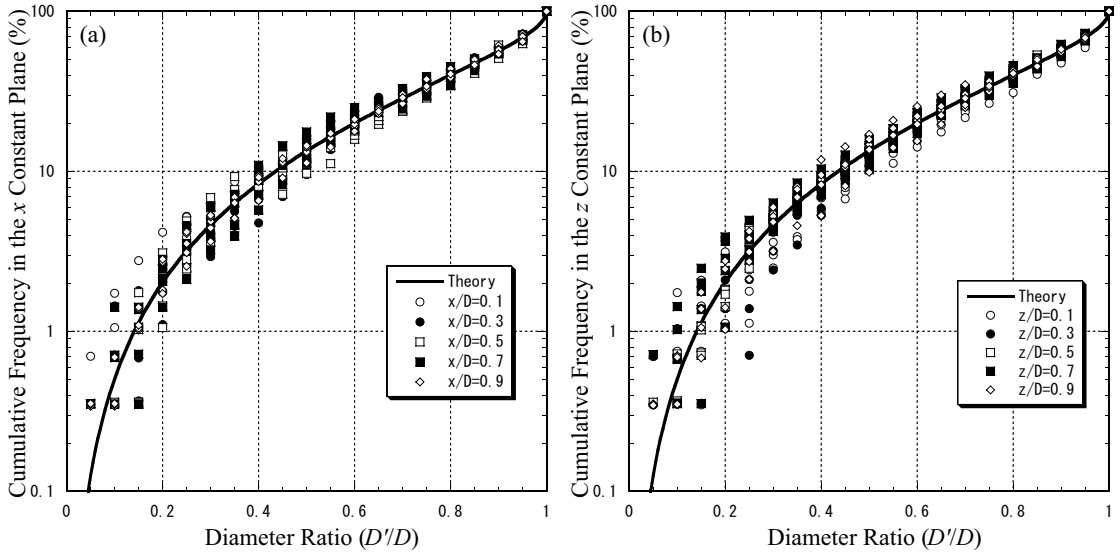


Figure 3. Cumulative frequencies of diameters of circles appearing on x and z cross sections created by cutting aggregate, corresponding to (a) and (b), respectively.

provided that the centers of packed balloons are distributed in a uniformly random manner [Debbas and Rumpf 1966; Bennett 1972].

Figures 3a and 3b show the cumulative frequencies of diameters of simulated random packings for all samples in terms of percentage for the x , $z = \text{constant}$ cross sections, respectively. (The figure for $y = \text{constant}$ is omitted because it is similar to that for $x = \text{constant}$). Here, the solid line is the theoretical prediction given by Equation (2). Each frequency is for each of five sectional planes x/D or $z/D = 0.1, 0.3, 0.5, 0.7$ and 0.9 . It is seen that the frequencies scatter more around the theoretical value for the smaller diameter ratio. The scattering in [Kato et al. 2002] is smaller than the present one, because there the frequencies were calculated by averaging over the five sectional planes. The cumulative frequency for the $z = \text{constant}$ planes is not different from that for the $x = \text{constant}$ plane. This implies that the centers of spheres are almost uniformly randomly distributed equally along the three mutually orthogonal directions, although the scattering on the $z = \text{constant}$ planes seems somewhat larger.

3.4. Distributions of branch orientations. A *branch* is defined as a segment connecting the centers of balloons in contact; the spheres are judged to be in contact if the length of branch is equal to or smaller than the diameter of sphere D , and not in contact if larger.

Although the sphere centers appear to be uniformly distributed as shown in the above subsection, it may be important to examine orientations of all branches.

Kato et al. [2002] examined distributions, not cumulative ones, of the branch orientations in the similar random packings. They divided the domain of zenithal and azimuthal angles into sub-domains by 10° step and calculated both the angular frequencies for each step, which were depicted in the form of a bar chart. They concluded that the peak in the zenithal distributions appears around 45° from the vertical

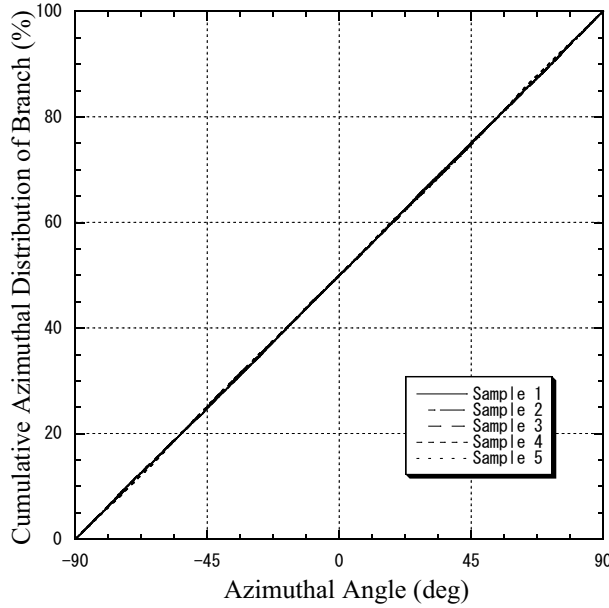


Figure 4. Cumulative azimuthal distributions of branch orientations.

line, while the distributions are uniform in the horizontal plane. In the present paper, we use both the cumulative azimuthal and zenithal distributions of branch orientations in more detail and better accuracy.

Before calculating both of the angular distributions, let us introduce the spherical polar coordinate (r, θ, φ) , where θ and φ are the zenithal and azimuthal angles, respectively. The zenithal angle is measured from the z -(vertical) axis, while the azimuthal angle is from the x -axis in the $x - y$ plane. The zenithal and azimuthal angles of the k -th branch ($k = 1, 2, \dots, N$, with N being the total number of branches) are denoted by θ_k and φ_k , respectively.

First, to examine the azimuthal distribution of branch orientations, consider the domain of

$$\varphi, \quad (-\pi/2, \pi/2].$$

If φ_k is in $(\pi/2, \pi]$, φ_k should be replaced by $\varphi_k - \pi$; similarly, that in $(-\pi, \pi/2]$ by $\varphi_k + \pi$. Furthermore, we rearrange the branches in ascending order of azimuthal angles in the range of $(-\pi/2, \pi/2]$. We define the cumulative frequency of the branch orientations in the following way:

$$F(\varphi) = \left(\frac{1}{\sum_{k=1}^N \sin \theta_k} \sum_{k=1}^N \sin \theta_k H(\varphi - \varphi_k) \right) \times 100(\%), \tag{3}$$

where $H(\varphi)$ is the Heaviside step function. $F(\varphi)$ with $\sin \theta_k$ as weight shows the cumulative percentage of the number of branches.

The cumulative azimuthal distribution defined above is calculated for all five samples. The results are shown in Figure 4. It can be seen that the cumulative distribution is almost exactly a straight line with gradient of 100/180. No sample seems to have any distinct characteristics. Thus, we conclude that the

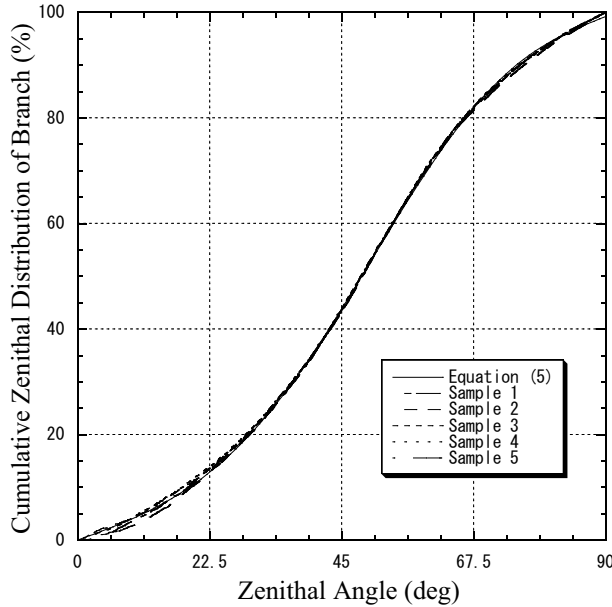


Figure 5. Cumulative zenithal distributions of branch orientations and their fitting curve given by Equation (5).

structure of the obtained random packings is isotropic in the horizontal $x - y$ plane in terms of the branch orientations.

Next, we examine the cumulative zenithal distribution of branch orientations. Similarly to the cumulative azimuthal distribution, θ_k should be replaced by $\theta_k - \pi/2$ if it is in $(\pi/2, \pi]$. We rearrange the branches in ascending order of azimuthal angles in the range of $(0, \pi/2]$. Since the circle made by $\theta = \theta_k$ on a sphere of unite radius is $2\pi \sin \theta_k$ in circumference, the cumulative frequency of branches per unit area of the sphere $F(\theta)$ is defined by

$$F(\theta) = \left(\frac{1}{\sum_{k=1}^N \frac{1}{\sin \theta_k}} \sum_{k=1}^N \frac{H(\theta - \theta_k)}{\sin \theta_k} \right) \times 100(\%). \tag{4}$$

The cumulative frequency is calculated and shown in Figure 5 for all samples. From this figure, it can be seen that the cumulative frequency distribution is not linear at all. Since the frequency distribution of branches is an odd function both about $\theta = 0^\circ$ and $\theta = 90^\circ$, we can express the distribution by the Fourier series of $\sin 2m\theta$. If we take into account only the first four terms, the distribution is

$$F(\theta) = a_0\theta + \sum_{m=1}^3 \frac{a_m}{2m} \sin 2m\theta, \tag{5}$$

$$a_0 = 63.043, \quad a_1 = -7.4308, \quad a_2 = -38.522, \quad a_3 = 7.5564,$$

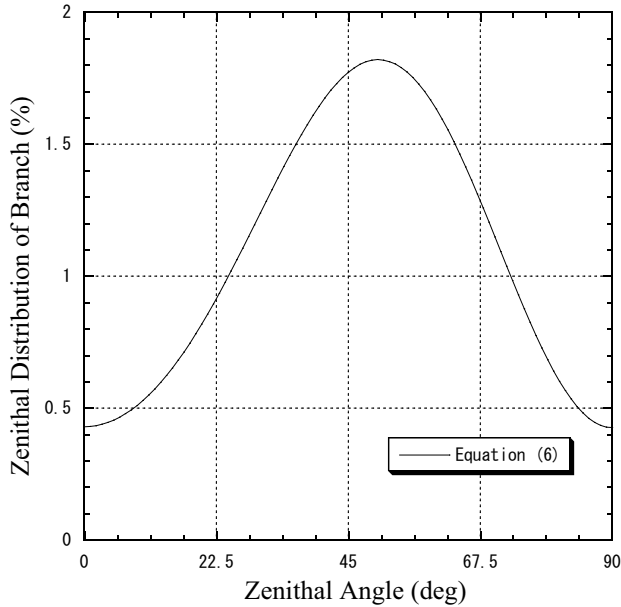


Figure 6. Zenithal distributions of branch orientations obtained from fitting curve, given by Equation (6).

where the four coefficients a_m have been determined by the least square method. The fitting curve of Equation (5) is drawn by a solid line in Figure 5; it cannot be distinguished from the original five curves, which reveals that the fitting is complete.

Differentiating Equation (5) with respect to θ , we obtain

$$\frac{dF(\theta)}{d\theta} = \sum_{m=0}^3 a_m \cos 2m\theta. \quad (6)$$

This equation shows the zenithal distribution of branches itself and its curve is drawn in Figure 6. It is seen that the maximum frequency exists between 50° and 60° while the minimal is at $\theta = 0^\circ$. By setting the derivative of Equation (6) to zero, $d^2F(\theta)/d\theta^2 = 0$, it is found that the frequency distribution has its peak at $\theta = 50.1^\circ$. This peak position is correct because the present evaluation is more precise than that given by Kato et al. [2002]; their peak position $\theta = 45^\circ$ was too rough.

From the arguments so far in this subsection, we can conclude that the packing structures are transversely isotropic, indicating the effect of gravity. This conclusion comes from the viewpoint of branch orientations.

4. Evaluation of spring constants of a sintered pair of balloons

4.1. Model of finite element. In the sintering of the balloon aggregate, necks are created around contact points of the balloons by mass diffusion; the necks grow up with sintering time. However, we do not simulate it, because we do not need to know its whole process, but need only to obtain the geometry of

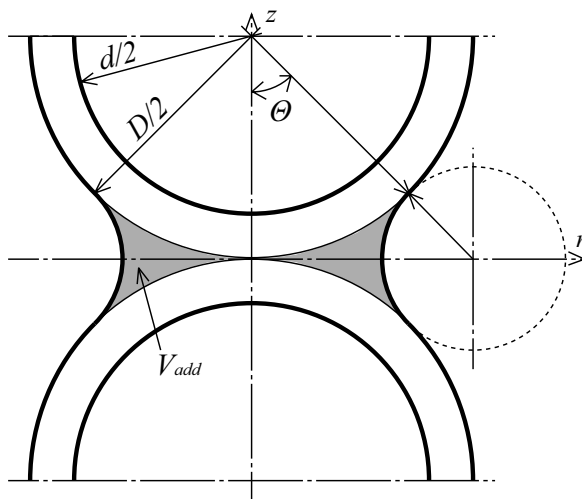


Figure 7. Geometry of a sintered (mass-added) balloon.

sintered state of aggregates. In order to represent the sintered geometry around the point of contact of paired balloons, we add around it some mass of the same substance as that of the balloons; the added volume is that bounded by the two contact balloons and a torus touching both of them, as shown in Figure 7. The geometry made by adding some mass is not exactly the same as that of real sintered balloon particles, but at least the shape of the void space in the mass-added spherical aggregate is much more similar to the real one than the infinitely long needle shape [Kurashige et al. 1992], which was assumed when the self-consistent models [Wu 1966; Berryman 1981] were applied. In Figure 7, D and d are the outer and inner diameters of the balloon, respectively. This added geometry was also used in [Sanders and Gibson 2003b; Sanders and Gibson 2003a].

The added volume per contact point can be calculated by simple geometrical consideration as

$$V_{\text{add}} = \frac{\pi}{4} \left(\frac{1 - \cos \Theta}{\cos \Theta} \right)^2 \left\{ 1 - \left(\frac{\pi}{2} - \Theta \right) \tan \Theta \right\} \times D^3, \tag{7}$$

where Θ is the angle depicted in Figure 7. We will adopt the angle Θ as a parameter of sintering degree; we call it a *sintering degree angle*. In [Sanders and Gibson 2003b; Sanders and Gibson 2003a], this is called a *bond angle*.

4.2. Results of FEM analysis. The spring has six degrees of freedom: one for elongation and torsion, and two for bending and shearing. The spring constants are defined in the same manner as in the conventional structural analysis. Consider an hour glass shape of the two sintered half balloons shown in Figure 7. Fix its one end and give some generalized displacement on the other end; then we calculate the corresponding generalized forces, which provide the required spring constants. As a function of sintering degree angle Θ , we can obtain K_{ex} and M_{tx} for axial elongation $u_x = 1 \times 10^{-5} D$ and torsional angle $\theta_r = 1 \times 10^{-2}$, respectively; K_{bx} and M_{bz} for bending angle and K_{sy} and M_{sz} for lateral displacement. With respect to these notations, K and M correspond to the associated force and moment (or torque); the

first subscript represents the deformation modes, while the second one describes displacement direction or rotation axis. Notice that $K_{sz} = K_{sy}$, $M_{sy} = M_{sz}$ and $K_{bz} = K_{by}$ because of the geometrical axisymmetry.

We employ triangular ring elements. Numbers of nodes and elements are 2640 and 4949, respectively, for the thickest-walled balloons, while the thinnest-walled model contains 9989 nodes and 18545 elements. The numbers of nodes and elements are enough to evaluate the spring constants, as has been checked for the case of the smallest sintering angle by varying these numbers.

The FEM calculations have been carried out setting the Poisson's ratio of the balloon material to 0.25. It is reported in [Sanders and Gibson 2003b] that varying the Poisson's ratio has a negligible effect on the overall response.

All the spring constants normalized by Young's modulus of the balloon substance E_s versus the sintering degree angle Θ are given in Figures 8a to 8f for various inner/outer diameter ratios d/D ranging from 0.0—0.9 with their increment being 0.1. Figures 8a and 8b show K_{ex} for the axial elongation and M_{tx} for the torsion, respectively. Figures 8c and 8d depict K_{bx} and M_{bz} for the bending. Shown in Figures 8e and 8f are K_{sy} and M_{sz} for the lateral displacement with no gradient at the end, respectively. For all inner/outer ratios, these spring constants monotonically increase with the sintering angle, seeming to converge to zero with vanishing angle. Notice that $K_{bx} = M_{sz}$ because of the reciprocal relation. We confirmed that these spring constants numerically satisfy this relation (see Figures 8c and 8f).

These spring constants will be used to evaluate the macroscopic, effective or overall elastic moduli of the aggregate in the next section. Since the average coordination number for the random packing constructed is approximately 6.02, and very near to 6 for the regular simple cubic packings, we will simulate the tensile and shear tests up to the range of $\Theta = 45^\circ$ in the next section. The lower limit is set to 3° .

5. Macroscopic material properties

5.1. Simulation method. Now, we will do tensile and shear tests of a *random network* of springs, with rigid hinges, of six degrees of freedom with the estimated spring constants. The random network is a mechanical model of sintered balloons. The model might be satisfactory if deformation of each hourglass-shaped spring is concentrated around its neck or near the sintered portion. This concentrated deformation for all springs would account for the overall deformation of the sintered aggregate. To confirm this is the case, we calculate strain energy density within the balloon pair from the above FEM analysis. The results are shown in Figure 9 for the case of a rather thin and well sintered balloon, $d/D = 0.9$ and $\Theta = 30^\circ$. From this figure, it is seen that the high density region is near the sintered portion. Although the results are not illustrated here, the strain energy density obtained by the further FEM calculation depicts that the deformation occurs only around the neck even in the case of a thin neck, e.g., $\Theta = 5^\circ$, independently of the balloon wall thickness. This argument reveals that both ends of the hourglass-shaped spring, which are far from the neck or sintered portion, is little deformed and plays a roll of the rigid hinge.

Since it was found that more spheres are in contact with others around the direction of $\theta = 50.1^\circ$ from the vertical line, the porous media modeled by the random networks are expected to be transversely isotropic in their elastic moduli. So, we carry out both the three different tensile tests and shear tests of the networks. From the tensile tests, we obtain the macroscopic Young's moduli E_x^* , E_y^* and E_z^* and

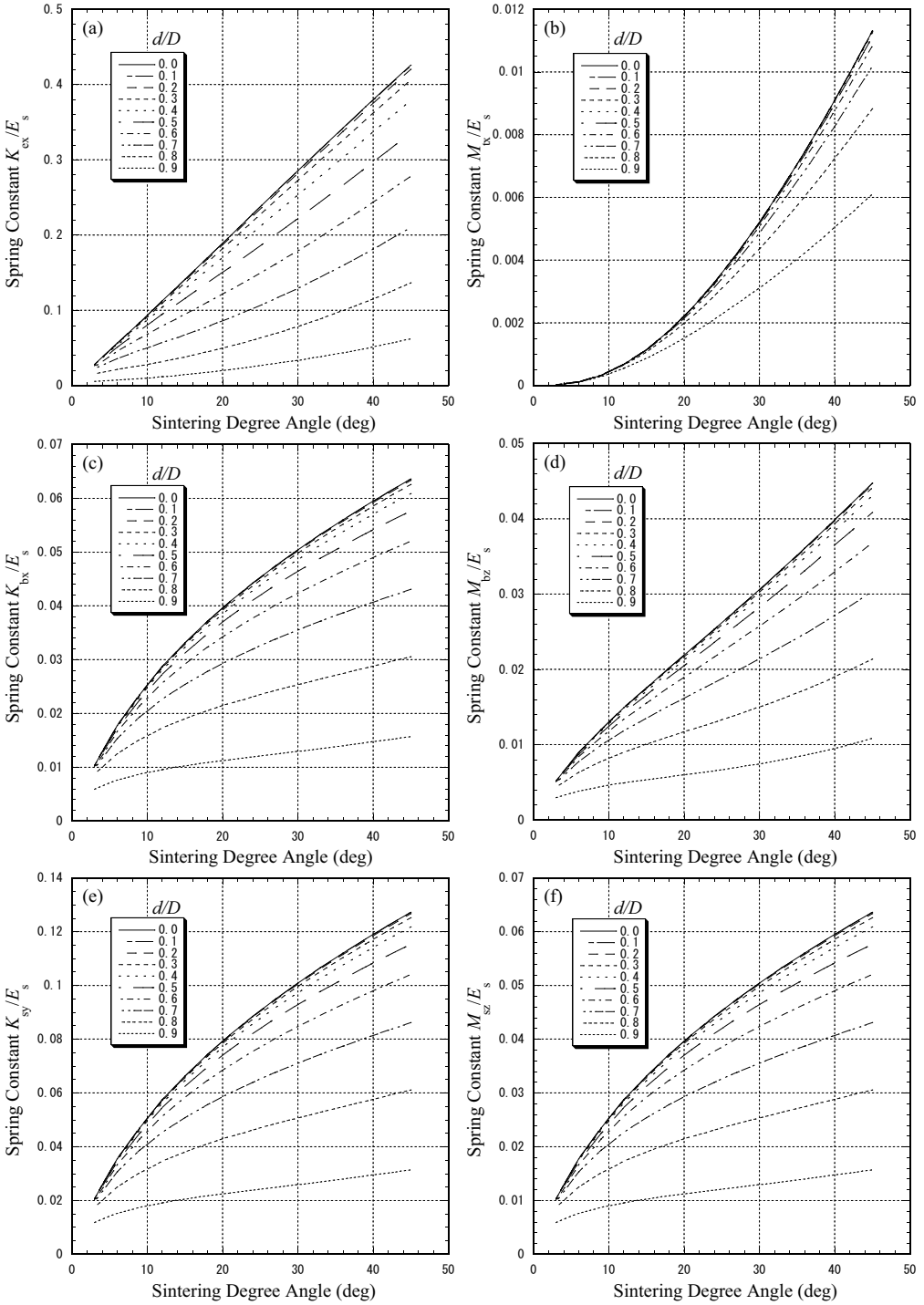


Figure 8. Spring constants versus sintering degree angle for (a) elongation; (b) torsion; (c) & (d) bending; and (e) & (f) shearing.

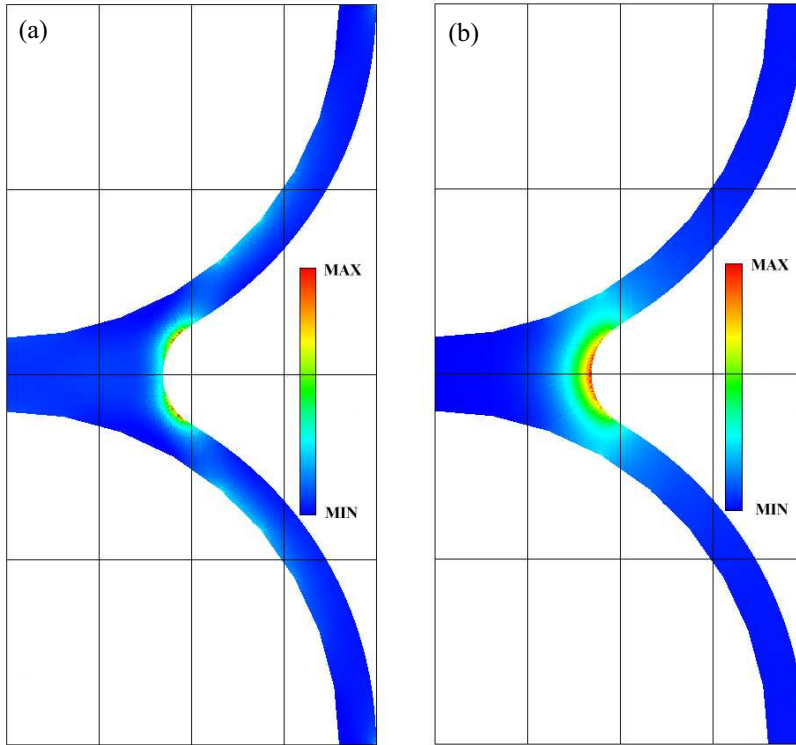


Figure 9. Distribution of strain energy density for (a) elongation and (b) torsion.

Poisson's ratios ν_{xy} , ν_{xz} , ν_{yx} , ν_{yz} , ν_{zx} , ν_{zy} according to their definition. Furthermore, we obtain G_{xy} , G_{xz} and G_{yz} from the shear tests.

We use a program for the 3D structural analysis, which is borrowed from a textbook [Beaufait et al. 1970], to simulate both of the tests.

5.2. Numerical results. Figures 10 and 11 respectively show macroscopic Young's moduli and shear moduli versus the total porosity in the sintered state; it should be noted that this porosity accounts for the hollow space of all the balloons as well, and that the added volume around all necks by sintering is subtracted from the whole void space. Although we have calculated all these moduli for all samples, we will present only graphs for sample 1, because we can see little difference among the samples for all diameter ratios. In all the figures, the ordinate presents their macroscopic Young's moduli E^* or shear moduli G^* normalized by Young's modulus of the balloon substance E_s ; d/D is adopted as a geometrical parameter.

Figure 10 shows the porosity dependence of Young's moduli E_x^*/E_s , while Figure 11 shows G_{xy}^*/E_s .

Comparing the results in Figure 10 with those in the figure for E_z^* omitted for brevity, we find that the value of E_z^* is roughly 20% larger than E_x^* over the entire region of porosity for all diameter ratios. This can be understood from the following facts: all springs lying in the direction of around 50.1° from the

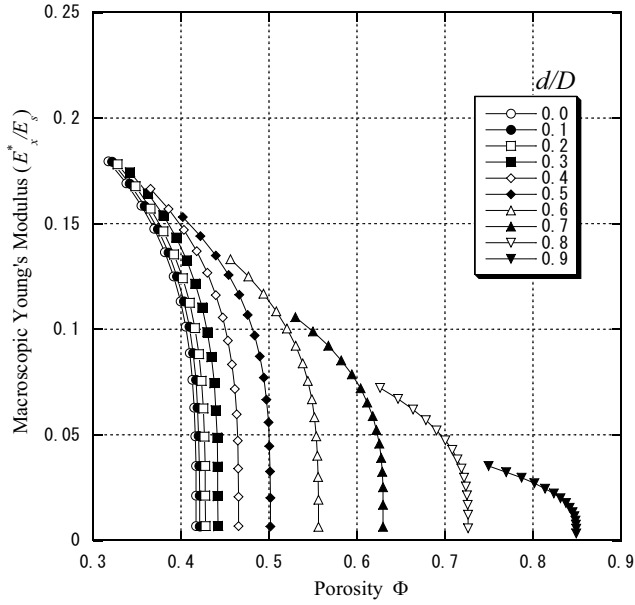


Figure 10. Young's moduli in isotropy plane versus porosity.

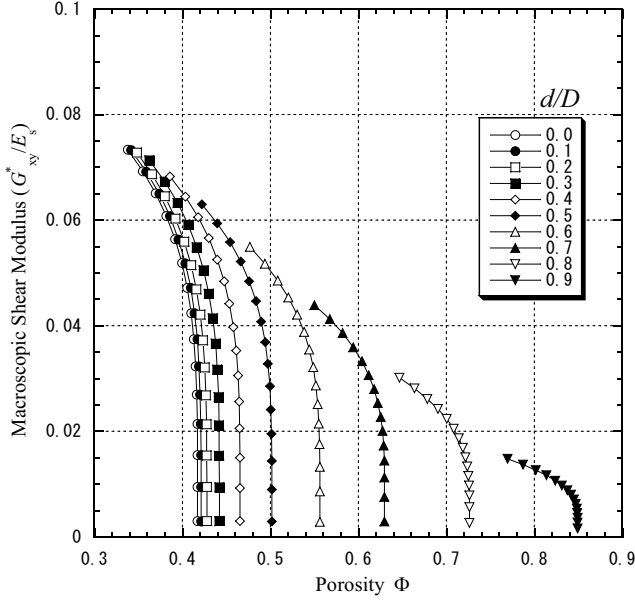


Figure 11. Shear moduli in isotropy plane versus porosity.

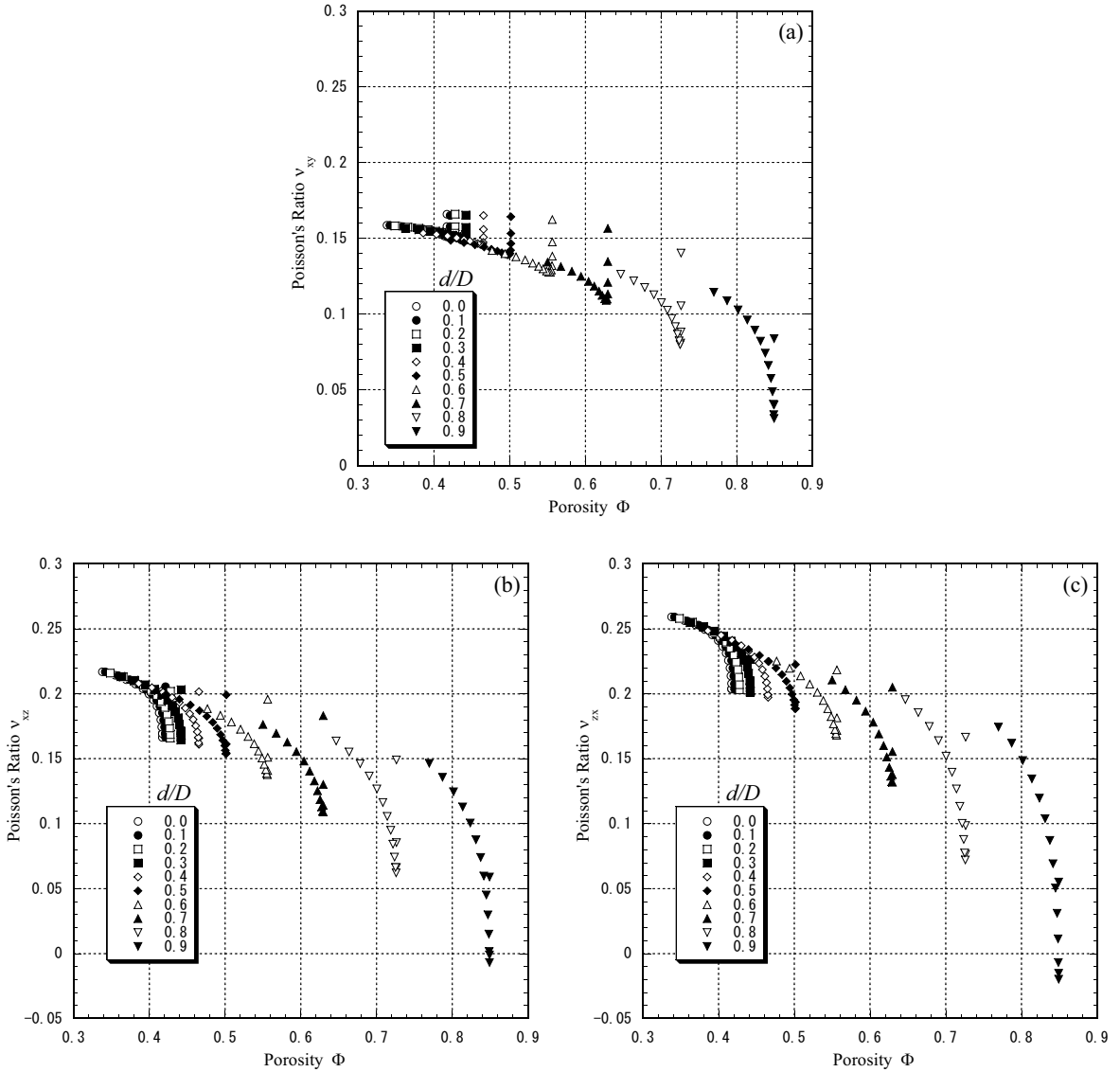


Figure 12. Three independent Poisson's ratios versus porosity: (a) ν_{xy} , (b) ν_{xz} and (c) ν_{zx} , for d/D ranging from 0.0—0.9

vertical line support forces in the z direction, while only some parts of them lying around the $x - z$ plane support forces in the x direction but some parts of them lying around the $y - z$ plane do not. Similarly, G_{xy}^* shown in Figure 11 is larger than G_{xz}^* , not shown, by more than approximately 30% over the whole range of porosity. Furthermore, we can numerically confirm, although not shown, that $E_x^* = E_y^*$ and $G_{xz}^* = G_{yz}^*$. These observations show again that the packing structure of the simulation is of transverse isotropy.

Next, we examine dependence of the elastic moduli on the degree of sintering and the diameter ratio; each marker on an individual curve in both the figures corresponds to the sintering degree angle ranging from 3° to 45° at the interval of 3° ; and progress toward the right on the curve makes the sintering degree low. For each constant diameter ratio, the smaller the degree of sintering is, the smaller the elastic moduli. Increases in the diameter ratio lower the curves as a whole, especially so for the higher degree of sintering. They also shift the curves toward higher porosity as a whole, because the balloon wall becomes thinner. The elastic moduli are little affected by the ratio if the ratio is smaller or for the thicker balloon wall. The elastic moduli lowers rapidly if the diameter ratio is larger than 0.5. The elastic modulus of the aggregate of hollow spheres with $d/D = 0.9$ is approximately a quarter of that of solid spheres. All these things are valid for all the figures, including the omitted. This is also true for Young's moduli E_y^* and for shear moduli G_{yz}^* , as confirmed by further calculation.

It should be added that both Young's and shear moduli go to zero as the sintering angle tends to zero for all the diameter ratios even for a finite value of the porosity; the porosity depends on the diameter ratio as well as on the sintering angle and packing structures.

In the above, we have mentioned that Young's modulus E_z^* and shear modulus G_{xz}^* are larger than E_x^* and G_{xy}^* by 20% and 30%, respectively. These moduli, however, depend on the diameter ratio and sintering angle in a very complicated manner, so the degree of anisotropy cannot be described by a simple multiplication factor. Therefore, it may be important for practical purposes to express all the elastic moduli in terms of a simple—that is, polynomial—expression of the two geometrical parameters d/D and Θ . The expressions are determined by the conventional least square method as follows:

$$\begin{aligned} \frac{E_x^*}{E_s} = & \left(4.24 \times 10^{-3} + 2.91 \times 10^{-3} \frac{d}{D} - 7.45 \times 10^{-3} \left(\frac{d}{D} \right)^2 \right) \Theta \\ & + \left(-5.42 \times 10^{-6} - 2.57 \times 10^{-5} \frac{d}{D} + 3.43 \times 10^{-5} \left(\frac{d}{D} \right)^2 \right) \Theta^2, \quad (8) \end{aligned}$$

$$\begin{aligned} \frac{E_z^*}{E_s} = & \left(5.28 \times 10^{-3} + 6.64 \times 10^{-3} \frac{d}{D} - 9.27 \times 10^{-3} \left(\frac{d}{D} \right)^2 \right) \Theta \\ & + \left(-1.16 \times 10^{-5} - 3.42 \times 10^{-5} \frac{d}{D} + 4.96 \times 10^{-5} \left(\frac{d}{D} \right)^2 \right) \Theta^2, \quad (9) \end{aligned}$$

$$\begin{aligned} \frac{G_{xy}^*}{E_s} = & \left(1.82 \times 10^{-3} + 1.27 \times 10^{-3} \frac{d}{D} - 3.16 \times 10^{-3} \left(\frac{d}{D} \right)^2 \right) \Theta \\ & + \left(-4.17 \times 10^{-6} - 1.15 \times 10^{-5} \frac{d}{D} + 1.6 \times 10^{-5} \left(\frac{d}{D} \right)^2 \right) \Theta^2, \quad (10) \end{aligned}$$

$$\begin{aligned} \frac{G_{xz}^*}{E_s} = & \left(2.25 \times 10^{-3} + 1.44 \times 10^{-3} \frac{d}{D} - 3.9 \times 10^{-3} \left(\frac{d}{D} \right)^2 \right) \Theta \\ & + \left(-2.07 \times 10^{-6} - 1.33 \times 10^{-5} \frac{d}{D} + 1.75 \times 10^{-5} \left(\frac{d}{D} \right)^2 \right) \Theta^2. \quad (11) \end{aligned}$$

All the equations are valid for $0.0 \leq d/D \leq 0.9$ and $3^\circ \leq \Theta \leq 45^\circ$ to within 8.5%.

Now, we will move on to discuss Poisson's ratio. The topic is a little more complicated than that of the Young's and shear moduli.

Figures 12a—c show the behavior of Poisson's ratios, ν_{xy} , ν_{xz} , ν_{zx} , respectively, against the total porosity for various inner/outer diameter ratios d/D ranging from 0.0—0.9 with their increment being 0.1; ν_{xy} means Poisson's ratio in the isotropy plane; ν_{xz} denotes the contraction ratio in the z axis under uniaxial tension in the x direction, while ν_{zx} the one in the x axis under z tension. The estimation of all Poisson's ratios here are based on the described definitions; in other words, not on some relations among other elastic moduli estimated.

The macroscopic Poisson's ratio obtained is smaller than that of the balloon substance $\nu = 0.25$ with exceptional cases of ν_{zx} , for very small porosity and small diameter ratios. It seems to be a general tendency that Poisson's ratio of porous materials is smaller than that of the original material. Few papers report on this topic but Sanders and Gibson [2003a] give much smaller Poisson's ratios for the BCC and FCC packings than the original one.

In Figure 12, it can be seen that the dependence of Poisson's ratios on porosity is similar to that of Young's and shear moduli. The ratios decrease with decreasing sintering degree angle for each diameter ratio. They are little affected by the diameter ratio if it is smaller (or for the thicker balloon wall), more precisely for $d/D < 0.4$. The Poisson's ratios decrease rapidly if the diameter ratio is larger than 0.5. Poisson's ratio ν_{xy} in the isotropy plane is the smallest, while ν_{zx} is the largest; ν_{xz} lies between them.

It is interesting that all Poisson's ratios sharply ascend with an approach of porosity to the nonsintered state. This is supported by the similar ascent for the solid sphere aggregates obtained by Kato et al. [2002], but it may be true on condition that these results are sufficient in their accuracy in the limit.

We have already given the formulae to estimate the four elastic moduli. The final modulus of five independent moduli for transversely isotropic media, if we select ν_{zx} , can be expressed in terms of d/D and Θ as follows:

$$\begin{aligned} \nu_{zx} = & \left(1.38 \times 10^{-1} + 3.19 \times 10^{-3} \left(\frac{d}{D} \right)^{2.2} - 5.41 \times 10^{-1} \left(\frac{d}{D} \right)^{5.2} \right) \Theta^{0.1} \\ & + \left(8.1 \times 10^{-3} - 2.16 \times 10^{-3} \left(\frac{d}{D} \right)^{1.2} + 1.18 \times 10^{-1} \left(\frac{d}{D} \right)^{6.8} \right) \Theta^{0.5}. \quad (12) \end{aligned}$$

The equation describes the results for $0.0 \leq d/D \leq 0.9$ and $9^\circ \leq \Theta \leq 45^\circ$ to within 10% in all cases. Note that the lower limit of the sintering degree is 9° to avoid the rapid ascent in the ratio in approaching the nonsintered state; the relative error is larger for the degree smaller than that. More important, the expression given by Equation (12) is not a polynomial of d/D and Θ , because the dependence of the ratio on these parameters is too complicated to express by means of a polynomial.

For the transverse isotropy in elasticity, the reciprocal relations

$$\begin{aligned} \frac{\nu_{xz}}{E_x^*} &= \frac{\nu_{zx}}{E_z^*}, \\ \frac{\nu_{yz}}{E_y^*} &= \frac{\nu_{zy}}{E_z^*} \end{aligned} \quad (13)$$

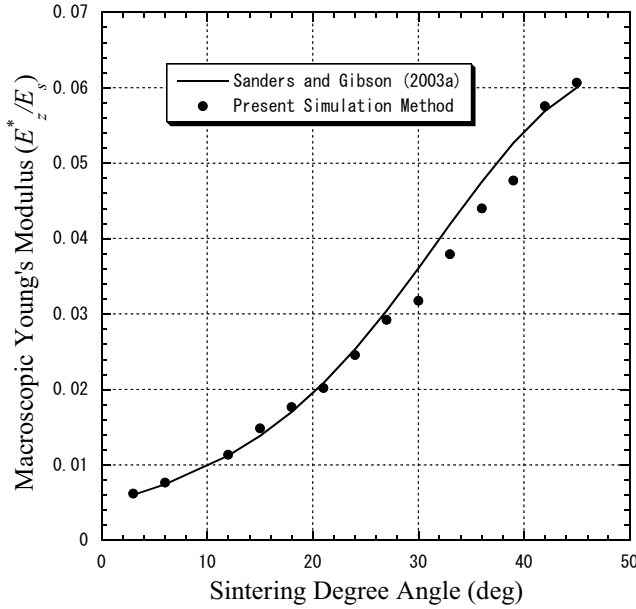


Figure 13. Relative Young's modulus in [100] for SC packing calculated by present method versus sintering degree angle together with that of [Sanders and Gibson 2003b].

must be fulfilled. Poisson's ratio ν_{xz} can be determined from the first relation. In addition, $\nu_{yz} = \nu_{xz}$ and $\nu_{zy} = \nu_{zx}$. We have confirmed that the present results almost satisfy these relations.

Poisson's ratio ν_{xy} can be expressed in terms of E_x^* and G_{xy}^* as

$$\nu_{xy} = \frac{E_x^*}{2G_{xy}^*} - 1. \quad (14)$$

However, Poisson's ratios ν_{xy} calculated by Equation (14) differ considerably from the simulated ones. If G_{xy}^* deviates by $\pm 5\%$ in Equation (14), the ν_{xy} varies by as much as 30%. Thus, we cannot use Equation (14) to determine Poisson's ratio ν_{xy} . Instead, it can be obtained from the following expression in terms of d/D and Θ in a manner similar to that of Equation (12):

$$\begin{aligned} \nu_{xy} = & \left(7.96 \times 10^{-2} - 2.24 \times 10^{-1} \left(\frac{d}{D} \right)^{1.9} + 3.48 \times 10^{-1} \left(\frac{d}{D} \right)^{2.5} \right) \Theta^{0.4} \\ & + \left(-1.01 \times 10^{-2} - 5.2 \times 10^{-2} \left(\frac{d}{D} \right)^{1.9} + 7.68 \times 10^{-2} \left(\frac{d}{D} \right)^{2.5} \right) \Theta^{0.8}. \quad (15) \end{aligned}$$

The equation describes the results for $0.0 \leq d/D \leq 0.9$ and $9^\circ \leq \Theta \leq 45^\circ$ to within 10% in all cases.

We have determined the in-plane shear modulus G_{xy}^* from the shear tests. However, we can calculate G_{xy}^* from only the tensile test using Equation (14). So we estimate G_{xy}^* from Equation (14). The results fit those from the shear test to within $\pm 5\%$.

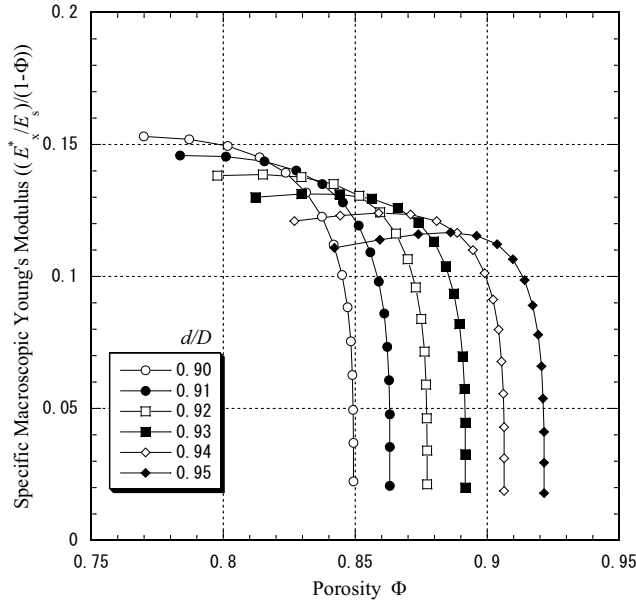


Figure 14. Specific Young's moduli in isotropy plane versus porosity for d/D ranging from 0.9—0.95.

6. Discussion

6.1. Validity of the proposed simulation method. We have simulated the elastic properties of randomly packed sintered balloons using the three-step simulation method. Here, we will confirm validity of the present method by applying it to the simple cubic packing.

[Sanders and Gibson 2003b] evaluated the mechanical properties of simple cubic packed hollow-sphere foams by the FEM analysis of a unit cell with the periodic boundary conditions. We calculate the relative Young's and shear moduli for the same SC packings by the present method. Figure 13 compares the relative Young's modulus for the SC packing calculated by the present method with that obtained from the formula given by [Sanders and Gibson 2003b]. A good agreement is seen between them for a wider range of the sintering degree angle, although Sanders and Gibson pointed out that their formula is applicable to the range from 10° to 40° .

6.2. Better design; larger specific moduli. One of the advantageous properties of the sintered balloon compacts is a high specific modulus. To consider the better design of compacts, we examine a specific relative modulus, which is defined as the relative effective modulus divided by the relative effective density of a sintered aggregate, where the term *relative* means the quantity divided by that of the original balloon material. In what follows, this term will be omitted in almost all cases. We have calculated the specific Young's and shear moduli for a wide range of the diameter ratios, but will not show all of them. In Figures 14 and 15, we will show the specific moduli only for the case of the diameter ratio larger than 0.9, because commercially available balloons have rather thin walls. Both figures illustrate the moduli in the isotropy plane.

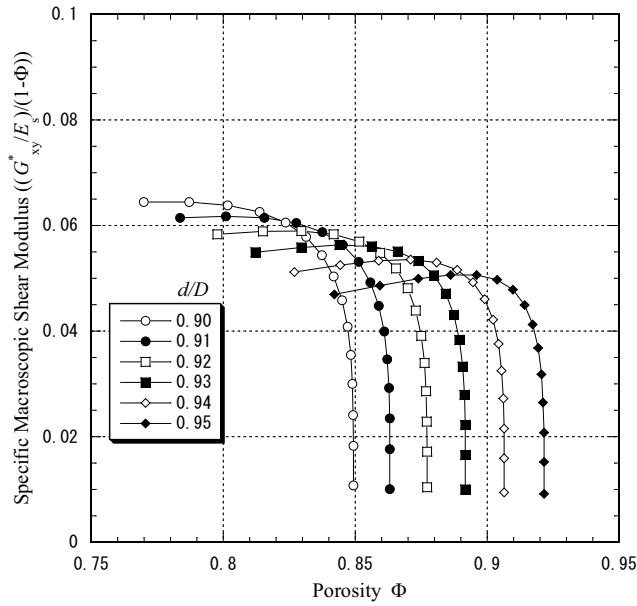


Figure 15. Specific shear moduli in isotropy plane versus porosity for d/D ranging from 0.9—0.95.

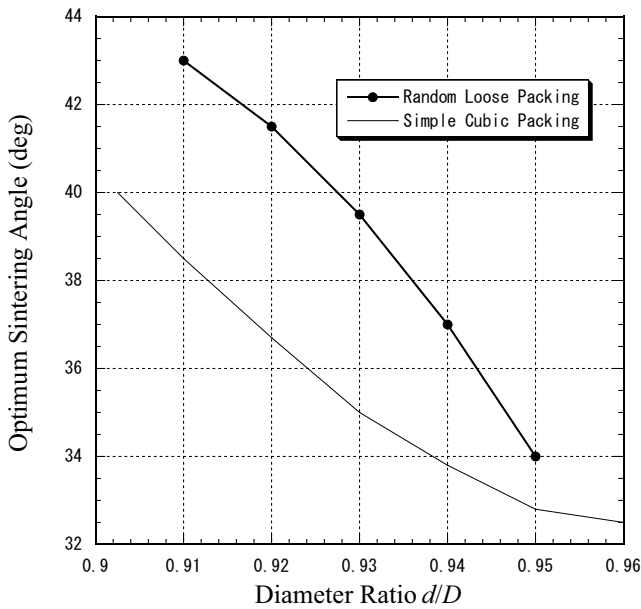


Figure 16. Optimum sintering degree angle plotted as a function of inner/outer diameter ratio for random and SC packings.

In the first place, we have to point out that the curves of the specific moduli against the porosity have a peak both for Young's and shear moduli in the isotropy plane. This is also true for those out of the plane.

It seems that there is a complicated mechanism behind the peak occurrence. Let us consider the case of Young's modulus, so that it may be enough to refer to only Figure 14 and, in addition, Figure 10 for comparison. See the curve of $d/D = 0.9$ in the latter figure; it can be seen that the modulus gradually slows down with increase in porosity for the smaller porosity range and rapidly falls down for the larger porosity, approaching the point contact state; the gradual slow-down is caused by deformation of upper and lower thin shell parts around a neck and this deformation is not strongly affected by the sintering degree, while the sharp falling is due to the concentrated deformation near a contact point because of the less sintered state. On the other hand, because the balloon thickness is constant on the given curve, decrease in the porosity implies increase in the sintering degree. This increase makes the specific density rise; dividing the modulus depicted in Figure 10 by this increased specific density leads to the smaller specific modulus shown in Figure 14.

In many cases, the largest specific modulus is the optimum porous material with respect to rigidity. Figure 16 shows the optimum sintering angle for the Young's modulus, in the z direction, of sintered random packings, plotted as a function of the diameter ratio together with that given for the SC regular packings by [Sanders and Gibson 2003b]. It is seen that the optimum angle of the random packings higher than that of SC. It should be added that the Young's and shear modulus (the latter is not shown) in the isotropy plane have almost the same optimum sintering angle and the same thing is true for the moduli out of the plane (not shown).

In the material design, therefore, we must select the best sintering condition from this result.

6.3. Negative Poisson's ratios. Here, we will back to Figures 12a—12c. Scrutinizing the figures, we find that the Poisson's ratios are negative in the case of very thin balloon wall and very low sintering degree. Because of this interesting characteristic, we have calculated the Poisson's ratios for the case of $d/D = 0.8$ or more in detail and shown them in Figures 17a—17c.

The negative Poisson's ratio implies that the uniaxial tensile (compressive) load applied on a body will lead to its expansion (contraction) in the direction orthogonal to the applied load.

A foam with negative Poisson's ratio was first fabricated by Lakes [1987]. In general, typical microstructures of foams with negative Poisson's ratio are of three types: inverted *re-entrant* cell shape [Lakes 1987; Friis et al. 1988], solid particles attached to each other by elastic strips [Lakes 1991; Prall and Lakes 1997] and nodes, connected by tensile springs, constrained by hinged inextensible rods [Evans and Caddock 1989]. However, the slightly sintered random balloon aggregates have none of such structures as mentioned just in the above.

Using the discrete element method for a random granular material, Bathurst and Rothenburg [1988] and Alzebdeh and Ostoja-Starzewski [1999] showed that effective Poisson's ratio for an aggregate of two-dimensional irregular particles is negative in some cases even when the individual Poisson's ratios of particles are positive. Bathurst and Rothenburg [1988] examined negative Poisson's ratio behavior by varying the ratio of a normal stiffness and shear stiffness between particles. They found that the Poisson's ratio is negative when the shear stiffness is higher than the normal stiffness.

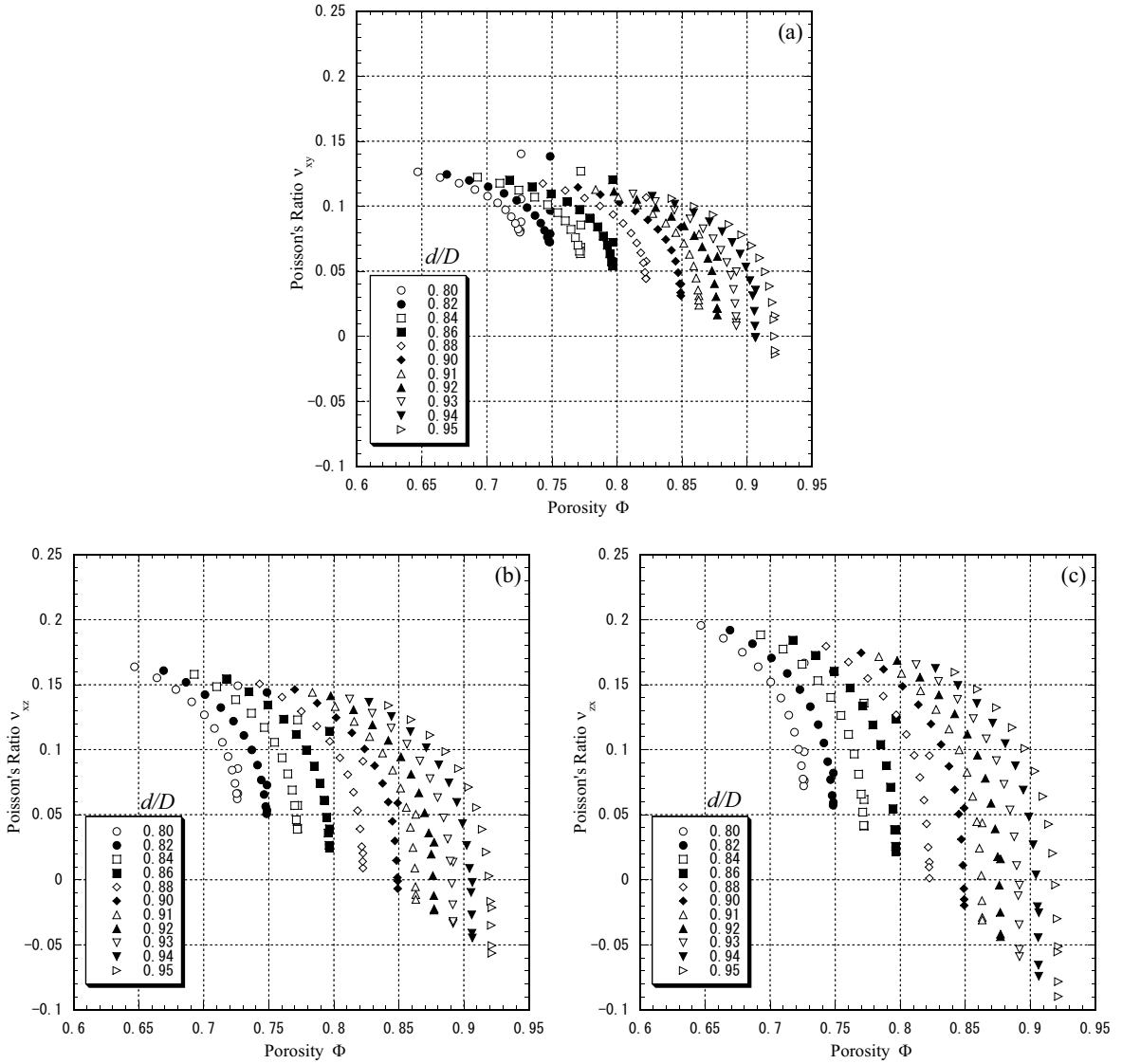


Figure 17. Three independent Poisson's ratios versus porosity: (a) ν_{xy} , (b) ν_{xz} and (c) ν_{zx} , for d/D ranging from 0.8 — 0.95.

Following their discussion, we pick up only the shear stiffness K_{sy} and normal stiffness K_{ex} of a connected balloon pair from the six spring constants calculated in the above FEM analysis. For a solid sphere of $d/D = 0.0$, if the sintering degree angle varies down from 45° to 3° (this descending degree means the shift of the markers toward the right or higher porosity on each curve in Figure 12), their ratio K_{sy}/K_{ex} changes from 0.2 to 0.37, whereas, for a thinner balloon of $d/D = 0.9$, the ratio changes from 0.5 to 2.1. A rough estimate of the variation of the Poisson's ratio against that of the stiffness ratio shows that the Poisson's ratios are negative when K_{sy}/K_{ex} is greater than 2. More precisely, this critical value

slightly differs from each other for the three independent Poisson's ratios; remember that we have the three distinct ratios because the sintered aggregates under consideration are transversely isotropic.

6.4. Comparison with SC packing. Finally, we compare the Young's modulus of the present random loose (RL) packing with that of the regular packing. Here, we select the SC packing from among regular packings, because the average coordination number for the random packing constructed is approximately 6 and very near to that for the regular SC packings, whose coordination number is exactly 6. It should be mentioned here that the present RL packing is of axial symmetry with the five independent elastic moduli while the SC packing is of cubic symmetry with the three.

In Figure 18, the Young's moduli, E_x^* and E_z^* , of the RL packing are plotted versus the sintering degree in the case of $d/D = 0.9$. Also shown in the figure are the Young's moduli in the $[1\ 0\ 0]$, $[1\ 1\ 0]$ and $[1\ 1\ 1]$ directions for the SC packing for the same d/D ; these moduli are calculated by using the standard transformations of elastic moduli from the three compliances, S_{11} , S_{12} and S_{44} , in the form of the fitting formula given by [Sanders and Gibson 2003b].

A good agreement is seen between the Young's modulus of the random packing in the z direction and that of the SC in the $[1\ 0\ 0]$ direction only when $\Theta < 20^\circ$. However, for $\Theta > 20^\circ$, the latter is greater than the former; the difference between them increases with increasing sintering degree angle. For the sintering degree angle $\Theta = 30^\circ$ and diameter ratio $d/D = 0.9$, the relative Young's modulus is 38% greater in the $[1\ 0\ 0]$ direction of SC than in the z direction for the RL packing. For the same sintering degree angle and diameter ratio, the relative Young's modulus in the $[1\ 1\ 0]$ and $[1\ 1\ 1]$ directions of SC are 1.86 and 2.11 times that in the z direction for the RL packing.

Figure 19 shows the Young's modulus versus porosity for both the RL and SC packings when we have $d/D = 0.9$ and 0.95. For each diameter ratio, the Young's modulus of RL packing is much smaller than that of the regular SC stacking, although the difference between them is rather small when the sintering degree angle is about 20° . It can be concluded that, although the average coordination number for the random packing constructed is very near to that for the regular SC packings, the macroscopic properties of the RL packing are significantly different from those of the packing in elasticity and anisotropy.

7. Conclusions

We have evaluated the macroscopic elastic properties of the sintered, randomly packed balloons for various degrees of sintering and for a wide range of the balloon wall thickness. The conclusions are summarized below:

- (1) The packing structure in the vertical direction is different from that in other directions. Branches lie more frequently around the direction of 50.1° from the vertical line than other directions, whereas they are uniformly distributed about the vertical line. The packings constructed by the present method are of transverse isotropy; in other words, the structure is affected by gravity.
- (2) The elastic moduli are little affected by the diameter ratio for thicker balloons whose ratio is less than about 0.4, while the properties of thinner balloons are very sensitive to the ratio; they rapidly decline with decreasing wall thickness of the balloons.
- (3) The moduli are larger in the direction of gravity than in the horizontal one, as expected for the gravity-affected packing structures.

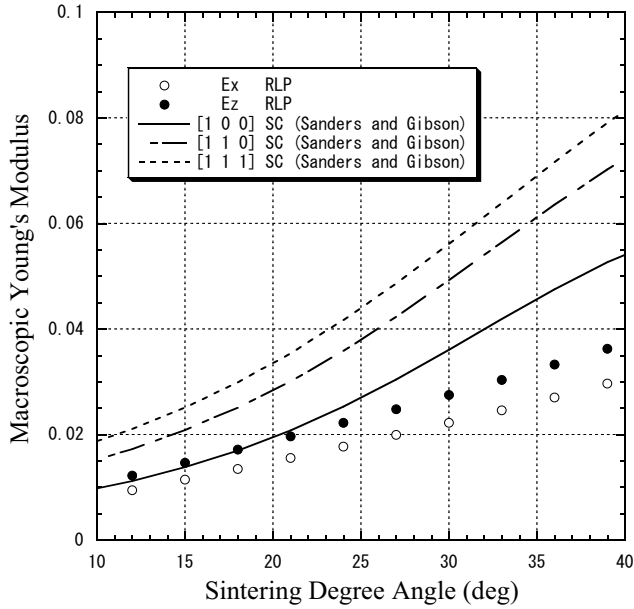


Figure 18. Relative Young's modulus plotted against sintering degree angle for random loose packing and SC packing.

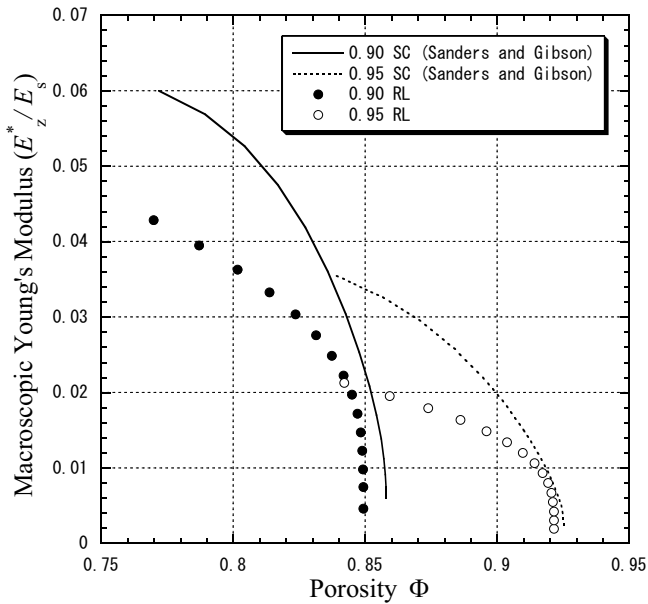


Figure 19. Relative Young's modulus plotted against porosity for random loose packing and SC packing for $d/D = 0.9$ and 0.95 .

- (4) The specific elastic moduli are found to have a peak when the balloons have a diameter ratio larger than about 0.9.
- (5) The Poisson's ratios are slightly negative in the case of very thin balloon walls and very low sintering degree.
- (6) The macroscopic properties of the RL packing are significantly different from those of the SC packing in elasticity and anisotropy, although the both packings have almost the same coordination number.

References

- [Alzebedeh and Ostoja-Starzewski 1999] K. Alzebedeh and M. Ostoja-Starzewski, "On a spring-network model and effective elastic moduli of granular materials", *J. Appl. Mech. (ASME)* **66** (1999), 172–180.
- [Andersen et al. 2000] O. Andersen, U. Wang, L. Schnerder, G. Stephani, and B. Kieback, "Novel metallic hollow sphere structures", *Adv. Eng. Mater.* **2**:4 (2000), 192–195.
- [Andrews et al. 1999] E. Andrews, W. S. Sanders, and L. J. Gibson, "Compressive and tensile behaviour of aluminum foams", *Mater. Sci. Eng. A* **270**:2 (1999), 113–124.
- [Ashby et al. 2000] M. F. Ashby, N. Evans, A. Fleck, L. J. Gibson, J. W. Hutchinson, and H. N. G. Wadley, *Metal foams: A design guide*, Butterworth-Heinemann, 2000.
- [Bardella and Genna 2001] L. Bardella and F. Genna, "On the elastic behavior of syntactic foams", *Int. J. Solids Struct.* **38**:40 (2001), 7235–7260.
- [Bathurst and Rothenburg 1988] R. J. Bathurst and L. Rothenburg, "Note on a random isotropic granular material with negative Poisson's ratio", *Int. J. Eng. Sci.* **26**:4 (1988), 373–383.
- [Beaufait et al. 1970] F. W. Beaufait, W. H. Rowan, Jr, P. G. Hoadley, and R. M. Hackett, *Computer methods of structural analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1970.
- [Bennett 1972] C. H. Bennett, "Serially deposited amorphous aggregates of hard spheres", *J. Appl. Phys.* **43**:6 (1972), 2727–2734.
- [Berryman 1981] J. G. Berryman, "Elastic wave propagation in fluid-saturated porous media", *J. Acoust. Soc. Am.* **69**:2 (1981), 416–424.
- [Debbas and Rumpf 1966] D. Debbas and H. Rumpf, "On the randomness of beds packed with spheres or irregular shaped particles", *Chem. Eng. Sci.* **21**:6–7 (1966), 583–608.
- [Evans and Caddock 1989] K. E. Evans and B. D. Caddock, "Microporous materials with negative Poisson's ratios, II: Mechanisms and interpretation", *J. Phys. D Appl. Phys.* **22**:14 (1989), 1883–1887.
- [Friis et al. 1988] R. S. Friis, R. S. Lakes, and J. B. Park, "Negative Poisson's ratio polymeric and metallic foams", *J. Mater. Sci.* **23**:12 (1988), 4406–4414.
- [Gasser et al. 2003] S. Gasser, F. Paun, A. Cayzeele, and Y. Bréchet, "Uniaxial tensile elastic properties of a regular stacking of brazed hollow spheres", *Scr. Mater.* **48**:12 (2003), 1617–1623.
- [Gasser et al. 2004a] S. Gasser, Y. Brechet, and F. Paun, "Materials design for acoustic liners: an example of tailored multifunctional materials", *Adv. Eng. Mater.* **6**:1–2 (2004), 97–102.
- [Gasser et al. 2004b] S. Gasser, F. Paun, and Y. Bréchet, "Finite element computation for the elastic properties of a regular stacking of hollow sphere", *Mater. Sci. Eng. A* **379**:1–2 (2004), 240–244.
- [Gibson and Ashby 1997] L. J. Gibson and M. F. Ashby, *Cellular solids: structure and properties*, 2nd ed., Cambridge University Press, Cambridge, 1997.
- [Haughey and Beveridge 1969] D. P. Haughey and G. S. G. Beveridge, "Structural properties of packed beds", *Can. J. Chem. Eng.* **47** (1969), 130–140.
- [Kato et al. 2002] H. Kato, C. Matsunaga, M. Kurashige, and K. Imai, "Anisotropy in packing structure and elasticity of sintering spherical particles", *JSME Int. J. A Mech. M.* **45**:4 (2002), 585–595.

- [Kurashige et al. 1992] M. Kurashige, K. Imaida, and Y. Goto, "Phase velocity measurement for three bulk waves in water-saturated sintered glass beads by an ultrasonic mode conversion method", pp. 508–513 in *Proceedings of the International Symposium on Impact Engineering* (Sendai), vol. 2, edited by I. Maekawa, 1992.
- [Kurashige et al. 1999a] M. Kurashige, T. Hayashi, and K. Imai, "Simulated effective elastic moduli and wave velocities in water-saturated sintered glass-beads", *Acta Mech.* **132**:1–4 (1999), 177–194.
- [Kurashige et al. 1999b] M. Kurashige, M. Mishima, and K. Imai, "Simulated effective thermal conductivity of sintered, randomly packed spheres and statistical structures of packings", *J. Therm. Stresses* **22**:7 (1999), 713–733.
- [Lakes 1987] R. S. Lakes, "Foam structures with negative Poisson's ratio", *Science* **235**:4792 (1987), 1038–1040.
- [Lakes 1991] R. S. Lakes, "Deformation mechanisms in negative Poisson's ratio: structural aspects", *J. Mater. Sci.* **26**:9 (1991), 2286–2292.
- [Marur 2005] P. R. Marur, "Effective elastic moduli of syntactic foams", *Mater. Lett.* **59**:14–15 (2005), 1954–1957.
- [Nolan and Kavanagh 1992] G. T. Nolan and P. E. Kavanagh, "Computer simulation of random packing of hard spheres", *Powder Technol.* **72**:2 (1992), 149–155.
- [Norris and Gojny 1990] B. Norris and F. J. Gojny, "Hollow metal sphere filled stabilized skin structures and method of making", United States Patent. No. 4925740, 1990.
- [Prall and Lakes 1997] D. Prall and R. S. Lakes, "Properties of a chiral honeycomb with a Poisson's ratio of -1 ", *Int. J. Mech. Sci.* **39**:3 (1997), 305–314.
- [Queheillalt et al. 2002] D. T. Queheillalt, D. J. Sypeck, and H. N. G. Wadley, "Ultrasonic characterization of cellular metal structures", *Mater. Sci. Eng. A* **323**:1–2 (2002), 138–147.
- [Sanders and Gibson 2003a] W. S. Sanders and L. J. Gibson, "Mechanics of BCC and FCC hollow-sphere foams", *Mater. Sci. Eng. A* **352**:1–2 (2003), 150–161.
- [Sanders and Gibson 2003b] W. S. Sanders and L. J. Gibson, "Mechanics of hollow sphere foams", *Mater. Sci. Eng. A* **347**:1–2 (2003), 70–85.
- [Taguchi et al. 2006] I. Taguchi, M. Kurashige, and K. Imai, "Effects of cubic container's wall or floor on random packing structures of spherical particles", *JSME Int. J. A Mech. M.* **49**:2 (2006), 265–272.
- [Torobin 1986] L. B. Torobin, "Metal microspheres, filamented hollow metal microspheres and articles produced therefrom", United States Patent. No. 4582534, 1986.
- [Torquato 2001] S. Torquato, *Random heterogeneous materials: microstructure and macroscopic properties*, Springer, New York, 2001.
- [Tory et al. 1973] E. M. Tory, B. H. Church, M. K. Tam, and M. Ratner, "Simulated random packing of equal sphere", *Can. J. Chem. Eng.* **51** (1973), 484–493.
- [Wadley 2002] H. N. G. Wadley, "Cellular metals manufacturing", *Adv. Eng. Mater.* **4**:10 (2002), 726–733.
- [Wichmann and Hill 1982] B. A. Wichmann and I. D. Hill, "Algorithm AS 183: An efficient and portable pseudo-random number generator", *Appl. Stat.* **31**:2 (1982), 188–190.
- [Wu 1966] T. T. Wu, "The effect of inclusion shape on the elastic moduli of a two-phase material", *Int. J. Solids Struct.* **2**:1 (1966), 1–8.

Received 23 Apr 2006. Accepted 31 Aug 2006.

ISAO TAGUCHI: isao_taguchi@pref.aomori.lg.jp

Graduate School of Engineering, Iwate University, Ueda 4-3-5, Morioka, Iwate 020-8551, Japan

MICHIO KURASHIGE: kurashige@iwate-u.ac.jp

Department of Mechanical Engineering, Iwate University, Ueda 4-3-5, Morioka, Iwate 020-8551, Japan

CYLINDRICAL INDENTATION INDUCED DEFORMATION IN FACE-CENTERED CUBIC METAL SINGLE CRYSTALS

YONG XUE GAN, YUKI SAITO AND XI CHEN

We investigate the deformation field induced by a cylindrical indentation on a face-centered cubic single crystal of aluminum or copper. We first present experimental measurements of the load–displacement curve and the crystal lattice rotation field (under plane strain condition) of an aluminum single crystal subject to indentation, together with related results for a copper crystal. Next, finite element simulations of the lattice rotation and displacement field associated with the cylindrical indentation are provided. The numerical and experimental results about lattice rotation features are compared with theoretical predictions based on the single crystal plasticity. Finally, the displacement fields obtained from the numerical solutions and experiments are compared. Both electron backscatter diffraction experiments using scanning electron microscopy and finite element simulations show the existence of different slip sector boundaries in the single crystals, in agreement with theoretical predictions of active slip systems and dislocation structures.

1. Introduction

As a localized impression testing technique, indentation has been extensively studied for evaluating materials properties such as the modulus [He et al. 2006], hardness [Zhang et al. 2004], fracture toughness [Tanaka et al. 2003], and creep properties [Wen et al. 2006]. Indentation has several advantages over other mechanical property testing methods, as it can be easily carried out on small specimens with minimum sample preparation [Sastry 2005a]. Compared with conical or wedge indenters, a cylindrical indenter can apply an approximately constant load to the contact region, resulting in relatively stable stress and deformation fields [Sastry 2005b]. Thus, cylindrical indentation is a very useful method to characterize the constitutive behavior of materials.

Because the mechanical properties of hard materials are difficult to obtain by conventional tension, compression or cyclic loading tests, the indentation method has recently been applied to such materials, for instance in [Guillou et al. 1993] (ceramics), [Seo et al. 2003] (surface passive films) and [Stevenson et al. 2001] (intermetallic compounds). Indentation has also been used for evaluating the deformation behavior of ductile materials [Zhu et al. 2004; Fujiwara and Otsuka 1999]. Yamada and Ikeda [1975] investigated the deformation mechanism of indentation onto the surface of a copper single crystal, finding subgrain formation on the surface and misorientation. In [Kobayashi et al. 1990], indentation tests using a steel ball were carried out on the (001), (011) and (111) faces of α -CuAl single crystals. The plastic

Keywords: indentation, single crystal, anisotropic plasticity, deformation field, lattice rotation map, numerical simulations. This work is supported by a research initiation fund from Department of Mechanical Engineering at The Cooper Union. The numerical work is supported by the National Science Foundation (CMS-0407743). The experimental work is also supported in part by Columbia Nanomechanics Research Center. We acknowledge the usage of the shared experimental facilities at Columbia University.

deformation mechanism of the surface layer under different indentation configurations was interpreted. Besides the details on slip trace distribution, the dislocation density around the indentation on the (111) and (211) planes was revealed by etching pits on the surface.

The load-displacement behavior during nanoindentation of electrodeposited single crystal copper nanowires (about 500 nm in size) was studied by Bansal et al. [2005], and compared with that of bulk nanocrystalline and annealed copper. The hardness value for extruded polycrystalline copper nanowires with 50 nm grain size was reported as 2.1 GPa; the value for single crystal copper nanowire was found to be about 1.8 GPa.

Understanding the anisotropic properties associated with indentation into face-centered cubic (FCC) single crystals is very important because the deformation field is determined by the contribution of each active slip system. Kobayashi et al. [1990] report that the slip tends to occur in the [110] direction on the (001) and (011) crystal planes if the indentation loading is applied on the (001) or (110) crystallographic plane; they also found that in the case of indentation on the (111) plane, the slip occurred on two sets of planes, located in truncated triangular pyramids, one diverging and the other converging into the single crystal. Due to plastic anisotropy, the hardness of FCC single crystals is orientation-dependent. Based on this property, a method for determining the orientation of a single crystal by indentation testing was reported in [Chang and Sheu 1992]. In that work, the orientation of the crystal was determined through an exam of the slip lines formed around the indented regions in Al-Li single crystals.

In this paper, we present the experimental, numerical and analytical results obtained from the investigation of cylindrical indentation into FCC single crystals. Aluminum and copper crystals are employed in the study because these metallic crystals possess typical elastic-plastic constitutive behaviors. The indentation patterns, such as the slip line traces generated by cylindrical indenters, are sufficiently regular and the deformation state is considerably stable.

Section 2 describes the materials and experimental procedures used. Section 3 deals with the finite element simulation method and its implementation. The indentation load-displacement relationship and the crystal lattice rotation field under plane strain conditions of an aluminum single crystal are presented in Section 4. Also in the same section, finite element (FE) simulation solutions to the lattice rotation and displacement field associated with the cylindrical indentation into face-centered cubic (FCC) single crystals are provided. Correlation among the predictions on the deformation state by the single crystal plasticity theory, the numerical FE solutions and the experimental results of the lattice rotation features is established. In the last part of Section 4, a comparison is made of the indentation displacement field obtained from numerical simulations and from experiments. Conclusions are summarized in Section 5.

2. Materials and experiments

Materials. Two types of FCC single crystals are employed in this work. The first is an aluminum single crystal of 99.999% purity, grown from melt with the seeded Bridgman technique. The direction of solidification is $[1\bar{1}4]$. The as-grown crystal was etched in a 20% (by weight) NaOH aqueous solution to remove the surface oxide layer formed at the high temperatures required for crystal growth. Laue X-ray diffraction was conducted to determine the crystallographic orientation to within $\pm 1^\circ$. After taking the X-ray image, we reoriented the Al single crystal and performed a series of cuts using an electric discharge

machine (EDM) so that the $[110]$ is the surface normal; $[\bar{1}10]$ is aligned horizontally; and $[001]$ is aligned up vertically.

The second type of FCC single crystal used is copper single crystal. Pure copper single crystal was supplied in cylindrical form with the axis oriented along $[001]$ crystallographic direction. The crystal was cut using the same electrical discharge machine, so that the (110) plane was exposed. Laue X-ray diffraction for this crystal was also conducted to confirm the orientation.

Sandpaper with various grits from No. 120 to No. 1200, supplied by Buehler, Lake Bluff, IL, were used in coarse polishing. Fabric polishing pads and diamond polishing compounds with $6\ \mu\text{m}$ and $1\ \mu\text{m}$ diamond particles, which were also acquired from Buehler, were used in fine polishing. Chemicals including hydrochloric acid and phosphoric acid in ACS standard purity were used for removing the surface oxide layer and for electrochemical polishing of the copper crystal; the chemicals were purchased from Alfa Aesar, Ward Hill, MD. The nickel(II) sulfate and nickel anode for electroplating nickel on the surface of the copper crystal were obtained from Alfa Aesar. Solvents such as trichloroethylene and acetone supplied by Alfa Aesar were used for surface degreasing of the aluminum and copper single crystals. Epoxy resin and the curing agent trademarked as DEVCON, used for bonding the single crystals, were acquired from ITW Performance Polymers, Riviera Beach, FL.

Specimen preparation. The copper and aluminum crystals were cut into $20 \times 10 \times 10\ \text{mm}^3$ blocks using the EDM. The surface to be indented was polished using grit 120, 240, 320 and 400 silicon carbide sandpaper. The surface of the single crystals was then mechanically polished following the procedures: The grit 600 sandpaper was used to grind the surface under minimum pressure along one direction until any deep scratches from the previous cutting process became invisible. During this process, water was used as lubricant and coolant to prevent the surface from being overheated. The purpose of this polishing procedure is to remove the possible deep deformation layer from the previous grinding processes. The grinding/polishing along one direction can prevent the rounded corner formation along the edges of the polished surface. The polishing direction was changed by 90° to continue polishing using the same grit 600 sand paper, but even less pressure was applied until the scratches from the previous polishing procedure became invisible. Water was supplied continuously for cooling and smooth grinding. The purpose of this polishing procedure is to remove the possible residual stress layer generated from previous polishing. The specimen was washed with tap water to prevent any coarse abrasive going into the next polishing procedure.

Subsequent mechanical polishing was conducted using a grit 1200 sand paper to polish the surface under minimum pressure along one direction until any rough scratches became invisible, lubricated by water. After this procedure, the polish direction is changed by 90° to continue polishing using a grit 1200 sand paper repetitively until the scratches from the previous procedure became invisible. This repeated procedure could help to further reduce the thickness of the surface deformation layer formed in the previous polishing procedures. In addition, changing the polishing direction helps to remove any texture built-up due to polishing along a fixed direction. After each polishing step, the specimen was washed with tap water to prevent any coarse abrasive going into the next polishing procedure.

The following procedures were applied for polishing the surface of the single crystal with diamond compound and using lapping oil as lubricant. The specimen is first polished with $6\ \mu\text{m}$ diamond paste with minimum pressure until no deep scratches were shown. The specimens were then washed in soapy

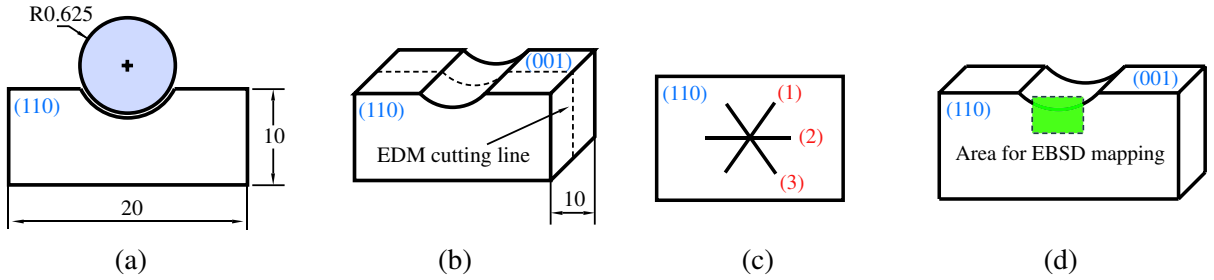


Figure 1. Schematic of cylindrical indentation experiment: (a) cylindrical indentation into Al single crystal, (b) EDM cutting for exposing the midsection of the crystal, (c) active slip systems in the single crystal under cylindrical indentation, (d) region for EBSD measurement.

water. Ultrasonic cleaning was applied to remove any attached particles from the specimen. After that, the specimens were polished using $1\ \mu\text{m}$ diamond paste with minimum pressure. The oil lubricant was applied frequently to keep a very smooth polishing condition. The polishing was done until no preferentially aligned scratches could be observed. Upon the completion of this polishing procedure, the surface of the aluminum and copper crystals show mirror luster with shiny reflection under light.

Indentation. Cylindrical indentation was performed on both aluminum and copper single crystal specimens. The indenter with a diameter of 0.625 mm as shown in Figure 1(a) was made of tungsten carbide bonded by a ferrous alloy. The indentation was under load control condition using a microindentation fixture. The loading direction was $[00\bar{1}]$ and the loading rate was approximately 5×10^{-4} mm/s. During indentation, the load and the indentation displacement data were recorded by a LabView 7.0 program and subsequently processed using a MATLAB code.

Electron microscopic examination. After the indentation, exposure of the mid-section of the single crystal specimens with plane strain deformation conditions was performed by an EDM cut as shown in Figure 1(b). After the cut, the aluminum single crystal was put into a compacted fixture and the indented area was painted with a layer of epoxy for protection. The indented (001) surface of the copper single crystal was electroplated with nickel to protect the indented region. The electrolyte and electroplating conditions for this work were adapted from the nickel-plating formula given in [Lowenheim 1978], and were as follows: nickel(II) sulfate hexahydrate, 250 g per liter; hydrochloric acid, 60 g (of 36.5% solution) per liter; temperature, 50°C ; current density, $500\ \text{A}/\text{m}^2$; time, 2 minutes. The anode was a nickel rod.

After electroplating, the copper crystal was carefully polished on the (110) plane following the same procedures described on page 559. After mechanical polish, the copper single crystal was first degreased using soapy water, then trichloroethylene, and finally acetone. The cleaned copper crystal was surface activated in a 10% HCl aqueous solution followed by rinsing in distilled water. After that, the surface of (110) of the single copper crystal was electropolished in an orthophosphoric acid solution. The composition of the solution and the detailed electropolishing procedures are given in [Ahmed et al. 1997] and [Morse 2002], respectively.

The procedures for aluminum crystal preparation are as follows: The exposed surface from the mid-section of the Al single crystal was protected by the epoxy coating and then coarse polished using 120,

240, 320 and 400 grit silicon carbide sandpaper with water as a lubricant. Further polishing using 600 and 1200 grit sandpaper was performed under minimum pressure. After that, fine polishing with 6-micron and 1-micron diamond compounds with oil as a lubricant was performed. After fine polishing with the diamond paste, electropolishing of the single crystal was conducted in an electrolyte containing 30% volume nitric acid (ACS purity, concentration 67%) and 70% volume methyl alcohol for 30 seconds at -20°C . Ultrasonic cleaning in solvents was performed after electropolishing. The specimens were then used for scanning electron microscopic (SEM) examination and electron backscatter diffraction (EBSD) measurements.

SEM examination and EBSD data acquisition at a specimen tilt angle of 70° were carried out on a JEOL 5600 scanning electron microscope. Figure 1(c) schematically shows the slip traces of the single crystals underneath the indented region. Such slip line features reveal three active slip systems as predicted by the slip line theory [Rice 1973]. Figure 1(d) is the schematic showing the region for the EBSD measurement. In Figure 2, left, the configuration of a specimen in the chamber of the microscope is shown. Figure 2, right, is an SEM micrograph taken from the indented aluminum single crystal. The area for the EBSD mapping is also shown. The typical step size for EBSD mapping is $3 \pm 0.1 \mu\text{m}$. The EBSD measurement data were analyzed using the HKL Channel 5 software. Backscatter electron (BSE) microscopic examination on selected region of copper single crystal specimens was performed at 0° specimen tilt angle using a Hitachi 4700 field emission electron microscope to determine the indentation displacement field, as will be discussed in more detail in Section 4.6.

3. Finite element simulations

Finite element analysis was performed using ABAQUS version 6.4-1. The user-material subroutine for single crystal plasticity developed by Huang [1991] and modified by Kysar [1997] was used in the implementation of finite element solutions for both aluminum and copper single crystals. The indentation contact, loading and unloading were set under displacement control to achieve better convergence than in the case of under load control. The critical shear stress for the aluminum single crystal, τ , was assumed to be 1.0 MPa. This value for copper single crystal was 58.0 MPa according to Huang [1991]. The element

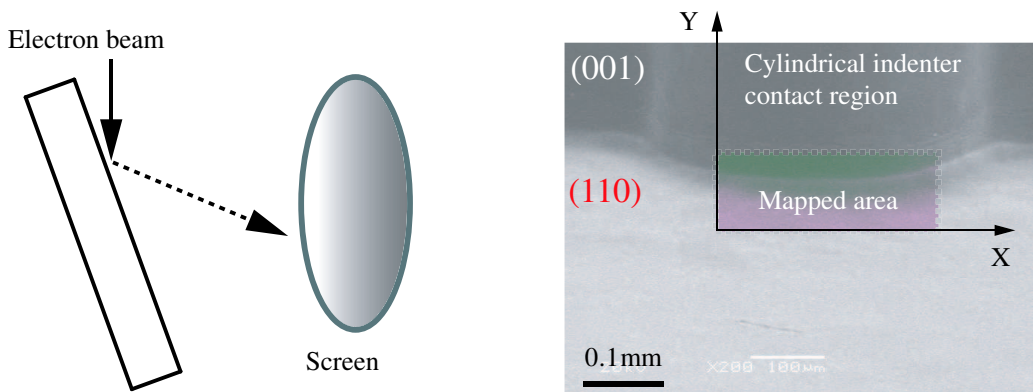


Figure 2. Illustrations of crystal lattice rotation measurement: specimen configuration and SEM image showing EBSD mapping region.

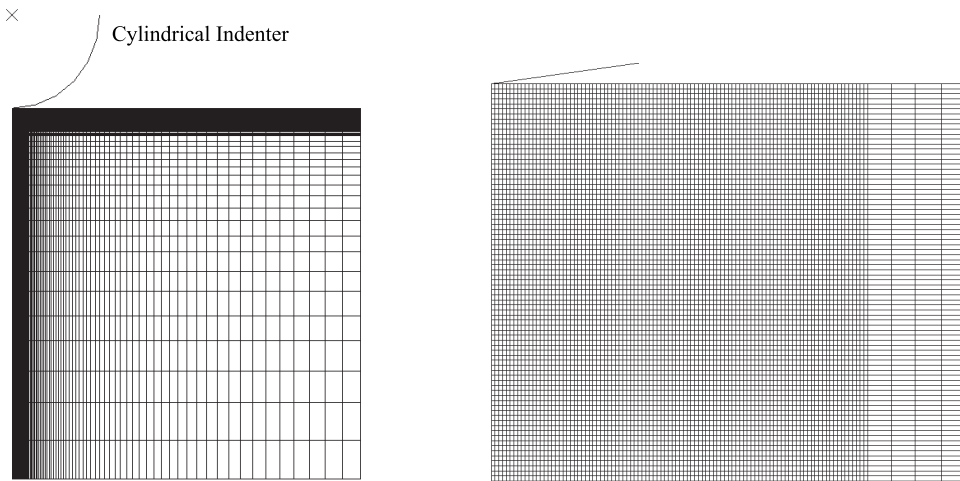


Figure 3. Schematic of the cylindrical indenter and the mesh for numerical analysis: global view (left) and detail near the tip of the indenter. (In the global view, the small mesh size collapses the lines into a band.)

used in the simulation was a plane strain reduced integration, hybrid element (CPE4RH) implemented in ABAQUS. The cylindrical indenter was defined as analytically rigid. Figure 3 shows the configuration of the cylindrical indenter and the mesh of the single crystals; note on the left-hand diagram that the density of the mesh increases drastically in the vicinity of the indenter tip. In the finite element analysis, the lattice rotation and displacements were solved incrementally by ABAQUS using a finite strain kinematic structure described by Huang [1991]. The constitutive properties were taken to be elastic-plastic. The power-law rate-dependent relationship was initially proposed by Hutchinson [1976], and described in detail in [Connolly and McHugh 1999; Huang 1991; Kysar 2001; Peirce et al. 1983; Savage et al. 2004]. The function related to the critical resolved shear stress of the k -th slip system was taken as a constant. The reference strain rate was $\dot{\gamma}_0 = 10^{-3} \text{ s}^{-1}$, and the rate sensitivity exponent $m = 50$. The Peirce–Asaro–Needleman hardening model [Peirce et al. 1983] was used in the simulation. The parameters related to the hardening were defined as in [Huang 1991]. During the simulation, the minimum iteration step used is 1×10^{-9} and the maximum step is 5×10^{-5} and approximately 20000 increments were performed in the loading simulation and unloading simulations, respectively.

4. Results and discussion

4.1. Load-displacement relation. Figure 4 shows the load-displacement curve for the aluminum single crystal, where the large hysteresis loop indicates the finite plastic deformation occurred under the cylindrical indentation. Since the deformation is dominated by plasticity, the elastic unloading region is fairly small (the right-hand part of Figure 4). The deformation behavior can be explained by the easy gliding (slip) property of an FCC single crystal under external loadings. In an FCC single crystal such as aluminum single crystal, there are 12 favorable slip systems. Under certain loading conditions, some of the slip systems become active. Thus, the plastic deformation dominated behavior is observed in

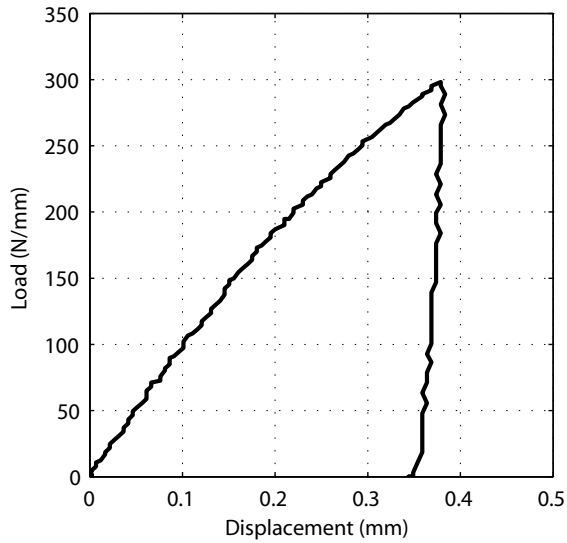


Figure 4. Load-displacement relation associated with cylindrical indentation.

Figure 4. Such plastic deformation-dominated behavior makes the experimental results of lattice rotation comparable with the slip line theory predictions, as will be discussed in detail in Section 4.4.

4.2. Crystal lattice rotation determined by EBSD measurement. Figure 5 is the out-of-plane crystal lattice rotation map for the FCC aluminum single crystal. In the indenter penetrated region, it can be seen that significant misorientation exists. This came from the severe plastic deformation of the indented plane: (001).

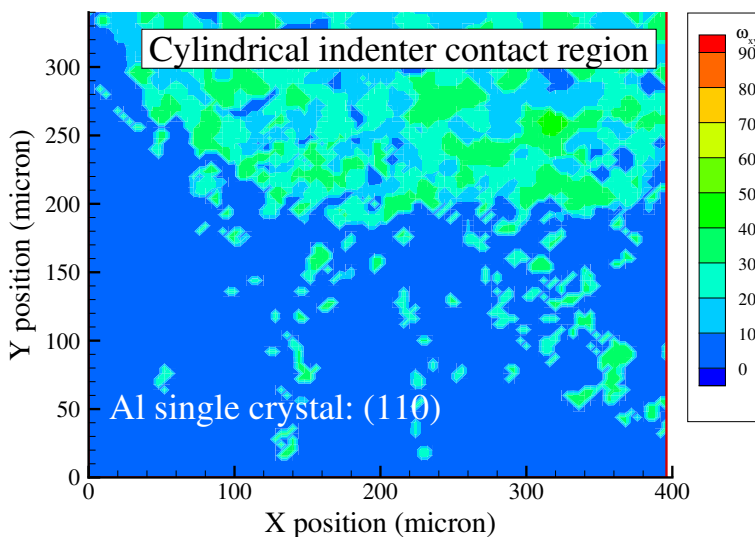


Figure 5. EBSD measurement results showing out-of-plane lattice rotation.

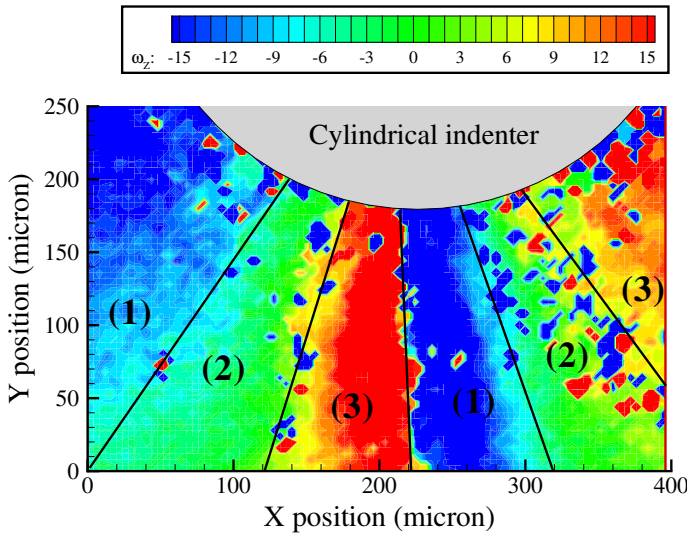


Figure 6. EBSD measurement results showing in-plane lattice rotation.

However, in the lower part of Figure 5, the original orientation of (110) plane was preserved; the out-of-plane rotation was essentially equal to zero in this region. This confirms that the midsection of the single crystal was under plane strain deformation conditions.

In Figure 6, the in-plane lattice rotation map is shown with three distinct sectors in a quadrant. The change in lattice rotation angle from one sector to another sector varies depending on the nature of the boundary. For example, from slip sector 1 to 2 and vice versa, and again from 2 to 3, the change in lattice rotation angle is about 15° . However, the change in lattice rotation is doubled when crossing the slip boundary between slip sector 1 and slip sector 3. That is why along the vertical sector boundary, as defined by Kysar et al. [2005], a “jump” in lattice rotation can be found, which was shown through experimental measurements in the region close to a microvoid within an FCC single crystal by Gan et al. [2006]. Such a sharp transition of lattice rotation along the slip boundary between the slip system 1 and slip system 3 is also revealed in this cylindrical indentation problem. The vertical line in the middle part of Figure 6 schematically shows the location at which sharp transition of crystal lattice rotation occurred. The lattice rotation results as shown in Figure 6 will be used to compare with that obtained from finite element analysis in Section 4.3.

The crystal lattice rotation results of the copper single crystal specimen are found to have the similar features as those for the aluminum single crystal. The out-of-plane rotation was essentially zero indicating that the plane strain deformation state held in the midsection of the copper single crystal. The three single slip systems are also revealed on the in-plane lattice rotation maps of the copper single crystal. Nevertheless, the in-plane lattice rotation map of the copper single crystal shows the change in rotation angle in the range of $\pm 20^\circ$ instead of $\pm 30^\circ$ for the aluminum single crystal as shown in Figure 6. Such a difference in the magnitude of lattice rotation angle may come from the difference in the critical shear strength of the two types of single crystals. The copper single crystal has a much higher critical shear strength than the aluminum single crystal. Thus the lattice rotation of the copper single crystal under the same indentation loading level is less than that of the aluminum single crystal. It is also noticed that the

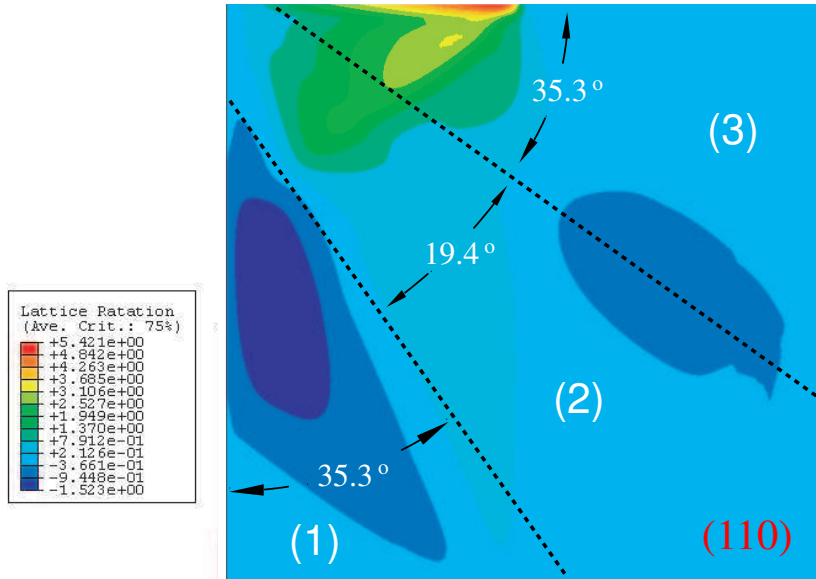


Figure 7. Numerical solution to crystal lattice rotation showing three deformation zones in a quadrant.

indenter penetration depth for the copper single crystal is much less than that for the aluminum single crystal.

4.3. Numerical solution to crystal lattice rotation. Figure 7 is the in-plane crystal lattice rotation map obtained from the finite element simulation. The unit of the lattice rotation angle is degree, which was calculated from one of the solution dependent variables (SDVs) in the ABAQUS finite element simulation program. In this case, SDV80 refers to the in-plane crystal lattice rotation. Some features of the in-plane lattice rotation from the simulation can be seen from this contour plot. In the left part of Figure 7, a region with negative lattice rotation can be found. This is due to the active dislocation movement of the slip system 1. The existence of such a negative lattice rotation region is also revealed by the experimental results as illustrated by Figure 6 in Section 4.2.

In the middle of Figure 7, the lattice rotation is close to zero. This is the region where slip system 2, with the slip direction parallel to $[\bar{1}10]$, is active. In the top-middle part of Figure 7 there is a positive lattice rotation region, caused by the active dislocation movement of slip system 3. All these features are comparable to the experimental results as shown in Figure 6. The in-plane lattice rotation predicted by the numerical simulation is in qualitative agreement with the EBSD experimental results in the sense that the three slip sectors in each quadrant are revealed. It is noted that the magnitude of the in-plane lattice rotation in the experiment is much larger than that in the simulation because a significantly larger strain was applied to the crystal specimen than in the simulation.

4.4. Correlation of the crystal lattice rotation features and the predictions by single crystal plasticity theory. The salient feature of the crystal lattice rotation as revealed by the experimental results in Section 4.2 and the numerical solutions in Section 4.3 can be explained by the single crystal plasticity theory [Schmid and Boas 1968]. In a single crystal, plastic deformation is anisotropic due to the motion of

dislocations within the crystal on discrete slip systems. Under external loading, whether a slip system is active or not is determined by Schmid’s law [1925] which states

$$n_i^{(k)} \sigma_{ij} s_j^{(k)} = \pm \tau^{(k)}, \tag{1}$$

where \mathbf{n} is the unit normal of the slip plane and \mathbf{s} is the unit vector parallel to the slip direction. σ_{ij} is the applied stress tensor and $\tau^{(k)}$ is the critical resolved shear stress of the slip system.

According to Rice [1987], in an FCC single crystal, certain slip systems may act cooperatively to allow plane strain deformation. In this case, the crystals were loaded along $[00\bar{1}]$ crystallographic direction, which causes the three slip systems active, as illustrated in Figure 8, left. There, $\mathbf{t}^{(k)}$ is the dislocation line vector of the k -th slip system with $k = 1, 2, 3$; $\mathbf{b}^{(k)}$ is the Burgers vector of the k -th slip system; $\mathbf{n}^{(k)}$ is the surface normal of the slip plane; and $\mathbf{s}^{(k)}$ is the slip direction vector. It is noted that slip system 1 corresponds to the effective slip system: $(\bar{1}\bar{1}1)[\bar{1}\bar{1}2]$, which is oriented at an in-plane angle of $\phi_1 = \tan^{-1}(\sqrt{2}) \approx 54.7^\circ$ counterclockwise relative to the $[\bar{1}\bar{1}0]$ direction. slip system 2 is parallel to the $[\bar{1}\bar{1}0]$ direction; the slip angle ϕ_2 is equal to 0. The notation slip system 3 refers to the effective slip system of $(\bar{1}\bar{1}1)[1\bar{1}2]$. It is oriented at an in-plane angle of $\phi_3 = -\tan^{-1}(\sqrt{2}) \approx -54.7^\circ$ relative to the $[\bar{1}\bar{1}0]$ direction.

Under plane strain conditions ($\sigma_{13} = \sigma_{23} = 0$), Equation (1) can be expanded as

$$(s_1 n_2 + s_2 n_1) \sigma_{12} + 2s_1 n_1 \frac{(\sigma_{11} - \sigma_{22})}{2} = \pm \tau. \tag{2}$$

If \mathbf{n} and \mathbf{s} in the crystal are projected onto the $x_1 - x_2$ plane of the crystal and rescaled as unit vectors \mathbf{N} and \mathbf{S} , respectively, Equation (2) can be rewritten as

$$(S_1 N_2 + S_2 N_1) \sigma_{12} + 2S_1 N_1 \frac{(\sigma_{11} - \sigma_{22})}{2} = \pm \beta_k \tau, \tag{3}$$

where β_k is a constant specific to the slip system under consideration such that (2) and (3) are equivalent

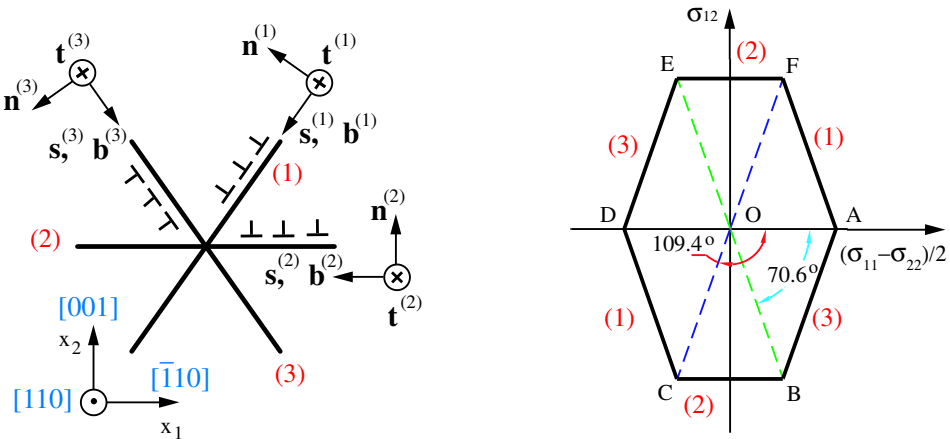


Figure 8. Schematic drawing for micromechanics analysis: configuration of the three active slip systems (left) and yield surface (right).

expressions. Rice [1987] noted that $\beta_1 = \beta_3 = 2/\sqrt{3}$ and that $\beta_2 = \sqrt{3}$ for face-centered cubic crystals. The sign for β_k depends on whether the slip system is activated in positive or negative sense.

Since $S_1 = \cos \phi_k$, $S_2 = \sin \phi_k$, $N_1 = -S_2$ and $N_2 = S_1$, Equation (3) can be simplified as

$$\sigma_{12} = \tan 2\phi_k \left(\frac{\sigma_{11} - \sigma_{22}}{2} \right) \pm \frac{\beta_k \tau}{\cos 2\phi_k}. \quad (4)$$

The above yield conditions can be plotted in stress space for each of the three effective slip systems. A hexagonal yield surface is obtained as shown in Figure 8, right. The positions of the vertices of the yield surface are as follows:

Vertex	A	B	C	D	E	F
$\frac{\sigma_{11} - \sigma_{22}}{2\tau}$	$\frac{\sqrt{6}}{2}$	$\frac{\sqrt{6}}{4}$	$-\frac{\sqrt{6}}{4}$	$-\frac{\sqrt{6}}{2}$	$-\frac{\sqrt{6}}{4}$	$\frac{\sqrt{6}}{4}$
σ_{12}	0	$-\sqrt{3}$	$-\sqrt{3}$	0	$\sqrt{3}$	$\sqrt{3}$

On the yield surface, the angle AOB equals to 70.6° . This defines the angular sector region in physical space. Since a θ degree angular region in physical space corresponds to a 2θ degree angular region in stress space, the region defined as slip system 3 takes 70.6° in the stress space as shown by the angle AOB in Figure 8, right. Thus, the slip system 3 can be determined in physical space as a circular region defined by $0 \leq \theta \leq 35.3^\circ$. Such a prediction to the existence of slip system 3 and its exact location in physical space from the single crystal plasticity theory can be confirmed by the experimental result of Figure 6. In the right-hand part of the figure, the positive in-plane crystal lattice rotation region (red-colored region) takes about a 35.3° angular region. In the right-hand part of Figure 7 (numerical solutions), a similar positive lattice rotation region can be seen. The angle COB shown in Figure 8, right, is the active region for slip system 2 in stress space. This region can be mapped into physical space as the angular region $35.3^\circ \leq \theta \leq 54.7^\circ$, which defines slip sector 2 in physical space. Similarly, the angular region $54.7^\circ \leq \theta \leq 90^\circ$ is determined as slip sector 3. The locations for slip sector 2 and slip sector 3 are also found in experiments. The experimental results are given in Figure 6. The numerical solutions as shown in Figure 7 reveals such single slip regions. Therefore, the results from experiments and simulations are in agreement with the analytical predictions.

It is also indicated that such a consistency is in a qualitative sense, because the single crystal plasticity theory assumes the crystal is rigid-ideally plastic. Also assumed here is that the yield surface as shown in Figure 8 holds at the incipient stage of yielding, where essentially no hardening behavior is considered. However, in the finite element simulation, a small strain hardening was imposed to ensure the numerical convergence. In the cylindrical indentation experiment, considerably large deformation was observed. Although the average behavior as revealed by the load-displacement relation in Figure 4 is elastic-plastic, localized strain hardening and strain gradient exist [Gan 2005]. That is why the regions related to the slip system 1 and slip system 2 shrinks in the EBSD map of Figure 6. The actual measured slip sectors from the active motion of the slip systems 1 and 2 are smaller than the analytical predictions due to the sink-in of the crystal under large scale deformation and associated strain hardening.

4.5. Numerical solutions to displacements. Finite element solutions for the 2-D displacement fields were obtained for FCC single crystals. The results from the simulations for aluminum single crystals are shown in Figure 9: horizontal displacement on the left and vertical displacement on the right. Qualitatively, both copper and aluminum have the same salient features in their indentation displacement fields. The deformed state of the crystal under the cylindrical indentation is shown in a magnified way and a factor of 10 was used to generate the plots in Figure 9. In the left half, the blue color stands for the zero displacement along x -axis, while the red colored region reveals the greatest displacement along positive x -axis. It can be seen that the region close to the indenter tip has zero horizontal movement, due to the symmetrical deformation of the crystals under indentation.

In Figure 9, right, yellow-green indicates zero displacement along the y -axis, blue indicates the negative vertical displacement and red positive vertical displacement. It is evident that materials adjacent to the indenter tip move downward, while the contact region far away from the indenter tip shows the pile-up feature of the single crystals as illustrated in the upper right part of the figure.

4.6. Comparison of displacement fields between numerical solutions and experimental results. The displacement fields associated with the cylindrical indentation into a copper single crystal were measured experimentally. As shown in Figure 10(a), on the (110) plane of a well-polished copper single crystal, regularly aligned patterns were generated by thin film technologies. First, a thin layer of photosensitive polymer (a positive photoresist, AZ5412) was spin coated on the (110) plane with a thickness of $0.6\ \mu\text{m}$. Next, a photoresist was baked on a hot plate at 105°C for 3 minutes. Subsequently, UV photolithography was employed over a mask for 15 seconds. The mask has regularly aligned $5\ \mu\text{m} \times 5\ \mu\text{m}$ chromium patterns with a separation distance of $12.5\ \mu\text{m}$. After UV exposure, a 1:4 diluted AZ400K aqueous solution was used as a developer to obtain the microscale arrays of pattern. The following step was to evaporate a gold thin film with a thickness of $0.02\ \mu\text{m}$ on the patterned polymeric layer. Since some part of the copper single crystal was covered by the polymer, while the other part was not covered by the polymer due to the photolithography process, gold coating only partially covered the (110) plane of the copper single crystal. By soaking the gold coated single crystal into acetone for 30 seconds, the polymer

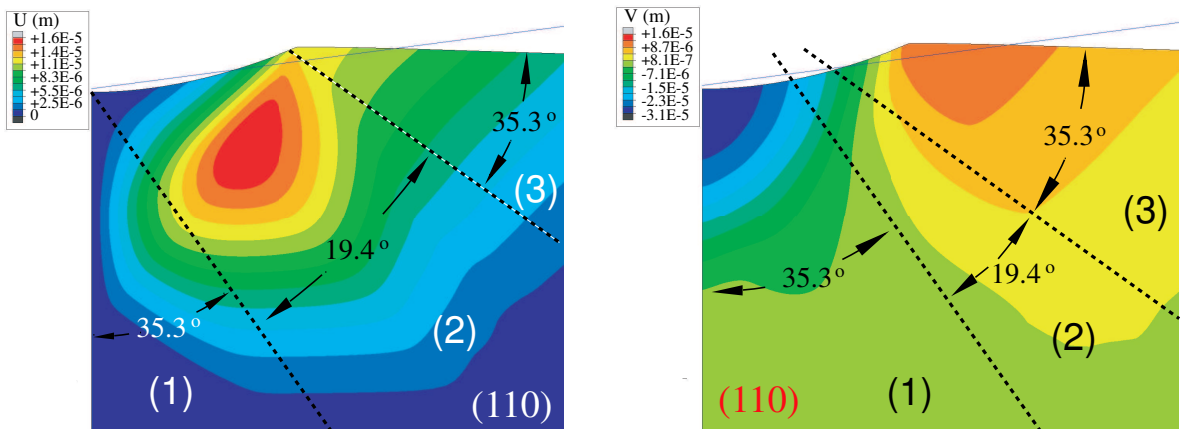


Figure 9. Horizontal (left) and vertical (right) components of numerical solution for the displacement field.

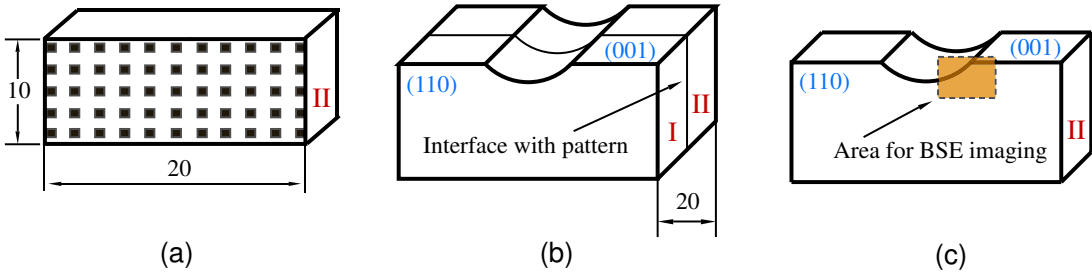


Figure 10. Experiments for measuring displacement field: (a) generating gold pattern on the (110) crystallographic plane of the single crystal, (b) cylindrical indentation into the two pieces of the single crystal, (c) imaging selected indentation area.

coating (photoresist) was peeled off from the surface of the single crystal. While in the part with gold directly being deposited on the single crystal, the gold coating was kept and the microscale square arrays formed, as schematically shown in Figure 10(a).

During cylindrical indentation, two pieces of single crystals with the same dimension were stacked together so that the two surfaces of the two pieces of crystals with microscale pattern were facing together as shown in Figure 10(b). This configuration kept the deformation state of the interface with microscale pattern under the plane strain conditions. After indentation, a selected region as shown in Figure 10(c) was examined using a Hitachi 4700 field emission electron microscope. A backscatter electron micrograph is shown in Figure 11.

The displacement fields as shown in Figure 11 are comparable to that in Figure 9. For example, the vertically aligned microscale patterns shift right with the increase of distance from the indenter tip.

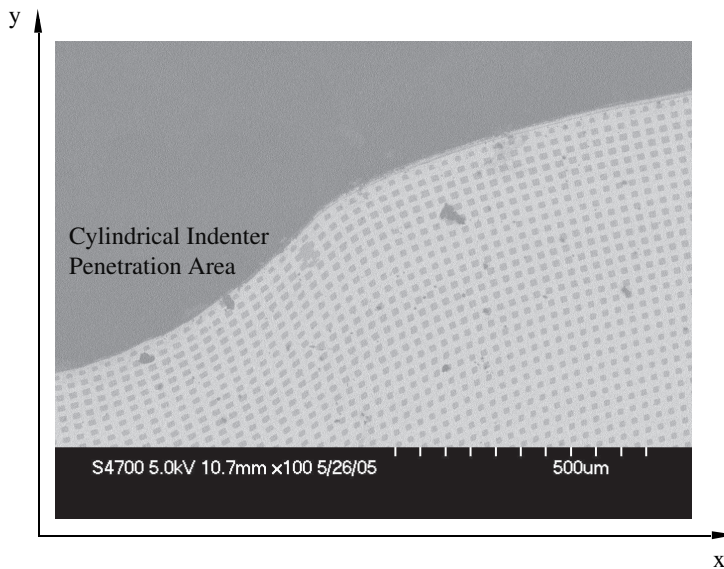


Figure 11. Backscatter scanning electron micrograph showing the plane strain deformation state.

This means that positive x -displacements exist, which is in agreement with the trends predicted by the finite element analysis results of Figure 9, left. In view of the vertical displacement, in the region underneath the cylindrical indenter, the experimental results are consistent with the numerical solutions. Negative vertical displacements are observed because of the sink-in of the single crystal under indentation loading along $-y$ direction. The difference in the vertical displacement results between the finite element simulations and the experiments is as follows: The numerical solutions predict a significant pile-up in the area underneath the right part of the indenter as shown by the positive displacement region in Figure 9, right. Nevertheless, the experimental results just show sink-in behavior, and no region of positive vertical displacement can be found in Figure 11.

Measurement of both the lattice rotation field and displacement field is essential for determining the stretching components and rotational components of a deformation gradient tensor. The results of the crystal lattice rotation in Figures 5 and 6 are helpful in determining the lattice curvature tensor as defined by Nye [1953]. From the lattice curvature tensor a correlation between the crystal lattice rotation tensor and the deformation gradient tensor can be established. For the displacement fields shown in Figure 11, they can be used to establish an explicit relationship with the stretching components of the deformation tensor. Therefore, the lattice rotation field along with the displacement field provides a significant amount of knowledge about the single crystals' plastic deformation states under cylindrical indentation.

5. Conclusions

Based on the studies of cylindrical indentation induced deformation fields in face-centered cubic aluminum and copper single crystals, the following concluding remarks can be made.

1. The indentation load-displacement relationship shows plastic deformation dominated behavior of the single crystals under cylindrical indentation.
2. Crystal lattice rotation field determined experimentally under plane strain conditions of the FCC single crystals reveals three slip sectors, which is in agreement with the qualitative features determined by finite element simulations. The lattice rotation features can also be correlated to the predictions to the deformation field based on the single crystal plasticity theory.
3. The displacement field determined by the numerical simulations also reveals the existence of different slip sector boundaries in the single crystals.
4. Comparison on the displacement field between the numerical solutions and the experimental results shows consistency in main features. The numerical solutions predict a significant pile-up in the area underneath the right part of the cylindrical indenter as shown by a positive vertical displacement region. However, the experimental results just show sink-in behavior; no region of positive vertical displacement was found.
5. Measurement of both lattice rotation field and displacement field is essential for determining the stretching components and rotational components of a deformation gradient tensor. The results of the crystal lattice rotation as obtained in this work is helpful to establish a correlation between the crystal lattice rotation tensor and the deformation gradient tensor. The displacement results can be used to determine the stretching components of the deformation tensor. Therefore, the lattice rotation field and the displacement field provide insight into the plastic deformation state of the FCC single crystals under cylindrical indentation.

References

- [Ahmed et al. 1997] J. Ahmed, A. J. Wilkinson, and S. G. Roberts, "Characterizing dislocation structures in bulk fatigued copper single crystals using electron channeling contrast imaging (ECCI)", *Philos. Mag. Lett.* **76**:4 (1997), 237–246.
- [Bansal et al. 2005] S. Bansal, E. Toimil-Molares, A. Saxena, and R. R. Tummala, "Nanoindentation of single crystal and polycrystalline copper nanowires", pp. 71–76 in *Proceedings of the 55th Electronic Components and Technology Conference* (Lake Buena Vista, FL), vol. 1, IEEE, Piscataway, NJ, 2005.
- [Chang and Sheu 1992] S. C. Chang and T. S. Sheu, "Determination of Al-Li crystal orientation by indentation", *Mater. Sci. Eng. A* **150**:1 (1992), L11–L12.
- [Connolly and McHugh 1999] P. Connolly and P. E. McHugh, "Fracture modelling of WC-Co hardmetals using crystal plasticity theory and the Gurson model", *Fatigue Fract. Eng. Mater. Struct.* **22**:1 (1999), 77–86.
- [Fujiwara and Otsuka 1999] M. Fujiwara and M. Otsuka, "Characterization of micro-indentation creep in β -Sn single crystals at elevated temperatures", *J. Jpn. Inst. Met.* **63** (1999), 760–769.
- [Gan 2005] Y. X. Gan, *High strain gradient plasticity of single crystals: theory, simulations and experiments*, Ph.D. thesis, Columbia University, New York, 2005.
- [Gan et al. 2006] Y. X. Gan, J. W. Kysar, and T. L. Morse, "Cylindrical void in a rigid-ideally plastic single crystal, II: Experiments and simulations", *Int. J. Plast.* **22**:1 (2006), 39–72.
- [Guillou et al. 1993] M.-O. Guillou, J. L. Henshall, and R. M. Hooper, "Indentation cyclic fatigue of single-crystal magnesium oxide", *J. Am. Ceram. Soc.* **76**:7 (1993), 1832–1836.
- [He et al. 2006] L. H. He, N. Fujisawa, and M. V. Swain, "Elastic modulus and stress-strain response of human enamel by nano-indentation", *Biomaterials* **27**:24 (2006), 4388–4398.
- [Huang 1991] Y. Huang, "A user-material subroutine incorporating single crystal plasticity in the ABAQUS finite element program", Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 1991. Mech report 178.
- [Hutchinson 1976] J. W. Hutchinson, "Bounds and self-consistent estimates for creep of polycrystalline materials", *Proc. R. Soc. Lond. A* **348**:1652 (1976), 101–127.
- [Kobayashi et al. 1990] S. Kobayashi, T. Harada, and S. Miura, "Mechanism of deformation in α Cu-Al single crystals by spherical indenter, I: Simple indentation", *J. Soc. Mater. Sci. (Japan)* **39** (1990), 126–131.
- [Kysar 1997] J. W. Kysar, "Addendum to 'A user-material subroutine incorporating single crystal plasticity in the ABAQUS finite element program: Mech report 178'", Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 1997.
- [Kysar 2001] J. W. Kysar, "Continuum simulations of directional dependence of crack growth along a copper/sapphire bicrystal interface, I: Experiments and crystal plasticity background", *J. Mech. Phys. Solids* **49**:5 (2001), 1099–1128.
- [Kysar et al. 2005] J. W. Kysar, Y. X. Gan, and G. Mendez-Arzuza, "Cylindrical void in a rigid-ideally plastic single crystal, I: Anisotropic slip line theory solution for face-centered cubic crystals", *Int. J. Plast.* **21**:8 (2005), 1481–1520.
- [Lowenheim 1978] A. F. Lowenheim, *Electroplating*, McGraw-Hill, New York, 1978.
- [Morse 2002] T. L. Morse, "Measuring lattice rotation induced by plastic deformation in single crystals", Columbia University, New York, 2002. MRSEC REU student research report.
- [Nye 1953] J. F. Nye, "Some geometrical relations in dislocation solids", *Acta Metall.* **1**:2 (1953), 153–162.
- [Peirce et al. 1983] D. Peirce, R. J. Asaro, and A. Needleman, "Material rate dependence and localised deformation in crystalline solids", *Acta Metall.* **31**:12 (1983), 1951–1976.
- [Rice 1973] J. R. Rice, "Plane strain slip line theory for anisotropic rigid-plastic materials", *J. Mech. Phys. Solids* **21**:2 (1973), 63–74.
- [Rice 1987] J. R. Rice, "Tensile crack tip fields in elastic-ideally plastic crystals", *Mech. Mater.* **6**:4 (1987), 317–335.
- [Sastry 2005a] D. H. Sastry, "Impression creep technique: an overview", *Mater. Sci. Eng. A* **409**:1–2 (2005), 67–75.
- [Sastry 2005b] D. H. Sastry, "On some investigations with the impression creep technique", Talk presented at 134th Annual Meeting and Exhibition of the TMS (San Francisco), 2005.

- [Savage et al. 2004] P. Savage, B. P. O. Donnell, P. E. McHugh, B. P. Murphy, and D. F. Quinn, "Coronary stent strut size dependent stress-strain response investigated using micromechanical finite element models", *Ann. Biomed. Eng.* **32**:2 (2004), 202–211.
- [Schmid 1925] E. Schmid, pp. 342 in *Proceedings of the First International Congress of Applied Mechanics* (Delft, 1924), edited by C. B. Biezeno and J. M. Burgers, Waltman, Delft, 1925.
- [Schmid and Boas 1968] E. Schmid and W. Boas, *Plasticity of crystals with special reference to metals*, Chapman and Hall, London, 1968.
- [Seo et al. 2003] M. Seo, M. Chiba, and Y. Kurata, "Nano-indentation to the passive metal surfaces in solution", *Corros. Eng.* **52** (2003), 5–11.
- [Stevenson et al. 2001] M. E. Stevenson, M. L. Weaver, and R. C. Bradt, "Microhardness anisotropy and the indentation size effect in single crystal NiAl", in *High-temperature ordered intermetallic alloys, IX: 2000 Fall MRS Symposium N* (Boston, 2000), edited by J. H. Schneibel et al., Materials Research Society Symposia Proceedings **646**, Materials Research Society, Warrendale, PA, 2001. Paper N6.4.
- [Tanaka et al. 2003] M. Tanaka, K. Higashida, H. Nakashima, H. Takagi, and M. Fujiwara, "Fracture toughness evaluated by indentation methods and its relation to surface energy in silicon single crystals", *Mater. Trans. (JIM)* **44**:4 (2003), 681–684.
- [Wen et al. 2006] S. P. Wen, F. Zeng, Y. Gao, and F. Pan, "Indentation creep behavior of nano-scale Ag/Co multilayers", *Scr. Mater.* **55**:2 (2006), 187–190.
- [Yamada and Ikeda 1975] A. Yamada and M. Ikeda, "Misorientation of crystal sub-grains at the indentation on Cu single crystal", *Bull. Jpn. Soc. Precis. Eng.* **9**:22 (1975), 53–54.
- [Zhang et al. 2004] Z. F. Zhang, Z. M. Sun, H. Zhang, and H. Hashimoto, "Micron-scale deformation and damage mechanisms of Ti_3SiC_2 crystals induced by indentation", *Adv. Eng. Mater.* **6**:12 (2004), 980–983.
- [Zhu et al. 2004] T. Zhu, J. Li, K. J. Van Vliet, S. Ogata, S. Yip, and S. Suresh, "Predictive modeling of nanoindentation-induced homogeneous dislocation nucleation in copper", *J. Mech. Phys. Solids* **52**:3 (2004), 691–724.

Received 24 Jul 2006. Accepted 9 Oct 2006.

YONG XUE GAN: gan@cooper.edu

Department of Mechanical Engineering, Albert Nerken School of Engineering, The Cooper Union for the Advancement of Science and Art, 51 Astor Place, New York, NY 10003, United States

and

Columbia Nanomechanics Research Center, Department of Civil Engineering and Mechanics, 622 Mudd Building, Columbia University, New York, NY 10027, United States

YUKI SAITO: ys2121@columbia.edu

Columbia Nanomechanics Research Center, Department of Civil Engineering and Mechanics, 622 Mudd Building, Columbia University, New York, NY 10027, United States

XI CHEN: xichen@civil.columbia.edu

Department of Civil Engineering and Engineering Mechanics, Fu Foundation School of Engineering and Applied Science, Columbia University, 500 West 120th Street, New York, NY 10027, United States

ON THE NUMERICS AND CORRELATION OF SCRATCH TESTING

FREDRIK WREDENBERG AND PER-LENNART LARSSON

A numerical strategy based on the finite element method and intended for an accurate analysis of the scratch test is presented. For simplicity, but not out of necessity, the material was described by classical von Mises elastoplasticity utilizing large deformation theory. Based on this strategy numerous results are presented and correlation of scratch properties is discussed within the framework of theory for indentation testing. Furthermore, the existence of a representative plastic strain in the spirit of Tabor was studied. The investigation also concerns details regarding frictional effects, normal vs. tangential scratch hardness, similarities and differences between indentation and scratch test characteristics as well as details regarding the behavior of local field variables at scratching.

1. Introduction

Indentation testing and scratch testing show many similar features. However, due to its relative simplicity, indentation testing has been used much more frequently when, for example, material characterization is at issue. A further reason for the popularity of indentation testing is that such tests are more easily analyzed than scratch tests, and consequently a great deal of knowledge has been gained over the years regarding the mechanical behavior at indentation. Most often, at least when modern experimental devices such as the nanoindenter (or other types of instrumented indentation devices) are at issue, sharp indenters are used for practical reasons. For such indenters semiempirical relations for material characterization were derived and used by the late 1940's and early 1950's, in particular for metals and alloys [Tabor 1951]. In short, from comprehensive experimental investigations [Tabor 1951] derives a relation

$$H = C\sigma(\epsilon_{\text{repr}}), \quad (1)$$

between the indentation hardness H , here defined as the mean contact pressure at indentation, and the material yield stress at a representative value of the accumulated (effective) plastic strain ϵ_{repr} . Furthermore in Equation (1), C is a constant that only depends on the geometry of the indenter. For a Vickers indenter, Tabor [1951] determines the values $C \approx 3$ and $\epsilon_{\text{repr}} \approx 8\%$ while Atkins and Tabor [1965] find $C \approx 2.54$ and $\epsilon_{\text{repr}} \approx 11\%$ for a conical indenter with an angle of $\beta = 22^\circ$ between the indenter and the undeformed surface, as shown in Figure 1. Based partly on the above discussed results, further progress is achieved by Johnson [1970; 1985], who shows from theoretical considerations that indentation testing on different materials can be well correlated by using a parameter

$$\Lambda = \frac{E \tan \beta}{(1 - \nu^2)\sigma_{\text{repr}}}, \quad (2)$$

Keywords: scratch test, hardness, friction, finite elements, contact.

The authors acknowledge the support through grant 621-2005-5803 from the Swedish Research Council.

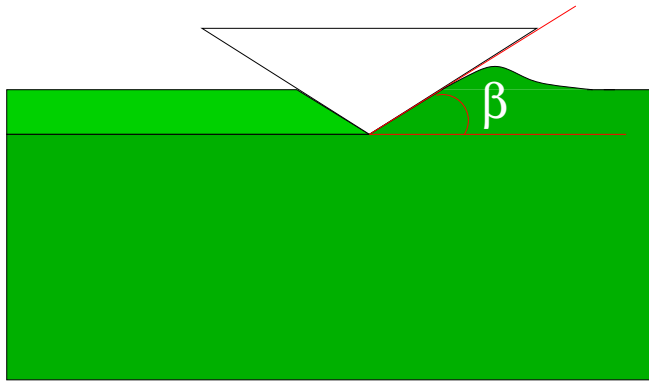


Figure 1. The angle β of a conical indenter/stylus.

where E is Young's modulus and ν is Poisson's ratio. Johnson [1970; 1985] also suggests that indentation properties for various materials will fall into one of three levels as shown in Figure 2. In level I, $\Lambda < 3$, very little plastic deformation occurs during the indentation test, and all global properties can be derived from an elastic analysis. In level II, $3 < \Lambda < 40$, an increasing amount of plastic deformation is present, and both the elastic and the plastic properties of the material will influence the outcome of a hardness test, according to Johnson [1970; 1985]. Based on the fact that the stress field just beneath the indenter in such a situation is almost hydrostatic, the process is very similar to the case of expansion of a spherical cavity in a large solid due to an internal pressure, and the formula

$$H = \frac{2}{3} \sigma_{\text{repr}} \left(1 + \ln \frac{E \tan \beta}{3(1 - \nu^2) \sigma_{\text{repr}}} \right) \quad (3)$$

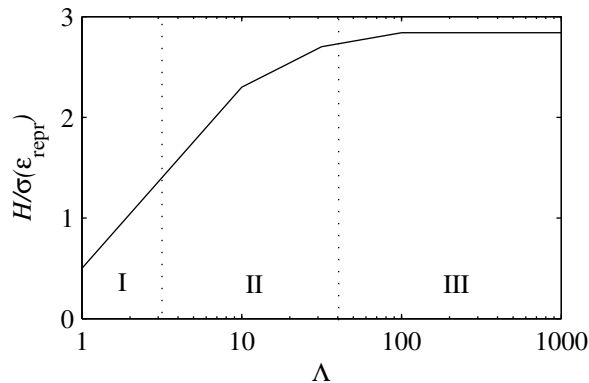


Figure 2. Sketch of the characteristic behavior of indentation hardness [Johnson 1985]. The indentation hardness H divided by the stress $\sigma(\epsilon_{\text{repr}})$ is plotted against the non-dimensional strain parameter Λ .

is derived for the hardness. Finally, in level III, $\Lambda > 40$, plastic deformation is present all over the contact area, and elasticity no longer influences the hardness value for the material. This is also the region pertinent to most standard engineering materials, such as steel and many aluminum and copper alloys, and it is also the region where Equation (1) applies.

The accuracy of the results above have been extensively tested (and improved) during the last ten to fifteen years by computational simulations of the indentation of materials with tailored constitutive properties that take advantage of modern computers and advanced numerical methods [Laursen and Simo 1992; Giannakopoulos et al. 1994; Larsson et al. 1996; Larsson 2001; Mata 2004]. In short, even though significant improvements have resulted from such numerical investigations it can be concluded that the validity of the overall findings by Tabor [1951] and Johnson [1970; 1985] have survived these scrutinizing studies. This is particularly so when it comes to such features as the existence of a representative strain at indentation, and the usefulness of the parameter in Equation (2) when it comes to correlation of indentation results [Dao et al. 2001; Larsson 2001].

Nowadays, scratch testing is also a well established technique for hardness testing. Historically, one of the earliest efforts was Mohs' hardness scale (1824), which is based on the fact that a harder material will leave a visible scratch on a softer material if rubbed against it. The scratch test may also be used for tribological testing, although it is being questioned for use on hard coatings [Bull 1999]. Throughout the coating industry, the scratch test is used for adhesion testing of coatings. The adhesion test usually consists of a loaded diamond tip that is drawn across a surface under continuously (or stepwise) increasing load. At some load a well-defined failure occurs and the critical load is found. This technique is used for the ranking of coating adhesion.

As indicated above, the fundamental knowledge about mechanical behavior at scratching is not nearly as developed as for indentation testing. It should be mentioned, however, that early mechanical analyses concerning different aspects of scratching of metals exist and have been presented by Goddard and Wilman [1962], Childs [1970], Vathaire et al. [1981], and Gilormini and Felder [1983]. Furthermore, as the scratch test is often used to determine characteristics of polymeric coatings, this issue is also under investigation [Briscoe et al. 1996; Gauthier et al. 2001]. It goes almost without saying, however, that due to the complexity of the boundary value problem, high accuracy of results at scratch analysis can only be obtained by relying on numerical methods, preferably the finite element method (FEM). In recent years quite a few such analyses have been presented. In this context, Bucaille et al. [2001] analyze cone scratching of ideally plastic materials, and their results also indicate that the parameter Λ as given in Figure 2 can indeed be used for correlating scratch parameters as well as indentation parameters. Furthermore, Bucaille et al. [2001] find, as could be expected, that the strain levels at scratching are much higher than at indentation, but as only ideally-plastic material behavior is at issue, this finding is not explored in connection with the concept of representative strains. Further progress based on FEM analyses of the scratch test is also achieved by Bucaille et al. [2004] for polymeric materials modeled using standard elastoplasticity, corresponding to stage II rheology Figure 2.

Although beyond the scope of this work, investigations have also been performed for severe scratching, that is, fracture, delamination, and similar phenomena during scratching, by Holmberg et al. [2003], Subhash and Zhang [2002], Malzbender and de With [2001], and Thouless [1998]. Bull [1991] and others find that a coating may fail in many modes including tensile cracking, buckling and spalling. To account for the cracking behavior, Subhash and Zhang [2002] propose an *elastic-plastic-cracking* model

both for indentation and scratching. In this model the material behavior is represented by tensile cracking and compressive yielding, where a material point is assumed to fail in tension and be deleted from the model if the crack opening displacement exceeds a critical value.

From the discussion above it should be clear that the mechanical behavior at scratching is very complicated. But, based on using the finite element method, scattered progress in the understanding of different features has been gained in recent years. It remains, however, to achieve a more complete understanding of the behavior of different local and global scratch parameters, including the effect from strain-hardening, friction, and the Johnson parameter Λ . This will be attempted in the present investigation, which will also include details regarding normal versus tangential scratch hardness, similarities and differences between indentation and scratch test characteristics, as well as details regarding the behavior of local field variables at scratching. In doing so, a numerical strategy based on the finite element method, and in particular the commercial FEM package [ABAQUS 2004], was relied upon to yield results of sufficient accuracy. For clarity and convenience, but not of necessity, the analysis was restricted to cone scratching of classical elastoplastic (von Mises) materials, since in such a case no characteristic length was introduced into the problem. Scratching of thin films was not analyzed in this initial work, but will be the subject of forthcoming analyses using the present numerical scheme.

2. Theoretical background

The present analysis of scratching using a sharp conical stylus concerns a problem where quasistatic and steady-state conditions prevail. Such a situation is of substantial practical importance and is relatively easily achieved, at least during a scratch test under controlled experimental conditions. It is important to emphasize that the interpretation of the experimental scratch results then becomes much simpler.

Assuming that quasistatic and steady-state conditions prevail, as in the case of normal indentation, the problem is self-similar with no characteristic length present. Consequently, the normal hardness

$$H_{\text{norm}} = F_{\text{norm}}/A_{\text{norm}}, \quad (4)$$

and the tangential hardness

$$H_{\text{tan}} = F_{\text{tan}}/A_{\text{tan}}, \quad (5)$$

as well as a ratio h/\sqrt{A} (where h is the scratch depth), will be constant during the loading sequence of a scratch test, and stresses and strains will be functions solely of material properties and the dimensionless variable x_i/\sqrt{A} (x_i being Cartesian coordinates, as shown in Figure 3). In this context, \sqrt{A} should be interpreted as a representative contact length and the indices *norm* and *tan* represent the normal and tangential components of the scratch quantities. Clearly, the fact that the normal and tangential scratch hardness are constant during a cone scratch test is valid for classical elastoplastic material behavior, as assumed here, but it fails for strain gradient material behavior, since a characteristic length is present in the constitutive equation [Fleck and Hutchinson 1993].

Due to the characteristics of the scratch test, at least as formulated presently, a possible route of attack would be to take advantage of the prevailing steady-state conditions. Accordingly, a standard steady-state transformation converting time derivatives into spatial derivatives, as is often used to analyze crack problems, would transform the problem to a stationary one, which is certainly more attractive when a numerical analysis is at issue. However, in the present study the numerical results clearly suggest that

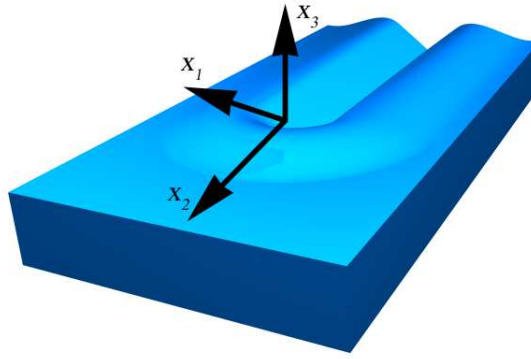


Figure 3. Cartesian coordinates x_i (following the tip of the stylus). Scratching is performed in the x_2 direction.

large deformation theory must be relied upon, and this is also the case for the corresponding normal indentation problem [Giannakopoulos et al. 1994; Larsson et al. 1996]. In addition, the strain levels at scratching are even higher, as shown by, for example, [Bucaille et al. 2001], as well as the present investigation. In such a case some of the boundary conditions resulting from a steady-state transformation become very complicated. Furthermore, the present numerical scheme will in the future be used to solve problems including film/substrate systems and in that case, steady-state conditions no longer apply. Accordingly, in the present study we relied upon a more straightforward FEM strategy, as will be discussed below.

As the constitutive specification, the incremental, rate-independent Prandtl–Reuss equation for classical von Mises plasticity with isotropic hardening is

$$\hat{\tau}_{ij} = \frac{E}{(1+\nu)} \left(\delta_{ik}\delta_{jl} + \frac{\nu}{(1-2\nu)}\delta_{ij}\delta_{kl} - \frac{3\tau'_{ij}\tau'_{kl} - (E/(1+\nu))}{2\tau_e^2\left(\frac{2}{3}K + \frac{E}{(1+\nu)}\right)} \right) D_{kl}, \quad (6)$$

in a large strain formulation. In Equation (6), δ_{ij} is the Kronecker identity tensor, D_{ij} is the rate of deformation, and $\hat{\tau}_{ij}$ is the Jaumann rate of the Kirchhoff stress τ_{ij} . The Kirchhoff stress is related to the Cauchy stress σ_{ij} as $\tau_{ij} = J\sigma_{ij}$, where J is the ratio of volume in the current state to volume in the previous state. Furthermore, τ_e and $\hat{\tau}_{ij}$ are the von Mises effective stress and deviatoric stress, respectively. Finally, K is the instantaneous slope of the uniaxial compressive Kirchhoff stress. Note that Equation (6) is only valid at plastic loading when $\tau_e = \tau(\epsilon_p)$, the initial yield stress being given by $\tau_Y = \tau(0)$. At elastic loading or unloading, a hypoelastic formulation of Hooke's law, pertinent to the first part of Equation (6), was relied upon. Obviously, within the present setting, kinematic hardening effects were not included in the analysis. Such effects could certainly have influenced the outcome of scratch test but would also have increased the number of required numerical computations substantially (due to an increased number of constitutive parameters) and would have made a straightforward interpretation of the results more difficult. For this reason, it was thought advisable as a first attempt to restrict the analysis to classical von Mises plasticity with isotropic hardening, especially since the loading part of a scratch test was of primary interest.

To summarize the governing equations, D_{ij} is connected with the material velocity \dot{u}_i , as

$$D_{ij} = \left(\frac{\partial \dot{u}_i}{\partial x_j} + \frac{\partial \dot{u}_j}{\partial x_i} \right), \quad (7)$$

while throughout the analysis, equilibrium equations in absence of body forces,

$$\frac{\partial \sigma_{ij}}{\partial x_j} = 0 \quad (8)$$

must be satisfied. Regarding boundary conditions, the surface outside the contact area was assumed to be traction free and within the area of contact unilateral kinematic constraints, given by the shape of the conical indenter/stylus depicted in Figure 1, and these assumptions had to be accounted for.

3. Numerical analysis

Scratching was simulated using the commercial FE-program [ABAQUS 2004] (see Figure 4 for the FE mesh). In the simulation a rigid conical stylus with attack angle $\beta = 22^\circ$ was first pressed normally into the mesh and then dragged tangentially across the surface, until steady state conditions were achieved. During the scratching process the stylus was held at a constant depth. The contact area and the reaction forces were recorded during the process. As specified above, the material was described by classical von Mises elasto-plasticity with isotropic hardening according to

$$\sigma(\epsilon_p) = \sigma_Y + \sigma_0 \epsilon_p^n, \quad (9)$$

where $\sigma(\epsilon_p)$ is the flow stress, σ_Y the initial yield stress, ϵ_p the effective plastic strain, and n the hardening exponent. The strength parameter σ_0 was adjusted so that the stress at $\epsilon_p = 8\%$ was twice that of the initial yield stress. This choice is not of importance when the numerical results are evaluated; rather, it was made by recalling the value of representative stress suggested by Tabor [1951], and it was convenient for correlating scratch/indentation results in the spirit of Johnson [1970; 1985]. As discussed above, it was found during simulations that the model experienced extremely large rotations/strains in the vicinity of the stylus. Thus, the use of large strain theory was deemed necessary.

Simulations for small values of Λ were performed with implicit time stepping and full integration elements. For larger Λ , explicit time stepping and reduced integration elements needed to be used. Full integration elements were used for small Λ to avoid the hour-glassing effect. It was also found necessary to use adaptive meshing as Λ gets larger, due to extreme element distortion. To decrease CPU time, the mesh was divided in half along the scratch, and symmetry conditions were applied. The mesh was constructed from 30284 eight node linear brick elements. These elements were chosen since they show a faster convergence with respect to mesh refinement than tetrahedral elements, and do not have the inherent contact problems of quadratic elements [ABAQUS 2004]. During the simulation some 40 elements were in contact with the stylus at any given time. The mesh domain size was $10 \times 10 \times 22.5 \text{ mm}^3$ and the scratch depth was 0.3 mm.

In Table 1 the convergence of the scratch hardness with respect to various parameters can be seen. Mass scaling was introduced to increase the time step in the explicit analysis and thus decrease the computational time. A mass scaling of 1000 was chosen for the analysis. The remeshing frequency was kept as low as possible, so as not to lose too much precision but still maintain the undistorted element

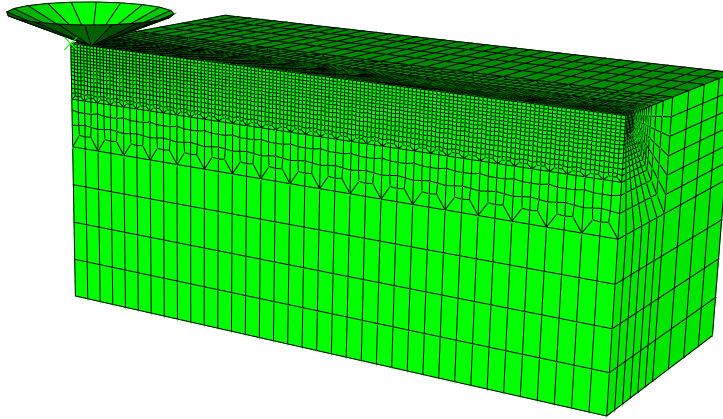


Figure 4. Finite elements mesh. Linear 8-node brick elements. Symmetry conditions apply to the surface in view. The stylus is also shown.

shape, since every remapping of the mesh causes some diffusion of the solution [ABAQUS 2004]. In Table 1 it can be seen that remeshing every 500th or 5000th time step made very little difference. The scratch depth was varied to ascertain that enough elements were in contact. Table 1 indicates that ≈ 40 elements in contact were sufficient, equaling a contact radius of approximately 1 mm. With the numerical precision secured, a parametric study was then performed for different values of Λ , n , and the interfacial coefficient of friction μ_i .

Furthermore, to simulate indentation, a finite element model developed by Larsson [2001] was used. The simulated indentation hardness, and also other indentation quantities, could then be related to the corresponding simulated scratch results.

4. Results and discussion

In this section the results from the numerical simulations are presented, and the results pertinent to the most important features related to scratching will be discussed in some detail. It should be stressed that most of the results below are pertinent to frictionless contact, but discussions about the influence

Parameter	Used	$H_{\text{norm}}/\text{MPa}$	Reference	$H_{\text{norm}}/\text{MPa}$
Mass scaling	1000	476	10000	480
Scratch depth (elements in contact)	0.3 (39)	476	0.4 (76)	472
Remeshing freq.	2e-3	476	2e-4	475

Table 1. Convergence analysis of scratch normal hardness H_{norm} at scratching, for elastic perfectly plastic material, $\Lambda = 100$.

of interfacial friction are included in each particular subchapter. Furthermore, if not stated otherwise, in most of the figures, when nondimensionalized variables are introduced, a representative stress σ ($\epsilon_p = 8\%$) was used, following Tabor [1951].

In the subchapters below, the following aspects concerning scratching and scratch testing are discussed: Section 4.1, the contact area at scratching, Section 4.2, scratch normal and tangential hardness and the concept of a representative strain, Section 4.3, the apparent coefficient of friction, and Section 4.4, scratch testing versus indentation testing. It should be mentioned that in the last subchapters the behavior of important field variables will also be discussed, and in particular those related to cracking.

4.1. The contact area at scratching. The contact area at indentation (and at scratching) is a very important parameter for material characterization, for one. Accordingly, it seems advisable to discuss this feature first of all in the context of analytical and numerical estimates, and in particular with the correlation of elastoplastic material parameters in mind.

An estimate of the normally projected contact area (A_{norm}) for large Λ is given by Goddard and Wilman [1962] as

$$A_{\text{norm}} = (w/2)^2 \frac{\pi}{2}, \quad (10)$$

where w is the scratch width (distance between the top levels of the residual groove). The tangentially projected contact area (A_{tan}) can be found through a simple geometric consideration as

$$A_{\text{tan}} = \frac{w^2}{4} \tan \beta. \quad (11)$$

These expressions are based on solely geometrical considerations, assuming that the contact area is semiconical.

In Figure 5 the actual contact area can be seen during simulation of scratching of an elastic perfectly plastic material. The corresponding ratio of the actual normally projected area, A_{real} , to the contact area estimated by Equation (10), is listed in Table 2. It is clear from Table 2 that the estimate Equation (10) worked well for large Λ materials, despite the fact that the contact areas shown in Figure 5 are not perfectly semiconical. This was indeed the case for both elastic perfectly plastic materials and strain hardening materials at stage III rheology. Not so, however, for smaller Λ , since contact then also occurred on the rear face of the stylus.

Λ ($\epsilon_{\text{repr}} = 8\%$)	10	100	1000
$n = 0$	1.39	1.10	1.10
$n = 0.1$	1.33	1.08	1.12
$n = 0.17$	1.20	1.00	1.09
$n = 0.33$	1.16	0.91	1.06
$n = 1$	1.43	1.05	0.91

Table 2. Ratio of simulated actual normal contact area to estimated normal contact area by Equation (10) for different Λ , for elastic plastic material with frictionless contact. Λ ($\epsilon_{\text{repr}} = 8\%$) = 1 yielded no measurable groove width, hence the estimate was not applicable.

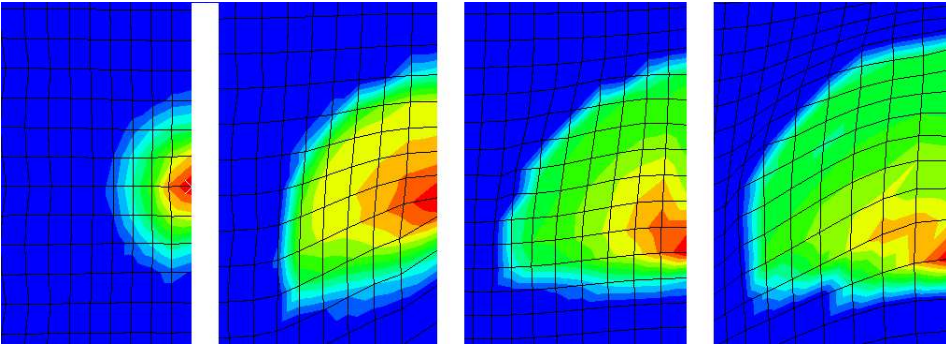


Figure 5. Contact pressure distribution at $\Lambda = 1$, $\Lambda = 10$, $\Lambda = 100$ and $\Lambda = 1000$, for elastic perfectly plastic material. Only half of the contact region is visible in the figure. The scratch direction is upwards.

Furthermore, it should be emphasized here that an estimate of the contact area based on purely geometric considerations, $A_{\text{nom}} = 1/2\pi h^2 \cot^2 \beta$, assuming a nominal semicircular normally projected contact area, is definitely not generally applicable during scratching, or during standard normal indentation testing (see Figure 6) [Larsson 2001]. It is of fundamental importance, as proven by corresponding analyses for indentation, to account correctly for piling up and sinking of material at the contact boundary when constitutive specifications are at issue. However, in this context, elastic effects are very hard to account for, and it was therefore thought advisable to present only results where rigid plasticity is dominant, as is shown in Figure 6. This figure also shows corresponding indentation results, and clearly the nominal area is far from a good approximation of the actual contact area during scratching even for predominantly plastic deformation.

4.2. Scratch normal and tangential hardness and the concept of a representative strain. In the following section the behavior of hardness values are considered. This concerns both the overall mechanical behavior and its correlation at material characterization.

In Figure 7 the scratch normal hardness is divided by the yield stress at a plastic strain level of 8% and plotted against Johnson's parameter Λ ($\epsilon_{\text{repr}} = 8\%$). It is obvious from Figure 7 that 8% was not a representative level of strain, since the curves do not coincide. Instead a different level of strain was sought. By minimizing the deviation of the curves in Figure 7 for $\Lambda(\epsilon_{\text{repr}} = 8\%) = 10, 100,$ and 1000 by means of the least-squares method, a representative level of plastic strain was found. The new representative level of plastic strain was found to be approximately 35%. For $\epsilon_{\text{repr}} = 35\%$ the stress $H/\sigma(\epsilon_{\text{repr}})$ was roughly constant for large Λ materials regardless of the strain hardening exponent n (see Figure 8). Of course, this is not an obvious choice for a representative level, since, for example, Wredenberg and Larsson [2005] showed that $\epsilon_{\text{repr}} = 57\%$ could also in some cases be used (except where $n = 1$), and in addition a representative strain of 35% and 57% gives $C \approx 2.5$ and $C \approx 2.4$, respectively (see Equation (1)). It is interesting to note that in previous studies of sharp indentation problems, representative strain levels close to 35% are suggested [Larsson et al. 1996; Larsson 2001], but this corresponds to a situation where a two-parameter model (with two representative strain levels)

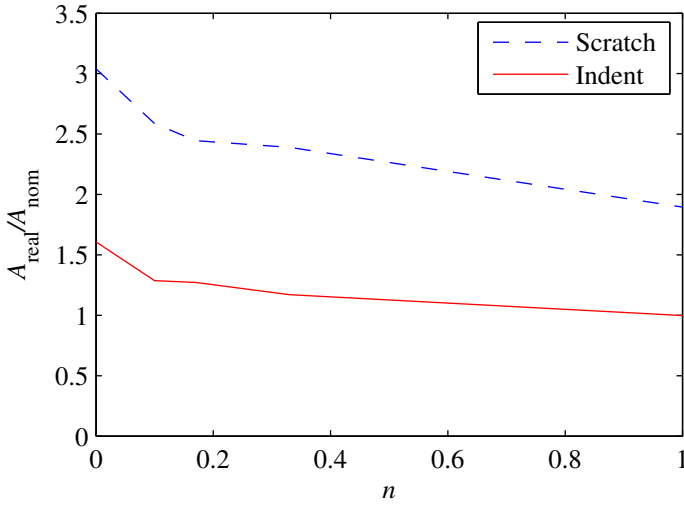


Figure 6. The actual normally projected contact area, A_{real} , divided by the nominal contact area, A_{nom} , plotted against the hardening parameter n . Results both for indentation and scratching. $\Lambda(\epsilon_{\text{repr}} = 8\%) = 1000$.

was used according to

$$H = C_1\sigma_1 + C_2\sigma_h. \tag{12}$$

In Equation (12), C_1 and C_2 are constants analogous to C in Equation (1), and σ_1 and σ_h are representative stresses at different values of the effective plastic strain. In short, it is obvious that the strain levels at scratching are significantly higher than the corresponding ones at indentation, indicating that the scratch

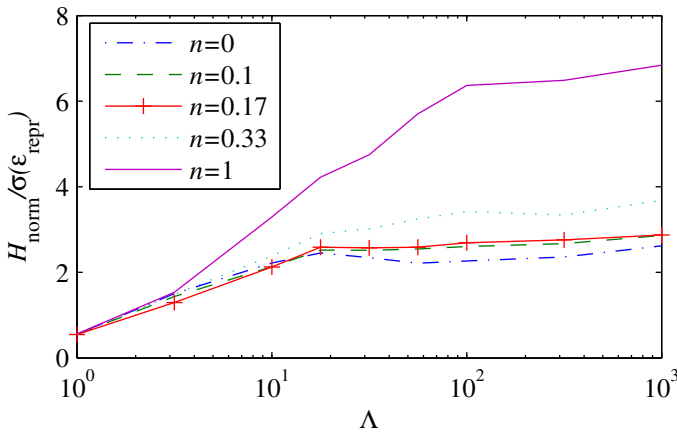


Figure 7. Scratch normal hardness divided by the yield stress, at a plastic strain of 8 %, for different $\Lambda(\epsilon_{\text{repr}} = 8\%)$ and n .

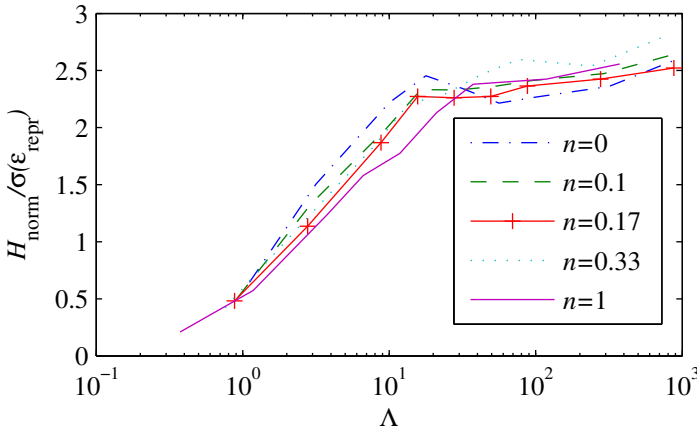


Figure 8. Scratch normal hardness divided by the yield stress at a plastic strain of 35 %, for different Λ ($\epsilon_{\text{repr}} = 35\%$) and n .

test is much less appropriate for standard material characterization, as will be discussed further below. Furthermore, the tangential hardness did not level out for large values of Λ (see Figure 9).

The scratch test is influenced by friction to a greater degree than is the standard indentation test. A numerical study was performed for $\Lambda(\epsilon_{\text{repr}} = 8\%) = 100$, investigating this effect. For this material the normal scratch hardness clearly decreased with increasing interfacial friction (see Figure 10). The opposite effect was noted for the tangential hardness (see Figure 11) This certainly complicates the interpretation of scratch experiments, particularly for material characterization based on the concept of a representative strain, since it is necessary to know both the dependence on friction and the coefficient of friction itself to draw any conclusion based on the scratch hardness. The representative level of plastic

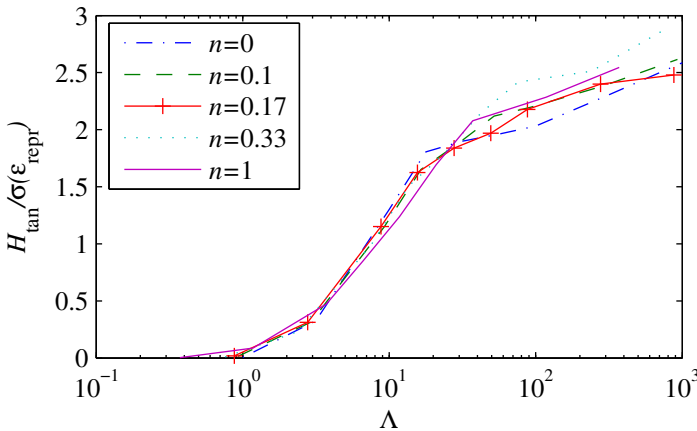


Figure 9. Scratch tangential hardness divided by the yield stress at a plastic strain of 35 % for different Λ ($\epsilon_{\text{repr}} = 35\%$) and n .

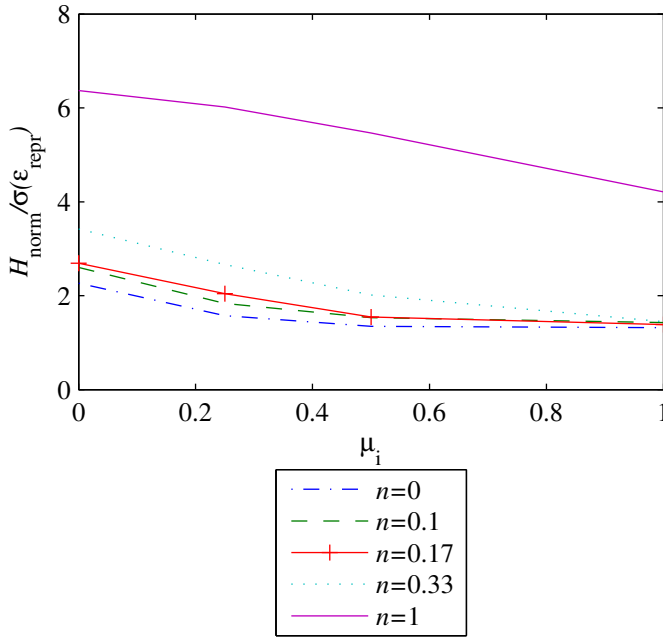


Figure 10. The influence of friction on normal hardness for strain hardening materials with different hardening n . Simulations with $\Lambda(\epsilon_{\text{repr}} = 8\%) = 100$.

strain varies with the coefficient of internal friction, which further complicates the matter. In fact, friction did increase ϵ_{repr} (for example, $\epsilon_{\text{repr}} \approx 0.5$ for $\mu_i = 0.2$) but it was still possible to determine accurate values for this variable. Interestingly, Felder and Bucaille [2006] find that the scratch hardness, when the contact area is calculated using Equation (10), is independent of the interfacial coefficient of friction. This feature was not studied in detail here, as a precise determination of the scratch width w was hard to achieve. However, preliminary results indicate that any frictional effect was within the margin of error.

Looking at the first principal stress distribution in Figures 17 and 18, it is clearly shown that the increasing shear exerted by the frictional force substantially increased the tensile forces in the wake of the stylus, potentially causing fracture.

It is also important to compare the values of the ratio of normal scratch hardness to tangential scratch hardness. This ratio may be expressed as

$$\frac{H_{\text{norm}}}{H_{\text{tan}}} = \frac{F_{\text{norm}} A_{\text{tan}}}{A_{\text{norm}} F_{\text{tan}}}. \quad (13)$$

The ratio $F_{\text{norm}}/F_{\text{tan}}$ may be replaced with $(\frac{2}{\pi} \tan \beta)^{-1}$ using Equation (17) below, assuming frictionless contact (see Section 4.3). The ratio of the tangentially projected area (A_{tan}) to the normally projected area (A_{norm}) can be written as

$$\frac{A_{\text{tan}}}{A_{\text{norm}}} = \frac{\frac{w^2}{4} \tan \beta}{(w/2)^2 \frac{\pi}{2}} = \frac{2}{\pi} \tan \beta \quad (14)$$

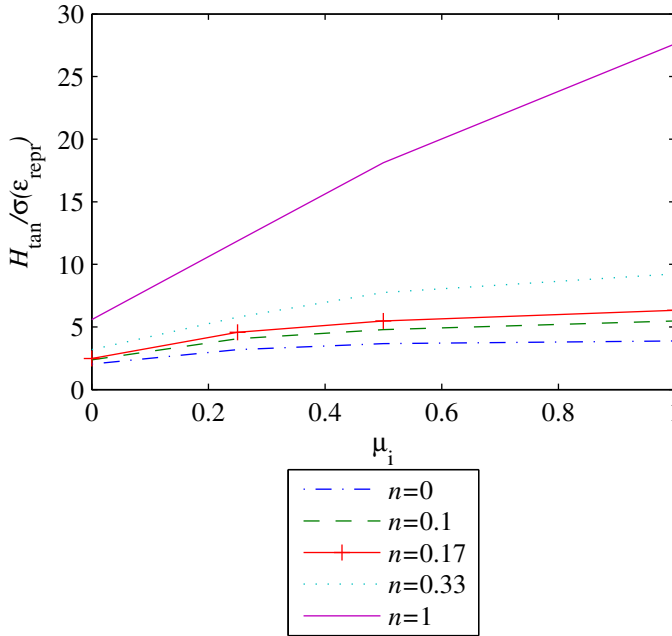


Figure 11. The influence of friction on tangential hardness for strain hardening materials with different hardening n . Simulations with $\Lambda(\epsilon_{repr} = 8\%) = 100$.

using Equation (10) and (11). Thus the hardness ratio simplifies to

$$\frac{H_{norm}}{H_{tan}} = \frac{\frac{2}{\pi} \tan \beta}{\frac{2}{\pi} \tan \beta} = 1. \tag{15}$$

This result can be seen in Figure 12, where the ratio of H_{norm}/H_{tan} goes to unity as the Johnson parameter Λ increases. For smaller Λ contact was no longer only on the front face and hence the ratio $H_{norm}/H_{tan} > 1$. This was due to the effect of the contact pressure on the rear face which decreases the tangential force F_{tan} , unlike the normal force F_{norm} which is increased by the contact force on the rear face. It is worth mentioning that Equation (15) applies to a conical stylus but not necessarily to styli of other shapes, since the derivation of Equation (15) relies on the independence of the contact pressure $p(r)$ from the angle φ [Subhash and Zhang 2002]. This is almost true for the conical stylus (as shown by the numerical calculations) but not necessarily for other shapes.

4.3. The apparent coefficient of friction. The tangential force divided by the normal force is known as the apparent coefficient of friction μ_0 or macroscopic friction at scratching. The apparent coefficient of friction is of interest since it will influence the amount of shear exerted on the specimen and consequently many other relevant scratch quantities.

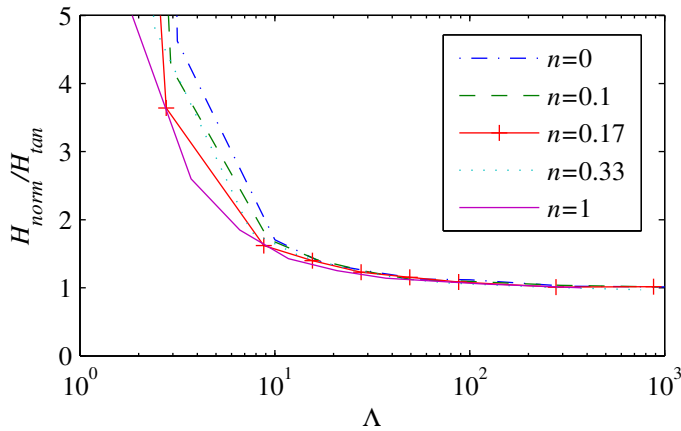


Figure 12. The ratio of normal hardness to tangential hardness at frictionless scratching. Note that the ratio goes to unity as Λ ($\epsilon_{\text{repr}} = 35\%$) increases.

The frictional force is often decomposed into an adhesive or interfacial part and a plowing part [Bowden and Tabor 1950]. This allows the apparent coefficient of friction μ_0 to be written as

$$\mu_0 = \mu_i + \mu_p, \quad (16)$$

where μ_i is the interfacial coefficient of friction, and μ_p is the plowing coefficient of friction. An estimate of the plowing coefficient of friction may be found by integrating the pressure vector over the surface. Assuming that the contact pressure is axi-symmetric over the contact area [Subhash and Zhang 2002] and that contact only occurs on the front face of the stylus with a constant contact radius [Goddard and Wilman 1962], the plowing coefficient of friction may be found by

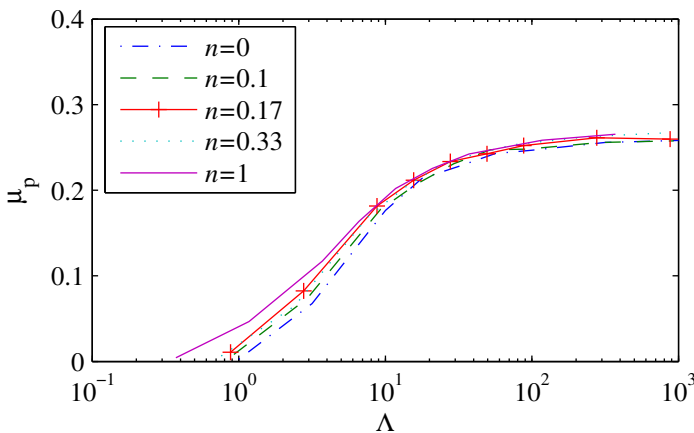


Figure 13. The plowing coefficient of friction. Λ is evaluated using $\epsilon_{\text{repr}} = 35\%$. Interfacial friction $\mu_i = 0$.

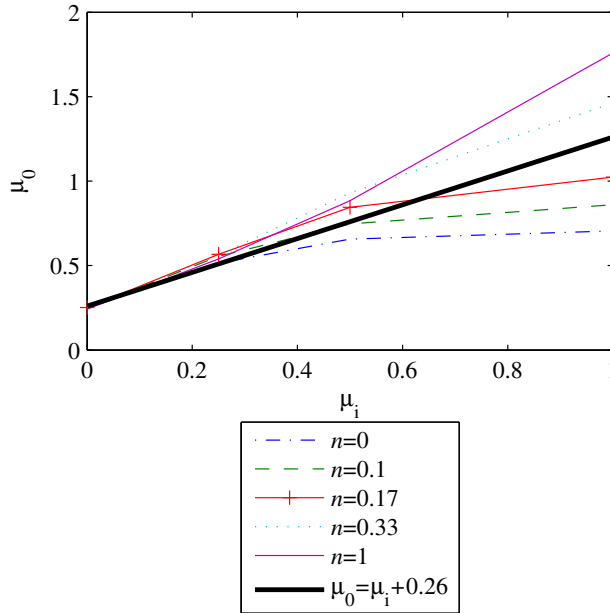


Figure 14. The apparent coefficient of friction μ_0 as a function of interfacial friction μ_i . Simulated with $n = 0, n = 0.1, n = 0.17, n = 0.33, n = 1$ and $\Lambda(\epsilon_{repr} = 8\%) = 100$. Equation (18) is plotted as a reference.

$$\mu_p = \frac{F_{tan}}{F_{norm}} = \frac{\mathbf{e}_t \cdot \int_A p(r) \mathbf{n} dA}{\mathbf{e}_n \cdot \int_A p(r) \mathbf{n} dA} = \tan(\beta) \frac{\int_A \cos(\varphi) p(r) dA}{\int_A p(r) dA} = \frac{2}{\pi} \tan \beta, \tag{17}$$

where $p(r)$ is the contact pressure, r is a radial coordinate, φ is the angle from the symmetry plane, $\mathbf{n} = (\sin(\beta) \cos(\varphi), \sin(\beta) \sin(\varphi), \cos \beta)$ is the surface normal vector, and \mathbf{e}_t and \mathbf{e}_n are vectors in the scratch direction and normal direction, respectively. Equation (17) gives $\mu_p \approx 0.26$ at frictionless contact for a cone with $\beta = 22^\circ$. Assuming a constant pressure distribution, Goddard and Wilman [1962] come to the same conclusion.

Previous simulations of frictionless scratching of elastic perfectly plastic materials ($n = 0$) show that the plowing coefficient of friction varies with Johnson’s parameter Λ [Bucaille et al. 2001]. In Figure 13 it can be seen that all materials (with different hardening exponents) follow the same master curve and that the plowing coefficient of friction levels out to an approximate value of 0.26, as predicted by Equation (17). Choosing a different ϵ_{repr} will only shift the curves along the x -axis. The coincidence of the curves in Figure 13 is, however, very robust with respect to the chosen representative strain. In fact, it is only the curve for $n = 1$ that was found to have a significant dependence on the chosen representative plastic strain.

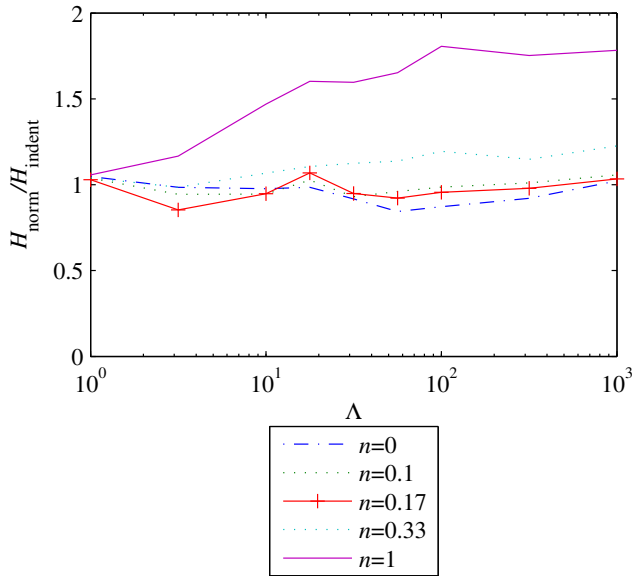


Figure 15. Ratio of scratch normal hardness to indentation hardness. Simulations with different n and Λ ($\epsilon_{\text{repr}} = 8\%$).

Such master curves could possibly be used for material characterization, in particular for polymers where Λ is small. It should be emphasized, though, that interfacial friction severely influences the apparent coefficient of friction, as shown in Figure 14. Thus it is of utmost importance to acquire an accurate value of the interfacial friction to be able to use the curve in Figure 13 for material characterization, which limits its practical usefulness.

To investigate the influence of hardening on the apparent coefficient of friction (when $\mu_i \neq 0$) a series of simulations for $\Lambda(\epsilon_{\text{repr}} = 8\%) = 100$ has been performed for a varying hardening parameter n . The results in Figure 14 show that for moderate levels of interfacial friction the apparent coefficient of friction μ_0 is independent of the hardening exponent n for nonfrictionless scratching. It is clear that this is not true for greater levels of interfacial friction. However, Figure 14 suggests that it is a reasonable approximation for the determination of the apparent coefficient of friction for levels of interfacial friction of up to 0.25. For materials with large Λ , where the plowing part of the apparent coefficient friction is constant (≈ 0.26 for a conical stylus with angle $\beta = 22^\circ$), Equation (16) may be simplified using Equation (17) to

$$\mu_0 = \mu_i + 0.26. \quad (18)$$

4.4. Scratch hardness vs. indentation hardness. The ratio of normal scratch hardness to normal indentation hardness is dependent on the hardening of the material in question. Since the scratch test is more sensitive to hardening than the indentation test, that is, has larger representative strain, a high level of hardening (n close to 1) gives a higher scratch hardness to indentation hardness ratio (see Figure 15). The opposite holds for low hardening or perfectly plastic materials. Thus for moderate levels of hardening the ratio of scratch normal hardness to indentation hardness is in the vicinity of 1. However, when interfacial

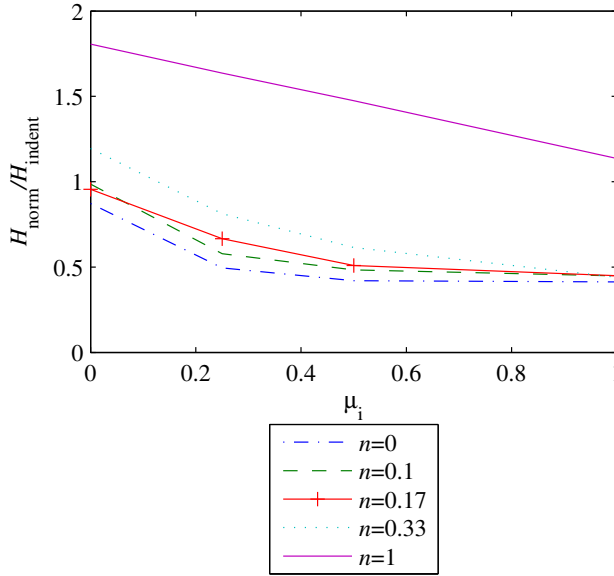


Figure 16. Ratio of scratch normal hardness to indentation hardness at varying interfacial friction μ_i . Simulations with different n and $\Lambda(\epsilon_{repr} = 8\%) = 100$.

friction was present as in Figure 16, the ratio of scratch normal hardness to indentation hardness was substantially decreased and no longer larger than 1.

The constant C in Figure 2 was, as stated earlier, found to be approximately 2.5 for scratching of predominately plastic materials. This is very close to the finding by Atkins and Tabor [1965] ($C \approx 2.54$) for conical indentation. However, it should be emphasized once again that the representative strain level is much higher at scratching (35% at scratching and 11% at indentation). This suggests that the ratio of scratch normal hardness to indentation hardness could simply be expressed as the ratio of the yield stress at 11% plastic strain to the yield stress at 35% plastic strain for materials undergoing mainly plastic deformation. However, this is not an exact relation, and in particular for higher levels of strain hardening this relation is not accurate.

It seems appropriate in this context to also discuss to some extent the behavior of field variables at scratching, particularly the stress fields, and specifically, to compare them to corresponding indentation quantities. In a practical situation this is pertinent to such features as crack initiation and growth, which are known to be of greater interest during scratching than during sharp (conical) indentation. In this discussion it is assumed in a straightforward manner that cracking is governed by a simple stress criterion such as

$$\sigma_1 = \sigma_B, \tag{19}$$

where σ_1 and σ_B are maximum principal stress and uniaxial failure stress, respectively. Accordingly, an approach based on fracture mechanics is not considered here.

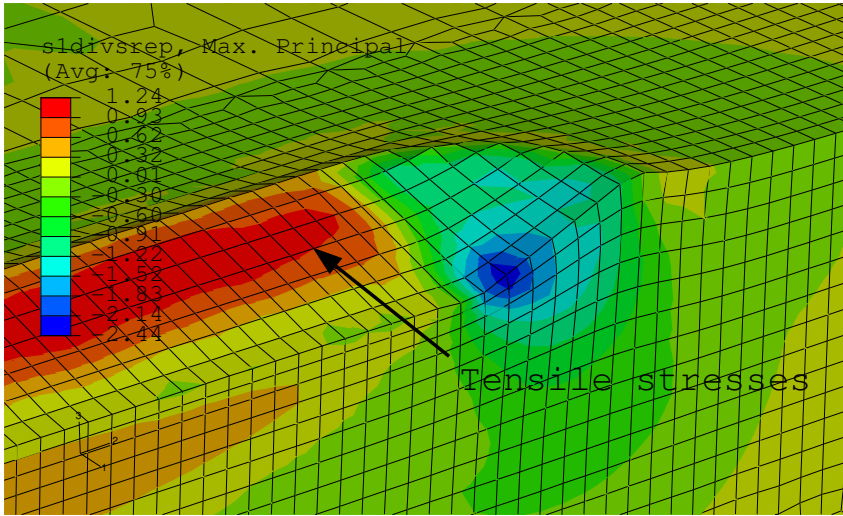


Figure 17. The first principal stress divided by the representative stress at frictionless scratching of a $\Lambda(\epsilon_{repr} = 8\%) = 100$ material with hardening exponent $n = 0.33$ and representative stress $\sigma(\epsilon_{repr} = 35\%) = 240$ MPa.

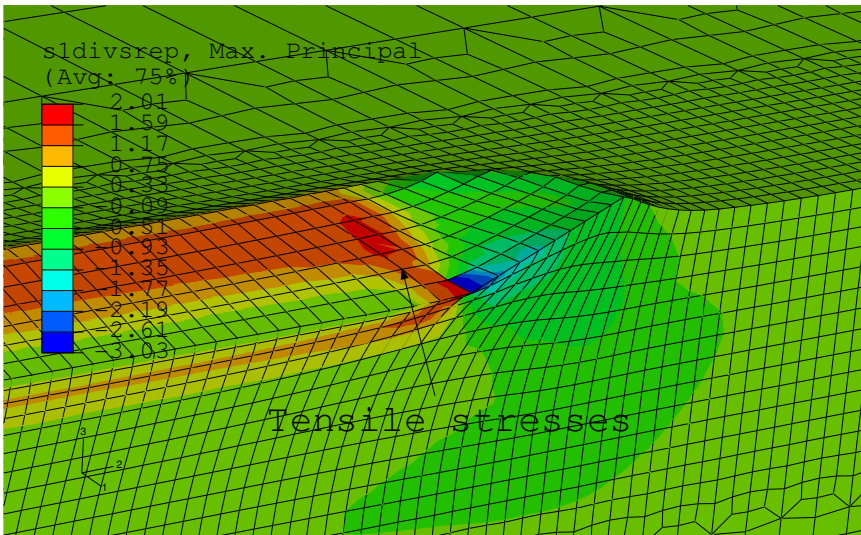


Figure 18. The first principal stress divided by the representative stress at scratching of a $\Lambda(\epsilon_{repr} = 8\%) = 100$ material with hardening exponent $n = 0.33$, representative stress $\sigma(\epsilon_{repr} = 35\%) = 240$ MPa and interfacial friction $\mu_i = 0.25$.

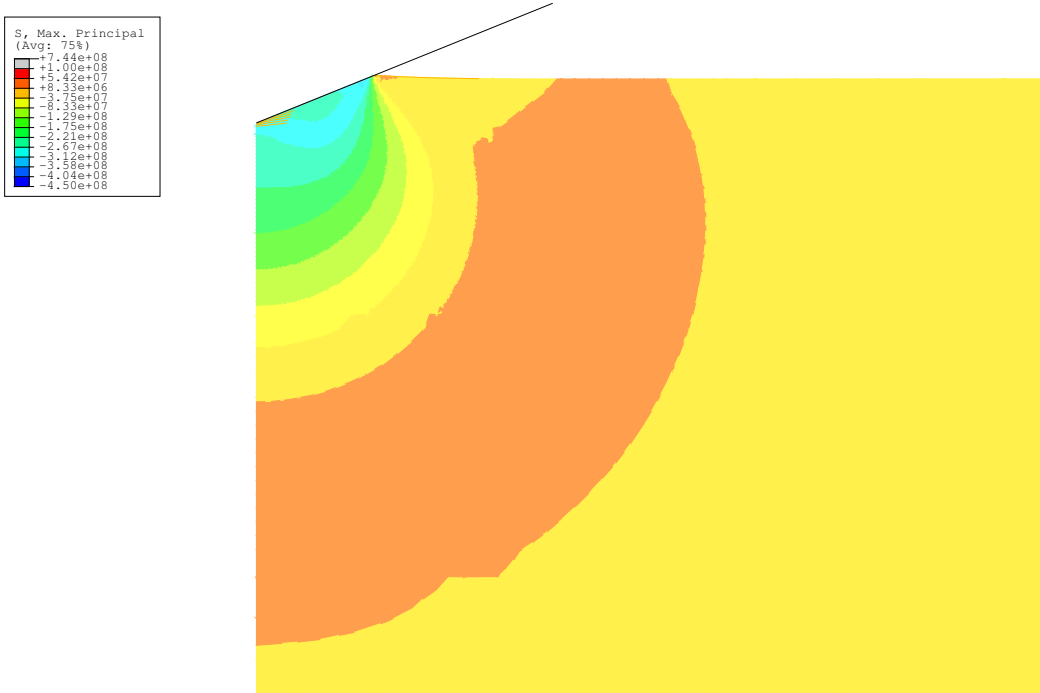


Figure 19. The first principal stress (in Pa) at frictionless indentation of a $\Lambda(\epsilon_{\text{repr}} = 8\%) = 100$ material with hardening exponent $n = 0.33$ and representative stress $\sigma(\epsilon_{\text{repr}} = 35\%) = 240$ MPa.

For this purpose, results for the maximum principal stress are shown for scratching in Figures 17 and 18, and for indentation in Figures 19 and 20, for frictionless as well as frictional contact ($\mu_i = 0.25$). The material in this case has a value of the Johnson parameter $\Lambda(\epsilon_{\text{repr}} = 8\%) = 100$, indicating mainly plastic but also some elastic deformation during scratching. As can be seen in Figures 17 and 18, the front of the stylus was exposed to compressive stresses during scratching as could be expected, while in the wake of the stylus, high tensile stresses developed, potentially causing fracture based on Equation (19). These tensile stresses substantially increased when frictional effects were accounted for, as shown in Figure 18.

In comparing the scratch results with corresponding indentation results, two features were immediately obvious from Figures 19 and 20. First, the tensile stresses at scratching were much higher than at indentation and, second, scratch values were much more influenced (that is, they increased) by friction. The same was also found, as indicated above, for global quantities such as hardness.

There are very few theoretical numerical studies concerned with the details of stress levels, specifically with cracking during scratching. However, recently Holmberg et al. [2003] presented a FEM study of scratching using a Rockwell indenter (a conical indenter with a spherical tip). In this study, scratching of a film/substrate system was considered (a single material combination, a TiN coating of a steel substrate, was explicitly investigated), which means that the present results and those of Holmberg et al. [2003] are not directly comparable. However, good qualitative agreement between the two sets of results are

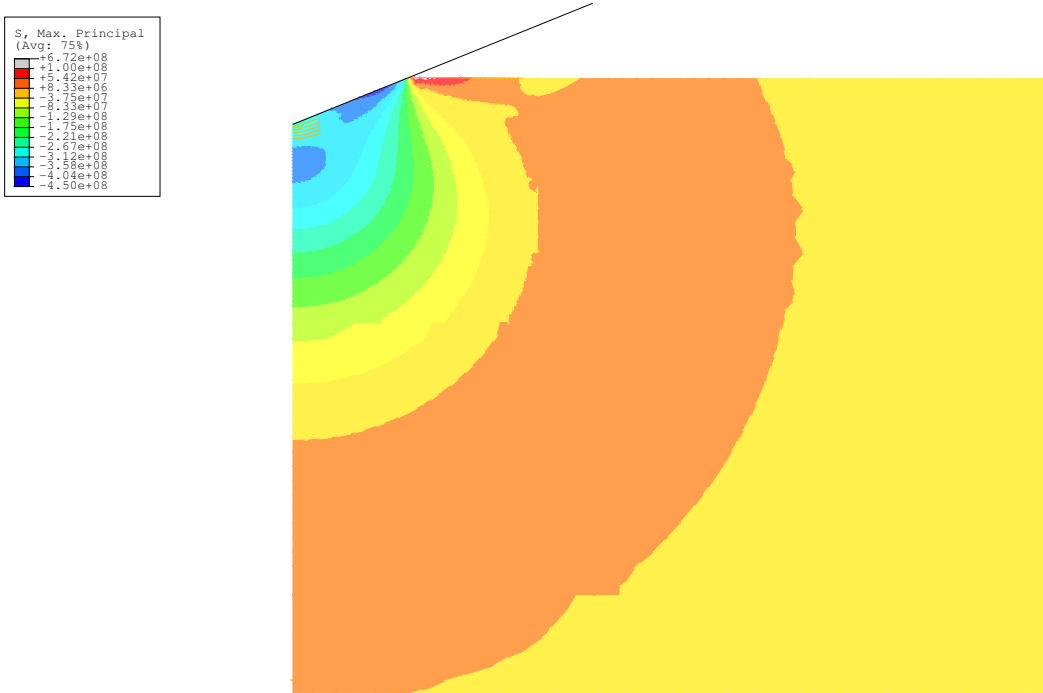


Figure 20. The first principal stress (in Pa) at indentation of a $\Lambda(\epsilon_{\text{repr}} = 8\%) = 100$ material with hardening exponent $n = 0.33$ and representative stress $\sigma(\epsilon_{\text{repr}} = 35\%) = 240$ MPa and interfacial friction $\mu_i = 0.25$.

found. For example, both investigations report a location of the maximum tensile stress in the wake of the stylus, which has also been noted in a number of previous studies.

5. Conclusions

The scratch test has been analyzed using a numerical approach based on the finite element method. The most important findings can be summarized as follows:

- The geometric approximation of the contact area given by Goddard and Wilman [1962] works well for materials where the plastic deformation is substantial. Not so, however, for materials with mainly elastic deformation.
- The concept of a representative strain holds true for scratching. The representative strain was found to be very large, approximately 35%.
- Interfacial friction will severely influence important scratch quantities.
- The numerical results suggest that other scratch quantities, such as the apparent coefficient of friction, can be used for material characterization.
- The normal scratch hardness and the tangential scratch hardness will be approximately equal for a conical stylus scratching a metallic material at low friction. As the friction increases this ratio will decrease since the normal hardness decreases and the tangential hardness increases.

- A useful first approximation of the apparent coefficient of friction can be found by adding the interfacial coefficient of friction to the master curve of friction (see Figure 13). For metallic materials this may be expressed as $\mu_0 = \mu_i + 0.26$. This explicit relation only holds for a conical stylus with angle $\beta = 22^\circ$. Obviously the opposite, to find the interfacial coefficient of friction μ_i from μ_0 , is also possible.
- The scratch normal hardness of a specimen is generally found to be equal to or higher than the indentation hardness of the same specimen, for frictionless contact. In case of a highly strain hardening material the scratch hardness will be substantially higher due to the much higher level of representative strain at scratching. Not so, however, when friction is present.

6. Acknowledgements

The authors wish to thank Professor Fred Nilsson for providing valuable advice and discussions, and for reading and commenting on the manuscript.

References

- [ABAQUS 2004] ABAQUS, *ABAQUS Manual v.6.4*, ABAQUS Inc., Pawtucket R.I. USA, 2004.
- [Atkins and Tabor 1965] A. G. Atkins and D. Tabor, "Plastic indentation in metals with cones", *Journal of The Mechanics and Physics of Solids* **13** (1965), 149–164.
- [Bowden and Tabor 1950] F. P. Bowden and D. Tabor, *The friction and lubrication of solids*, Clarendon press, 1950.
- [Briscoe et al. 1996] B. J. Briscoe, P. D. Evans, E. Pelillo, and S. K. Sinha, "Scratching maps for polymers", *Wear* **200** (1996), 137–147.
- [Bucaille et al. 2001] J. L. Bucaille, E. Felder, and G. Hochstetter, "Mechanical analysis of the scratch test on elastic and perfectly plastic materials with three-dimensional finite element modeling", *Wear* **249** (2001), 422–432.
- [Bucaille et al. 2004] J. L. Bucaille, E. Felder, and G. Hochstetter, "Experimental and three-dimensional finite element study of scratch test of polymers at large deformations", *Journal of Tribology* **126** (2004), 372–379.
- [Bull 1991] S. J. Bull, "Failure modes in scratch adhesion testing", *Surface and Coatings Technology* **50** (1991), 25–32.
- [Bull 1999] S. J. Bull, "Can scratch testing be used as a model for the abrasive wear of hard coatings?", *Wear* **233-235** (1999), 412–423.
- [Childs 1970] T. H. C. Childs, "The sliding of rigid cones over metals in high adhesion conditions", *International Journal of Mechanical Sciences* **12** (1970), 393–403.
- [Dao et al. 2001] M. Dao, N. Chollacoop, K. J. V. Vliet, T. A. Venkatesh, and S. Suresh, "Computational modeling of the forward and reverse problems in instrumented sharp indentation", *Acta Materialia* **49** (2001), 3899–3918.
- [Felder and Bucaille 2006] E. Felder and J. L. Bucaille, "Mechanical analysis of the scratching of metals and polymers with conical indenters at moderate and large strains", *Tribology International* **39** (2006), 70–87.
- [Fleck and Hutchinson 1993] N. A. Fleck and J. W. Hutchinson, "A phenomenological theory for strain gradient effects in plasticity", *Journal of the Mechanics and Physics of Solids, Volume* **41** (1993), 1825–1857.
- [Gauthier et al. 2001] C. Gauthier, S. Lafaye, and R. Schirrer, "Elastic recovery of a scratch in a polymeric surface: experiments and analysis", *Tribology International* **34** (2001), 469–479.
- [Giannakopoulos et al. 1994] A. E. Giannakopoulos, P.-L. Larsson, and R. Vestergaard, "Analysis of Vickers indentation", *International journal of Solids and Structures* **31** (1994), 2679–2708.
- [Gilormini and Felder 1983] P. Gilormini and E. Felder, "Theoretical and experimental study of the ploughing of a rigid-plastic semi-infinite body by a rigid pyramidal indenter", *Wear* **88** (1983), 195–206.
- [Goddard and Wilman 1962] J. Goddard and H. Wilman, "A theory of friction and wear during the abrasion of metals", *Wear* **5** (1962), 114–135.

- [Holmberg et al. 2003] K. Holmberg, A. Laukkanen, H. Ronkainen, K. Wallin, and S. Varjus, “A model for stresses, crack generation and fracture toughness calculation in scratched TiN-coated steel surfaces”, *Wear* **254** (2003), 278–291.
- [Johnson 1970] K. L. Johnson, “The correlation of indentation experiments”, *Journal of The Mechanics and Physics of Solids* **18** (1970), 115–126.
- [Johnson 1985] K. L. Johnson, *Contact mechanics*, Cambridge University Press, UK, 1985.
- [Larsson 2001] P.-L. Larsson, “Investigation of sharp contact at rigid-plastic conditions”, *International Journal of Mechanical Sciences* **43** (2001), 895–920.
- [Larsson et al. 1996] P.-L. Larsson, A. E. Giannakopoulos, E. Söderlund, D. J. Rowcliffe, and R. Vestergaard, “Analysis of Berkovich indentation”, *International Journal of Solids and Structures* **33** (1996), 221–248.
- [Laursen and Simo 1992] T. A. Laursen and J. C. Simo, “A study of microindentation using finite elements”, *Journal of Materials Research* **7** (1992), 618–626.
- [Malzbender and de With 2001] J. Malzbender and G. de With, “Analysis of scratch testing of organic-inorganic coatings on glass”, *Thin Solid Films* **386** (2001), 68–78.
- [Mata 2004] M. Mata, *Continuum analysis of sharp indentation experiments in metallic materials: theory and finite element simulations.*, Ph.D. thesis, Universitat Politècnica de Catalunya, 2004.
- [Subhash and Zhang 2002] G. Subhash and W. Zhang, “Investigation of the overall friction coefficient in single-pass scratch test”, *Wear* **252** (2002), 123–134.
- [Tabor 1951] D. Tabor, *Hardness of metals*, Cambridge University Press, UK, 1951.
- [Thouless 1998] M. D. Thouless, “An analysis of spalling in the microscratch test”, *Engineering Fracture Mechanics* **61** (1998), 75–81.
- [Vathaire et al. 1981] M. D. Vathaire, F. Delamare, and E. Felder, “An upper bound model of ploughing by a pyramidal indenter”, *Wear* **66** (1981), 55–64.
- [Wredenberg and Larsson 2005] F. Wredenberg and P.-L. Larsson, “Experimental and numerical analysis of the scratch test”, *WIT Transactions on Engineering Sciences* **49** (2005), 251–260.

Received 27 Sep 2006. Accepted 29 Nov 2006.

FREDRIK WREDENBERG: fredrik@hallf.kth.se
KTH Solid Mechanics, Royal Institute of Technology, S-10044, Stockholm, Sweden

PER-LENNART LARSSON: pelle@hallf.kth.se
KTH Solid Mechanics, Royal Institute of Technology, S-10044, Stockholm, Sweden

SUBMISSION GUIDELINES

ORIGINALITY

Authors may submit manuscripts in PDF format on-line. Submission of a manuscript acknowledges that the manuscript is *original and has neither previously, nor simultaneously, in whole or in part, been submitted elsewhere*. Information regarding the preparation of manuscripts is provided below. Correspondence by email is requested for convenience and speed. For further information, write to:

Marie-Louise Steele
Division of Mechanics and Computation
Durand Building, Room 262
Stanford University
Stanford CA 94305

LANGUAGE

Manuscripts must be in English. A brief abstract of about 150 words or less must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate), and email address if available. A home-page URL is optional.

FORMAT

Authors are encouraged to use \LaTeX and the standard article class, but submissions in other varieties of \TeX , and, exceptionally in other formats, are acceptable. Electronic submissions are strongly encouraged in PDF format only; after the refereeing process we will ask you to submit all source material.

REFERENCES

Bibliographical references should be listed alphabetically at the end of the paper and include the title of the article. All references in the bibliography should be cited in the text. The use of \BIBTeX is preferred but not required. Tags will be converted to the house format (see a current issue for examples), however, in the manuscript, the citation should be by first author's last name and year of publication, e.g. "as shown by Kramer, et al. (1994)". Links will be provided to all literature with known web locations and authors are encouraged to provide their own links on top of the ones provided by the editorial process.

FIGURES

Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple, or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to:

production@mathscipub.org

Each figure should be captioned and numbered so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript with all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables.

WHITE SPACE

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

PROOFS

Page proofs will be made available to authors (or to the designated corresponding author) at a web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Journal of Mechanics of Materials and Structures

Volume 2, N° 3 March 2007

A proposed method for fatigue crack detection and monitoring using the breathing crack phenomenon and wavelet analysis	V. K. NGUYEN AND O. A. OLATUNBOSUN	399
Mixed piezoelectric plate elements with continuous transverse electric displacements	E. CARRERA AND C. FAGIANO	421
A semi-infinite higher-order displacement discontinuity method and its application to the quasistatic analysis of radial cracks produced by blasting	H. HOSSEINI_NASAB AND M. F. MARJI	439
In-situ optimized PWAS phased arrays for Lamb wave structural health monitoring	L. YU AND V. GIURGIUTIU	459
One-dimensional thermoelastic waves in elastic half-space with dual phase-lag effects	S. KR. ROYCHOUdhURI	489
Ferromagnetic shape memory effects in an iron palladium alloy	J. CUI AND T. SHIELD	505
Macroscopic elastic properties of randomly packed balloons	I. TAGUCHI AND M. KURASHIGE	529
Cylindrical indentation induced deformation in face-centered cubic metal single crystals	Y. X. GAN, Y. SAITO AND X. CHEN	557
On the numerics and correlation of scratch testing	F. WREDENBERG AND P.-L. LARSSON	573



1559-3959(200703)2:3;1-B