Journal of
Mechanics of
Materials and Structures

JOURNAL OF MECHANICS OF MATERIALS AND STRUCTURES

msp.org/jomms

Founded by Charles R. Steele and Marie-Louise Steele

msp

INDENTATION AND RESIDUAL STRESS IN THE AXIALLY SYMMETRIC ELASTOPLASTIC CONTACT PROBLEM

TIAN-HU HAO

Most theoretical studies of mechanical indention, going back to Boussinesq, Hertz, and later Sneddon, address the relations between indenter pressure, indention size, and stress components. However, the relationship between indention and residual stress is also interesting. Here we use the Dorris and Nemat-Nasser method to derive a relation between the indention and the residual stress components for an axisymmetric load.

1. Introduction

The elastic contact problem plays a key role in interpreting experimental results of indention. This study was first considered by Boussinesq and Hertz in the late nineteenth century, and later Sneddon made major contributions. These authors derived general relationships among the load, the displacement, and the contact area for an axisymmetric indenter.

Also of interest is the relationship between the indention and the residual stress. This has been addressed for example in [Suresh and Giannakopoulos 1998], where it is stated that the residual stress cannot be determined using the loading theory of elasticity. In [Hao 2006] we made some progress in the study of the problem in the framework of the theory of finite elasticity. The paper continues that investigation, by considering the important unloading case. As in the previous paper, we derive the elastoplastic deformation is derived using the DNN method [Dorris and Nemat-Nasser 1980]. The elastic deformation is eliminated from the total deformation, leading to the residual plastic deformation. Thus we determine the relation between the residual stress and the residual plastic deformation.

2. Analysis of the axially symmetrical finite elastic-plastic case

Following Dorris and Nemat-Nasser we write, for the axially symmetrical case,

$$D_{ab} = 0.5(v_{a,b} + v_{b,a}), \quad D_{\theta\theta} = v_2/x_2 \tag{1}$$

where x_1 and x_2 denote z and r, v_a is the increment displacement, $v_{a,b} = \partial v_a/\partial x_b$ and D_{ab}; $D_{\theta\theta}$ are the components of the rate tensor. Note that in [Dorris and Nemat-Nasser 1980] v_a is the velocity, but in this paper, it is the incremental displacement, whose dimension is length. Since we are dealing with small deformations superposed on initial stress body, the increment displacement v_a of the small deformation is a small part of the whole displacement.

Still following Dorris and Nemat-Nasser, we use the current configuration as the reference one. We deal with the first Piola–Kirchhoff stress increment $\delta\sigma_{ab}$, where the index a denotes the direction of the stress and the index b denotes the normal to the surface subjected to $\delta\sigma_{ab}$ in the reference configuration (note that $\delta\sigma_{ab} \neq \delta\sigma_{ba}$). Only the incompressible case is considered; the compressible case can be derived from the results of the incompressible case and it will be studied in another paper.

We turn to the constitutive equations, still following [Dorris and Nemat-Nasser 1980]. Similarly to what is done in Appendix A — compare A — we can write

$$\delta\sigma_{ab} = \Phi_{abce}D_{ce} + P\delta_{ab} - D_{ac}T_{cb} - D_{bc}T_{ca} + v_{a,c}T_{cb}, \tag{2}$$

where Φ_{abce} will be discussed later, P is an unknown scalar function (hydrostatic pressure), and T_{ca} is the Cauchy stress. For the flow theory, the constitutive equation is

$$\Phi_{abce} = (2\mu\delta_{ac}\delta_{be} - 6A\mu^2 ST'_{ab}T'_{ce}/\overline{T}^2), \tag{3}$$

where μ is the elastic constant, S equals $(2h/3 + 2\mu)^{-1}$, T'_{ab} and D'_{ab} are the deviatoric parts of the Cauchy stress T_{ab} and rate D_{ab}, the scalar A is defined as 1 if $T'_{ab}D'_{ab} \geq 0$ and as 0 if $T'_{ab}D'_{ab} < 0$, and $\overline{T}^2 = 1.5T'_{ab}T'_{ab}$.

The value of h is given by

$$h = \left(\frac{1}{E_t} - \frac{1}{E}\right)^{-1},$$

where E is the initial Young's modulus and E_t is the instantaneous tangent modulus. E_t equals $d\sigma/d\epsilon = En(\sigma/\sigma_y)^{1-1/n}$ for $0 \leq n \leq 1$, σ is the true stress, σ_y is the yield stress and ϵ is the logarithmic strain (σ and ϵ are of simple tension or compression).

From total deformation theory, one has

$$\Phi_{abce} = 2\mu(\overline{\gamma}/\gamma_0)^{n-1}\delta_{ac}\delta_{be} - \frac{3h(1-n)}{n}T'_{ab}T'_{ce}/\overline{T}^2 \tag{4}$$

for some n satisfying $0 \leq n \leq 1$; here $\overline{\gamma}$ is the effective increment strain and γ_0 is a reference increment strain.

Because (3) and (4) are in similar form, from now on, only the flow theory case is discussed. When $A = 0$, i.e., $T'_{ab}D'_{ab} < 0$ or $\sigma < \sigma_y$, we are in the unloading case or the elastic case, and we can deal with this problem as in [Hao 2006]. The case $A = 1$, i.e., $T'_{ab}D'_{ab} \geq 0$ and $\sigma > \sigma_y$, is the loading case, to be considered in this paper. As the cone indention causes compressive stresses, in order to satisfy the requirement of the loading case $A = 1$, the residual stress must also be compressive.

Following [Dorris and Nemat-Nasser 1980], in view of the constitutive equations, one obtains

$$\delta\sigma_{22} = 2\mu v_{2,2} + 2\mu^2 Sv_{1,1} + P, \tag{5}$$

$$\delta\sigma_{11} = (2\mu - 4\mu^2 S - T)v_{1,1} + P, \tag{6}$$

$$\delta\sigma_{21} = (2\mu - T)(v_{1,2} + v_{2,1})/2 + Tv_{2,1}, \tag{7}$$

$$\delta\sigma_{12} = (2\mu - T)(v_{1,2} + v_{2,1})/2, \tag{8}$$

$$\delta\sigma_{\theta\theta} = 2\mu v_2/r + 2\mu^2 Sv_{1,1} + P, \tag{9}$$

where T is the residual stress.

The homogeneous residual stress σ_R is equal to T. When the xy-plane is parallel to the surface, the residual stresses σ_x and σ_y are also equal to T. The first equilibrium equation is

$$\delta\sigma_{22,2} + (\delta\sigma_{22} - \delta\sigma_{\theta\theta})/x_2 + \delta\sigma_{21,1} = 0. \tag{10}$$

From the calculations in Appendix B, one obtains

$$2\mu v_{2,221} + 2\mu^2 S v_{1,121} + P_{,21} + (1/x_2)2\mu(v_{2,21} - v_{2,1}/x_2) + (2\mu - T)(v_{2,111} + v_{1,121})/2 + T v_{2,111} = 0 \tag{11}$$

The other equilibrium equation is

$$\delta\sigma_{11,1} + \delta\sigma_{12,2} + \delta\sigma_{12}/x_2 = 0. \tag{12}$$

Also in view of Appendix B, one has

$$(2\mu - 8\mu^2 S - T)v_{1,112}/2 + P_{,12} + (2\mu - T)(v_{1,222} + v_{1,22}/x_2 - v_{1,2}/x_2^2)/2 = 0, \tag{13}$$

Considering (11) and (13) and eliminating $P_{,12}$, one has

$$2\mu v_{2,221} + 2\mu^2 S v_{1,121} + (1/x_2)2\mu(v_{2,21} - v_{2,1}/x_2) + (2\mu - T)(v_{2,111} + v_{1,211})/2 + T v_{2,111}$$
$$- (2\mu - 8\mu^2 S - T)v_{1,112}/2 - (2\mu - T)(v_{1,222} + v_{1,22}/x_2 - v_{1,2}/x_2^2)/2 = 0. \tag{14}$$

Let $L(v_2) = v_{2,22} + v_{2,2}/x_2 - v_2/x_2^2$ and $L(v_{1,2}) = v_{1,222} + v_{1,22}/x_2 - v_{1,2}/x_2^2$. Substituting into (14), one obtains

$$2\mu L(v_2)_{,1} + 6\mu^2 S v_{1,112} + (2\mu + T)v_{2,111}/2 - (2\mu - T)L(v_{1,2})/2 = 0. \tag{15}$$

Let v_2 be F_{11}, where F_{11} is $\partial^2 F/\partial x_1^2$. In view of Appendix C, one has

$$(4\mu - 12\mu^2 S)L(F)_{11} + (2\mu + T)L^0 F_{1111} + (2\mu - T)L^2(F) = 0. \tag{16}$$

Let $G(s, x_1)$ be $\int_0^\infty x_2 F(x_1, x_2)J_1(sx_2)\,dx_2$ which is the Hankel transform of $F(x_1, x_2)$ with order 1 [Sneddon 1951]. Therefore, one has

$$4\mu(1 - 3\mu S)\int_0^\infty x_2 L(F)_{11} J_1(sx_2)\,dx_2$$
$$+ (2\mu + T)\int_0^\infty x_2(F)_{1111} J_1(sx_2)\,dx_2 + (2\mu - T)\int_0^\infty x_2 L^2(F)J_1(sx_2)\,dx_2 = 0. \tag{17}$$

If $x_2 \to 0$ and ∞, $x_2 F \to 0$, we have

$$\int_0^\infty x_2 L(F)J_1(sx_2)\,dx_2 = -s^2 G(s, x_1),$$
$$\int_0^\infty x_2 L^2(F)J_1(sx_2)\,dx_2 = -s^2\int_0^\infty x_2 L(F)J_1(sx_2)\,dx_2 = s^4 G(s, x_1). \tag{18}$$

Substituting (18) into (17), one obtains

$$-4\mu s^2(1 - 3\mu S)G(s, x_1)_{,11} + (2\mu + T)G(s, x_1)_{,1111} + s^4(2\mu - T)G(s, x_1) = 0 \tag{19}$$

Let $G(s, x_1) = N(s)\exp(mx_1)$, where m is a function of s. We obtain $G(s, x_1)_{,11} = N(s)\exp(mx_1)m^2$, $G(s, x_1)_{,1111} = N(s)\exp(mx_1)m^4$ and

$$s^4(1 - Q) - 2s^2(1 - 3\mu S)m^2 + (1 + Q)m^4 = 0,$$
$$m^2 = r^2 s^2 = s^2\{(1 - 3\mu S) \pm [Q^2 - 6\mu S - 9\mu^2 S^2]^{1/2}\}/(1 + Q), \tag{20}$$

where $r^2 = \{(1 - 3\mu S) \pm [Q^2 - 6\mu S - 9\mu^2 S^2]^{1/2}\}/(1 + Q)$ and $Q = T/2\mu$. One can deal only with the case where there are two different real positive roots r_1^2 and r_2^2. It can be proved that the same results will be obtained in the case with two conjugate complex roots.

Letting $x_1 \to \infty$, $v_2 \to 0$, $F_1 \to 0$, $r_1, r_2 > 0$, similar to [Hao 2006], one obtains

$$G(s, x_1) = \int_0^\infty x_2 F(x_1, x_2) J_1(sx_2)\, dx_2 = N_1(s)\exp(m_1 x_1) + N_2(s)\exp(m_2 x_1)$$
$$= N_1(s)\exp(-r_1 sx_1) + N_2(s)\exp(-r_2 sx_1). \tag{21}$$

Considering $x_1 = 0$, $\delta\sigma_{21} = 0$, in view of Appendix C, one has

$$G(s, x_1) = N_1(s)e^{-r_1 sx_1} + N_2(s)e^{-r_2 sx_1} = N_1(s)(e^{-r_1 sx_1} - Ue^{-r_2 sx_1}) \tag{22}$$

where

$$U = \frac{r_1^3 + r_1(2\mu - T)/(2\mu + T)}{r_2^3 + r_2(2\mu - T)/(2\mu + T)}.$$

Now, the stress component $\delta\sigma_{11}$ and v_1 are discussed. According to Appendix C, one has

$$\delta\sigma_{11} = (2\mu - T)[U(r_2^2 + 1) - (r_1^2 + 1)]\int_0^\infty sJ_0(sx_2)s^3 N_1(s)\, ds/2, \tag{23}$$

$$v_1 = (r_1 - Ur_2)\int_0^\infty sJ_0(sx_2)s^2 N_1(s)\, ds. \tag{24}$$

The boundary conditions are

$$(r_1 - Ur_2)a^4 \int_0^\infty s^3 J_0(sx_2)N_1(s)\, ds = a^4[v_1(x_2)]_{x_1=0} \quad \text{for } x_1 = 0,\ a \ge x_2 \ge 0,$$
$$\int_0^\infty s^4 J_0(sx_2)N_1(s)\, ds = 0 \qquad\qquad \text{for } x_1 = 0,\ x_2 > a, \tag{25}$$

where a is the radius of contact area, which will be discussed in detail later. Finally,

$$p^3 N_1(s) = p^3 N_1(p/a) = f(p), \quad s = p/a. \tag{26}$$

3. The circular cone and the residual stress

We now turn to a circular cone on a half-space and consider the residual stress. Let α be the angle of the circular cone (the angle between the axis of symmetry and the generatrix). Then

$$[v_1(x_2)]_{x_1=0} = b + a\cot\alpha(1 - x_2/a) \quad \text{for } a \ge x_2 \ge 0 \tag{27}$$

and

$$a^4[v_1(x_2)]_{x_1=0} = a^4[b + a\cot\alpha(1 - x_2/a)] = (r_1 - Ur_2)(A_0 + A_1 x_2/a) \quad \text{for } a \geq x_2 \geq 0 \qquad (28)$$

On the foundation of that the stress component $\delta\sigma_{11}$ is finite at the edge of the punch, the relation between b and a can be obtained. Similar to [Hao 2006], we obtain

$$f(P) =$$
$$\frac{1}{\sqrt{\pi}}\left\{ A_0\left(\cos p + p\int_0^1 u\sin(pu)\,du\right)\frac{\Gamma(1)}{\Gamma(3/2)} + A_1\left(\cos p + p\int_0^1 u^2\sin(pu)\,du\right)\frac{\Gamma(3/2)}{\Gamma(2)}\right\}. \qquad (29)$$

According to [Hao 2006] and Appendix D, for the compressive force R on the cone, one obtains

$$R = \pi(2\mu - T)a^2\cot\alpha\,\frac{r_2 r_1(r_1 r_2 - \beta) + (r_2^2 + r_1 r_2 + r_1^2) + \beta}{2r_2 r_1(r_2^2 + 2r_1 r_2 + r_1^2)^{1/2}}, \qquad (30)$$

where $\beta = \dfrac{2\mu - T}{2\mu + T}$ and r_1, r_2 can be calculated from the equalities $r_2^2 r_1^2 = \beta$, $r_2^2 + r_1^2 = \dfrac{2\mu - 6\mu^2 S}{2\mu + T}$.

The contact radius a is

$$a = \left(\frac{R[r_2 r_1(r_2^2 + 2r_1 r_2 + r_1^2)^{1/2}]}{\pi(2\mu - T)[r_2 r_1(r_1 r_2 - \beta) + (r_2^2 + r_1 r_2 + r_1^2) + \beta]\cot\alpha}\right)^{1/2}, \qquad (31)$$

from which the contact area πa^2 is easily calculated. The penetration depth is

$$v_1(x_1, x_2)_{x_1=0, x_2=0} = \tfrac{1}{2}\pi a\cot\alpha = \frac{1}{2}\left(\frac{\pi R\cot\alpha[r_2 r_1(r_2^2 + 2r_1 r_2 + r_1^2)^{1/2}]}{(2\mu - T)[r_2 r_1(r_1 r_2 - \beta) + (r_2^2 + r_1 r_2 + r_1^2) + \beta]}\right)^{1/2}. \qquad (32)$$

4. The unloading of the indenter

Lastly, as the residual stress is determined by the indention, the unloading of the indenter is now discussed. For the unloading case, let A be 0 in (3), i.e., $S = 0$ in (20). One obtains two roots r_1, r_2. Then, all other related values (penetration depth, contact area) of the purely elastic case are obtained.

For an example, the penetration depth in the unloading case is studied.

$$h = 0.5\{\pi R\cot\alpha[r_2 r_1(r_2 + r_1)]/(2\mu - T)[r_2 r_1(r_1 r_2 - r_2^2 r_1^2) + (r_2^2 + r_1 r_2 + r_1^2) + r_2^2 r_1^2]\}^{1/2} \qquad (33)$$

where h is the penetration depth of the purely elastic case and the values of r_1, r_2 can be obtained where $S = 0$ in (20). Subtracting it from the elastoplastic case, one obtains the penetration depth for the residual plastic deformation case.

5. Results and explanation

As an illustration, we take the specific example considered in [Hao 2006]. The parameters are $\mu = 10$–$100\,\text{GPa}$, $\sigma_y = $ yield stress $= 200\,\text{MPa}$, $R = 0.46\,\text{kg}$ and $\alpha = \pi/12$. The results are in Figures 1–3.

In the figure, N is a function of the elastic shear modulus μ and the plastic constant $k = E_t/E$ (recall that E is the initial Young's modulus and E_t is the instantaneous modulus). This parameter N equals $\{(\mu^{1/2}/3\mu_0)^k k^{(1-k)}\}$, where $\mu_0 = 7\,\text{Gpa}$. When $E_t = 0$ i.e. $k = 0$ or $N = 0$, the material is soft. When $k = 1$ or $N = (\mu^{1/2}/3\mu_0)$, the material is tough. Therefore, the parameter N is a coefficient to determine
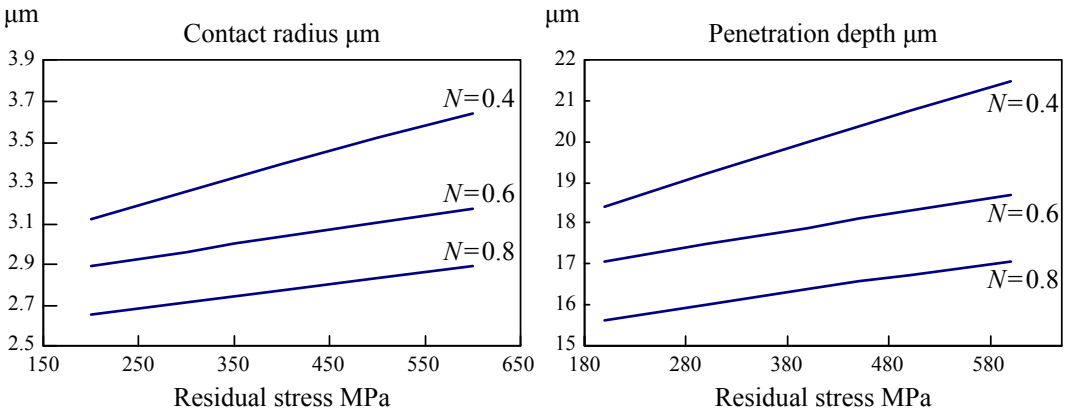
Figure 1. Elastoplastic case: contact radius (left) and penetration depth (right) versus residual stress.
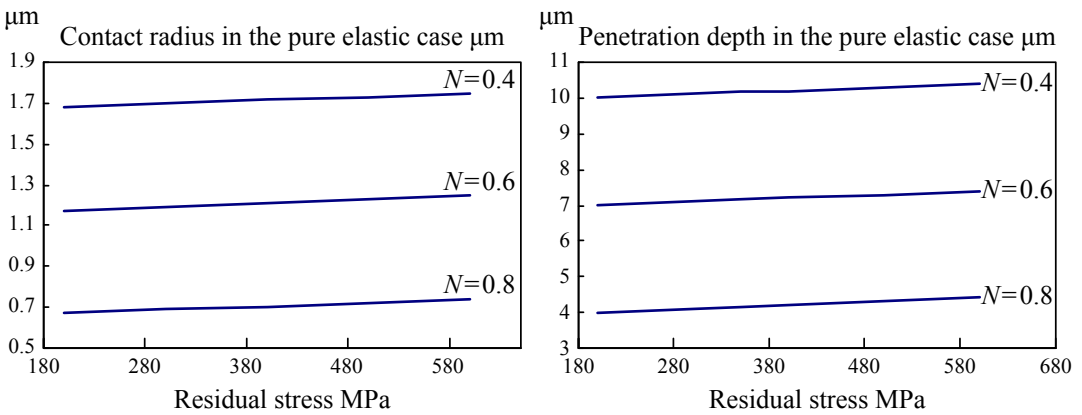

Figure 2. Purely elastic case: contact radius (left) and penetration depth (right) versus residual stress.
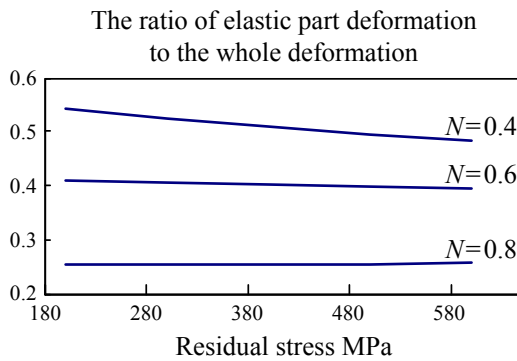

Figure 3. Ratio of elastic part deformation to the whole deformation, plotted versus residual stress.

the softness or toughness of materials. Figure 1 shows that in the elastoplastic case, the contact radius and the penetration depth vary with the residual stress. However, in the purely elastic case ($k = 1$), when N is constant, we see in Figure 2 that the contact radius and penetration depth are nearly constant. These figures show that the shear modulus is almost an identified factor to determine the contact area, the contact radius and the penetration depth. These figures also show that the effect of Jaumann rate is not notable for this case. Therefore, in the purely elastic case, when the Jaumann rate is even considered, we hardly determine the residual stress according to the indention. This means that in order to determine the residual stress the plastic deformation must be considered. Figure 3 shows that when N is a constant, the larger the residual stress is, the smaller the ratio of elastic part deformation to whole deformation becomes. Naturally, the larger the residual stress becomes, the larger the plastic part deformation also becomes.

6. Conclusions

In this paper, the axially symmetrical elastoplastic contact problem and its application are studied. The relation among the penetration depth, the contact radius and the residual stress has been determined. Besides, the unloading case is considered. The essence of this method is to deviate from the linear theory to consider Jaumann rate. When deviating from the linear theory a little, this difficult problem can be studied easily. For an example, when studying the effects of air inside crack in the piezoelectric materials, we must deal with the opening crack after deformation because before deformation, the crack is closed and no air can be in it. Therefore, in this case, the replacement of the boundary after deformation by that before deformation leads to great deviation. In order to avoid it, we use the approximate boundary after deformation. The approximate boundary after deformation is the boundary before deformation plus the displacement. Naturally, the displacement is found for the body before deformation as in the theory of elasticity. On the basis of this consideration, the semipermeable boundary condition was suggested in [Hao and Shen 1994].

Appendix A. Jaumann rate increment, reference and current configurations

According to [Dorris and Nemat-Nasser 1980], for an incompressible body, the Jaumann rate is

$$\dot{\sigma}_{ab} = DT_{ab}/Dt + (u_{a,c} - D'_{ac})T_{cb} - T_{ac}D'_{cb}, \qquad D'_{ab} = 0.5(u_{a,b} + u_{b,a})$$

where u_a is the velocity component, σ_{ab} is the Piola stress component and T_{ab} is the Cauchy stress component. For convenience, only the hypo-elastic solid is dealt with

$$\dot{\sigma}_{ba} = (2\mu D'_{ab} + \dot{P}\delta_{ab}) + (u_{a,c} - D'_{ac})T_{cb} - T_{ac}D'_{cb}$$

Using the variables $\delta\sigma_{ba}$ and D_{ab} from (1) to replace $\dot{\sigma}_{ba}$ and D'_{ab} from A, we obtain

$$\delta\sigma_{ba} = (2\mu D_{ab} + P\delta_{ab}) + (v_{a,c} - D_{ac})T_{cb} - T_{ac}D_{cb}$$

where $D_{ab} = 0.5(v_{a,b} + v_{b,a})$, v_a is the increment displacement component and $D_{\theta\theta} = v_2/x_2$. There are two configurations in our work. The configuration of the body after deformation is the current configuration. That before deformation can be the reference configuration.

Appendix B. Derivation of (11) and (13)

$$\delta\sigma_{22,2} + (\delta\sigma_{22} - \delta\sigma_{\theta\theta})/x_2 + \delta\sigma_{21,1} = 0$$

$$2\mu v_{2,22} + 2\mu^2 S v_{1,12} + P_{,2} + (1/x_2)2\mu(v_{2,2} - v_2/x_2) + (2\mu - T)(v_{2,11} + v_{1,12})/2 + T v_{2,11} = 0$$

$$2\mu v_{2,221} + 2\mu^2 S v_{1,121} + P_{,21} + (1/x_2)2\mu(v_{2,21} - v_{2,1}/x_2) + (2\mu - T)(v_{2,111} + v_{1,121})/2 + T v_{2,111} = 0$$

$$\delta\sigma_{11,1} + \delta\sigma_{12,2} + \delta\sigma_{12}/x_2 = 0$$

$$(2\mu - 4\mu^2 S - T)v_{1,11} + P_{,1} + (2\mu - T)(v_{2,12} + v_{1,22})/2 + (2\mu - T)(v_{2,1} + v_{1,2})/2x_2 = 0$$

In order to consider the incompressible equation $v_{1,1} + v_{2,2} + v_2/x_2 = 0$, which can become $v_{1,11} + v_{2,12} + v_{2,1}/x_2 = 0$, we have

$$(2\mu - 8\mu^2 S - T)v_{1,11}/2 + P_{,1} + (2\mu - T)[(v_{2,12} + v_{1,11} + v_{1,22})/2 + (v_{2,1} + v_{1,2})/2x_2] = 0 \quad (34)$$

$$(2\mu - 8\mu^2 S - T)v_{1,11}/2 + P_{,1} + (2\mu - T)(v_{1,22} + v_{1,2}/x_2)/2 = 0 \quad (35)$$

$$(2\mu - 8\mu^2 S - T)v_{1,112}/2 + P_{,12} + (2\mu - T)(v_{1,222} + v_{1,22}/x_2 - v_{1,2}/x_2^2)/2 = 0 \quad (36)$$

Appendix C. Derivation of (16), $\delta\sigma_{11}$ and v_1

Letting $v_2 = F_{11}$, from the incompressible equation $v_{1,1} + v_{2,2} + v_2/x_2 = 0$, we have

$$v_{1,1} = (-v_{2,2} - v_2/x_2) = (-F_2 - F/x_2)_{,11}, \quad v_1 = -(F_2 + F/x_2)_{,1}$$

$$v_{1,2} = (-F_2 - F/x_2)_{,12} = -(F_{22} + F_2/x_2 - F/x_2^2)_{,1} = -L(F)_{,1}, \quad v_{1,12} = -L(F)_{,11}, \quad (37)$$

$$v_{1,112} = -L(v_2)_1 = -L(F)_{111}$$

We have

$$2\mu L(F)_{,111} - 6\mu^2 S L(F)_{111} + (2\mu + T)F_{,11111}/2 + (2\mu - T)L^2(F)_{,1}/2 = 0$$

$$(4\mu - 12\mu^2 S)L(F)_{11} + (2\mu + T)F_{1111} + (2\mu - T)L^2(F) = 0$$

$$\delta\sigma_{21} = (2\mu - T)(v_{1,2} + v_{2,1})/2 + T v_{2,1} = (2\mu - T)[-(F_{,12} + F_{,1}/x_2)_{,2} + F_{,111}]/2 + T F_{,111}$$

$$= (2\mu - T)[F_{,111} - L(F)_1]/2 + T F_{,111} = [(2\mu + T)F_{,111} - (2\mu - T)L(F)_1]/2$$

Considering $G(s, x_1) = N_1(s)e^{-r_1 s x_1} + N_2(s)e^{-r_2 s x_1}$ and $\int_0^\infty x_2 L(F)J_1(sx_2)\, dx_2 = -s^2 G(s, x_1)$, one obtains

$$\int_0^\infty x_2 \delta\sigma_{21} J_1(sx_2)\, dx_2 = (2\mu + T)G(s, x_1)_{,111}/2 + (2\mu - T)s^2 G(s, x_1)_{,1}/2$$

$$\delta\sigma_{21} = [(2\mu + T)/2] \int_0^\infty s\{G(s, x_1)_{111} + [(2\mu - T)/(2\mu + T)]s^2 G(s, x_1)_1\}J_1(sx_2)\, ds$$

Considering $G(s, x_1) = N_1(s)e^{-r_1 s x_1} + N_2(s)e^{-r_2 s x_1}$, $G(s, x_1)_1 = -[N_1(s)r_1 s e^{-r_1 s x_1} + N_2(s)r_2 s e^{-r_2 s x_1}]$, $G(s, x_1)_{11} = N_1(s)r_1^2 s^2 e^{-r_1 s x_1} + N_2(s)r_2^2 s^2 e^{-r_2 s x_1}$, $G(s, x_1)_{111} = -[N_1(s)r_1^3 s^3 e^{-r_1 s x_1} + N_2(s)r_2^3 s^3 e^{-r_2 s x_1}]$, one has

$$\delta\sigma_{21} = -[(2\mu+T)/2]\int_0^\infty s\{N_1(s)r_1^3 s^3 e^{-r_1 s x_1} + N_2(s)r_2^3 s^3 e^{-r_2 s x_1} + [(2\mu-T)/(2\mu+T)]$$
$$s^2[N_1(s)r_1 s e^{-r_1 s x_1} + N_2(s)r_2 s e^{-r_2 s x_1}]\}J_1(sx_2)\,ds \quad (38)$$

If $x_1 = 0$, we get from (38)

$$\delta\sigma_{21}^= - [(2\mu+T)/2]$$
$$\times \int_0^\infty s\{N_1(s)r_1^3 s^3 + N_2(s)r_2^3 s^3 + [(2\mu-T)/(2\mu+T)]s^2[N_1(s)r_1 s + N_2(s)r_2 s]\}J_1(sx_2)\,ds$$

Considering $x_1 = 0$, $\delta\sigma_{21} = 0$, one has

$$N_1(s)r_1^3 + N_2(s)r_2^3 + [(2\mu-T)/(2\mu+T)][N_1(s)r_1 + N_2(s)r_2]\} = 0$$
$$N_1(s)r_1^3 + [(2\mu-T)/(2\mu+T)]N_1(s)r_1 + N_2(s)r_2^3 + [(2\mu-T)/(2\mu+T)]N_2(s)r_2 = 0$$
$$N_1(s)[r_1^3 + (2\mu-T)/(2\mu+T)r_1] + N_2(s)[r_2^3 + (2\mu-T)/(2\mu+T)r_2] = 0$$
$$N_2(s) = -N_1(s)[r_1^3 + (2\mu-T)/(2\mu+T)r_1]/[r_2^3 + (2\mu-T)/(2\mu+T)r_2]$$
$$G(s,x_1) = N_1(s)\{e^{-r_1 s x_1} - e^{-r_2 s x_1}[r_1^3 + (2\mu-T)/(2\mu+T)r_1]/[r_2^3 + (2\mu-T)/(2\mu+T)r_2]\}$$
$$G(s,x_1) = N_1(s)(e^{-r_1 x_1} - Ue^{-r_2 s x_1})$$

where

$$U = \frac{r_1^3 + (2\mu-T)/(2\mu+T)r_1}{r_2^3 + (2\mu-T)/(2\mu+T)r_2} = (r_1/r_2)\frac{r_1^2 + (2\mu-T)/(2\mu+T)}{r_2^2 + (2\mu-T)/(2\mu+T)} = (r_1/r_2)\frac{r_1^2 + \beta}{r_2^2 + \beta} \quad (39)$$

and $\beta = (2\mu-T)/(2\mu+T) = r_2^2 r_1^2$.

In view of (35), one has $(2\mu - 8\mu^2 S - T)v_{1,11}/2 + P_{,1} + (2\mu-T)(v_{1,22} + v_{1,2}/r)/2 = 0$. Considering $v_{1=} - (F_{,2} + F/x_2)_{,1}$, one obtains

$$-(2\mu - 8\mu^2 S - T)(F_{,2} + F/x_2)_{,111}/2 + P_{,1} - (2\mu-T)[(F_{,2} + F/x_2)_{,221} + (F_{,2} + F/x_2)_{,21}/x_2]/2 = 0$$
$$P = (2\mu - 8\mu^2 S - T)(F_{,2} + F/x_2)_{,11}/2 + (2\mu-T)[(F_{,2} + F/x_2)_{,22} + (F_{,2} + F/x_2)_{,2}/x_2]/2$$
$$P = (2\mu - 8\mu^2 S - T)(F_{,2} + F/x_2)_{,11}/2 + (2\mu-T)[L(F)_{,2} + L(F)/x_2]/2$$

In view of the (6), one has

$$\delta\sigma_{11} = (2\mu - 4\mu^2 S - T)v_{1,1} + P$$
$$= -(2\mu - 4\mu^2 S - T)(F_{,2} + F/x_2)_{,11} + (2\mu - 8\mu^2 S - T)(F_{,2} + F/x_2)_{,11}/2$$
$$+ (2\mu-T)[L(F)_{,2} + L(F)/r]/2$$
$$= -(2\mu-T)(F_{,2} + F/x_2)_{,11}/2 + (2\mu-T)[L(F)_{,2} + L(F)/r]/2$$
$$= (2\mu-T)[-(F_{,2} + F/x_2)_{,11} + L(F)_{,2} + L(F)/x_2]/2$$
$$\delta\sigma_{11,2} = (2\mu-T)[-(F_{,2} + F/x_2)_{,211} + L(F)_{,22} + L(F)_{,2}/x_2 - L(F)/x_2^2]/2$$
$$= (2\mu-T)[-L(F)_{,11}/2 + L^2(F)]/2$$

Now, we discuss the order-one Hankel transform of the preceding quantity.

$$\int_0^\infty x_2 J_1(sx_2)\delta\sigma_{11,2}\,dx_2$$
$$= (2\mu - T)\int_0^\infty x_2 J_1(sx_2)[-L(F),_{11}/2 + L^2(F)]/2\,dx_2$$
$$= (2\mu - T)\frac{1}{2}\left[-\int_0^\infty x_2 J_1(sx_2)L(F),_{11}\,dx_2 + \int_0^\infty x_2 J_1(sx_2)L^2(F)\,dx_2\right]$$
$$= (2\mu - T)[s^2 G_{11} + s^4 G]/2 = (2\mu - T)s^4 N_1(s)[(r_1^2+1)e^{-r_1 sx_1} - U(r_2^2+1)e^{-r_2 sx_1}]$$

Its Hankel retransform is

$$\delta\sigma_{11,2} = (2\mu - T)\int_0^\infty s J_1(sx_2)s^4 N_1(s)[(r_1^2+1)e^{-r_1 sx_1} - U(r_2^2+1)e^{-r_2 sx_1}]\,ds/2 \tag{40}$$

For $x_1 = 0$, from (40), one has

$$\delta\sigma_{11,2} = (2\mu - T)[(r_1^2+1) - U(r_2^2+1)]\int_0^\infty s J_1(sx_2)s^4 N_1(s)\,ds/2$$

$$\delta\sigma_{11} = (2\mu - T)[(r_1^2+1) - U(r_2^2+1)]\int_0^\infty s \int J_1(sx_2)\,dx_2 s^4 N_1(s)\,ds/2$$

Considering $dJ_0(u)/du = -J_1(u)$, one has

$$\delta\sigma_{11} = (2\mu - T)[U(r_2^2+1) - (r_1^2+1)]\int_0^\infty s J_0(sx_2)s^3 N_1(s)\,ds/2$$

According to (37), one has

$$v_{1,2} = -(F,_2 + F/x_2),_{12} = -L(F),_1 \tag{41}$$

$$\int_0^\infty x_2 J_1(\xi x_2)v_{1,2}\,dx_2 = -\int_0^\infty x_2 J_1(\xi x_2)L(F),_1\,dx_2 = s^2\left[\int_0^\infty x_2 J_1(\xi x_2)F\,dx_2\right]_{,1} = s^2 G,_1 \tag{42}$$

$$v_{1,2} = -\int_0^\infty s J_1(sx_2)s^3 N_1(s)(r_1 e^{-r_1 sx_1} - U r_2 e^{-r_2 sx_1})\,ds \tag{43}$$

$$v_{1,2} = -\int_0^\infty s J_1(sx_2)s^3 N_1(s)(r_1 - U r_2)\,ds \quad v_1 = -\int_0^\infty s \int J_1(sx_2)\,dx_2 s^3 N_1(s)(r_1 - U r_2)\,ds \tag{44}$$

$$v_1 = (r_1 - U r_2)\int_0^\infty s J_0(sx_2)s^2 N_1(s)\,ds \tag{45}$$

Appendix D. The circular cone

Letting $a^4[v_1(x_2)]_{x_1=0} = a^4[b + a\cot\alpha(1-\rho)] = (r_1 - U r_2)g(\rho) = (r_1 - U r_2)(A_0 + A_1\rho)$, one has

$$A_0 = (b + a\cot\alpha)a^4/(r_1 - U r_2), \quad A_1 = -a^5\cot\alpha/(r_1 - U r_2) \quad \rho = x_2/a$$

where α is the angle of the circular cone (the angle between the axis of symmetry and the generatrix).

Considering (29), one obtains from [Gradshteyn and Ryzhik 1965]

$$f(p) = 2(A_0/\pi + A_1/2)\sin p/p + A_1(\cos p - 1)/p^2$$

As $p^3 N_1 = f(p)$ and $(sa) = p$, one has

$$p^3 N_1 = 2(A_0/\pi + A_1/2)\sin p/p + A_1(\cos p - 1)/p^2 \tag{46}$$

$$\delta\sigma_{11} = (2\mu - T)[U(r_2^2 + 1) - (r_1^2 + 1)]\int_0^\infty s J_0(sx_2)s^3 N_1(s)\,ds/2 \tag{47}$$

$$\delta\sigma_{11} = (2\mu - T)[U(r_2^2 + 1) - (r_1^2 + 1)]\int_0^\infty s J_0(sx_2)\left[(2A_0/\pi + A_1)\frac{\sin(sa)}{sa} + A_1\frac{\cos(sa) - 1}{(sa)^2}\right]ds/2 \tag{48}$$

As the integral $\int_0^\infty J_0(p)\sin p\,dp$ is divergence, for the finiteness of stress component $\delta\sigma_{11}$ at the edge of the punch, we have $(2A_0/\pi + A_1) = 0$, that is, $b = a\cot\alpha(\pi/2 - 1)$. Hence

$$v(x_2)_{x_1=x_2=0} = b + a\cot\alpha = 0.5\pi a\cot\alpha$$

Because $\int_0^\infty J_0(pp)(\cos p - 1)/p\,dp = -\cosh^{-1}(1/\rho)$, one has

$$\delta\sigma_{11} = (2\mu - T)[U(r_2^2 + 1) - (r_1^2 + 1)]A_1 a^{-5}\int_0^\infty sa J_0(sax_2/a)[(\cos(sa) - 1)/(sa)^2]\,dsa/2 \tag{49}$$

$$\delta\sigma_{11} = -(2\mu - T)[U(r_2^2 + 1) - (r_1^2 + 1)]a^{-5}A_1\cosh^{-1}(a/x_2)/2 \tag{50}$$

As $\int_0^a[\cosh^{-1}(a/x_2)x_2]\,dx_2 = 0.5a^2$ and $A_1 = -a^5\cot\alpha/(r_1 - Ur_2)$, one obtains for the compressive force R on the cone

$$R = -2\pi\int_0^a[\delta\sigma_{11}]_{x_1=o}x_2\,dx_2 = 2\pi(2\mu - T)[U(r_2^2 + 1) - (r_1^2 + 1)]a^{-5}A_1\int_0^a[\cosh^{-1}(a/x_2)x_2]\,dx_2/2$$

$$= \pi(2\mu - T)[U(r_2^2 + 1) - (r_1^2 + 1)]a^{-3}A_1/2 = \pi(2\mu - T)[(r_1^2 + 1) - U(r_2^2 + 1)]a^2\frac{\cot\alpha}{2r_1 - 2Ur_2}$$

From (39), we know that $U = (r_1/r_2)\dfrac{r_1^2 + \beta}{r_2^2 + \beta}$, where $\beta = (2\mu - T)/(2\mu + T) = r_2^2 r_1^2$. Then, one has

$$(r_1 - Ur_2) = r_1\{1 - [r_1^2 + r_2^2 r_1^2]/[r_2^2 + r_2^2 r_1^2]\} = r_1\{r_2^2 - r_1^2\}/[r_2^2 + r_2^2 r_1^2] \tag{51}$$

and

$$(r_1^2 + 1 - Ur_2^2 - U) = \{r_1^2 + 1 - (r_1 r_2)[r_1^2 + r_2^2 r_1^2]/[r_2^2 + r_2^2 r_1^2] - (r_1/r_2)[r_1^2 + r_2^2 r_1^2]/[r_2^2 + r_2^2 r_1^2]\}$$

$$= \frac{(r_1^2 + 1)[r_2^2 + r_2^2 r_1^2] - (r_1 r_2)[r_1^2 + r_2^2 r_1^2] - (r_1/r_2)[r_1^2 + r_2^2 r_1^2]}{r_2^2 + r_2^2 r_1^2}$$

$$= \frac{r_1^2 r_2^2 + r_2^2 r_1^2 - (r_1 r_2)[r_1^2 + r_2^2 r_1^2]}{r_2^2 + r_2^2 r_1^2} + \frac{[r_2^2 + r_2^2 r_1^2] - (r_1/r_2)[r_1^2 + r_2^2 r_1^2]}{r_2^2 + r_2^2 r_1^2}$$

$$= \frac{r_1[r_1 r_2(r_2 - r_1) + (r_1 - r_2)r_2^2 r_1^2] + [r_2^2 - (r_1/r_2)r_1^2] + r_2^2 r_1^2 - (r_1/r_2)r_2^2 r_1^2}{r_2^2 + r_2^2 r_1^2}$$

$$= (r_2 - r_1)\{r_1[r_1 r_2 - r_2^2 r_1^2] + [r_2^2 + r_1 r_2 + r_1^2]/r_2 + r_2^2 r_1^2/r_2\}/[r_2^2 + r_2^2 r_1^2]$$

$$= (r_2 - r_1)r_2 r_1 r_1 r_2 - r_2^2 r_1^2 + [r_2^2 + r_1 r_2 + r_1^2] + r_2^2 r_1^2\}/[(r_2^2 + r_2^2 r_1^2)r_2] \tag{52}$$

Considering the equations (51) and (52), one knows

$$\frac{(r_1^2+1)-U(r_2^2+1)}{r_1-Ur_2} = \frac{(r_2-r_1)[r_2r_1(r_1r_2-r_2^2r_1^2)+(r_2^2+r_1r_2+r_1^2)+r_2^2r_1^2]/[(r_2^2+r_2^2r_1^2)r_2]}{r_1\{r_2^2-r_1^2\}/(r_2^2+r_2^2r_1^2)}$$

$$= \{-r_2^3r_1^3+(r_2^2+r_1r_2+r_1^2)+2r_2^2r_1^2\}/r_2r_1(r_2+r_1)$$

$$= \{-r_2^3r_1^3+[r_2^2+r_1r_2+r_1^2]+2r_2^2r_1^2\}/[r_2r_1(r_2^2+2r_1r_2+r_1^2)^{1/2}]$$

Considering r_1 and r_2 are the two positive roots of the equation $p_1+p_2m^2/s^2+p_3m^4/s^4=0$, where $p_3=(1+Q)$, $p_2=-(2-6\mu S)$, $p_1=(1-Q)$ and $Q=T/2\mu$, one has

$$r_2r_1=(p_1/p_3)^{1/2}=(1-Q)^{1/2}/(1+Q)^{1/2}=(2\mu-T)^{1/2}/(2\mu+T)^{1/2},$$

$$r_2^2+r_1^2=-p_2/p_3=(2-6\mu S)/(1+Q)=(2\mu-6\mu^2S)/(2\mu+T)$$

where S can be found from (3). Then, one obtains (30). The contact radius a is then given by (31) and the penetration depth by (32). Finally, from the expression for a and the values $(r_2^2+r_1^2)=-p_2/p_3$, $r_1r_2=(p_1/p_3)^{1/2}$, $\beta=(p_1/p_3)$ and $T/2\mu=Q$, one obtains

$$\pi a^2=(R/2\mu)\{(p_1/p_3)^{1/2}[2(p_1/p_3)^{1/2}-p_2/p_3]^{1/2}\}$$

$$/\{(1-Q)[(p_1/p_3-p_1^{3/2}/p_3^{3/2})+(p_1/p_3)^{1/2}-p_2/p_3+p_1/p_3]\cot\alpha\} \quad (53)$$

Substituting $-p_2/p_3=+2(1-3\mu S)/(1+Q)$, $(p_1/p_3)=(1-Q)/(1+Q)$, $(p_1/p_3)^{1/2}=(1-Q)^{1/2}/(1+Q)^{1/2}$ into (53), one has

$$\pi a^2=(R/2\mu)tg\alpha\{[1/(1-Q^2)^{1/2}][2(1-Q)^{1/2}/(1+Q)^{1/2}+2(1-3\mu S)/(1+Q)]^{1/2}\}$$

$$/\{(1-Q)^{1/2}/(1+Q)^{1/2}[(1-Q)^{1/2}/(1+Q)^{1/2}-(1-Q)/(1+Q)]$$

$$+[(1-Q)^{1/2}/(1+Q)^{1/2}+2(1-3\mu S)/(1+Q)]+(1-Q)/(1+Q)\} \quad (54)$$

References

[Dorris and Nemat-Nasser 1980] J. F. Dorris and S. Nemat-Nasser, "Instability of a layer on a half space", *J. Appl. Mech.* (*ASME*) **47** (1980), 304–312.

[Gradshteyn and Ryzhik 1965] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series and products*, Scripta Technica, Academic Press, New York, 1965.

[Hao 2006] T. H. Hao, "Axially symmetrical contact problem of finite elasticity and its application to estimating residual stresses by cone indentation", *J. Mech. Mater. Struct.* **1**:8 (2006), 1367–1384.

[Hao and Shen 1994] T. H. Hao and Z. Y. Shen, "A new electric boundary condition of electric fracture mechanics and its application", *Eng. Fract. Mech.* **47**:6 (1994), 793–802.

[Sneddon 1951] I. N. Sneddon, *Fourier transforms*, McGraw-Hill, New York, 1951.

[Suresh and Giannakopoulos 1998] S. Suresh and A. E. Giannakopoulos, "A new method for estimating residual stresses by instrumented sharp indentation", *Acta Mater.* **46** (1998), 5755–5767.

TIAN-HU HAO: haoth0000@yahoo.com.cn
State Key Lab for Modification of Polymer Materials and Chemical Fibers, Donghua University, P.O. Box 220, Shanghai, 200051, China

msp

FORM FINDING OF TENSEGRITY STRUCTURES USING FINITE ELEMENTS AND MATHEMATICAL PROGRAMMING

KATALIN K. KLINKA, VINICIUS F. ARCARO AND DARIO GASPARINI

We show that the minimization of total potential energy is the general principle behind the well-known rule of maximizing some lengths of a truss mechanism to define a tensegrity. Moreover, the latter rule is a special case, due to the usual high values of the modulus of elasticity. An innovative mathematical model is presented for finding the form of tensegrity structures, based on the finite element method and on mathematical programming. A special line element that shows constant stress for any displacement of its nodes is used to define a prestressed equilibrium configuration. Form finding is formulated as an unconstrained nonlinear programming problem, where the objective function is the total potential energy and the displacements of the nodal points are the unknowns. A connection is made with the geometric shape minimization problem, defined by a constrained nonlinear programming problem. A quasi-Newton method is used, which avoids the evaluation of the tangent stiffness matrix.

1. Introduction

Maxwell [1864] wrote: "In those cases in which stiffness can be produced with a smaller number of lines, certain conditions must be fulfilled, rendering the case one of a maximum or minimum value of one or more of its lines. The stiffness of such frames is of an inferior order, as a small disturbing force may produce a displacement infinite in comparison with itself". In [Calladine 1978], the author who made the connection between tensegrity structures and the exceptions to Maxwell's rule writes that presumably Maxwell intended to refer to a maximum or minimum value of the length of one or more of its lines. An explanation for Maxwell's obscure remark about maximum or minimum values based on the principle of minimum total potential energy is presented. A review of the important literature related to form finding methods for tensegrity structures is given in [Tibert and Pellegrino 2003] and more recently in [Hernández Juan and Mirats Tur 2008]. These methods can be classified into kinematical and statical methods. This text concentrates on the total potential energy minimization method for form finding. A special line element that shows constant stress for any displacement of its nodes is used to define a prestressed equilibrium configuration. The form finding is formulated as an unconstrained nonlinear programming problem, where the objective function is the total potential energy and the displacements of the nodal points are the unknowns. Another approach, which minimizes the total potential energy by modifying the lengths of selected elements, is described in [Pagitz and Miratz Tur 2009]. A quasi-Newton method is used, which avoids the evaluation of the tangent stiffness matrix. An interesting connection is made between minimizing the total potential energy, which is defined by an unconstrained nonlinear programming problem, and the geometric shape minimization problem, which is defined by a constrained nonlinear programming problem. The strain energy for a line element can be interpreted as

Keywords: cable, element, line, minimization, nonlinear, optimization, tensegrity.

a penalty function, as it imposes resistance for changing the length of the element. The total potential energy minimization method for the analysis of cable structures was first described in [Pietrzak 1978]. The following conventions apply unless otherwise specified or made clear by the context. A Greek letter expresses a scalar. A lower case letter represents a column vector.

2. Line element definition

Figure 1 shows the geometry of the element. The nodes are labeled 1 and 2; a superscript associates a variable with its corresponding node. The nodal displacements, given by vectors u, transform the element from its initial configuration to its final configuration. The strain is assumed constant along the element.
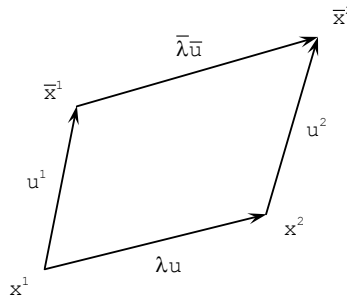
Figure 1. Line element.

3. Engineering strain

The vector u is a unity vector. Note that λ represents the undeformed length of the element. The nodal displacements vectors are numbered according to its node numbers. The deformed length can be written as follows:

$$z = \frac{u^2 - u^1}{\lambda} \qquad \Rightarrow \qquad \bar{\lambda}\bar{u} = \lambda u + \lambda z, \tag{1}$$

$$\delta = 2u^T z + z^T z \qquad \Rightarrow \qquad \bar{\lambda} = \lambda\sqrt{1+\delta}. \tag{2}$$

The unit vector parallel to the element in its final configuration can be written as

$$\bar{u} = \frac{u+z}{\sqrt{1+\delta}}. \tag{3}$$

The engineering strain can be written as

$$\epsilon = \frac{\bar{\lambda} - \lambda}{\lambda}. \tag{4}$$

Inaccuracy often results from severe cancellation that occurs when nearly equal values are subtracted [Goldberg 1991]. In order to avoid it, the previous expression should be evaluated as

$$\epsilon = \frac{\delta}{\sqrt{1+\delta}+1}. \tag{5}$$

4. Variable stress element

Considering σ as the conjugate stress to the engineering strain ϵ and α as the undeformed area of the element, the potential strain energy and its gradient with respect to the nodal displacements can be written as

$$\phi = \alpha\lambda \int_0^\epsilon \sigma(\xi)\, d\xi, \tag{6}$$

$$\frac{\partial \phi}{\partial u_i^1} = -\alpha\sigma(\epsilon)\bar{u}_i, \quad \frac{\partial \phi}{\partial u_i^2} = +\alpha\sigma(\epsilon)\bar{u}_i. \tag{7}$$

The gradient can be interpreted as internal forces acting on nodes of the element.

Stress and strain. Consider stress as a linear function of strain with E as the modulus of elasticity. The potential strain energy can be written as

$$\phi = \frac{\alpha E (\bar{\lambda} - \lambda)^2}{2\lambda}. \tag{8}$$

The strain energy can be interpreted as a penalty function with the modulus of elasticity as the penalty parameter. The modulus of elasticity, which is usually large, imposes resistance to changing the length of the elements.

5. Constant stress element

A constant stress element can be defined by imposing a constant stress σ. The potential strain energy can be written as

$$\phi = \alpha\lambda \int_0^\epsilon \sigma\, d\xi = \alpha\sigma(\bar{\lambda} - \lambda). \tag{9}$$

The potential strain energy is equal to the force multiplied by the relative displacement between the nodes. In the expression for the strain energy, the undeformed length can be eliminated because it does not depend on the nodal displacements. Its permanence in the expression would only add constants, one for each element, to the total potential strain energy function. To minimize a function plus a constant is equivalent to minimize the function only. Therefore, the potential strain energy can be replaced by

$$\phi = \alpha\sigma\bar{\lambda}. \tag{10}$$

The strain energy is simply the final length of the element multiplied by the imposed constant force. The gradient with respect to the nodal displacements can be written as

$$\frac{\partial \phi}{\partial u_i^1} = -\alpha\sigma\bar{u}_i, \quad \frac{\partial \phi}{\partial u_i^2} = +\alpha\sigma\bar{u}_i. \tag{11}$$

Note that the components of the gradient are given by a scalar (stress multiplied by area) multiplying the unit vector parallel to the element in its final configuration. The gradient can be interpreted as internal forces with constant modulus acting on nodes of the element. The element shows constant stress for any displacement of its nodes. A similar element was described in [Meek 1971]. The element was called variable initial length element.

6. Form finding

The initial configuration of a tensegrity structure is defined as the configuration of zero nodal displacements for all its nodes. A form finding strategy can be defined as: Starting with an initial configuration, select some elements as constant stress elements by specifying a stress value. Find the prestressed equilibrium configuration by minimizing the total potential strain energy.

7. Equilibrium configuration

Considering C as the set of constant stress elements, V as the set of variable stress elements and u as the vector of unknown displacements, the total potential strain energy function and its gradient can be written as

$$\pi(u) = \sum_{e \in C} \phi_e + \sum_{e \in V} \phi_e, \tag{12}$$

$$\nabla \pi(u) = \sum_{e \in C} \nabla \phi_e + \sum_{e \in V} \nabla \phi_e. \tag{13}$$

The stable equilibrium configurations correspond to local minimum points of the total potential energy function, which in the absence of external forces reduces to the total potential strain energy function. In order to find the local minimum points of a nonlinear multivariate function, the general strategy that can be used is: Choose a starting point and move in a given direction such that the function decreases. Find the minimum point in this direction and use it as a new starting point. Continue this way until a local minimum point is reached. The problem of finding the minimum points of a nonlinear multivariate function is replaced by a sequence of sub problems, each one consisting of finding the minimum of a univariate nonlinear function. In the quasi-Newton methods, starting with the unit matrix, a positive definite approximation to the inverse of the Hessian matrix is updated at each iteration. This update is made using only values of the gradient vector. A direction such that the function decreases is calculated as minus the product of this approximation of the inverse of the Hessian matrix and the gradient vector calculated at the starting point of each iteration. Consequently, it is not necessary to solve any system of equations. Moreover, the analytical derivation of an expression for the Hessian matrix is not necessary. Note that by minimizing the total potential energy function it is almost impossible to find an unstable equilibrium configuration, which corresponds to a local maximum point. The only exception is that it is possible to find a saddle point, that is, the point is a local minimum and also a local maximum. However, even in this improbable situation, a direction of negative curvature to continue toward a local minimum point can be found as described in [Gill and Murray 1974]. It is important to emphasize that minimizing total potential energy to find equilibrium configurations does not require support constraints to prevent rigid body motion. The computer code uses the limited memory BFGS to tackle large-scale problems as described in [Nocedal and Wright 2006]. It also employs a line search procedure through cubic interpolation [loc. cit.].

7.1. *Geometrical shape minimization.* Due to the fact that the modulus of elasticity is usually a big number, the problem of minimizing the total potential strain energy can be interpreted as an equality constrained nonlinear programming problem converted to an unconstrained nonlinear programming problem

by the quadratic penalty method. This interpretation leads to an extension of the mathematical model for geometric shape minimization described in [Arcaro and Klinka 2009].

Special case 1: A structure with the same modulus of elasticity for all elements, area equal to 1 for all elements and stress equal to 1 (tension) for all constant stress elements. Minimizing the total potential strain energy can be interpreted as minimizing the sum of the lengths of the constant stress elements while keeping the lengths of the variable stress elements.

$$\text{Min } \pi(u) = + \sum_{e \in C} \bar{\lambda} + \frac{E}{2} \sum_{e \in V} \frac{(\bar{\lambda} - \lambda)^2}{\lambda} \tag{14}$$

Special case 2: A structure with the same modulus of elasticity for all elements, area equal to 1 for all elements and stress equal to -1 (compression) for all constant stress elements. Minimizing the total potential strain energy can be interpreted as maximizing the sum of the lengths of the constant stress elements while keeping the lengths of the variable stress elements.

$$\text{Min } \pi(u) = - \sum_{e \in C} \bar{\lambda} + \frac{E}{2} \sum_{e \in V} \frac{(\bar{\lambda} - \lambda)^2}{\lambda} \tag{15}$$

8. Examples

Elements shown in red are in compression. Elements shown in blue are in tension. Constant stress elements are shown in green in the initial configuration.

Example 1. A straight prismoid with height = 3. The bottom and top regular triangles are inscribed in a circle of radius = 1. It is composed by 3 constant stress elements and 9 variable stress elements.

Special case 1: Figure 2 shows the initial shape on the left, the final shape with $E = 1000$ on the center and the final shape with $E = 10$ on the right. The constant stress elements are shown in blue in the final configuration. The top triangle rotates 150 degrees clockwise relative to the bottom triangle.

Special case 2: Figure 3 shows the initial shape on the left, the final shape with $E = 1000$ on the center and the final shape with $E = 10$ on the right. The constant stress elements are shown in red in the final
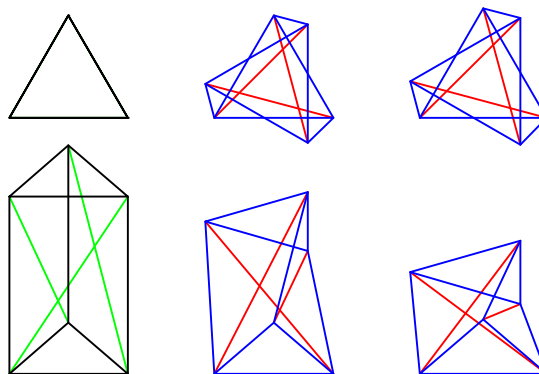
Figure 2. Initial and prestressed configurations for Example 1, special case 1.
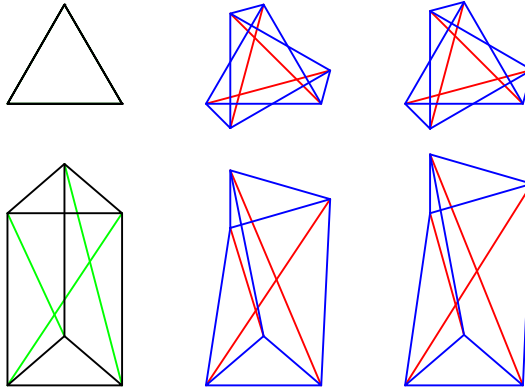
Figure 3. Initial and prestressed configurations for Example 1, special case 2.

	Initial	$E = 1000$	$E = 10$
$\sigma = +1$	3.4641	2.3473	1.5184
$\sigma = -1$	3.4641	3.5329	3.7782

Table 1. Initial and prestressed configurations for Example 1.

configuration. The top triangle rotates 30 degrees counterclockwise relative to the bottom triangle.

Table 1 shows the lengths of the constant stress elements in the initial and final configurations.

Example 2. Figure 4 shows the geometry of a sculpture called the stella octangula, which was proposed by Hungarian architect, sculptor and author David Georges Emmerich. An extensive description of his works is given in [Chassagnoux 2006]. An analysis of this structure, using the dynamic relaxation method, is described in [Motro 2011]. Recently, a modified dynamic relaxation algorithm for the analysis of tensegrity structures was proposed in [Ali et al. 2011].

The geometry consists of 18 elements with length equal to s and 6 diagonal elements with length equal to $s\sqrt{3}$. The connectivity of the diagonal elements is also shown in Figure 4. The coordinates of

Elem	Node	Node
3	4	11
6	5	10
9	6	12
12	7	1
15	8	3
18	9	2

Figure 4. The stella octangula: geometry and connectivity of diagonal elements.

Node	Coord-X	Coord-Y	Coord-Z
1	$-s/2$	$-r/2$	h
2	$s/2$	$-r/2$	h
3	0	r	h
4	0	$-2r$	h
5	s	r	h
6	$-s$	r	h
7	s	$-r$	$-h$
8	$-s$	$-r$	$-h$
9	0	$2r$	$-h$
10	0	$-r$	$-h$
11	$-s/2$	$r/2$	$-h$
12	$s/2$	$r/2$	$-h$

Table 2. Coordinates of the vertices.

the vertices appear in Table 2, where we have set

$$r = \frac{s}{\sqrt{3}}, \quad h = \frac{s}{\sqrt{6}}. \tag{16}$$

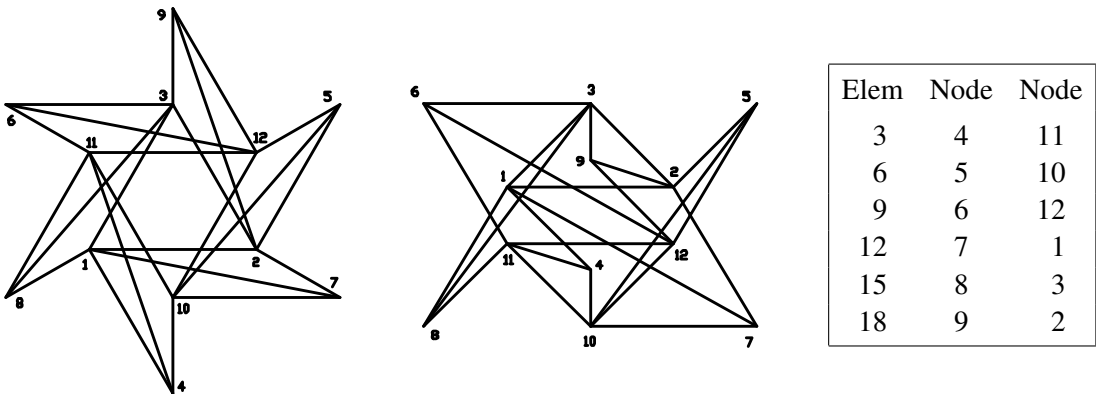A stella octangula with parameter $s = 1$, $E = 1000$ and all elements with area = 1. There are support constraints on nodes 1, 2 and 3 to prevent rigid body motion. According the definition given in [Zhang et al. 2006], a regular tensegrity is characterized by equal length for the elements in tension and by equal length for the elements in compression. A nonregular tensegrity can be generated by imposing different stress values for selected elements of a regular tensegrity. The regular tensegrity can be recovered by imposing equal stress values for the same selected elements on the previously generated nonregular tensegrity. Another approach to generate a nonregular tensegrity, which is based on the dynamic relaxation method, is presented in [Tibert and Pellegrino 2003].

Generating the nonregular tensegrity: The stress values for the diagonal elements of the regular stella octangula and the lengths for the diagonal elements of its prestressed configuration are shown in the left half of Table 3.

Elem	Stress	Length		Elem	Stress	Length
3	-1.25	1.4573		3	-1.00	1.7343
6	-1.50	1.5664		6	-1.00	1.7345
9	-1.75	1.6312		9	-1.00	1.7348
12	-2.00	1.8578		12	-1.00	1.7351
15	-2.25	1.8899		15	-1.00	1.7353
18	-2.50	1.8914		18	-1.00	1.7357

Table 3. Nonregular (left) and regular (right) stella octangula.
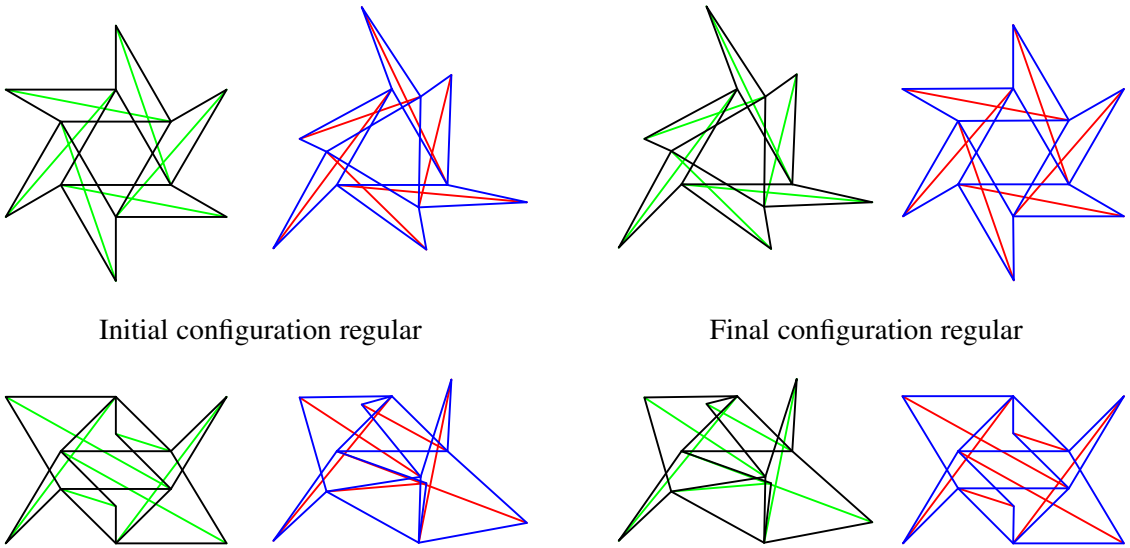
Figure 5. Stella octangula: Initial and prestressed configurations when the initial configuration is regular (left block), and when the final configuration is regular (right block).

Figure 5 shows the initial configuration (regular stella octangula) in the first column and its prestressed configuration (nonregular stella octangula) in the second column.

Recovering the regular tensegrity: We can also apply the procedure using as input the nonregular stella octangula obtained in the example immediately above. The stress values for the diagonal elements of the nonregular stella octangula and the lengths for the diagonal elements of its prestressed configuration are shown in the right half of Table 3.

Figure 5 shows the initial configuration (nonregular stella octangula) in the third column its prestressed configuration (regular stella octangula) in the last column.

Conclusions

The principle of minimum total potential energy is a fundamental one in physics, and lies at the basis of the mathematical model presented here for form finding of tensegrity structures, using the finite element method and mathematical programming. The proposed approach can generate a nonregular tensegrity starting from a regular tensegrity. Additionally, the regular tensegrity can be recovered by imposing equal stresses on the previously selected constant stress elements of the nonregular tensegrity. The use of a quasi-Newton method to minimize the total potential energy function has several advantages over solving the equilibrium equations in nonlinear mechanics: It allows the analysis of under constrained structures even without support constraints to prevent rigid body motion. It is not necessary to derive the tangent stiffness matrix. It is not necessary to solve any system of equations. It can handle large-scale problems with efficiency.

References

[Ali et al. 2011] N. B. H. Ali, L. Rhode-Barbarigos, and I. F. C. Smith, "Analysis of clustered tensegrity structures using a modified dynamic relaxation algorithm", *Int. J. Solids Struct.* **48**:5 (2011), 637–647.

[Arcaro and Klinka 2009] V. F. Arcaro and K. K. Klinka, "Finite element analysis for geometrical shape minimization", *J. IASS* **50**:2 (2009), 79–86.

[Calladine 1978] C. R. Calladine, "Buckminster Fuller's 'tensegrity' structures and Clerk Maxwell's rules for the construction of stiff frames", *Int. J. Solids Struct.* **14**:2 (1978), 161–172.

[Chassagnoux 2006] A. Chassagnoux, "David Georges Emmerich: professor of morphology", *Int. J. Space Struct.* **21**:1 (2006), 59–71.

[Gill and Murray 1974] P. E. Gill and W. Murray, "Newton-type methods for unconstrained and linearly constrained optimization", *Math. Program.* **7**:1 (1974), 311–350.

[Goldberg 1991] D. Goldberg, "What every computer scientist should know about floating-point arithmetic", *ACM Comput. Surv.* **23**:1 (1991), 5–48.

[Hernández Juan and Mirats Tur 2008] S. Hernández Juan and J. M. Mirats Tur, "Tensegrity frameworks: static analysis review", *Mech. Mach. Theory* **43**:7 (2008), 859–881.

[Maxwell 1864] J. C. Maxwell, "On the calculation of the equilibrium and stiffness of frames", *Philos. Mag.* (6) **27**:182 (1864), 294–299.

[Meek 1971] J. L. Meek, *Matrix structural analysis*, McGraw-Hill, New York, 1971.

[Motro 2011] R. Motro, "Structural morphology of tensegrity systems", *Meccanica* (*Milano*) **46**:1 (2011), 27–40.

[Nocedal and Wright 2006] J. Nocedal and S. J. Wright, *Numerical optimization*, 2nd ed., Springer, New York, 2006.

[Pagitz and Miratz Tur 2009] M. Pagitz and J. M. Miratz Tur, "Finite element based form-finding algorithm for tensegrity structures", *Int. J. Solids Struct.* **46**:17 (2009), 3235–3240.

[Pietrzak 1978] J. Pietrzak, "Matrix formulation of static analysis of cable structures", *Comput. Struct.* **9**:1 (1978), 39–42.

[Tibert and Pellegrino 2003] A. G. Tibert and S. Pellegrino, "Review of form-finding methods for tensegrity structures", *Int. J. Space Struct.* **18**:4 (2003), 209–223.

[Zhang et al. 2006] L. Zhang, B. Maurin, and R. Motro, "Form-finding of nonregular tensegrity systems", *J. Struct. Eng.* (*ASCE*) **132**:9 (2006), 1435–1440.

KATALIN K. KLINKA: katalin@klinka.hu
Department of Structural Engineering, Budapest University of Technology and Economics, 2 Bertalan Lajos, Budapest, 1111, Hungary

VINICIUS F. ARCARO: vinicius.arcaro@arcaro.org
Institute for Membrane and Shell Technologies, Bauhausstrasse 8, D-06846 Dessau, Germany

DARIO GASPARINI: dag6@cwru.edu
Civil Engineering Department, Case Western Reserve University, 2182 Adelbert Rd, Cleveland OH 44106, United States

msp

EXPERIMENTAL AND ANALYTICAL INVESTIGATION OF THE BEHAVIOR OF DIAPHRAGM-THROUGH JOINTS OF CONCRETE-FILLED TUBULAR COLUMNS

RONG BIN, CHEN ZHIHUA, ZHANG RUOYU, APOSTOLOS FAFITIS AND YANG NAN

Static tensile loading experiments and nonlinear finite element analysis were carried out to study the mechanical properties and failure modes of diaphragm-through joints of concrete-filled square steel tubular columns. Comparison between experimental data and finite element analysis revealed that the FE predictions of failure modes, load-displacement curves and bearing capacity agree with the test results. It was found that the tensile load from the steel beam flange is mainly shared by the square steel tube and the diaphragm. The plastic zone of the tube appears around the intersections of the tube and the diaphragm whereas the diaphragm plastic zone appears along the cross-section lines enclosed by the steel square tube. Calculation models of yield lines on square steel tube and diaphragm are established based on distribution pattern of plastic zone, and an analytical method for the design of such joints is proposed as well. The experimentally obtained bearing capacities of the tested specimens are in good agreement with the analytically computed capacities.

1. Introduction

Frame structures with concrete-filled square steel tubular columns and H-shaped steel beams have been increasingly used in structural engineering. Inner diaphragm joints and diaphragm-through joints, shown in Figure 1, are two forms of rigid beam-column connections that are used in many applications. In the diaphragm-through method the square tube is disconnected and two diaphragms are placed at the level of the upper and lower flange of the steel beam. The beam is then connected to these diaphragms.

For the inner diaphragm joints, a large number of experimental and theoretical studies [Park et al. 2010; Han et al. 2003; 2008; Fukumoto and Morita 2005; Ricles et al. 2004] have been carried out and calculation methods and formulae [Sasaki et al. 1995; Nie et al. 2005; Zhou et al. 2005] are proposed for the flexural bearing capacity and the shear bearing capacity of inner diaphragm joints. However, for diaphragm-through joints, there is relatively limited research focusing mainly on the seismic performance. Nishiyama et al. [2004] carried low-reversed cyclic loading experiments on 10 specimens of diaphragm-through joints to investigate the failure process. Chen et al. [2006] compared the hysteretic behaviors of inner diaphragm joints and diaphragm-through joints based on low-reversed cyclic loading experimental results and Jiang et al. [2009] studied ductility, energy dissipation and rigidity degradation of diaphragm-through joints by low-reversed cyclic loading experiments.

The flexural bearing capacity is calculated based on the assumption that it depends on the tensile capacity of the panel zone as shown in Figure 2 [Matsui 1985; Morino and Tsuda 2002], and it can be
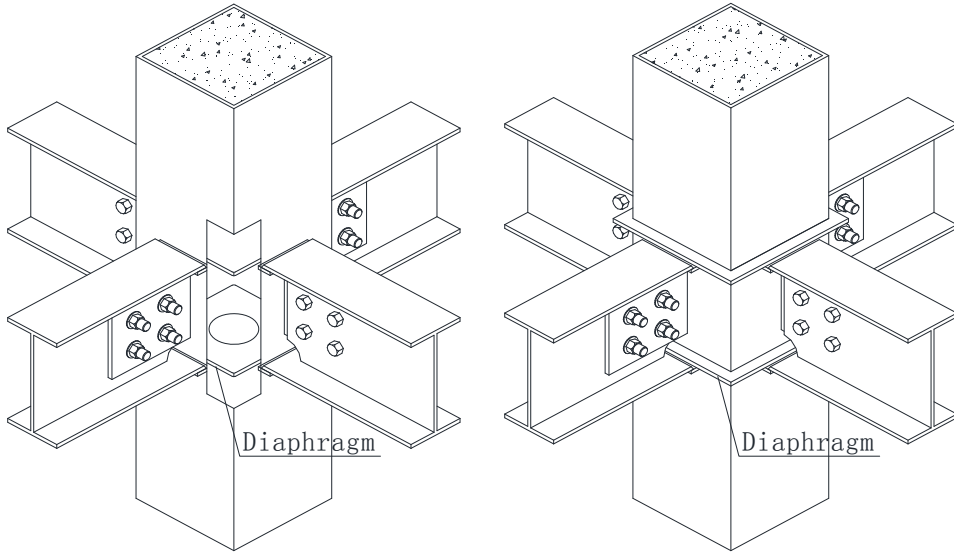
Figure 1. Joint forms: inner diaphragm joint (left) and diaphragm-through joint (right).
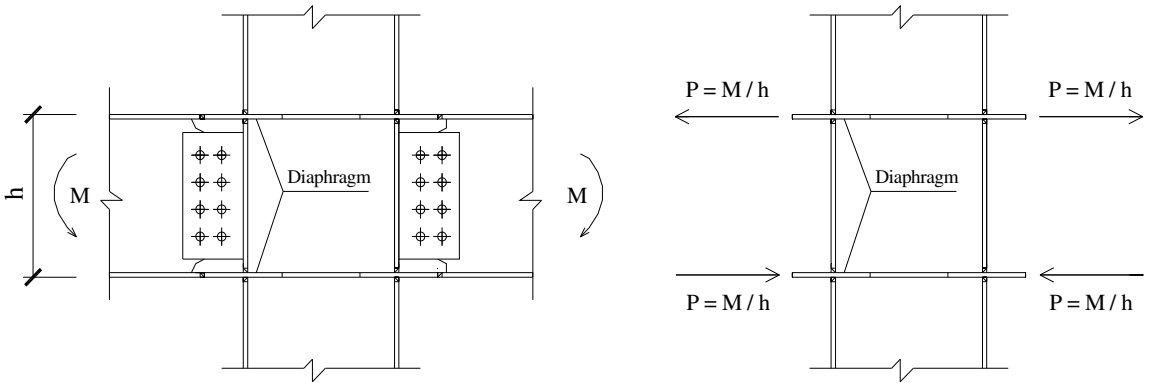
Figure 2. Calculation model for flexural bearing capacity of diaphragm-through joints.

expressed as

$$M = P(h - t) \tag{1}$$

where P is the tensile capacity of the panel zone, h is the distance between outer faces of the flanges and t is the thickness of the flange.

In current practice [Fukumoto and Morita 2005] the tensile capacity is evaluated based on two assumed yielding mechanisms of the panel zone. In the first it is assumed that section 1-1, shown in Figure 3, is the yield location and then the yield strength of panel zone is

$$P_{y1}^d = (D + 2h_d - d)t_d f_y^d \tag{2}$$
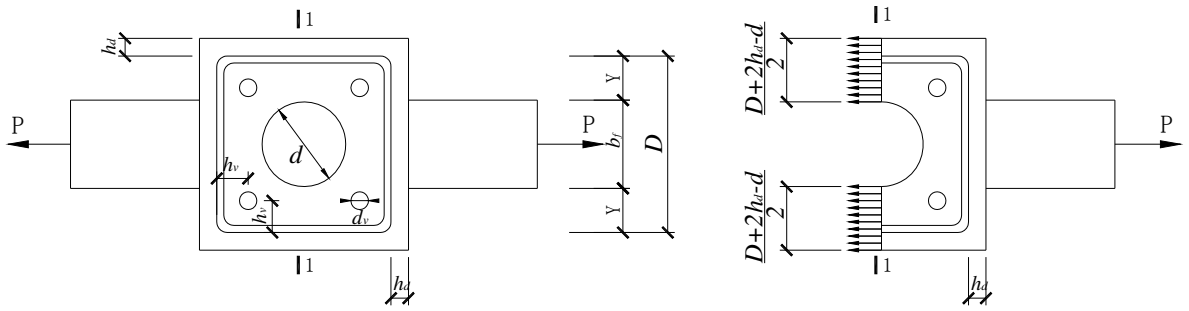
Figure 3. Yield mechanism 1 of diaphragm-through joint.
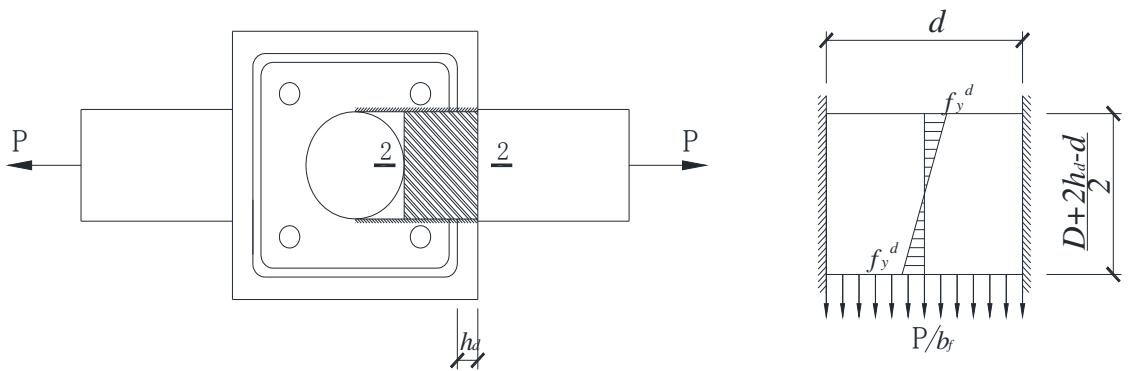
Figure 4. Yield mechanism 2 of diaphragm-through joint.

where D is the width of the square steel tube, h_d is the cantilevered length of the diaphragm, d is the diameter of the central opening for concrete casting, t_d is the thickness of the diaphragm and f_y^d is the yield strength.

In the second mechanism the part of the diaphragm near the central opening is treated as a fixed beam under uniformly distributed load as shown in Figure 4. The width of the beam cross-section equals the thickness of the diaphragm, shown in Figure 4. Assuming that yielding will take place at section 2-2, the yield strength is

$$P_{y2}^d = (D + 2h_d - d)^2 \frac{b_f \cdot t_d}{d^2} f_y^d \tag{3}$$

where b_f is the width of the flange. Note that (3) is a simplifying approximation ignoring the fact that this beam is usually a deep beam.

From (2) and (3), the yield tensile capacity of the panel zone can be expressed as

$$P_y^d = \min(P_{y1}^d, P_{y2}^d). \tag{4a}$$

Substituting the ultimate strength f_u^d of the diaphragm for the yield strength f_y^d of (4a), the ultimate tensile capacity of the panel zone can be expressed as

$$P_u^d = \min(P_{u1}^d, P_{u2}^d). \tag{4b}$$

No experimental data of static tensile loading for diaphragm-through joints were found in the literature. The above mentioned calculation methods for the tensile capacity of diaphragm-through joints ignore the contribution of the tensile capacity of steel tube and the concrete infill. Therefore the yielding mechanism models and the correctness of the results need to be verified by static tensile loading experiments.

The objectives of this investigation are to study the maximum load-bearing capacity and mechanical properties of diaphragm-through joint subjected to static tensile load. The analytical method mentioned above for bearing capacity of diaphragm-through joint under static tensile load is to be examined as well. The main parts of this paper are: firstly to report the experimental results on 8 diaphragm-through joints under static tensile load, secondly to analyze the mechanisms of the diaphragm-through joints subjected to static tensile load by using finite element analysis and thirdly, to establish the yield mechanism model of the joint and to develop an analytical method for the determination of the bearing capacity of diaphragm-through joints under static tensile load.

2. Experimental study

Materials.

(1) Steel tube

Two types of square hollow steel tubes are used to manufacture the specimens. The dimensions of these steel tubes were 250 mm × 250 mm. The thickness was 8 mm and 10 mm respectively. To determine the steel properties of these square tubes, tension coupons are cut and tested in tension. The properties of the tubes are shown in Table 1.

(2) Diaphragm and flange

Two types of steel plates are used to manufacture the diaphragms and the beam flanges with thicknesses 10 mm and 12 mm. The properties of the steel plates are shown in Table 2.

(3) Concrete

Two types of concrete marked by C20 and C40 are used. To determine the concrete material properties, $150 \times 150 \times 150$ mm cubes were cast and cured in conditions similar to that of the experiment and tested in compression. The concrete properties are shown in Table 3.

Dimension (mm)	Thickness (mm)	Yield strength f_y (N/mm^2)	Ultimate strength f_u (N/mm^2)	Modulus of elasticity E (10^5 N/mm^2)	Elongation δ (%)
250 × 250	8	284.7	409.8	2.036	33.6
	10	316.1	463.3	2.180	38.7

Table 1. Material properties of steel tubes.

Thickness (mm)	Yield strength f_y (N/mm^2)	Ultimate strength f_u (N/mm^2)	Modulus of elasticity E (10^5 N/mm^2)	Elongation δ (%)
10	302.8	421.4	1.952	37.4
12	305.2	429.8	2.167	34.6

Table 2. Material properties of steel plates.

| Concrete | Mix proportion (kg/m^3) | | | Average crushing strength (MPa) | Modulus of elasticity ($\times 10^4$ MPa) |
	Cement	Sand	Coarse aggregate		
C20	350	690	1160	25.3	3.46
C40	460	720	1036	47.6	3.73

Table 3. Material properties of concrete.
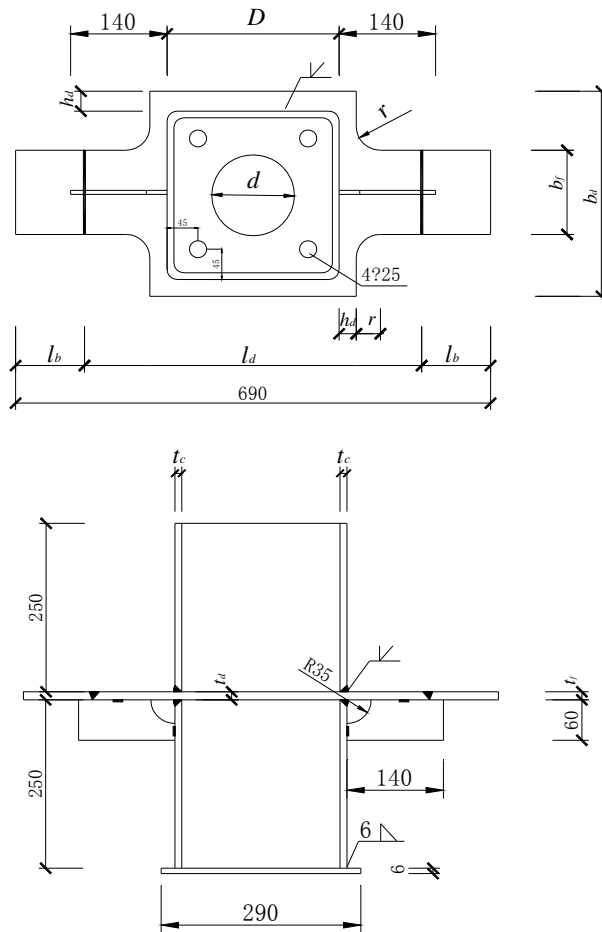
Figure 5. Specimen of tensile loading experiment.

Specimens. Eight diaphragm-through joints of the concrete-filled square steel tubular columns were designed for static tensile loading, as shown in Figure 5. In these specimens, the flange of the steel beam and the diaphragm are connected by butt weld. The stress concentration and the welding stress may have a negative effect on mechanical properties of diaphragm-through joints. Therefore, the arc fillet eliminates the stress concentration effect and sets these butt welds at some distance from the column.

After cutting, the surface is treated by surface grinding machine. The steel tubes, diaphragms and beam flanges are thoroughly wire brushed to remove the rust and loose debris that may be present during

Specimen	Steel tube $D \times t_c$	Flange $l_b \times b_f \times t_b$	Diaphragm $l_d \times b_d \times t_d$	d	r	h_d	Concrete
TR1	250×8	$100 \times 125 \times 10$	$490 \times 300 \times 10$	120	35	25	C40
TR2	250×8	$100 \times 125 \times 10$	$490 \times 300 \times 10$	140	35	25	C40
TR3	250×8	$100 \times 125 \times 10$	$490 \times 300 \times 10$	160	35	25	C40
TR4	250×8	$100 \times 125 \times 12$	$490 \times 300 \times 12$	160	35	25	C40
TR5	250×10	$100 \times 125 \times 12$	$490 \times 300 \times 12$	160	35	25	C40
TR6	250×8	$100 \times 125 \times 10$	$490 \times 300 \times 10$	140	50	25	C40
TR7	250×10	$100 \times 125 \times 12$	$490 \times 320 \times 12$	160	35	45	C40
TR8	250×10	$100 \times 125 \times 12$	$490 \times 300 \times 12$	160	35	25	C20

Table 4. Specimen parameters. All dimensions are in millimeters.

cutting. A steel plate is placed at the bottom of each steel tube prior to filling concrete. The concrete is cured for 28 days. A layer of cement mortar is used to flush the concrete top surface with the steel tube.

The parameters of each specimen are listed in Table 4. The meaning of the symbols D, t_c, l_b, b_f, t_b, l_d, b_d, t_d, d, r and h_d are shown in Figure 5.

Experimental set-up. The specimens of diaphragm-through joints are stretched with clamping conditions at both flange ends using a 1000 kN capacity axial tensile testing machine. A load interval of 50 kN (about one-tenth of the estimated carrying load capacity) is used. Each load interval is maintained for 10 minutes. The response of the specimens is recorded continuously in order to obtain the load-displacement curve. When the load-displacement curve slope changes, it means the specimen enters the yield stage and the loading increment is decreased to 10 kN until failure. The experimental set-up is shown in Figure 6.
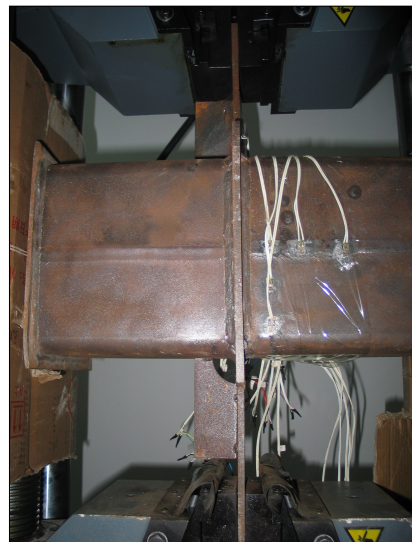

Figure 6. Experimental set-up: axial tensile testing machine (left) and tensile loading (right).

Specimen	Yield load P_y^e (kN)	Failure load P_u^e (kN)	Failure displacement Δ (mm)	Failure mode	Description of failure mode
TR1	361.9	505.8	13.7	flange	Fracture at steel beam flange
TR2	350.7	500.9	14.6	flange	Fracture at steel beam flange
TR3	337.5	462.8	14.7	joint	Local buckling at steel tube Stretching of diaphragm
TR4	396.4	538.9	11.7	joint	Local buckling at steel tube Stretching of diaphragm
TR5	407.8	606.1	17.4	joint	Local buckling at steel tube Stretching of diaphragm
TR6	361.9	504.4	14.1	flange	Fracture at steel beam flange
TR7	416.7	610.0	14.6	joint	Local buckling at steel tube Stretching of diaphragm
TR8	405.7	604.7	18.9	joint	Local buckling at steel tube Stretching of diaphragm

Table 5. Bearing capacity and failure mode.

Results and discussion.

Failure modes. The specimens are loaded up to failure to investigate the mechanical properties and the bearing capacities of the diaphragm-through joints under static tensile loading. A summary of the experimental results such as bearing capacity, displacement and failure modes is given in Table 5. In Table 5, the failure load P_u^e is defined as the maximum strength of specimen. The yield load P_y^e is obtained by a graphical method [Nie et al. 2008]. The failure modes of all specimens can be divided into two types: failure of the flange and failure of the joint (as described Table 5).

In the case of specimen TR6, at the initial stage, there were no obvious deformations in the specimen. When loaded further, plastic necking started to appear at the beam flange. Finally, at failure load, the beam flange was pulled off while there were still no obvious deformations in the diaphragm and the steel tube, as shown in Figure 7, top left. Similar failure was observed in the case of specimens TR1 and TR2.

In the case of specimen TR5, the diaphragm was stretched in the loading direction and the steel tube exhibited signs of local buckling. At the failure point, the deformations of the diaphragm and the steel tube and local buckling were obvious, as shown in Figure 7, top right, and the specimen could not take any more load. After the removal of the concrete from steel tube at the end of the static tensile loading experiment, it could be seen that the circular opening in the center of diaphragm, which is used for concrete casting, is stretch to ellipse, as shown in Figure 7, bottom. Similar failure was observed in the case of specimens TR3, TR4, TR7 and TR8.

Load-displacement curves. All the specimens exhibit the smooth load-displacement curves shown in Figure 8. From these load-displacement curves, it can be seen that specimens have similar behavior. The curves exhibit elastic behavior at the initial stage followed by an inelastic response. All specimens show a clear plastic plateau before the failure point as well as a better ductility.

Figure 7. Failure modes. Clockwise from top left: fracture at steel beam flange; local buckling at steel tube; stretching of diaphragm.
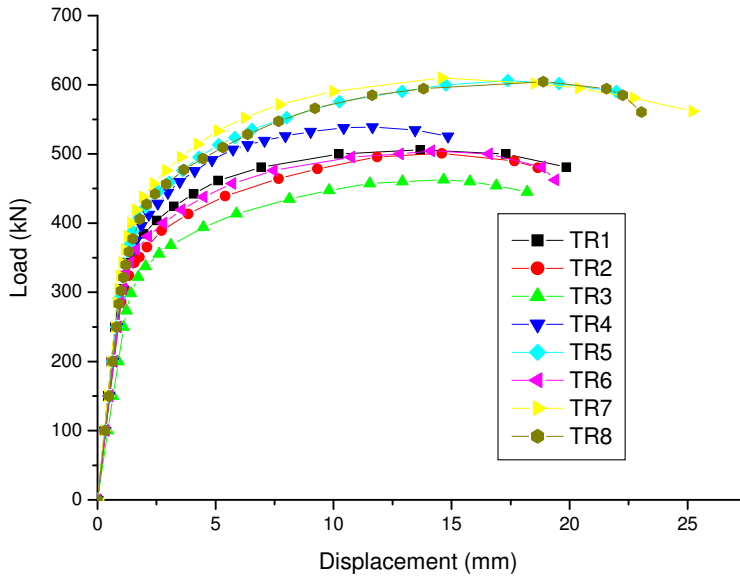
Figure 8. Load-displacement curves.
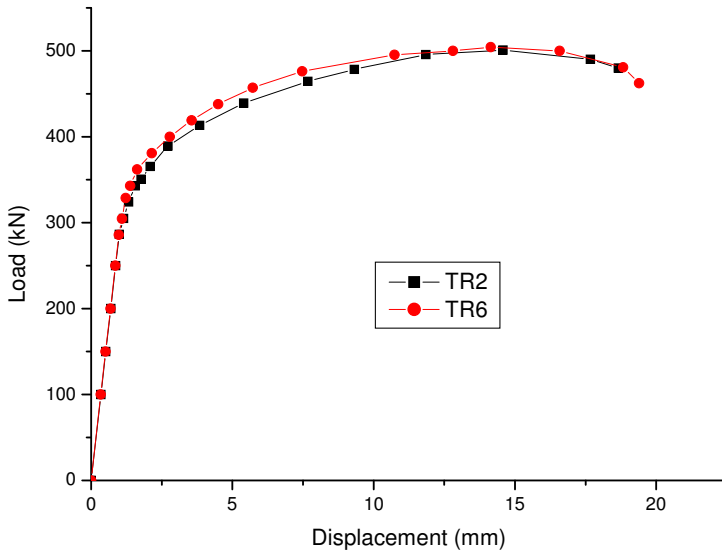
Figure 9. Effect of arc fillet.

The specimens were designed with different parameters such as thickness of steel tube, dimension of diaphragm, radius of arc fillet and concrete and a parametric study was conducted as discussed below.

(1) *Fillet*. In the case of specimens TR2 and TR6, the dimensions and material properties are the same except for the radius *r* of arc fillet, which is 35 mm for TR2 and 50 mm for TR6. Figure 9 shows the load-displacement comparison between specimens TR2 and TR6. It can be seen that the change of the arc fillet radius from 35 mm to 50 mm does not have an obvious influence in the failure load and the ductility of specimens TR2 and TR6.
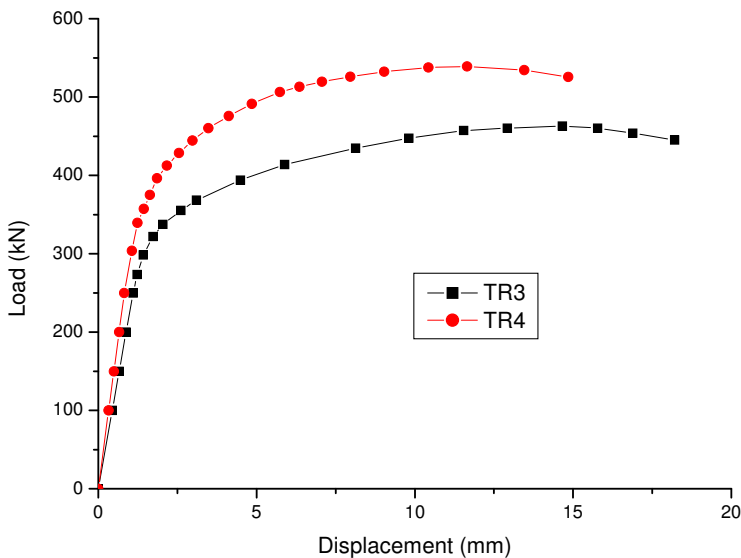
Figure 10. Effect of diaphragm thickness.
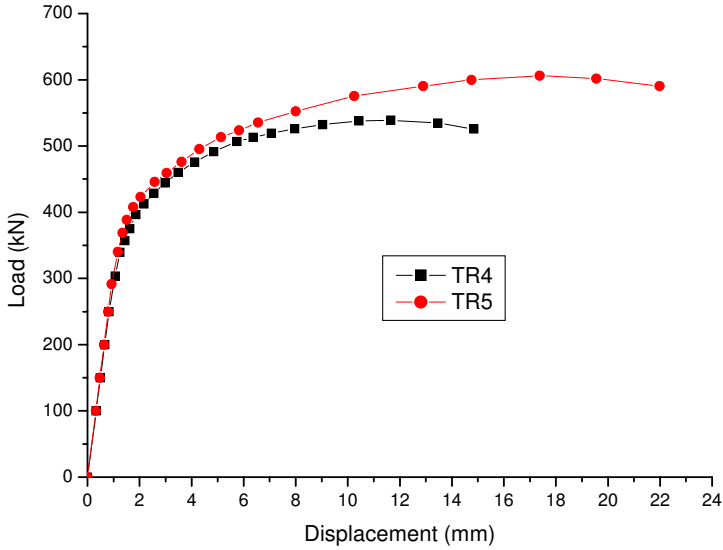
Figure 11. Effect of steel tube thickness.

(2) *Thickness of diaphragm.* In the case of specimens TR3 and TR4, dimensions and material properties are the same except for the thickness t_d of the diaphragm, which is 10 mm for TR3 and 12 mm for TR4. Figure 10 shows the load-displacement comparison between specimens TR3 and TR4. It can bee seen that the larger thickness of the diaphragm lead to a larger failure load.

(3) *Thickness of steel tube.* In the case of specimens TR4 and TR5, dimensions and material properties are the same except for the thickness t_c of steel tube, which is 8 mm for specimen TR4 and 10 mm in TR5. Figure 11 shows the load-displacement comparison between specimens TR4 and TR5. It can be seen that larger thickness of steel tube lead to larger failure load.
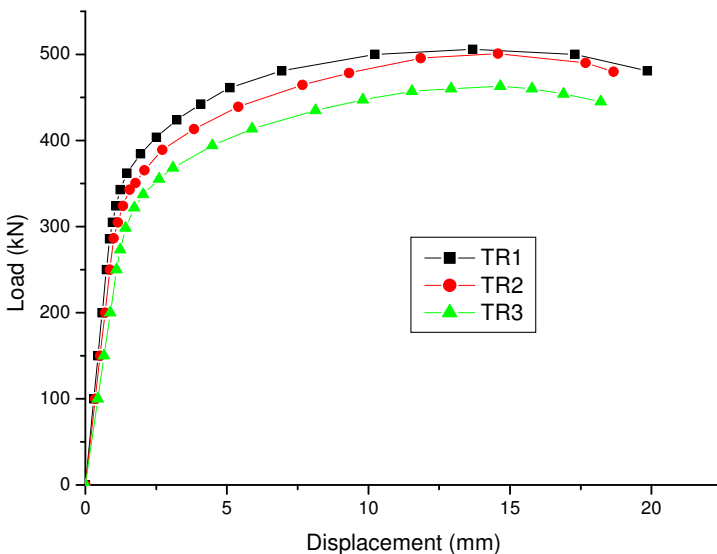
Figure 12. Effect of diameter of central opening in diaphragm.
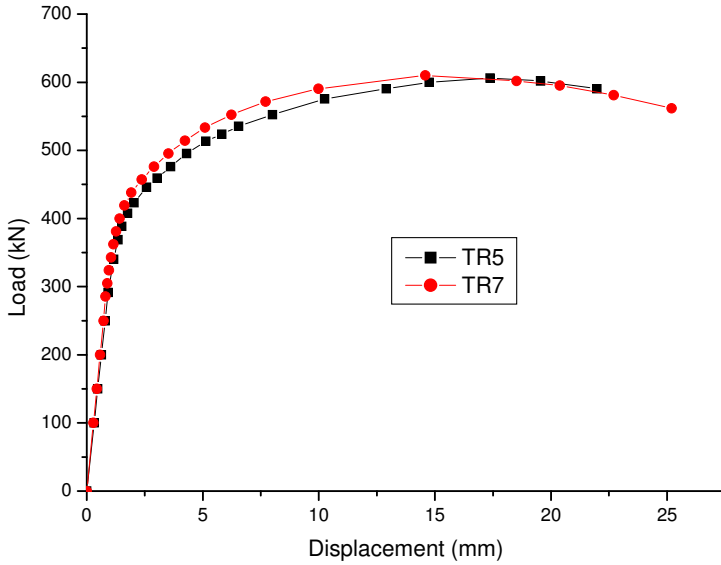
Figure 13. Effect of cantilever length of diaphragm.

(4) *Diameter of central opening in diaphragm.* In the case of specimens TR1, TR2 and TR3, dimensions and material properties are the same except for the diameter *d* of central opening in diaphragm. The diameters *d* of central opening in diaphragms for specimen TR1, TR2 and TR3 are 120 mm, 140 mm and 160 mm, respectively. Figure 12 shows the load-displacement comparison among specimens TR1, TR2 and TR3. It can be seen that larger diameter of central opening in diaphragm lead to smaller failure load.
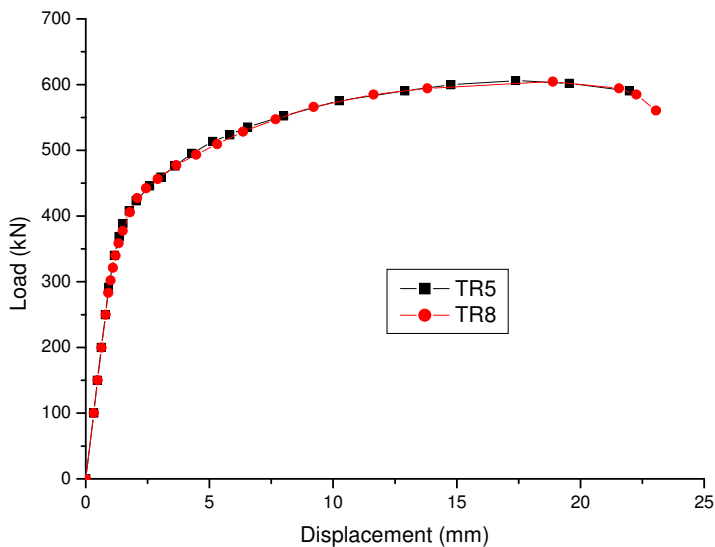
Figure 14. Effect of concrete grade.

Specimen	Yield bearing capacity			Ultimate bearing capacity		
	P_y^e (kN)	P_y^d (kN)	P_y^d/P_y^e	P_u^e (kN)	P_u^d (kN)	P_u^d/P_u^e
TR1	361.9	582.6	1.61	505.8	810.8	1.60
TR2	350.7	466.1	1.33	500.9	648.6	1.29
TR3	337.5	278.8	0.83	462.8	388.0	0.84
TR4	396.4	335.3	0.85	538.9	472.2	0.88
TR5	407.8	335.3	0.82	606.1	472.2	0.78
TR6	361.9	466.1	1.29	504.4	648.6	1.29
TR7	416.7	554.3	1.33	610.0	780.6	1.28
TR8	405.7	335.3	0.83	604.7	472.2	0.78

Table 6. Bearing capacity comparison.

(5) *Cantilever length of diaphragm.* In the case of specimens TR5, and TR7, dimensions and material properties are the same except for the cantilever length h_d of diaphragm, which is 25 mm for TR5 and 45 mm for TR7. Figure 13 shows the load-displacement comparison among specimens TR5 and TR7. It can be seen that the change of cantilever length of diaphragm from 25 mm to 45 mm does not have any obvious influence in failure load and ductility of specimens TR5 and TR7.

(6) *Concrete.* In the case of specimens TR5 and TR8, dimensions and material properties are the same except for the grade of concrete, which is C40 for TR5 and C20 for TR8. Figure 14 shows the load-displacement comparison between specimens TR5 and TR8. It can be seen that the change of grade of concrete from C40 to C20 does not have a noticeable influence on the failure load and ductility of specimens TR5 and TR8.

Bearing capacity comparison. Table 6 presents the experimentally obtained yield tensile capacity P_y^e and ultimate tensile capacity P_u^e of the specimens together with the corresponding computed values P_y^d and P_u^d using (4). For the yield tensile capacity, the difference between experimental results and calculated values ranges from 15% to 61%. For the ultimate tensile capacity, the difference ranges from 12% to 60%. These differences between experimental and calculated values are big. For specimens TR4 and TR5, experiments indicate that an increase of thickness of steel tube increases the bearing capacity. However, the calculated values of TR4 and TR5 are the same because the contribution of the steel tube to the bearing capacity is not reflected in (4). The computed bearing capacity of specimen TR5 is smaller than that of TR7 because specimen TR7 has a larger diaphragm but the experimental bearing capacity of these two specimens is almost the same.

These tests show that the correlation between the predictions of (4) and the experimentally obtained values is rather poor. Therefore there is a need for a better understanding of the yield mechanism of the diaphragm-through joint and an improved method to estimate the yield and ultimate capacity of the joint.

3. Finite element analysis

A three-dimensional finite element model is developed to predict the failure mode, bearing capacity and load-deformation behavior of diaphragm-through joints and compare these predictions with test results.
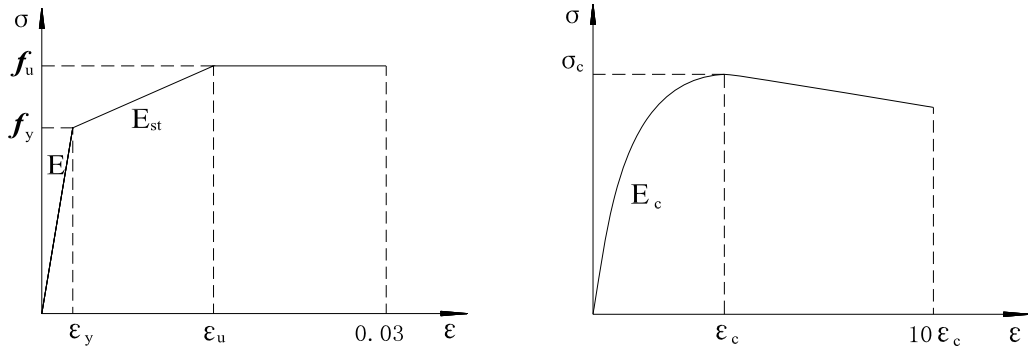
Figure 15. Constitutive law of steel (left) and concrete (right).

Finite element model. The three-dimensional finite element models of diaphragm-through joint specimens were created using the finite element package ANSYS 11.0. The finite element models have the same size and the material properties with the specimens of the static tensile loading experiments as described before.

The main components that need to be modeled in order to simulate the behavior of the diaphragm-through joint are the steel tube, the diaphragm, the beam flange, the infilled concrete and the interface between the concrete and the steel tube. In addition, the choice of the element type, the initial geometric configuration, the boundary conditions and the load application are also important in simulating the diaphragm-through joint and they are also discussed in the following.

(1) *Modeling of the steel tube, diaphragm and beam flange.* The element Solid 95 is used to model steel tube, the diaphragm and the beam flange. The constitutive law is assumed elasto-plastic with Poisson's ratio equal to 0.3. The idealized stress-strain curve [Nie et al. 2008] used in the numerical analysis is shown in Figure 15, left, where $\epsilon_y = f_y/E$, $\epsilon_u = 10(f_u - f_y)/E$.

(2) *Modeling of concrete.* The three-dimensional 8-node element Solid 65 is adopted to model the infilled concrete. The uniaxial stress-strain curve [Chen et al. 2009] shown in Figure 15, right, is adopted with Poisson's ratio equal to 0.2. The constitutive law has two branches. The ascending branch is assumed parabolic up to a strain ϵ_c equal to 0.003 and the descending branch is linear.

(3) *Modeling of the concrete-steel tube interface.* The contact action between the steel and the concrete is modeled by the contact elements Targe 170 and Conta 173. These surface-to-surface contact elements consist of two matching contact faces of steel and concrete elements. The friction between the two faces is maintained as long as the surfaces remain in contact. The coefficient of friction is taken as 0.25. These contact elements allows the surfaces to separate under the influence of tensile force. However, the contact elements are not allowed to penetrate each other.

(4) *Initial geometric configuration.* Since initial geometrical deformations are generally present in practice in diaphragm-through joint specimens, the effect of such deformations is considered in this simulation. For this purpose, a 1% initial deformation resembling the first-order buckling mode has been introduced to the finite element modeling.
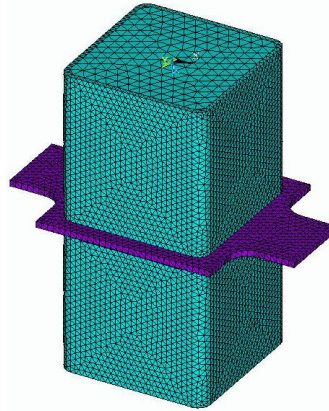
Figure 16. Finite element model.

(5) *Boundary conditions and load application.* The tensile load is transferred through the flange of the steel beam acting on both sides of the diaphragm. To simulate this action one side of the diaphragm was fixed and the tensile load was applied on the other end.

    The load was applied as static uniform load using displacement control at each node of the loaded surface, and the displacement increments were identical to the increments of the experimental investigation. The finite element models are shown in Figure 16.

Numerical results. The failure modes of the diaphragm-through joints, the load-displacement curves and the bearing capacity obtained by the numerical simulation are presented and discussed below.

(1) *Failure modes.* The finite element analysis shows that the failure modes of specimens are similar. The failure modes of specimens TR3 are shown in Figure 17. It can be seen that local buckling appears at the steel tube in agreement with the experimental results.

(2) *Load-displacement curves.* Figure 18 shows the static tensile load-displacement curves compared
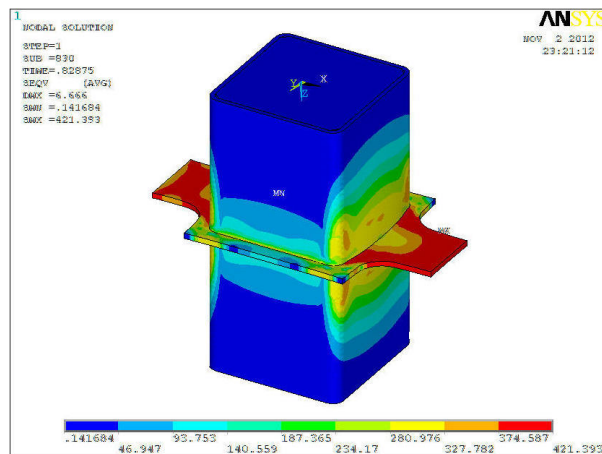
Figure 17. Failure mode of TR3.
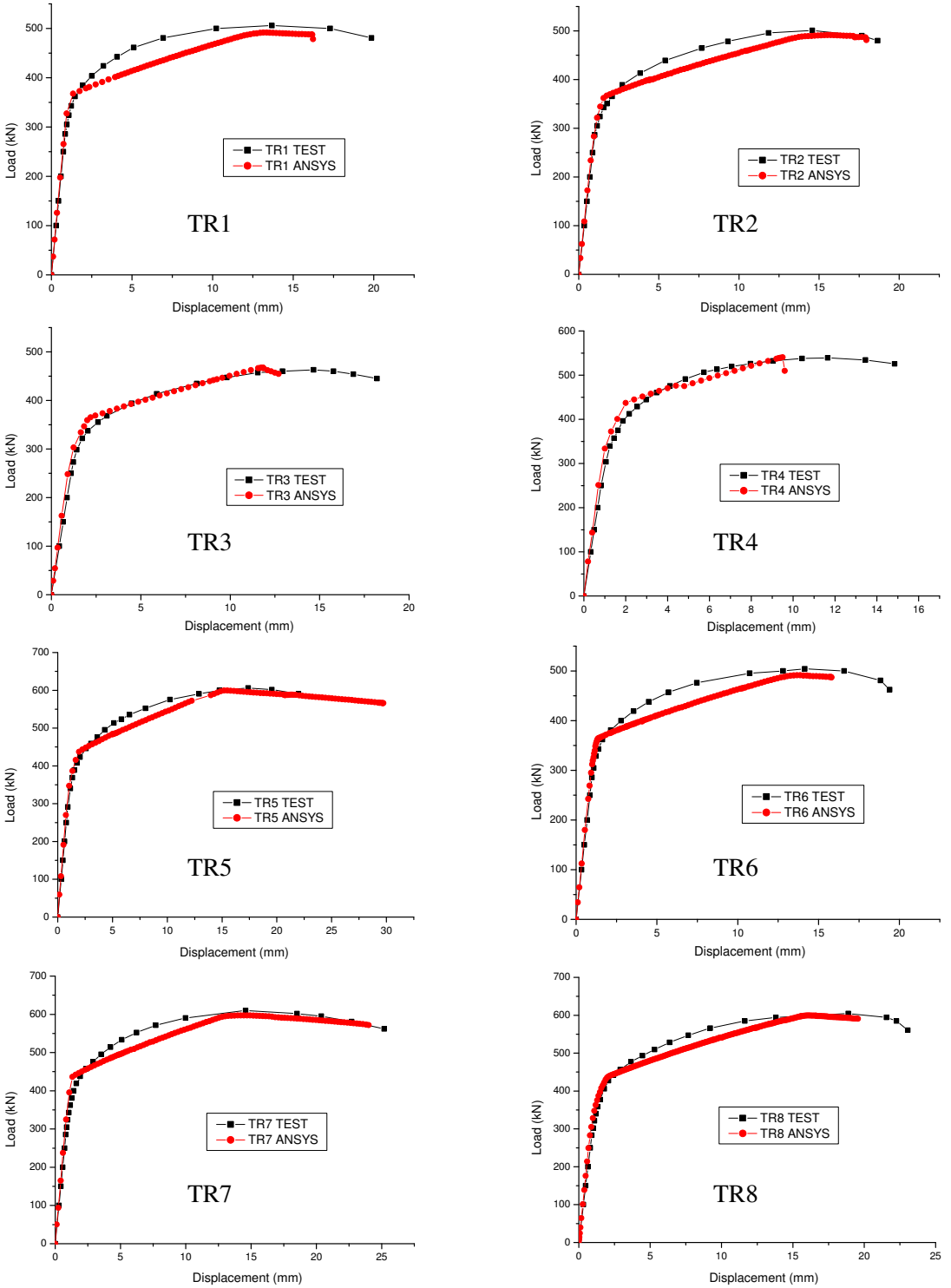
Figure 18. Comparison of load-displacement curves for TR1–TR8.

Specimen	Yield bearing capacity			Ultimate bearing capacity		
	P_y^e (kN)	P_y^f (kN)	P_y^f / P_y^e	P_u^e (kN)	P_u^f (kN)	P_u^f / P_u^e
TR1	361.9	367.4	1.02	505.8	514.6	1.02
TR2	350.7	362.2	1.03	500.9	512.3	1.02
TR3	337.5	347.7	1.03	462.8	474.4	1.03
TR4	396.4	413.5	1.04	538.9	557.4	1.03
TR5	407.8	421.5	1.03	606.1	596.3	0.98
TR6	361.9	367.5	1.02	504.4	517.5	1.02
TR7	416.7	436.3	1.05	610.0	597.2	0.98
TR8	405.7	419.2	1.03	604.7	595.4	0.98

Table 7. Comparison of load-displacement curves.

with the experimental results of specimens. The load-displacement curves obtained by numerical simulation for all the diaphragm-through joints exhibited linear elastic behavior at the initial stage followed by inelastic behavior when the load was further increased. All load-displacement curves are in good agreement with the experimental ones.

(3) *Bearing capacity.* The yield load P_y^f and the ultimate load P_u^f obtained by the above curves are shown in Table 7, and compared with experimental results. For the yield tensile capacity, the difference between experimental and FE results ranges from 2% to 5%. For the ultimate tensile capacity, it ranges from 2% to 3%. It can be concluded that the finite element analysis results and experimental results are in good agreement and thus finite element analysis will give reliable predictions.

4. Yield mechanism

The load transfer mechanism. The finite element analysis and the static tensile loading experiment show that, as the load increases, the load is transferred to the tube and the concrete core. The stress contour of specimen TR3 at failure is shown in Figure 19. It can be seen that there is a large stress distribution
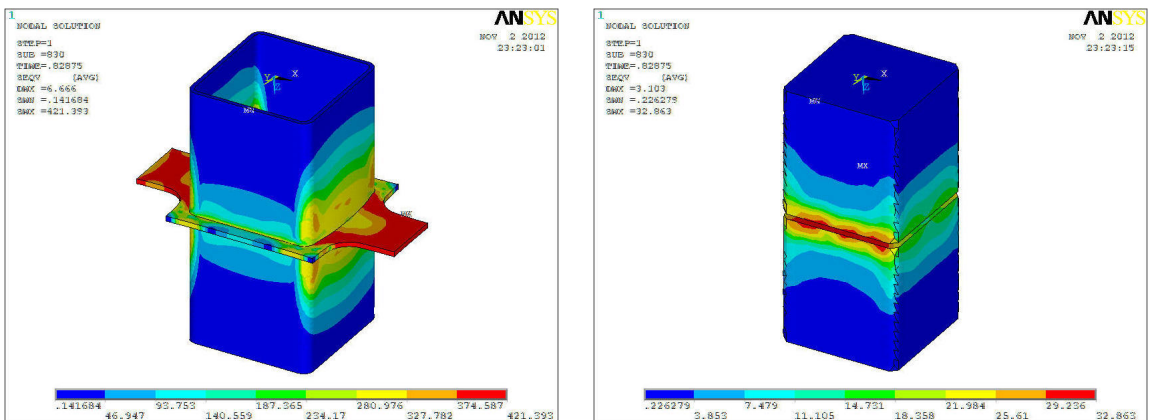
Figure 19. Stress contour of TR3: steel tube and diaphragm (left); concrete (right).
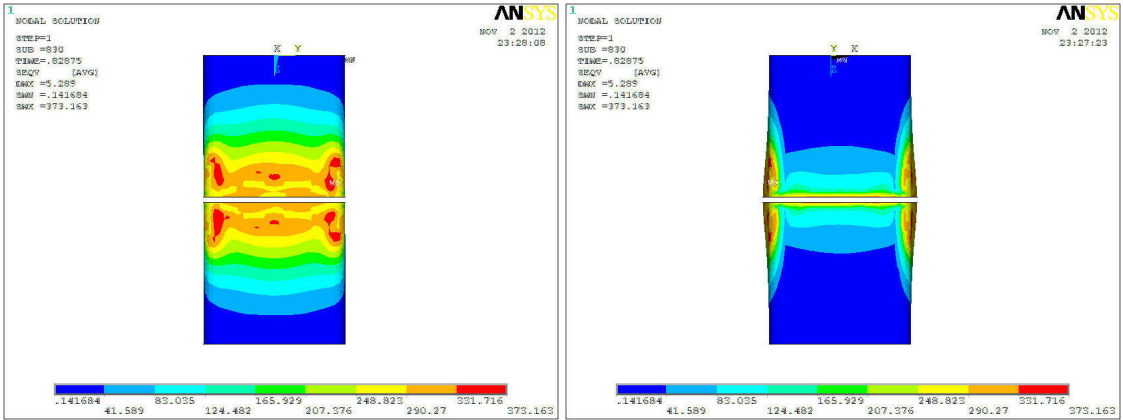
Figure 20. Stress contour in steel tube of TR3: façade (left) and side (right).

zone in the diaphragm and the nearby steel tube while the stress level of the concrete core is lower with a smaller distribution zone. It seems that the tensile load from the steel beam flange is transferred mainly to the square steel tube and the diaphragm.

The yield mechanism of the steel tube. The distribution of the plastic zone in the steel tube of specimen TR3, as obtained by the finite element analysis, is shown in Figure 20. The stress level of the steel tube is higher along the loading direction with a larger distribution area, while in the perpendicular direction the stress level is lower with a smaller distribution area. Plastic zone appears mainly in the intersection of steel tube and diaphragm. Similar results were obtained for the other specimens.

Based on these results a rational procedure for the calculation of the tensile load carried by the steel tube is the yield line model shown in Figure 21.

Let M_{yw} and M_{yc} be the yield moment per unit length of the butt weld and the steel tube respectively and $M_{ya} = \min\{M_{yw}, M_{yc}\}$. By the principle of virtual work ($\partial P_y^c / \partial X = 0$), the distance X is

$$X = \sqrt{(D + Y - t_c)Y/2}. \tag{5}$$
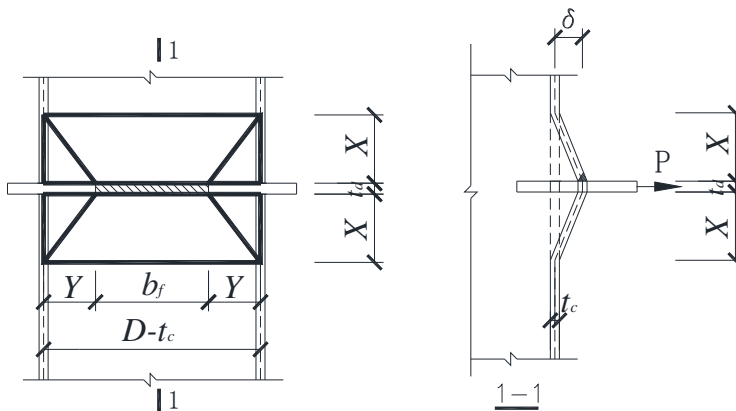
Figure 21. Calculation model of yield lines for steel tube.
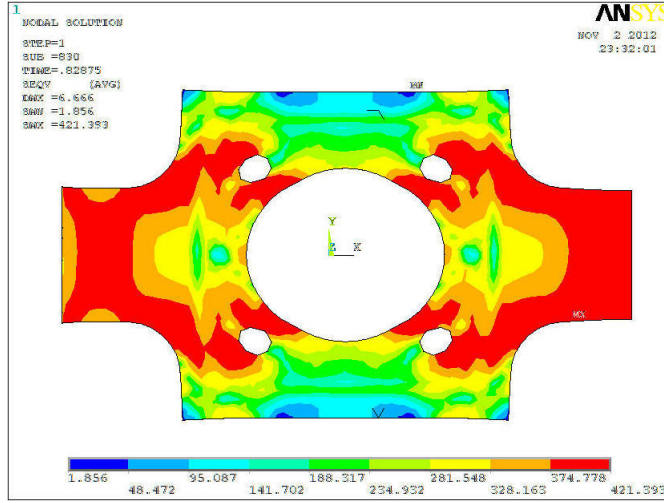
Figure 22. Stress contour in diaphragm of TR3.

And then

$$P_c^y = \frac{4(M_{ya} + M_{yc})X}{Y} + \frac{4M_{yc}(D + Y - t_c)}{X} \tag{6}$$

where D is the width of the tube and t_c is the thickness of the walls of the tube.

The yield mechanism of the diaphragm. The stress distribution of the plastic zone in the diaphragm is shown in Figure 22. The stress level of the diaphragm is higher along the loading direction. In the perpendicular direction the stress level is lower, especially in the cantilevered section. From the stress contour of the specimen of Figure 22 as well as the other analyzed specimens, it seems that the tensile load from the flange is mainly transferred to the section of the diaphragm enclosed by the steel tube and a plastic zone appears along the diagonals of the cross-section lines enclosed by the tube.

Based on these observations a computation model of the yielding mechanism of the diaphragm is shown in Figure 23 where it is assumed that yielding will take place along the diagonals of the diaphragm [Lu 1997].

The yield tensile strength P_d^y can be calculated from Figure 23 using

$$P_d^y = \sqrt{2}t_d f_{yd}(l_2 + l_1/2) \tag{7}$$

where t_d is the thickness of the diaphragm and f_{yd} is the yield stress of the diaphragm.

Verification of the analytical method. The total ultimate bearing capacity of diaphragm-through joint under static tensile load is given by (6) and (7) if the yield strength is substituted by the ultimate

$$P_u = \frac{4(M_{ua} + M_{uc})X}{Y} + \frac{4M_{uc}(D + Y - t_c)}{X} + \sqrt{2}t_d f_{ud}(l_2 + l_1/2). \tag{8}$$

In order to verify the proposed analytical method, a comparison, as exhibited in Table 8, is made between the static tensile experimental results of diaphragm-through joints and the calculated ultimate tensile capacity P_u. For specimens TR3, TR4, T5, TR7 and TR8 with failure of joint, P_u is calculated using (8).
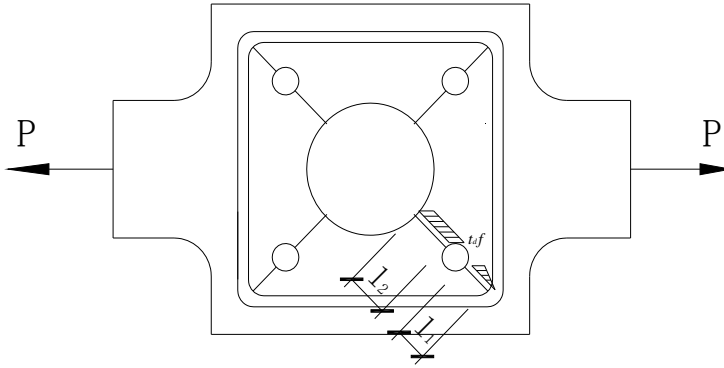
Figure 23. Calculation model of yield lines for diaphragm.

Specimen		Ultimate bearing capacity		
		P_u^e (kN)	P_u (kN)	P_u/P_u^e
Failure of flange	TR1	505.8	526.8	1.04
	TR2	500.9	526.8	1.05
	TR6	504.4	526.8	1.04
Failure of joint	TR3	462.8	425.5	0.92
	TR4	538.9	482.8	0.90
	TR5	606.1	580.4	0.96
	TR7	610.0	580.4	0.95
	TR8	604.7	580.4	0.96

Table 8. Comparison of test and computed bearing capacities.

For specimens TR1, TR2 and TR6 with failure of flange, $P_u = f_u \cdot t \cdot b_f$, where, f_u is the ultimate strength of the flange, t is the thickness of the flange and b_f is the width of the flange. According to (8), the calculated ultimate tensile capacity P_u of specimens TR1, TR2 and TR6 is 559.3 kN, 529.6 kN and 529.6 kN respectively, which is larger than $P_u = f_u \cdot t \cdot b_f$. As shown in Table 8, the computed capacities are in good agreement with the experimental.

5. Conclusions

Through static tensile loading experiments and finite element simulations, the mechanical properties and the bearing capacity of diaphragm-through joint of concrete-filled square steel tubular column under static tensile load are investigated in this paper.

The static tensile loading experiments show that there are two types of failure modes: the failure of the flange and the failure of the joint. The contribution of the infilled concrete is small and can be neglected.

Comprehensive comparison of finite element analysis and experimental results are in good agreement. Based on the findings of this analytical and experimental investigation, a rational computational procedure for the evaluation of the ultimate capacity is proposed. This procedure takes into account the

contribution of the diaphragm-through as well as the steel tube which, unlike the contribution of the infilled concrete, was found to be considerable.

The predictions of the proposed procedure are in good agreement with the experimental data of this investigation.

Acknowledgments

References

[Chen et al. 2006] Y.-Y. Chen, G. Li, L. Zhuang, M.-X. Huang, and J. Li, "Experimental study on the seismic performance of diaphragm through type joint connecting H beam and tube column in steel frame", *Prog. Steel Build. Struct.* **8**:1 (2006), 23–30. In Chinese.

[Chen et al. 2009] Z.-H. Chen, B. Rong, and A. Fafitis, "Axial compression stability of a crisscross section column composed of concrete-filled square steel tubes", *J. Mech. Mater. Struct.* **4**:10 (2009), 1787–1799.

[Fukumoto and Morita 2005] T. Fukumoto and K. Morita, "Elasto-plastic behavior of panel zone in steel beam-to-concrete filled steel tube column moment connections", *J. Struct. Eng.* (*ASCE*) **131**:7 (2005), 1841–1853.

[Han et al. 2003] L. H. Han, Y. F. Yang, and Z. Tao, "Concrete-filled thin-walled steel SHS and RHS beam-columns subjected to cyclic loading", *Thin-Walled Struct.* **41**:9 (2003), 801–833.

[Han et al. 2008] L. H. Han, W. D. Wang, and X. L. Zhao, "Behaviour of steel beam to concrete-filled SHS column frames: finite element model and verifications", *Eng. Struct.* **30**:6 (2008), 1647–1658.

[Jiang et al. 2009] X.-L. Jiang, J.-K. Miao, and Z.-H. Chen, "Experiment on seismic performance of diaphragm-through joint between concrete-filled square steel tubular column and steel beam", *J. Tianjin Univ.* **42**:3 (2009), 194–200. In Chinese.

[Lu 1997] L. H. Lu, *The static strength of I-beam to rectangular hollow section column connections*, Ph.D. thesis, Delft University, Delft, 1997, http://repository.tudelft.nl/view/ir/uuid:80f96639-3f74-4231-8dab-af93cdb97a5f.

[Matsui 1985] C. Matsui, "Strength and behavior of frames with concrete filled square steel tubular columns under earthquake loading", pp. 143–146 in *Proceedings of the International Specialty Conference on Concrete Filled Steel Tubular Structures* (Harbin, 1985), 1985.

[Morino and Tsuda 2002] S. Morino and K. Tsuda, "Design and construction of concrete-filled steel tube column system in Japan", *Earthq. Eng. Eng. Seismol.* **4**:1 (2002), 51–73.

[Nie et al. 2005] J.-G. Nie, K. Qin, and G.-B. Zhang, "Experimental research and theoretical analysis on flexural capacity of connections for concrete-filled steel square tubular columns with inner diaphragms", *J. Archit. Civ. Eng.* **22**:1 (2005), 42–49. In Chinese.

[Nie et al. 2008] J.-G. Nie, K. Qin, and C. S. Cai, "Seismic behavior of connections composed of CFSSTCs and steel-concrete composite beams — finite element analysis", *J. Constr. Steel Res.* **64**:6 (2008), 680–688.

[Nishiyama et al. 2004] I. Nishiyama, T. Fujimoto, and T. Fukumoto, "Inelastic force-deformation response of joint shear panels in beam-column moment connections to concrete-filled tubes", *J. Struct. Eng.* (*ASCE*) **130**:2 (2004), 244–252.

[Park et al. 2010] S. H. Park, S. M. Choi, Y. S. Kim, Y. Park, and J. H. Kim, "Hysteresis behavior of concrete filled square steel tube column-to-beam partially restrained composite connections", *J. Constr. Steel Res.* **66**:7 (2010), 9433–953.

[Ricles et al. 2004] J. M. Ricles, S. W. Peng, and L. W. Lu, "Seismic behavior of composite concrete filled steel tube column-wide flange beam moment connections", *J. Struct. Eng.* (*ASCE*) **130**:2 (2004), 223–232.

[Sasaki et al. 1995] S. Sasaki, M. Teraoka, K. Morita, and T. Fujiwara, "Structural behavior of concrete-filled square tubular column with partial penetration weld corner seam to steel H-beam connections", pp. 33–40 in *Structural steel: PSSC '95, 4th*

Pacific Structural Steel Conference: 2, Structural connections (Singapore, 1995), edited by N. E. Shanmugam and Y. S. Choo, Pergamon, Oxford, 1995.

[Zhou et al. 2005] T. H. Zhou, S. F. Nie, L. F. Lu, and B. K. He, "Design of concrete-filled square tube column and steel beam joint with internal diaphragms", *J. Build. Struct.* **26**:5 (2005), 23–29. In Chinese.

RONG BIN: rongbin2010@hotmail.com
School of Civil Engineering and Key Laboratory of Coast Civil Structure Safety, Tianjin University, Tianjin, 300072, China

CHEN ZHIHUA: zhchen@tju.edu.cn
School of Civil Engineering and Key Laboratory of Coast Civil Structure Safety, Tianjin University, Tianjin 300072, China

ZHANG RUOYU: zryu@tju.edu.cn
School of Naval and Ocean Engineering, Tianjin University, Tianjin 300072, China

APOSTOLOS FAFITIS: fafitis@asu.edu
Department of Civil and Environmental Engineering, Arizona State University, Tempe, AZ 85287, United States

YANG NAN: yangnangre@gmail.com
School of Civil Engineering, Tianjin University, Tianjin 300072, China

msp

BUCKLING AND POSTBUCKLING BEHAVIOR OF FUNCTIONALLY GRADED TIMOSHENKO MICROBEAMS BASED ON THE STRAIN GRADIENT THEORY

Reza Ansari, Mostafa Faghih Shojaei, Vahid Mohammadi,
Raheb Gholami and Mohammad Ali Darabi

Presented herein is a comprehensive study on the buckling and postbuckling analysis of microbeams made of functionally graded materials (FGMs) based on the modified strain gradient theory. The present model is developed in the skeleton of the Timoshenko beam theory and the von Karman geometric non-linearity, and enables one to consider size effects through introducing material length scale parameters. Also, the current model can be reduced to the modified couple stress and classical models if two or all material length scale parameters are set equal to zero, respectively. Utilizing a power law function, the volume fraction of the ceramic and metal phases of the functionally graded microbeam is expressed. The stability equations and corresponding boundary conditions are derived using Hamilton's principle and then solved through the generalized differential quadrature (GDQ) method in conjunction with a direct approach without linearization. The effects of the length scale parameter, slenderness ratio, material gradient index and boundary conditions on the buckling and postbuckling behavior of microbeams are carefully studied. Furthermore, the non-dimensional critical axial load of microbeams predicted by modified strain gradient and classical theories for the first three postbuckling modes is investigated and it is observed that the classical theory underestimates the non-dimensional critical axial load, especially at higher postbuckling modes. In addition, the influence of imperfections on the deflection of microbeams in prebuckled and postbuckled states is discussed.

1. Introduction

Microstructures made of FGMs are becoming hot research areas, since they pave the way for achieving highly sensitive and most desired micro-electromechanical systems [Hasanyan et al. 2008; Witvrouw and Mehta 2005; Fu et al. 2004].

As the size-dependent deformation behavior is detected in the micro-torsion and micro-bending experiments of microbeams, the size effects must be incorporated in the study of FGM microstructures. The conventional continuum mechanics is not able to capture size effects; hence, developing size-dependent elasticity theories has become an important issue. In this regard, several non-classical continuum theories such as the strain gradient elasticity, couple stress elasticity, nonlocal elasticity and the surface elasticity have been proposed [Aifantis 1999; Mindlin and Tiersten 1962; Eringen 1972; Gurtin et al. 1998]. The two first size-dependent theories encompass some higher-order stress constituents in addition to the classical stresses.

One type of the higher-order continuum theories reported in Mindlin and Tiersten 1962; Koiter 1964; Toupin 1962; 1964] is the couple stress theory which includes two material length scale parameters in addition to classical material constants for isotropic elastic materials. These two material length scale parameters are in relation to the underlying microstructure of material and are so complicated to determine. Yang et al. [2002] proposed the modified couple stress theory (MCST) and facilitated applying couple stress theory by considering only one material length scale parameter beside two classical material constants. A variational formulation for the MCST was also presented by Park and Gao [2008]. This theory has been extensively used to interpret size effects on the vibrational, bending and buckling behavior of microstructures [Ma et al. 2008; 2010; 2011; Tsiatas 2009; Xia et al. 2010; Park and Gao 2006].

The strain gradient theory proposed by Fleck and Hutchinson [1993] is the extension of Mindlin's formulation [Mindlin 1964; 1965; Mindlin and Eshel 1968] which considers only the first derivative of the strain tensor. In comparison with the couple stress theory, this theory comprises some higher-order stress constituents beside the classical and couple stresses. Lam et al. [2003] introduced the modified strain gradient theory (MSGT) through considering three higher-order material constants related to dilatation gradient, deviatoric gradient and symmetric rotation gradient tensors. It is noted that the MSGT is also a simplified version of the Mindlin's general theory [Mindlin 1964; 1965; Mindlin and Eshel 1968] which can be developed by omitting the second-order derivatives of the strain components in the elastic strain energy. Numerous studies have been accomplished to investigate the size-dependent static bending and vibrational analysis of microstructures based on the MSGT. The static deformation and vibrational behavior of an Euler–Bernoulli microbeam was studied by Kong et al. [2009] using the MSGT. Wang et al. [2012] developed a micro-scaled Timoshenko beam model based on the MSGT and Hamilton's principle. They also studied the static bending and free vibration of a simply-supported Timoshenko microbeam. In another work, Kahrobaiyan et al. [2011], based on the MSGT, proposed a nonlinear size-dependent Euler–Bernoulli beam model and studied the nonlinear size-dependent static bending of a hinged-hinged microbeam. In a recent study, based on the MSGT, Ansari et al. [2012a] developed a nonlinear size-dependent Timoshenko microbeam model and examined the influences of geometric parameters, Poisson's ratio and material length scale parameters on the vibrational behavior of microbeams.

A literature review shows that the majority of studies on microstructures are concerned with the homogeneous materials. However, in recent years, several efforts have been tended to study the size-dependent mechanical behavior of microstructures made of FGMs [Asghari et al. 2010; 2011; Reddy and Jinseok 2012; Mirzavand and Eslami 2011; Ke et al. 2012; Mohammadi-Alasti et al. 2011; Ke and Wang 2011; Ansari et al. 2012b; 2011]. In this direction, utilizing the MCST, Asghari et al. [2010] studied the static and vibrational behavior of Euler–Bernoulli FGM microbeams. It was observed that when the ratio of the beam's characteristic size to the material length scale parameter is low, the results obtained by the MCST have a significant difference with those obtained by the classical theory (CT). Ke and Wang [2011] investigated size-dependent dynamic stability of FGM microbeams based on the MCST and Timoshenko beam theory. They found that when the thickness of beam is equal to the material length scale parameter, the size effect on the dynamic stability characteristics is prominent. Ansari et al. [2012b] investigated the nonlinear free vibration behavior of FGM microbeams based on the MSGT and the Timoshenko beam theory. They measured the influences of important parameters on the vibrational response of the microbeams.

In this paper, based on the MSGT, the buckling and postbuckling of FGM microbeams have been studied. To this end, the volume fraction of the ceramic and metal phases of FGM microbeams is expressed by using the power law function. In the framework of nonlinear Timoshenko beam theory, the stability equations and associated boundary conditions are derived employing Hamilton's principle. Then, the higher-order governing differential equations are discretized along with three different boundary conditions by using the GDQ method in conjunction with a direct approach without linearization. The current article provides an accurate insight into the effects of length scale parameter, slenderness ratio, material gradient index, and boundary conditions on the buckling and postbuckling behavior of FGM microbeams. The present model explicitly considers the net effect of imperfections of FGM microbeams which is owing to fabrication defects, geometric irregularities, and non-ideal loading. Recognizing the influence of imperfections is an important matter, since imperfections in structures are inescapable and may lead to considerable variations in the stability response. By considering imperfections, instead of sudden bifurcation at a critical axial load, the microbeams show a slight out-of-plane deflection from the prebuckled state to a postbuckled one.

2. Modified strain gradient theory

Based on the modified strain gradient theory of Lam et al. [2003], the stored strain energy U_m in a continuum constructed by a linear elastic material occupying region Ω undergoing infinitesimal deformations is given as

$$U_m = \frac{1}{2} \int_\Omega (\sigma_{ij}\epsilon_{ij} + p_i\gamma_i + \tau_{ijk}^{(1)}\eta_{ijk}^{(1)} + m_{ij}^s\chi_{ij}^s)\, dv, \tag{1a}$$

$$\epsilon_{ij} = \tfrac{1}{2}(u_{i,j} + u_{j,i}), \tag{1b}$$

$$\gamma_i = \epsilon_{mm,i}, \tag{1c}$$

$$\eta_{ijk}^{(1)} = \eta_{ijk}^{(1)} = \eta_{ijk}^s - \tfrac{1}{5}(\delta_{ij}\eta_{mmk}^s + \delta_{jk}\eta_{mmi}^s + \delta_{ki}\eta_{mmj}^s), \qquad \eta_{ijk}^s = \tfrac{1}{3}(\epsilon_{jk,i} + \epsilon_{ki,j} + \epsilon_{ij,k}), \tag{1d}$$

$$\chi_{ij}^s = \tfrac{1}{2}(\theta_{i,j} + \theta_{j,i}), \qquad \theta_i = \tfrac{1}{2}(\mathrm{curl}(u))_i \tag{1e}$$

where ϵ_{ij} denotes the strain tensor and u_i is the components of displacement vector u. Also, γ_i, $\eta_{ijk}^{(1)}$, χ_{ij}^s are dilatation gradient, deviatoric stretch gradient, and symmetric rotation gradient tensors, respectively [Lam et al. 2003]. θ_i stands for the infinitesimal rotation vector θ, and δ is the Kronecker delta.

The constitutive equations corresponding to a linear isotropic elastic material expressed by kinematic parameters effective on the strain energy density are [Lam et al. 2003; Timoshenko and Goodier 1970]

$$\sigma_{ij} = \lambda\mathrm{tr}(\epsilon)\delta_{ij} + 2\mu\epsilon_{ij}, \qquad p_i = 2\mu l_0^2\gamma_i, \qquad \tau_{ijk}^{(1)} = 2\mu l_1^2\eta_{ijk}^{(1)}, \qquad m_{ij}^s = 2\mu l_2^2\chi_{ij}^s \tag{2}$$

in which p, $\tau^{(1)}$, and m^s present the higher-order stresses. Also, l_0, l_1, l_2 stand for three independent material length scale parameters related to the dilatation gradients, deviatoric stretch gradients and symmetric rotation gradients, respectively. Also, in the constitutive equation of the classical stress, λ and μ are two classical material constants of bulk and shear modules, respectively, and are given as [Timoshenko and Goodier 1970; Ke et al. 2010]

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \qquad \mu = \frac{E}{2(1+\nu)}. \tag{3}$$
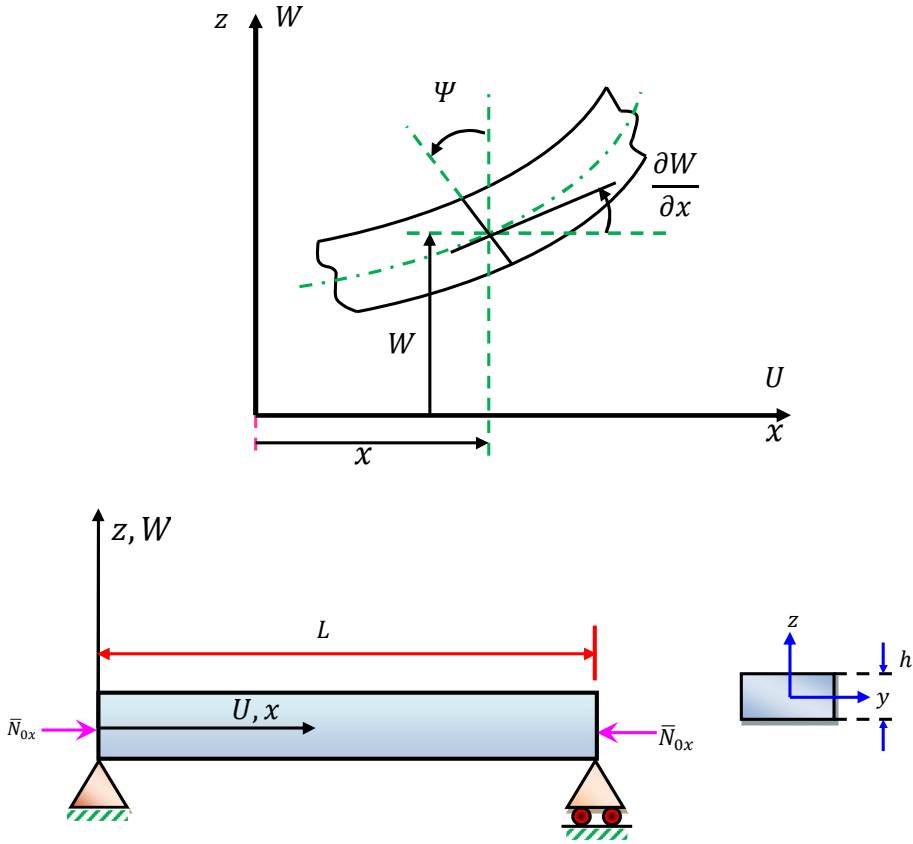
Figure 1. Schematic of a Timoshenko beam: kinematic parameters, coordinate system, geometry and loading.

Accordingly, the theory of [Lam et al. 2003] encompasses higher-order stresses beside the classical stress. These tensors are specified by three independent material length scale parameters and two classical material constants.

3. Material properties of FGM microbeams

As shown in Figure 1, consider a FGM microbeam made of a mixture of ceramic and metal with length L and thickness h undergoing an axial compressive load $\bar{N}_{0x}$. Also, the kinematic parameters and geometry and loading are demonstrated in this figure. The materials at bottom surface ($z = -h/2$) and top surface ($z = h/2$) of the microbeam are supposed to be metal-rich and ceramic-rich, respectively.

Effective material characteristics of the FGM microbeam such as Young's modulus (E) and Poisson's ratio (ν) can be achieved as the following relations

$$E(z) = (E_c - E_m)V_f(z) + E_m, \tag{4}$$

$$\nu(z) = (\nu_c - \nu_m)V_f(z) + \nu_m, \tag{5}$$

where the subscripts m and c signify metal and ceramic phases, respectively. Amongst the different

existing functions to describe the variation of the volume fractions of constituents, herein a simple power law function is used to explain the volume fraction of the ceramic and metal phases as (see [Fares et al. 2009])

$$V_f(z) = \left(\frac{1}{2} + \frac{z}{h} \right)^k, \tag{6}$$

where k represents the power law index.

4. Mathematical formulation of stability equations and associated boundary conditions

Based on the Timoshenko beam formulation, by introducing $U(x)$, $W(x)$, $W_s(x)$ and $\Psi(x)$ as the axial displacement of the center of sections, lateral deflection of the beam, lateral pre-deformation owing to imperfection, and the rotation angle of the cross section with respect to the vertical direction, the kinematics of microbeams are given as

$$u_1 = U(x) - z\Psi(x), \qquad u_2 = 0, \qquad u_3 = W(x) - W_s(x). \tag{7}$$

It is noteworthy to mention that in this study the net imperfection is treated as a pre-deflection of microbeam in unloaded state. By considering a Timoshenko microbeam under tiny slopes after deflection and possible finite transverse deflection W, the nonlinear strain-displacement relations are approximated by the von Karman relation as

$$\epsilon_{11} = \frac{du_1}{dx} + \frac{1}{2}\left[\left(\frac{dW}{dx}\right)^2 - \left(\frac{dW_s}{dx}\right)^2 \right] = \frac{dU}{dx} - z\frac{d\Psi}{dx} + \frac{1}{2}\left[\left(\frac{dW}{dx}\right)^2 - \left(\frac{dW_s}{dx}\right)^2 \right], \quad \epsilon_{13} = \frac{1}{2}\left(\frac{dW}{dx} - \frac{dW_s}{dx} - \Psi \right). \tag{8}$$

Also, the non-zero constituents of $\boldsymbol{\theta}$, $\boldsymbol{\chi}^s$, $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}^{(1)}$ can be obtained through introducing (7) and (8) into (1c)–(1e) as

$$\theta_2 = -\frac{1}{2}\left(\Psi + \frac{dW}{dx} - \frac{dW_s}{dx} \right), \tag{9a}$$

$$\chi_{12}^s = \chi_{21}^s = -\frac{1}{4}\left(\frac{d\Psi}{dx} + \frac{d^2W}{dx^2} - \frac{d^2W_s}{dx^2} \right), \tag{9b}$$

$$\gamma_1 = \frac{d^2U}{dx^2} - z\frac{d^2\Psi}{dx^2} + \frac{dW}{dx}\cdot\frac{d^2W}{dx^2} - \frac{dW_s}{dx}\cdot\frac{d^2W_s}{dx^2}, \qquad \gamma_3 = -\frac{d\Psi}{dx}, \tag{9c}$$

$$\eta_{111}^{(1)} = \frac{2}{5}\left(\frac{d^2U}{dx^2} - z\frac{d^2\Psi}{dx^2} + \frac{dW}{dx}\cdot\frac{d^2W}{dx^2} - \frac{dW_s}{dx}\cdot\frac{d^2W_s}{dx^2} \right), \qquad \eta_{333}^{(1)} = -\frac{1}{5}\left(\frac{d^2W}{dx^2} - \frac{d^2W_s}{dx^2} - 2\frac{d\Psi}{dx} \right),$$

$$\eta_{113}^{(1)} = \eta_{311}^{(1)} = \eta_{131}^{(1)} = \frac{4}{15}\left(\frac{d^2W}{dx^2} - \frac{d^2W_s}{dx^2} - 2\frac{d\Psi}{dx} \right), \tag{9d}$$

$$\eta_{223}^{(1)} = \eta_{322}^{(1)} = \eta_{232}^{(1)} = -\frac{1}{15}\left(\frac{d^2W}{dx^2} - \frac{d^2W_s}{dx^2} - 2\frac{d\Psi}{dx} \right),$$

$$\eta_{221}^{(1)} = \eta_{212}^{(1)} = \eta_{122}^{(1)} = \eta_{313}^{(1)} = \eta_{133}^{(1)} = \eta_{331}^{(1)} = -\frac{1}{5}\left(\frac{d^2U}{dx^2} - z\frac{d^2\Psi}{dx^2} + \frac{dW}{dx}\cdot\frac{d^2W}{dx^2} - \frac{dW_s}{dx}\cdot\frac{d^2W_s}{dx^2} \right).$$

Finally, by inserting (8) and (9) into (2), the major constituents of the symmetric section of the stress tensor and nonzero constituents of the higher-order stresses are achieved. The strain energy due to the

variation of the classical and higher order stresses with respect to the initial configuration Π_s can be written as

$$
\begin{aligned}
\Pi_s = \frac{1}{2} \int_0^L \Bigg\{ & N_1 \left(\frac{dU}{dx} + \frac{1}{2} \left[\left(\frac{dW}{dx} \right)^2 - \left(\frac{dW_s}{dx} \right)^2 \right] \right) - M_1 \left(\frac{d\Psi}{dx} \right) + Q_1 \left(\frac{dW}{dx} - \frac{dW_s}{dx} - \Psi \right) - P_3 \left(\frac{d\Psi}{dx} \right) \\
& + (P_1 + T_{111}^\tau) \left(\frac{d^2U}{dx^2} + \frac{dW}{dx} \cdot \frac{d^2W}{dx^2} - \frac{dW_s}{dx} \cdot \frac{d^2W_s}{dx^2} \right) - (M_1^p + M_{111}^\tau) \left(\frac{d^2\Psi}{dx^2} \right) \\
& - \frac{4}{3} T_{333}^\tau \left(\frac{d^2W}{dx^2} - \frac{d^2W_s}{dx^2} - 2\frac{d\Psi}{dx} \right) - \frac{Y_{12}}{2} \left(\frac{d\Psi}{dx} + \frac{d^2W}{dx^2} - \frac{d^2W_s}{dx^2} \right) \Bigg\} dx, \quad (10)
\end{aligned}
$$

in which the normal resultant force N_1, shear force Q_1, bending moment M_1, couple moment Y_{12} and other higher-order resultants force and higher order moments in a section attributable to higher-order stresses acting on the section are introduced as follows:

$$
N_1 = \int_A \sigma_{11} \, dA = A_{11} \left(\frac{dU}{dx} + \frac{1}{2} \left[\left(\frac{dW}{dx} \right)^2 - \left(\frac{dW_s}{dx} \right)^2 \right] \right) - B_{11} \frac{d\Psi}{dx},
$$

$$
M_1 = \int_A \sigma_{11} z \, dA = B_{11} \left(\frac{dU}{dx} + \frac{1}{2} \left[\left(\frac{dW}{dx} \right)^2 - \left(\frac{dW_s}{dx} \right)^2 \right] \right) - D_{11} \frac{d\Psi}{dx}, \quad (11a)
$$

$$
Q_1 = \int_A \sigma_{13} \, dA = k_s A_{55} \left(\frac{dW}{dx} - \frac{dW_s}{dx} - \Psi \right), \qquad Y_{12} = \int_A m_{12}^s \, dA = -\frac{l_2^2}{2} A_{55} \left(\frac{d\Psi}{dx} + \frac{d^2W}{dx^2} - \frac{d^2W_s}{dx^2} \right),
$$

$$
P_1 = \int_A p_1 \, dA = 2l_0^2 \left(A_{55} \left[\frac{d^2U}{dx^2} + \frac{dW}{dx} \cdot \frac{d^2W}{dx^2} - \frac{dW_s}{dx} \cdot \frac{d^2W_s}{dx^2} \right] - B_{55} \frac{d^2\Psi}{dx^2} \right),
$$

$$
P_3 = \int_A p_3 \, dA = -2 A_{55} l_0^2 \frac{d\Psi}{dx}, \quad (11b)
$$

$$
T_{111}^\tau = \int_A \tau_{111}^{(1)} \, dA = \frac{4l_1^2}{5} \left(A_{55} \left[\frac{d^2U}{dx^2} + \frac{dW}{dx} \cdot \frac{d^2W}{dx^2} - \frac{dW_s}{dx} \cdot \frac{d^2W_s}{dx^2} \right] - B_{55} \frac{d^2\Psi}{dx^2} \right), \quad (11c)
$$

$$
T_{333}^\tau = \int_A \tau_{333}^{(1)} \, dA = -\frac{2l_1^2}{5} A_{55} \left(\frac{d^2W}{dx^2} - \frac{d^2W_s}{dx^2} - 2\frac{d\Psi}{dx} \right), \quad (11d)
$$

$$
M_1^p = \int_A p_1 z \, dA = 2l_0^2 \left(B_{55} \left[\frac{d^2U}{dx^2} + \frac{dW}{dx} \cdot \frac{d^2W}{dx^2} - \frac{dW_s}{dx} \cdot \frac{d^2W_s}{dx^2} \right] - D_{55} \frac{d^2\Psi}{dx^2} \right), \quad (11e)
$$

$$
M_{111}^\tau = \int_A \tau_{111}^{(1)} z \, dA = \frac{4l_1^2}{5} \left(B_{55} \left[\frac{d^2U}{dx^2} + \frac{dW}{dx} \cdot \frac{d^2W}{dx^2} - \frac{dW_s}{dx} \cdot \frac{d^2W_s}{dx^2} \right] - D_{55} \frac{d^2\Psi}{dx^2} \right) \quad (11f)
$$

where A stands for the cross-sectional areas of the microbeam. The symbol k_s appearing in shear force Q_1 is a correction factor to consider the non-uniformity of shear strain over the microbeam cross-section [Ke et al. 2012; Reddy 2007]. Also, the stiffness constituents in these relations are defined as

$$
\{A_{11}, B_{11}, D_{11}\} = \int_{-h/2}^{h/2} (\lambda(z) + 2\mu(z))\{1, z, z^2\} \, dz, \quad \{A_{55}, B_{55}, D_{55}\} = \int_{-h/2}^{h/2} \mu(z)\{1, z, z^2\} \, dz. \quad (12)
$$

The work Π_P done by the axial force $\bar{N}_{0x}$ can be also written as

$$\Pi_P = \frac{1}{2} \int_0^L \bar{N}_{0x} \left(\frac{dW}{dx} \right)^2 dx. \tag{13}$$

Using (10) and (13) and implementing the principle of virtual work, the governing stability equations and corresponding boundary conditions of an FGM Timoshenko microbeam can be obtained as

$$\frac{dN_1}{dx} - \frac{d^2(P_1 + T^\tau_{111})}{dx^2} = 0, \tag{14a}$$

$$\frac{d}{dx} \left\{ \left(N_1 - \bar{N}_{0x} - \frac{d(P_1 + T^\tau_{111})}{dx} \right) \cdot \frac{dW}{dx} \right\} + \frac{dQ_1}{dx} + \frac{4}{3} \frac{d^2 T^\tau_{333}}{dx^2} + \frac{1}{2} \frac{d^2 Y_{12}}{dx^2} = 0, \tag{14b}$$

$$Q_1 - \frac{d(M_1 + P_3)}{dx} + \frac{d^2(M^p_1 + M^\tau_{111})}{dx^2} + \frac{8}{3} \frac{dT^\tau_{333}}{dx} = 0, \tag{14c}$$

$$\left. \left(N_1 - \frac{d(P_1 + T^\tau_{111})}{dx} \right) \right|_{x=0,L} = 0 \quad \text{or} \quad \delta U|_{x=0,L} = 0, \tag{14d}$$

$$\left. \left((N_1 - \bar{N}_{0x}) \frac{dW}{dx} + Q_1 - \frac{d(P_1 + T^\tau_{111})}{dx} \frac{d^2 W}{dx^2} + \frac{4}{3} \frac{dT^\tau_{333}}{dx} + \frac{1}{2} \frac{dY_{12}}{dx} \right) \right|_{x=0,L} = 0 \quad \text{or} \quad \delta W|_{x=0,L} = 0, \tag{14e}$$

$$\left. \left(M_1 + P_3 - \frac{d(M^p_1 + M^\tau_{111})}{dx} - \frac{8}{3} T^\tau_{333} + \frac{Y_{12}}{2} \right) \right|_{x=0,L} = 0 \quad \text{or} \quad \delta \Psi|_{x=0,L} = 0, \tag{14f}$$

$$(P_1 + T^\tau_{111})|_{x=0,L} = 0 \quad \text{or} \quad \left. \delta \left(\frac{dU}{dx} \right) \right|_{x=0,L} = 0, \tag{14g}$$

$$\left. \left(-(P_1 + T^\tau_{111}) \frac{dW}{dx} + \frac{4}{3} T^\tau_{333} + \frac{Y_{12}}{2} \right) \right|_{x=0,L} = 0 \quad \text{or} \quad \left. \delta \left(\frac{dW}{dx} \right) \right|_{x=0,L} = 0, \tag{14h}$$

$$(M^p_1 + M^\tau_{111})|_{x=0,L} = 0 \quad \text{or} \quad \left. \delta \left(\frac{d\Psi}{dx} \right) \right|_{x=0,L} = 0. \tag{14i}$$

Accordingly, based on the MSGT, the stability equations and associated boundary conditions of a size-dependent FGM microbeam are derived. In effect, these relations can be reduced to the stability equations and associated boundary conditions of a FGM Timoshenko microbeam achieved by the MCST and CT, if two or three material length scale parameters set to be zero, respectively [Ke et al. 2012].

We introduce the parameters

$$\xi = \frac{x}{L}, \quad \eta = \frac{L}{h}, \quad (u, w, w_s) = \frac{(U, W, W_s)}{h}, \quad \psi = \Psi, \quad N_0 = \frac{\bar{N}_{0x}}{A_{110}}, \quad (\ell_0, \ell_1, \ell_2) = \frac{(l_0, l_1, l_2)}{h}, \tag{15a}$$

$$(a_{11}, a_{55}, b_{11}, b_{55}, d_{11}, d_{55}) = \left(\frac{A_{11}}{A_{110}}, \frac{A_{55}}{A_{110}}, \frac{B_{11}}{A_{110}h}, \frac{B_{55}}{A_{110}h}, \frac{D_{11}}{A_{110}h^2}, \frac{D_{55}}{A_{110}h^2} \right), \tag{15b}$$

$$c_1 = (2\ell_0^2 + \tfrac{4}{5}\ell_1^2), \quad c_2 = (\tfrac{8}{15}\ell_1^2 + \tfrac{1}{4}\ell_2^2), \quad c_3 = (\tfrac{16}{15}\ell_1^2 - \tfrac{1}{4}\ell_2^2), \quad c_4 = (2\ell_0^2 + \tfrac{32}{15}\ell_1^2 + \tfrac{1}{4}\ell_2^2) \tag{15c}$$

which A_{110} denotes the value of A_{11} of a homogeneous metal beam. The stability equations of the FGM microbeam can be rewritten as

$$a_{11}\frac{d^2u}{d\xi^2}-\frac{a_{55}c_1}{\eta^2}\frac{d^4u}{d\xi^4}-b_{11}\frac{d^2\psi}{d\xi^2}+\frac{b_{55}c_1}{\eta^2}\frac{d^4\psi}{d\xi^4}+\bar{Z}_1=0, \qquad (16\text{a})$$

$$k_sa_{55}\left(\frac{d^2w}{d\xi^2}-\frac{d^2w_s}{d\xi^2}-\eta\frac{d\psi}{d\xi}\right)-\frac{a_{55}c_2}{\eta^2}\left(\frac{d^4w}{d\xi^4}-\frac{d^4w_s}{d\xi^4}\right)+\frac{a_{55}c_3}{\eta}\frac{d^3\psi}{d\xi^3}-N_0\frac{d^2w}{d\xi^2}+\bar{Z}_2=0, \qquad (16\text{b})$$

$$k_sa_{55}\eta\left(\frac{dw}{d\xi}-\frac{dw_s}{d\xi}-\eta\psi\right)-b_{11}\frac{d^2u}{d\xi^2}+(d_{11}+c_4a_{55})\frac{d^2\psi}{d\xi^2}+\frac{a_{55}c_3}{\eta}\left(\frac{\partial^3w}{d\xi^3}-\frac{d^3w_s}{d\xi^3}\right)$$

$$+\frac{c_1}{\eta^2}\left[b_{55}\frac{d^4u}{d\xi^4}-d_{55}\frac{d^4\psi}{d\xi^4}\right]+\bar{Z}_3=0, \qquad (16\text{c})$$

where

$$\bar{Z}_1=\frac{a_{11}}{\eta}\frac{dw}{d\xi}\frac{d^2w}{d\xi^2}-\frac{a_{55}c_1}{\eta^3}\left(\frac{dw}{d\xi}\frac{d^4w}{d\xi^4}+3\frac{d^2w}{d\xi^2}\frac{d^3w}{d\xi^3}\right)-\frac{a_{11}}{\eta}\frac{dw_s}{d\xi}\frac{d^2w_s}{d\xi^2}$$

$$+\frac{a_{55}c_1}{\eta^3}\left(\frac{dw_s}{d\xi}\frac{d^4w_s}{d\xi^4}+3\frac{d^2w_s}{d\xi^2}\frac{d^3w_s}{d\xi^3}\right), \qquad (17\text{a})$$

$$\bar{Z}_2=\frac{a_{11}}{\eta}\left[\frac{du}{d\xi}\frac{d^2w}{d\xi^2}+\frac{d^2u}{d\xi^2}\frac{dw}{\partial\xi}+\frac{3}{2\eta}\left(\frac{dw}{d\xi}\right)^2\frac{d^2w}{d\xi^2}\right]+\frac{b_{55}c_1}{\eta^3}\left[\frac{dw}{d\xi}\frac{d^4\psi}{d\xi^4}+\frac{d^2w}{d\xi^2}\frac{d^3\psi}{d\xi^3}\right]$$

$$-\frac{a_{55}c_1}{\eta^3}\left[\frac{d^4u}{d\xi^4}\frac{dw}{d\xi}+\frac{d^3u}{d\xi^3}\frac{d^2w}{d\xi^2}+\frac{1}{\eta}\left(\frac{d^2w}{d\xi^2}\right)^3+\frac{4}{\eta}\frac{dw}{d\xi}\frac{d^2w}{d\xi^2}\frac{d^3w}{d\xi^3}+\frac{1}{\eta}\left(\frac{dw}{d\xi}\right)^2\frac{d^4w}{d\xi^4}\right]$$

$$-\frac{b_{11}}{\eta}\left(\frac{d^2w}{d\xi^2}\frac{d\psi}{d\xi}+\frac{dw}{d\xi}\frac{d^2\psi}{d\xi^2}\right)-\frac{a_{11}}{2\eta^2}\left(\frac{dw_s}{d\xi}\right)^2\frac{\partial^2w}{d\xi^2}-\frac{a_{11}}{\eta^2}\frac{dw}{d\xi}\frac{dw_s}{d\xi}\frac{d^2w_s}{d\xi^2}$$

$$+\frac{a_{55}c_1}{\eta^4}\left(\frac{d^2w}{d\xi^2}\frac{dw_s}{d\xi}\frac{d^2w_s}{d\xi^2}+3\frac{d^2w_s}{d\xi^2}\frac{dw}{d\xi}\frac{d^3w_s}{d\xi^3}+\left(\frac{d^2w_s}{d\xi^2}\right)^2\frac{d^2w}{d\xi^2}+\frac{dw}{d\xi}\frac{dw_s}{d\xi}\frac{d^4w_s}{d\xi^4}\right), \qquad (17\text{b})$$

$$\bar{Z}_3=\frac{b_{55}c_1}{\eta^3}\left[3\frac{d^2w}{d\xi^2}\frac{d^3w}{d\xi^3}+\frac{dw}{d\xi}\frac{\partial^4w}{d\xi^4}\right]-\frac{b_{11}}{\eta}\frac{dw}{d\xi}\frac{d^2w}{d\xi^2}+\frac{b_{11}}{\eta}\frac{dw_s}{d\xi}\frac{d^2w_s}{d\xi^2}$$

$$-\frac{b_{55}c_1}{\eta^3}\left(\frac{dw_s}{d\xi}\frac{d^4w_s}{d\xi^4}+3\frac{d^2w_s}{d\xi^2}\frac{d^3w_s}{d\xi^3}\right). \qquad (17\text{c})$$

Also, depending on the type of end conditions, the boundary conditions of the FGM microbeam can be as follows:

For clamped (C) end condition

$$u=w=\psi=\frac{du}{d\xi}=\frac{dw}{d\xi}=\frac{d\psi}{d\xi}=0. \qquad (18)$$

For simply-supported (SS) end condition

$$u = w = a_{55}\left(\frac{d^2u}{d\xi^2} + \frac{1}{\eta}\frac{dw}{d\xi}\frac{d^2w}{d\xi^2}\right) - b_{55}\frac{d^2\psi}{d\xi^2}$$

$$= b_{55}\left(\frac{d^2u}{d\xi^2} + \frac{1}{\eta}\frac{dw}{d\xi}\frac{d^2w}{d\xi^2} - \frac{1}{\eta}\frac{dw_s}{d\xi}\frac{d^2w_s}{d\xi^2}\right) - d_{55}\frac{d^2\psi}{d\xi^2} = 0, \qquad (19a)$$

$$-c_1\left(\frac{a_{55}}{\eta^3}\frac{d^2w}{d\xi^2}\left(\frac{dw}{d\xi}\right)^2 - \frac{a_{55}}{\eta^3}\frac{d^2w_s}{d\xi^2}\frac{dw_s}{d\xi}\frac{dw}{d\xi} + \frac{a_{55}}{\eta^2}\frac{dw}{d\xi}\frac{d^2u}{d\xi^2} - \frac{b_{55}}{\eta^2}\frac{dw}{d\xi}\frac{d^2\psi}{d\xi^2}\right)$$

$$- \frac{a_{55}c_2}{\eta^3}\frac{d^2w}{d\xi^2} + a_{55}c_3\frac{d\psi}{d\xi} = 0, \qquad (19b)$$

$$b_{11}\left(\frac{du}{d\xi} + \frac{1}{2\eta}\left(\frac{dw}{d\xi}\right)^2 - \frac{1}{2\eta}\left(\frac{dw_s}{d\xi}\right)^2\right) - d_{11}\frac{d\psi}{d\xi} + \frac{d_{55}c_1}{\eta^2}\frac{d^3\psi}{d\xi^3} + \frac{a_{55}c_3}{\eta}\frac{d^2w}{d\xi^2} - a_{55}c_4\frac{d\psi}{d\xi}$$

$$+ c_1\left(-\frac{b_{55}}{\eta^2}\frac{d^3u}{d\xi^3} - \frac{b_{55}}{\eta^3}\frac{dw}{d\xi}\frac{d^3w}{d\xi^3} + \frac{b_{55}}{\eta^3}\frac{dw_s}{d\xi}\frac{d^3w_s}{d\xi^3} - \frac{b_{55}}{\eta^3}\left(\frac{d^2w}{d\xi^2}\right)^2 + \frac{b_{55}}{\eta^3}\left(\frac{d^2w_s}{d\xi^2}\right)^2\right) = 0. \qquad (19c)$$

5. Numerical solution

Various numerical techniques are available to solve the present stability equations and associated boundary conditions. Herein, the GDQ method [Shu 2000] is utilized to discretize the stability equations and boundary conditions. This technique has exhibited a great potential in solving large partial differential equations [Ansari et al. 2012a; 2012b]. In this work, for sake of briefness, we avoid presenting the discretized counterparts of stability equations and corresponding boundary conditions. Substituting the equations of the boundary conditions into the equations of the system in the boundaries leads to a set of nonlinear equations of the domain coupled with the boundary as

$$\boldsymbol{F} : \mathbb{R}^{3n+1} \to \mathbb{R}^{3n}, \qquad \boldsymbol{F}(\boldsymbol{v}, N_0) = 0, \qquad \boldsymbol{v} = \{\boldsymbol{u}^T, \boldsymbol{w}^T, \boldsymbol{\psi}^T\}^T \qquad (20)$$

in which n is the number of grid points in the GDQ discritization; $\boldsymbol{v}$ and N_0 denote the field variable vector dictating the buckling deformation and the buckling load, respectively. This nonlinear load-deflection equation is rather large and can be hardly treated through a linearization scheme. Therefore, another solution strategy is adopted herein which does not need any linearization. Since the buckling load N_0 itself is unknown, the number of unknown variables is one more than the number of the equations appearing in (20). To rectify this problem, a normalizing equation is added to (20) to convert the present eigenvalue problem to a set of nonlinear equations of the form

$$\begin{aligned} \boldsymbol{F}(\boldsymbol{v}, N_0) &= 0, \\ \boldsymbol{v}^T\boldsymbol{v} - c &= 0. \end{aligned} \qquad (21)$$

Now, the preceding nonlinear equations can be solved by implementing the Newton method, provided that the initial values are selected appropriately. To this end, the equations are solved first by dropping the nonlinear terms. Then, the linear response is imparted to the nonlinear equations in Newton's method as initial values so as to obtain eigenpairs of the nonlinear model.

6. Results and discussion

In this section, the buckling and postbuckling analyses of FGM Timoshenko microbeams undergoing an axial compressive load with different boundary conditions are represented. FGM microbeams made of a mixture of aluminum (Al) and ceramic (SiC) with material properties of $E_m = 70$ GPa, $\nu_m = 0.3$, $\rho_m = 2702$ kg/m^3 for Al, and $E_c = 427$ GPa, $\nu_c = 0.17$, $\rho_c = 3100$ kg/m^3 for SiC are considered. The materials at bottom surface and top surface of the microbeam are considered to be metal-rich and ceramic-rich, respectively.

In order to evaluate the length scale parameters, it is necessary to have experimental data of a homogeneous epoxy or FGM microbeam. Lam et al. [2003] experimentally approximated the length scale parameter of an isotropic homogeneous microbeam as $l = 17.6\,\mu$m. To our best knowledge, there is no experimental report concerned with FGM microbeams in the literature. To have a quantitative analysis on the size effect of the FGM microbeams, the values of each length scale parameter are set to be $l = 15\,\mu$m.

In the following subsections, based upon the above-mentioned assumptions, firstly, the size-dependent buckling behavior of FGM microbeams is described; then the postbuckling analysis of FGM microbeams is performed. Numerical analyses are accomplished for FGM microbeams with three commonly-used end conditions including SS-SS, C-SS and C-C boundary conditions.

Buckling analysis. Table 1 reports the first five non-dimensional buckling loads predicted by the MSGT for C-C, C-SS, and SS-SS microbeams with different material gradient indexes. It is seen that non-dimensional buckling loads decreases with the rise of material gradient index. Hence, the stability of

Boundary conditions	Mode	Ceramic	$k = 0.1$	$k = 1$	$k = 10$	Metal
	1	0.1119	0.107	0.0864	0.067	0.0622
	2	0.1756	0.1679	0.1351	0.104	0.0967
C-C	3	0.2704	0.2585	0.2075	0.1588	0.1477
	4	0.3321	0.3174	0.2544	0.194	0.1805
	5	0.4190	0.4004	0.3206	0.2439	0.2269
	1	0.0651	0.0623	0.0504	0.0393	0.0364
	2	0.1477	0.1412	0.1139	0.0879	0.0817
C-SS	3	0.2292	0.2191	0.1761	0.1351	0.1255
	4	0.3035	0.2901	0.2327	0.1777	0.1653
	5	0.3786	0.3618	0.2898	0.2207	0.2053
	1	0.035	0.0335	0.0273	0.0212	0.0197
	2	0.1119	0.107	0.0864	0.067	0.0622
SS-SS	3	0.1939	0.1854	0.1492	0.1148	0.1067
	4	0.2704	0.2585	0.2075	0.1588	0.1477
	5	0.3441	0.3289	0.2636	0.201	0.1869

Table 1. First five non-dimensional buckling loads for different gradient material indexes (k) of FGM microbeams with C-C, C-SS, and SS-SS end conditions ($h/l = 2$, $L/h = 12$).

Boundary conditions	Mode	$h/l=1$	$h/l=2$	$h/l=3$	$h/l=6$	$h/l=8$	$h/l=16$	CT
C-C	1	0.1989	0.0864	0.0525	0.0276	0.0235	0.0194	0.0178
	2	0.2891	0.1351	0.0882	0.0497	0.0429	0.0359	0.0331
	3	0.4468	0.2075	0.141	0.0837	0.0731	0.0618	0.0573
	4	0.583	0.2544	0.1744	0.1075	0.0947	0.081	0.0755
	5	0.7831	0.3206	0.2181	0.1363	0.1207	0.1038	0.097
C-SS	1	0.1284	0.0504	0.0294	0.0149	0.0126	0.0103	0.0094
	2	0.2479	0.1139	0.0721	0.0392	0.0337	0.028	0.0258
	3	0.3727	0.1761	0.1178	0.0682	0.0592	0.0497	0.046
	4	0.5158	0.2327	0.1593	0.0966	0.0848	0.0721	0.0671
	5	0.6836	0.2898	0.1985	0.1234	0.109	0.0934	0.0872
SS-SS	1	0.0792	0.0273	0.0153	0.0075	0.0063	0.0052	0.0047
	2	0.1989	0.0864	0.0526	0.0276	0.0235	0.0194	0.0178
	3	0.3159	0.1492	0.0975	0.0548	0.0473	0.0395	0.0364
	4	0.4468	0.2075	0.141	0.0837	0.0731	0.0618	0.0573
	5	0.6013	0.2636	0.181	0.1113	0.098	0.0837	0.078

Table 2. First five non-dimensional buckling loads for different non-dimensional length scale parameters (h/l) of FGM microbeams with C-C, C-SS, and S-SS end conditions ($k = 1$, $L/h = 12$).

fully ceramic microbeams is larger than that of microbeams with lower material gradient indexes. Also, it is seen that the values of the critical buckling load in microbeams with C-C end conditions are more than those for SS-SS and C-SS counterparts which signifies that C-C microbeams are more stable than other counterparts.

Table 2 is represented to highlight the effect of the non-dimensional length scale parameter h/l on the first five non-dimensional buckling loads of FGM microbeams with three different boundary conditions. It is seen that the value of non-dimensional buckling load reduces with the increase of the non-dimensional length scale parameter, in a way that at lower non-dimensional length scale parameters, this decrement is very considerable, while when the values increases, there is no prominent changes. Accordingly, the CT underestimates the non-dimensional critical buckling loads, especially at lower non-dimensional length scale parameters.

Table 3 reveals the first five non-dimensional buckling loads corresponding to different slenderness ratios L/h of FGM microbeams with three boundary conditions. It is shown that an increase in the slenderness ratio leads to lower values of non-dimensional buckling loads in a way that this reduction is more evident in higher buckling modes. Accordingly, the possibility of buckling increases at higher slenderness ratios and microbeams with lower slenderness ratios can resist more axial compressive loads.

Figure 2 compares the non-dimensional critical buckling loads predicted by the MSGT, MCST and CT versus the non-dimensional length scale parameter h/l. The figure is plotted for FGM microbeams with SS-SS, C-SS and C-C boundary conditions. With an increase in non-dimensional length scale parameter, the non-dimensional critical buckling load shows a downward trend. Also, it is seen that non-dimensional

Boundary conditions	Mode	$L/h = 8$	$L/h = 10$	$L/h = 12$	$L/h = 16$	$L/h = 20$
	1	0.0992	0.0709	0.0525	0.0316	0.0209
	2	0.1476	0.1128	0.0882	0.057	0.0392
C-C	3	0.2195	0.1739	0.141	0.0966	0.0691
	4	0.2629	0.2108	0.1744	0.1249	0.0928
	5	0.3262	0.2613	0.2181	0.1605	0.1225
	1	0.0597	0.0408	0.0294	0.0171	0.0111
	2	0.1271	0.0944	0.0721	0.045	0.0303
C-SS	3	0.1892	0.1477	0.1178	0.0785	0.055
	4	0.2428	0.1941	0.1593	0.1119	0.0816
	5	0.297	0.2385	0.1985	0.1443	0.1087
	1	0.0331	0.0217	0.0153	0.0086	0.0055
	2	0.0992	0.0709	0.0526	0.0317	0.0209
SS-SS	3	0.1633	0.1248	0.0975	0.0629	0.0432
	4	0.2195	0.1739	0.141	0.0966	0.0691
	5	0.2725	0.2187	0.181	0.1295	0.0961

Table 3. First five non-dimensional buckling loads for different slenderness ratios (L/h) of FGM microbeams with C-C, C-SS, and SS-SS end conditions ($h/l = 3$, $k = 1$).

critical buckling loads predicted by the CT and MCST are lower than those of the MSGT; in other words, the CT and MCST underestimate the non-dimensional critical buckling loads. Moreover, the difference between the values predicted by classical and non-classical theories get really pronounced at small values of non-dimensional length scale parameter which confirms the importance of using the MSGT theory at lower non-dimensional length scale parameters. However, it is seen that as the values of the non-dimensional length scale parameter grow, the difference between the values predicted by classical and non-classical theories fades. Furthermore, non-dimensional critical buckling loads of the microbeam with C-C end conditions are significantly more than those for other counterparts.

Postbuckling analysis. Figures 3 to 6 are represented to provide an accurate insight into the postbuckling behavior of FGM microbeams with C-C, C-SS and SS-SS boundary conditions. All figures illustrate the non-dimensional deflection of FGM microbeams versus the non-dimensional applied axial load for FGM microbeams with C-C, C-SS and SS-SS boundary conditions.

Figure 3 is depicted to highlight the influence of the material gradient index k on the postbuckling behavior of FGM microbeams with C-C, C-SS and SS-SS boundary conditions. As can be seen, the non-dimensional deflection soars sharply with increasing the applied axial load. An increase in the material gradient index shifts the graphs to the left-hand side in all types of boundary conditions and induces more large postbuckled deflections which indicates that the stability decreases. Moreover, when the type of boundary conditions comes to SS-SS microbeams, the graphs get closer which signify that the effect of material gradient index is more prominent in microbeams with C-C end conditions. Besides, the slope of graphs in SS-SS microbeams is much higher than that of microbeams with C-SS and C-C end conditions.
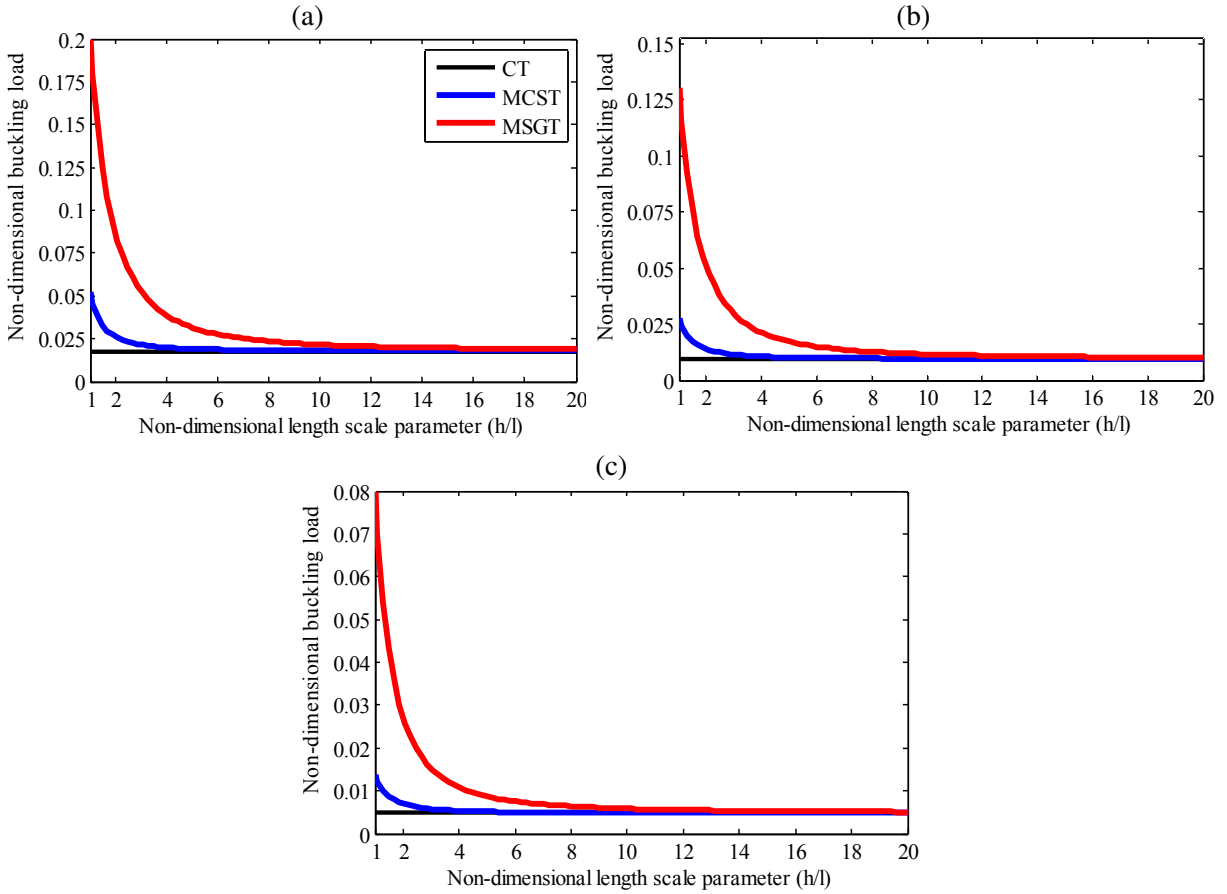
Figure 2. Non-dimensional critical buckling loads predicted by the MSGT, MCST and CT versus non-dimensional length scale parameter (h/l) of FGM microbeams with C-C, C-SS, and SS-SS end conditions ($k = 1$, $L/h = 12$) (a) C-C microbeam, (b) C-SS microbeam, and (c) SS-SS microbeam.

Effect of the non-dimensional length scale parameter h/l on the non-dimensional deflection of FGM microbeams with three types of boundary conditions is investigated in Figure 4. It is observed that a decrease in the value of non-dimensional length scale parameter shifts the graphs to the right-hand side. In other words, CT underestimates the non-dimensional critical axial load. Also, by comparing these three plots, it is seen that the difference between graphs is really significant in C-C microbeams which signal the necessity of using the MSGT theory in studying postbuckling behavior of microbeams with C-C end conditions compared to other end conditions.

Figure 5 reveals the influence of slenderness ratio L/h on the postbuckling behavior of FGM microbeams with C-C, C-SS and SS-SS boundary conditions. With increasing the slenderness ratio of FGM microbeams, the graphs move to the left-hand side which indicates that the stability declines. Also, it is observed that at higher slenderness ratios, a small rise in the applied axial load leads to a sharp increase in the non-dimensional deflection, while at lower slenderness ratios, the non-dimensional

Figure 3. Non-dimensional deflection versus the non-dimensional axial load for different material gradient indexes (k) of FGM microbeams with C-C, C-SS, and SS-SS end conditions ($h/l = 2$, $L/h = 12$) (a) C-C microbeam, (b) C-SS microbeam, and (c) SS-SS microbeam.

deflection increases gradually. So, microbeams with lower slenderness ratios are much stronger than those with larger ones. Furthermore, the graphs are comparatively closer in microbeams with SS-SS end conditions. Hence, microbeams with C-C boundary conditions are more strongly affected by slenderness ratio compared to SS-SS and C-SS counterparts.

Figure 6 compares the non-dimensional deflection of FGM microbeams predicted by the MSGT and CT versus non-dimensional axial load for the first three modes. As it was mentioned earlier, here it is seen that the CT underestimates the non-dimensional critical axial load, especially at higher modes which confirms the necessity of using the MSGT theory at higher postbuckling modes.

Figure 7 illustrates the effect of the imperfection on the non-dimensional deflection of the microbeams versus non-dimensional applied load for first three postbuckling modes of FGM microbeams with different end conditions. Imperfection is treated here by the pre-deformation W_s of the microbeam in its unstressed state. γ is a small dimensionless parameter proportional to the maximum pre-deformation $W_s^{\max}$ at a certain characteristic length L, i.e., $W_s^{\max} = \gamma L$. This implies that the magnitude of imperfection (the proportionality factor γ) is reasonably scaled with the structure size. To achieve higher postbuckling deformation mode, the linear buckling deformation with the maximum amplitude of $W_s^{\max}$ corresponding to that mode is chosen as the pre-deformation. As it is demonstrated in this figure, for ideal loading and geometry $\gamma = 0$, there is no deflection at prebuckling region, while as imperfection
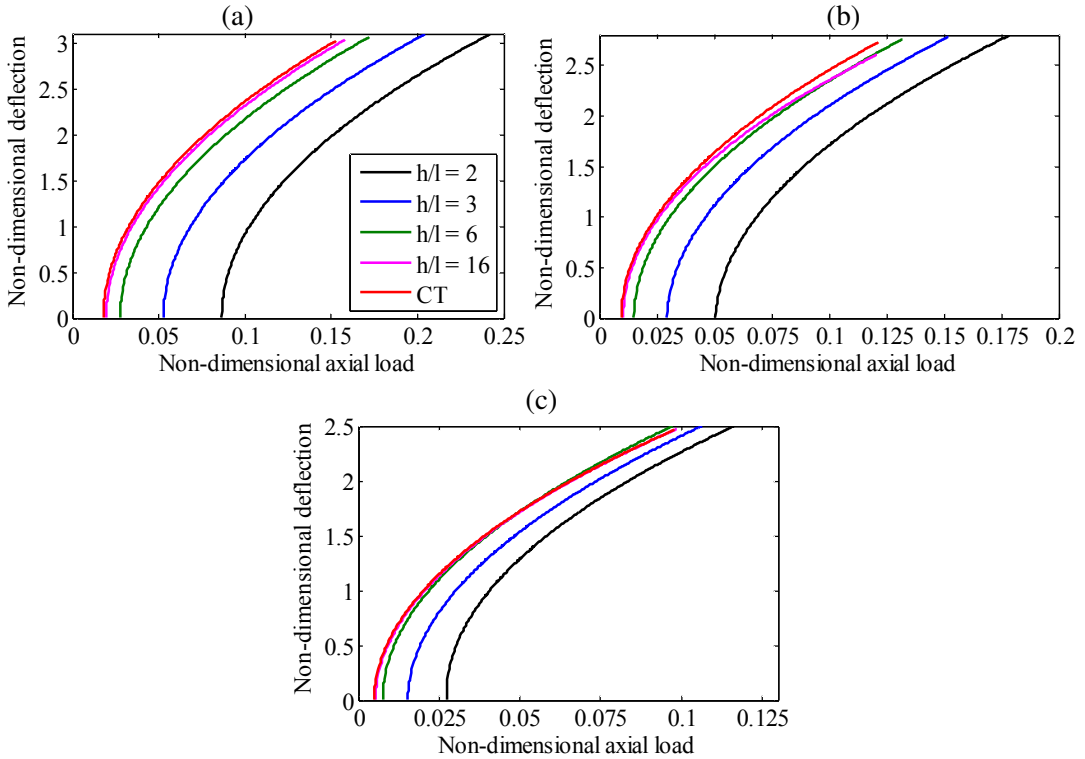
Figure 4. Non-dimensional deflection versus the non-dimensional axial load for different non-dimensional length scale parameter (h/l) of FGM microbeams with C-C, C-SS, and SS-SS end conditions ($k = 1$, $L/h = 12$) (a) C-C microbeam, (b) C-SS microbeam, and (c) SS-SS microbeam.

grows, the deviation between the ideal microbeam and imperfect counterparts continue to increase. This deviation even grows at higher modes and considerably affects the buckling deflection, especially at critical axial loads. In addition, it is evident that the difference between ideal and non-ideal beam is more prominent at microbeams with C-C end supports which signify that C-C microbeams are more sensitive to imperfections compared to other counterparts.

7. Conclusion

In this work, the buckling and postbuckling of FGM microbeams were studied based on the MSGT. By employing the power law function, the volume fraction of the ceramic and metal phases of the FGM microbeams was described. According to the nonlinear Timoshenko beam theory and Hamilton's principle, the higher-order governing differential equations and corresponding boundary conditions were derived and solved through the GDQ method in conjunction with a direct approach without linearization. The effect of imperfection on the non-dimensional deflection of microbeams was considered. The effects of the length scale parameter, slenderness ratio, material gradient index and boundary conditions on the critical buckling load and postbuckling deflection of FGM Timoshenko microbeams were carefully
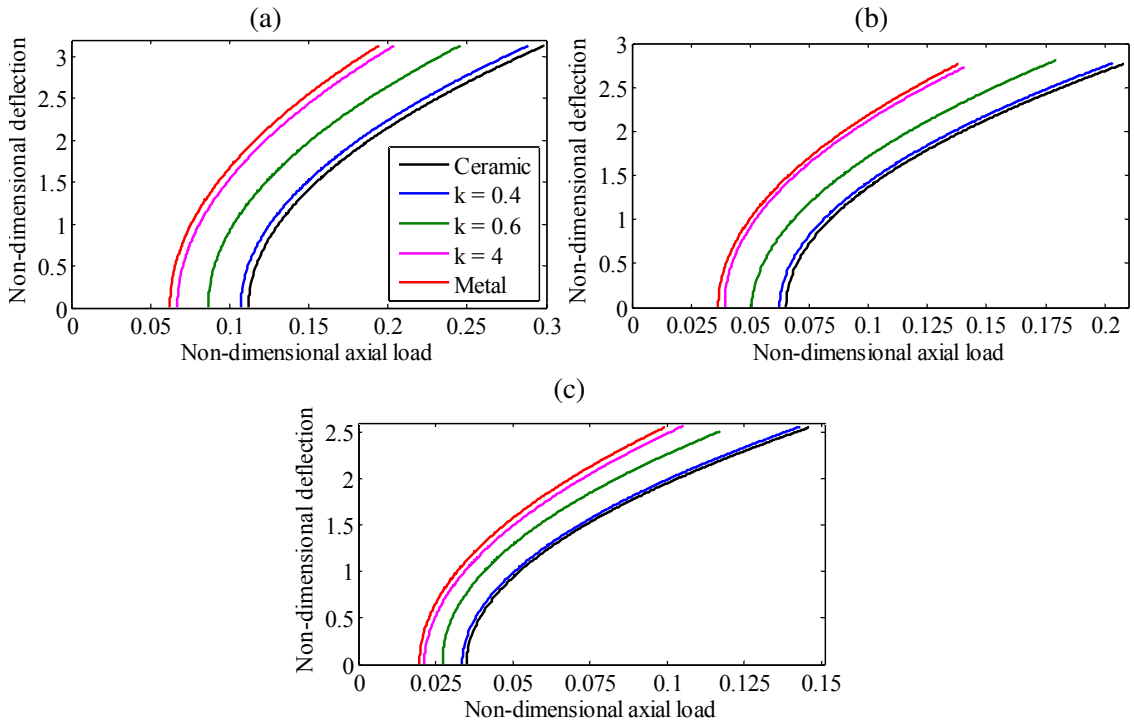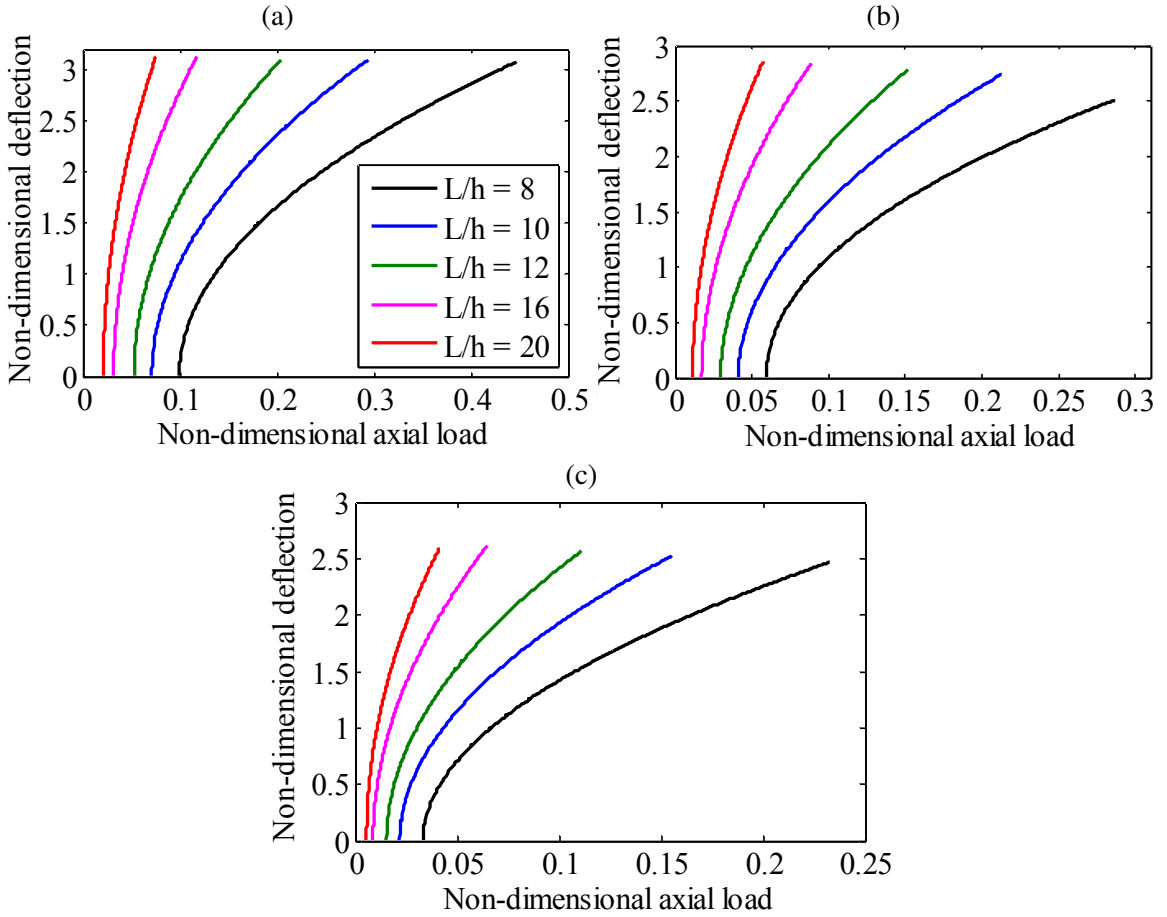
Figure 5. Non-dimensional deflection versus the non-dimensional axial load for different slenderness ratios L/h of FGM microbeams with C-C, C-SS, and SS-SS end conditions ($k = 1$, $h/l = 3$) (a) C-C microbeam, (b) C-SS microbeam, and (c) SS-SS microbeam.

investigated. It was observed that with an increase in material gradient index, the non-dimensional critical buckling loads and postbuckled deflections decreases and increases, respectively. With raising the non-dimensional length scale parameter, the value of non-dimensional critical buckling load reduces; hence, the CT and MCST underestimate the non-dimensional critical buckling loads, especially for C-C microbeams at lower non-dimensional length scale parameters. The larger slenderness ratio, lower values of non-dimensional critical buckling loads; so, microbeams with lower slenderness ratios can resist more axial compressive loads. The effects of material gradient index and slenderness ratio are more prominent in microbeams with C-C end conditions compared to SS-SS and C-SS counterparts. Also, it was seen that the values of the critical buckling load in microbeams with C-C end conditions are more that those for SS-SS and C-SS counterparts. In addition, we found that the imperfection considerably affects the buckling deflection of FGM microbeams, especially at critical axial loads.
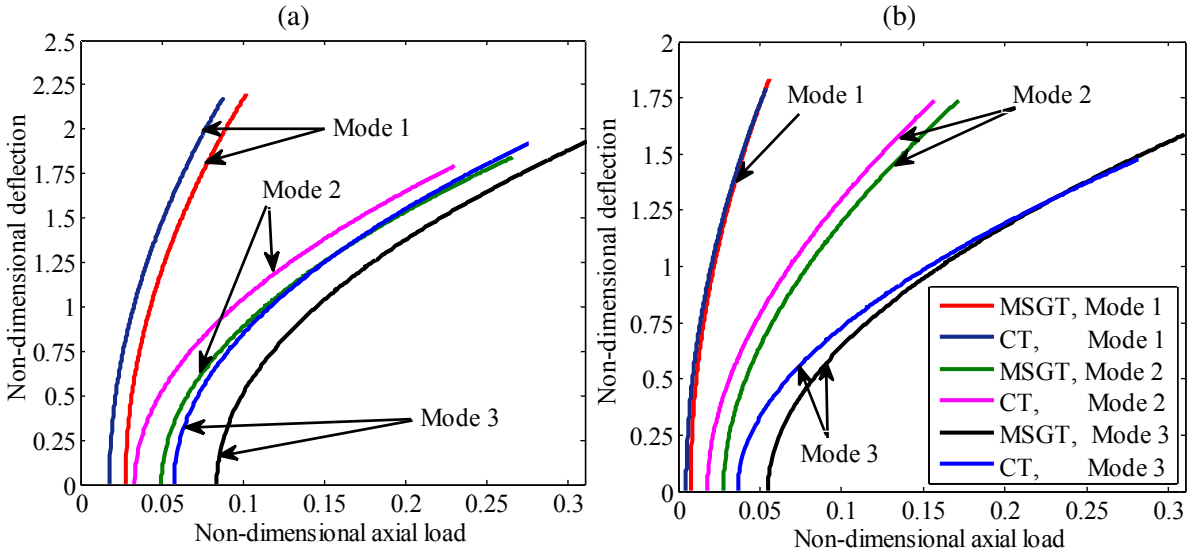
Figure 6. Non-dimensional deflection of FGM microbeams predicted by the MSGT and CT versus non-dimensional axial load for the first three postbuckling modes of FGM microbeams with C-C and SS-SS end conditions ($k = 1$, $h/l = 6$, $L/h = 12$) (a) C-C microbeam, and (b) SS-SS microbeam.
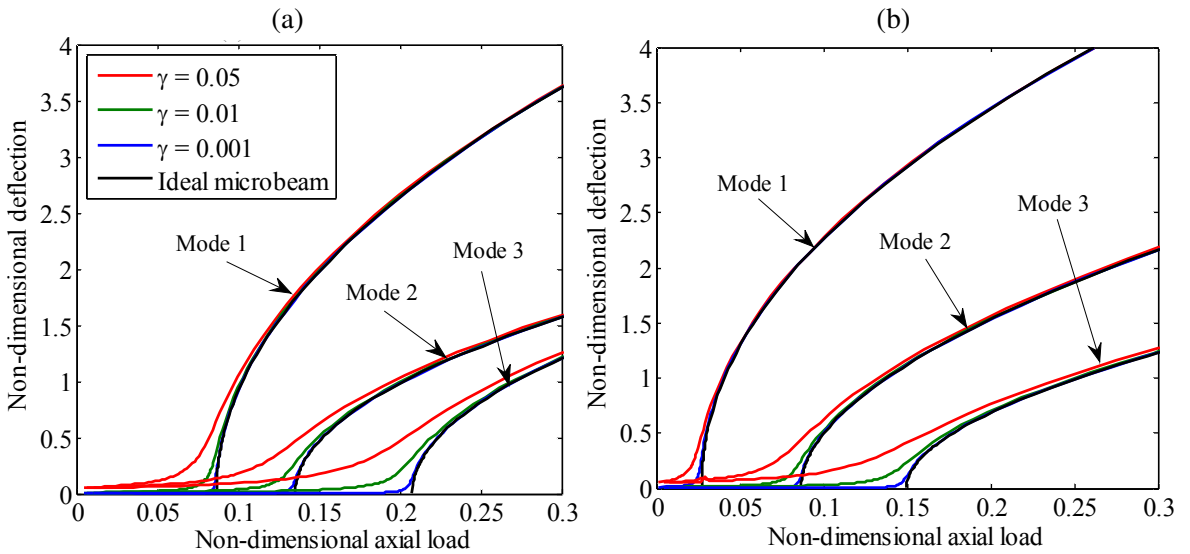

Figure 7. Effect of the imperfection on the non-dimensional deflection versus non-dimensional axial load for first three postbuckling modes of FGM microbeams with C-C and SS-SS end conditions ($L/h = 12$, $k = 1$, $h/l = 2$) (a) C-C microbeam, and (b) SS-SS microbeam.

References

[Aifantis 1999] E. C. Aifantis, "Strain gradient interpretation of size effects", *Int. J. Fract.* **95**:1–4 (1999), 299–314.

[Ansari et al. 2011] R. Ansari, R. Gholami, and S. Sahmani, "Free vibration analysis of size-dependent functionally graded microbeams based on the strain gradient Timoshenko beam theory", *Compos. Struct.* **94** (2011), 221–228.

[Ansari et al. 2012a] R. Ansari, R. Gholami, and M. A. Darabi, "A nonlinear Timoshenko beam formulation based on a strain gradient theory", *J. Mech. Mater. Struct.* **7**:2 (2012), 195–212.

[Ansari et al. 2012b] R. Ansari, R. Gholami, and S. Sahmani, "Study of small scale effects on the nonlinear vibration response of functionally graded Timoshenko microbeams based on the strain gradient theory", *J. Comput. Nonlinear Dyn.* (*ASME*) **7**:3 (2012), Art. ID #031009.

[Asghari et al. 2010] M. Asghari, M. T. Ahmadian, M. H. Kahrobaiyan, and M. Rahaeifard, "On the size-dependent behavior of functionally graded micro-beams", *Mater. Des.* **31** (2010), 2324–2329.

[Asghari et al. 2011] M. Asghari, M. Rahaeifard, M. H. Kahrobaiyan, and M. T. Ahmadian, "The modified couple stress functionally graded Timoshenko beam formulation", *Mater. Des.* **32** (2011), 1435–1443.

[Eringen 1972] A. C. Eringen, "Nonlocal polar elastic continua", *Int. J. Eng. Sci.* **10**:1 (1972), 1–16.

[Fares et al. 2009] M. E. Fares, M. K. Elmarghany, and D. Atta, "An efficient and simple refined theory for bending and vibration of functionally graded plates", *Compos. Struct.* **91** (2009), 296–305.

[Fleck and Hutchinson 1993] N. A. Fleck and J. W. Hutchinson, "A phenomenological theory for strain gradient effects in plasticity", *J. Mech. Phys. Solids* **41**:12 (1993), 1825–1857.

[Fu et al. 2004] Y. Q. Fu, H. J. Du, W. M. Huang, S. Zhang, and M. Hu, "TiNi-based thin films in MEMS applications: a review", *Sens. Actuators A Phys.* **112** (2004), 395–408.

[Gurtin et al. 1998] M. E. Gurtin, J. Weissmuller, and F. Larche, "A general theory of curved deformable interfaces in solids at equilibrium", *Philos. Mag. A* **78**:5 (1998), 1093–1109.

[Hasanyan et al. 2008] D. J. Hasanyan, R. C. Batra, and R. C. Harutyunyan, "Pull-in instabilities in functionally graded microthermoelectromechanical systems", *J. Therm. Stresses* **31** (2008), 1006–1021.

[Kahrobaiyan et al. 2011] M. H. Kahrobaiyan, M. Asghari, M. Rahaeifard, and M. T. Ahmadian, "A nonlinear strain gradient beam formulation", *Int. J. Eng. Sci.* **49**:11 (2011), 1256–1267.

[Ke and Wang 2011] L.-L. Ke and Y.-S. Wang, "Size effect on dynamic stability of functionally graded microbeams based on a modified couple stress theory", *Compos. Struct.* **93**:2 (2011), 342–350.

[Ke et al. 2010] L.-L. Ke, J. Yang, and S. Kitipornchai, "An analytical study on the nonlinear vibration of functionally graded beams", *Meccanica* (*Milano*) **45**:6 (2010), 743–752.

[Ke et al. 2012] L.-L. Ke, Y.-S. Wang, J. Yang, and S. Kitipornchai, "Nonlinear free vibration of size-dependent functionally graded microbeams", *Int. J. Eng. Sci.* **50**:1 (2012), 256–267.

[Koiter 1964] W. T. Koiter, "Couple-stresses in the theory of elasticity: I and II", *P. K. Ned. Akad. Wetensc. B* **67** (1964), 17–29 and 30–44.

[Kong et al. 2009] S. Kong, S. Zhou, Z. Nie, and K. Wang, "Static and dynamic analysis of micro beams based on strain gradient elasticity theory", *Int. J. Eng. Sci.* **47**:4 (2009), 487–498.

[Lam et al. 2003] D. C. C. Lam, F. Yang, A. C. M. Chong, J. Wang, and P. Tong, "Experiments and theory in strain gradient elasticity", *J. Mech. Phys. Solids* **51**:8 (2003), 1477–1508.

[Ma et al. 2008] H. M. Ma, X.-L. Gao, and J. N. Reddy, "A microstructure-dependent Timoshenko beam model based on a modified couple stress theory", *J. Mech. Phys. Solids* **56**:12 (2008), 3379–3391.

[Ma et al. 2010] H. M. Ma, X.-L. Gao, and J. N. Reddy, "A nonclassical Reddy–Levinson beam model based on the a modified couple stress theory", *J. Multiscale Comput. Eng.* **8** (2010), 167–180.

[Ma et al. 2011] H. M. Ma, X.-L. Gao, and J. N. Reddy, "A non-classical Mindlin plate model based on a modified couple stress theory", *Acta Mech.* **220**:1–4 (2011), 217–235.

[Mindlin 1964] R. D. Mindlin, "Micro-structure in linear elasticity", *Arch. Ration. Mech. Anal.* **16** (1964), 51–78.

[Mindlin 1965] R. D. Mindlin, "Second gradient of strain and surface tension in linear elasticity", *Int. J. Solids Struct.* **1**:4 (1965), 417–438.

[Mindlin and Eshel 1968] R. D. Mindlin and N. N. Eshel, "On first strain-gradient theories in linear elasticity", *Int. J. Solids Struct.* **4** (1968), 109–124.

[Mindlin and Tiersten 1962] R. D. Mindlin and H. F. Tiersten, "Effects of couple-stresses in linear elasticity", *Arch. Ration. Mech. Anal.* **11** (1962), 415–448.

[Mirzavand and Eslami 2011] B. Mirzavand and M. R. Eslami, "A closed-form solution for thermal buckling of piezoelectric FGM rectangular plates with temperature-dependent properties", *Acta Mech.* **218** (2011), 87–101.

[Mohammadi-Alasti et al. 2011] B. Mohammadi-Alasti, G. Rezazadeh, A. M. Borgheei, S. Minaei, and R. Habibifar, "On the mechanical behavior of a functionally graded micro-beam subjected to a thermal moment and nonlinear electrostatic pressure", *Compos. Struct.* **93** (2011), 1516–1525.

[Park and Gao 2006] S. K. Park and X.-L. Gao, "Bernoulli–Euler beam model based on a modified couple stress theory", *J. Micromech. Microeng.* **16** (2006), 2355–2359.

[Park and Gao 2008] S. K. Park and X.-L. Gao, "Variational formulation of a modified couple stress theory and its application to a simple shear problem", *Z. Angew. Math. Phys.* **59**:5 (2008), 904–917.

[Reddy 2007] J. N. Reddy, *Theory and analysis of elastic plates and shells*, 2nd ed., CRC, Boca Raton, FL, 2007.

[Reddy and Jinseok 2012] J. N. Reddy and K. Jinseok, "A nonlinear modified couple stress-based third-order theory of functionally graded plates", *Compos. Struct.* **94** (2012), 1128–1143.

[Shu 2000] C. Shu, *Differential quadrature and its application in engineering*, Springer, London, 2000.

[Timoshenko and Goodier 1970] S. Timoshenko and J. N. Goodier, *Theory of elasticity*, 3rd ed., McGraw-Hill, New York, 1970.

[Toupin 1962] R. A. Toupin, "Elastic materials with couple-stresses", *Arch. Ration. Mech. Anal.* **11**:1 (1962), 385–414.

[Toupin 1964] R. A. Toupin, "Theories of elasticity with couple-stress", *Arch. Ration. Mech. Anal.* **17** (1964), 85–112.

[Tsiatas 2009] G. C. Tsiatas, "A new Kirchhoff plate model based on a modified couple stress theory", *Int. J. Solids Struct.* **46** (2009), 2757–2764.

[Wang et al. 2012] B. Wang, J. Zhao, and S. Zhou, "A micro scale Timoshenko beam model based on strain gradient elasticity theory", *Eur. J. Mech. A Solids* **29** (2012), 591–599.

[Witvrouw and Mehta 2005] A. Witvrouw and A. Mehta, "The use of functionally graded poly-SiGe layers for MEMS applications", *Mater. Sci. Forum* **492–493** (2005), 255–260.

[Xia et al. 2010] W. Xia, L. Wang, and L. Yin, "Nonlinear non-classical microscale beams: static bending, postbuckling and free vibration", *Int. J. Eng. Sci.* **48**:12 (2010), 2044–2053.

[Yang et al. 2002] F. Yang, A. C. M. Chong, D. C. C. Lam, and P. Tong, "Couple stress based strain gradient theory for elasticity", *Int. J. Solids Struct.* **39**:10 (2002), 2731–2743.

REZA ANSARI: r_ansari@guilan.ac.ir
Department of Mechanical Engineering, University of Guilan, P.O. Box 3756, Rasht, Iran

MOSTAFA FAGHIH SHOJAEI: mfsh@live.com
Department of Mechanical Engineering, University of Guilan, P.O. Box 3756, Rasht, Iran

VAHID MOHAMMADI: v.n.m@live.com
Department of Mechanical Engineering, University of Guilan, P.O. Box 3756, Rasht, Iran

RAHEB GHOLAMI: rahebgholami@gmail.com
Department of Mechanical Engineering, University of Guilan, P.O. Box 3756, Rasht, Iran

MOHAMMAD ALI DARABI: mad.fluid.darabi@gmail.com
Department of Mechanical Engineering, University of Guilan, P.O. Box 3756, Rasht, Iran

msp

MEASUREMENT OF ELASTIC PROPERTIES OF AISI 52100 ALLOY STEEL BY ULTRASONIC NONDESTRUCTIVE METHODS

MOHAMMAD HAMIDNIA AND FARHANG HONARVAR

In certain applications, it is important to measure the mechanical properties of steel with high accuracy. These measurements are usually performed by destructive processes, in which the test sample is destroyed. Destructive techniques cannot detect the minuscule changes made in the mechanical properties of steel during heat treatment processes. Ultrasonic nondestructive testing is an alternative method that can be used for measuring the mechanical properties of steel more precisely. In this paper, the ultrasonic method is used for measuring the elastic properties of AISI 52100 steel samples which are heat treated at different levels. Each sample has its specific microstructure and hardness due to the heat treatment process it has gone through. The elastic properties of each sample are obtained by measuring the velocities of longitudinal and shear waves in each sample. Comparison of the results obtained from ultrasonic measurements with those available in reference tables shows that the ultrasonic technique can measure the elastic properties of AISI 52100 samples with high accuracy. The ultrasonic technique also shows that it can identify the tempered martensite embrittlement zone in AISI 52100 samples.

1. Introduction

Most mechanical parts used in industrial applications are subjected to some sort of heat treatment. It is essential to verify if the heat-treatment process is done properly and if the component has acquired the desired properties after processing [Badidi Bouda et al. 2000]. Destructive techniques, which are based on variations of physical and mechanical properties of the material, cannot clearly distinguish the slight differences in the microstructure of the material [Vasudevan and Palanichamy 2002]. However, the study of ultrasonic wave propagation in metals and composites can provide information on the microstructure as well as the mechanical and physical properties of the material. The changes made in the atomic structure of materials during a heat-treatment process have been studied before [Ringer et al. 1997] but in this work, we only consider the overall changes observed in the ultrasonic and mechanical properties of the material. Ultrasonic techniques are used not only to detect discontinuities, such as voids, cracks, and inclusions, but also to evaluate such material characteristics as microstructure, grain size, yield strength, fracture toughness, volume fraction of second phase, and residual stress [Kumar et al. 2003]. Quantitative assessment of microstructural changes and mechanical properties can be carried out by measuring ultrasonic parameters such as wave velocity and attenuation [Birks et al. 1991]. The ultrasonic wave velocity depends on elastic constants and density of the body while the attenuation depends on microstructure and crystalline defects [Rokhlin and Matikas 1996; Carreon et al. 2009].

The relationship between ultrasonic and mechanical properties has been studied for various materials, including tempered CA-15 martensitic stainless steel [Hsu et al. 2004], SiC ceramics [ASTM E664-93

2000], and Nimonic alloys [Murthy et al. 2008; 2009]. A few studies have also been conducted on the dependency of ultrasonic wave velocity and microstructure of materials. Some of the materials considered for this purpose are: low-alloy steels [ASTM E92-82 2003], austenitic stainless steels [Vasudevan and Palanichamy 2002; Hakan Gür and Orkun Tuncer 2005], AISI 4140 and 5140 steels, SAE 1040 and 4140 steels [Hakan Gür and Çam 2007], alumina–zirconia ceramics [Rokhlin and Matikas 1996], and aluminum alloy 2024 [Rosen et al. 1985]. The wave velocity and attenuation of SAE 52100 samples quenched at different temperatures have also been studied by Papadakis [1970].

In this paper, we are interested in evaluating the potential of the ultrasonic nondestructive evaluation technique in the measurement of the elastic constants of AISI 52100 alloy steel samples which are hardened to different levels. Because this alloy has enormous industrial applications, the right choice of heat treatment can result in the most appropriate mechanical properties. Our purpose is to demonstrate that small changes in the mechanical properties of samples can be measured by ultrasonic testing. Any other alloy could also be used for this purpose. A number of samples are prepared and processed by different heat treatment schemes. The heat treatment processes are defined based on reference handbooks and advice from industrial experts. Ultrasonic velocity and attenuation measurements have been used to accurately measure the elastic constants, such as the Young's modulus and Poisson's ratio, of these samples in order to differentiate them from the raw sample and from one another.

2. Experiments

Preparation of specimens. The chemical composition of the AISI 52100 alloy steel (DIN 100Cr6 or DIN 1.3505/Ball and Roller bearing steels) used in the present study is given in Table 1. Six cylindrical coupon specimens, 30 mm in diameter and 6 mm in thickness, were prepared from a bar stock of AISI 52100 alloy steel. In order to compare the ultrasonic characteristics and elastic properties of these specimens, they were prepared with dimensions as similar as possible [ASTM E664-93 2000]. The test surfaces of all specimens were finished with a parallelism tolerance of 0.2 mm.

Heat treatment. From the six specimens, one (sample 1) was intact and the remaining five (2–6) were subjected to different heat treatment processes. The details of the heat treatment processes conducted on specimens 2–6 are schematically shown in Figure 1.

All five specimens were first subjected to stress-relief heat treatment in two stages. The temperatures and holding times of these two stress-relief heat treatment stages are shown in Figure 1. After the stress-relief process, all specimens were austenitized to produce austenite with a homogeneous composition of carbon and other alloying elements. The two main factors in austenitization are the temperature and the time, which were selected by the aid of a temperature-time-transformation diagram and experience [Verhoeven 2005].

All specimens were heated to 850 °C at a heating rate of 50 °C/hour and held at this temperature for 30 minutes. They were then quenched in salt bath furnaces to arrive at a temperature of 180 °C. Good heat transfer and natural convection in the salt bath furnaces was used to promote uniformity of specimen

$$C = 1\% \quad Si = 0.25\% \quad Mn = 0.35\% \quad Cr = 1.5\% \quad P = 0.015\%$$

Table 1. Chemical composition of AISI 52100 by weight.
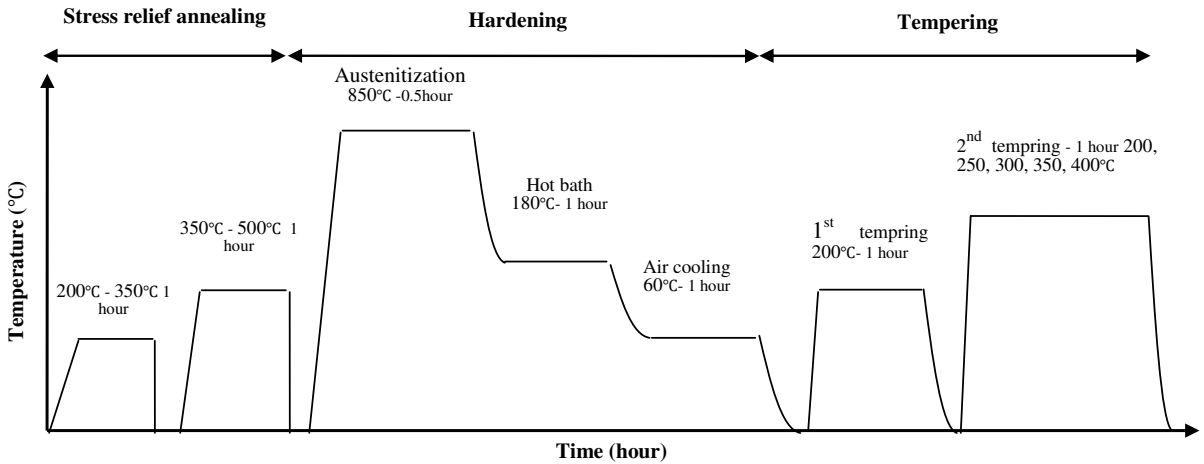
Figure 1. Heat treatment processes.

temperatures [ASM 4 1991]. The specimens were then cooled in air to arrive at 60 °C. Cooling in air at a moderate rate is to prevent large differences in temperature between the outside surface and the center core. All specimens were tempered in two stages: first, all specimens were reheated up to 200 °C for 1 hour, and second, specimens 1–5 were reheated at temperatures of 200 °C, 250 °C, 300 °C, 350 °C, and 400 °C for 1 hour, respectively. The temperature uncertainty of the oven was ±20 °C. Tempering was accomplished by soaking the entire specimen in the furnace for enough time to bring the tempering mechanism to the desired point of completion.

The mechanical properties of tempered steel are affected by a number of variables, including: tempering temperature, duration of tempering, cooling rate from the tempering temperature, and composition of the steel [ASM 4 1991]. In order to investigate the effect of tempering temperature on the ultrasonic and mechanical properties of the specimens, the duration of tempering, cooling rate, and composition of the material were held constant and all other parameters in the stress relief annealing and hardening process were kept as identical as possible.

Hardness measurement. Hardness levels of the specimens were obtained by a Vickers hardness tester using a 10 kg load. A small indentation in the polished surface of each specimen was made using a standard procedure. For each specimen, at least five indentations were made under similar conditions and the average of the diagonal lengths of the indentations were used to compute the Vickers hardness number using the following equation [ASTM E92-82 2003]:

$$VHN = \frac{1.854}{d^2}F \tag{1}$$

where d is the diagonal length of the indentation in mm and F is the applied load in kg. The accuracy in measuring d was estimated to be ±2 μm.

Density measurements. Density was measured using Archimedes' principle:

$$\rho = \frac{W_a}{W_a - W_b}\rho_b, \tag{2}$$

where W_a and W_b are the weights of the specimen in air and water, respectively, and ρ_b is the density of water. All weight measurements were made using a digital weighing scale. The uncertainty in these measurements was ± 0.1 mg. The experiments were repeated five times for each specimen in order to obtain a reliable density value.

Ultrasonic velocity measurements. Figure 2 shows the configuration of the experimental setup for the contact and immersion ultrasonic testing methods. A pulser/receiver (Panametrics 5072PR), a 100 MHz A/D converter (CompuScope14100) with 14-bit resolution and a PC were employed for recording ultrasonic signals.

Ultrasonic wave velocity measurements were carried out for both longitudinal and transverse waves. The longitudinal wave velocity was measured by a 0.25 inch diameter, 10 MHz immersion focused probe (Panametrics V312-N-SU) and shear wave velocity was measured by a 0.5 inch diameter, 5 MHz contact probe (Panametrics V155). Honey was used for coupling the normal beam shear wave probe to the specimens. The time of flight of the longitudinal and transverse waves in each specimen was measured between the second and third backwall echoes. The ultrasonic wave velocity (V) was determined using the following equation [ASTM E494-95 1995]:

$$V = \frac{k \times 2d}{t_m - t_n},\tag{3}$$

where d is the specimen thickness, k is the difference between the echo indices m and n, and t_m and t_n are the transit times of the m-th and n-th echoes, respectively. Figure 3 shows typical specimen time traces of the longitudinal and transverse waves. In the top graph, the backwall echoes of the longitudinal wave probe are not very strong and this is because the normal beam probe is a focused transducer. Due to its smaller wavelength, the shear wave is more sensitive to changes in microstructure [Palanichamy et al. 1995; Murthy et al. 2009].

Variations in ultrasonic wave velocity due to changes in material characteristics and microstructure (except changes in texture) are usually small, and therefore accurate velocity measurements are required for detecting these changes.
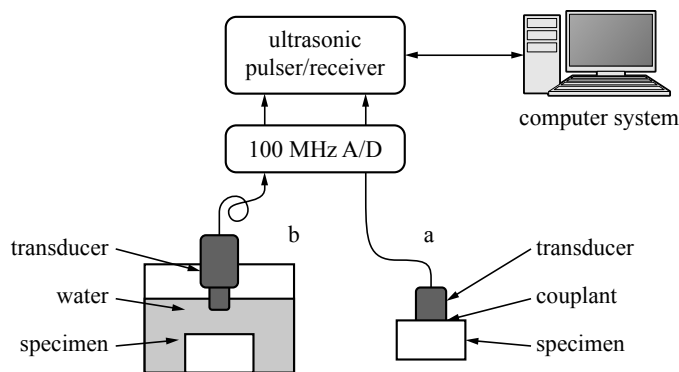
Figure 2. Schematic of the experimental setup: (a) contact method and (b) immersion method.
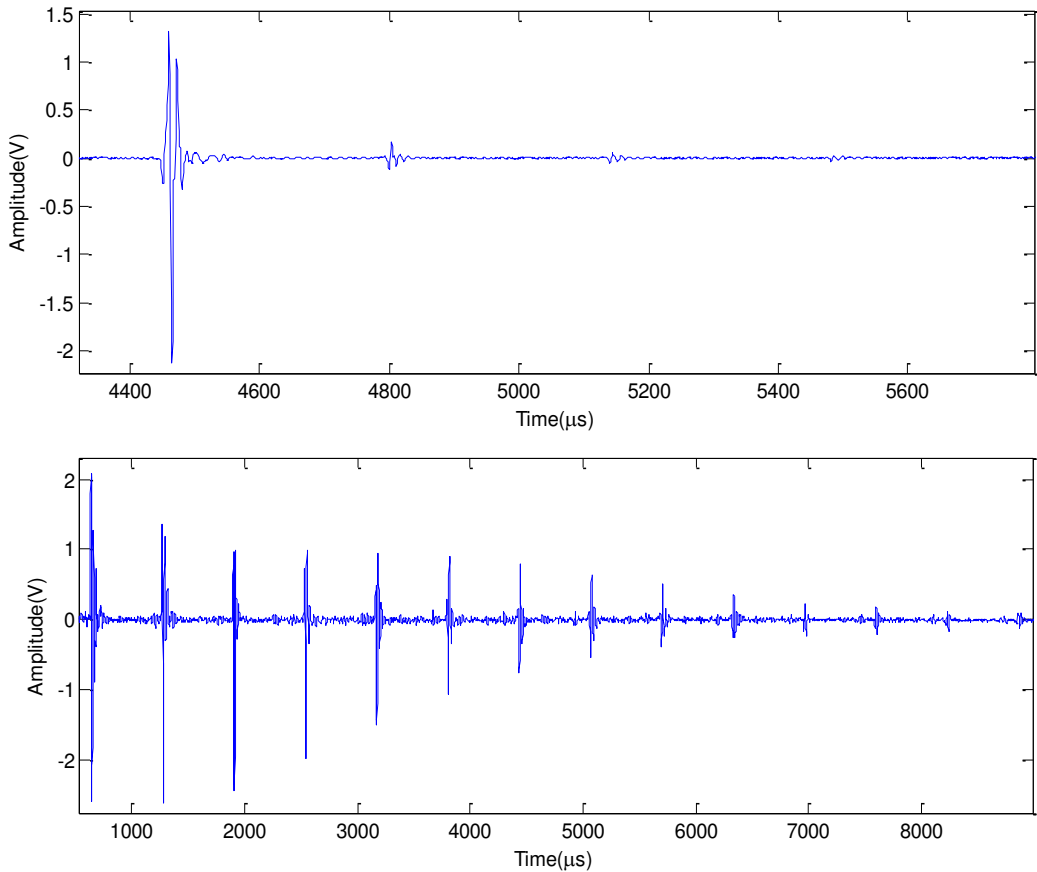
Figure 3. Specimen time traces for longitudinal (top) and transverse (bottom) waves.

The time delay between two consecutive echoes can be measured by three different time-delay measurement techniques: phase-slope, overlap, and cross-correlation [dos Santos Mesquita and Halldórsdóttir 2005]. Among these three techniques, the cross-correlation technique does not require explicit criteria for accepting or rejecting specific features in echoes affected by distortion or low signal-to-noise ratio. Therefore, in this study, the cross-correlation method is used in order to eliminate the need for the somewhat arbitrary criteria required for the other two techniques. Cross-correlation can be performed between two time signals, $x(t)$ and $y(t)$, using computer algorithms which implement the cross-correlation function $R_{xy}(t)$, defined as [Hull et al. 1985]:

$$R_{xy}(t) = \lim_{T \to \infty} \frac{1}{T} \int_0^T x(\tau) y(t + \tau) \, d(\tau), \tag{4}$$

where t is the time index of signals x and y, and τ is the time delay of associated components.

Ultrasonic attenuation measurements. Measurement of attenuation coefficients in materials is useful in applications such as comparison of heat treatments of different lots of material or assessment of environmental degradation of materials [Birks et al. 1991]. The attenuation coefficient of the ultrasonic wave is affected by absorption and scattering. Absorption can be due to dislocation damping, magnetic

resistance, or thermal elasticity, and scattering can be due to grain boundaries, voids, inclusions, second-phase particles, and cracks [Hsu et al. 2004].

The ultrasonic attenuation coefficient (α) is calculated from the equation

$$\alpha \ (\text{dB/mm}) = \frac{20\log A_1/A_2}{2T} \tag{5}$$

(see [ASTM E664-93 2000]), where A_1 and A_2 are peak amplitudes of the first and second transmitted echoes, and T is the test specimen thickness in millimeters.

Determination of elastic properties. Elastic properties are determined by the aggregate response of the interatomic forces among all atoms of the metal. Therefore, the presence of small quantities of solute atoms in dilute alloys or their rearrangement by heat treatment will have relatively little effect on the absolute values of elastic constants [Brandes and Brook 1998]. The determination of longitudinal and transverse wave velocities in an isotropic material makes it possible to approximately calculate the elastic constants using the following formulae [Birks et al. 1991; ASTM E494-95 1995]:

$$\sigma = \frac{1 - 2(V_s/V_l)^2}{2(1 - (V_s/V_l)^2)}, \quad E = \frac{\rho V_s(3V_l^2 - 4V_s^2)}{V_l^2 - V_s^2}, \quad G = \rho V_s^2, \quad K = \rho\left(V_l^2 - \tfrac{4}{3}V_s^2\right),$$

where σ is the Poisson's ratio and E, G, and K, are the Young's, shear, and bulk moduli, respectively. V_l and V_s are the ultrasonic longitudinal and transverse wave velocities, respectively. The above equations are usually used for material characterization by ultrasonic wave velocity measurements as well as for studying the effect of metallurgical parameters such as precipitation on metals in isotropic materials [ASTM E92-82 2003].

3. Results and discussion

Ultrasonic characteristics and elastic properties of the specimens were found by using the equations above. The corresponding values for the six AISI 52100 alloy steel specimens are listed in Tables 2 and 3.

Figure 4 shows the effect of tempering temperature on hardness. During tempering, martensite decomposes into a mixture of ferrite and cementite leading to a decrease in volume as tempering temperature increases. Percentage of martensitic structure as well as hardness decrease with an increase in tempering temperature. All tempered specimens had hardness values greater than the untreated specimen.

Sample	Density (kg/m³)	Longitudinal velocity (m/s)	Transverse velocity (m/s)	Attenuation (dB/mm)	Z (kg/m³·s)
1	7322	5845.2	3182.5	0.124	42798.98
2	7289	6066.9	3225.5	0.155	44221.64
3	7280	5910.4	3210.4	0.133	43024.94
4	7374	5918.3	3222.2	0.054	43644.04
5	7356	5931.4	3224.2	0.182	43634.42
6	7346	5986.6	3281.4	0.206	43978.67

Table 2. Ultrasonic characteristics of the specimens. Sample 6 was untreated. (*Z* is the acoustic impedance.)

Sample	Vickers hardness	Poisson's ratio	Young's modulus (GPa)	Shear modulus (GPa)	Bulk modulus (GPa)
1	849.1	0.289	191.23	74.16	151.29
2	715.6	0.303	197.61	75.83	167.18
3	692.1	0.291	193.67	75.02	154.26
4	631.4	0.289	197.43	76.56	156.21
5	579.4	0.290	197.35	76.47	156.84
6	275.3	0.285	203.33	79.10	157.81

Table 3. Elastic properties of specimens. Sample 6 was untreated.

The first two parts of Figure 5 show the effect of tempering temperature on longitudinal and transverse wave velocities. Transverse and longitudinal wave velocities increase continuously as the tempering temperature increases, except in the tempered martensite embrittlement (TME) range (260–370 °C or 500–700 °F). The TME range is also known as 500 °F embrittlement or one-step temper embrittlement [Verhoeven 2005]. Both transverse and longitudinal wave velocities show sudden increases in the TME range.

sample number	tempering temperature
1	200°C
2	250°C
3	300°C
4	350°C
5	400°C
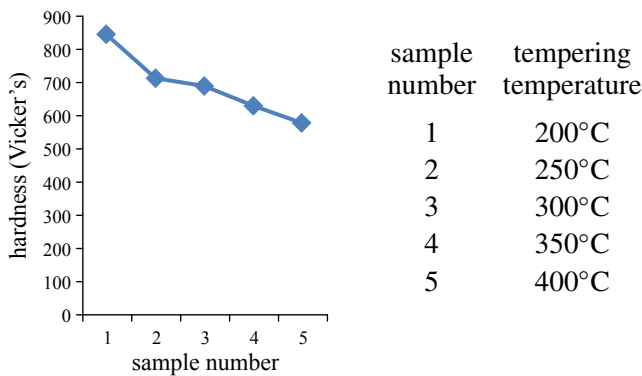
Figure 4. Effect of tempering temperature on hardness.
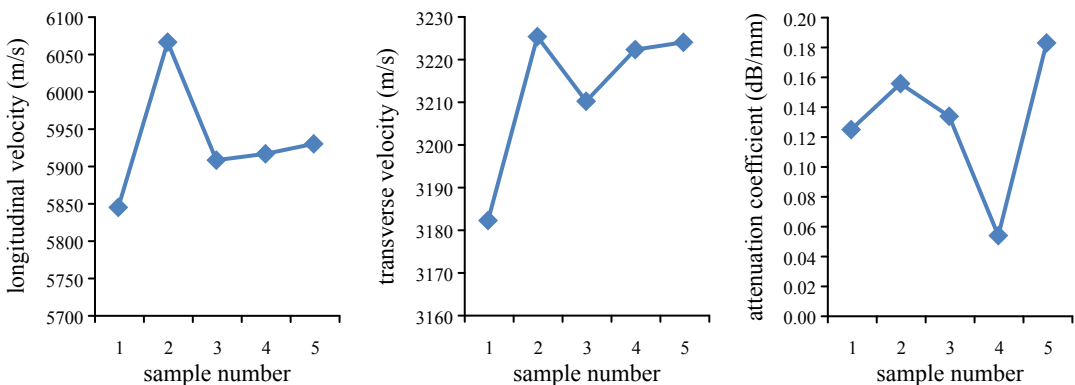
Figure 5. Effect of tempering temperature on ultrasonic properties. From left to right, longitudinal wave velocity, transverse wave velocity, and attenuation coefficient.
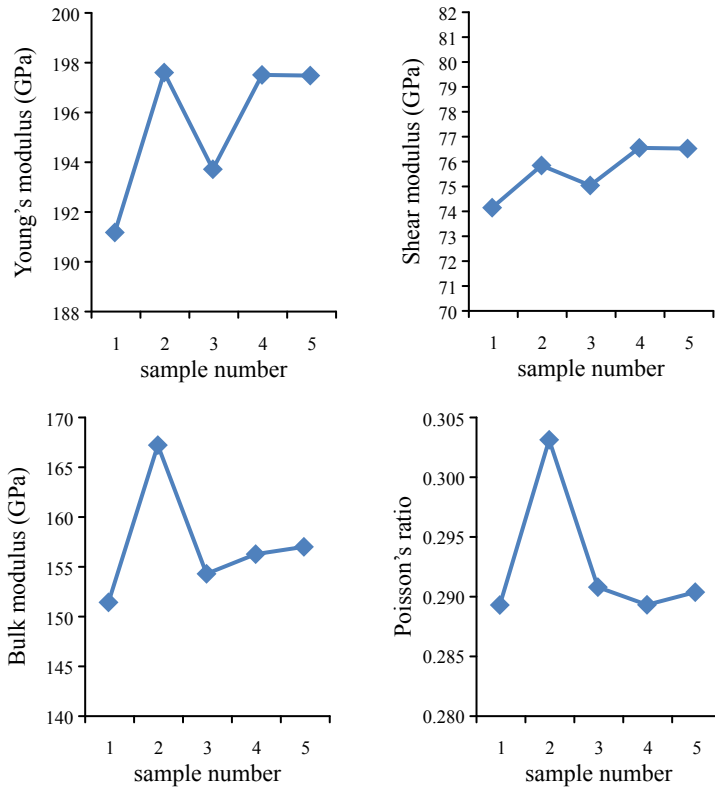
Figure 6. Effect of tempering temperature on elastic constants. Top: Young's modulus and shear modulus. Bottom: bulk modulus and Poisson's ratio.

Figure 5, right, shows the effect of tempering temperature on the attenuation of ultrasonic waves. It is observed that the attenuation data is quite scattered and no particular relationship can be established between the wave attenuation and tempering temperature. The attenuation coefficient is controlled by two phenomena: scattering and absorption. The existence of transition carbide, low-carbon martensite, and dislocations in the range 200–250 °C increase the scattering attenuation. In the temperature range 250–350 °C, cementite and ferrite appear and replace the aforementioned components which leads to a drop in the attenuation factor. Beyond the tempering temperature of 350 °C, the grain size increases and consequently the scattering attenuation also increases.

Figure 6 shows the effect of tempering temperature on elastic properties. The changes in the values of the Young's, shear, and bulk moduli with changes in tempering temperature are similar. They all increase with increase of the tempering temperature. Figure 6d shows the effect of tempering temperature on the Poisson's ratio. In Figure 6d, the Poisson's ratios of the tempered samples are approximately constant except in the TME range (also see Table 3).

Comparing Figures 4, 5 and 6, it can be observed that by increasing the tempering temperature, hardness decreases while elastic constants (the Young's, shear, and bulk moduli) and ultrasonic properties (longitudinal and transverse wave velocities and attenuation) increase except in the TME range, where the process is reversed. The major effect of tempering is elimination of many of the smaller laths and
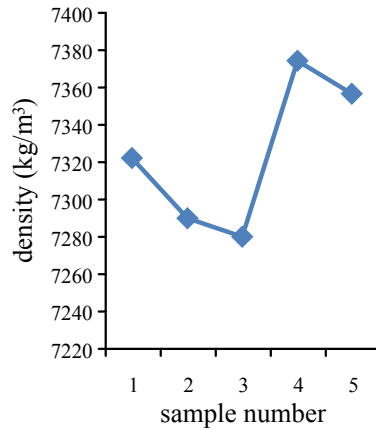
Figure 7. Effect of tempering temperature on density.

production of coarse, spherical cementite particles at the prior austenite grain boundaries and within packets [Krauss 1989]. During tempering, decomposition of martensite into ferrite and carbides creates both structural and dimensional changes. In the temperature range of 100–250 °C, transition carbide, epsilon carbide, and eta carbide form and the carbon content of matrix martensite decreases to about 0.25%. Consequently, the bct structure in the equilibrium structure of martensite tends to transform into a bcc structure.

Figure 7 shows the effect of tempering temperature on density. Because of the existence of martensite and dislocations in this range, the volume of specimens is increased and density is decreased. By increasing the tempering temperature, dislocation density is effectively lowered and any residual dislocations within laths rearrange themselves into low-angle boundaries within the equiaxed grains.

In the temperature range of 200–300 °C, the retained austenite transforms to ferrite and cementite. In plain carbon steels and low-alloy steels, the retained austenite transforms to bainite with an increase in volume and a decrease in density. By eliminating the retained austenite, volume decreases and density increases.

In the temperature range of 250–350 °C, the transition carbide and low-carbon martensite are replaced by cementite and ferrite [Krauss 1989].

It has already been shown that the formation of ferrite, coarse pearlite, fine pearlite, cementite, bainite, and martensite phases in some steels (medium-carbon low alloy steels [Krauss 1989], AISI 4140 [Palanichamy et al. 1995], 38 NiCr Mo4 [ASTM E494-95 1995]) is continuously associated with an increase in hardness and a decrease in ultrasonic velocity.

Our results show that there are more than one microstructure in each specimen and by increasing the tempering temperature, the percentage of cementite and ferrite in microstructure increases and leads to improvement in the ultrasonic and elastic properties of the tempered specimens. The abnormal behavior of mechanical and ultrasonic properties in the TME range can be due to different factors, some of the most important of which are [ASM 1 1990]:

(1) Initiation of cementite precipitation at TME range. Low-temperature precipitated cementite is long and plate-like and is formed at grain boundaries.

(2) Impurity elements, chiefly phosphorus, which segregate to the prior austenite grain boundaries during austenitization.

(3) The presence of small amounts of retained austenite films between the laths of martensite.

(4) Decomposition of interlath-retained austenite to interlath cementite.

Conclusions

The purpose of this research was to derive relationships between tempering temperature and the ultrasonic and elastic properties of AISI 52100 alloy steel. The elastic properties were obtained by measuring the longitudinal and transverse wave velocities of heat-treated AISI 52100 samples. The results show that the ultrasonic technique can identify any slight changes in the microstructure and mechanical properties of AISI 52100 alloy steel. Accordingly, one can choose the most suitable tempering temperature range for manufacturing components with specific elastic properties.

The abnormal behavior of the mechanical and ultrasonic properties of AISI 52100 alloy steel at the tempered martensite embrittlement range was noticed and possible reasons for this kind of behavior were suggested. As a future task, this behavior could be studied more precisely.

References

[ASM 1 1990] *Properties and selection: irons, steels, and high performance alloys*, ASM Handbook **1**, ASM International, Materials Park, OH, 1990.

[ASM 4 1991] *Heat treating*, ASM Handbook **4**, ASM International, Materials Park, OH, 1991.

[ASTM E494-95 1995] "Standard practice for measuring ultrasonic velocity in materials", ASTM Standard E494-95, ASTM International, 1995, Available at http://tinyurl.com/astm-E494-95. Replaced by E494 - 10.

[ASTM E664-93 2000] "Standard practice for the measurement of the apparent attenuation of longitudinal ultrasonic waves by immersion method", ASTM Standard E664-93, ASTM International, 2000, Available at http://tinyurl.com/agbehwf. Superseded by E664M - 10.

[ASTM E92-82 2003] "Standard test method for Vickers hardness of metallic materials", ASTM Standard E92-82, ASTM International, 2003, Available at http://tinyurl.com/astm-E92-82. Replaced by E384 - 11e1, Standard test for Knoop and Vickers hardness of materials.

[Badidi Bouda et al. 2000] A. Badidi Bouda, A. Benchaala, and K. Alem, "Ultrasonic characterization of materials hardness", *Ultrasonics* **38**:1–8 (2000), 224–227.

[Birks et al. 1991] A. S. Birks, R. E. Green, Jr., and P. McIntire (editors), *Ultrasonic testing*, 2nd ed., Nondestructive Testing Handbook **7**, American Society for Nondestructive Testing, Columbus, OH, 1991.

[Brandes and Brook 1998] E. A. Brandes and G. B. Brook (editors), *Smithells metals reference book*, 7th ed., Butterworth-Heinemann, Oxford, 1998.

[Carreon et al. 2009] H. Carreon, A. Ruiz, A. Medina, G. Barrera, and J. Zarate, "Characterization of the alumina-zirconia ceramic system by ultrasonic velocity measurements", *Mater. Charact.* **60**:8 (2009), 875–881.

[Hakan Gür and Çam 2007] C. Hakan Gür and I. Çam, "Comparison of magnetic Barkhausen noise and ultrasonic velocity measurements for microstructure evaluation of SAE 1040 and SAE 4140 steels", *Mater. Charact.* **58**:5 (2007), 447–454.

[Hakan Gür and Orkun Tuncer 2005] C. Hakan Gür and B. Orkun Tuncer, "Characterization of microstructural phases of steels by sound velocity measurement", *Mater. Charact.* **55**:2 (2005), 160–166.

[Hsu et al. 2004] C.-H. Hsu, H.-Y. Teng, and Y.-J. Chen, "Relationship between ultrasonic characteristics and mechanical properties of tempered martensitic stainless steel", *J. Mater. Eng. Perform.* **13**:5 (2004), 593–598.

[Hull et al. 1985] D. R. Hull, H. E. Kautz, and A. Vary, "Ultrasonic velocity measurement using phase-slope and cross-correlation methods", *Mater. Eval.* **43** (1985), 1455–1460. NASA Technical Memo 83794.

[Krauss 1989] G. Krauss, *Steels: heat treatment and processing principles*, ASM International, Materials Park, OH, 1989.

[Kumar et al. 2003] A. Kumar, T. Jayakumar, B. Raj, and K. K. Ray, "Characterization of solutionizing behavior in VT14 titanium alloy using ultrasonic velocity and attenuation measurements", *Mater. Sci. Eng. A* **360**:1–2 (2003), 58–64.

[Murthy et al. 2008] G. V. S. Murthy, S. Ghosh, M. Das, G. Das, and R. N. Ghosh, "Correlation between ultrasonic velocity and indentation-based mechanical properties with microstructure in Nimonic 263", *Mater. Sci. Eng. A* **488**:1–2 (2008), 398–405.

[Murthy et al. 2009] G. V. S. Murthy, G. Sridhar, A. Kumar, and T. Jayakum, "Characterization of intermetallic precipitates in a Nimonic alloy by ultrasonic velocity measurements", *Mater. Charact.* **60**:3 (2009), 234–239.

[Palanichamy et al. 1995] P. Palanichamy, A. Joseph, T. Jayakumar, and B. Raj, "Ultrasonic velocity measurements for estimation of grain size in austenitic stainless steel", *NDT & E Int.* **28**:3 (1995), 179–185.

[Papadakis 1970] E. P. Papadakis, "Ultrasonic attenuation and velocity in SAE 52100 steel quenched from various temperatures", *Metall. Trans.* **1**:4 (1970), 1053–1057.

[Ringer et al. 1997] S. P. Ringer, T. Sakurai, and I. J. Polmear, "Origins of hardening in aged Al-Cu-Mg-(Ag) alloys", *Acta Mater.* **45**:9 (1997), 3731–3744.

[Rokhlin and Matikas 1996] S. I. Rokhlin and T. E. Matikas, "Ultrasonic characterization of surfaces and interphases", *Mater. Res. Soc. Bull.* **21**:10 (1996), 22–29.

[Rosen et al. 1985] M. Rosen, L. Ives, S. Ridder, F. Biancaniello, and R. Mehrabian, "Correlation between ultrasonic and hardness measurements in aged aluminum alloy 2024", *Mater. Sci. Eng.* **74**:1 (1985), 1–10.

[dos Santos Mesquita and Halldórsdóttir 2005] M. dos Santos Mesquita and S. Halldórsdóttir, "Data analysis: filtering, cross-correlation, coherence and applications to geophysical data using Matlab", 2005, Available at http://www.geofys.uu.se/cj/geofysdatabeh/notes/report_with_matlab.pdf.

[Vasudevan and Palanichamy 2002] M. Vasudevan and P. Palanichamy, "Characterization of microstructural changes during annealing of cold worked austenitic stainless steel using ultrasonic velocity measurements and correlation with mechanical properties", *J. Mater. Eng. Perform.* **11**:2 (2002), 169–176.

[Verhoeven 2005] J. D. Verhoeven, *Metallurgy of steel for bladesmiths & others who heat treat and forge steel*, Iowa State University, Ames, IA, 2005.

MOHAMMAD HAMIDNIA: m_hamidnia74@yahoo.com
NDE Lab., Faculty of Mechanical Engineering, K. N. Toosi University of Technology, Pardis St., MollaSadra Ave., Vanak Sq., Tehran, Iran

FARHANG HONARVAR: honarvar@mie.utoronto.ca
NDE Lab, Faculty of Mechanical Engineering, K. N. Toosi University of Technology, Pardis St., MollaSadra Ave., Vanak Sq., Tehran, Iran

msp

BOUNDARY INTEGRAL EQUATION FOR
NOTCH PROBLEMS IN AN ELASTIC HALF-PLANE
BASED ON GREEN'S FUNCTION METHOD

Y. Z. CHEN

This paper studies a boundary integral equation (BIE) for notch problems in an elastic half-plane based on Green's function method. The boundary along the half-plane is traction-free. A fundamental solution is suggested, which is composed of a principal part and a complementary part. The process for evaluating the complementary part from the principal part is similar to the Green's function method for Laplace's equation. After using the Somigliana identity or Betti's reciprocal theorem between the field of the fundamental solution and the physical field, the displacements at the domain point are obtained. Letting the domain point approach the boundary point and using the generalized Sokhotski–Plemelj formula, a BIE of the notch problem for a traction-free half-plane boundary is obtained. The accuracy of the suggested technique is examined. Computed results for elliptic notches and a square notch with rounded corner are presented in the paper.

1. Introduction

It is well known that except for in some simple cases the solution for elastic problems cannot be obtained in a closed form. Therefore, researchers research the numerical solutions of elastic problems. In fact, some boundary integral equations (BIEs) in plane elasticity were suggested in [Muskhelishvili 1953]. However, even though some integral equations could then be formulated, it was difficult to solve them using simpler computation instruments, such as by hand. The situation changed significantly after the computer became available. Researchers developed two important computation techniques in elasticity: the finite element method and the boundary integral equation method.

Boundary integral equations (BIEs) in elasticity were proposed by several pioneer researchers [Rizzo 1967; Cruse 1969; Jaswon and Symm 1977]. The history of the development of BIEs was summarized recently in [Cheng and Cheng 2005]. In the direct formulation of BIEs, the dual formulations are significant. In the first version of the BIE, the boundary displacements are related to the boundary displacement and traction through some integral operators. In the second version of the BIE, the boundary tractions are related to the boundary displacement and traction through some integral operators. Generally, the singularity of the kernels of the second version is higher by one order than its counterpart in the first version [Brebbia et al. 1984; Hong and Chen 1988].

The use of complex integral equations has some proven advantages for solving two-dimensional potential and elastic problems [Hromadka and Lai 1987; Linkov and Mogilevskaya 1994; Mogilevskaya and Linkov 1998; Linkov 2002; Chen and Chen 2004]. Recently, a formulation of indirect BIEs in plane elasticity using single or double layer potentials and complex variables was suggested [Chen et al. 2010].

Continuous and discontinuous properties of the displacements and tractions of a moving point across the boundary play an important role in the formulation of BIEs. These properties can be investigated exactly through the use of complex variables.

The stress concentration problem around a notch is an important topic in the strength of materials. Many problems in this field were solved in [Savin 1961]. However, the scale of study and accuracy of computation in early year (1951) was not expected, simply because no modern computer was available at that time. A special boundary element formulation for multiple circular hole problems in an infinite plate was suggested in [Chen 1985]. The multiple circular hole problem for an elastic half-plane and the periodic group circular hole problems were solved by using the series expansion variational method [Chen 1994; Chen and Lin 2007].

An analytical solution has been given for the class of problems of an elastic half plane with a circular cavity loaded on the cavity boundary. For the case of a uniform radial stress at the cavity boundary the solution can be given in a closed form [Verruijt 1998]. A complex boundary integral equation method has been presented for the problem of an infinite isotropic elastic plane containing multiple circular holes. The suggested method was reduced to solve a hypersingular integral equation [Wang et al. 2003]. It is seen from the derivation that the method is effective for solving the problem of multiple circular holes. Similarly, a semianalytical method was presented for solving the problem of an isotropic elastic half-plane containing multiple circular holes of arbitrary sizes [Dejoie et al. 2006]. A boundary element formulation for cracked anisotropic elastic half-planes was proposed. In the formulation, the complex potentials were composed of two parts, a principal part and a complementary part [Chen and Cheung 1990; Pan et al. 1997].

This paper studies a boundary integral equation for notch problems in an elastic half-plane based on Green's function method. The boundary along the half-plane is traction-free. The first step in the derivation is to formulate a fundamental solution. In the usual BIE, the fundamental solution corresponds to a particular solution for concentrated forces applied at a point. Thus, this solution does not satisfy the traction-free condition along the boundary of the half-plane.

In the present study, the pair of complex potentials for concentrated forces in the usual formulation is called the principal part, and is denoted by $\phi_p(z)$ and $\psi_p(z)$. The complex potentials for the traction-free half-plane boundary can be investigated in the form of $\phi(z) = \phi_p(z) + \phi_c(z)$ and $\psi(z) = \psi_p(z) + \psi_c(z)$. The purpose of introducing the part composed of $\phi_c(z)$ and $\psi_c(z)$ is to eliminate the tractions along the half-plane boundary caused by $\phi_p(z)$ and $\psi_p(z)$. The pair of complex potentials $\phi_c(z)$ and $\psi_c(z)$ is called the complimentary part in the whole complex potentials. The pair $\phi_c(z)$ and $\psi_c(z)$ can be derived from the pair $\phi_p(z)$ and $\psi_p(z)$ in an exact way, which is similar to the Green's function method in the solution of Laplace's equation. After using the Somigliana identity or Betti's reciprocal theorem between the field of the fundamental solution and the physical field, the displacements at the domain point are obtained. Letting the domain point approach the boundary point and using the generalized Sokhotski–Plemelj formula, a BIE for notch problems for a traction-free half-plane boundary is obtained. Since the fundamental solution satisfies the traction-free condition along the boundary of the half-plane in advance, there is no need to set the equation for points along the half-plane boundary. The formulation is a general one, without regard to the configuration of the notch. A discretization procedure is suggested, and the accuracy of the suggested method is examined. Computed results for elliptic notches and a square notch with rounded corners are presented in the paper.

2. Formulation of BIE for notch problems in an elastic half-plane based on Green's function method

Background on the complex variable method of plane elasticity. The complex variable function method plays an important role in plane elasticity. The fundamentals of this method are introduced here. In the method, the stresses $(\sigma_x, \sigma_y, \sigma_{xy})$, the resultant forces (X, Y) and the displacements (u, v) are expressed in terms of complex potentials $\phi(z)$ and $\psi(z)$ such that [Muskhelishvili 1953]

$$\sigma_x + \sigma_y = 4 \operatorname{Re} \phi'(z), \quad \sigma_y - \sigma_x + 2i\sigma_{xy} = 2[\bar{z}\phi''(z) + \psi'(z)], \tag{1}$$

$$f = -Y + iX = \phi(z) + z\overline{\phi'(z)} + \overline{\psi(z)}, \tag{2}$$

$$2G(u + iv) = \kappa\phi(z) - z\overline{\phi'(z)} - \overline{\psi(z)}, \tag{3}$$

where a bar over a function denotes the conjugated value for the function, G is the shear modulus of elasticity, $\kappa = (3 - \nu)/(1 + \nu)$ in the plane stress problem, $\kappa = 3 - 4\nu$ in the plane strain problem, and ν is the Poisson's ratio. Sometimes, the displacements u and v are denoted by u_1 and u_2, the stresses σ_x, σ_y, and σ_{xy} by σ_1, σ_2, and σ_{12}, and the coordinates x and y by x_1 and x_2. For the sake of the following derivations, an equation for finding a particular derivative is introduced as follows [Muskhelishvili 1953; Savruk 1981; Chen et al. 2003]:

$$\frac{d}{dz}\{f(z)\overline{g(z)}\} = f'(z)\overline{g(z)} + \frac{d\bar{z}}{dz}(f(z)\overline{g'(z)}). \tag{4}$$

In (4), $f(z)$ and $g(z)$ denote some analytic functions. The derivative in (4) is called the derivative in a specified direction (DISD).

Except for the physical quantities mentioned above, from (2) and (3) two DISDs are introduced as follows [Savruk 1981; Chen et al. 2003]:

$$J_1(z) = \frac{d}{dz}\{-Y + iX\} = \phi'(z) + \overline{\phi'(z)} + \frac{d\bar{z}}{dz}(z\overline{\phi''(z)} + \overline{\psi'(z)}) = \sigma_N + i\sigma_{NT}, \tag{5}$$

$$J_2(z) = 2G\frac{d}{dz}\{u + iv\} = \kappa\phi'(z) - \overline{\phi'(z)} - \frac{d\bar{z}}{dz}(z\overline{\phi''(z)} + \overline{\psi'(z)}) = (\kappa + 1)\phi'(z) - J_1. \tag{6}$$

It is easy to verify that σ_N and σ_{NT} in $J_1 = \sigma_N + i\sigma_{NT}$ denote the normal and shear tractions along the segment from z to $z + dz$. Secondly, the J_1 and J_2 values depend not only on the position of a point z, but also on the direction of the segment $d\bar{z}/dz$.

The physical meaning of $J_1(z) = d/dz\{-Y + iX\} = \sigma_N + i\sigma_{NT}$ and $J_2(z) = 2Gd/dz\{u + iv\}$ is introduced in Figure 1. Since $-Y + iX$ (see (2)) represents the resultant force function defined along the interval from z to $z + dz$ (Figure 1), we can prove

$$J_1(z) = \frac{d}{dz}\{-Y + iX\} = \lim_{dz \to 0} \frac{\{-Y + iX\}\big|_{\text{at } z+dz} - \{-Y + iX\}\big|_{\text{at } z}}{dz} = \sigma_N + i\sigma_{NT}. \tag{7}$$

Thus, the $J_1(z)$ value represents the traction $\sigma_N + i\sigma_{NT}$ applied along the interval from z to $z + dz$ (Figure 1).
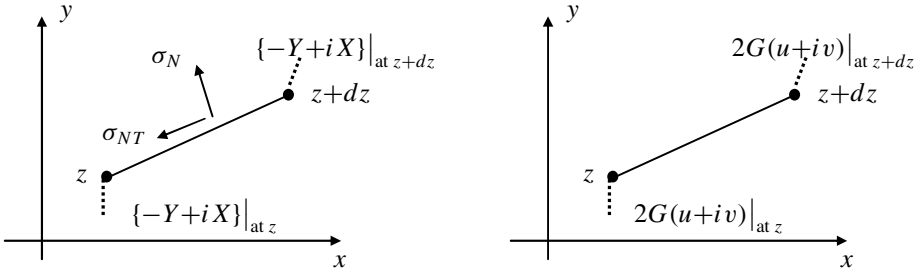
Figure 1. Left: physical meaning of the component $J_1(z)$. Right: physical meaning of the component $J_2(z)$.

Similarly, since $2G(u + iv)$ represents the displacement function defined along the interval from z to $z + dz$, we can define

$$J_2(z) = 2G\frac{d}{dz}\{u + iv\} = 2G \lim_{dz \to 0} \frac{(u + iv)\big|_{\text{at } z+dz} - \{u + iv\}\big|_{\text{at } z}}{dz}. \tag{8}$$

Thus, the $J_2(z)$ value represents the deformation state along the interval from z to $z + dz$ (Figure 1). In fact, from the value of $J_2(z)$, we can get the elongation and the rotation of the interval from z to $z + dz$.

We can find a particular usage of (5) for work done by two stress fields. Suppose there are two stress fields, which are denoted by the subscripts α and β. Thus, we can formulate the following equality:

$$(u - iv)^{(\alpha)}\big(-i J_1^{(\beta)}(z)\big)\, dz = (u - iv)^{(\alpha)}\, d(X + iY)^{(\beta)} \tag{9}$$

and

$$\text{Re}\big\{-i(u - iv)^{(\alpha)} J_1^{(\beta)}(z)\, dz\big\} = u^{(\alpha)}\, dX^{(\beta)} + v^{(\alpha)}\, dY^{(\beta)}. \tag{10}$$

Equation (10) represents the work done along the interval from z to $z + dz$ for displacement for the α field to the traction for the β field. Equation (10) is particularly important in the case of using Betti's reciprocal theorem between two stress fields.

Formulation of BIE for notch problems in an elastic half-plane. The formulation of a fundamental solution plays an important role in the formulation of BIE for notch problems in an elastic half-plane. The fundamental solution satisfies the following conditions (see Figure 2, top left):

(1) There is a singular point $z = \tau$ in the upper half-plane with a concentrated force (P_x, P_y).

(2) The solution satisfies the traction-free condition along the boundary of the half-plane.

The original problem depicted in Figure 2, top left, can be considered as a superposition of the two stress fields defined by the diagrams in the right column of Figure 2. Therefore, it is natural to find the complex potentials for the fundamental solution in the form

$$\phi(z) = \phi_p(z) + \phi_c(z), \quad \psi(z) = \psi_p(z) + \psi_c(z). \tag{11}$$

In (11), the pair $\phi_p(z)$ and $\psi_p(z)$ is called the principal part of the whole complex potentials, which is related to the fundamental field caused by concentrated force at the point $z = \tau$ of the upper half-plane

$\phi(z) = \phi_p(z) + \phi_c(z)$
$\psi(z) = \psi_p(z) + \psi_c(z)$

$\phi_p(z)$, $\psi_p(z)$ singular in the upper plane with singular point τ in the upper plane
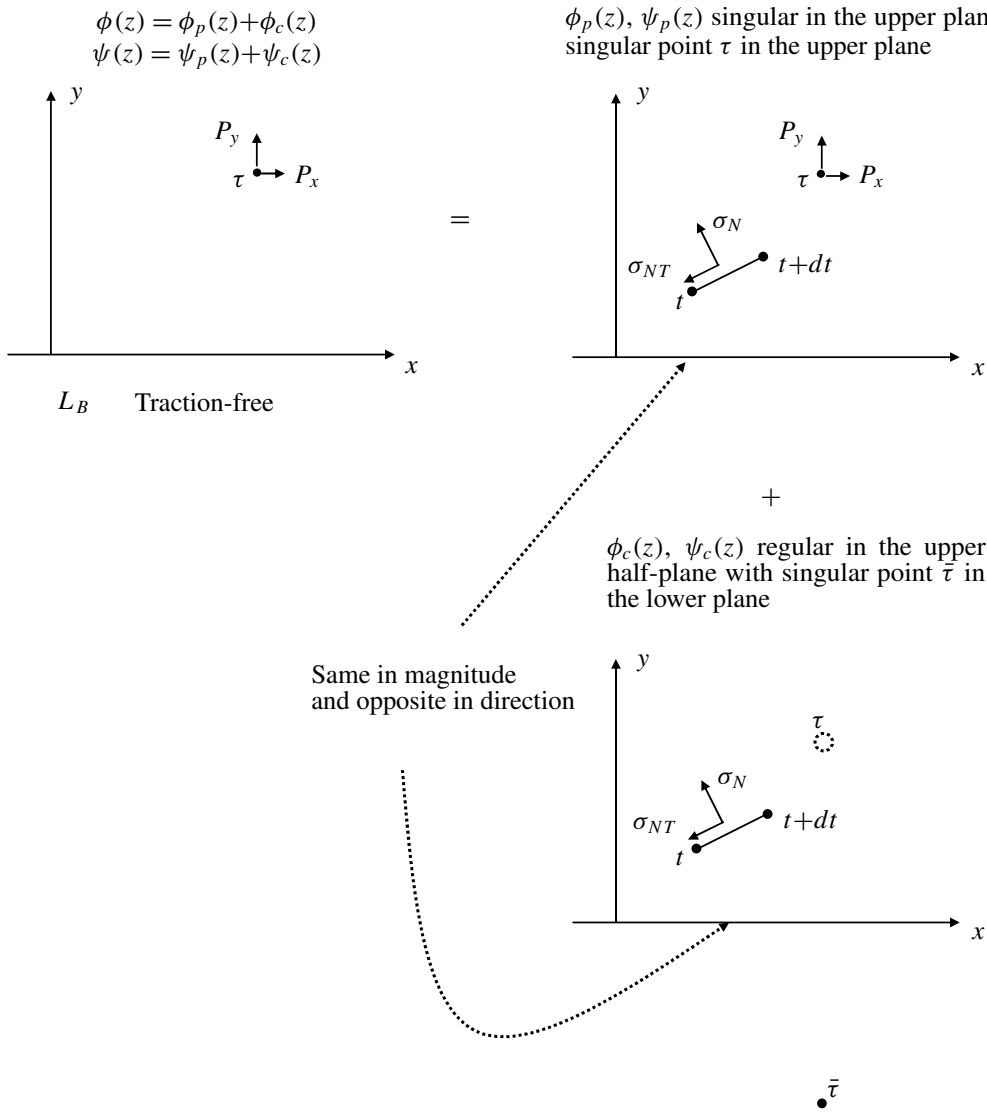
Figure 2. Construction of the fundamental solution.

(Figure 2, top right). The relevant complex potentials are as follows [Muskhelishvili 1953; Savruk 1981; Chen et al. 2003]:

$$\phi_p(z) = F \ln(z - \tau), \quad \phi'_p(z) = \frac{F}{z - \tau}, \quad \phi''_p(z) = -\frac{F}{(z - \tau)^2}, \tag{12}$$

$$\psi_p(z) = -\kappa \bar{F} \ln(z - \tau) - \frac{F \bar{\tau}}{z - \tau}, \quad \psi'_p(z) = -\frac{\kappa \bar{F}}{z - \tau} + \frac{F \bar{\tau}}{(z - \tau)^2}, \tag{13}$$

where

$$F = -\frac{P_x + i P_y}{2\pi(\kappa + 1)}. \tag{14}$$

In (14), $P_x + i P_y$ is the concentrated force applied at the point $z = \tau$ in Figure 2, top right.

The complex potentials shown by (12) and (13) are defined in the full infinite plane. From (5), (12), and (13) we can evaluate the relevant boundary traction at the point t with the increase dt as follows:

$$(\sigma_N + i\sigma_{NT})_{*p} = \frac{F}{t - \tau} + \frac{\bar{F}}{\bar{t} - \bar{\tau}} + \frac{d\bar{t}}{dt}\left(-\frac{\kappa F}{\bar{t} - \bar{\tau}} - \frac{\bar{F}(t - \tau)}{(\bar{t} - \bar{\tau})^2}\right). \tag{15}$$

Similarly, from (3), (12), and (13), we can evaluate the relevant displacement at point t as follows:

$$2G(u - iv)_{*p} = 2\kappa \bar{F} \ln|t - \tau| - F\frac{\bar{t} - \bar{\tau}}{t - \tau}. \tag{16}$$

In (15) and (16), the subscript $*p$ denotes that the arguments are derived from the principal part of the complex potentials. It is emphasized here that the pair $\phi_p(z)$ and $\psi_p(z)$ does not satisfy the traction-free condition along the boundary of the half-plane L_b (Figure 2).

In (11), the pair $\phi_c(z)$ and $\psi_c(z)$ is called the complementary part, or the regular part, of the whole complex potentials (Figure 2, bottom). The role of the pair $\phi_c(z)$ and $\psi_c(z)$ is to eliminate the boundary traction along the boundary of the half-plane caused by the pair $\phi_p(z)$ and $\psi_p(z)$ (Figure 2, right column).

Since z has a real value along the boundary of the half-plane, we have

$$z = \bar{z} \quad (\text{for } z \in L_B). \tag{17}$$

In (17), L_B denotes the boundary of the half-plane (Figure 2).

It is seen that the traction-free condition along the boundary of the half-plane (or $\sigma_N + i\sigma_{NT} = 0$) along $z \in L_B$ is equivalent to the vanishing resultant force function along the same boundary (Figure 2). By using (2) and (17), this condition can be expressed as

$$\overline{\phi(z)} + z\phi'(z) + \psi(z) = 0 \quad (\text{for } z \in L_B), \tag{18}$$

or, in an alternative form,

$$\overline{\phi_p(z)} + \overline{\phi_c(z)} + z\phi'_p(z) + \psi_p(z) + z\phi'_c(z) + \psi_c(z) = 0 \quad (\text{for } z \in L_B). \tag{19}$$

Similarly to the image method in electrostatics, substituting (12) and (13) into (19) yields

$$\phi_c(z) = \kappa F \ln(z - \bar{\tau}) + \frac{\bar{F}(\tau - \bar{\tau})}{z - \bar{\tau}}, \quad \phi'_c(z) = \frac{\kappa F}{z - \bar{\tau}} - \frac{\bar{F}(\tau - \bar{\tau})}{(z - \bar{\tau})^2},$$

$$\phi''_c(z) = -\frac{\kappa F}{(z - \bar{\tau})^2} + \frac{2\bar{F}(\tau - \bar{\tau})}{(z - \bar{\tau})^3}, \tag{20}$$

and

$$\psi_c(z) = -\bar{F}\ln(z - \bar{\tau}) - \frac{\kappa F\bar{\tau}}{z - \bar{\tau}} + \frac{\bar{F}(\tau - \bar{\tau})}{z - \bar{\tau}} + \frac{\bar{F}(\tau - \bar{\tau})\bar{\tau}}{(z - \bar{\tau})^2},$$

$$\psi'_c(z) = -\frac{\bar{F}}{z - \bar{\tau}} + \frac{\kappa F\bar{\tau}}{(z - \bar{\tau})^2} - \frac{\bar{F}(\tau - \bar{\tau})}{(z - \bar{\tau})^2} - \frac{2\bar{F}(\tau - \bar{\tau})\bar{\tau}}{(z - \bar{\tau})^3}. \tag{21}$$

The relevant derivation for obtaining the pair $\phi_c(z)$ and $\psi_c(z)$ form the pair $\phi_p(z)$ and $\psi_p(z)$ can be found in [Savruk 1981; Chen and Cheung 1990; Chen et al. 2003; 2009]. The complex potentials $\phi_c(z)$

and $\psi_c(z)$ have a singular point $z = \bar{\tau}$ at the lower half-plane, or they represent a regular function in the upper half-plane (Figure 2, bottom right).

From (5), (20), and (21), we can evaluate the relevant traction at the point t with increasing dt (dt a complex value) as follows (Figure 2):

$$(\sigma_N + i\sigma_{NT})_{*c} = \bar{F}c_1(t, \tau) + Fc_2(t, \tau), \tag{22}$$

where

$$c_1(t, \tau) = -\frac{\tau - \bar{\tau}}{(t - \bar{\tau})^2} + \frac{\kappa}{\bar{t} - \tau} - \frac{d\bar{t}}{dt}\frac{\kappa(t - \tau)}{(\bar{t} - \tau)^2},$$

$$c_2(t, \tau) = \frac{\kappa}{t - \bar{\tau}} + \frac{\tau - \bar{\tau}}{(\bar{t} - \tau)^2} + \frac{d\bar{t}}{dt}\left(-\frac{1}{\bar{t} - \tau} + \frac{\tau - \bar{\tau}}{(\bar{t} - \tau)^2} - \frac{2(t - \tau)(\tau - \bar{\tau})}{(\bar{t} - \tau)^3}\right). \tag{23}$$

Similarly, from (3), (20), and (21), we have

$$2G(u - iv)_{*c} = \bar{F}d_1(t, \tau) + Fd_2(t, \tau), \tag{24}$$

where

$$d_1(t, \tau) = 2\kappa^2 \ln|t - \bar{\tau}| - (\kappa^2 - 1) \ln(t - \bar{\tau}) - \frac{\tau - \bar{\tau}}{t - \bar{\tau}} + \frac{(\bar{t} - \bar{\tau})(\tau - \bar{\tau})}{(t - \bar{\tau})^2},$$

$$d_2(t, \tau) = -\frac{\kappa(\tau - \bar{\tau})}{\bar{t} - \tau} - \frac{\kappa(\bar{t} - \bar{\tau})}{t - \bar{\tau}}. \tag{25}$$

In (22) and (24), the subscript $*c$ denotes that the arguments are derived from the complementary part of the complex potentials. In the functions $c_1(t, \tau)$, $c_2(t, \tau)$, $d_1(t, \tau)$, and $d_2(t, \tau)$, the variable t is always located in the upper half-plane and the variable $\bar{\tau}$ in lower half-plane. Thus, for example, the term $1/(t - \bar{\tau})$ is regular.

After using Betti's reciprocal theorem, or the Somigliana identity, between the field of the fundamental solution (or the stress field shown by Figure 2, top left) and the physical field (or the stress field shown by Figure 3), we have

$$P_x u(\tau) + P_y v(\tau) + \mathrm{Re}\left(\int_\Gamma (u - iv)(dX + i\, dY)_*\right) = \mathrm{Re}\left(\int_\Gamma (u - iv)_*(dX + i\, dY)\right) \quad (\tau \in S^-), \tag{26}$$

where

$$(dX + i\, dY)_* = (dX + i\, dY)_{*p} + (dX + i\, dY)_{*c}, \tag{27}$$

$$(u - iv)_* = (u - iv)_{*p} + (u - iv)_{*c}. \tag{28}$$

In (26), Γ denotes the notch contour (Figure 3).

In addition, we can define

$$(\sigma_N + i\sigma_{NT})_* = (\sigma_N + i\sigma_{NT})_{*p} + (\sigma_N + i\sigma_{NT})_{*c}. \tag{29}$$

In (26), the left-hand term represents the work done by traction in the field of the fundamental solution (Figure 2, top left) to the displacement of the physical field (Figure 3). In addition, the right-hand term represents the work done by traction in the physical field to the displacement in the field of the fundamental solution.
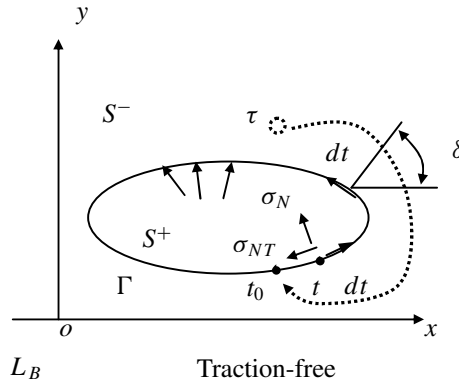
Figure 3. A physical stress field for a notch in the half-plane.

In the derivation, $dX + i\,dY$ denotes the force applied on the segment dt (Figures 2 and 3). From Figure 3 or (5), we can find

$$dX + i\,dY = (\sigma_N + i\sigma_{NT})e^{i(\pi+\delta)}\,ds = -(\sigma_N + i\sigma_{NT})e^{i\delta}\,ds = i(\sigma_N + i\sigma_{NT})\,dt. \tag{30}$$

By using (27)–(30) and substituting those expressions for $(\sigma_N + i\sigma_{NT})_{*p}$, $2G(u - iv)_{*p}$, $(\sigma_N + i\sigma_{NT})_{*c}$, and $2G(u - iv)_{*c}$ into (26), (26) can be written in an explicit form:

$$
\begin{aligned}
P_x u(\tau) &+ P_y v(\tau) + \operatorname{Re} \int_\Gamma i\left(\frac{F}{t-\tau} + \frac{\bar{F}}{\bar{t}-\bar{\tau}} + \frac{d\bar{t}}{dt}\left(-\frac{\kappa F}{\bar{t}-\bar{\tau}} - \frac{\bar{F}(t-\tau)}{(\bar{t}-\bar{\tau})^2}\right)\right)(u - iv)\,dt \\
&+ \operatorname{Re} \int_\Gamma i(\bar{F}c_1(t,\tau) + Fc_2(t,\tau))(u - iv)\,dt \\
&= \frac{1}{2G}\operatorname{Re} \int_\Gamma i\left(2\kappa\bar{F}\ln|t-\tau| - F\frac{\bar{t}-\bar{\tau}}{t-\tau}\right)(\sigma_N + i\sigma_{NT})\,dt \\
&+ \frac{1}{2G}\operatorname{Re} \int_\Gamma i(\bar{F}d_1(t,\tau) + Fd_2(t,\tau))(\sigma_N + i\sigma_{NT})\,dt, \qquad\qquad \tau \in S^-.
\end{aligned} \tag{31}
$$

Note that the point $\tau(\tau \in S^-)$ now is a domain point (Figure 3).

In the following analysis, one can let

$$U(t) = u(t) + iv(t), \quad Q(t) = \sigma_N(t) + i\sigma_{NT}(t) \quad (t \in \Gamma). \tag{32}$$

In (31), if we let $P_x = 1$, $P_y = 0$, and $F = -1/2\pi(\kappa + 1)$, we can find an equation with $u(\tau)$. Similarly, if we let $P_x = 0$, $P_y = 1$, and $F = -i/2\pi(\kappa + 1)$, we can find an equation with $v(\tau)$. Thus, we will find

$$
\begin{aligned}
u(\tau) + iv(\tau) &+ B_1 i \int_\Gamma \left(-\frac{1}{\bar{t}-\bar{\tau}} + \frac{d\bar{t}}{dt}\frac{(t-\tau)}{(\bar{t}-\bar{\tau})^2}\right)\overline{U(t)}\,dt + B_1 i \int_\Gamma \left(\frac{1}{\bar{t}-\bar{\tau}} - \frac{dt}{d\bar{t}}\frac{\kappa}{t-\tau}\right)U(t)\,d\bar{t} \\
&+ B_1 i \int_\Gamma C_1(t,\tau)\overline{U(t)}\,dt + B_1 i \int_\Gamma C_2(t,\tau)U(t)\,d\bar{t} \\
&= B_2 i \int_\Gamma \left(-2\kappa\ln|t-\tau|Q(t)\,dt - \frac{t-\tau}{\bar{t}-\bar{\tau}}\overline{Q(t)}\,d\bar{t}\right) \\
&+ B_2 i \int_\Gamma (D_1(t,\tau)Q(t)\,dt + D_2(t,\tau)\overline{Q(t)}\,d\bar{t}), \qquad\qquad \tau \in S^-,
\end{aligned} \tag{33}
$$

where

$$C_1(t, \tau) = -c_1(t, \tau) = \frac{\tau - \bar{\tau}}{(t - \bar{\tau})^2} - \frac{\kappa}{\bar{t} - \tau} + \frac{d\bar{t}}{dt} \frac{\kappa(t - \tau)}{(\bar{t} - \tau)^2},$$

$$C_2(t, \tau) = \overline{c_2(t, \tau)} = \frac{\kappa}{\bar{t} - \tau} - \frac{\tau - \bar{\tau}}{(t - \bar{\tau})^2} + \frac{dt}{d\bar{t}}\left(-\frac{1}{t - \bar{\tau}} - \frac{\tau - \bar{\tau}}{(t - \bar{\tau})^2} + \frac{2(\bar{t} - \bar{\tau})(\tau - \bar{\tau})}{(t - \bar{\tau})^3}\right),$$

(34)

$$D_1(t, \tau) = -d_1(t, \tau) = -2\kappa^2 \ln|t - \bar{\tau}| + (\kappa^2 - 1)\ln(t - \bar{\tau}) + \frac{\tau - \bar{\tau}}{t - \bar{\tau}} - \frac{(\bar{t} - \bar{\tau})(\tau - \bar{\tau})}{(t - \bar{\tau})^2},$$

$$D_2(t, \tau) = \overline{d_2(t, \tau)} = \frac{\kappa(\tau - \bar{\tau})}{t - \bar{\tau}} - \frac{\kappa(t - \tau)}{\bar{t} - \tau},$$

(35)

$$B_1 = \frac{1}{2\pi(\kappa + 1)}, \qquad B_2 = \frac{1}{4\pi G(\kappa + 1)}.$$

(36)

The derivation (31)–(36) is requires care, but not difficult. The relevant derivation can be found in Appendix A.

Equation (33) can be rewritten as

$$U(\tau) + B_1 i \int_\Gamma \left(-\frac{(\kappa - 1)}{t - \tau} U(t)\, dt + L_1(t, \tau)U(t)\, dt - L_2(t, \tau)\overline{U(t)}\, dt\right)$$

$$+ B_1 i \int_\Gamma C_1(t, \tau)\overline{U(t)}\, dt + B_1 i \int_\Gamma C_2(t, \tau)U(t)\, d\bar{t}$$

$$= B_2 i \int_\Gamma \left(-2\kappa \ln|t - \tau| Q(t)\, dt - \frac{t - \tau}{\bar{t} - \bar{\tau}}\overline{Q(t)}\, d\bar{t}\right)$$

$$+ B_2 i \int_\Gamma (D_1(t, \tau)Q(t)\, dt + D_2(t, \tau)\overline{Q(t)}\, d\bar{t}), \qquad \tau \in S^-, \qquad (37)$$

where

$$L_1(t, \tau) = -\frac{d}{dt}\left\{\ln\frac{t - \tau}{\bar{t} - \bar{\tau}}\right\} = -\frac{1}{t - \tau} + \frac{1}{\bar{t} - \bar{\tau}}\frac{d\bar{t}}{dt}, \quad L_2(t, \tau) = \frac{d}{dt}\left\{\frac{t - \tau}{\bar{t} - \bar{\tau}}\right\} = \frac{1}{\bar{t} - \bar{\tau}} - \frac{t - \tau}{(\bar{t} - \bar{\tau})^2}\frac{d\bar{t}}{dt}. \quad (38)$$

In (37), letting $\tau \to t_0$ ($\tau \in S^-$ is a domain point, $t_0 \in \Gamma$ a boundary point) and using the generalized Sokhotski–Plemelj formulae shown in Appendix B and the results in Appendix C, yields

$$\frac{U(t_0)}{2} + B_1 i \int_\Gamma \left(-\frac{\kappa - 1}{t - t_0} U(t)\, dt + L_1(t, t_0)U(t)\, dt - L_2(t, t_0)\overline{U(t)}\, dt\right)$$

$$+ B_1 i \int_\Gamma C_1(t, t_0)\overline{U(t)}\, dt + B_1 i \int_\Gamma C_2(t, t_0)U(t)\, d\bar{t}$$

$$= B_2 i \int_\Gamma \left(-2\kappa \ln|t - t_0| Q(t)\, dt - \frac{t - t_0}{\bar{t} - \bar{t}_0}\overline{Q(t)}\, d\bar{t}\right)$$

$$+ B_2 i \int_\Gamma (D_1(t, t_0)Q(t)\, dt + D_2(t, t_0)\overline{Q(t)}\, d\bar{t}), \qquad t_0 \in \Gamma. \qquad (39)$$

In (39), the closed elliptic contour Γ corresponds to the closed contour L shown in Figure 7, middle (see page 978).

Note that, when taking the limit process $\tau \to t_0$, (from $\tau \in S^-$ to $t_0 \in \Gamma$), the following three integrals in (37),

$$B_1 i \int_\Gamma (-\frac{(\kappa - 1)}{t - \tau} U(t) \, dt), \quad B_1 i \int_\Gamma (L_1(t, \tau) U(t) \, dt), \quad B_1 i \int_\Gamma (-L_2(t, \tau) \overline{U(t)} \, dt),$$

have jump values $-\pi B_1(\kappa - 1)U(t_0)$, $-2\pi B_1 U(t_0)$, and 0, respectively (see the Sokhotski–Plemelj formulae shown in (B.2)–(B.4) in Appendix B and the results in Appendix C). Thus, the three integrals at the left-hand side of (37) have a jump value $-\pi B_1(\kappa + 1)U(t_0) = -U(t_0)/2$ (equal to the sum of the above-mentioned three values). Finally, the term in (37) becomes $U(t_0)/2$ (note that $U(t_0)/2 = U(t_0) - (U(t_0)/2)$). In the real variable BIE, this property has been obtained previously [Brebbia et al. 1984]. However, in this paper the property is obtained in a more explicit way by using the generalized Sokhotski–Plemelj formulae shown in (B.2)–(B.4) in Appendix B and the results in Appendix C.

Equation (39) represents the complex variable boundary integral equation for a notch problem in an half-plane with a traction-free boundary of the half-plane, which is based on Betti's reciprocal theorem or the Somigliana identity. It is seen that the formulation of the fundamental solution is similar to the Green's function method in the solution of Laplace's equation.

3. Numerical examples

The numerical solution technique is introduced below. After discretization, from (39) the BIE can be written in the form

$$M_U U(t) = M_Q Q(t) \quad (t \text{ denotes the discrete points along notch boundary}), \tag{40}$$

where M_U is a matrix acting upon the displacement vector $U(t)$ and M_Q is a matrix acting upon the traction vector $Q(t)$. In the traction boundary value problem, the traction vector $Q(t)$ is given beforehand, and the displacement vector $U(t)$ can be evaluated from the algebraic equation derived from (40).

In the computation, M divisions for the elliptic contour are used. Without loss of generality, one boundary element is assumed on the interval $|s| \leq d$. In addition, the following linear shape function for $U(s)$ (or $Q(s)$) on the interval $|s| \leq d$ is used:

$$U(s) = f_1 \frac{d - s}{2d} + f_2 \frac{d + s}{2d} \quad \left(\text{with } U(s)\big|_{s=-d} = f_1 \text{ and } U(s)\big|_{s=d} = f_2\right). \tag{41}$$

Substituting the shape function for $U(s)$ into the first term of the first integral of (39), if the observation point t_0 is in the middle of the interval, the integral will be reduced to find the following singular integral:

$$\int_{-d}^{d} \frac{U(s) \, ds}{s} = f_2 - f_1. \tag{42}$$

In addition, if the point t_0 is located at the middle of the neighboring interval, the relevant integral can be evaluated by some integration rule, for example, the Simpson quadrature rule. In addition, the second and third terms of the first and second integrals in the left-hand side of (39) belong to some regular integral. Those integrals can be integrated by some quadrature rule. Thus, all the elements in the matrix M_U can be evaluated. Similarly, we can evaluate all the elements in the matrix M_Q, which is derived from the right-hand side of (39).

Three numerical examples are introduced below. The first example is devoted to examining the efficiency and accuracy of the suggested method, where a known result based on a closed-form solution is obtained in advance. The second example is devoted to evaluating the circumference stress σ_T for an ellipsoidal contour, where some computed results can be compared with those obtained by other researchers. The third example is devoted to evaluating the circumference stress σ_T for a square contour with rounded corners, where the computed results are obtained for the first time in this paper.

Example 1. The first example is devoted to examining the efficiency and accuracy of the suggested method. An elliptic notch with two axes a and b is located in the upper half-plane, and e is the distance from the lower edge of the ellipse to the boundary of the half-plane (Figure 4).

The following particular solution is suggested, and the relevant complex potentials take the form

$$\phi(z) = \phi_p(z) + \phi_c(z), \quad \psi(z) = \psi_p(z) + \psi_c(z), \tag{43}$$

where

$$\phi_p(z) = c_1 a^2 p \frac{1}{z - z_c}, \qquad\qquad \psi_p(z) = c_2 a^2 p \frac{1}{z - z_c}, \tag{44}$$

$$\phi_c(z) = \bar{c}_1 a^2 p \frac{z}{(z - \bar{z}_c)^2} - \bar{c}_2 a^2 p \frac{1}{z - \bar{z}_c}, \quad \psi_c(z) = \bar{c}_1 a^2 p \left(-\frac{1}{z - \bar{z}_c} + \frac{z(z + \bar{z}_c)}{(z - \bar{z}_c)^3} \right) - \bar{c}_2 a^2 p \frac{z}{(z - \bar{z}_c)^2}, \tag{45}$$

with $z_c = (b + e)i = hi$.

In (44) and (46), c_1 and c_2 are some constants, and p denotes some parameter which has dimensions of stress. The tractions applied on the elliptic contour can be evaluated from (1) accordingly (Figure 4). In addition, $b/a = 0.5$, $e/b = 0.2$, $c_1 = 1 + 0.5i$, and $c_2 = 2 + 1.5i$ are used in computation. It is easy to verify that the complex potentials shown by (43)–(45) satisfy the traction-free condition along the boundary of the half-plane (or L_b). The derivation for $\phi_c(z)$ and $\psi_c(z)$ from the assumed $\phi_p(z)$ and $\psi_p(z)$ can be found in [Chen and Cheung 1990; Chen et al. 2003; 2009].

As mentioned above, after discretization to (39), the BIE can be written in the form

$$M_U U(t) = M_Q Q(t) \quad (t \text{ denoting the discrete points}), \tag{46}$$

where M_U is a matrix acting upon the displacement vector $U(t)$, and M_Q is a matrix acting upon the traction vector $Q(t)$. In the computation, $M = 180$ divisions for the elliptic contour are used in computation.
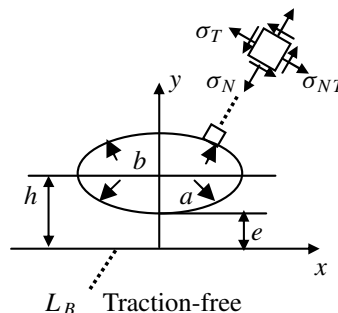

Figure 4. An elliptic notch in the half-plane under some loading along the notch contour.

In the examination, U_{ex} (Q_{ex}) denotes the displacement (traction) obtained from the exact solution with the complex potentials shown by (43)–(45). In the meantime, U_{num} (Q_{num}) denotes the displacement (traction) obtained from a numerical solution.

In the examination of the Dirichlet problem, from the input data $U_{\text{ex}}(t)$ and (46), the following algebraic equation is obtained:

$$\boldsymbol{M_Q}\,\boldsymbol{Q}(t) = \boldsymbol{M_U}\,\boldsymbol{U}_{\text{ex}}(t). \tag{47}$$

From (47), we can obtain the numerical solution for the vector $Q_{\text{num}}(t)$. In addition, the numerical solution $Q_{\text{num}}(t)$ is compared with the exact one. After computation, we find the error

$$\frac{\max|Q_{\text{num}}(t_j) - Q_{\text{ex}}(t_j)|}{\max|Q_{\text{ex}}(t_j)|} = 0.23\%$$

from many discrete points.

In the examination of the Neumann problem, from the input data $Q_{\text{ex}}(t)$ and (46), the following algebraic equation is obtained:

$$\boldsymbol{M_U}\,\boldsymbol{U}(t) = \boldsymbol{M_Q}\,\boldsymbol{Q}_{\text{ex}}(t). \tag{48}$$

From (48), we can obtain the numerical solution for the vector $U_{\text{num}}(t)$. In addition, the numerical solution $U_{\text{num}}(t)$ is compared with the exact one. After computation, we find the error

$$\frac{\max|U_{\text{num}}(t_j) - U_{\text{ex}}(t_j)|}{\max|U_{\text{ex}}(t_j)|} = 1.11\%$$

from many discrete points.

For the examination of the circumference stress σ_T component, the following technique is suggested. In the plane strain case, the strain component ϵ_T (in the T direction) can be expressed as (see Figure 4)

$$\epsilon_T = \frac{1}{E}(\sigma_T(1 - \nu^2) - \nu(1 + \nu)\sigma_N), \tag{49}$$

or

$$\sigma_T = \frac{E\epsilon_T + \nu(1 + \nu)\sigma_N}{1 - \nu^2}, \tag{50}$$

where E is the Young's modulus of elasticity. In (49), the component σ_N is from input data, and ϵ_T is the strain in the T direction which can be evaluated from the obtained displacement on the boundary. Thus, the values of σ_T at discrete points are obtainable.

In addition, the numerical solution $\sigma_{T,\text{num}}(t)$ is compared with the exact one. After computation, we find the error $\max|\sigma_{T,\text{num}}(t_j) - \sigma_{T,\text{ex}}(t_j)|/\max|\sigma_{T,\text{ex}}(t_j)| = 1.17\%$ from many discrete points. From the above-mentioned results we see that the suggested technique provides an accurate result in this example.

Example 2. The second example is devoted to evaluating the circumference stress σ_T for an elliptic contour with two axes a and b located in the upper half-plane, where e is the distance from the lower edge of the ellipse to the boundary of the half-plane (Figure 5). The remote loading is $\sigma_x = p$, and the boundary of the half-plane is traction-free. As before, $M = 180$ divisions are used in the discretization.

The normalized stress is denoted by f_T ($f_T = \sigma_T/p$). Table 1 shows the computed results for f_T ($f_T = \sigma_T/p$) under the following conditions:
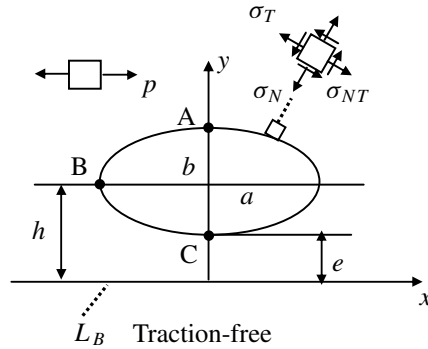
Figure 5. An elliptic notch in the half-plane under the remote tension $\sigma_x = p$.

$b/a = 0.5$										
Point	$e/b = 0.1$	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
A	2.361	2.307	2.270	2.243	2.221	2.202	2.187	2.173	2.161	2.151
B	−0.681	−0.945	−1.092	−1.183	−1.239	−1.275	−1.296	−1.308	−1.314	−1.315
C	7.178	5.154	4.358	3.890	3.573	3.342	3.163	3.021	2.905	2.808
	$e/b = 1$	2	3	4	5	6	7	8	9	10
A	2.151	2.090	2.058	2.040	2.031	2.024	2.022	2.018	2.016	2.014
B	−1.315	−1.246	−1.173	−1.123	−1.096	−1.069	−1.056	−1.045	−1.036	−1.030
C	2.808	2.327	2.170	2.100	2.070	2.049	2.038	2.029	2.024	2.020
$b/a = 1$										
Point	$e/b = 0.1$	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
A	3.626	3.510	3.435	3.381	3.338	3.303	3.275	3.250	3.229	3.211
B	−0.840	−1.035	−1.129	−1.178	−1.203	−1.214	−1.217	−1.215	−1.210	−1.203
C	9.314	6.739	5.659	5.040	4.632	4.344	4.128	3.961	3.830	3.724
	$e/b = 1$	2	3	4	5	6	7	8	9	10
A	3.211	3.110	3.064	3.040	3.031	3.024	3.026	3.022	3.019	3.017
B	−1.203	−1.124	−1.075	−1.048	−1.038	−1.026	−1.024	−1.020	−1.016	−1.014
C	3.724	3.253	3.128	3.071	3.055	3.039	3.033	3.027	3.023	3.019
$b/a = 2$										
Point	$e/b = 0.1$	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
A	6.096	5.864	5.720	5.617	5.539	5.477	5.426	5.384	5.348	5.317
B	−0.946	−1.080	−1.133	−1.154	−1.159	−1.157	−1.152	−1.144	−1.136	−1.128
C	12.893	9.415	8.027	7.264	6.778	6.443	6.200	6.013	5.870	5.756
	$e/b = 1$	2	3	4	5	6	7	8	9	10
A	5.317	5.148	5.083	5.049	5.039	5.031	5.041	5.036	5.032	5.029
B	−1.128	−1.065	−1.037	−1.022	−1.020	−1.015	−1.017	−1.015	−1.013	−1.012
C	5.756	5.263	5.136	5.074	5.066	5.048	5.046	5.039	5.035	5.031

Table 1. The normalized circumference stress f_T ($f_T = \sigma_T/p$) at some points for an elliptic notch (see Figure 5).

(1) $b/a = 0.5, 1, 2,$

(2) $e/b = 0.1, 0.2, \ldots 1$ and $1, 2, \ldots 10$, and

(3) at the boundary points A, B, and C (Figure 5),

From Table 1 we see that, for the small e/b case, the stress concentration factor can reach a larger value. For example, for the following three cases:

(1) $e/b = 0.1, b/a = 0.5,$

(2) $e/b = 0.1, b/a = 1.0$, and

(3) $e/b = 0.1, b/a = 2,$

at the point C we have f_T values 7.178, 9.314, and 12.893, respectively.

For purposes of comparison, the computed results for f_T under the condition $b/a = 1$ and $e/b = 0.185, 0.337, 0.543, 0.811, 1.151, 1.577, 2.107,$ and 2.762 are plotted in Table 2. It is found from the tabulated results that the computed results coincide with those obtained in [Dejoie et al. 2006].

Example 3. The third example is devoted to evaluating the circumference stress σ_T for a square contour with rounded corners. The width of the square contour is $2a$, and the radius of the rounded corners is $0.5a$ (Figure 6). Under the conditions:

(1) $e/a = 0.1, 0.2, \ldots, 1$, and $1, 2, \ldots, 10$ and

(2) at the boundary points A, B, C, D, and E,

the computed results for f_T ($f_T = \sigma_T/p$) are plotted in Table 3.

From Table 3 we see that, for the small e/b case, the stress concentration factor can reach a larger value. For example, for the two cases $e/a = 0.1$ and 0.2, at the point E we have f_T values 5.213 and 4.212, respectively.

Point	$e/b = 0.185$	0.337	0.543	0.811	1.151	1.577	2.107	2.762
A∗	3.520	3.413	3.321	3.248	3.188	3.141	3.102	3.079
A∗1	3.524	3.414		3.249		3.142	3.104	3.076
A∗2	3.293	3.302	3.274	3.229	3.181	3.139	3.103	3.075
A∗3	3.362	3.266	3.201	3.152	3.115	3.087	3.065	3.048
B∗	−1.012	−1.146	−1.205	−1.211	−1.191	−1.156	−1.122	−1.090
B∗1	−1.104	−1.151		−1.215		−1.157	−1.121	−1.089
B∗2	−0.730	−1.010	−1.148	−1.190	−1.182	−1.153	−1.119	−1.089
C∗	6.975	5.388	4.492	3.941	3.598	3.377	3.242	3.153
C∗1	6.960	5.390		3.944		3.377	3.238	3.151
C∗2	5.614	4.982	4.365	3.900	3.581	3.372	3.236	3.150
C∗3	5.064	4.366	3.919	3.609	3.396	3.254	3.162	3.108

Table 2. Comparison results for the normalized circumference stress $f_T = \sigma_T/p$ at some points for a circular hole ($b/a = 1$). See Figure 5. ∗ indicates data from the present study, ∗1 from [Dejoie et al. 2006], ∗2 from [Chen 1994], and ∗3 from [Savin 1961].
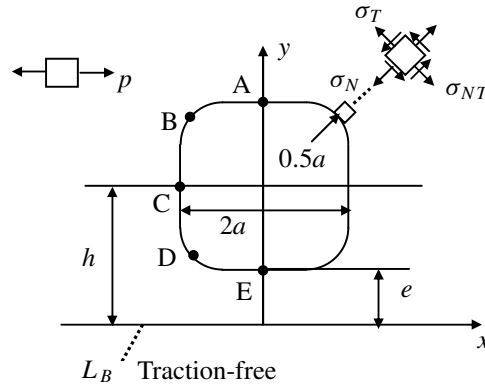
Figure 6. A square notch with rounded corners in the half-plane under the remote tension $\sigma_x = p$.

Point	$e/b = 0.1$	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
A	2.328	2.242	2.191	2.154	2.130	2.109	2.092	2.078	2.066	2.056
B	2.544	2.315	2.173	2.073	2.000	1.942	1.895	1.858	1.827	1.800
C	−0.483	−0.662	−0.751	−0.801	−0.828	−0.844	−0.852	−0.855	−0.855	−0.854
D	−1.427	−0.668	−0.123	0.245	0.504	0.693	0.837	0.952	1.043	1.117
E	5.213	4.212	3.703	3.378	3.141	2.959	2.813	2.696	2.598	2.517
Point	$e/b = 1$	2	3	4	5	6	7	8	9	10
A	2.056	1.992	1.967	1.953	1.940	1.937	1.928	1.925	1.923	1.921
B	1.800	1.669	1.627	1.609	1.595	1.591	1.584	1.582	1.581	1.580
C	−0.854	−0.818	−0.789	−0.773	−0.768	−0.765	−0.765	−0.762	−0.759	−0.758
D	1.117	1.446	1.527	1.558	1.565	1.573	1.572	1.574	1.575	1.575
E	2.517	2.143	2.031	1.985	1.956	1.946	1.936	1.931	1.927	1.925

Table 3. The normalized circumference stress f_T ($f_T = \sigma_T/p$) at some points for a square notch with rounded corners (see Figure 6).

4. Conclusions

Formulation of the fundamental solution based on Green's function method plays an important role in the present study. Since two stress fields for the fundamental solution and for the physical problem satisfy the traction-free condition along the boundary of the half-plane, we can formulate the BIE on the boundary of the notch only. In addition, the computed results for the case of a circular hole prove that the suggested BIE provides accurate results.

Appendix A: Derivation of (33) from (31)

The following is useful for the derivation of (33) from (31). We may assume a value as follows:

$$C = \text{Re}(i\,\overline{H}s) \quad \text{(where } H = -P_x - iP_y \text{ and } s \text{ is a complex value).} \tag{A.1}$$
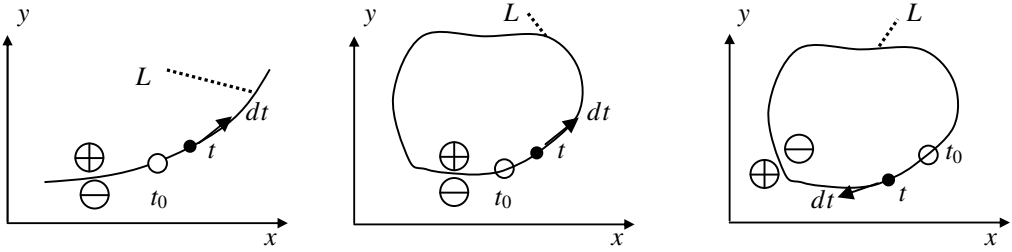
Figure 7. Left: a curve L in an infinite plane. Middle: a closed contour L in an infinite plane with dt in the counterclockwise direction. Right: a closed contour L in an infinite plane with dt in the clockwise direction.

Thus, we will find

$$C\big|_{P_x=1, P_y=0} + i C\big|_{P_x=0, P_y=1} = \mathrm{Re}(-is) + i\,\mathrm{Re}(-s) = \mathrm{Re}(-is) + i\,\mathrm{Im}(-is) = -is. \tag{A.2}$$

For example, one term in (31) can be expressed as follows:

$$S = \frac{1}{2G}\,\mathrm{Re}\int_\Gamma i\,(2\kappa\,\overline{F}\ln|t-\tau|)(\sigma_N + i\sigma_{NT})\,dt = B_2\,\mathrm{Re}\int_\Gamma i\,\overline{H}(2\kappa\ln|t-\tau|)(\sigma_N + i\sigma_{NT})\,dt. \tag{A.3}$$

By using (A.2), we have

$$S\big|_{P_x=1, P_y=0} + i S\big|_{P_x=0, P_y=1} = B_2 i \int_\Gamma (-2\kappa\ln|t-\tau|)(\sigma_N + i\sigma_{NT})\,dt. \tag{A.4}$$

Similarly, we can assume

$$D = \mathrm{Re}(i H t) \quad \text{(where } H = -P_x - i P_y, \text{ and } t \text{ is a complex value).} \tag{A.5}$$

Similarly, we have

$$D\big|_{P_x=1, P_y=0} + i D\big|_{P_x=0, P_y=1} = \mathrm{Re}(-it) + i\,\mathrm{Re}(t) = \mathrm{Re}(i\bar{t}) + i\,\mathrm{Re}(\bar{t}) = \mathrm{Re}(i\bar{t}) + i\,\mathrm{Im}(i\bar{t}) = i\bar{t}. \tag{A.6}$$

For example, one term in (31) can be expressed as follows:

$$T = \frac{1}{2G}\,\mathrm{Re}\int_\Gamma i\left(-F\frac{\bar{t}-\bar{\tau}}{t-\tau}\right)(\sigma_N + i\sigma_{NT})\,dt = B_2\,\mathrm{Re}\int_\Gamma i\left(-H\frac{\bar{t}-\bar{\tau}}{t-\tau}\right)(\sigma_N + i\sigma_{NT})\,dt. \tag{A.7}$$

By using (A.6), we have

$$T\big|_{P_x=1, P_y=0} + i T\big|_{P_x=0, P_y=1} = B_2 i \int_\Gamma -\frac{t-\tau}{\bar{t}-\bar{\tau}}(\sigma_N - i\sigma_{NT})\,d\bar{t}. \tag{A.8}$$

Appendix B: The generalized Sokhotski–Plemelj formulae

In plane elasticity, the following integrals are useful [Muskhelishvili 1953; Savruk 1981; Chen et al. 2003]:

$$F(z) = \frac{1}{2\pi i}\int_L \frac{f(t)\,dt}{t-z}, \quad G(z) = \frac{1}{2\pi i}\int_L \frac{g(t)\,d\bar{t}}{t-z}, \quad H(z, \bar{z}) = \frac{1}{2\pi i}\int_L \frac{\bar{t}-\bar{z}}{(t-z)^2}h(t)\,dt, \tag{B.1}$$

where L is a smooth curve or a closed contour as shown in Figure 7, and the integrals should be understood in the sense of the principal value. Also, we assume that the functions $f(t)$, $g(t)$, and $h(t)$ satisfy the Hölder condition [Muskhelishvili 1953]. Sometimes, the functions $f(t)$, $g(t)$, and $h(t)$ are called the density functions hereafter. Clearly, the first two integrals in (B.1) are analytic functions, and the last one is not. The first integral in (B.1) is the well-known Cauchy integral.

Generally speaking, these integrals take different values when $z \to t_0^+$ and $z \to t_0^-$ ($t_0 \in L$). The limit values of these functions from the upper and lower sides of the curve L are found to be [Muskhelishvili 1953; Savruk 1981; Chen et al. 2003]

$$F^{\pm}(t_0) = \pm \frac{f(t_0)}{2} + \frac{1}{2\pi i} \int_L \frac{f(t)\, dt}{t - t_0}, \tag{B.2}$$

$$G^{\pm}(t_0) = \pm \frac{g(t_0)}{2} \frac{d\bar{t}_0}{dt_0} + \frac{1}{2\pi i} \int_L \frac{g(t)\, d\bar{t}}{t - t_0}, \tag{B.3}$$

$$H^{\pm}(t_0, \bar{t}_0) = \pm \frac{h(t_0)}{2} \frac{d\bar{t}_0}{dt_0} + \frac{1}{2\pi i} \int_L \frac{\bar{t} - \bar{t}_0}{(t - t_0)^2} h(t)\, dt. \tag{B.4}$$

Note that the notations $f(t)$, $g(t)$, $h(t)$, $F(z)$, $G(z)$, and $H(z, \bar{z})$ used in (B.1) have no relation with those mentioned in other places.

Appendix C: Properties of some integrals
with kernel functions $L_1(t, z)$ and $L_2(t, z)$ defined by (38)

The two integrals with kernel functions $L_1(t, z)$ and $L_2(t, z)$ shown by (38) are written in the form

$$W_1(z) = \frac{1}{2\pi i} \int_{\Gamma} L_1(t, z) f(t)\, dt \quad (z \in S^+ \text{ or } z \in S^-), \tag{C.1}$$

$$W_2(z) = \frac{1}{2\pi i} \int_{\Gamma} L_2(t, z) f(t)\, dt \quad (z \in S^+ \text{ or } z \in S^-), \tag{C.2}$$

where

$$L_1(t, z) = -\frac{d}{dt} \left\{ \ln \frac{t - z}{\bar{t} - \bar{z}} \right\} = -\frac{1}{t - z} + \frac{1}{\bar{t} - \bar{z}} \frac{d\bar{t}}{dt}, \tag{C.3}$$

$$L_2(t, z) = \frac{d}{dt} \left\{ \frac{t - z}{\bar{t} - \bar{z}} \right\} = \frac{1}{\bar{t} - \bar{z}} - \frac{t - z}{(\bar{t} - \bar{z})^2} \frac{d\bar{t}}{dt}. \tag{C.4}$$

In (C.1) and (C.2), Γ denotes a closed contour and $f(t)$ is an arbitrary function. If dt goes forward in a counterclockwise direction, S^+ and S^- are the inside finite region and outside infinite region, respectively (Figure 3).

In (C.1) and (C.2), letting $z \to t_0$ ($z \in S^+$, $t_0 \in \Gamma$) and $z \to t_0$ ($z \in S^-$, $t_0 \in \Gamma$), and using the generalized Sokhotski–Plemelj formulae shown in Appendix B, we will find

$$W_1^{\pm}(t_0) = \mp f(t_0) + \frac{1}{2\pi i} \int_{\Gamma} L_1(t, t_0) f(t)\, dt \quad (t_0 \in \Gamma), \tag{C.5}$$

$$W_2^{\pm}(t_0) = \frac{1}{2\pi i} \int_{\Gamma} L_2(t, t_0) f(t)\, dt \quad (t_0 \in \Gamma). \tag{C.6}$$

We can prove the assertion shown by (C.5) as follows. In fact, we can rewrite $W_1(z)$ as

$$W_1(z) = I_1 + I_2 \quad (z \in S^+ \text{ or } z \in S^-), \tag{C.7}$$

where

$$I_1(z) = \frac{1}{2\pi i} \int_\Gamma \left(-\frac{1}{t-z} \right) f(t)\, dt \qquad (z \in S^+ \text{ or } z \in S^-), \tag{C.8}$$

$$I_2(z) = \frac{1}{2\pi i} \int_\Gamma \left(\frac{1}{\bar{t}-\bar{z}} \frac{d\bar{t}}{dt} \right) f(t)\, dt \quad (z \in S^+ \text{ or } z \in S^-). \tag{C.9}$$

For convenience in derivation, we can define

$$I_3(z) = -\overline{I_2(z)} = \frac{1}{2\pi i} \int_\Gamma \frac{1}{t-z} \overline{f(t)}\, dt \quad (z \in S^+ \text{ or } z \in S^-). \tag{C.10}$$

In (C.8) and (C.10), letting $z \to t_0$ ($z \in S^+$, $t_0 \in \Gamma$) and $z \to t_0$ ($z \in S^-$, $t_0 \in \Gamma$), and using the generalized Sokhotski–Plemelj formulae shown in Appendix B, we will find

$$I_1^\pm(t_0) = \mp \frac{f(t_0)}{2} + \frac{1}{2\pi i} \int_\Gamma \frac{-1}{t-t_0} f(t)\, dt \quad (t_0 \in \Gamma), \tag{C.11}$$

$$I_3^\pm(t_0) = -\overline{I_2^\pm(t_0)} = \pm \frac{\overline{f(t_0)}}{2} + \frac{1}{2\pi i} \int_\Gamma \frac{1}{t-t_0} \overline{f(t)}\, dt \quad (t_0 \in \Gamma). \tag{C.12}$$

In addition, from (C.10) and (C.12) we will find

$$I_2^\pm(t_0) = -\overline{I_3^\pm(t_0)} = \mp \frac{f(t_0)}{2} + \frac{1}{2\pi i} \int_\Gamma \left(\frac{1}{\bar{t}-\bar{t}_0} \frac{d\bar{t}}{dt} \right) f(t)\, dt \quad (t_0 \in \Gamma). \tag{C.13}$$

From (C.7), (C.11), and (C.13), the validity of (C.5) is proved. Similarly, we can prove the validity of (C.6).

Acknowledgment

References

[Brebbia et al. 1984] C. A. Brebbia, J. C. F. Telles, and L. C. Wrobel, *Boundary element techniques: theory and applications in engineering*, Springer, Heidelberg, 1984.

[Chen 1985] Y. Z. Chen, "A special boundary-element formulation for multiple-circular-hole problems in an infinite plate", *Comput. Methods Appl. Mech. Eng.* **50**:3 (1985), 263–273.

[Chen 1994] Y. Z. Chen, "Multiple circular hole problem for an elastic half-plane", *Comput. Struct.* **52** (1994), 1277–1281.

[Chen and Chen 2004] J. T. Chen and Y. W. Chen, "Dual boundary element analysis using complex variables for potential problems with or without a degenerate boundary", *Eng. Anal. Bound. Elem.* **24** (2004), 671–684.

[Chen and Cheung 1990] Y. Z. Chen and Y. K. Cheung, "New integral equation approach for the crack problem in elastic half-plane", *Int. J. Fract.* **46** (1990), 57–69.

[Chen and Lin 2007] Y. Z. Chen and X. Y. Lin, "Solution of periodic group circular hole problems by using series expansion variational method", *Int. J. Numer. Methods Eng.* **69** (2007), 1405–1422.

[Chen et al. 2003] Y. Z. Chen, N. Hasebe, and K. Y. Lee, *Multiple crack problems in elasticity*, Advances in Damage Mechanics **4**, WIT Press, Southampton, 2003.

[Chen et al. 2009] Y. Z. Chen, X. Y. Lin, and Z. X. Wang, "Numerical solution for curved crack problem in elastic half-plane using hypersingular integral equation", *Philos. Mag.* **86** (2009), 2239–2253.

[Chen et al. 2010] Y. Z. Chen, X. Y. Lin, and Z. X. Wang, "Formulation of indirect BIEs in plane elasticity using single or double layer potentials and complex variable", *Eng. Anal. Bound. Elem.* **34**:4 (2010), 337–351.

[Cheng and Cheng 2005] A. H. D. Cheng and D. T. Cheng, "Heritage and early history of the boundary element method", *Eng. Anal. Bound. Elem.* **29** (2005), 268–302.

[Cruse 1969] T. A. Cruse, "Numerical solutions in three-dimensional elastostatics", *Int. J. Solids Struct.* **5** (1969), 1259–1274.

[Dejoie et al. 2006] A. Dejoie, S. G. Mogilevskaya, and S. L. Crouch, "A boundary integral method for multiple circular holes in an elastic half-plane", *Eng. Anal. Bound. Elem.* **30** (2006), 450–464.

[Hong and Chen 1988] H. K. Hong and J. T. Chen, "Derivations of integral equations of elasticity", *J. Eng. Mech.* (*ASCE*) **114** (1988), 1028–1044.

[Hromadka and Lai 1987] T. V. Hromadka, II and C. Lai, *The complex variable boundary element method in engineering analysis*, Springer, New York, 1987.

[Jaswon and Symm 1977] M. A. Jaswon and G. T. Symm, *Integral equation methods in potential theory and elastostatics*, Academic Press, London, 1977.

[Linkov 2002] A. M. Linkov, *Boundary integral equations in elasticity theory*, Solid Mechanics and its Applications **99**, Kluwer, Dordrecht, 2002.

[Linkov and Mogilevskaya 1994] A. M. Linkov and S. G. Mogilevskaya, "Complex hypersingular integrals and integral equations in plane elasticity", *Acta Mech.* **105**:1-4 (1994), 189–205.

[Mogilevskaya and Linkov 1998] S. G. Mogilevskaya and A. M. Linkov, "Complex fundamental solutions and complex variables boundary element method in elasticity", *Comput. Mech.* **22**:1 (1998), 88–92.

[Muskhelishvili 1953] N. I. Muskhelishvili, *Some basic problems of the mathematical theory of elasticity*, Noordhoff, Groningen, 1953.

[Pan et al. 1997] E. Pan, C. S. Chen, and B. Amadei, "A BEM formulation for anisotropic half-plane problems", *Eng. Anal. Bound. Elem.* **20** (1997), 185–195.

[Rizzo 1967] F. J. Rizzo, "An integral equation approach to boundary value problems in classical elastostatics", *Quart. Appl. Math.* **25** (1967), 83–95.

[Savin 1961] G. N. Savin, *Stress concentration around holes*, International Series of Monographs in Aeronautics and Astronautics, Division I: Solid and Structural Mechanics **1**, Pergamon, New York, 1961.

[Savruk 1981] M. P. Savruk, Двухмерные задачи упругости для тел с трещинами, Naukova Dumka, Kiev, 1981.

[Verruijt 1998] A. Verruijt, "Deformations of an elastic half plane with a circular cavity", *Int. J. Solids Struct.* **35** (1998), 2795–2804.

[Wang et al. 2003] J. L. Wang, S. L. Crouch, and S. G. Mogilevskaya, "A complex boundary integral method for multiple circular holes in an infinite plane", *Eng. Anal. Bound. Elem.* **27** (2003), 789–802.

Y. Z. CHEN: chens@ujs.edu.cn
Division of Engineering Mechanics, Jiangsu University, Xue Fu Road 301, Zhenjiang, Jiangsu 212013, China

msp

INTERNAL STRUCTURES AND INTERNAL VARIABLES IN SOLIDS

JÜRI ENGELBRECHT AND ARKADI BEREZOVSKI

The response of many materials (metals, alloys, composites, etc.) to external loading may be essentially influenced by an existing or emerging internal structure at smaller scales which must be taken into account. For this purpose the concept of dual internal variables can be used in order to describe the effect of internal fields. In this paper it is shown that dual internal variable theory is sufficiently general to model cases like micromorphic elasticity and the influence of microtemperature. Based on the material (canonical) balance equations for material momentum and energy, this approach extends single internal variable theory. The resulting governing equations are not limited by first-order reaction-diffusion equations, as is typical for single internal variable theory. Hyperbolic governing equations for internal variables provide the description of the interaction of waves at the macro and microlevels.

1. Introduction

The observed complexity in the dynamic behavior of solids is due to the fact that solids are inherently microstructured. *Internal structures* appear at different length scales and often coexist at more than one length scale within the same solid. Description of the microstructure's influence on the macromotion is therefore a necessary step for both theory and practice. However, as pointed out in [Kirchner and Steinmann 2005], "there is no unique answer to the question how the microstructure influence can be accounted for in a continuum mechanical model." In general terms, microstructures in elastic solids can be either natural or man-made. Man-made microstructures are usually completely ordered (like laminates), and then their influence can be taken into account by direct numerical computations by making use of well-defined geometry. In contrast, the influence of disordered microstructures (both natural and man-made) leads to certain internal fields, which affect the macroscopic behavior; the modeling of these fields may be an effective way to understand the complexity of such solids.

Over the past five decades, a number of advanced generalized continuum theories have been introduced to take into account the influence of structural inhomogeneities on the macroscopic behavior of materials (see overview in [Maugin 2011b]). The mathematical structure of such theories [Capriz 1989] includes a coarse-grained morphological descriptor introduced to describe the morphology of the material element [Mariano and Stazi 2005], which represents certain additional independent fields [Mariano 2002]. For example, considering the material element as a cell able to deform independently of the rest of the body, Mindlin [1964] in fact introduced a second-order symmetric tensor as a morphological descriptor. The relevant continuum theory is called micromorphic [Eringen and Suhubi 1964a; 1964b].

Recently, it has been found that the time evolution laws for the averaged conserved dynamical variables

derived from a molecular dynamic model follow exactly the balance laws in a micromorphic continuum [Chen and Lee 2003a]. Moreover, a multiscale field theory has been constructed for the concurrent atomic-continuum modeling of materials [Chen and Lee 2005; 2006]. Micromorphic theory is thus considered the most successful top-down formulation of a two-level continuum model, in which the deformation is composed of the macroscopic continuous deformation and the internal microscopic deformation of the inner structure [Hirschberger et al. 2007; Forest 2009; Grammenoudis and Tsakmakis 2009; Wang and Lee 2010; Gonella et al. 2011].

Alternatively, one can introduce *internal variables* to describe microstructural effects. Application of internal variables to the description of microstructural influence on global motion has a long history. As pointed out in [Truesdell 1984], Duhem was the first to introduce what are now called internal state variables. In the 1940s, Bridgman proposed an introduction of "a large scale thermodynamic parameter of state" [Bridgman 1961], which extends the state space. The thermodynamic theory of internal variables [Coleman and Gurtin 1967] had presupposed only first-order evolution equations for the internal variables and did not include their gradients. Accounting for the gradients leads to the weakly nonlocal theory [Maugin and Muschik 1994; Maugin 1999], which can be also enriched by the extra entropy flux [Maugin 1990]. The comprehensive theory of the internal state variables was presented recently in [Maugin 2006].

Internal variables are usually responsible for dissipative processes and must satisfy only the second law of thermodynamics. It is hoped that a few aggregate internal variables will adequately describe the influence of a microstructure [Rice 1971; Muschik 1990; Maugin and Muschik 1994]. However, as noted in [Kirchner and Steinmann 2005], "it is neither a priori known which specific features of the microstructure characterize such a macroscopic internal variable, nor whether the macroscopic behavior is described sufficiently accurately by such a quantity."

At the same time, the recently developed concept of dual internal variables [Ván et al. 2008] permits retrieving Mindlin micromorphic theory [Berezovski et al. 2011b] on the basis of the material formulation of continuum mechanics [Maugin 1993; 2011a]. This means that the dual internal variable approach is at least as powerful as the widely accepted micromorphic description. The material formulation takes internal variables into account naturally and consistently. Moreover, the structure of the governing equations for the microfields ensues directly from the Clausius–Duhem inequality, and is not considered as granted like it is in multifield theories [Capriz 1989] or derived from the requirement of invariance of the external power actions [Mariano 2002].

It has been noted that the question of how thermal effects should be taken into account in higher-order theories needs further analysis [Forest and Aifantis 2010]. In the developed dual internal variable theory temperature effects can be included consistently [Berezovski et al. 2011c].

The objective of the present paper is to demonstrate how different models of *internal structure* are related to a systematic thermomechanical method of describing *internal fields* in solids, that we call the *dual internal variable* approach. This theory is the direct extension of the comprehensive single internal variable theory [Maugin 2006]. The theory is presented here in the tensorial form generalizing preliminary results in a one-dimensional setting [Berezovski et al. 2009a; 2009b].

2. Thermomechanics with dual internal variables

2A. *Piola–Kirchhoff formulation.* The starting point in the formulation of thermomechanics with dual internal variables is the standard continuum mechanics. To formulate the balance laws of continuum

mechanics, we consider the motion of a body as a time-parametrized mapping χ, which connects a material point X in the reference configuration and its position x in the actual configuration in Euclidean physical space:

$$x = \chi(X, t). \tag{2-1}$$

Accordingly, the deformation gradient is defined by

$$F = \left.\frac{\partial \chi}{\partial X}\right|_t = \nabla_R \chi. \tag{2-2}$$

At any regular material point X in a continuous body in the presence of a body force f_0 per unit reference volume, the local balance laws in the so-called Piola–Kirchhoff formulation include the conservation of mass,

$$\left.\frac{\partial \rho_0}{\partial t}\right|_X = 0, \tag{2-3}$$

the balance of linear momentum,

$$\left.\frac{\partial (\rho_0 v)}{\partial t}\right|_X - \mathrm{Div}_R \, T = f_0, \tag{2-4}$$

and the energy conservation equation,

$$\left.\frac{\partial (K + E)}{\partial t}\right|_X - \nabla_R \cdot (T \cdot v - Q) = f_0 \cdot v. \tag{2-5}$$

The balance laws (2-3)–(2-5) are complemented by the second law of thermodynamics:

$$\left.\frac{\partial S}{\partial t}\right|_X + \nabla_R \cdot S \geq 0, \quad S = (Q/\theta) + J. \tag{2-6}$$

Here ρ_0 is the mass density in the reference configuration, $v = \partial \chi / \partial t|_X$ is the physical velocity, T is the first Piola–Kirchhoff stress tensor, $K = \frac{1}{2}\rho_0 v^2$ is the kinetic energy per unit reference volume, E is the internal energy per unit reference volume, Q is the material heat flux, S is the entropy density per unit reference volume, θ is the absolute temperature, S is the entropy flux, and the "extra entropy flux" J vanishes in most cases, although this is not a basic requirement, and $d/dt = \partial/\partial t|_X$ or a superimposed dot denotes the material time derivative.

2B. *Material (canonical) formulation.* It is well known that balance laws (2-4) and (2-5) can be represented in the so-called material (canonical) formulation [Maugin 1993; 2011a]. The canonical form of energy conservation for sufficiently smooth fields at any regular material point X in the body has the form [Maugin 2006; Maugin and Berezovski 2008]

$$\left.\frac{\partial (S\theta)}{\partial t}\right|_X + \nabla_R \cdot Q = h^{\mathrm{int}}, \quad h^{\mathrm{int}} := T : \dot{F} - \left.\frac{\partial W}{\partial t}\right|_X. \tag{2-7}$$

Here $W = E - \theta S$ is the Helmholtz free energy function.

Correspondingly, the canonical (material) momentum balance equation in the presence of a body force f_0 per unit reference volume is written as [Maugin 2006; Maugin and Berezovski 2008]

$$\left.\frac{\partial P}{\partial t}\right|_X - \mathrm{Div}_R \, b = f^{\mathrm{int}} + f^{\mathrm{ext}} + f^{\mathrm{inh}}, \tag{2-8}$$

where the *material momentum* $\boldsymbol{P}$, the material *Eshelby stress* $\boldsymbol{b}$, the material *inhomogeneity force* $\boldsymbol{f}^{\text{inh}}$, the material *external* (or body) force $\boldsymbol{f}^{\text{ext}}$, and the material *internal* force $\boldsymbol{f}^{\text{int}}$ are defined by

$$\boldsymbol{P} := -\rho_0 \boldsymbol{v} \cdot \boldsymbol{F}, \quad \boldsymbol{b} = -(L\boldsymbol{I}_R + \boldsymbol{T}.\boldsymbol{F}), \quad L = K - W, \tag{2-9}$$

$$\boldsymbol{f}^{\text{inh}} := \left.\frac{\partial L}{\partial \boldsymbol{X}}\right|_{\text{expl}} \equiv \left.\frac{\partial L}{\partial \boldsymbol{X}}\right|_{\text{fixed fields}} = \left(\tfrac{1}{2}\boldsymbol{v}^2\right)\nabla_R \rho_0 - \left.\frac{\partial W}{\partial \boldsymbol{X}}\right|_{\text{expl}}, \tag{2-10}$$

$$\boldsymbol{f}^{\text{ext}} := -\boldsymbol{f}_0 \cdot \boldsymbol{F}, \quad \boldsymbol{f}^{\text{int}} = \boldsymbol{T} : (\nabla_R \boldsymbol{F})^T - \nabla_R W\big|_{\text{impl}}. \tag{2-11}$$

Here a subscript expl means the material gradient keeping the fields fixed (and thus extracting the explicit dependence on $\boldsymbol{X}$), while impl indicates the material gradient taken only through the fields present in the function. The dot notation is used for the product of two tensors.

The second law (2-6) multiplied by θ yields the Clausius–Duhem inequality:

$$-\left(\frac{\partial W}{\partial t} + S\frac{\partial \theta}{\partial t}\right)\bigg|_X + \boldsymbol{T} : \dot{\boldsymbol{F}} + \nabla_R \cdot (\theta \boldsymbol{J}) - \boldsymbol{S} \cdot \nabla_R \theta \geq 0. \tag{2-12}$$

The canonical equations (2-7) and (2-8) are the most general equations for momentum and energy we can write down without a specification of the full dependency of free energy [Maugin 2006; 2011a; Maugin and Berezovski 2008]. These equations provide the basis of the description of a microstructured medium by means of the introduction of additional internal fields.

2C. *Dual internal variables.* The role of internal variables in continuum mechanics is described in [Maugin 1990; 2006; Maugin and Muschik 1994] bearing in mind dissipative processes. As was shown in [Ván et al. 2008], a unified treatment of both dissipative and nondissipative internal processes is possible in the framework of *dual* internal variable theory. This approach is an extension of classical single internal variable theory. Moreover, the limitation of governing equations to the first order is avoided by the concept of dual internal variables [Ván et al. 2008]. In what follows, the application of dual internal variables to thermomechanics [Berezovski et al. 2011b] is recalled shortly.

Thus, in the framework of the phenomenological continuum theory, it is assumed that the influence of a microstructure on the overall macroscopic motion of a body can be taken into account by the introduction of internal variables, which we associate with the integral distributed effect of the microstructure. The free energy W depends (in addition to the deformation gradient and temperature) on two internal variables, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, each of which is a second-order tensor, and their gradients

$$W = \overline{W}(\boldsymbol{F}, \theta, \boldsymbol{\alpha}, \nabla_R \boldsymbol{\alpha}, \boldsymbol{\beta}, \nabla_R \boldsymbol{\beta}). \tag{2-13}$$

The inclusion of gradients into the state space is related to a weak nonlocality of the theory. In this case, the equations of state are given by

$$\boldsymbol{T} = \frac{\partial \overline{W}}{\partial \boldsymbol{F}}, \quad S = -\frac{\partial \overline{W}}{\partial \theta}, \quad \boldsymbol{A} := -\frac{\partial \overline{W}}{\partial \boldsymbol{\alpha}}, \quad \mathscr{A} := -\frac{\partial \overline{W}}{\partial \nabla_R \boldsymbol{\alpha}}, \quad \boldsymbol{B} := -\frac{\partial \overline{W}}{\partial \boldsymbol{\beta}}, \quad \mathscr{B} := -\frac{\partial \overline{W}}{\partial \nabla_R \boldsymbol{\beta}}. \tag{2-14}$$

The main advantage of the material formulation of continuum mechanics is that the canonical equations

of momentum and energy keep their forms [Berezovski et al. 2011b]:

$$\left.\frac{\partial \boldsymbol{P}}{\partial t}\right|_X - \mathrm{Div}_R\,\tilde{\boldsymbol{b}} = \boldsymbol{f}^{\mathrm{th}} + \tilde{\boldsymbol{f}}^{\mathrm{intr}},$$
$$\left.\frac{\partial (S\theta)}{\partial t}\right|_X + \nabla_R \cdot \tilde{\boldsymbol{Q}} = h^{\mathrm{th}} + \tilde{h}^{\mathrm{intr}}, \tag{2-15}$$

with the modified Eshelby stress tensor

$$\tilde{\boldsymbol{b}} = -(L\mathbf{1}_R + \boldsymbol{T}\,.\,\boldsymbol{F} - \mathscr{A}:(\nabla_R\boldsymbol{\alpha})^T - \mathscr{B}:(\nabla_R\boldsymbol{\beta})^T), \tag{2-16}$$

the thermal source terms

$$\boldsymbol{f}^{\mathrm{th}} = S\nabla_R\theta, \quad h^{\mathrm{th}} = S\dot{\theta}, \tag{2-17}$$

and the intrinsic source terms

$$\tilde{\boldsymbol{f}}^{\mathrm{intr}} := \tilde{\mathscr{A}}:\nabla_R\boldsymbol{\alpha} + \tilde{\mathscr{B}}:\nabla_R\boldsymbol{\beta}, \quad \tilde{h}^{\mathrm{intr}} := \tilde{\mathscr{A}}:\dot{\boldsymbol{\alpha}} + \tilde{\mathscr{B}}:\dot{\boldsymbol{\beta}}. \tag{2-18}$$

In the above equations the following definitions are used

$$\tilde{\mathscr{A}} := -\left(\frac{\partial \overline{W}}{\partial \boldsymbol{\alpha}} - \mathrm{Div}_R\,\frac{\partial \overline{W}}{\partial (\nabla_R\boldsymbol{\alpha})}\right) = \boldsymbol{A} - \mathrm{Div}_R\,\mathscr{A}, \tag{2-19}$$

$$\tilde{\mathscr{B}} := -\left(\frac{\partial \overline{W}}{\partial \boldsymbol{\beta}} - \mathrm{Div}_R\,\frac{\partial \overline{W}}{\partial (\nabla_R\boldsymbol{\beta})}\right) = \boldsymbol{B} - \mathrm{Div}_R\,\mathscr{B}, \tag{2-20}$$

$$\tilde{\boldsymbol{S}} = \theta^{-1}\tilde{\boldsymbol{Q}}, \quad \tilde{\boldsymbol{Q}} = \boldsymbol{Q} - \mathscr{A}:\dot{\boldsymbol{\alpha}} - \mathscr{B}:\dot{\boldsymbol{\beta}}. \tag{2-21}$$

In this formulation the Eshelby stress complies with its role of grasping all the effects presenting gradients since the material gradients of internal variables play a role parallel to that of the deformation gradient $\boldsymbol{F}$.

2D. *Governing equations for internal variables.* Following the scheme originally developed in [Maugin 1990] for materials with diffusive dissipative processes described by means of internal variables of state, we chose the nonzero extra entropy flux in the form

$$\boldsymbol{J} = -\theta^{-1}\mathscr{A}:\dot{\boldsymbol{\alpha}} - \theta^{-1}\mathscr{B}:\dot{\boldsymbol{\beta}}. \tag{2-22}$$

The dissipation inequality (2-12) is then reduced to

$$\Phi = \tilde{\mathscr{A}}:\dot{\boldsymbol{\alpha}} + \tilde{\mathscr{B}}:\dot{\boldsymbol{\beta}} - \tilde{\boldsymbol{S}}\nabla_R\theta \geq 0, \tag{2-23}$$

and contains both intrinsic and thermal parts. The thermal part of this inequality can be satisfied by modification of the Fourier law [Berezovski et al. 2011c], which is nothing but the standard proportionality of the heat flux with respect to the temperature gradient

$$\boldsymbol{Q} - \mathscr{A}:\dot{\boldsymbol{\alpha}} - \mathscr{B}:\dot{\boldsymbol{\beta}} = -a^2\nabla_R\theta. \tag{2-24}$$

The intrinsic part of the dissipation inequality (2-23) depends solely on internal fields:

$$\tilde{h}^{\mathrm{intr}} := \tilde{\mathscr{A}}:\dot{\boldsymbol{\alpha}} + \tilde{\mathscr{B}}:\dot{\boldsymbol{\beta}} \geq 0. \tag{2-25}$$

The governing equations for the internal variables $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ yield from (2-25):

$$\begin{pmatrix} \dot{\boldsymbol{\alpha}} \\ \dot{\boldsymbol{\beta}} \end{pmatrix} = \boldsymbol{R} \begin{pmatrix} \tilde{\mathscr{A}} \\ \tilde{\mathscr{B}} \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} \dot{\boldsymbol{\alpha}} \\ \dot{\boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{R}^{11} & \boldsymbol{R}^{12} \\ \boldsymbol{R}^{21} & \boldsymbol{R}^{22} \end{pmatrix} \begin{pmatrix} \tilde{\mathscr{A}} \\ \tilde{\mathscr{B}} \end{pmatrix}, \tag{2-26}$$

where components $\boldsymbol{R}^{11}, \ldots, \boldsymbol{R}^{22}$ of the linear operator $\boldsymbol{R}$ are dependent on state variables [Gurtin 1996].

2E. *Remarks*. The governing equations for the internal variables, (2-26), complete the basic theory of the thermomechanics with dual internal variables. Formally, this theory is the direct extension of the comprehensive single internal variable theory [Maugin 2006]. This is a weakly nonlocal theory in the material formulation enriched by the extra entropy flux similarly to the single internal variable theory [Ván et al. 2008]. However, the introduction of the additional internal variables may change the mathematical structure of the theory. As it was demonstrated in [Berezovski et al. 2011c], in addition to the dissipative part, the dual internal variables contribute also to the reversible Poisson structure [Mielke 2011]. This is the origin of hyperbolic governing equations for the dual internal variables that generalize the internal variable theory significantly.

The given formulation of the theory is certainly of a general character and should be specified to describe the particular influence of an internal structure. We will demonstrate the ability of the theory on certain examples focusing on the explicit form of the governing equations for the internal variables. The governing equations coupled with the balance of momentum and energy constitute the corresponding continuum model with the influence of internal structure. Though each particular case is characterized by an explicit form of the free energy, we will keep the free energy as general as possible.

3. First example: The pure dissipative case

Representing the linear operator $\boldsymbol{R}$ as the sum of symmetric and skew-symmetric components $\boldsymbol{R} = (\boldsymbol{R} + \boldsymbol{R}^T)/2 + (\boldsymbol{R} - \boldsymbol{R}^T)/2$, that is,

$$\begin{pmatrix} \dot{\boldsymbol{\alpha}} \\ \dot{\boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{R}^{11} & (\boldsymbol{R}^{12} + \boldsymbol{R}^{21})/2 \\ (\boldsymbol{R}^{21} + \boldsymbol{R}^{12})/2 & \boldsymbol{R}^{22} \end{pmatrix} \begin{pmatrix} \tilde{\mathscr{A}} \\ \tilde{\mathscr{B}} \end{pmatrix} + \begin{pmatrix} 0 & (\boldsymbol{R}^{12} - \boldsymbol{R}^{21})/2 \\ (\boldsymbol{R}^{21} - \boldsymbol{R}^{12})/2 & 0 \end{pmatrix} \begin{pmatrix} \tilde{\mathscr{A}} \\ \tilde{\mathscr{B}} \end{pmatrix}, \tag{3-1}$$

we can see that the symmetry of the linear operator $\boldsymbol{R}$, which is equivalent to the Onsager reciprocity relations $\boldsymbol{R}^{12} = \boldsymbol{R}^{21}$, leads to the elimination of the antisymmetric part of the linear operator $\boldsymbol{R}$. In this case, we return to the classical situation, where internal variables are fully independent, dissipative, and governed by reaction-diffusion equations [Coleman and Gurtin 1967; Maugin and Muschik 1994; Maugin 2006].

In fact, the governing equations for the internal variables in this case,

$$\dot{\boldsymbol{\alpha}} = \boldsymbol{R}^{11} . \tilde{\mathscr{A}}, \quad \dot{\boldsymbol{\beta}} = \boldsymbol{R}^{22} . \tilde{\mathscr{B}}, \tag{3-2}$$

automatically provide the nonnegativity of the intrinsic part of the dissipation inequality (2-25),

$$\tilde{h}^{\text{intr}} := \tilde{\mathscr{A}} : \dot{\boldsymbol{\alpha}} + \tilde{\mathscr{B}} : \dot{\boldsymbol{\beta}} \geq 0, \tag{3-3}$$

if $\boldsymbol{R}^{11}$ and $\boldsymbol{R}^{22}$ are positive definite. The reaction-diffusion-like equations obtained,

$$\dot{\boldsymbol{\alpha}} = \boldsymbol{R}^{11} . (\boldsymbol{A} - \text{Div}_R \mathscr{A}) \quad \text{and} \quad \dot{\boldsymbol{\beta}} = \boldsymbol{R}^{22} . (\boldsymbol{B} - \text{Div}_R \mathscr{B}), \tag{3-4}$$

can be found under different names in numerous applications. For example, if the free energy depends on a scalar internal variable α as

$$W = \overline{W}(\ldots, \alpha, \nabla_R \alpha) = f(\ldots, \alpha) + \tfrac{1}{2} D(\nabla \alpha)^2, \tag{3-5}$$

we arrive at the Ginzburg–Landau (or the Allen–Cahn) equation (see [Cross and Hohenberg 1993])

$$\frac{1}{k}\dot{\alpha} = D\nabla^2 \alpha - f'(\alpha), \tag{3-6}$$

where $f'(\alpha)$ denotes the derivative with respect to α.

As we can conclude, the single internal variable theory is a particular case of the theory with dual internal variables. In the case of a pure dissipative internal structure, there is no need to go beyond the single internal variable theory. However, we have no reason to assume the symmetry of the linear operator $\boldsymbol{R}$ in the case of arbitrary internal structures.

4. Second example: The pure nondissipative case

The pure nondissipative case corresponds to the skew-symmetric matrix $\boldsymbol{R}$, which implies $\boldsymbol{R}^{11} = \boldsymbol{R}^{22} = \boldsymbol{0}$ and the Casimir reciprocity relations $\boldsymbol{R}^{12} = -\boldsymbol{R}^{21}$. The governing equations for the dual internal variables are fully coupled:

$$\dot{\boldsymbol{\alpha}} = \boldsymbol{R}^{12}.\tilde{\mathcal{B}} = \boldsymbol{R}^{12}.(\boldsymbol{B} - \mathrm{Div}_R \, \mathcal{B}), \tag{4-1}$$

$$\dot{\boldsymbol{\beta}} = -\boldsymbol{R}^{12}.\tilde{\mathcal{A}} = -\boldsymbol{R}^{12}.(\boldsymbol{A} - \mathrm{Div}_R \, \mathcal{A}), \tag{4-2}$$

and the dissipation $\tilde{h}^{\mathrm{intr}}$ vanishes. In this case, the evolution of one internal variable is driven by another one that manifests the duality between the internal variables.

To be more specific, let us consider a simple case with $\mathcal{B} = \boldsymbol{0}$, which means that the free energy function $\overline{W}$ is independent of $\nabla_R \boldsymbol{\beta}$. Then governing equation for the first internal variable (4-1) is reduced to

$$\dot{\boldsymbol{\alpha}} = \boldsymbol{R}^{12}.\boldsymbol{B}. \tag{4-3}$$

Assuming further a quadratic dependence of the free energy function with respect to the internal variable $\boldsymbol{\beta}$

$$\boldsymbol{B} := -\frac{\partial \overline{W}}{\partial \boldsymbol{\beta}} = -b\boldsymbol{\beta}, \tag{4-4}$$

we reduce (4-3) even more:

$$\dot{\boldsymbol{\alpha}} = -b\boldsymbol{R}^{12}.\boldsymbol{\beta}. \tag{4-5}$$

Substituting from (4-5) into (4-2), we arrive at a hyperbolic governing equation for the primary internal variable $\boldsymbol{\alpha}$:

$$\ddot{\boldsymbol{\alpha}} = (b\boldsymbol{R}^{12}.\boldsymbol{R}^{12}).\tilde{\mathcal{A}} = (b\boldsymbol{R}^{12}.\boldsymbol{R}^{12}).\left(-\frac{\partial \overline{W}}{\partial \boldsymbol{\alpha}} + \mathrm{Div}_R \, \frac{\partial \overline{W}}{\partial (\nabla_R \boldsymbol{\alpha})}\right). \tag{4-6}$$

To exemplify the latter, still sufficiently general governing equation, we need to identify the abstract primary internal variable $\boldsymbol{\alpha}$ with a certain well-known (micro)field variable. We will demonstrate such a representation on the example of Mindlin micromorphic theory [Mindlin 1964].

4A. *Mindlin micromorphic theory.* In the framework of Mindlin micromorphic theory [Mindlin 1964], each material point is endowed with three translational degrees of freedom $\boldsymbol{u}$ and a second-order microdeformation tensor $\boldsymbol{\psi}$ with nine independent components. In the case of centrosymmetric, isotropic materials, the equations of motion in terms of stresses [Mindlin 1964] can be represented in the form

$$\rho\dot{\boldsymbol{v}} = \operatorname{div}(\boldsymbol{\sigma} + \boldsymbol{\tau}) + \boldsymbol{f}, \tag{4-7}$$

$$\boldsymbol{I}.\ddot{\boldsymbol{\psi}} = \operatorname{div}\boldsymbol{\mu} + \boldsymbol{\tau} + \Phi, \tag{4-8}$$

where $\boldsymbol{I}$ is a microinertia tensor, $\boldsymbol{f}$ is the body force, and Φ is the double force per unit volume. The corresponding stress tensors, namely, the Cauchy stress $\boldsymbol{\sigma}$, the relative stress $\boldsymbol{\tau}$, and the double stress $\boldsymbol{\mu}$ are defined, respectively, as derivatives of the free energy with respect to the classical strain tensor $\boldsymbol{\varepsilon}$, the relative deformation tensor $\boldsymbol{\gamma}$, and the microdeformation gradient $\boldsymbol{\varkappa}$ [Mindlin 1964]:

$$\boldsymbol{\sigma} \equiv \frac{\partial W}{\partial \boldsymbol{\varepsilon}}, \quad \boldsymbol{\tau} \equiv \frac{\partial W}{\partial \boldsymbol{\gamma}}, \quad \boldsymbol{\mu} \equiv \frac{\partial W}{\partial \boldsymbol{\varkappa}}, \tag{4-9}$$

where

$$\boldsymbol{\varepsilon} \equiv \tfrac{1}{2}(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T), \quad \boldsymbol{\gamma} \equiv \nabla\boldsymbol{u} - \boldsymbol{\psi}, \quad \boldsymbol{\varkappa} \equiv \nabla\boldsymbol{\psi}. \tag{4-10}$$

The free energy density W is supposed to be a homogeneous, quadratic function of forty-two variables $\boldsymbol{\varepsilon}, \boldsymbol{\gamma}, \boldsymbol{\varkappa}$ [Mindlin 1964]. It should be noted that the balances of linear momentum both at the micro and macro level ((4-7) and (4-8), respectively) are introduced independently.

4B. *Rearrangement.* As it is shown in [Berezovski et al. 2011b], the constitutive relations in the micromorphic Mindlin theory can be represented in terms of the distortion $\nabla\boldsymbol{u}$ and microdeformation tensor $\boldsymbol{\psi}$. Accordingly, the stresses are represented as follows:

$$\boldsymbol{\sigma}' \equiv \frac{\partial W}{\partial \nabla\boldsymbol{u}}, \quad \boldsymbol{\tau}' \equiv \frac{\partial W}{\partial \boldsymbol{\psi}}. \tag{4-11}$$

The double stress remains unchanged. The equations of motion (4-7) and (4-8) then take on the form [Berezovski et al. 2011b]

$$\rho\dot{\boldsymbol{v}} = \operatorname{div}\boldsymbol{\sigma}' + \boldsymbol{f}, \tag{4-12}$$

$$\boldsymbol{I}.\ddot{\boldsymbol{\psi}} = \operatorname{div}\boldsymbol{\mu} - \boldsymbol{\tau}' + \Phi. \tag{4-13}$$

The change of sign on the right-hand side of the governing equation for the microdeformation (4-13) follows from the opposite signs of $\boldsymbol{\gamma}$ and $\boldsymbol{\psi}$ (see (4-10)$_2$).

4C. *Microdeformation tensor as an internal variable.* Now we consider the microdeformation tensor $\boldsymbol{\psi}$ as an internal variable $\boldsymbol{\alpha}$ and apply the formalism developed above. The microdeformation gradient $\boldsymbol{\varkappa}$ plays the role of the gradient of the internal variable $\boldsymbol{\alpha}$, and we introduce a dual internal variable $\boldsymbol{\beta}$ in the same way as previously described. In the nondissipative case, the dual internal variable $\boldsymbol{\beta}$ is auxiliary and does not affect the calculation of the derivatives of free energy with respect to microdeformation and double stress. Therefore, the governing equation for the internal variable $\boldsymbol{\alpha}$ follows from (4-6):

$$\ddot{\boldsymbol{\alpha}} = (b\boldsymbol{R}^{12}.\boldsymbol{R}^{12}).\left(-\frac{\partial W}{\partial \boldsymbol{\alpha}} + \operatorname{Div}\frac{\partial W}{\partial (\nabla\boldsymbol{\alpha})}\right). \tag{4-14}$$

Identifying the internal variable $\boldsymbol{\alpha}$ with the microdeformation tensor $\boldsymbol{\psi}$, the latter governing equation takes on the form

$$(b\boldsymbol{R}^{12}.\boldsymbol{R}^{12})^{-1}.\ddot{\boldsymbol{\psi}} = \left(-\frac{\partial W}{\partial \psi_{jk}} + \text{Div}\,\frac{\partial W}{\partial(\nabla\psi_{jk})}\right) = \text{div}\,\boldsymbol{\mu} - \boldsymbol{\tau}'. \tag{4-15}$$

As one can see, the governing equation for the microdeformation (4-15) is practically the same as the equation of motion at the microlevel (4-13) in the rearranged Mindlin theory. The external double force cannot appear in the internal variable theory [Berezovski et al. 2011b]. It should be noted that equation of motion (4-15) is not postulated, but follows from the dissipation inequality for the chosen functional dependence of the free energy in the considered nondissipative case. Clearly, this approach can be applied to microstrain and micropolar theories (with corresponding changes) since they are particular cases of the micromorphic theory [Eringen 1999].

5. Third example: A more general case (microtemperature effects)

Now we consider a more complicated case, including dissipative and nondissipative contributions. We choose the matrix $\boldsymbol{R}$ in (2-26) as follows:

$$\boldsymbol{R} = \begin{pmatrix} 0 & \boldsymbol{R}^{12} \\ -\boldsymbol{R}^{12} & \boldsymbol{R}^{22} \end{pmatrix}, \tag{5-1}$$

with the Casimir reciprocity relations. The governing equations for the dual variables are coupled:

$$\dot{\boldsymbol{\alpha}} = \boldsymbol{R}^{12}.(\boldsymbol{B} - \text{Div}_R\,\mathcal{B}), \tag{5-2}$$

$$\dot{\boldsymbol{\beta}} = -\boldsymbol{R}^{12}.(\boldsymbol{A} - \text{Div}_R\,\mathcal{A}) + \boldsymbol{R}^{22}.(\boldsymbol{B} - \text{Div}_R\,\mathcal{B}). \tag{5-3}$$

Here the intrinsic part of the dissipation $\tilde{h}^{\text{intr}}$ is positive. Considering again the simple case with the free energy function $\overline{W}$ independent of $\nabla_R\boldsymbol{\beta}$ and the quadratic dependence of the free energy function with respect to the internal variable $\boldsymbol{\beta}$, we again have for the primary internal variable

$$\dot{\boldsymbol{\alpha}} = -b\boldsymbol{R}^{12}.\boldsymbol{\beta}. \tag{5-4}$$

Substituting from (5-4) into (5-3), we obtain the governing equation for the primary internal variable $\boldsymbol{\alpha}$:

$$\ddot{\boldsymbol{\alpha}} - \boldsymbol{R}^{22}.(\boldsymbol{R}^{12})^{-1}\dot{\boldsymbol{\alpha}} = (b\boldsymbol{R}^{12}.\boldsymbol{R}^{12}).\left(-\frac{\partial\overline{W}}{\partial\boldsymbol{\alpha}} + \text{Div}_R\,\frac{\partial\overline{W}}{\partial(\nabla_R\boldsymbol{\alpha})}\right), \tag{5-5}$$

which is a Cattaneo–Vernotte type hyperbolic equation [Joseph and Preziosi 1989]. As shown by [Berezovski et al. 2011c] on the example of one-dimensional thermoelasticity, the primary internal variable $\boldsymbol{\alpha}$ can be identified in this case with the microtemperature. In this context, it is understood as a fluctuation of the macrotemperature due to the influence of the existing microstructure. The governing equation (5-5) is coupled with the canonical equations (2-15), because the modified Eshelby tensor $\tilde{\boldsymbol{b}}$ and entropy flux $\tilde{\boldsymbol{S}}$ include contributions by internal variables. In fact, the energy conservation equation (2-15)$_2$ can be represented in the form

$$\theta\frac{\partial S}{\partial t}\bigg|_X + \nabla_R \cdot \tilde{\boldsymbol{Q}} = \tilde{h}^{\text{intr}}. \tag{5-6}$$

Due to definition of entropy (2-14)$_2$, its time derivative can be calculated as

$$\frac{\partial S}{\partial t}\bigg|_X = -\frac{\partial^2 W}{\partial t\,\partial\theta}\bigg|_X = -\frac{\partial^2 W}{\partial \boldsymbol{F}\,\partial\theta}\bigg|_X : \dot{\boldsymbol{F}} - \frac{\partial^2 W}{\partial\theta^2}\bigg|_X \dot{\theta},\tag{5-7}$$

because the internal variables are independent of temperature. Thus the energy balance equation (5-6) can be represented as

$$c\dot{\theta} + \nabla_R \cdot \tilde{\boldsymbol{Q}} = \theta\boldsymbol{M} : \dot{\boldsymbol{F}} + \tilde{h}^{\text{intr}},\tag{5-8}$$

with

$$c = -\theta\frac{\partial^2 W}{\partial\theta^2}\bigg|_X, \quad \boldsymbol{M} = \frac{\partial^2 W}{\partial\boldsymbol{F}\,\partial\theta}\bigg|_X.\tag{5-9}$$

The final form of the heat conduction equation follows from the definition of $\tilde{\boldsymbol{Q}}$ (2-21), the modified Fourier's law (2-24), and the expression for the intrinsic heat source (2-25):

$$c\dot{\theta} = \nabla_R^2\theta + \theta\boldsymbol{M} : \dot{\boldsymbol{F}} + \tilde{\mathscr{A}} : \dot{\boldsymbol{\alpha}} + \tilde{\mathscr{B}} : \dot{\boldsymbol{\beta}}.\tag{5-10}$$

This means that the heat conduction equation is still parabolic, but coupled with the stress field and internal variables. The equation of motion is the same as in the Piola–Kirchhoff formulation (2-4), but due to the definition of the stress in dual internal variable theory (2-14)$_1$, the stress tensor may contain both thermal and internal parts depending on the constitutive relation, which is not imposed yet. The coupling between the equation of motion, energy balance (5-10), and governing equation for internal variables (5-5) may induce a wave-like propagation of the macrotemperature even in the case of a parabolic equation for the macrotemperature [Berezovski et al. 2011c]. The process of wave interaction at the micro and macro levels can be described as follows. Microtemperature perturbations are induced by a macrodeformation due to the heterogeneity in the presence of a microstructure. These perturbations, propagating with finite speed, can induce, in turn, corresponding changes in the macrotemperature. At last, the induced changes in the macrotemperature affect the macrodeformations once more, and so on.

6. Conclusions and discussion

To sum up, we have shown that the influence of the internal structure of materials under external loading can be modeled by internal fields using the concept of dual internal variables. As one can see, dual internal variable theory is sufficiently general to comprise the micromorphic elasticity and the microtemperature influence in addition to classical single internal variable reaction-diffusion-type applications. Based on the canonical balance equations for material momentum and energy, the weakly nonlocal dual internal variable approach leads to governing equations for internal variables, which ensue directly from the dissipation inequality. The resulting governing equations are not limited by first-order reaction-diffusion equations, as is typical for single internal variable theory. The resulting hyperbolic governing equations provide a description of the interaction of waves at the macro and microlevels. In summary, the dual internal variable approach offers a unified description of dissipative and nondissipative internal processes in solids in the framework of continuum mechanics. In contrast to other theories, the following features should be emphasized:

- The governing equations for the internal variables result from the dissipation inequality and therefore are thermodynamically consistent.

- The governing equations for the internal variables are not restricted by first-order differential equations, that is, they may include second-order derivatives responsible for wave motion.

- The boundary conditions for the internal variables are determined by zero extra entropy flux at a boundary, which is a natural condition for internal variables.

At the same time, the presented theory is not a theory of everything. It is difficult to expect that the internal variable theory will describe, for example, volume double and couple forces that may appear in the general micromorphic case, because these are external by definition. This theory cannot handle any "internal" fields, which can be controlled at the boundaries. Such a field should be treated as an internal degree of freedom. In the corresponding case, the balance equations for the internal field should be postulated or derived separately.

The dual internal variable theory is presented in the general framework of the material formulation of continuum mechanics without specifying the small strain approximation. However, the example of nondissipative micromorphic elasticity is presented following the original Mindlin form with the small strain approximation to demonstrate explicitly the similarity of the structure of the governing equations for internal fields. The introduced microtemperature adheres to the micromorphic approach [Forest and Aifantis 2010]. It can be interpreted physically as fluctuations about the mean temperature. The microtemperature itself cannot produce essential influence, but its gradients can do that [Forest and Amestoy 2008]. The introduction of the microtemperature is accompanied by the modification of the Fourier law providing the coupling between the macrotemperature and its fluctuations. The modification of the Fourier law yields from the standard choice of the proportionality of the modified heat flux to the temperature gradient in order to satisfy the thermal part of the dissipation inequality.

As for any high-order theory, the application of the described dual internal variable theory to practical problems depends on the specification of the free energy function, which involves many material constants. There are several methods to determine the material constants, such as classical homogenization [Nemat-Nasser and Hori 1993] or the more recent asymptotic homogenization methods [Peerlings and Fleck 2004; Fish and Fan 2008; Pindera et al. 2009].

Chen and Lee [2003b] proposed an approach to determine the material constants for micromorphic elastic solids through the phonon dispersion relations obtained by atomistic calculations or experimental measurements. Later, Zeng et al. [2006] extended the approach to determine material constants in nonlocal micromorphic theory. A similar approach was proposed in [Maranganti and Sharma 2007; Jakata and Every 2008] to determine the coefficients of higher-order derivatives in strain-gradient elasticity. An attempt was made to solve the mathematically well-posed inverse problems for the determination of material parameters for micromorphic elasticity in [Janno and Engelbrecht 2011].

To be able to compare theoretical and experimental dispersion curves, the dispersive wave equations for elastic solids with internal structure were retrieved, analyzed, and unified in [Berezovski et al. 2011a]. It was also shown that for a medium consisting of more than one phase of microstructure, additional internal variables are necessary in order to describe the local deformation and yield more accurate dispersion curves [Huang and Sun 2008; Berezovski et al. 2010]. Clearly, efforts to understand the influence of internal fields on macrobehavior lead us to better understanding of the mechanics of materials.

References

[Berezovski et al. 2009a] A. Berezovski, J. Engelbrecht, and G. A. Maugin, "Internal variables and generalized continuum theories", pp. 149–158 in *IUTAM Symposium on Progress in the Theory and Numerics of Configurational Mechanics*, edited by P. Steinmann, IUTAM Bookser. **17**, Springer, Berlin, 2009.

[Berezovski et al. 2009b] A. Berezovski, J. Engelbrecht, and G. A. Maugin, "One-dimensional microstructure dynamics", pp. 21–28 in *Mechanics of Microstructured Solids*, edited by J.-F. Ganghoffer and F. Pastrone, Lecture Notes in Applied and Computational Mechanics **46**, Springer, Berlin, 2009.

[Berezovski et al. 2010] A. Berezovski, J. Engelbrecht, and T. Peets, "Multiscale modeling of microstructured solids", *Mech. Res. Comm.* **37**:6 (2010), 531–534.

[Berezovski et al. 2011a] A. Berezovski, J. Engelbrecht, and M. Berezovski, "Waves in microstructured solids: A unified viewpoint of modeling", *Acta Mech.* **220**:1-4 (2011), 349–363.

[Berezovski et al. 2011b] A. Berezovski, J. Engelbrecht, and G. A. Maugin, "Generalized thermomechanics with dual internal variables", *Arch. Appl. Mech.* **81**:2 (2011), 229–240.

[Berezovski et al. 2011c] A. Berezovski, J. Engelbrecht, and G. A. Maugin, "Thermoelasticity with dual internal variables", *J. Thermal Stresses* **34**:5-6 (2011), 413–430.

[Bridgman 1961] P. W. Bridgman, *The nature of thermodynamics*, Harvard University Press, Cambridge, MA, 1961.

[Capriz 1989] G. Capriz, *Continua with microstructure*, Springer Tracts in Natural Philosophy **35**, Springer, New York, 1989.

[Chen and Lee 2003a] Y. Chen and J. D. Lee, "Connecting molecular dynamics to micromorphic theory, II: Balance laws", *Physica A* **322**:1–4 (2003), 377–392.

[Chen and Lee 2003b] Y. Chen and J. D. Lee, "Determining material constants in micromorphic theory through phonon dispersion relations", *Int. J. Eng. Sci.* **41**:8 (2003), 871–886.

[Chen and Lee 2005] Y. Chen and J. D. Lee, "Atomistic formulation of a multiscale field theory for nano/micro solids", *Phil. Mag.* **85**:33–35 (2005), 4095–4126.

[Chen and Lee 2006] Y. Chen and J. D. Lee, "Conservation laws at nano/micro scales", *J. Mech. Mater. Struct.* **1**:4 (2006), 681–704.

[Coleman and Gurtin 1967] B. D. Coleman and M. E. Gurtin, "Thermodynamics with internal state variables", *J. Chem. Phys.* **47**:2 (1967), 597–613.

[Cross and Hohenberg 1993] M. C. Cross and P. C. Hohenberg, "Pattern formation outside of equilibrium", *Rev. Mod. Phys.* **65**:3 (1993), 851–1112.

[Eringen 1999] A. C. Eringen, *Microcontinuum field theories, I: Foundations and solids*, Springer, New York, 1999.

[Eringen and Suhubi 1964a] A. C. Eringen and E. S. Suhubi, "Nonlinear theory of micro-elastic solids, II", *Int. J. Eng. Sci.* **2**:4 (1964), 389–404.

[Eringen and Suhubi 1964b] A. C. Eringen and E. S. Suhubi, "Nonlinear theory of simple micro-elastic solids, I", *Int. J. Eng. Sci.* **2**:2 (1964), 189–203.

[Fish and Fan 2008] J. Fish and R. Fan, "Mathematical homogenization of nonperiodic heterogeneous media subjected to large deformation transient loading", *Int. J. Numer. Meth. Eng.* **76**:7 (2008), 1044–1064.

[Forest 2009] S. Forest, "Micromorphic approach for gradient elasticity, viscoplasticity, and damage", *J. Eng. Mech.* **135**:3 (2009), 117–131.

[Forest and Aifantis 2010] S. Forest and E. C. Aifantis, "Some links between recent gradient thermo-elasto-plasticity theories and the thermomechanics of generalized continua", *Int. J. Solids Struct.* **47**:25-26 (2010), 3367–3376.

[Forest and Amestoy 2008] S. Forest and M. Amestoy, "Hypertemperature in thermoelastic solids", *C. R. Mecanique* **336**:4 (2008), 347–353.

[Gonella et al. 2011] S. Gonella, M. S. Greene, and W. K. Liu, "Characterization of heterogeneous solids via wave methods in computational microelasticity", *J. Mech. Phys. Solids* **59**:5 (2011), 959–974.

[Grammenoudis and Tsakmakis 2009] P. Grammenoudis and C. Tsakmakis, "Micromorphic continuum, I: Strain and stress tensors and their associated rates", *Int. J. Non-Linear Mech.* **44**:9 (2009), 943–956.

[Gurtin 1996] M. E. Gurtin, "Generalized Ginzburg–Landau and Cahn–Hilliard equations based on a microforce balance", *Phys. D* **92**:3-4 (1996), 178–192.

[Hirschberger et al. 2007] C. B. Hirschberger, E. Kuhl, and P. Steinmann, "On deformational and configurational mechanics of micromorphic hyperelasticity–theory and computation", *Comput. Methods Appl. Mech. Engrg.* **196**:41-44 (2007), 4027–4044.

[Huang and Sun 2008] G. L. Huang and C. T. Sun, "A higher-order continuum model for elastic media with multiphased microstructure", *Mech. Adv. Mater. Struct.* **15**:8 (2008), 550–557.

[Jakata and Every 2008] K. Jakata and A. G. Every, "Determination of the dispersive elastic constants of the cubic crystals Ge, Si, GaAs, and InSb", *Phys. Rev. B* **77**:17 (2008), 174301.

[Janno and Engelbrecht 2011] J. Janno and J. Engelbrecht, *Microstructured materials: Inverse problems*, Springer, Berlin, 2011.

[Joseph and Preziosi 1989] D. D. Joseph and L. Preziosi, "Heat waves", *Rev. Modern Phys.* **61**:1 (1989), 41–73.

[Kirchner and Steinmann 2005] N. Kirchner and P. Steinmann, "A unifying treatise on variational principles for gradient and micromorphic continua", *Phil. Mag.* **85**:33–35 (2005), 3875–3895.

[Maranganti and Sharma 2007] R. Maranganti and P. Sharma, "A novel atomistic approach to determine strain-gradient elasticity constants", *J. Mech. Phys. Solids* **55**:9 (2007), 1823–1852.

[Mariano 2002] P. M. Mariano, "Multifield theories in mechanics of solids", *Adv. Appl. Mech.* **38** (2002), 1–93.

[Mariano and Stazi 2005] P. M. Mariano and F. L. Stazi, "Computational aspects of the mechanics of complex materials", *Arch. Comput. Methods Eng.* **12**:4 (2005), 391–478.

[Maugin 1990] G. A. Maugin, "Internal variables and dissipative structures", *J. Non-Equilib. Thermodyn.* **15**:2 (1990), 173–192.

[Maugin 1993] G. A. Maugin, *Material inhomogeneities in elasticity*, Applied Mathematics and Mathematical Computation **3**, Chapman & Hall, London, 1993.

[Maugin 1999] G. A. Maugin, *The thermomechanics of nonlinear irreversible behaviors: An introduction*, World Scientific, Singapore, 1999.

[Maugin 2006] G. A. Maugin, "On the thermomechanics of continuous media with diffusion and/or weak nonlocality", *Arch. Appl. Mech.* **75**:10-12 (2006), 723–738.

[Maugin 2011a] G. A. Maugin, *Configurational forces: Thermomechanics, physics, mathematics, and numerics*, Chapman & Hall, Boca Raton, FL, 2011.

[Maugin 2011b] G. A. Maugin, "A historical perspective of generalized continuum mechanics", pp. 3–19 in *Mechanics of generalized continua*, edited by H. Altenbach et al., Advanced Structured Materials **7**, Springer, Berlin, 2011.

[Maugin and Berezovski 2008] G. A. Maugin and A. Berezovski, "Introduction to the thermomechanics of configurational forces", *Atti dell'Accademia Peloritana dei Pericolanti* **86**:S1 (2008), 1–17.

[Maugin and Muschik 1994] G. A. Maugin and W. Muschik, "Thermodynamics with internal variables, I: General concepts", *J. Non-Equilib. Thermodyn.* **19**:3 (1994), 217–249.

[Mielke 2011] A. Mielke, "Formulation of thermoelastic dissipative material behavior using GENERIC", *Contin. Mech. Thermodyn.* **23**:3 (2011), 233–256.

[Mindlin 1964] R. D. Mindlin, "Micro-structure in linear elasticity", *Arch. Rational Mech. Anal.* **16** (1964), 51–78.

[Muschik 1990] W. Muschik, *Aspects of nonequilibrium thermodynamics: Six lectures on fundamentals and methods*, Series in Theoretical and Applied Mechanics **9**, World Scientific, Teaneck, NJ, 1990.

[Nemat-Nasser and Hori 1993] S. Nemat-Nasser and M. Hori, *Micromechanics: Overall properties of heterogeneous materials*, Elsevier, Amsterdam, 1993.

[Peerlings and Fleck 2004] R. H. J. Peerlings and N. A. Fleck, "Computational evaluation of strain gradient elasticity constants", *Int. J. Multiscale Comp. Eng.* **2**:4 (2004), 599–619.

[Pindera et al. 2009] M.-J. Pindera, H. Khatam, A. S. Drago, and Y. Bansal, "Micromechanics of spatially uniform heterogeneous media: A critical review and emerging approaches", *Composites Part B* **40**:5 (2009), 349–378.

[Rice 1971] J. R. Rice, "Inelastic constitutive relations for solids: An internal-variable theory and its application to metal plasticity", *J. Mech. Phys. Solids* **19**:6 (1971), 433–455.

[Truesdell 1984] C. Truesdell, *Rational thermodynamics*, 2nd ed., Springer, New York, 1984.

[Ván et al. 2008] P. Ván, A. Berezovski, and J. Engelbrecht, "Internal variables and dynamic degrees of freedom", *J. Non-Equilibrium Thermodyn.* **33**:3 (2008), 235–254.

[Wang and Lee 2010] X. Wang and J. D. Lee, "Micromorphic theory: A gateway to nano world", *Int. J. Smart Nano Mater.* **1**:2 (2010), 115–135.

[Zeng et al. 2006] X. Zeng, Y. Chen, and J. D. Lee, "Determining material constants in nonlocal micromorphic theory through phonon dispersion relations", *Int. J. Eng. Sci.* **44**:18-19 (2006), 1334–1345.

JÜRI ENGELBRECHT: je@ioc.ee
Centre for Nonlinear Studies, Institute of Cybernetics at Tallinn University of Technology, Akadeemia tee 21, 12618 Tallinn, Estonia

ARKADI BEREZOVSKI: Arkadi.Berezovski@cs.ioc.ee
Centre for Nonlinear Studies, Institute of Cybernetics at Tallinn University of Technology, Akadeemia tee 21, 12618 Tallinn, Estonia

msp

THE INVERSE PROBLEM OF SEISMIC FAULT DETERMINATION USING PART TIME MEASUREMENTS

HUY DUONG BUI, ANDREI CONSTANTINESCU AND HUBERT MAIGRE

This paper shows how a part time measurement T of the acceleration on the ground can be done to avoid the consideration of reflections of seismic waves on the Moho discontinuity. A bounded hemispherical solid is considered with radius greater than $c_p T$, but smaller than the Moho depth, in order to get the null boundary conditions on the hemispherical surface. The inverse problem to determine the fault plane and the time-dependent fault geometry is well-defined and solved in closed form by the reciprocity gap functional method. Two formulae are established for the components of the fault slip, which involve time and spatial Fourier transforms of some quantities related to the reciprocity gap functional. A forward initial boundary value problem in elastodynamics enables us to get the internal wave field, the shear stress as well as the normal stress on the fault. This full solution would be useful for understanding the friction mechanism on the fault during a real strike, whenever the initial uniform tectonic stresses are known from geophysical considerations.

1. Introduction

The determination of the characteristics of an earthquake from measurements of the acceleration[1]

$$\partial_t \partial_t (\boldsymbol{u})(\boldsymbol{x}, t)$$

on the stress-free ground G is one important seismic inverse problem in geophysics [Aki and Richards 1980]. Mathematically, one has to determine the fault history $\Sigma(t)$, its plane and geometry, the slip and the shear stress released from ground measurements, without considering the physical friction mechanisms at the origin of the strike. Physical models of earthquakes are not considered in this paper (for a review of such models, see [Madariaga et al. 2000]). We need a model which is generally an elastic one, bounded or unbounded and some assumption on the seismic process. Due to tectonic stresses, it is generally assumed that the release of the shear stress, completely or incompletely, corresponds to some unknown amount of tangential slip $[[\boldsymbol{u}]]$. In some cases, the fault plane is already known, for example when the slip occurs in existing plane (San Andreas fault) or when the slip plane emerges on the ground. The analysis of data in view of the inversion is simpler in these cases.

Before introducing the aim of the paper, we briefly review some methods of earthquake inverse problems in order to see what can be improved.

In simple cases listed above, some *forward* models in infinite elastic medium can give good prediction of the acceleration signals to be compared to measurements [Campillo 1983; Cochard and Rice 1997; Vallée 2003; Peyrat et al. 2004; Lapusta et al. 2001]. True *inverse* problems are divided into two kinds.

Keywords: earthquake, inverse problem, reciprocity gap, elastic wave, fault history.

[1]We use the notation $\partial_t = \partial/\partial t$.

In the first one, a feedback between forward models of faults and observations is made to determine the optimal solution [Das and Kostrov 1990; Das and Suhadolc 1996]. Most papers on inverse problems in seismology are based on this optimization method (or geometry control method). The method has some advantage, since it is based on numerical approaches for which powerful finite element software exist. However, there are also some limitations because of the ill-posedness of inverse problems, for which analytical solutions are missing. According to Das and Suhadolc [1996], the least-squares functional (for fitting the computed acceleration from a guess fault S with the measured one a)

$$J(S(\tau)) = \|\partial_\tau \partial_\tau \boldsymbol{u} - \boldsymbol{a}\|_u^2 + \alpha \|\partial_\tau \boldsymbol{u}\|_{\partial S}^2, \tag{1}$$

$$\Sigma(t) = \arg \min_{S(\tau)} J(S(\tau)), \tag{2}$$

with possible additional regularizing term with $\alpha > 0$ to bound the velocity of the fault front ∂S, has a flat minimum in the $\boldsymbol{u}$ space.

The same authors give a clear statement on such an optimal solution: "*Even if the fitting of data seems to be quite good, the faulting process is poorly reproduced, so that in the real case, it would be difficult to know when one has obtained a correct solution*."

In the second kind of method, introduced by the authors and their colleagues for many kinds of inverse problems in solid and fracture mechanics, ranging from elastostatics [Andrieux et al. 1999] to heat diffusion in solids, acoustics and elastodynamics, with the earthquake-like analysis (see the reviews in [Bui 2006]), the original *nonlinear* inverse problems with respect to the unknowns considered *together*, is put in a special variational form which enables us to determine the unknowns *separately* in a certain order: first the normal to the fault plane N, then the fault plane and finally the fault geometry. Each subproblem is characterized by the use of a specific adjoint field. More importantly, the subvariational equations are *linear* or quasilinear (this word will be explained in Section 3) with respect to the considered unknowns. Therefore, the original nonlinear inverse problem is decomposed into *successive linear* inverse problems which are solved analytically in a certain order.

In previous works, models have been worked out for some earthquake inverse problems, which differ by physical considerations and methods of solution. First, to avoid reflection of waves in the underground discontinuities, by considering the near field only and assuming certain shape of the source, one can obtain signals to be compared to observations [Campillo 1983]. An earthquake *like* model assumes that the bounded solid is entirely elastic and that data are available in its whole boundary. Such a model does not apply to real earthquake, because data cannot be obtained in the underground, but has the advantage of giving an exact solution to the fault history $\Sigma(t)$ [Bui et al. 2005]. The second approximate model which also determines the full history of the fault considers a semi infinite elastic medium, for which it is necessary to specify the far-field radiated by the earthquake [Vallée 2003], which can be estimated approximately. Both models give the entire fault history described by the displacement jump across the fault $[\![\boldsymbol{u}]\!] = \boldsymbol{u}^+ - \boldsymbol{u}^-$, a function of $(\boldsymbol{x}, t)$ and suppose an infinite duration of the acceleration measurement $\boldsymbol{a}$. The method is based on the use of adjoint fields $\boldsymbol{v}(\boldsymbol{x}, t)$ to establish a variational equation for the current discontinuous elastodynamic field $\boldsymbol{u}(\boldsymbol{x}, t)$ called *reciprocity gap functional* (Section 2).

Regarding the actual seismic process, the physical models above are not realistic. In a real process, waves travel in the crust, then reflect on and diffract through the Mohorovičić discontinuity; see Figure 1. Suppose that the crust has a depth of 60 km, P-waves from a point source near the ground will take
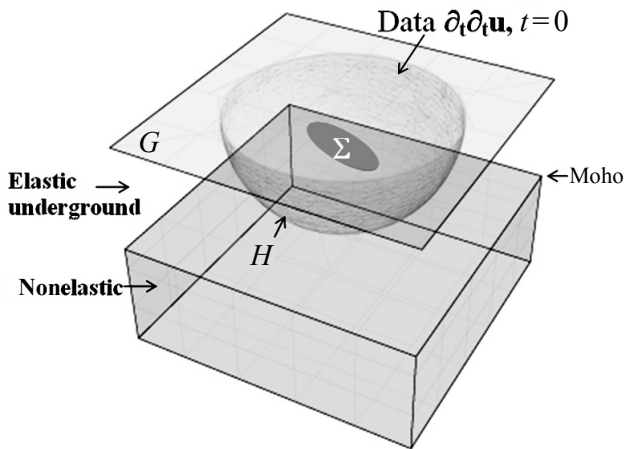
Figure 1. The seismic model. Part time measurement of the acceleration on the stress-free ground G, inside a circular area of radius $c_p T$ during the time T; H is a hemispherical surface tangent to the Mohorovičić discontinuity containing a small fault Σ near the ground. All seismic signals are inside the elastic hemisphere, for $0 < t < T$.

about 10 s to reach the Moho discontinuity. Beyond this time, after 20 s, reflected waves arrive on the ground so that measured data correspond to so many sources, the fault and the Moho interface. So the question is: what part of the signal can be used for the reconstruction of the wave due to the strike only? A two-dimensional simulation of a sudden release of shear stress over a small fault located near the free surface is shown in Figure 2. It is observed that complex waves, P-waves, S-waves and intermediate waves are generated, as predicted [Hudson 1980]. Numerical solutions are much more complex when the reflection of waves occurs at the interface. Ghost solutions are observed near the ground which explain the difficulty encountered in numerical optimization methods [Das and Suhadolc 1996].
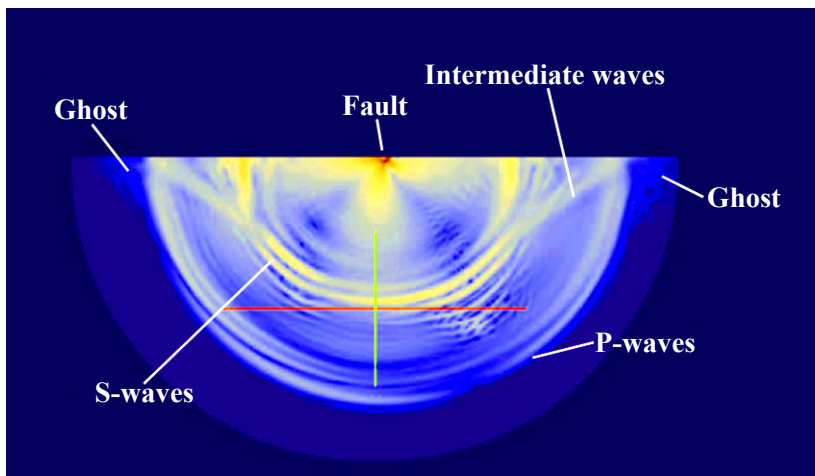

Figure 2. A two-dimensional numerical simulation of an earthquake by the sudden release of shear stress on a small fault near the ground, showing ghost solution.

Now, consider a finite duration T such that P-waves from the source do not reach the Moho discontinuity, the depth of which is more or less known from geophysical considerations. Take any volume or a hemisphere with boundary H in the elastic region above the Moho discontinuity, with radius larger than $c_p T$ but smaller than the Moho discontinuity depth, containing the perturbation field. We then know the boundary fields on H:

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{and} \quad \boldsymbol{T} = \boldsymbol{\sigma}[\boldsymbol{u}].\boldsymbol{n} = \boldsymbol{0} \quad \text{on } H \times [0, T]. \tag{3}$$

The aim of the paper is to show how to solve analytically the nonlinear inverse problem, using *partial* information on the ground acceleration, conditions of (3) and the *a priori* knowledge on the planar character of slip fault. No uncertainties of data are considered in this paper. The smaller the time T, the shorter the history of the strike which can be recovered. We shall assume first that acceleration data are due to some rapid release of shear stresses on the moving fault faces. The assumption is based on the experimental evidence of shear cracks faster than the shear wave speed which has been reported in [Rosakis et al. 1999], $c_s < V \leq \sqrt{2}c_s$. Actually, the only assumed boundary condition on the fault is that it is a planar shear fault. We do not know how the shear stresses are released, or what the fault speed is. We shall show that a rigorous mathematical analysis will solve the inverse problem in closed form, giving the fault plane in Section 3. Section 4 discusses a proposal to determine the fault history $\Sigma(t)$ and establishes two formulae giving the components of the displacement jump explicitly. The solution allows the determination of the release stress on the fault via a forward initial and boundary value problem in elastodynamics. A complete solution of the inverse problem, giving both the shear stress and the slip components on the fault, would be helpful for understanding the friction law mechanism. This paper examines the mathematical aspects of such an inverse problem.

2. The reciprocity gap functional

Since waves travel in the crust of finite depth we shall consider a *finite* duration T for the analysis of the inverse problem. The physical model is based on the following assumptions. At initial time, a sudden shear slip or a release of (unknown) shear stress over the fault faces $\Sigma^{\pm}$ generates elastic waves which can be measured on the stress-free ground. No stresses are applied to the external boundary of the solid. Such conditions have been considered in our earthquake-like model [Bui et al. 2005]. No reflections of waves occur yet at the Moho discontinuty during the time T and (3) holds. No other assumptions will be needed except that measurements of the acceleration (a time-dependent vector field) on the ground are done on the perturbed area up to T.

For the self-consistency of the paper, we reestablish the general expression of the reciprocity gap functional R and its particular expression for the present study for a finite duration time of measurements. Let Ω be the elastic region before the quake, with boundary $\partial\Omega = G \cup H$. After the quake, the sound region is $\Omega' = \Omega \setminus \Sigma$, with boundary $\partial\Omega' = G \cup H \cup \Sigma^{\pm}$. We shall denote the region occupied by the current wave by Ω''. Its boundary is the convex hull generated by spherical P-waves of radius $c_p T$ emitted by all points of the fault.

The current elastodynamic displacement field $\boldsymbol{u}(\boldsymbol{x}, t)$ due to either the release of shear stress on $\Sigma^{\pm}$ or the imposition of the tangential jump $[\![\boldsymbol{u}]\!]$ on a plane of normal $\boldsymbol{N}$ satisfies the equation of motion, initial conditions and boundary conditions (L is the isotropic elastic moduli tensor, ρ is the density)

as follows:

$$\text{div}(L : \text{grad}(\boldsymbol{u})) - \rho \partial_t \partial_t \boldsymbol{u} = \boldsymbol{0} \quad \text{on } (\Omega \setminus \Sigma) \times [0, T], \tag{4}$$

$$\boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{0}, \quad \partial_t \boldsymbol{u}(\boldsymbol{x}, 0) = \boldsymbol{0} \quad \text{in } (\Omega \setminus \Sigma), \tag{5}$$

$$\partial_t \partial_t \boldsymbol{u} = \boldsymbol{a} \quad \text{on } G \times [0, T], \tag{6}$$

$$\boldsymbol{T} = \sigma[\boldsymbol{u}].\boldsymbol{n} = \boldsymbol{0} \quad \text{on } G \times [0, T], \tag{7}$$

$$\boldsymbol{T} = \sigma[\boldsymbol{u}].\boldsymbol{n} = \boldsymbol{0} \quad \text{on } H \times [0, T], \tag{8}$$

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } H \times [0, T], \tag{9}$$

$$[[\boldsymbol{u}]].\boldsymbol{n} = \boldsymbol{0} \quad \text{on } \Sigma^{\pm}. \tag{10}$$

The adjoint field $\boldsymbol{v}$ is continuous throughout Ω and satisfies the wave equation and the final conditions

$$\text{div}(L : \text{grad}(\boldsymbol{v})) - \rho \partial_t \partial_t \boldsymbol{v} = \boldsymbol{0} \quad \text{in } \Omega \times [0, T], \tag{11}$$

$$\boldsymbol{v}(\boldsymbol{x}, t = T) = \boldsymbol{0}, \quad \partial_t \boldsymbol{v}(\boldsymbol{x}, t = T) = \boldsymbol{0} \quad \text{in } \Omega. \tag{12}$$

Multiply (4) by $\boldsymbol{v}$, (11) by $\boldsymbol{u}$, subtract the result, then integrate the final result in $\Omega' \times [0, T]$ where $\boldsymbol{N}$ is the normal to the fault ($\boldsymbol{N}$ is also the normal to Σ^-)

$$\int_0^T \int_\Sigma [[\boldsymbol{u}]].\sigma[\boldsymbol{v}].\boldsymbol{N} \, dt \, dS + \int_0^T \int_{\Omega'} \rho \frac{\partial}{\partial t}(\boldsymbol{u}\partial_t \boldsymbol{v} - \boldsymbol{v}\partial_t \boldsymbol{u}) \, dt \, dV = R(\text{data}, \boldsymbol{v}), \tag{13}$$

where

$$R(\text{data}, \boldsymbol{v}) \overset{\text{def}}{=} \int_0^T \int_G \boldsymbol{u}.\sigma[\boldsymbol{v}].\boldsymbol{n} \, dt \, dS. \tag{14}$$

We have

$$\int_0^T \int_{\Omega'} \rho \frac{\partial}{\partial t}(\boldsymbol{u}.\partial_t \boldsymbol{v} - \boldsymbol{v}.\partial_t \boldsymbol{u}) \, dt \, dV = \int_{\Omega'} dV \rho(\boldsymbol{u}.\partial_t \boldsymbol{v} - \boldsymbol{v}.\partial_t \boldsymbol{u}) \, |_0^T. \tag{15}$$

Using the initial conditions on $\boldsymbol{u}$ and the final conditions on $\boldsymbol{v}$ we obtain

$$\int_{\Omega'} dV \rho(\boldsymbol{u}.\partial_t \boldsymbol{v} - \boldsymbol{v}.\partial_t \boldsymbol{u}) \, |_0^T = 0. \tag{16}$$

Finally the variational equation for the inverse problem is given by

$$\int_0^T \int_\Sigma [[\boldsymbol{u}]].\sigma[\boldsymbol{v}].\boldsymbol{N} \, dt \, dS = R(\text{data}, \boldsymbol{v}) \quad \text{for all } \boldsymbol{v}. \tag{17}$$

$R(\text{data}, \boldsymbol{v})$ is called the *reciprocity gap functional*, which measures the *symmetry loss* between $\boldsymbol{u}$ and $\boldsymbol{v}$. It depends on the acceleration data $\boldsymbol{a}$ on the ground. By a time integration of $\boldsymbol{a}$, twice, we get the surface field $\boldsymbol{u}(\boldsymbol{x}, t)$ over G for the right side of (14).

As a matter of fact, the origin of time has been taken at the origin of the hidden source strike. Recorded acceleration signals start when the first strike arrives at the ground with some delay t_1. The true origin of time takes account of this delay, equal to $t_1 = (c_p/c_s - 1)\Delta$ with the difference Δ between the arrival times of the first S-wave and P-wave.

3. Adjoint fields to determine the fault plane

The adjoint fields $v(x, t)$ to be considered will depend on parameters (scalar, vector, tensor). They are either plane S-shear waves or plane P-waves, with the moving front Γ propagating in the p direction, with respective velocity $c = c_s$ or $c = c_p$. There is a smooth transition of $v(x, t)$ from the zero value behind the wave to the one in front of the wave, in a narrow band of width $w > 0$, $w \to 0$. Adjoint P-waves are discussed later. Adjoint S-waves polarized in the direction k are proportional to the regularized (smooth) Heaviside function $Y^w(x.p/c - t - \tau)$, where τ is a parameter fixing the initial position of the wave front, denoted by Γ^τ, w is a positive number tending to zero ($Y(y) = 1$ for $y > w$ and $Y(y) = 0$ for $y < 0$)

$$v(x, t; \tau, p) = Y^{+0}(x.p/c_s - t - \tau)k. \tag{18}$$

They satisfy the homogeneous wave equation in $(\Omega \setminus \Gamma) \times [0, T]$ and depend on the scalar τ and vectors p and do not have discontinuities across the fault. The adjoint stress in the band tends to an impulse as the width $w > 0$ tends to zero. R becomes the *instantaneous reciprocity gap* introduced in [Bui et al. 2004].

The fault plane. The choice of the propagating direction p for adjoint S-waves must satisfy (16). Figures 3 and 4(b) show that adjoint S-waves can be considered near the top A of the fault only, with fronts more or less inclined with the angle α with respect to the horizontal plane, with the maximum angle α_{max} which depends on the depth d of point A. The existence of a maximum angle is explained by the fact that an S-wave is slower than a P-wave and that possibly the final condition (12) cannot be satisfied for large angle. If the front has a steep angle, the S-wave at time T has a nonempty intersection with the domain Ω'', so (16) is not satisfied:

$$0 \le \alpha < \alpha_{max} = a \sin \frac{c_s}{c_p} - a \sin \frac{d}{T c_p}. \tag{19}$$
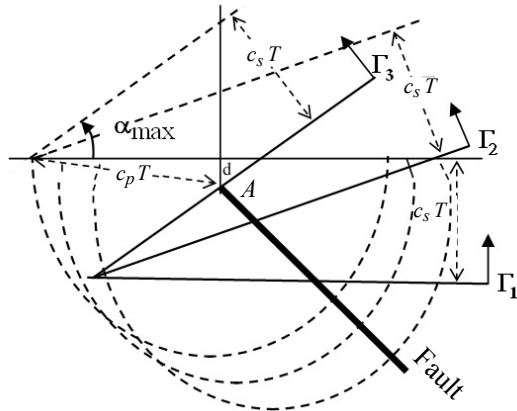

Figure 3. Adjoint S-waves propagating upwards. The angles between the propagating vectors and the vertical depend on the depth d of the nearest point A to the ground. The maximun angle is for the front Γ_3, the minimum (zero value) for the front Γ_1.
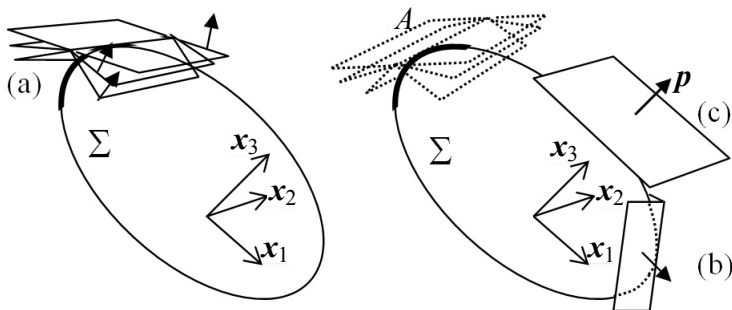
Figure 4. Left: adjoint S-waves propagating upwards (a) with different p under some angle constraints for determining the tangents to the fault front near point A. Right: P-waves of arbitrary propagation vectors (b) can also be used for determining the fault plane. A P-wave parallel to the fault plane (c) is used to determine the fault geometry.

Adjoint P-waves of arbitrary propagation vectors p can also be considered for determining the fault plane; see label (b) in Figure 4. Once the fault plane has been determined, a particular wave parallel to the fault can be used for determining the geometry; see label (c) in the same figure.

The apex of point A can be determined by geometrical considerations of the first waves reaching the ground. It can also be determined directly by horizontal adjoint waves of type Γ_1 by changing τ. First choose τ so that the fault is entirely behind the wave front, thus $\sigma[v] = 0$ on the fault or the integral in the left side of (17) vanishes, hence $R = 0$. Then change τ so that $R \neq 0$. The transition of R from zero (for large τ) to a non zero value indicates that the adjoint wave front is tangent to the fault boundary.

Let us consider now a wave front of the type Γ_2 of Figure 3, the initial position of which is near to point A. By changing τ we change its initial position and by changing p we change its orientation, also to get the transition from zero to a nonzero value of R revealing the contact point and the tangents to the fault front $\partial\Sigma$. Therefore by trials and errors with this *zero-crossing* method, we determine a small three-dimensional curve of the fault front near the point A. This is sufficient to get the entire fault plane.

Remark that we do not make the assumption of a global convex shape of the fault plane, but only a locally convex front; see Figure 4. The zero-crossing method is the same as the one considered for a line crack in 2 dimensions [Bui et al. 2004], which is illustrated in Figure 5 where the zero-crossing has been appreciated with a small number. Having determined the fault plane, we take the axes Ox_1x_2 on the fault plane and $N = e^3$.

In our previous method to determine the normal N, we considered an adjoint field such that the stress $\sigma[v]$ is constant everywhere, so that (17) can be written as

$$N.\sigma[v].\left(\int_0^{-\tau} \int_\Sigma [[u]] \, dt \, dS \right) = R(\text{data}, v) \quad \text{for all } v,$$

which is quasilinear in N, because Σ still depends on N. However, using an argument given in [Andrieux et al. 1999] we can show that two independent data are necessary and sufficient to determine the direction of the normal which is linked to the eigenvectors of the tensor Z^{ij} of components $R(v^{ij})$, without knowing Σ and $[[u]]$. Unfortunately, this analytical method is useless for the seismic inverse problem because the seismic event is unique.
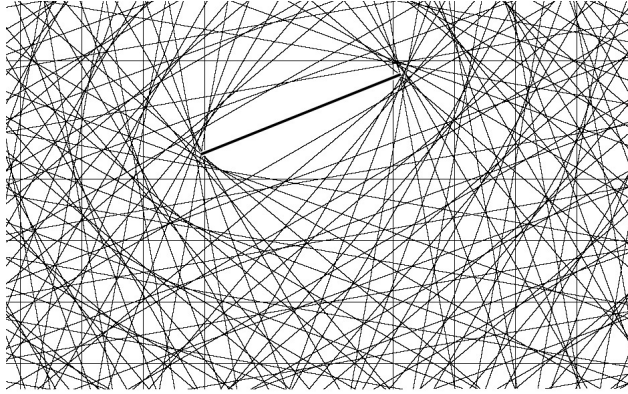
Figure 5. Instantaneous adjoint S-waves for determining the convex hull of a line crack
[Bui et al. 2004].

4. The partial fault history

We need another kind of adjoint P-waves and S-waves, which allows the components $[[u_1]]$, $[[u_2]]$ to be calculated separately, but under some additional assumptions. Moreover, new adjoint waves allow the determination of the partial fault history $\Sigma(t)$ for $0 < t < T$.

The reciprocity gap functional is rewritten below

$$\int_0^T \int_\Sigma [[\boldsymbol{u}]].\sigma[\boldsymbol{v}].\boldsymbol{N} \, dt \, dS + \int_{\Omega'} dV[\rho(\boldsymbol{u}.\partial_t \boldsymbol{v} - \boldsymbol{v}.\partial_t \boldsymbol{u})] \mid_0^T = R(\text{data}, \boldsymbol{v}) \quad \text{for all } \boldsymbol{v}. \tag{20}$$

The additional assumption. In the second integral, denoted by K, the term for $t = 0$ already vanishes because of the initial conditions on $\boldsymbol{u}$. We make the assumption that second term for $t = T$ can be omitted for a particular adjoint field. More precisely, for fixed T and for any given positive number ϵ, *as small as* we want, we can find a particular adjoint field such that

$$K = \int_{\Omega'} dV[\rho(\boldsymbol{u}.\partial_t \boldsymbol{v} - \boldsymbol{v}.\partial_t \boldsymbol{u})] \mid_0^T = O(\epsilon) \tag{21}$$

instead of $K = 0$, (16). The additional assumption is based on two arguments: firstly a natural and small damping of the medium exists so that the wave equation is only an approximated one,[2] secondly adjoint fields $\boldsymbol{v}$ having a good behavior at large time (rapid decay with time) can be easily found.

We study the geometry of the moving fault $\Sigma(t)$ with the coordinates attached to the fault plane $x_3 = 0$. Consider functions F depending on parameters ($\boldsymbol{s} \in \mathbb{R}^2$, $q \in \mathbb{R}$), satisfying the scalar wave equation for $0 < t < T$ (and $F = 0$ otherwise) of the form

$$F(\boldsymbol{x}, t; \boldsymbol{s}, q, c) = \exp((iq - k)t) \exp(i\boldsymbol{s}.\boldsymbol{x}) \exp\left(x_3 \left[\|s\|^2 + \frac{(iq - k)^2}{c^2}\right]^{1/2}\right), \tag{22}$$

---

[2]To take account of a small damping, a term such as $-\eta \partial_t \boldsymbol{u}$ ($\eta > 0$) can be added to the left side of (4) and the corresponding antidamping term $+\eta \partial_t \boldsymbol{v}$ to (11), the assumption becomes $\int_{\Omega'} dV[\rho(\boldsymbol{u}.\partial_t \boldsymbol{v} - \boldsymbol{v}.\partial_t \boldsymbol{u}) - \eta \boldsymbol{u}.\boldsymbol{v}] \mid^{t=T} = O(\epsilon)$.

with positive $k > 0$ chosen so that $\exp(-kT) = \epsilon$ and $c = c_p$ for P-waves and $c = c_s$ for S-waves.[3] Adjoint waves that are proportional to F have a good behavior at large time T as $\exp(-kT)$. The term $\exp(x_3)\|s\|^2$ is finite in the bounded medium. Two adjoint waves are considered, one for P-waves and the other for S-waves:

$$\boldsymbol{v}^{(p)} = \operatorname{grad} F(\boldsymbol{x}, t; \boldsymbol{s}, q, c_p), \quad \boldsymbol{v}^{(s)} = \operatorname{curl} F(\boldsymbol{x}, t; \boldsymbol{s}, q, c_s)\boldsymbol{e}^3. \tag{23}$$

The components of the slip. Equation (17) gives two linear systems of equations, with $x_3 = 0$ on the left sides of the equations, for $\boldsymbol{D} \stackrel{\text{def}}{=} [[\boldsymbol{u}]]$:

$$\int_{x_3=0} \int_0^T [is_1 D_1 + is_2 D_2] \exp((iq - k)t) \exp(i\boldsymbol{s}.\boldsymbol{x})\, dt\, dS$$
$$= \frac{1}{2\mu}\left[\|\boldsymbol{s}\|^2 + \frac{(iq-k)^2}{c_p^2}\right]^{-1/2} R(\text{data}, \boldsymbol{v}^{(p)}(\boldsymbol{s}, q, c_p)) \stackrel{\text{def}}{=} iG^{(1)}(\text{data}, \boldsymbol{s}, q, c_p),$$

$$\int_{x_3=0} \int_0^T [is_2 D_1 - is_1 D_2] \exp((iq - k)t) \exp(i\boldsymbol{s}.\boldsymbol{x})\, dt\, dS$$
$$= \frac{1}{\mu}\left[\|\boldsymbol{s}\|^2 + \frac{(iq - k)^2}{c_s^2}\right]^{-1/2} R(\text{data}, \boldsymbol{v}^{(s)}(\boldsymbol{s}, q, c_s)) \stackrel{\text{def}}{=} iG^{(2)}(\text{data}, \boldsymbol{s}, q, c_s).$$

The equations defines $G^{(1)}$ and $G^{(2)}$ in terms of $R(\text{data}, \boldsymbol{v}^{(p)})$ and $R(\text{data}, \boldsymbol{v}^{(s)})$, respectively. We see that the combinations $(s_1 G^{(1)} + s_2 G^{(2)})$ and $(s_2 G^{(1)} - s_1 G^{(2)})$ are proportional to $\|\boldsymbol{s}\|^2$, so we can divide them by $\|\boldsymbol{s}\|^2$. Now define function $E(t) = \exp(-kt)$ for $0 < t < T$ and $E(t) = 0$ otherwise. It follows that the double integral can be extended to the whole space-time. We obtain

$$\int_{x_3=0} \int_{-\infty}^{+\infty} D_1(\boldsymbol{x}, t) E(t) \exp(iqt) \exp(i\boldsymbol{s}.\boldsymbol{x})\, dt\, dS = (s_1 G^{(1)} + s_2 G^{(2)})/\|\boldsymbol{s}\|^2, \tag{24}$$

$$\int_{x_3=0} \int_{-\infty}^{+\infty} D_2(\boldsymbol{x}, t) E(t) \exp(iqt) \exp(i\boldsymbol{s}.\boldsymbol{x})\, dt\, dS = (s_2 G^{(1)} - s_1 G^{(2)})/\|\boldsymbol{s}\|^2. \tag{25}$$

Since $\boldsymbol{s} \in \mathbb{R}^2$ and $q \in \mathbb{R}$, the left sides of (24) and (25) are the double time Fourier transform F_t and spatial Fourier transform F_x of $D_1(\boldsymbol{x}, t)E(t)$ and $D_2(\boldsymbol{x}, t)E(t)$ respectively. Their right sides are known functions of data and $\boldsymbol{s}, q$.

Based on our previous works, which considered adjoint waves of the same exponential type, we know that the reciprocity gap functional is the spatial Fourier transform of a *compact support* function. For a similar inverse problem, the proof (based on the Wiener–Paley theorem) can be found in [Ben Abda and Bui 2003].

Finally, by double inverse Fourier transforms in the image spaces $(\boldsymbol{s}, q)$, we obtain each component of the displacement jump and the fault history given by its support function, $\Sigma = \operatorname{supp}([[u_1]]) \cup \operatorname{supp}([[u_2]])$, for $\epsilon = \exp(-kT) \to 0$:

$$[[u_1]](\boldsymbol{x}, t)E(t) = F_q^{-1} F_s^{-1} \frac{s_1 G^{(1)} + s_2 G^{(2)}}{\|\boldsymbol{s}\|^2}, \quad [[u_2]](\boldsymbol{x}, t)E(t) = F_q^{-1} F_s^{-1} \frac{s_2 G^{(1)} - s_1 G^{(2)}}{\|\boldsymbol{s}\|^2}. \tag{26}$$

---

[3]In the case of damping, we take

$$F(\boldsymbol{x}, t; \boldsymbol{s}, q, c) = \exp((iq - k)t) \exp(i\boldsymbol{s}.\boldsymbol{x}) \exp\left(x_3[\|\boldsymbol{s}\|^2 + \frac{(iq-k)^2}{c^2} - \eta(iq - k)]^{1/2}\right).$$

The released stresses. Knowing the displacement jump on the fault, the acceleration on the ground, the null displacement on H, with the initial conditions, we are left with a well-posed forward problem of elastodynamics which determines the displacement field inside the volume Ω' and the release stresses on the fault.

The knowledge of both the shear stress and the slip on the fault, could be helpful for understanding the friction law mechanism. The solution of the inverse problem and the forward one give the stress field $\sigma[u]$, in particular the normal stress σ_{33} and the shear stress on the fault σ_{3m}, along the direction m of the fault slip. These stress components are only variations of the stress in the strike. The actual stresses to be considered in the friction law are the sum of the released stress $\sigma[u]$ and the tectonic one Σ^{tect} provided by geophysical considerations only.

Remarks. 1. For an infinite medium or infinite time measurement $c_p T = \infty$, we have $K = 0$. The solution is the same given by (26).

2. For an explicit solution $\partial_t \partial_t u$ and exact data a in the residual J of (1), the numerical calculation of the slip via fast Fourier transforms in space and time, followed by the calculation of the theoretical acceleration on the ground via the elastodynamic solution, always involve errors. Therefore the residual J does not vanish. This behavior is also related to the remark of Das and Suhadolc about the flat minimum of the residual (quoted on page 998).

5. Conclusions

To avoid reflections of waves in the Moho discontinuity, one generally considers either near-fields or infinite medium for studying earthquake inverse problems. We have shown that these assumptions are not necessary. Firstly, we can consider a bounded solid above the Moho discontinuity. Secondly, we have shown that a part time T of measurement of the acceleration on the ground, before seismic waves reach the boundary of the solid considered, leads to a rigorous mathematical analysis of the equations and provides the solution to the inverse problem: the fault plane and its time-dependent geometry are determined. We establish the formulae giving the components of the fault slip in terms of the ground data during the time T. Having solved the inverse problem, a forward initial and boundary value problem in elastodynamics can provide the field solution and consequently the shear stress released on the fault.

The explicit solution of the earthquake inverse problem, obtained by the reciprocity gap method, provides both the slip field and the shear stress (and also the normal stress) on the fault and would be useful to understand the seismic mechanisms, particularly to find the friction law on the active fault.

Numerical applications of the proposed method are beyond the scope of this paper and will appear in a forthcoming paper.

References

[Aki and Richards 1980] K. Aki and P. Richards, *Qauntitative seismology, theory and methods*, W.H. Freeman and Cie, New York, 1980.

[Andrieux et al. 1999] S. Andrieux, A. Ben Abda, and H. D. Bui, "Reciprocity principle and crack identification", *Inverse Problems* **15**:1 (1999), 59–65.

[Ben Abda and Bui 2003] A. Ben Abda and H. D. Bui, "Planar crack identification for the transient heat equation", *J. Inverse Ill-Posed Probl.* **11**:1 (2003), 27–31.

[Bui 2006] H. D. Bui, *Fracture mechanics: inverse problems and solutions*, Solid Mechanics and Its Applications **139**, Springer, 2006.

[Bui et al. 2004] H. D. Bui, A. Constantinescu, and H. Maigre, "Numerical identification of linear cracks in 2D elastodynamics using the instantaneous reciprocity gap", *Inverse Problems* **20**:4 (2004), 993–1001.

[Bui et al. 2005] H. D. Bui, A. Constantinescu, and H. Maigre, "An exact inversion formula from determining a planar fault from boundary measurements", *J. Inverse Ill-Posed Probl.* **13**:3-6 (2005), 553–565.

[Campillo 1983] M. Campillo, "Numerical evaluation of the near-field high-frequency radiation from quasi-dynamic circular fault", *Bulletin Seismological Society of America* **73** (1983), 723–734.

[Cochard and Rice 1997] A. Cochard and J. R. Rice, "A spectral method for numerical elastodynamic fracture analysis without spatial replication of the rupture event", *J. Mech. Phys. Solids* **45**:8 (1997), 1395–1418.

[Das and Kostrov 1990] S. P. Das and B. V. Kostrov, "Inversion for seismic slip rate and distribution with stabilizing constraints. Applications to the 1986 Andreanov Islands earthquake", *J. Geophys. Res.* **95** (1990), 6899–6913.

[Das and Suhadolc 1996] S. Das and P. Suhadolc, "On the inverse problem for earthquake rupture. The Haskell-type source model", *Journal of Geophysical Research* **101**:B3 (1996), 5725–5738.

[Hudson 1980] J. A. Hudson, *The excitation and propagation of elastic waves*, Cambridge Monographs on Mechanics and Applied Mathematics, Cambridge University Press, 1980.

[Lapusta et al. 2001] N. Lapusta, J. Rice, Y. Ben-Zion, and G. Zheng, "Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction", *Geophys. Res.* **105** (2001), 23765–23789.

[Madariaga et al. 2000] R. Madariaga, S. Peyrat, and K. B. Olsen, "Rupture dynamics in 3D. A review", pp. 89–110 in *Problems in Geophysics for the New Millennium*, edited by E. Boschi et al., Editrice Compositori, Italy, 2000.

[Peyrat et al. 2004] S. Peyrat, K. Olsen, and R. Madariaga, "Which dynamic rupture parameters can be estimated from strong ground motion and geodesic data?", *Pure and Applied Geophysics* **161** (2004), 2155–2169.

[Rosakis et al. 1999] A. J. Rosakis, O. Samudrala, and D. Coker, "Cracks facter than the shear wave speed", *Sciences* **284**:5418 (1999), 1337–1340.

[Vallée 2003] M. Vallée, *Kinematic analysis of the earthquake with far-fields. Methods and solution*, Ph.D. thesis, Université de Grenoble, 2003.

HUY DUONG BUI: hdbui37@yahoo.fr
Laboratoire de Mécanique des Solides, CNRS UMR 7649, École Polytechnique ParisTech, Route de Saclay, 91128 Palaiseau, France

and

LaMSID/UMR EDF-CNRS-CEA 2832, EDF, Clamart, France, UME Mecanique, ENSTA ParisTech, France

ANDREI CONSTANTINESCU: andrei.constantinescu@lms.polytechnique.fr
Laboratoire de Mécanique des Solides, CNRS UMR 7649, École Polytechnique ParisTech, Route de Saclay, 91128 Palaiseau, France

HUBERT MAIGRE: hubert.maigre@insa-lyon.fr
INSA-Lyon, Villeurbanne, 69621 Lyon, France

SUBMISSION GUIDELINES

ORIGINALITY

Authors may submit manuscripts in PDF format online at the Submissions page. 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 one of the Chief Editors:

Davide Bigoni   bigoni@ing.unitn.it
Iwona Jasiuk   ijasiuk@me.concordia.ca
Yasuhide Shindo   shindo@material.tohoku.ac.jp

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 can use their preferred manuscript-preparation software, including for example Microsoft Word or any variant of TeX. The journal itself is produced in LaTeX, so accepted articles prepared using other software will be converted to LaTeX at production time. Authors wishing to prepare their document in LaTeX can follow the example file at www.jomms.net (but the use of other class files is acceptable). At submission time only a PDF file is required. After acceptance, authors must submit all source material (see especially Figures below).

REFERENCES

Bibliographical references should be complete, including article titles and page ranges. 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, for submission you may use the format of your choice. Links will be provided to all literature with known web locations; authors can supply their own links in addition to those provided by the editorial process.

FIGURES

Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector format for all diagrams and graphs in your manuscript: vector EPS or vector PDF files are the most useful. (EPS stands for Encapsulated PostScript.)

Most drawing and graphing packages—Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.—allow the user to save files in one of these formats. Make sure that what you're saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with as many details as you can about how your graphics were generated.

Please also include the original data for any plots. This is particularly important if you are unable to save Excel-generated plots in vector format. Saving them as bitmaps is not useful; please send the Excel (.xls) spreadsheets instead. Bundle your figure files into a single archive (using zip, tar, rar or other format of your choice) and upload on the link you been given at acceptance time.

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