JOURNAL OF RESEARCH of the National Bureau of Standards
Vol. 86, No. 1, January–February 1981

An Isoparametric Finite Element Model for Large-Strain Elastostatics

David S. Malkus*

Department of Mathematics, Illinois Institute of Technology, Chicago, IL 60616

and

E. R. Fuller, Jr.†

National Bureau of Standards, Washington, DC 20234

August 13, 1980

This paper describes a simple finite element model for large-strain elastostatics. The realization of the model in a small-scale computer-code is described. The purpose of the model is to produce test problems for research on the application of penalty techniques in nonlinear elasticity. For this reason the code must balance the requirements of reasonable flexibility with those of computational economy. The current code employs multilinear isoparametric elements. The model is capable of generalization to a variety of element types. The solution method employed is that of incremental loading combined with the Newton-Raphson method. Symmetric, banded systems of equations are produced which are solved in-core. Two- and three-dimensional symmetric bodies which are isoparametric images of a reference "brick" may be modeled. An example comparing two- and three-dimensional models of a "dogbone"—shaped A.S.T.M. rubber tensile-test specimen is presented. The results shed some light on the nature of stress-concentrations which occur in specimens of this geometry.

Key words: Checkerboard pressure; elastostatics; finite elements; in-core solver; isoparametrics; Mooney-Rivlin material; penalty method; plane-stress; nonlinear equation-solvers; strain invariants; tensile-test specimen.

1. Introduction

The finite element solution of the equations of elastostatic equilibrium for incompressible bodies poses special problems not encountered in the compressible case. These problems devolve from the fact that the constraint of material incompressibility turns the problem into a constrained minimization problem. There is a long history of attempts to find computationally efficient methods to impose the constraint within the context of the finite element method [1–5].[1] One of the most promising approaches has proved to be the "penalty method" [4–14]. This technique is one in which the discrete equations of the finite element method are modified to be able to incorporate a very large bulk modulus. Methods were devised [3] by which this could be done without incurring the disastrous loss of accuracy which accompanied early attempts to allow large bulk moduli in the finite element equations [1]. The state of the practice of penalty methods is now refined enough to allow the bulk modulus to be so large that practically any level of compressiblity can be achieved in this fashion. Many problems in incompressible elasticity and fluid dynamics are being successfully attacked using penalties [4–14].

The theory of penalty methods has not, however, caught up with practice. The stability and accuracy of finite element penalty formulations depend crucially on the choice of element types and numerical integra-

*The research described in this paper was begun while the first author was a National Research Council Postdoctoral Research Associate in the Mathematical Analysis Section of the National Bureau of Standards.

† Center for Materials Science, National Measurement Laboratory

[1] Figures in brackets indicate literature references at the end of this paper.

tion formulas [3–8]. Mathematical theories which can give predicted convergence rates and stability criteria exist for problems which are linear or those for which at least the constraint equations are linear [13,14]. Most of the successful applications of penalty methods are in such problems [6,8,11,12,14]. When the constraint is nonlinear—such as the constraint of material incompressibility under finite strain—the situation is quite different. Therefore problems of large-strain, incompressible elasticity are currently under intense scrutiny both in theory and in numerical experiments [9,10,15–19].

The purpose of this paper is to describe a relatively simple finite element model designed for the investigation of penalty methods in finite elasticity. The idea in development of the model is to balance the need for sufficient generality with the need for computational economy. Problems with moderately distorted geometry may be solved so that the interaction between isoparametric transformations and various numerical integration schemes may be observed. On the other hand, since it is desired to run numerous test cases over and over again varying key parameters, the model incorporates only the simplest of isoparametric mesh descriptions. The model is realized in a small-scale code in which all equations can be solved in-core.

The most important feature of the current model is that it has parallel two- and three-dimensional versions. This is designed to confront one of the major obstacles to numerical experimentation in finite elasticity: the unavailability of exact solutions for all but the most simplified problems. The current model allows computation of solutions for thin bodies using a fully three-dimensional formulation. This can be compared with plane-stress computations for the same body. In the plane-stress case, the incompressibility is imposed exactly in the model. In addition, the plane-stress equations assume a linear displacement variation through the thickness of the body. This can be matched exactly by the 3-D model using but one element through the thickness. Therefore the degree of success of the penalty method in imposing incompressibility can be evaluated by comparsion of the 2- and 3-D results. In this paper, we will deal with a simple, but interesting numerical test problem which illustrates such a comparison.

2. Finite element formulation

The kind of material properties we shall incorporate in our model are those of rubber-like, hyperelastic, isotropic materials. We adopt a total Lagrangian description similar to [19]. The static equilibria of such idealized materials are solutions to variational boundary-value problems similar to those described in [20]. We have found that we can derive the finite element equations of equilibrium very compactly if we base our derivations on differentiation with respect to displacement-gradients, as opposed to Green's strains [19,21]. For this reason we are naturally led to the use of the Lagrange stress tensor in our constitutive equation.

2.1. The variational formulation

Let u_i be the displacement field of a body with reference configuration Ω_o in a Cartesian system with coordinates, x^i. If $u_{i,j} = (\partial u_i/\partial x_j)$ is the displacement-gradient tensor, we shall make use of the deformation-gradient tensor

$$J_{ij} = \delta_{ij} + u_{i,j} \tag{1}$$

and the right Cauchy-Green tensor

$$G_{ij} = J_{ki}J_{kj} \tag{2}$$

The constitutive equation for the materials under consideration will be of the form

$$\tau_{ij} = \frac{\partial U}{\partial J_{ji}} = \frac{\partial U}{\partial u_{j,i}} \tag{3}$$

where τ_{ij} is the Lagrange stress tensor and U an energy per unit volume in the undeformed state.

The energy density, U, is assumed to be expressed as a function of the three principal invariants of G_{ij}. If we denote by ψ_{ij} the matrix of cofactors,

$$\psi_{ij} \equiv Co(J_{ij}) \tag{4}$$

the three principal invariants of G_{ij} are given by:

$$\begin{aligned} \mathrm{I} &= \mathrm{Trace}\,[G_{ij}] = J_{ki}J_{ki} \\ \mathrm{II} &= \mathrm{Trace}\,[Co(G_{ij})] = \psi_{ij}\psi_{ij} \\ \mathrm{III} &= |G_{ij}| = |J_{ij}|^2 \end{aligned} \tag{5}$$

where $|\bullet|$ denotes the determinant.

We consider problems whose solutions are characterized by the minima of the variational principle

$$\pi(u_i) = \int_{\Omega_o} [U(\mathrm{I},\mathrm{II},\mathrm{III}) - g_i u_i] dx^1 dx^2 dx^3 - \int_{\Gamma_\sigma^o} f_i u_i ds \tag{6}$$

where g_i is the body-force and f_i the surface tractions, both referred to the undeformed state. The surface of Ω_o, denoted Γ_o, is divided into a load-surface Γ_σ^o and a surface with prescribed displacements, Γ_u^o. All loads are assumed to be dead loads.

2.2. Isoparametric finite elements

We shall consider a particular kind of finite element mesh. It will be composed of only one type of element, which is defined on a reference domain Ω_R (the unit cube, for example), and is mapped into its location in the global mesh by a transformation based on the element shape-functions defined on Ω_R. A variety of moderately complex bodies may be modeled in this fashion, and with the development of some simple notation, the model can be described without much difficulty.

The process of constructing a finite element assemblage involves piecing together small domains to form a large domain. Similarly, the global trial functions are the unions of element trial functions which are supported (i.e., are nonzero) only on individual elements. The global trial functions we describe are compatible for elasticity problems [22,23]. This means that they are continuous across interelement boundaries, which is assured by requiring that the values of the trial function agree on shared boundaries.

a. The element transformations

Let the i^{th} component of the local coordinates of Ω_R be ξ^i, and select ξ_n^i as coordinates of the n^{th} of n_R nodal points of the prototypical element. Construct n_R shape functions, ϕ_r, defined for all values of ξ^i. These shape functions are polynomials on Ω_R, such that

$$\begin{cases} \phi_r(\xi^i) = 0 & \xi^i \notin \Omega_R \\ \phi_r(\xi_s^i) = \delta_{rs} & r,s = 1, \ldots, n_R \end{cases} \tag{7}$$

according to the procedure outlined in [22,23] and illustrated by an example in figure 1.

For each of the $1 \leqslant e \leqslant N_e$ elements in the global mesh, define a vector-valued map

$$T_i^e \equiv t_{ir}^e \phi_r \tag{8}$$

The summation convention is used on the non-tensorial index $r = 1, \ldots, n_R$, but is never used on the index e. Equation (8) defines the isoparametric transformation of each element from Ω_R to its location in the global mesh. Note that the global nodal locations are then

$$T_i^e(\xi_r^j) = t_{ir}^e \tag{9}$$

Examples of typical transformations are illustrated in figure 2.

We assume that the t_{ir}^e are chosen to satisfy certain basic properties: (1) that compatibility for elasticity problems is satisfied by the global shape functions, (2) that the totality of the transformations T_i^e defines an

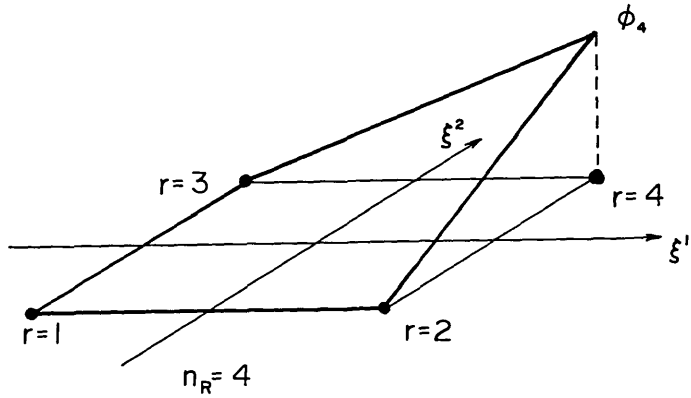

FIGURE 1. A finite element shape function on a reference four-node element.

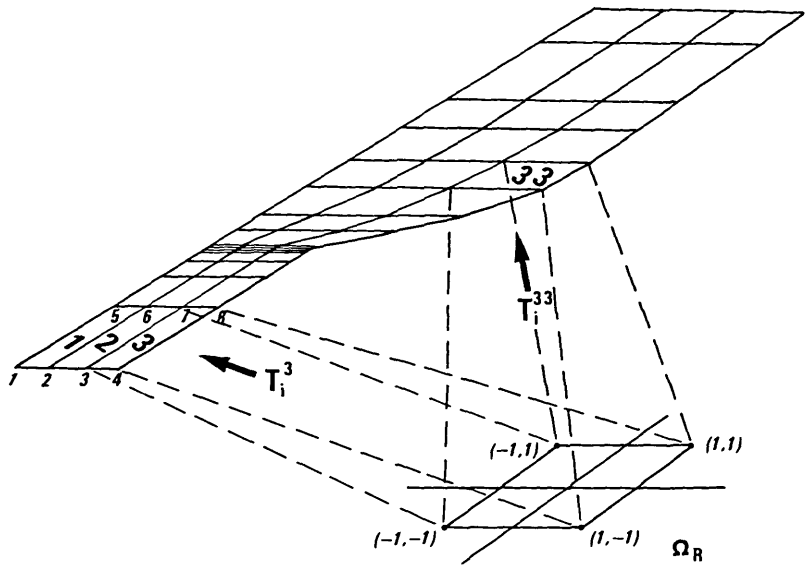

FIGURE 2. The transformation by T_i^e from the reference element on Ω_R to the global mesh on Ω_o.

assemblage of elements such that their domain $\overline{\Omega}_o$ with boundary $\overline{\Gamma}^o$ approximates Ω_o well enough so that $\Omega_o - \overline{\Omega}_o$ is negligible for the purposes of the model, (3) that T_i^e is an invertible transformation at every point of element e, with inverse S_i^e. The inverse S_i^e is extended to $\overline{\Omega}_o$ by

$$S_i^e(x^j) = \begin{cases} \xi^i, \ni T_j^e(\xi^i) = x^j, \ x^j \in \text{element } e \\ \text{any point not in } \Omega_R, \text{ otherwise} \end{cases} \tag{10}$$

The meshes which we will employ in our model are simple enough that the connectivity of the global mesh can be described by a closed form, integer-valued function, $A(e,r)$, called the "nodal alignment function". To each element e, $1 \leqslant e \leqslant N$, and r, $1 \leqslant r \leqslant n_R$, A assigns the global ordering number of the node with coordinates $t_{i,r}$. For more complicated meshes, A is well-defined, but may be determined from a table of data, rather than an arithmetic expression. In terms of $A(e,r)$, the global shape-functions have the form

$$\Phi_I = \bigcup_{\substack{e,r \\ \ni A(e,r)=I}} \{\phi_r[S_i^e(\bullet)]\} \tag{11}$$

In terms of (11), it can be seen that the purpose of the extension of S_i^e defined in (10) is merely to assure via (7) that whenever x^j is not in any of the elements with $A(e,r) = I$, then $\Phi_I(x^j)$ is well-defined and equal to zero.

A typical Φ_I and its construction is illustrated in figure 3. The finite element trial space, S^h, is then the space of all trial functions which satisfy the displacement boundary conditions on $\bar{\Gamma}_u^*$ and have the form

$$u_i(x^j) = u_i^I \Phi_I(x^j) \tag{12}$$

The sum on I is taken over all $1 \leq I \leq N$ global ordering numbers, and u_i^I is the i^{th} component of the nodal-value vector at node I. Equations (7) and (11) imply that if x_I^j are the coordinates of the I^{th} node in the global ordering,

$$u_i(x_I^j) = u_i^I \tag{13}$$

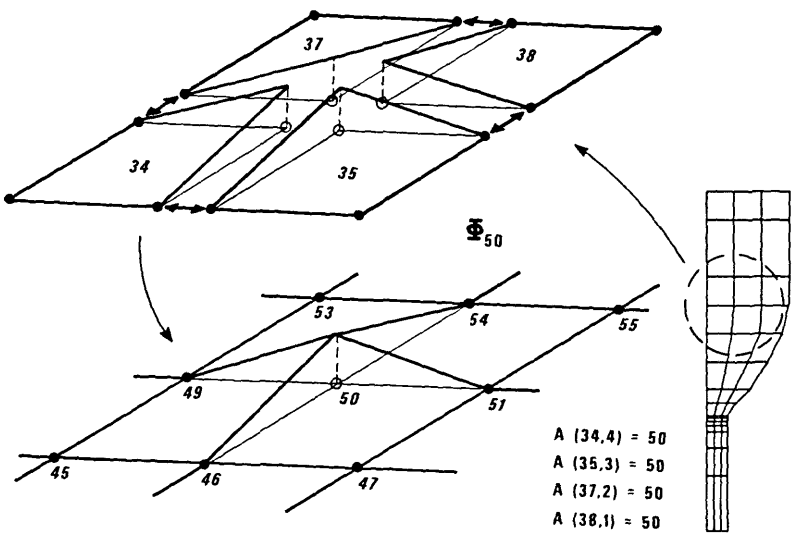

FIGURE 3. Alignment of local and global numberings, showing the role of $A(e,r)$ in piecing together the global shape function Φ_{50} from the ϕ_r on Ω_R.

The restriction of u_i to element e can be determined from the vectors of local nodal values in element e, u_i^{er}, $r = 1, \ldots, n_R$:

$$u_i^{er} = u_i^I, \qquad A(e,r) = I \tag{14}$$

Then by (11) through (14), the restrictions of u_i and $u_{i,j}$ to element e, denoted by, u_i^e and $u_{i,j}^e$, as functions of the reference coordinates, are

$$\begin{cases} u_i^e(x^j) = u_i^e[T_j^e(\xi^k)] = u_i^{er}\,\phi_r(\xi^k) \\ u_{i,j}^e(x^k) = [(\partial T^e)^{-1} u_i^{er} \dfrac{\partial \phi_r}{\partial \xi^l} C_{lj}^e](\xi^k) \end{cases} \tag{15}$$

where $\xi^k = S_k^e(x^j)$ and ∂T^e and C_{ij}^e are defined by

$$\begin{cases} \partial T^e_{ij} = \partial T^e_i / \partial \xi^j \\ \partial T^e = |\partial T^e_{ij}| \\ C^e_{ij} = Co\,(\partial T^e_{ji}) \end{cases} \tag{16}$$

To avoid confusion, differentiation with respect to the x^i will use the comma notation, and differentiation with respect to the ξ^i will use $\partial/\partial\xi^i$.

b. The element energy expressions

For U evaluated at any $u_i \in S^h$

$$\int_{\Omega_o} U\, dx^1 dx^2 dx^3 = \Sigma_e \int_{\Omega_R} U_e\, \partial T^e d\xi^1 d\xi^2 d\xi^3 \tag{17}$$

where

$$\begin{cases} U_e = U(\mathrm{I}_e, \mathrm{II}_e, \mathrm{III}_e) \\ \mathrm{I}_e = \mathrm{I}\,[u^e_{i,j}(\xi^k)], \text{ etc.} \end{cases} \tag{18}$$

The integral is evaluated by a quadrature rule with weights w_k and evaluation points η^i_k on Ω_R. We define elemental weights on e by

$$w^e_k = w_k\, \partial T^e(\eta^i_k) \tag{19}$$

The final energy expression is then

$$\int_{\Omega_o} U\, dx^1 dx^2 dx^3 \cong \Sigma_e \Sigma_k\, w^e_k\, U_e(\eta^i_k) \tag{20}$$

c. Modifications for penalty methods

The numerical integration scheme of (20) is adequate for a compressible energy density. For incompressible or nearly incompressible energy densities, following [8–10,15,19], we have

$$U(\mathrm{I,II,III}) = W(\mathrm{I,II}) + \frac{z}{2} G^2\,(\mathrm{III}-1) \tag{21}$$

where z is a bulk modulus scale factor; z may also be interpreted as the penalty enforcing III–1 to be arbitrarily close to zero for material incompressibility [8–10,15,19]. In the plane-stress case J_{ij} is assumed to have a form which enforces III = 1 so that $U(\mathrm{I,II,III}) = W(\mathrm{I,II})$. This will be fully described in section 3.2. It allows the numerical integration scheme of (20) to be applied in the plane-stress case.

For the 3-D model, two integration formulas are required, the one in (20) applied to the W terms in (21), and a "reduced" formula, with weights r_l and evaluation points ζ^i_l applied to the G^2 terms [8–10,22]. Equation 19 has analogue:

$$r^e_l = r_l\, \partial T^e(\zeta^i_l) \tag{22}$$

and (20) is replaced by:

$$\int_{\Omega_o} U\, dx^1 dx^2 dx^3 \cong \Sigma_e \left[\Sigma_k w^e_k W_e(\eta^i_k) + \frac{z}{2} \Sigma_l\, r^e_l G^2_e(\zeta^i_l)\right] \tag{23}$$

where W_e and G_e are defined as in (18).

2.3. Derivatives of the energy-density

The load terms in (6) from body and surface forces may be treated exactly as in linear finite element analysis [22,23]. Combining this with (20) or (23) gives a numerically integrated variational principle to be minimized over S^h. We will denote this principle by the same $\pi(u_i)$ and understand that throughout the following discussion the minimum is to be taken over S^h. The equations of equilibrium of the finite element model are then

$$F_i^I \equiv \frac{\partial \pi}{\partial u_i^I} = \sum_{\substack{e,r \\ \ni A(e,r) = I}} \Sigma_k \, w_k^e \frac{\partial W_e}{\partial u_i^{er}} (\eta_k^l) + f_i^I = 0$$

$$i = 1, 2 \text{ and } I = 1, 2, \ldots, N \tag{24}$$

in the plane stress case, and

$$F_i^I \equiv \frac{\partial \pi}{\partial u_i^I} = \sum_{\substack{e,r \\ \ni A(e,r) = I}} \left[\Sigma_k \, w_k^e \frac{\partial W_e}{\partial u_i^{er}} (\eta_k^l) + \frac{z}{2} \Sigma_t \frac{\partial}{\partial u_i^{er}} [G_e^2] (\zeta_t^l) \right] + f_i^I = 0 \tag{25}$$

$$i = 1, 2, 3 \text{ and } I = 1, 2, \ldots, N$$

For the 3-D case.

$$f_i^I = \sum_{\substack{e,r \\ \ni A(e,r) = I}} f_i^{er} \tag{26}$$

where f_i^{er} is the i^{th} component of the distributed nodal force from body and surface forces acting on node r of element e [22,23]. Thus f_i^I is just the usual finite element load vector, and (24–26) express the fact that the assembly of the equilibrium equations of the nonlinear model obeys the same rule as assembly of load vectors in linear analysis [22,23].

The equations of equilibrium, $F_i^I = 0$, will be solved by a combination of incremental loading with the Newton-Raphson method [22]. More details of these procedures will be described in section 4. Both incremental loading and the Newton-Raphson method require calculation of the "tangent stiffness matrix" [22],

$$\mathbf{K}_T = \left[\frac{\partial F_i^I}{\partial u_j^J} \right] \tag{27}$$

Since $\mathbf{K}_T$ is the matrix of second derivatives with respect to u_i^I of π, which is a continuously differentiable function of the nodal-values, $\mathbf{K}_T$ is a symmetric "hessian" matrix.

$$\frac{\partial F_i^I}{\partial u_j^J} = \frac{\partial F_j^J}{\partial u_i^I} = \sum_{\substack{e,r,s \\ \ni \substack{A(e,r)=I \\ A(e,s)=J}}} \Sigma_k w_k^e \frac{\partial^2 W_e}{\partial u_i^{er} \partial u_j^{es}} (\eta_k^l) \tag{28}$$

in the plane-stress case and

$$\frac{\partial F_i^I}{\partial u_j^J} = \frac{\partial F_j^J}{\partial u_i^I} = \sum_{\substack{e,r,s \\ \ni \substack{A(e,r)=I \\ A(e,s)=J}}} \left[\Sigma_k w_k^e \frac{\partial^2 W_e}{\partial u_i^{er} \partial u_j^{es}} (\eta_k^m) + \frac{z}{2} \Sigma_t r_t \frac{\partial^2}{\partial u_i^{er} \partial u_j^{es}} [G_e^2] (\zeta_t^m) \right] \tag{29}$$

in the 3-D case.

In other words, (27–29) imply that $\mathbf{K}_T$ assembles by the same rule by which stiffness matrices assemble in linear analysis [22,23]. Beyond that, $\mathbf{K}_T$ will have the same banded structure which a stiffness matrix in

linear analysis would have, using the same mesh [22,23]. This follows because the band-structure is determined by

$$\omega \equiv \max_{\substack{e,r,s \\ A(e,r)=I \\ \ni A(e,s)=J}} |I-J| \tag{30}$$

$\mathbf{K}_T$ will have $(2\omega+1)d$ bands, where $d = 3$ for the 3-D model and $d=2$ for plane-stress. Clearly ω is determined by the form of A and not the form of the energy expression. Boundary conditions are also imposed on $\bar{\Gamma}_u$ in the same manner as in linear analysis [22,23].

a. Separation of tensorial factors

In what follows, let U_e represent generically either a compressible element energy density, $U(\mathrm{I}_e,\mathrm{II}_e,\mathrm{III}_e)$, or one of the terms of an incompressible element energy density, W_e or G_e. From the previous section we may conclude that

$$\begin{cases} \dfrac{\partial U_e}{\partial u_i^{er}} = \dfrac{\partial U_e}{\partial u_{k,\ell}^e}\dfrac{\partial u_{k,\ell}^e}{\partial u_i^{er}} = \dfrac{\partial U_e}{\partial u_{i,j}^e}(\partial T^e)^{-1}\dfrac{\partial \phi_r}{\partial \xi^m}C_{mj}^e \\ \dfrac{\partial^2 U_e}{\partial u_i^{er}\partial u_k^{es}} = \dfrac{\partial^2 U_e}{\partial u_{i,j}^e \partial u_{k,\ell}^e}(\partial T^e)^{-2}\dfrac{\partial \phi_r}{\partial \xi^m}\dfrac{\partial \phi_s}{\partial \xi^n}C_{mj}^e C_{n\ell}^e \end{cases} \tag{31}$$

We point out again that there is no summation on e; in fact, since in what follows we will always be concerned with elemental expressions, we will drop the index e. At this point the strain-energy derivatives are seen to be the sum of terms which are separated into factors, the second of which,

$$(\partial T)^{-1}\frac{\partial \phi_r}{\partial \xi^m}C_{mj} \text{ or } (\partial T)^{-2}\frac{\partial \phi_r}{\partial \xi^m}\frac{\partial \phi_s}{\partial \xi^n}C_{mj}C_{n\ell} \tag{32}$$

is determined from the elemental isoparametrics and local shape functions, and is no different in character from the corresponding factors in linear isoparametric problems [22,23]. What distinguishes the finite elasticity case is the complexity of the first factors. It is these derivatives of the energy density with respect to displacement-gradients to which we now turn our attention. Actually, we may further restrict our attention to the derivatives of strain-invariants, since

$$\begin{cases} U_{ij} = \dfrac{\partial U}{\partial u_{i,j}} = \dfrac{\partial U}{\partial \mathrm{I}}\dfrac{\partial \mathrm{I}}{\partial u_{i,j}} + \dfrac{\partial U}{\partial \mathrm{II}}\dfrac{\partial \mathrm{II}}{\partial u_{i,j}} + \dfrac{\partial U}{\partial \mathrm{III}}\dfrac{\partial \mathrm{III}}{\partial u_{i,j}} \\ U_{ijk\ell} = \dfrac{\partial^2 U}{\partial u_{i,j}\partial u_{k,\ell}} = \dfrac{\partial^2 U}{\partial \mathrm{I}^2}\dfrac{\partial \mathrm{I}}{\partial u_{i,j}}\dfrac{\partial \mathrm{I}}{\partial u_{k,\ell}} + \dfrac{\partial U}{\partial \mathrm{I}}\dfrac{\partial^2 \mathrm{I}}{\partial u_{i,j}\partial u_{k,\ell}} \\ \qquad + \text{similar terms for II and III} \end{cases} \tag{33}$$

$\dfrac{\partial U}{\partial \mathrm{I}}, \dfrac{\partial^2 U}{\partial \mathrm{I}^2}, \dfrac{\partial U}{\partial \mathrm{II}} \ldots$, etc., depend on the form of the energy density for a specific material and are not usually difficult to evaluate. So the problem is to compute the derivatives of the invariants.

b. Derivatives of the strain-invariants

Derivatives of the strain invariants, I_{ij}, $\mathrm{I}_{ijk\ell}$, II_{ij}, ..., etc., are defined in a fashion similar to U_{ij} and $U_{ijk\ell}$. The expression for I_{ij} follows immediately from equation (5), as does III_{ij}, by remembering that the derivative of a determinant with respect to the matrix itself is the matrix of cofactors. Next we observe that the derivative of the matrix of cofactors with respect to the matrix itself can be written

$$\frac{\partial \psi_{ij}}{\partial J_{k\ell}} = \epsilon_{ikm}\epsilon_{j\ell n}J_{mn} \tag{34}$$

where ϵ_{ijk} is the Cartesian permutation symbol. Using the chain-rule and collecting terms by application of symmetries gives an expression for II_{ij}. So we have

$$\begin{cases} I_{ij} = 2J_{ij} \\ II_{ij} = 2\epsilon_{ikm}\epsilon_{jln}J_{mn}\psi_{kl} \\ III_{ij} = 2|J_{ij}|\psi_{ij} \end{cases} \tag{35}$$

Calculating I_{ijkl} is trivial. Using the chain-rule and the identity $\epsilon_{ijp}\epsilon_{klp} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{kj}$ leads readily to an expression for II_{ijkl}. To find III_{ijkl}, use can be made of the inverse of J_{ij}, $(J_{ij})^{-1} = \psi_{ji}/|J_{kl}|$. Then III_{ij} becomes

$$III_{ij} = 2|J_{ij}|^2(J_{ji})^{-1} \tag{36}$$

(note the transposition). Making use of the identity

$$\frac{\partial(J_{ji})^{-1}}{\partial J_{kl}} = -(J_{li})^{-1}(J_{jk})^{-1} = \psi_{il}\psi_{kj}/|J_{mn}|^2 \tag{37}$$

leads to a simple expression for III_{ijkl}. Accordingly we have

$$\begin{cases} I_{ijkl} = 2\delta_{ik}\delta_{jl} \\ II_{ijkl} = 2\epsilon_{ikm}\epsilon_{jln}\psi_{mn} + 2\delta_{ik}\delta_{jl}J_{mn}J_{mn} \\ \qquad\qquad - 2\delta_{ik}J_{ml}J_{mj} - 2\delta_{jl}J_{im}J_{km} + 2J_{ij}J_{kl} \\ III_{ijkl} = 4\psi_{ij}\psi_{kl} - 2\psi_{il}\psi_{kj} \end{cases} \tag{38}$$

Although the terms with the ϵ_{ijk} may look forbidding, they may be computed using no floating point operations other than "change sign," and the same calculations can be used for II_{ij} and II_{ijkl}. Alternative expressions (of comparable complexity) for II_{ij} and II_{ijkl} can be derived without ϵ_{ijk} using the identity: $II = (1/2)(I^2 - G_{ij}G_{ij})$.

3. Two- and three-dimensional problems

In this section we describe the assumptions which are made in reducing a three-dimensional problem in finite elasticity to a two-dimensional problem. We describe how this reduction is incorporatd in the framework of the results of section 2. We also will describe the similarity and differences between the element shape-functions and the isoparametric transformations in the two- and three-dimensional cases. It will be easier to describe the implementation of the isoparametric transformatlons in two-dimensions. The three-dimensional implementation can be viewed as a straight-forward generalization of the 2-D case. Finally, as will be explained in more detail in section 5, linear elastic problems serve an important purpose in computer code verification. They are also interesting in their own right. We therefore conclude this section with a discussion of two- and three-dimensional linear elasticity problems. The linear two-dimensional model incorporates compressible, nearly incompressible and incompressible materials in states of plane stress or plane strain. Penalties are used to enforce plane-strain incompressibility.

3.1. Three-dimensional problems

The results of section 2 were derived in full three-dimensional form. They may be applied directly to the coding of the element matrix, element load-vector and assembly subroutines [22,23]. Details of how the mesh is described and how the isoparametric transformations are generated will be described later in this section.

But the two-dimensional, plane-stress model requires additional development. It may be thought of as a three-dimensional model in which specific simplifying assumptions are imposed on the stress and strain tensors. To see how this is accomplished, it is useful to look first at those tensors in the fully three-dimensional case and observe an example of a particular material model.

a. The general incompressible formulation

In the 3-D case J_{ij}, G_{ij}, τ_{ij} and all the tensors of the previous section are of dimension three in each rank. The Lagrange stress tensor can be written in terms of (35)

$$\tau_{ij} = \frac{\partial U}{\partial u_{j,i}} = \frac{\partial U}{\partial \mathrm{I}} \mathrm{I}_{ji} + \frac{\partial U}{\partial \mathrm{II}} \mathrm{II}_{ji} + \frac{\partial U}{\partial \mathrm{III}} \mathrm{III}_{ji} \tag{39}$$

For incompressible, 3-D problems, U is assumed to have the form (21), so (39) becomes

$$\tau_{ij} = \frac{\partial W}{\partial \mathrm{I}} \mathrm{I}_{ji} + \frac{\partial W}{\partial \mathrm{II}} \mathrm{II}_{ji} + zG(\mathrm{III}-1) \frac{\partial G}{\partial \mathrm{III}} \mathrm{III}_{ji} \tag{40}$$

Recall that z is a large penalty parameter forcing $G(\mathrm{III}-1)$ to be small in the finite element solution. Therefore the pressure is given by

$$h = zG(\mathrm{III}-1) \frac{\partial G}{\partial \mathrm{III}} \tag{41}$$

so that (40) gives an incompressible constitutive equation in the $z \to \infty$ limit.

For stress calculations from the finite element model, it is more physically meaningful to compute the Euler stress tensor, referred to the strained body:

$$\sigma'_{ij} = J_{ki}\tau_{kj} / |J_{mn}| \tag{42}$$

Equation (35) and the fact that $(J_{ji})^{-1} = \psi_{ij} / |J_{ij}|$ with (41–42) imply

$$\sigma'_{ij} = \left[\frac{\partial W}{\partial \mathrm{I}} J_{ki}\mathrm{I}_{jk} + \frac{\partial W}{\partial \mathrm{II}} J_{ki}\mathrm{II}_{jk} + 2h\mathrm{III}\delta_{ij}\right] / |J_{mn}| \tag{43}$$

Then denote by B_{ij} the left Cauchy-Green tensor and by C_{ij} its inverse

$$\begin{aligned} B_{ij} &= J_{ik}J_{jk} \\ C_{ij} &= \psi_{kj}\psi_{ki} / |J_{mn}|^2 \end{aligned} \tag{44}$$

We can rewrite II as

$$\mathrm{II} = C_{ii} |G_{mn}| \tag{45}$$

from which it easily follows using (37) that

$$J_{ki}\mathrm{II}_{jk} = 2\mathrm{III}[C_{kk}\delta_{ij} - C_{ij}] \tag{46}$$

which leads to

$$\sigma'_{ij} = 2\sqrt{\mathrm{III}} \left[\frac{\partial W}{\partial \mathrm{I}} B_{ij} - \frac{\partial W}{\partial \mathrm{II}} C_{ij} + \bar{h}\, \delta_{ij}\right] \tag{47}$$

in which

$$\bar{h} = \mathrm{III}h + \frac{\partial W}{\partial \mathrm{II}} C_{ii} \tag{48}$$

Finally it is desired to have a stress-free undeformed state, but for $J_{ij} = \delta_{ij}$ (47) gives

$$\sigma'_{ij} = h_o \delta_{ij} = \left(\frac{\partial W}{\partial \mathrm{I}} + 2\frac{\partial W}{\partial \mathrm{II}}\right)\delta_{ij} \tag{49}$$

so letting $H = \bar{h} - h_o$

$$\sigma_{ij} = 2\sqrt{\mathrm{III}}\left[\frac{\partial W}{\partial \mathrm{I}} B_{ij} - \frac{\partial W}{\partial \mathrm{II}} C_{ij} + H\delta_{ij}\right] \tag{50}$$

All constituents of (50) are easy to compute given J_{ij}. Symmetry implies only six of the components of σ_{ij} need be calculated.

b. Application to Mooney-Rivlin materials

A simple constitutive equation of the form (40) can be obtained by taking

$$\begin{cases} \dfrac{\partial W}{\partial \mathrm{I}} = c_1, \quad \dfrac{\partial W}{\partial \mathrm{II}} = c_2 \\ \\ G(x) = x \end{cases} \tag{51}$$

This is a penalty function version of the Mooney-Rivlin constitutive equation [9,10,19, and 21]; c_1 and c_2 are material constants. This equation is intended to model exactly incompressible materials, and in fact, it has some continuum mechanical inadequacies if significant compressive deformation is allowed [10]. Therefore it should only be employed with large values of z. The exactly incompressible, plane-stress version of the constitutive model has been found to give reasonable agreement with experimental results in some problems [21]. How well the 3-D penalty version succeeds will be discussed below.

The symmetric stress tensor for Mooney-Rivlin material is given by

$$\begin{cases} \sigma_{ij} = 2\sqrt{\mathrm{III}}\,[c_1 B_{ij} - c_2 C_{ij} + H\delta_{ij}] \\ \\ H = z(\mathrm{III}-1)\mathrm{III} + c_2 C_{ii} - c_1 - 2c_2 \end{cases} \tag{52}$$

It should be noted that this differs slightly from the H used elsewhere [9,19,21], because $c_2 C_{ii}\delta_{ij}$ is considered a pressure term. This means that H defined in (52) has the value $c_2 - c_1$ in the undeformed state. [In [19], since the $c_2 C_{ii}\delta_{ij}$ term is included with the deviatoric terms, the pressure in the undeformed state is $-c_1 - 2c_2$.] A quantity of more interest than H, then, is the pressure above the undeformed state:

$$\bar{H} = z(\mathrm{III}-1)\mathrm{III} + c_2(C_{ii}-3) \tag{53}$$

3.2. Finite plane stress

a. The general incompressible formulation

The assumptions in the plane-stress model are that

1. The body is very thin, with uniform thickness, $2t$, in the x^3-coordinate. How thin the body must be will be dealt with further below.
2. The applied forces are in the x^1-x^2 plane; the upper and lower surfaces in the x^1-x^2 plane are free surfaces.
3. $\sigma_{31} = \sigma_{32} = \sigma_{33} = 0$ throughout the thickness, i.e. for $-t \leq x^3 \leq t$.

4. The deformation is essentially biaxial, i.e.

$$\begin{aligned} u_1(x^1,x^2,x^3) &= u_1(x^1,x^2) \\ u_2(x^1,x^2,x^3) &= u_2(x^1,x^2) \\ u_3(x^1,x^2,x^3) &= [\Theta(x^1,x^2)-1]x^3 \end{aligned} \tag{54}$$

for an unknown function $\Theta(x^1,x^2)$.

5. The body is exactly incompressible.

This model is discussed in more detail in [21]. An example of such a deformation is shown schematically in figure 4.

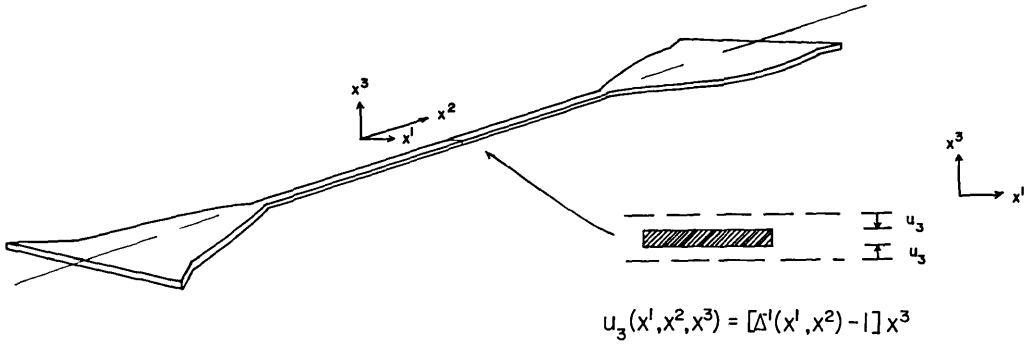

FIGURE 4. Plane-stress assumptions for the deformation through the thickness of a thin body.

These assumptions imply that the deformation-gradients have the form

$$J_{ij} = \begin{bmatrix} 1+u_{1,1} & u_{1,2} & 0 \\ u_{2,1} & 1+u_{2,2} & 0 \\ x^3\Theta_{,1} & x^3\Theta_{,2} & \Theta \end{bmatrix} \tag{55}$$

Assumption 1 is further specified by the requirement that the body be thin enough for $t\Theta_{,1}$ and $t\Theta_{,2}$ to be neglected. Note that this is not an absolute requirement on the thickness of the body, since if $\Theta_{,1}$ and $\Theta_{,2}$ are small enough, it can compensate for a larger t. On the other hand, very rapid spatial variations of $\Theta(x^1,x^2)$ could invalidate the assumption, even for small t.

Let $K_{\alpha\beta}$ be defined for $\alpha,\beta = 1,2$

$$K_{\alpha\beta} = \begin{bmatrix} 1+u_{1,1} & u_{1,2} \\ u_{2,1} & 1+u_{2,2} \end{bmatrix} \tag{56}$$

We will use Greek subscripts to indicate tensors which are of dimension 2 in each rank. Imposing the thinness assumption leads to

$$J_{ij} = [K_{\alpha\beta} \oplus \Theta] \tag{57}$$

where " $\oplus$ " means "matrix direct sum" of the 2x2 matrix, $K_{\alpha\beta}$, with the 1x1 matrix, Θ. That is

$$J_{ij} = \left[\begin{array}{cc:c} K_{\alpha\beta} & & \begin{matrix}0\\0\end{matrix} \\ \hdashline \begin{matrix}0 & 0\end{matrix} & & \Theta \end{array}\right] \tag{58}$$

Assumption 5 is enforced by taking $\Theta = \Delta^{-1}(x^1,x^2)$ where

$$\Delta(x^1,x^2) = |K_{\alpha\beta}| \tag{59}$$

Now we can rewrite the invariants of G_{ij} and their derivatives, using (35–38) and the special form (58). In (58) we have mixed tensor and matrix notation in an obvious way. In what follows, we wish to construct tensors of rank 4 from tensors of rank 2 given in matrix notation. We adopt the convention that a matrix with subscripts appended on the lower right defines a tensor of rank 2 with components given by the matrix entries. The first subscript is the row index and the second the column index, as usual. For example, in terms of (58), ψ_{ij} of cofactors of J_{ij} is

$$\psi_{ij} = \{[\Delta^{-1}CoK_{\alpha\beta}] \oplus \Delta\}_{ij}. \tag{60}$$

Similarly

$$\left\{\begin{aligned}
\mathrm{I} &= K_{\alpha\beta}K_{\alpha\beta} + \Delta^{-2} \\
\mathrm{II} &= \Delta^{-2}CoK_{\alpha\beta}CoK_{\alpha\beta} + \Delta^{2} \\
\mathrm{I}_{ij} &= 2[K_{\alpha\beta} \oplus \Delta^{-1}]_{ij} \\
\mathrm{II}_{ij} &= 2\{[\Delta^{-2}K_{\alpha\beta} + \Delta CoK_{\alpha\beta}] \oplus [\Delta^{-1}CoK_{\alpha\beta}CoK_{\alpha\beta}]\}_{ij} \\
\mathrm{III}_{ij} &= 2\{[\Delta^{-1}CoK_{\alpha\beta}] \oplus \Delta\}_{ij} \\
\mathrm{I}_{ijkl} &= 2\delta_{ik}\delta_{jl} \\
\mathrm{II}_{ijkl} &= 2\{\epsilon_{ikm}\epsilon_{jln}\psi_{mn} + \delta_{ik}\delta_{jl}\mathrm{I} \\
&\quad - \delta_{ik}[K_{\gamma\alpha}K_{\gamma\beta} \oplus \Delta^{-2}]_{jl} - \delta_{jl}[K_{\alpha\gamma}K_{\beta\gamma} \oplus \Delta^{-2}]_{ik} \\
&\quad + J_{ij}J_{kl}]\}
\end{aligned}\right. \tag{61}$$

Equation (58) has made III = 1 and the energy density is not a function of III, so III_{ijkl} is not needed. III_{ij} would only be needed in the event one wished to calculate τ_{ij}. By enforcing plane-stress assumption 3, the pressure can be determined directly. It should be observed that (58) and (59) already imply that $\sigma_{31} = \sigma_{32} = 0$, since (58) implies

$$\left\{\begin{aligned}
B_{ij} &= [K_{\alpha\gamma}K_{\beta\gamma} \oplus \Delta^{-2}]_{ij} \\
C_{ij} &= [\Delta^{-2}CoK_{\gamma\alpha}CoK_{\gamma\beta} \oplus \Delta^{2}]_{ij}
\end{aligned}\right. \tag{62}$$

Substituting (62) in (50), setting $\sigma_{33} = 0$ and solving for H implies

$$H = -\frac{\partial W}{\partial \mathrm{I}}\Delta^{-2} + \frac{\partial W}{\partial \mathrm{II}}\Delta^{2} \tag{63}$$

The reduced constitutive law may be now written

$$\left\{\begin{aligned}
\sigma_{ij} &= 2\left[\frac{\partial W}{\partial \mathrm{I}}B_{ij} - \frac{\partial W}{\partial \mathrm{II}}C_{ij} + H\,\delta_{ij}\right] \\
H &= -\frac{\partial W}{\partial \mathrm{I}}\Delta^{-2} + \frac{\partial W}{\partial \mathrm{II}}\Delta^{2}
\end{aligned}\right. \tag{64}$$

Assumption 3 is enforced by choosing plane finite elements of the type illustrated in figures 1-3 and by replacing every occurrence of $u_{\alpha,3}$ or $u_{3,\alpha}$ by zero ($\alpha = 1,2$) and $u_{3,3}$ by Θ in the energy expressions of section 2.2. Any explicit occurrence of x^3 is replaced by $x^3 = 0$, since we assume negligible variation of all quantities through the thickness. The x^3-integral in the multiple integrals of the energy expressions integrates to a factor of $2t$ in the energy expressions. This may be ignored since this will not change the minimizing displacement field. The result of the reduction is that there is no dependence on x^3 or t in the plane-stress model.

A reduced displacement field is used which has only two components, u_α; u_3 is a function of u_1 and u_2. Therefore in section 2.3, all derivatives with respect to nodal values of u_3 vanish. But it is important to note that in (31), terms of the form $\dfrac{\partial u_{3,3}}{\partial u_\alpha^{er}}$ do not vanish. They have a new form, introduced by the plane-stress assumption of the form of $u_3(x^1,x^2)$. For notational convenience, define

$$\phi_{rm} \equiv (\partial T)^{-1} \frac{\partial \phi_r}{\partial \xi^k} C_{km} \tag{65}$$

This is the r^{th} element shape function's derivative with respect to Cartesian spatial variable x^m, and for plane stress $m = 1$ or 2. Let $L = $ I, II or III and u_α^r $\alpha = 1$ or 2 be the α^{th} component of the r^{th} local nodal value in an element

$$\left\{\begin{aligned} \frac{\partial L}{\partial u_\alpha^r} &= \frac{\partial L}{\partial u_{i,j}} \frac{\partial u_{i,j}}{\partial u_\alpha^r} = L_{ij} \frac{\partial u_{i,j}}{\partial u_\alpha^r} \\ \frac{\partial^2 L}{\partial u_\alpha^r \partial u_\beta^s} &= \frac{\partial^2 L}{\partial u_{i,j} \partial u_{k,\ell}} \frac{\partial u_{i,j}}{\partial u_\alpha^r} \frac{\partial u_{k,\ell}}{\partial u_\beta^s} + L_{ij} \frac{\partial^2 u_{i,j}}{\partial u_\alpha^r \partial u_\beta^s} \\ &= L_{ijk\ell} \frac{\partial u_{i,j}}{\partial u_\alpha^r} \frac{\partial u_{k,\ell}}{\partial u_\beta^s} + L_{ij} \frac{\partial^2 u_{i,j}}{\partial u_\alpha^r \partial u_\beta^s} \end{aligned}\right. \tag{66}$$

Since trial functions are linear in the nodal-value coefficients u_α^r, if i and $j \neq 3$ we have

$$\frac{\partial u_{i,j}}{\partial u_\alpha^r} = \delta_{i\alpha}\phi_{rj}, \quad \frac{\partial^2 u_{i,j}}{\partial u_\alpha^r \partial u_\beta^s} = 0 \tag{67}$$

$u_{3,3}$ is not linear in u_α^r and is a function of the in-plane displacement gradients. Therefore contributions to the chain-rule sums in (66) from $i,j,k,$ and $\ell = 1$ or 2 can be computed directly from (61), and (67) which hold for these values of the subscripts $i,j,k,$ and ℓ. The contributions to (66) from $i = j = 3$ and/or $k = \ell = 3$ may be found by using L_{ij} and $L_{ijk\ell}$ determined by (61) and the following facts:

$$\left\{\begin{aligned} \frac{\partial u_{3,3}}{\partial u_\alpha^r} &= \frac{\partial \Delta^{-1}}{\partial u_\alpha^r} = -\Delta^{-2} \frac{\partial \Delta}{\partial u_\alpha^r} \\ \frac{\partial \Delta}{\partial u_\alpha^r} &= (-1)^{\alpha+1}(\phi_{r\alpha} J_{3-\alpha\ 3-\alpha} - \phi_{r\ 3-\alpha} J_{3-\alpha\ \alpha}) \\ &\qquad \text{(no sum on } \alpha) \\ \frac{\partial^2 u_{3,3}}{\partial u_\alpha^r \partial u_\beta^s} &= 2\Delta^{-3} \frac{\partial \Delta}{\partial u_\alpha^r} \frac{\partial \Delta}{\partial u_\beta^s} - \Delta^{-2} \frac{\partial^2 \Delta}{\partial u_\alpha^r \partial u_\beta^s} \\ \frac{\partial^2 \Delta}{\partial u_\alpha^r \partial u_\beta^s} &= (1-\delta_{\alpha\beta})(-1)^{\alpha+1}(\phi_{r\alpha}\phi_{s\beta} - \phi_{r\beta}\phi_{s\alpha}) \\ &\qquad \text{(no sum on } \alpha,\beta) \end{aligned}\right. \tag{68}$$

b. Application to Mooney-Rivlin materials

When the constitutive equation with $\dfrac{\partial W}{\partial I}$ and $\dfrac{\partial W}{\partial II}$ given as in (51) is employed in the plane-stress case (64) becomes

$$\left\{\begin{aligned} \sigma_{ij} &= 2[c_1 B_{ij} - c_2 C_{ij} + H\delta_{ij}] \\ H &= -c_1 \Delta^{-2} + c_2 \Delta^2 \end{aligned}\right. \tag{69}$$

$\bar{H}$, the pressure above the undeformed state, given by

$$\bar{H} = c_1(1-\Delta^{-2}) + c_2(\Delta^2-1) \tag{70}$$

3.3. Shape functions and isoparametrics

Here we discuss the specific choice of shape functions, ϕ_r, and the isoparametric transformations, T_i^e, which are described generally in section 2.2. It will be easier to describe the two-dimensional case first and think of the 3-D case as the appropriate generalization.

a. Two dimensions

The reference element, Ω_R, is the square in the $\xi^1-\xi^2$ plane with $-1 \le \xi^i \le 1$. On this domain,

$$\phi_r(\xi^1,\xi^2) = \frac{(\xi^1+\xi_r^1)(\xi^2+\xi_r^2)}{4\,\xi_r^1\,\xi_r^2} \tag{71}$$

where $(\xi_1^1, \xi_1^2) = (-1,-1)$, $(\xi_2^1, \xi_2^2) = (1,-1)$, $(\xi_3^1, \xi_3^2) = (-1,1)$ and $(\xi_4^1, \xi_4^2) = (1,1)$. These element shape functions look exactly like the one illustrated in figure 1.

To construct the isoparametric transformation, T_i^e, all that is needed is to determine the nodal locations, t_{ir}^e, of (8). Then the isoparametric transformation of Ω_R is determined throughout element e via (8). The nodal locations t_{ir}^e give the $i = 1,2$ coordinates in the global mesh of the nodes $r = 1,2,3,4$ of element e. From this information, the rest of the computations implementing the ideas of section 2 can be carried out once the nodal alignment function, $A(e,i)$, is defined. To do this, we chose to restrict our attention to curved bodies which are described as a class of differentiable, invertible tensor transformations of the reference "brick" of elements illustrated in figure 5. It is clear that the nodal alignment of such a body is the same as the nodal alignment of the reference brick. Therefore the global element and nodal numbering schemes and $A(e,i)$ may be determined from the reference brick. We have chosen the numbering scheme illustrated in figure 5. In terms of these numberings, $A(e,i)$ can be expressed as a simple, closed form, integer expression.

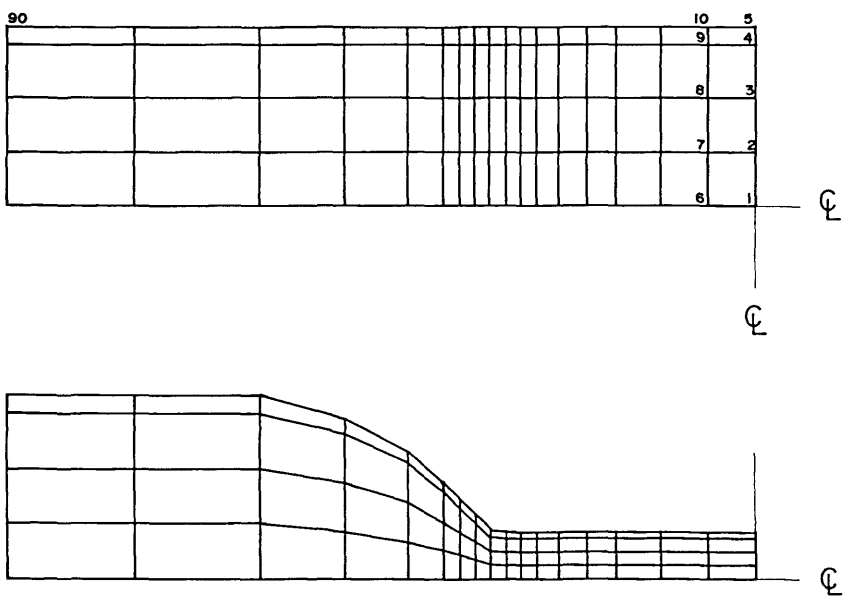

FIGURE 5. A two-dimensional reference brick before and after curving. Numbers indicate beginning and ending of global numbering scheme.

As a further simplification, we consider only bodies which have symmetry lines on the lower boundary and rightmost boundary as illustrated in figure 5. The transformation of the reference brick is generated by defining a cubic spline [24] which interpolates to given data defining the location of the upper boundary of the curved body (the leftmost boundary is not curved). This spline is used to define a transformation which varies linearly in the interior of the reference brick in such a way that the reference brick and curved body have the same axes of symmetry. This means that the coordinates, $t_{i,r}^c$, of the nodal locations in the curved body can be obtained by multiplying the appropriate coordinate of the corresponding nodes in the reference brick by a factor which depends on the given cubic spline and the distance of that node from the symmetry line in the reference brick.

b. Three dimensions

The ϕ_r are trilinear polynomials on Ω_R, which is the cube such that $-1 \leq \xi^i \leq 1$.

$$\phi_r(\xi^1,\xi^2,\xi^3) = \frac{(\xi^1+\xi_r^1)(\xi^2+\xi_r^2)(\xi^3+\xi_r^3)}{8\,\xi_r^1\,\xi_r^2\xi_r^3} \tag{72}$$

where ξ_r^i are the coordinates of the corners of Ω_R.

The $t_{i,r}^c$ are determined analogously to the way they are determined in 2-D. In 3-D, the back, rightmost and bottom surfaces of the reference brick are symmetry surfaces, as illustrated in figure 6. The leftmost surface is not curved. The remaining two surfaces are defined by interpolating splines and define transformations which vary linearly in the interior in such a way that the symmetry planes of reference and curved domain coincide. In the current code a further simplification is obtained by requiring that the surface splines be tensor products of one-dimensional splines. This restricts the type of surface which can be described but allows the use of a simple, one-dimensional spline code.

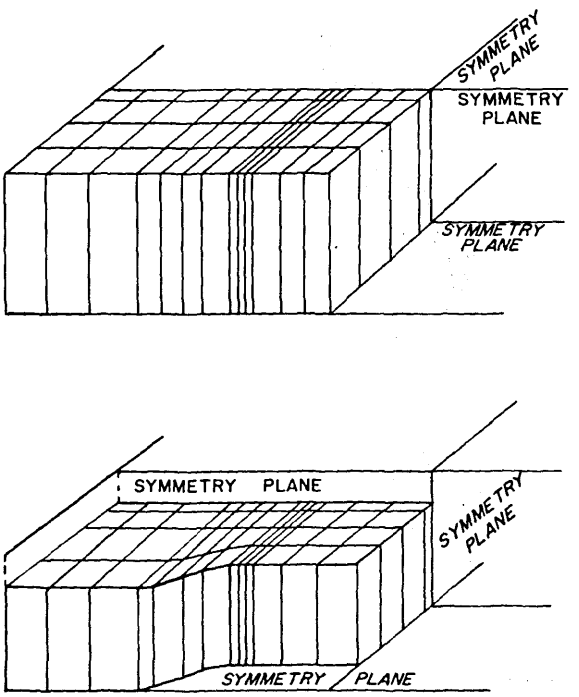

FIGURE 6. The three dimensional generalization of the curving procedure illustrated in figure 5.

3.4. Linear elasticity

The current code can compute finite element solutions to standard 2- and 3-D linear isotropic elasticity problems [22]. The same shape-functions and mesh generation are used as for the nonlinear problems. The linear solutions can be compared to small-strain nonlinear solutions. This allows determination of the load and/or strain levels at which the linear and nonlinear elastic models depart significantly in their predictions. It is also interesting to observe the nature of the differences in the linear and nonlinear solutions. A simple analysis shows that the Mooney constitutive laws (52) and (69) have shear modulus

$$\mu = 2(c_1 + c_2) \tag{73}$$

in the zero-strain limit. The incompressibility condition determines the other constitutive parameter of linear, isotropic elasticity which give a correspondence between the material properties of the linear and nonlinear models in the small-strain limit.

For comparison with nonlinear plane-stress solutions, it is natural to look at linear plane-stress solutions. As with the nonlinear constitutive laws, the linear plane-stress model enforces incompressibility exactly by assumption, without penalties. However, by a well-known re-interpretation of the material properties, every plane-stress solution is also a plane-strain solution [25]. Incompressible plane-strain requires a penalty to enforce zero volume change. Therefore the current code can be used to evaluate the behavior of two-dimensional elements with penalties, at least in linear elasticity.

a. Three-dimensional linear problems

Three-dimensional linear elasticity with a penalty to enforce incompressibility is treated exactly as described in [5]. In terms of the infinitesimal strain tensor $e_{ij} = 1/2(u_{i,j} + u_{j,i})$, the constitutive law is

$$\sigma_{ij} = 2\mu\{e_{ij} + (ze_{ii})\delta_{ij}\} \tag{74}$$

μ is the shear modulus and z is the penalty parameter

$$z = \frac{\nu}{1-2\nu} \tag{75}$$

ν is Poisson's ratio, which has the value 1/2 for exact incompressibility. By taking

$$h = \frac{z}{\nu} e_{ii} \tag{76}$$

equation (74) becomes

$$\sigma_{ij} = 2\mu \{e_{ij} + \nu h\delta_{ij}\} \tag{77}$$

which is the usual constitutive law for linear, isotropic elasticity, which is valid for incompressible material if $\nu = 1/2$ [1].

b. Linear plane-stress and plane-strain

The linear plane-strain constitutive law is used to model the cross-section of a body which is very long in the x^3-direction [25]. The problem is two-dimensional since it is assumed that $u_i(x^1,x^2,x^3) = u_i(x^1,x^2)$ and $e_{33} = 0$. Substituting this into (74–77) gives a 2-D constitutive law for which

$$h = \frac{z}{\nu}(e_{11} + e_{22}) \tag{78}$$

Then as $\nu \to 1/2$, z is a penalty parameter enforcing $e_{11} + e_{22} = 0$ in the incompressible limit. The energy expressions and the finite element model are very similar to 3-D linear elasticity, but $e_{33} = 0$, and there is no x^3-dependence. If we take $\alpha,\beta = 1,2$ the plane-strain constitutive law is

$$\sigma_{\alpha\beta} = 2\mu\{e_{\alpha\beta} + \frac{\nu}{1-2\nu}\delta_{\alpha\beta}e_{\gamma\gamma}\} \tag{79}$$

For any value of $\nu \neq 1/2$, if we re-interpret a solution obtained using (79) by defining

$$\begin{aligned} \nu &= \frac{\eta}{1+\eta} \\ e_{33} &= -\frac{\eta}{1-\eta}(e_{11} + e_{22}) \end{aligned} \tag{80}$$

it is easy to see that we have a linear plane-stress solution satisfying $\sigma_{33} = 0$, with apparent Poisson's ratio η, and shear modulus μ [25]. In particular, if we want the plane-stress solution to be incompressible, we choose $\eta = 1/2$. This corresponds to $\nu = 1/3$ in (79), or $z = 1$. The current code then can be used to obtain an incompressible plane-stress solution by setting $\nu = 1/3$, and an incompressible plane-strain solution by the penalty approach, taking ν close to $1/2$, but not exactly equal to $1/2$. Since they involve quite different geometric assumptions, the two kinds of 2-D incompressible solutions usually are quite different. The same kind of comparisons between the plane-stress solutions and 3-D solutions are possible as were for nonlinear elasticity.

4. Nonlinear equation-solving procedures

One of the principles which has guided us in the construction of this finite element model is that it should retain as much of the structure of a linear elasticity model as possible. Other than in the assumption of somewhat simplified loading, this has been done without making any restriction on the degree of nonlinearity in the energy expressions. The point is that this can be done because the energy principle (6) which characterizes solutions in our chosen continuum model admits to the generalized Rayleigh-Ritz procedures described above. Because trial function are linear combinations of the basis functions, and because of the linearity of the integral in the energy expressions—features which are common to linear and nonlinear elastic models—the assembly of elemental expressions and much of the generation of these expressions are essentially linear procedures. Much of the current computer code is the same code which would be generated to solve linear elasticity problems alone.

To solve the nonlinear equations (24) or (25), many of the common procedures involve generation of a sequence of linearizations using the tangent stiffness matrix (27). A sequence of linear systems is solved with $\mathbf{K}_T$ on the left-hand side. We have seen that $\mathbf{K}_T$ has the same banded structure in the linear and nonlinear case. Furthermore, it is a consequence of variational calculus [26] that at the solution to (24) or (25), $\mathbf{K}_T$ is positive-definite (we assume our problem is well-posed). It is a guiding principle of the nonlinear iteration schemes described here to retain positive definiteness of $\mathbf{K}_T$ throughout the whole sequence of iterates. First of all, this will allow even more common code between the linear and nonlinear procedures: the same stable positive-definite, linear-equation solvers can be used. Second—and in our view more important—is that departure of the sequence of iterates from the region of positive-definiteness of $\mathbf{K}_T$ has serious implications for the stability of the whole solution process—both in a numerical and analytical sense. Loss of positive definiteness of $\mathbf{K}_T$ means that the numerical inversion procedure is less stable in the sense that small rounding errors are more likely to be magnified [23]. But potentially more serious is the fact that loss of positive-definiteness of $\mathbf{K}_T$ is an indication that the current iterate in the nonlinear scheme is outside the region of local convexity surrounding the desired solution [21]. The iteration scheme runs a danger of becoming "lost" and failing to converge.

The numerical results described in the next section show that the possibility of the iteration scheme becoming lost in a region of nonconvexity of the energy is sometimes a very real one—in spite of the precau-

tions we take. One of the most important conclusions which can be drawn from the current state of research into penalty methods is finite elasticity is a negative one: penalties greatly increase the likelihood of departure from the convex region. Before we discuss this further we describe the iteration schemes we employ.

4.1. Incremental loading

Iteration schemes require an "initial guess" solution at which to begin. The degree of nonlinearity in equation (24) or (25) is, to a large extent, dependent on the intensity of the loads as reflected in the magnitude of the load terms. A useful strategy in producing an initial guess at the desired load-level is to begin from the unloaded state with $u_i \equiv 0$, and increment the load gradually to the desired level. At each step of this "incremental loading" process [21,22], an approximate solution is generated which is an initial guess for the next load-level. When the final load level is achieved, the approximate solution at that load can be used as an initial guess for another iterative method which produces the final, refined solution. The strength of this strategy lies in the fact that for small enough load increments, the approximate solution at each load stays in a region of convexity of the energy. Furthermore, if the errors which accumulate during the loading process tend to cause departure from the region of convexity, the user has three options: (1) refine the load increment; (2) use the other iterative method to eliminate the accumulating errors at intermediate load-levels; or (3) combine (1) and (2).

Let u represent the nodal-value vector of displacements in (24) or (25). Let the strain energy terms in these equations be denoted by the vector $P(u)$ and the load terms by F. If $0 \leq \lambda \leq 1$ is a scalar, at any stage of the incremental loading process we solve the nonlinear equation

$$P(u + \Delta u) + (\lambda + \Delta\lambda)F = 0 \tag{81}$$

where $\Delta\lambda F$ is the load increment, and Δu is the change in the displacement field in moving from solution u at load-level λF to solution $u + \Delta u$ at load $(\lambda + \Delta\lambda)F$. Expanding $P(u + \Delta u)$ in a multivariate Taylor series and retaining only first-order terms in Δu, along with the definition of $\mathbf{K}_T = \mathbf{K}_T(u)$ in (27) leads to

$$P(u) + \mathbf{K}_T(u)\Delta u + \lambda F + \Delta\lambda F = 0 \tag{82}$$

If u is the exact solution at the previous load-level or u has been obtained by a previous step of incremental loading, $P(u) + \lambda F = 0[(\Delta\lambda)^2]$. This gives an algorithm with $0(\Delta\lambda)$ accumulated error:

$$\left\{\begin{aligned} u^{(0)} &= 0, \qquad \lambda_0 = 0 \\ \lambda_i &= \lambda_{i-1} + \Delta\lambda \\ \Delta u &= -\Delta\lambda \mathbf{K}_T^{-1}(u^{(i-1)})F \\ u^{(i)} &= u^{(i-1)} + \Delta u \end{aligned}\right. \tag{83}$$

The iterations in (83) can be carried out until $\lambda = 1$, or interrupted before then so that the current $u^{(i)}$ can be replaced by a refined $u^{(i)}$. The algorithms which do the refinement (such as the Newton-Raphson method described in the next section) tend to be more costly than one step of incremental loading. Therefore in the attempt to remain in the region of convexity, a judicious balance between load-increment, $\Delta\lambda$, and number of intermediate refinements may be called for. In our experience with the Mooney-Rivlin constitutive law, convexity can be maintained with very large $\Delta\lambda$, so that no intermediate refinement is needed, in the 2-D plane-stress model. On the other hand, efficient computations in the 3-D case require a careful balance between choice of $\Delta\lambda$ and the number and frequency of intermediate refinements. This is due to the sensitivity of the incompressibility constraint equations to small volume changes [9,10].

4.2. Newton-Raphson iterations

Let λF denote the current load-level, for $0 < \lambda \leqslant 1$. If u is an initial guess at this load-level, obtained either by incremental loading or other means, we wish to compute a correction, Δu, so that

$$P(u + \Delta u) + \lambda F = 0 \tag{84}$$

Again making a Taylor series expansion for $P(u + \Delta u)$ and neglecting second order terms in Δu yields the Newton-Raphson method:

$$\begin{cases} u^{(0)} = u \\ \Delta u = -\mathbf{K}_T^{-1}(u^{(i-1)})[P(u^{(i-1)}) + \lambda F] \\ u^{(i)} = u^{(i-1)} + \Delta u \end{cases} \tag{85}$$

The neglect of second order terms in (85) means that several iterations may be needed. It can be shown that in any vector norm $\|u^{(i+1)} - u^{(i)}\| = 0(\|u^{(i)} - u^{(i-1)}\|^2)$ [27], if $u^{(i-1)}$ is "sufficiently close" to the exact solution. This means that convergence can be very rapid. But it also implies that no prior guarantees can be made if $u^{(i-1)}$ is far from the exact solution. The algorithm can get lost in a nonconvex region, even if it begins with u in the desired region of convexity. We have found that with the 2-D problems, almost any u obtained by incremental loading (even with $\Delta\lambda$ large) is sufficiently close, right from the start of (85), to initiate quadratic or near quadratic convergence. Furthermore, even if u is far enough from the solution so that the ultimate quadratic convergence rate is not achieved from the start, the iterates of (85) converge more slowly until $u^{(i-1)}$ is close enough to give quadratic convergence.

In 3-D problems, we find that care must be taken in choosing $\Delta\lambda$ so that the initial guess, u, is accurate enough for (85) not to get lost. If this is done successfully then the ultimate quadratic convergence rate is guaranteed. The dominant computational cost in (85) is the assembly of $\mathbf{K}_T$. 2-D problems are so cheap and Newton's method so robust that there is little to be gained over the employment of (83) and (85) as described here. But for 3-D problems with penalties, our experience indicates that the sensitivity of the iterative schemes and the dominant computational cost of assembling $\mathbf{K}_T$ warrant the investigation of other iteration schemes. A likely candidate may be "modified Newton's method" in which $\mathbf{K}_T$ is reassembled less often than once every step [21,22].

4.3. Interpolatory mesh refinement

Another way to generate an initial guess either for the next load increment in (83) or to initiate (85) is to take a solution to (24) or (25) which is essentially fully converged on a given mesh and interpolate that solution to a finer mesh. This is done by taking the values given by (12) at the new nodal locations to define nodal values of a new trial function in the form (12). Of course, the new trial function has more degrees of freedom, so in general it will not represent an equilibrium solution to (24) or (25). The hope is that it will be a good initial guess to such a solution. A strategy which has proved successful in our problems is to combine load increments (83) and any required refinements by Newton's method (85) on a crude mesh until $\lambda = 1$. At the final load-level we refine the mesh in several steps—each generating an initial guess for the next which is refined by Newton's method. We have not performed a critical comparison of this strategy with others which are possible. The interpolatory refinements could be mixed with iterations of (83) or (85) in a variety of ways, and it would take careful study to determine the most efficient strategy. The reason we have used this particular strategy is that it provides a sequence of refinements at the final load-level which can give some idea of convergence of the model as the mesh is refined.

Again we find that the 3-D model is much more sensitive to the mesh refinement scheme than the 2-D model, because of the penalty. The sensitivity is reflected in the fact that the interpolated initial guess is often outside of the region of convexity. As described in [10], this can be traced to the fact that volumetric integration points, ζ_r, of (25) and (29) on the refined mesh do not overlay the ζ_r on the cruder mesh. This leads to inaccurate computation of the pressure, h, via (41). A smoothing technique for h is described in [10],

which is based on the equivalence theorem of [8]. It is employed as a standard feature of the 3-D code and appears to work as long as the refinement is not too drastic.

4.4. The in-core solvers

The first priority in the choice of linear equation-solving procedures for the systems in (83) or (85) is that they take advantage of the symmetric, band-matrix structure of $\mathbf{K}_T$, described in (30). We store $\mathbf{K}_T$ as $d(\omega + 1)$ vectors of length dN. Since it is hoped that the application of the procedures outlined earlier in this section will lead to positive-definite matrices, we employ elimination procedures with no interchanges [23,27–29]. There are several algorithms which can reduce the matrix to a compact triangular form which overlays the original matrix bands and require no additional storage space [23,27–29]. We employ two such algorithms. The first is the Cholesky method [23,27–29]; this will only work for positive-definite matrices. If any iterate of (83) or (85) departs from the convex region, the Cholesky method will fail, the iterations will cease, and our program will produce a suitable error message.

The Cholesky method is used unless it is determined that there is no reasonable load-refinement strategy to avoid departure from the convex region. In that case we employ the symmetric column elimination procedure described in [23]. This will usually provide the desired triangular decomposition, though, strictly speaking, the existence of the decomposition is not guaranteed without the use of interchanges [23,28]. Interchanges are not allowed if the band structure of the decomposition is to be preserved, and violation of band structure is incompatible with the goal of maximizing the size of $\mathbf{K}_T$ which can be decomposed in-core. Even when the decomposition without interchanges exists for an indefinite $\mathbf{K}_T$, the decomposition process is potentially unstable in a numerical sense [28]. When this difficulty is combined with the tendency of Newton's method to become lost if convexity is lost, it means that the symmetric elimination must be used with caution and as a last resort, when convexity cannot be maintained. If the iteration procedures (83) and (85) survive the loss of convexity, and the iterates converge, eventually convexity must be regained [26]. When this happens, then there is no essential difference numerically or theoretically between the symmetric elimination and the Cholesky decomposition, as long as convexity is maintained. We find that in 3-D problems, we often cannot avoid loss of convexity in the iteration scheme. On the other hand, if the load increments, $\Delta\lambda$, are not too large and the mesh refinements not too drastic, then the loss of convexity can be survived and solutions obtained. Therefore the symmetric elimination provides an indispensible way of recovering from loss of convexity, which would be impossible with the Cholesky method.

5. The stretching of rubber sheets

To illustrate the type of problem the model described here is designed to solve, we give an example of the stretching of a rubber sheet. This is similar to a problem described in [21], except that we have used the geometry of an A.S.T.M. standard tensile-test specimen for rubber sheet [30]. Our major concern was to investigate the application of penalty methods to impose the incompressibility constraint; therefore the emphasis in the results reported here is on the success or failure of the modelling process. However, we have been able to make some observations—largely qualitative—about the nature of the strain field in such specimens.

5.1. The A.S.T.M. tensile-test specimen

Figure 7 illustrates a top view of the A.S.T.M. tensile-test specimen described in [30] for the testing of the failure strength of rubber sheets. The dimensions are given in the lower portion of the figure. These specimens are referred to as "dumbbells" by the A.S.T.M., but this does not imply axisymmetry. Rather, the specimens are cut from a sheet which is typically one to several millimeters thick. Therefore, the plane-stress assumption described earlier seems appropriate. The 3-D model assumes a thickness of 2.5 mm.

In the upper portion of figure 7, a fairly crude irregular mesh is depicted. Experimentation with various mesh subdivisions allows concentration of subdivision in areas where spatial variations of strain is the largest. The large elements are placed in regions where the deformation is close to simple extension, which can be represented exactly, even by large elements of the type described in section 3.3.

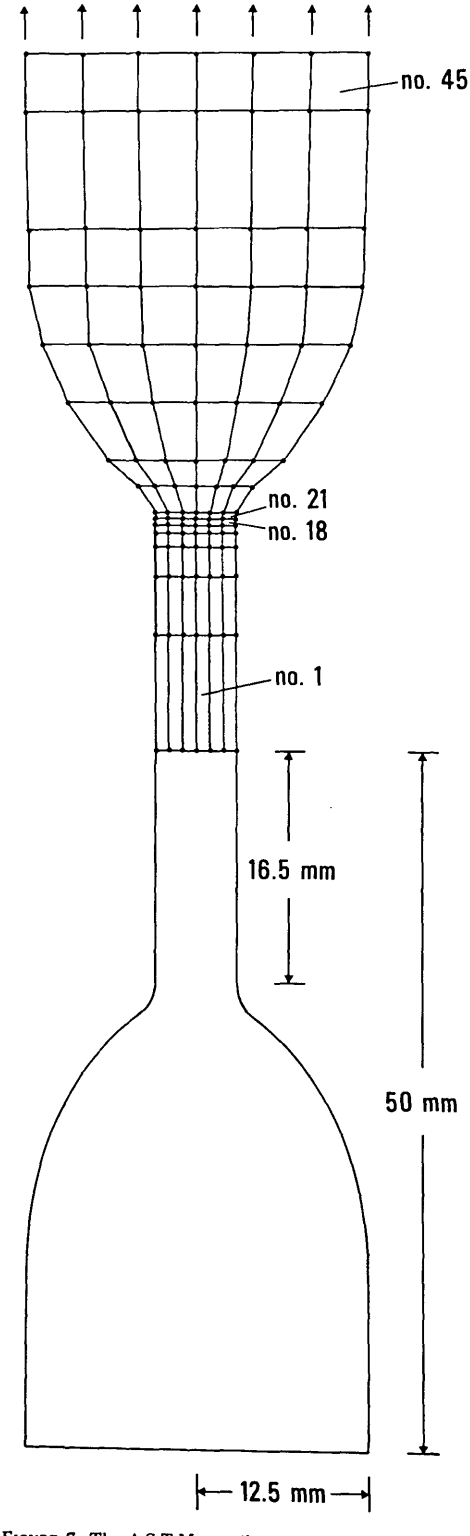

FIGURE 7. The A.S.T.M. tensile test specimen, lower half showing Ω with exact dimensions. Upper half shows a typical non-uniform finite element mesh giving Ω_e.

5.2. Two- and three-dimensional models

a. Symmetry

As has been illustrated in figures 5 and 6, there are certain symmetry assumptions in the current implementation of our model. Both 2- and 3-D versions assume symmetry of the deformation between the upper and lower portions of figure 7, as implied by figures 5 and 6. Also as implied by figures 5 and 6, symmetry of solutions between right and left halves of figure 7 is assumed, so that in the 2-D case, only the upper right quadrant is discretized. In the 3-D model symmetry is assumed along the midplane as implied by figure 6, so that only one octant of the body is actually discretized.

b. Boundary conditions

The boundary conditions along symmetry lines and planes are enforced as they are in linear analysis [22,23]. Displacements into, and certain shears along the symmetry lines and planes are assumed to be zero. The physical boundary conditions are a fairly simplified idealization of possible attachments to physical gripping devices (for example, see [31]). The conditions amount to a simple specification of end displacements to give various total extensions of the specimen. In 3-D, this roughly corresponds to gluing the whole end of the specimen to a relatively rigid surface. In the 2-D model, since the only end displacement available is on the midplane, the 2-D boundary condition is roughly like gluing the centerline of the end of the specimen to a rigid surface. The rest of the end is free to slip on the support. The 2- and 3-D physical boundary conditions are illustrated in figure 8.

More realistic boundary conditions and interfacing with gripping devices can be envisaged, and could be incorporated in the model. In [32] methods for dealing with more realistic grips are described. However, for the purposes of testing penalty methods we did not feel that such considerations were crucially important.

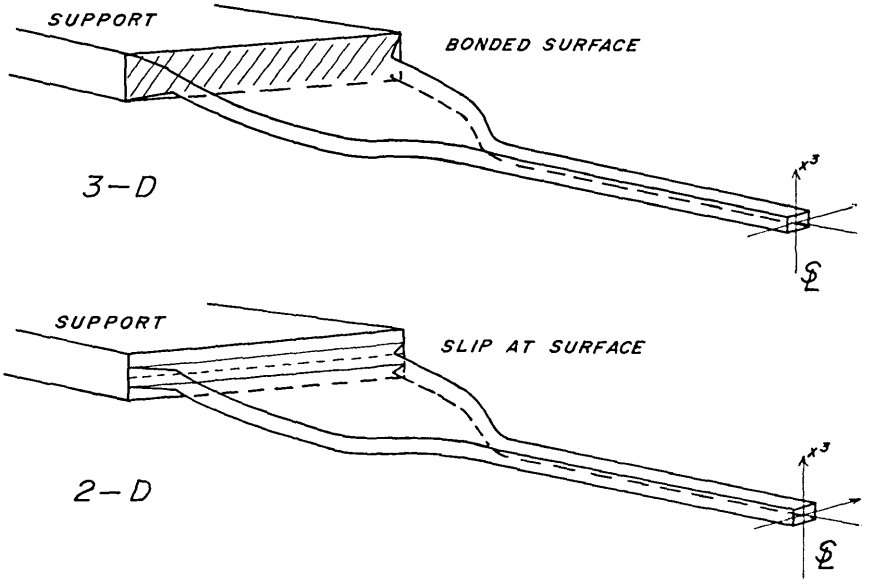

FIGURE 8. Boundary conditions for 3-D (upper)—bonding to a rigid support, and for 2-D (lower)—bonding of centerline only.

c. The use of penalties

The penalty method of section 2.2 was employed to impose near incompressibility. The Mooney-Rivlin law (52) was used with $c_2/c_1 = 0.2$ and $z = 50\ c_1$. The actual value of c_1 serves only to scale the stresses, so c_1 can be used to give appropriate dimensional units. This value of the penalty leads to a typical compressibility error of about 2-4 percent. For example, the most highly refined 3-D mesh of figure 7 has three elements

across the half-width and fifteen along the length and one through the thickness ($3 \times 15 \times 1$). For the element in this mesh with the highest strain component, III = 1.0344 instead of III = 1.

A penalty is also used to enforce the displacement boundary condition on the end of the specimen [14]. If Γ_d denotes the surface of the specimen which is bonded to the support, then the following term is appended to the variational principle (6):

$$z_d \int_{\Gamma_d} (u_i - u_i^*)(u_i - u_i^*) ds \quad \text{(sum on } i\text{)} \tag{86}$$

where z_d is the boundary penalty and u_i^* is the specified displacement vector. In practice it has been found that z_d can be taken quite large. In the current examples we have taken $z_d = 10^4$. The boundary integral in (86) is easy to evaluate [22,23]. Employment of this boundary penalty means that all essential b.c. are homogenous [22,23].

5.3. Code verification procedures

Because exact solutions to the problem we describe here are not available, special attention must be paid to insuring that the computer code is functioning correctly. Of course, it is not really possible to give an ironclad guarantee of the correctness of the code, but we have used three tests which assure us, beyond reasonable doubt, that the results we present here are artifacts of the model and not coding blunders.

a. The small-strain limit

As described in section 3.4, our code has the capability to produce linearly elastic solutions with the same geometry as for the nonlinear results. A simple asymptotic analysis of the Mooney-Rivlin law and the variational principle (6) using that law shows that when squares of the components of $G_{ij} - \delta_{ij}$ are small enough to be neglected, linear elasticity and nonlinear elasticity should give nearly the same results. We found that our nonlinear code reproduced the 2- and 3-D linear displacements to 4 digits when the total extension of the specimen was on the order of 0.5 percent. The differences between the two solutions were typically in the second digit for extensions in the range of 1-2 percent and progressively more noticeable for higher extensions. These results are in total accordance with what is to be expected from the asymptotic analysis.

b. Quadratic convergence of Newton's method

The fact that the small-strain limit is obtained from the nonlinear code does not guarantee that the code is working correctly. If an error is made in the coding of energy-density terms of higher order in displacement-gradients, this might not affect the small-strain limit. In the course of debugging this code and other nonlinear codes, it has been our experience that the quadratic convergence of Newton's method described in section 4.2 can only be obtained if $\mathbf{K}_T(u)$ is derived and coded correctly. What the obtaining of quadratic convergence means is that the derivatives computed according to section 2.3 are the correct second derivatives (29) of the first derivatives (24) of the variational principle (6). It does not completely assure that the first derivatives are derived correctly. The fact that, when solutions are obtained, $\mathbf{K}_T(u)$ is invariably positive definite suggests that the first derivatives are also correct.

Table 1 shows the obtaining of quadratic convergence in a 2-D mesh (pt. A.) and a 3-D mesh (pt. B.). Part C. of table 1 shows an example of a sequence of iterations for a 3-D mesh becoming "lost." All values in the table are the square roots of the sum of the squares of the components of F_i^λ. The 2-D results are from a 7×26 element mesh discretizing the first quadrant of the specimen, extended to a total extension of 200 percent. Ten load increments (83) of 20 percent were performed on a regular 3×5 mesh with no intermediate refinement. Then the 3×5 fully extended displacements were refined by Newton's method and interpolated to a 4×7 regular mesh. The same procedure was followed to proceed to a 5×10 mesh, a 5×12 mesh, then a 6×15 mesh, and finally these results were interpolated to the irregular 7×26 mesh. It is the refinement by Newton's method of these last interpolated results which is given in table 1A.

TABLE 1. *Quadratic Convergence of Newton's method as reflected in the decline of the rms residual of the equilibrium equations A and B. C shows iterations becoming lost.*

	A.	B.	C.
iter. no.	7 × 26 2-D	3 × 15 × 1 3-D	3 × 8 × 1 3-D
1	0.422×10^{1}	0.551×10^{3}	0.900×10^{3}
2	0.588×10^{0}	0.271×10^{2}	0.794×10^{2}
3	0.262×10^{-1}	0.665×10^{1}	0.529×10^{5}
4	0.652×10^{-4}	0.883×10^{0}	0.370×10^{8}
5	0.627×10^{-9}	0.108×10^{0}	0.130×10^{8}
6	0.445×10^{-14}	0.585×10^{-3}	0.176×10^{8}

Table 1B is obtained from the performance of 10 load increments of 10 percent on a $3 \times 8 \times 1$ element 3-D mesh, to give a total extension of 100 percent. Newton's method was used to refine the results after every two load increments. Interpolatory refinement with Newton iterations was used to go to a $3 \times 10 \times 1$ and then a $3 \times 12 \times 1$ regular mesh. Finally, the results were interpolated to the irregular $3 \times 15 \times 1$ mesh whose top view is given in figure 7. It is the refinement by Newton's method of these last interpolated results which is given in table 1B. The $\mathbf{LDL}^{T}$ column solver is required for the procedures involved in table 1B, since convexity is lost after the first pair of load increments. It is always regained by the Newton iteration, but is lost at the next load increments, and just after each interpolatory refinement.

Table 1C shows what happens when an attempt is made to carry out the load increments described for table 1B on a $3 \times 5 \times 1$ mesh instead of a $3 \times 8 \times 1$ mesh. The loss of convexity is so severe that when an attempt is made to interpolate the displacements to a $3 \times 8 \times 1$ mesh and refine by Newton's method the results of table 1C are obtained. The ability to obtain quadratic convergence with a careful choice of load increments leads us to believe that the code is correct, but that remaining close enough to the convex region is a delicate matter in the 3-D model.

c. Convergence with mesh refinement

Figure 9 shows values of the extensional strain J_{22} and shear strain J_{21} for a specimen extended to a total extension of 100 percent. The results are taken from a 2-D calculation with 26 elements along the length, with smaller elements concentrated at the top of the neck. The smallest elements are on the order of a quarter the size of the smallest elements in figure 7 and the larger elements of the refined 7×26 mesh are in the simple extension portion of the neck and are half the length of the corresponding elements in figure 7. Figure 9 plots values of J_{22} and J_{21} at a position indicated by the arrow on the cruder 3×15 mesh (both halves shown) of figure 10. This is where the strain-gradients are largest. With the same lengthwise subdivision of the specimen into 26 pieces, results are compared for meshes which have 3, 4, 6 and 7 elements across the neck. The 7×26 mesh is obtained from the 6×26 by subdividing the outer element in two. Also plotted on figure 9 are the results of the $3 \times 15 \times 1$ 3-D mesh, whose top view is given in figure 7.

The results show that to a large extent, the results from all the meshes overlay the same curve, with the largest deviations coming—as expected-from the $3 \times 15 \times 1$ and 3×26 meshes. Therefore we are led to conclude that convergence is taking place and that, qualitatively at least, even the cruder meshes give good results to graphical accuracy. A similar conclusion can be drawn from a comparison of tables 2 and 3 in which values of the pressure H of (70) for a 7×26 and 3×15 mesh are compared—likewise for a total extension of 100 percent.

We believe that the circumstantial evidence gathered from the obtaining of the small-strain limit, quadratic convergence of Newton's method, and convergence with mesh refinement make it highly unlikely that our numerical results are an artifact of coding blunders. We have run exactly parallel two- and three-dimensional versions of the 3×15 mesh of figure 7 ($3 \times 15 \times 1$ in the 3-D case). While we have not run as many 3-D convergence tests as 2-D convergence tests, we find that the agreement in displacements and strain is in general as good as the agreement illustrated in figure 9. We are convinced that convergence of displacements with mesh refinement is taking place in the 3-D model. What is happening to the pressures in table 4 is another—and very interesting—matter, which will be discussed below.

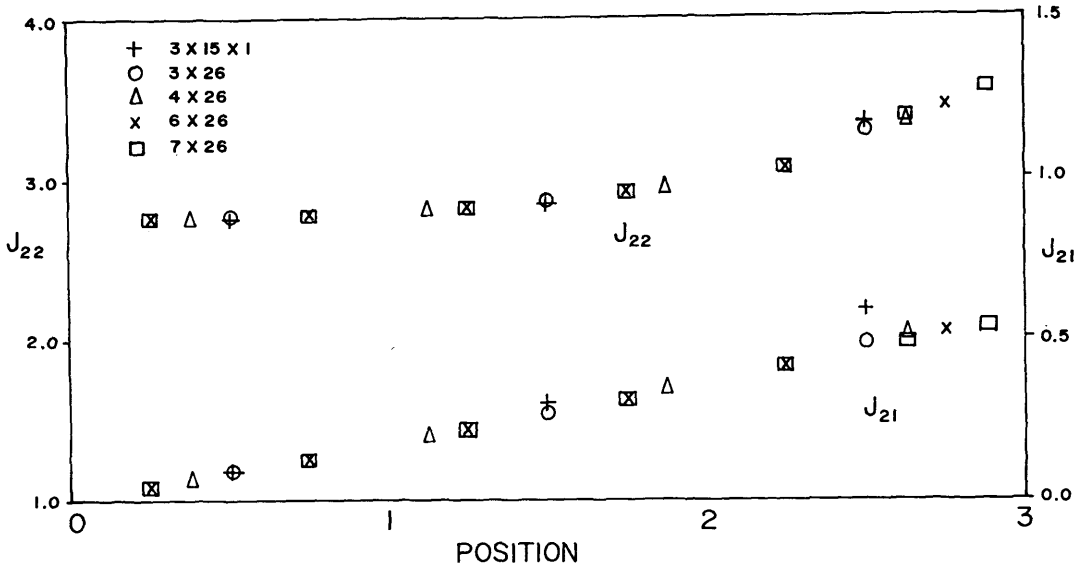

FIGURE 9. Extensional strain (upper) and shear-strain (lower) sampled at element centroids in stress-concentration region for various meshes. Values given on cross-section of neck.

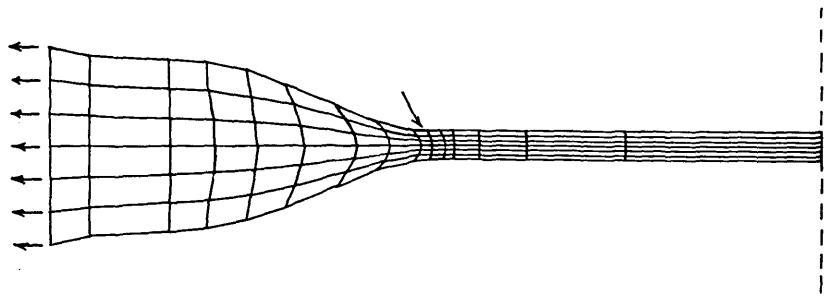

FIGURE10. Mesh of figure 7 deformed to total extension of 100 percent. Arrow marks stress-concentration.

TABLE 2. *Piecewise constant pressures along a cross-section of the neck in the stress-concentration region. Values taken from a refined 2-D mesh (7 × 26). Positions taken at center point of individual elements.*

Position	0.25	0.75	1.25	1.75	2.25	2.625	2.875
Pressure	0.537	0.533	0.523	0.508	0.480	0.452	0.438

TABLE 3. *Similar values to table 2, taken from a crude 2-D mesh.*

Position	0.50	1.50	2.50	Average value
Pressure	0.534	0.514	0.498	0.515

TABLE 4. *Values of the pressure taken from a 3-D mesh along cross-section near the stress-concentration region.*

Position	0.50	1.50	2.50	Average value
Pressure	0.621	0.454	0.563	0.546

5.4. Discussion of results

a. The plane-stress results

Our experience shows that the use of Newton's method with incremental loading described in section 4., when combined with the isoparametric plane-stress model described in this paper, produces a robust and computationally inexpensive algorithm. However, we must point out that the geometric assumptions and boundary conditions used in the current study advise a cautious interpretation of our results. We do think that we can draw some broad conclusions about the behavior of the A.S.T.M. speciments based on the qualitative picture which emerges from our 2-D results.

The picture which emerges is that the A.S.T.M. specimen geometry does what it is intended to do: namely concentrate the extension of the specimen in the neck region and away from the grips. The geometry does produce a region of concentrated deformation at the position of the arrow in figure 10, but our results indicate that the concentration is not severe. The data we cite to support these points, some of which are plotted in figure 9, are taken from our 7×26 mesh—not the cruder mesh of figure 10.

To illustrate the concentration of extension in the neck, we look at the value of the extension ratio, J_{22}, in the element nearest the center of the neck as compared to the total extension ratio of the specimen (based on grip separation). For 10 percent total extension of the specimen, the neck value of J_{22} is 7.2 percent higher than the total extension ratio of 1.10. For 100 percent total extension, the neck has a J_{22} 51 percent higher than the total extension ratio 2.00, and for total extension of 200 percent, J_{22} is 67 percent higher on the neck than the total extension ratio of 3.00.

Turning to the region of concentrated deformation, we find that there is a maximum of J_{22} at the point indicated by the arrow in figure 10. For 10 percent total extension, the maximum extension ratio is 5.5 percent higher than the extension ratio in the neck. For 100 percent total extension, the maximum extension ratio is 10 percent higher than in the neck, and for 200 percent total extension, the maximum is 20 percent higher than in the neck.

The extension ratio in the region of concentrated deformation seems to approach a value of 20 percent above the neck value in the limit as total extension is increased. The extension ratio J_{22} is not a rotationally invariant quantity, and one might raise the question as to whether the concentration of deformation grows unboundedly with total extension ratio by increasing shear-strain concentration. We look at the energy-density, which from (50), (51) and (64) can be seen to be $W(\mathrm{I}, \mathrm{II}) = c_1(\mathrm{I}-3) + c_2(\mathrm{II}-3)$, if one requires that the energy-density of the undeformed state is zero. Again we compare neck values of W with the maximum value. For 10 percent total extension W is 92 percent higher at its maximum than in the neck; for 100 percent total extension, it is 57 percent higher; and for 200 percent total extension, it is only 53 percent higher. So the deformation-energy concentration actually decreases in intensity with increasing extension ratio. Thus the shear-strain intensity cannot be growing unboundedly. In fact, if we look at the quantity $J_{21}/(J_{22}-1)$, we find that for 10 percent total extension, the maximum value is 27 percent above the value in the neck. For 100 percent total extension, the intensity is 21 percent over the neck value, and for 200 percent total extension, the intensity is 19 percent over the neck value. So in fact shear-deformation appears to become relatively less important as the total extension increases.

The A.S.T.M. standard [30] specifies that the total extension of the neck be measured visually, using marks on the neck. If the results we present here represent an adequate qualitative picture of the specimen, we would expect the specimen to fail in the neck, just below the wide portion. The A.S.T.M. standard requires that neck extension be recorded at the moment of failure. Our results indicate that there is a built-in safety factor of about 20 percent, in that the actual extension ratio near the point of faliure could have been about 20 percent higher than recorded. Shear deformation near the point of such a failure would not appear to be a major factor.

b. The three-dimensional results

For reasons discussed in detail earlier, the 3-D incremental loading/Newton algorithm combined with the isoparametric finite element/penalty model described here cannot be classified as a robust or computationally inexpensive algorithm. One would expect 3-D nonlinear finite element models to be dramatically more expensive than 2-D models in general. But one would also hope that the nonlinear iteration scheme could be made more reliable, so that situations like the one illustrated in table 1C do not lead to the waste

of expensive iterations. But there is a more basic theoretical question raised by the 3-D results, which should be answered before much effort is spent on the optimization of iteration schemes. This is a question of the stability and accuracy of the penalty or equivalent mixed formulation [8, 13–16].

Table 5 shows values of the pressure in a cross-section of elements near the stress-concentration region of a 3 × 10 × 1 mesh. The elements are rather large, and the centroids from which the pressures are taken are further down the neck towards the middle of the specimen than the sample points of tables 2 and 3, therefore the pressure is lower. But note that there is a smooth variation of pressure. Ths contrasts sharply with the oscillation of pressure across the cross-section in the stress-concentration region shown in table 4. These results are taken from the irregular 3 × 15 × 1 mesh whose top view is given in figure 7. We have already seen that J_{22} and J_{21} vary smoothly and are reasonably accurate in these elements from figure 9. In general, a detailed comparison of the 3 × 15 × 1 results shows that u_i and J_{ij} compare well with those from the 2-D 7 × 26 mesh—particularly in comparison to the contrast of the results in tables 2 and 4. What we observe there is that the 3-D pressures alone are affected by a "checkerboard" pattern of high and low values in the stress-concentration region. This is illustrated in figures 11 and 12. The pattern of high-low-high on one cross-section is countered by a reverse pattern on adjacent cross-sections. The average pressure value along a checkerboarded cross-section seems to agree much better with average values taken from the 2-D model, as is illustrated in figure 12 and tables 3 and 4. This leads us to speculate that averaging techniques may be able to remove the checkerboard.

TABLE 5. *Values of the pressure taken from a cross-section near the stress concentration for a cruder 3–D mesh.*

Position	0.50	1.50	2.50
Pressure	0.446	0.443	0.421

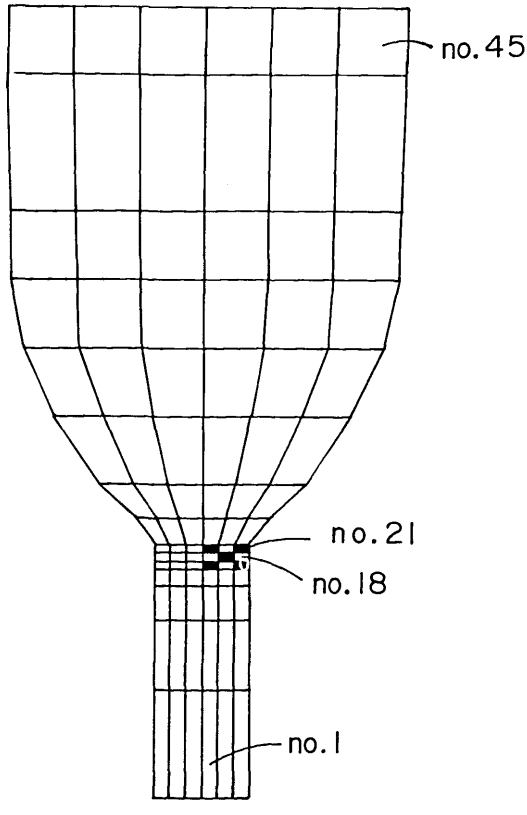

FIGURE 11. Top view of a 3-D mesh—qualitative picture of checkerboard pressures.

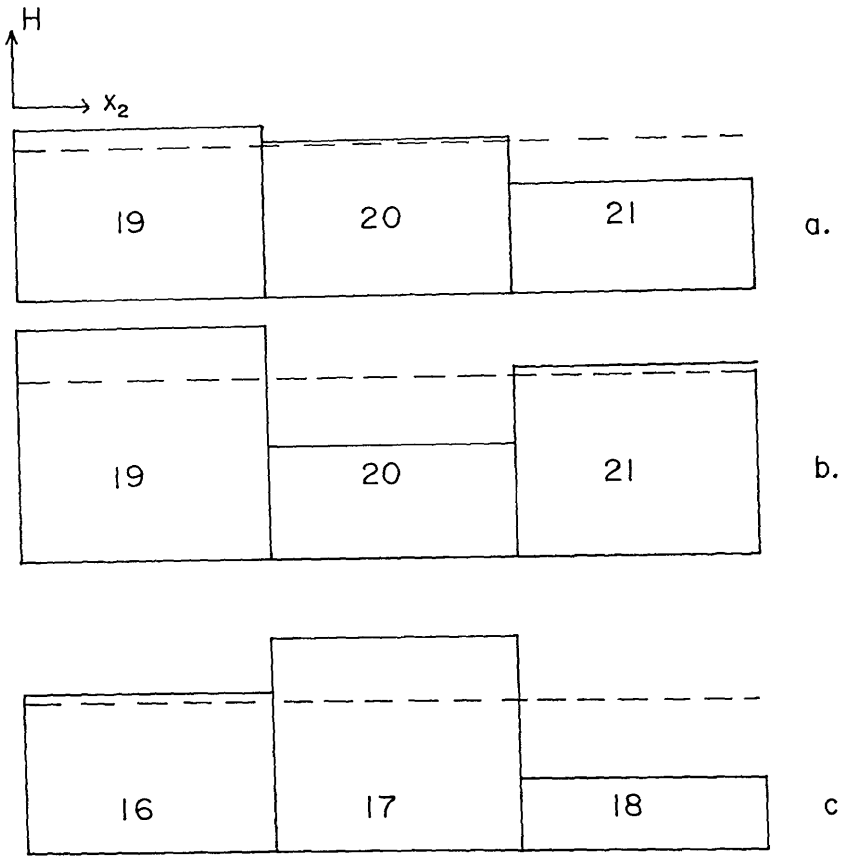

FIGURE 12. Piecewise-constant pressures on a cross-section of the neck: (a) Accurate pressures; (b) checkerboard pressures; (c) checkerboard variation on adjacent cross-section to (b).

In view of the success obtained using this trilinear isoparametric element in a variety of penalty or mixed formulations [2, 7, 8, 11], we are led to conclude that the checkerboard pressure problem is a relatively rare occurrence. But, particularly since this phenomenon occurs when the mesh is refined, we are led to doubt that our element can satisfy the basic stability requirement of [13–16]. This matter is under investigation by the first author and others [14, 15, 34]. Techniques to overcome these difficulties using modified penalty procedures are also being investigated [18, 19, 33, 34]. Though the appearance of checkerboard pressure keeps us from drawing any physical conclusions based on our 3-D results, we are convinced that it will prove extremely useful to future theoretical investigations to be able to reliably reproduce this checkerboard phenomenon in such a basically simple problem.

6. Conclusions

The purpose in the development of the finite element model described here is to generate test problems for research on penalty methods in finite elasticity. As such it has proved very successful. In spite of the simplicity of the model, it has been able to isolate two important areas which will require further investigation if finite element penalty methods are to become productive tools in the analysis of incompressible materials under finite strain. The first area is that of the sensitivity of iteration schemes to small volume changes. Progress has been made in this direction, but still more is required [9, 10]. Perhaps more important is the question of element stability raised by the appearance of checkerboard pressure modes. The stability condition is known [13–17], but few elements seem to satisfy it—whether in penalty or equivalent mixed formulations [8–10]. On the other hand, even for elements which apparently do not satisfy it, the

checkerboard modes may be removable and often do not occur. More precise mathematical predictions need to be provided for these phenomena. It would be very useful to be able to aid such theoretical investigations with a code with more general element capabilities than the current one. Clearly the model admits to a variety of element types. Beyond that, new developments indicate that there are other penalty formulations which may be more stable than the reduced/selective integration techniques incorporated in the current model [19, 33, 34]. Therefore the model should be enhanced to include these possibilities.

Finally it is tempting to think of ways in which to make the code itself more sophisticated. The mesh generator could be based on the unions of several reference bricks, making the variety of bodies describable much wider. Load increment/refinement strategies could be made adaptive, in order to maintain convexity as much as possible without human intervention. More physically realistic boundary conditions and non-conservative loads could be incorporated. Many other constitutive laws fit within the framework of the model and would require only minor coding to implement. However, these enhancements belong further down the road, after some of the challenging questions raised by the current model are resolved.

The finite element model and the computer code described here were developed while the first author was a Postdoctoral Research Associate at NBS. He would like to express his appreciation to that organization for the opportunity to carry on this research, which continues to be a fruitful avenue of investigation. He would like to express his gratitude to his advisors, E.A. Kearsley and J.T. Fong and thank Applied Mathematics Division Chief, B. Colvin, for his support and encouragement.

7. References

[1] Herrmann, L. R., Elasticity equations for incompressible and nearly incompressible materials by a variational theorem, AIAA J. **3** 1896-1900 (1965).

[2] Hughes, T. J. R. and Allik, H., Finite elements for compressible and incompressible continua, Proc. Symp. Civil Engng., Vanderbilt Univ., Nashville, TN, 27-62 (1962).

[3] Fried, I., Finite element analysis of incompressible material by residual energy balancing. Int. J. Solids Structs. **10**, 993-1002 (1974).

[4] Malkus, D. S., Finite Element Analysis of Incompressible Solids, Ph.D. Thesis, Boston University, 1975.

[5] Malkus, D. S., A finite element displacement model valid for any value of the compressibility, Int. J. Solids Structs., **12**, 731-738 (1976).

[6] Malkus, D. S., Calculation of hole error by finite element methods, Proc. 8th Int. Cong. on Rheology, 618-619 (1977).

[7] Hughes, T. J. R., Equivalence of finite elements for nearly incompressible elasticity, J. Appl. Mech. **44**, 181-183 (1977).

[8] Malkus, D. S. and Hughes, T. J. R., Mixed finite element methods—reduced and selective integration: A unification of concepts, Comp. Meth. Appl. Mech. Engng., **15**, 63-81 (1978).

[9] Malkus, D. S., Penalty methods in the finite element analysis of fluids and structures, Nuc. Engng. Design, **57**, 441-448 (1980).

[10] Malkus, D. S., Finite elements with penalties in nonlinear elasticity, Int. J. Num. Meth. Engng., to appear.

[11] Hughes, T. J. R., Liu, W. K. and Brooks, A., Finite element analysis of incompressible viscous flows by the penalty function formulation, J. Comp. Phys., **30**, 1-60 (1979).

[12] Bernstein, B., Kadivar, M. K. and Malkus, D. S., Steady flow of memory fluids with finite elements: Two test problems, J. Comp. Meth. Appl. Mech. Engng., to appear.

[13] Babuska, I., Oden, J. T. and Lee, J. K., Mixed-hybrid finite element approximations of second-order elliptic boundary-value problems, Comp. Meth. Appl. Mech. Engng., **11**, 175-206 (1977).

[14] Bercovier, M., Perturbation of mixed variational problems. Application to mixed finite element methods, R.A.I.R.O. Analyse numerique/Numerical Analysis, **12**, 211-236 (1978).

[15] Oden, J. T., A theory of penalty methods for finite element approximations of highly nonlinear problems in continuum mechanics, Comp. Structs. **8**, 445-449 (1978).

[16] Malkus, D. S., Incompressible elements: The LBB condition and the discrete eigenstructure, Proc. U.S.-Europe Workshop on Finite Elements in Nonlinear Structural Mechanics, Springer-Verlag (1980).

[17] Hughes, T. J. R., Current trends in finite element research, Appl. Mech. Rev., to appear.

[18] Nagtegaal, J. C. and de Jong, J. E., Some computational aspects of elastic-plastic large strain analysis, research report M.S. 10.910, Technical University, Eindhoven, The Netherlands (1979).

[19] Cescotto, S. and Fonder, G., A finite element approach for large strains of nearly incompressible rubber-like materials, Int. J. Solids Structs. **15**, 589-605 (1979).

[20] Oden, J. T., Existence theory for a class of problems in nonlinear elasticity, J. Math. Anal. Applics. **69**, 51-83 (1979).

[21] Oden, J. T., *Finite Elements of Nonlinear Continua*, McGraw-Hill, New York (1972).

[22] Zienkiewicz, O. C., *The Finite Element Method*, Third Edition McGraw-Hill, London (1977).

[23] Bathe, K. J. and Wilson, E. L., *Numerical Methods in Finite Element Analysis,* Prentice-Hall, Englewood Cliffs, NJ (1976).
[24] de Boor, C., *A Practical Guide to Splines,* Springer-Verlag, New York (1978).
[25] Fung, Y, C., *Foundations of Solid Mechanics,* Prentice-Hall, Englewood Cliffs, NJ (1965).
[26] Gel'Fand, I. M. and Fomin, S. V., *Calculus of Variations,* Prentice-Hall, Englewood Cliffs, NJ (1963).
[27] Dahlquist, G. and Bjorck, A, *Numerical Methods,* Prentice-Hall, Englewood Cliffs, NJ (1974).
[28] Wilkinson, J. H., *The Algebraic Eigenvalue Problem,* Clarendon Press, Oxford (1965).
[29] Schwarz, H. R., Rutishauser, H., and Stiefel, E., *Numerical Analysis of Symmetric Matrices,* Prentice-Hall, Englewood Cliffs, NJ (1973).
[30] Standard test methods for rubber properties in tension (A.S.T.M. Standard Test Method D412-75), in 1976 Annual Book of ASTM Standards, Part 37, Rubber, Natural and Synthetic—General Test Methods; Carbon Black, American Society for Testing and Materials, Philadelphia, 79-80 (1976).
[31] Dewey, B. R., Finite element analysis of creep and plasticity tensile-test specimens, Experimental Mech. **16**, 16-20 (1976).
[32] Mitchell, R. A., Woolley, R. M., and Halsey, N., High-strength end fittings for FRP rod and rope, J. Eng. Mech. Div., ASCE, **100**, 687-706 (1974).
[33] Hughes, T. J. R., Generalization of selective integration procedures to anisotropic and nonlinear media, Int. J. Num. Meth. Engng., to appear.
[34] Hughes, T. J. R. and Malkus, D. S., manuscript in preparation on a generalized mixed method/penalty method equivalence theorem.