PERGAMON

International Journal of Solids and Structures 37 (2000) 735–760

INTERNATIONAL JOURNAL OF SOLIDS and STRUCTURES

www.elsevier.com/locate/ijsolstr

A computational model for shape memory alloys

Sanjay Govindjee*, Garrett J. Hall

Structural Engineering, Mechanics, and Materials, Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720-1710, USA

Received 23 August 1998; in revised form 2 February 1999

Abstract

An investigation into the computational aspects of a multi-well mixture approach to shape memory modeling is undertaken with the goals of determining its qualitative behavior as well as its efficiency in a numerical setting. A basic rate dependent model for the transformation is first introduced, followed by a discussion of the steps taken to implement the constitution in discrete form. Numerical simulations demonstrate the quantitative and qualitative response of shape memory alloy structural systems to various thermal and mechanical cycles. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Shape memory effect; Martensitic transformation; Phase transformation

1. Introduction

Although there is a significant body of literature treating various aspects of transformations in shape memory alloys, the phenomena is sufficiently complex such that any computational treatment attempting to capture the macroscopic response of a structural system composed of these materials necessarily requires significant simplifications to result in a tractable problem. Thus, the majority of previous attempts within the *computational* arena have followed a phase space approach which, with the advent of robust integration schemes, has proven to be an efficient technique for modeling one-dimensional constitutive response (for examples, see Brinson and Lammering, 1993; Brinson and Bekker, 1997; Govindjee and Kasper, 1997, 1999). However, the primary disadvantage of such an approach is that it is not readily apparent how one should extend the model to capture multidimensional material response; see Lubliner and Auricchio (1996) for one possible method. To overcome this difficulty, in this paper we investigate the computational aspects of a model derived primarily from the work of Achenbach (1989), Achenbach and Müller (1985), Müller and Xu (1991), Abeyaratne et al. (1994), and Abeyaratne and

* Corresponding author. Fax: +1-510-643-8928.
E-mail address: sanjay@ce.berkeley.edu (S. Govindjee)

PII: S0020-7683(99)00048-7

Kim (1997). This constitutive model, which is generically referred to in the literature as a multi-well approach, uses the principles of statistical physics to estimate the transformation rates between variants in a shape memory alloy. As such, the evolution of the system depends upon energetic considerations which are valid independent of dimension, although in the present paper we limit our investigation to a one-dimensional application. Rather than attempting to improve upon the models suggested in the literature, the primary goal of this work is to investigate the feasibility of implementing multi-well models in conjunction with mixture theory within a computational setting such that complex structural systems can be analyzed with reasonable efficiency.

An outline of the presented work is as follows. In Section 2, the kinematic and essential constitutive assumptions are presented in a general setting. A concrete example is then given in Section 3, where constitutive equations for a three variant (one dimensional) shape memory alloy are specified. This leads into a discussion of the computational aspects of implementing the given model within a finite element setting, followed by Section 4 in which several simulations are presented which demonstrate the behavior of the model. Appendix A details some considerations on matrix exponentiation; Appendix B gives details of the model's algorithmic tangent; and Appendix C gives an extension of the model to plasticity.

2. Constitutive framework

2.1. Preliminaries

Let Ω_0 and Ω be the reference and spatial configurations of a bounded deformable continuum comprised of $\alpha \in \{1,2,3,\ldots,v\}$ continua moving in $\mathbb{R}^3$ space. Such a system may be idealized as a mixture in the classical sense (Atkin and Craine, 1976), where each constituent enjoys its own path in space, time, temperature, etc., whereas the history of the body is taken to be a functional of the constituent paths. Typically, each of the constituents is allowed a motion such that a spatial point is simultaneously occupied in the manner

$$\boldsymbol{x} = \boldsymbol{\varphi}^{\alpha}(\boldsymbol{X}^{\alpha}, t) \tag{1}$$

which leads naturally to the existence of constituent velocities $\mathbf{v}^{\alpha}$ and the concept of diffusion. However, for the purpose of simulating the macroscopic effects of changes in lattice structure under the assumption of first-order diffusionless displacive transformations, it is reasonable to consider a pointwise homogeneous mixture. Specifically, it is assumed there exists for each point $\boldsymbol{X} \in \Omega_0$ a spatial position $\boldsymbol{x} \in \Omega$, an absolute temperature θ, and a set of volume fractions ξ^{α}, the histories of which may be described by the sufficiently smooth functions

$$\boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{X},t) \quad \theta = \hat{\theta}(\boldsymbol{X},t) \quad \xi^{\alpha} = \hat{\xi^{\alpha}}(\boldsymbol{X},t). \tag{2}$$

In order to correctly simulate motions with large rotational components, the Green–Lagrange tensor ($\boldsymbol{E}$) is taken as the strain measure for all subsequent calculations. It is defined from the deformation gradient ($\boldsymbol{F}$) in the usual manner:

$$\boldsymbol{F} = \frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{X}} \quad \boldsymbol{E} = \frac{1}{2}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F} - \boldsymbol{I}). \tag{3}$$

2.2. Conservation of mass

At every point in the spatial continuum the volume fractions are defined by $\xi^\alpha = \lim_{V\to 0}(V^\alpha/V)$, from which is it apparent that the local volume fractions are subject to the restrictions

$$\sum_\alpha \xi^\alpha = 1; \quad 0 \le \xi^\alpha \le 1. \tag{4}$$

Then, assigning to each member of the mixture a reference density ρ_0^α which maps to a unique density ρ^α in the spatial configuration, one may compute the mass of the body, m, in the following manner:

$$m = \int_\Omega \sum_\alpha \rho^\alpha \xi^\alpha \, \mathrm{d}v. \tag{5}$$

If the mass flux across the bounding surface of the body $\partial\Omega$ is zero, then conservation of mass dictates

$$\dot{m} = \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \sum_\alpha \rho^\alpha \xi^\alpha \, \mathrm{d}v = 0. \tag{6}$$

Neglecting the possibility of diffusion in accordance with the previous kinematic assumptions, a single velocity vector $\mathbf{v}$ describes the motion at each point $\mathbf{v} = \hat{\mathbf{v}}(\boldsymbol{x}, t)$. Using this fact, the transport theorem as applied to the mixture is expressed as

$$\dot{m} = \int_\Omega \sum_\alpha [\dot{\overline{\rho^\alpha \xi^\alpha}} + \rho^\alpha \xi^\alpha \operatorname{div}(\mathbf{v})] \, \mathrm{d}v = 0. \tag{7}$$

Letting $J = \det(\boldsymbol{F})$ represent the Jacobian of the deformation gradient, we have the following basic relations

$$\rho_0^\alpha = \rho^\alpha J \quad \frac{\mathrm{d}}{\mathrm{d}t} J = \operatorname{div}(\mathbf{v}) J. \tag{8}$$

This allows Eq. (7) to be rewritten as

$$\dot{m} = \int_\Omega \sum_\alpha \rho^\alpha \dot{\xi}^\alpha \, \mathrm{d}v = 0 \tag{9}$$

from which it is concluded that at every point in the body $\Sigma_\alpha \rho^\alpha \dot{\xi}^\alpha = 0$. Of primary interest are materials in which $\rho^\alpha = \rho^\beta$ for all α, β, in which case conservation of mass reduces to

$$\sum_\alpha \dot{\xi}^\alpha = 0. \tag{10}$$

In what follows, Eqs. (4) and (10) serve as fundamental restrictions on the values ξ^α and transformation rates $\dot{\xi}^\alpha$ which the volume fractions may assume. Furthermore, the superscripts on the densities will be dropped, as we will assume equal densities.

2.3. Potential functions

In constructing a free energy function for the mixture described thus far, it is assumed that the individual behavior of each constituent is sufficiently captured by a thermoelastic material model. The

individual functions are denoted by

$$\psi^\alpha = \hat{\psi}^\alpha(\boldsymbol{E}, \theta) \tag{11}$$

where $\rho_0\psi^\alpha$ is the Helmholtz free energy per unit volume. In considering the free energy of the mixture, it is apparent that there are at least two sources which contribute to the total. The first is the contribution stored in the individual constituents, and the second is due to distortional energy accumulated at the interfacial boundaries. Denoting the second contribution by $\epsilon = \hat{\epsilon}(\boldsymbol{E}, \theta, \xi^\alpha)$, the total free energy may be approximated by the function

$$\Psi = \sum_\alpha \xi^\alpha \psi^\alpha + \epsilon \tag{12}$$

where it is clear that $\Psi = \hat{\Psi}(\boldsymbol{E}, \theta, \xi^\alpha)$. Previous authors on the subject of constitutive modeling for shape memory alloys have taken different approaches in considering an interfacial energy function. For example, Achenbach (1989) derived a form based on the probabilistic argument that the surface energy is maximal half way through a transformation between two variants, and hence the energy was approximated as a constant times the product of the relative fractions involved. On the other hand, Boyd and Lagoudas (1996a, b) chose to neglect the energy stored at the interface in their model. In order to simplify the computational aspects of simulating shape memory alloy behavior, we have initially chosen the latter approach, namely, to consider $\epsilon \approx 0$ for the present investigation. This leaves

$$\rho_0\Psi = \sum_\alpha \xi^\alpha \rho_0 \psi^\alpha \tag{13}$$

as the chosen Helmholtz free energy function for the mixture. One may then construct a Gibbs function by the following transformation

$$G = -\sup_E \{\boldsymbol{S}:\boldsymbol{E} - \rho_0\Psi\} \tag{14}$$

so that $G = \hat{G}(\boldsymbol{S}, \theta)$ and $\boldsymbol{S}$ is the effective second Piola–Kirchhoff stress measure for the mixture. This may also be rewritten as

$$G = -\sup_E \left\{ \sum_\alpha \xi^\alpha (\boldsymbol{S}:\boldsymbol{E} - \rho_0\psi^\alpha) \right\}. \tag{15}$$

A framework for modeling the transformation between constituents is considered next.

2.4. Transformation kinetics

While several methods have been suggested in the literature for modeling the transformation phenomena in shape memory alloys, we have chosen to follow the approach taken by Achenbach (1989) and later adopted by Abeyaratne et al. (1994). The approach is based on statistical physics, wherein it is postulated that the rate of transformation between constituents is proportional to the net probability that one phase will overcome the energetic barrier required to transform to a second phase. Thus, the thermally activated process will occur at a rate

$$\dot{\xi}^{\alpha} = \sum_{\beta}^{\beta\neq\alpha} \omega(P_{\beta\alpha}\xi^{\beta} - P_{\alpha\beta}\xi^{\alpha}) \tag{16}$$

where $P_{\alpha\beta}$ estimates the transition probability of constituent α transforming to constituent β, with an attempt frequency of ω. Since Eq. (16) is inhomogeneous in time it will lead to a model that displays dependencies on the loading rate. At this point we do not propose a specific form for $P_{\alpha\beta}$; this is addressed in the following section. Of current interest is the resulting structure, which by summation of rates is expressible in matrix notation as

$$\dot{\boldsymbol{\xi}} = \mathbf{Q}\boldsymbol{\xi}. \tag{17}$$

From an implementation viewpoint, it is noteworthy that properties of systems of differential equations which have this specific form have been treated extensively in the literature of probability and statistics, wherein **Q** is referred to as the infinitesimal generator of a Markov process (Stewart, 1994). It is also worth noting that Eq. (17) exactly satisfies the balance of mass constraint given in Eq. (10).

In the following section, specific functions are chosen for the constituent free energy functions ψ^{α}, and the transformation probabilities $P_{\alpha\beta}$, thereby closing the set of constitutive equations.

3. Example: A three variant shape memory alloy

With the underlying assumptions described in Section 2, we now consider applying the theory to problems which can reasonably be characterized as one-dimensional with regard to constitutive response. Furthermore, we consider an idealized polycrystalline material which is composed of three structures: an austenitic parent phase, and two variants of a martensitic phase. The austenite phase will be denoted by 'a', and the martensite twins will be denoted by + or − in accordance with the sign of their associated Bain strain. The stresses and strains which appear in the following equations will be understood to be the one dimensional (scalar) counterparts to the second Piola–Kirchhoff stress tensor and the Green–Lagrange strain tensor.

3.1. Helmholtz, Gibbs, and potential energy functions

In one dimension, the three constituent materials may be described by free energy functions (per unit volume) of the form (Abeyaratne et al., 1994)

$$\rho_0\psi^{\mathrm{a}} = \tfrac{1}{2}\mathbb{C}^{\mathrm{a}}\boldsymbol{E}^2 - \zeta\mathbb{C}^{\mathrm{a}}\boldsymbol{E}(\theta-\theta_0) + \rho_0 c\theta(1-\log(\theta/\theta_0)) \tag{18}$$

$$\rho_0\psi^{-} = \tfrac{1}{2}\mathbb{C}^{-}(\boldsymbol{E}-\boldsymbol{E}_{\mathrm{T}}^{-})^2 - \zeta\mathbb{C}^{-}(\boldsymbol{E}-\boldsymbol{E}_{\mathrm{T}}^{-})(\theta-\theta_0) + \rho_0 c\theta(1-\log(\theta/\theta_0)) - \rho_0\lambda_{\mathrm{T}}(1-\theta/\theta_0) \tag{19}$$

$$\rho_0\psi^{+} = \tfrac{1}{2}\mathbb{C}^{+}(\boldsymbol{E}-\boldsymbol{E}_{\mathrm{T}}^{+})^2 - \zeta\mathbb{C}^{+}(\boldsymbol{E}-\boldsymbol{E}_{\mathrm{T}}^{+})(\theta-\theta_0) + \rho_0 c\theta(1-\log(\theta/\theta_0)) - \rho_0\lambda_{\mathrm{T}}(1-\theta/\theta_0) \tag{20}$$

where it is envisaged that ψ^{a} corresponds with the parent phase while ψ^{-} and ψ^{+} correspond to the two twin variants of the martensitic phase. It is implied that the heat capacities, c, and the thermal expansion coefficients, ζ, of the three constituents are constant and equal; the latent heat, λ_{T}, of the twins match; $\mathbb{C}^{\alpha}$ are the stiffnesses; $\boldsymbol{E}_{\mathrm{T}}^{\alpha}$ are the martensitic variant Bain strains; and θ_0 is the reference temperature. A free energy function for the mixture is formulated in accordance with Eq. (13), resulting in

$$\Psi = \hat{\Psi}(\boldsymbol{E},\theta,\xi^\alpha) = \sum_\alpha \xi^\alpha \psi^\alpha = \xi^{\rm a}\psi^{\rm a} + \xi^-\psi^- + \xi^+\psi^+. \tag{21}$$

Consequently, in accordance with mixture theory (Atkin and Craine, 1976), the effective second Piola–Kirchhoff stress $\boldsymbol{S}$ is computed directly from the Helmholtz Potential and may be expressed as a sum of contributions from each phase.

$$\boldsymbol{S} = \xi^{\rm a}\rho_0\frac{\partial\psi^{\rm a}}{\partial\boldsymbol{E}} + \xi^-\rho_0\frac{\partial\psi^-}{\partial\boldsymbol{E}} + \xi^+\rho_0\frac{\partial\psi^+}{\partial\boldsymbol{E}} \tag{22}$$

$$\boldsymbol{S} = \xi^-\mathbb{C}^-[\boldsymbol{E} - \boldsymbol{E}^-_{\rm T} - \zeta(\theta-\theta_0)] + \xi^{\rm a}\mathbb{C}^{\rm a}[\boldsymbol{E} - \zeta(\theta-\theta_0)] + \xi^+\mathbb{C}^+[\boldsymbol{E} - \boldsymbol{E}^+_{\rm T} - \zeta(\theta-\theta_0)]. \tag{23}$$

Likewise, the effective modulus may be computed from the free energy as

$$\rho_0\frac{\partial^2\Psi}{\partial\boldsymbol{E}^2} = \sum_\alpha \rho_0\frac{\partial^2\psi^\alpha}{\partial\boldsymbol{E}^2} = \xi^-\mathbb{C}^- + \xi^{\rm a}\mathbb{C}^{\rm a} + \xi^+\mathbb{C}^+ = \mathbb{C} \tag{24}$$

thereby allowing the stress relation to be restated in the following compact manner.

$$\boldsymbol{S} = \mathbb{C}[\boldsymbol{E} - \zeta(\theta-\theta_0)] - \xi^-\mathbb{C}^-\boldsymbol{E}^-_{\rm T} - \xi^+\mathbb{C}^+\boldsymbol{E}^+_{\rm T}. \tag{25}$$

Next, a three variant Gibbs function for the mixture is defined as described in Eq. (15), resulting in

$$G = -\sup_E\left\{\sum_\alpha \xi^\alpha(\boldsymbol{S}:\boldsymbol{E} - \rho_0\psi^\alpha)\right\} = -\sup_E\left\{\sum_\alpha \xi^\alpha g^\alpha\right\} \tag{26}$$

where $\hat{g^\alpha}(\boldsymbol{S},\ \boldsymbol{E},\ \theta) = \boldsymbol{S}:\boldsymbol{E} - \rho_0\psi^\alpha(\boldsymbol{E},\ \theta)$ defines a potential energy function for constituent α. In particular, the relation

$$G = \hat{G}(\boldsymbol{S},\ \theta) = -\sup_E\{\hat{g^\alpha}(\boldsymbol{S},\ \theta,\ \boldsymbol{E})\} \tag{27}$$

is valid when $\xi^\alpha = 1$ and the mixture at the associated point is comprised of only a single variant. Under an applied stress field these potential functions are considered indicators of the driving potential for phase transformation. As such, they play an important role in the evolution of the mixture as investigated next.

3.2. Transition probability

An explicit statement of Eqs. (16) and (17) for the three variant problem is given by

$$\begin{bmatrix} \dot{\xi}^{\rm a} \\ \dot{\xi}^- \\ \dot{\xi}^+ \end{bmatrix} = \omega \begin{bmatrix} -(P_{\rm a-} + P_{\rm a+}) & P_{\rm -a} & P_{\rm +a} \\ P_{\rm a-} & -(P_{\rm -a} + P_{\rm -+}) & P_{\rm +-} \\ P_{\rm a+} & P_{\rm -+} & -(P_{\rm +a} + P_{\rm +-}) \end{bmatrix} \begin{bmatrix} \xi^{\rm a} \\ \xi^- \\ \xi^+ \end{bmatrix} \tag{28}$$

where explicit vector–matrix notation has been employed. The specification of estimates of the transformation probabilities $P_{\alpha\beta}$ essentially follows from the work of Abeyaratne et al. (1994) as well as Achenbach and Müller (1985). Given a representative volume possibly comprised of any two constituents α and β, the rate at which phase α transforms to phase β is given by a classical first order

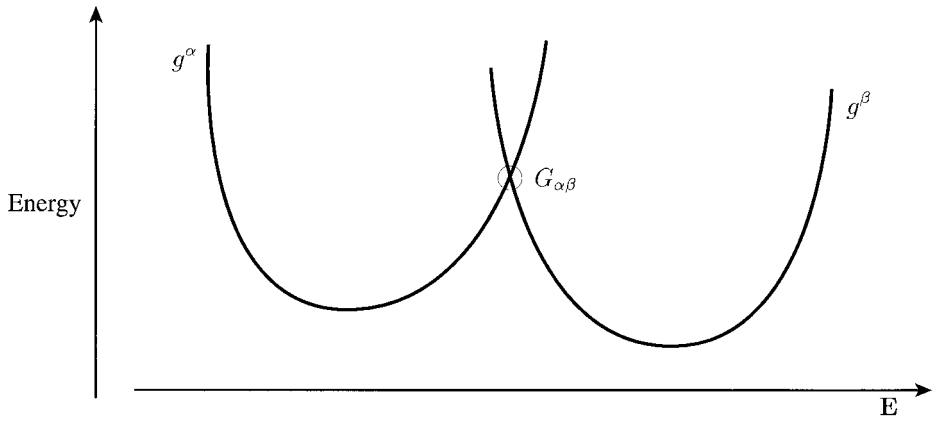


Fig. 1. Graphical interpretation of barrier energy for a fixed stress value.

or Eyring model as

$$P_{\alpha\beta} = \exp\left[\frac{-\Delta v b_{\alpha\beta}}{k\theta}\right], \tag{29}$$

where k is Boltzmann's constant, Δv is the volume occupied by the transforming region, θ is the absolute temperature, and $b_{\alpha\beta}$ is the energetic barrier for the transition from α to β. The value of this barrier is estimated simply as

$$b_{\alpha\beta} = \hat{b}_{\alpha\beta}(\boldsymbol{S},\theta) = G_{\alpha\beta} - G(\boldsymbol{S},\theta,\xi^\alpha = 1) \geq 0 \tag{30}$$

where $G_{\alpha\beta}$ is defined as the energy at which the potential functions g^α and g^β are equal in value for a fixed value of stress. The graphical interpretation which is typically portrayed in the referenced works is a plot of the potential functions g^α such as that shown in Fig. 1. In this figure, the circled intersection locates $G_{\alpha\beta}$. For the present case of a three variant system, the above method can be used to compute the six required probability terms which make up $\mathbf{Q}$.

Eqs. (28)–(30), when combined with the framework presented in Section 2, are sufficient to define the phase transformation rates. The thermodynamic consistency of the model is governed by the Clausius–Planck inequality which reduces in the present case to the requirement that

$$-\int_{t_1}^{t_2} \frac{1}{\theta} \frac{\partial G}{\partial \boldsymbol{\xi}} \dot{\boldsymbol{\xi}} \geq 0, \tag{31}$$

where t_1, t_2 are equilibrium states. We next turn our attention to computational aspects of the model.

3.3. Time integration

In what follows, it is assumed that the model is to be implemented within a traditional strain driven finite element code. This allows one to consider the constitutive model at a Gauss point as an independent entity within an unspecified array of integration points. At time t_{n+1} the constitution will

have available to it the current strain and temperature values ($\boldsymbol{E}_{n+1}$, θ_{n+1}) as well as the history variables ξ^{α}_{n}. In return, the constitution must provide a stress value $\boldsymbol{S}_{n+1}$, updated volume fractions ξ^{α}_{n+1}, and the algorithmic tangent[1] $\mathbb{C}^{\text{alg}}_{n+1}=(\partial \boldsymbol{S}_{n+1}/\partial \boldsymbol{E}_{n+1})$. The subscripts n and $n+1$ denote quantities evaluated at times t_n and $t_{n+1}=t_n+\Delta t$, $\Delta t>0$, respectively. Note for simplicity we will consider the temperature as an assigned field. In moving the (one dimensional) discrete form of the constitution from time t_n to time t_{n+1}, we begin by writing the effective modulus, stress and rate matrix at time t_{n+1}:

$$\mathbb{C}_{n+1}=\xi^{-}_{n+1}\mathbb{C}^{-}+\xi^{\mathrm{a}}_{n+1}\mathbb{C}^{\mathrm{a}}+\xi^{+}_{n+1}\mathbb{C}^{+} \tag{32}$$

$$\boldsymbol{S}_{n+1}=\mathbb{C}_{n+1}[\boldsymbol{E}_{n+1}-\zeta(\theta_{n+1}-\theta_0)]-\xi^{-}_{n+1}\mathbb{C}^{-}\boldsymbol{E}^{-}_{\mathrm{T}}-\xi^{+}_{n+1}\mathbb{C}^{+}\boldsymbol{E}^{+}_{\mathrm{T}} \tag{33}$$

$$\mathbf{Q}_{n+1}=\hat{\mathbf{Q}}(\boldsymbol{S}_{n+1},\,\theta_{n+1}). \tag{34}$$

Turning to the transformation equations, we first express the solution as the exact form (for constant $\mathbf{Q}$)

$$\boldsymbol{\xi}(\Delta t)=\exp[\mathbf{Q}\Delta t]\boldsymbol{\xi}_n. \tag{35}$$

Then, by applying 'backward Euler', it is required that the discrete form of Eq. (35) be satisfied at the end of each timestep. This requirement is stated as

$$\boldsymbol{\xi}_{n+1}=\exp[\tilde{\mathbf{Q}}_{n+1}]\boldsymbol{\xi}_n, \tag{36}$$

where $\tilde{\mathbf{Q}}_{n+1}=(\Delta t)\mathbf{Q}_{n+1}$ is a nonlinear function of the fractions ξ^{α}. The motivation for integrating this set in the specified manner is threefold. The first lies in the highly nonlinear relationship between $\mathbf{Q}$ and the volume fractions, hence the need for an accurate integration scheme. The second is the small size of the system ($\dim(\mathbf{Q})=v$ where v is the number of variants) and the need to compute the transient solution, thereby making the computation of the exponential a reasonable approach. The third reason is based on the properties of $\mathbf{Q}$ and a comparison of several methods for solving problems with this specific form. However, as is discussed further in Appendix A, a more efficient technique will prove beneficial in multidimensions where the number of variants is increased.

In order to solve for the unknowns in the implicit integration, it is convenient to formulate a residual function $\mathbf{f}$ defined as

$$\mathbf{f}(\boldsymbol{\xi})=\boldsymbol{\xi}-\exp[\tilde{\mathbf{Q}}(\hat{\boldsymbol{S}}(\boldsymbol{\xi}))]\boldsymbol{\xi}_n, \tag{37}$$

where the nonlinearity of the equation set is apparent from the noted function dependencies. The determination of a self-consistent solution of Eqs. (32)–(36) is accomplished when

$$\mathbf{f}(\boldsymbol{\xi}_{n+1})=\mathbf{0} \tag{38}$$

is satisfied to some acceptable tolerance. An outline of the approach taken to solve for $\boldsymbol{\xi}_{n+1}$ is given by Algorithm 1, where it is seen that the technique is of the Newton–Raphson type.

[1] See Simo and Taylor (1986) for a discussion of consistent tangent operators; see also Appendix B.

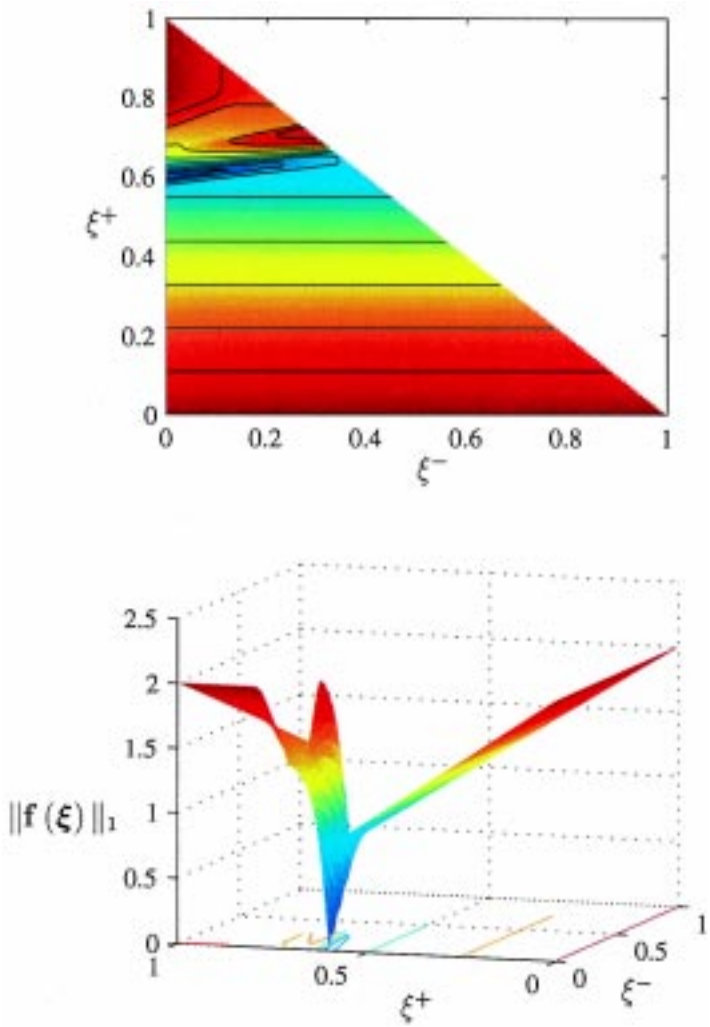


Fig. 2. Plot of the norm of the residual function over the phase space domain. The upper figure provides a contour plot of the surface shown below.

Algorithm 1 (Constitutive iteration)**.**

1. Input at time t^{n+1}:

$$\{\boldsymbol{E}_{n+1}, \theta_{n+1}, \boldsymbol{\xi}_n\} \tag{39}$$

2. Initialize values:

$$\boldsymbol{S} = \hat{\boldsymbol{S}}(\boldsymbol{E}_{n+1}, \theta_{n+1}, \boldsymbol{\xi}_n) \tag{40}$$

$$\boldsymbol{\xi}_k = \boldsymbol{\xi}_n \tag{41}$$

3. Compute the residual function:

$$\mathbf{Q} = \hat{\mathbf{Q}}(\boldsymbol{S}, \theta_{n+1}) \tag{42}$$

$$\mathbf{f}_k = \xi_k - \exp[\Delta t \mathbf{Q}]\xi_n \tag{43}$$

4. IF $\|\mathbf{f}\|_\infty <$ TOL THEN

$$\xi_{n+1} = \xi_k \tag{44}$$

$$\boldsymbol{S}_{n+1} = \boldsymbol{S} \tag{45}$$

$$\mathbb{C}^{\text{alg}}_{n+1} = \frac{\partial \boldsymbol{S}_{n+1}}{\partial \boldsymbol{E}_{n+1}} \tag{46}$$

$$\text{RETURN} \Longrightarrow \text{END} \tag{47}$$

5. ELSE

$$\xi_{k+1} = \xi_k - \kappa[\nabla_\xi \mathbf{f}_k]^{-1}\mathbf{f}_k \tag{48}$$

$$\boldsymbol{S} = \hat{\boldsymbol{S}}(\boldsymbol{E}_{n+1}, \theta_{n+1}, \xi_{k+1}) \tag{49}$$

$$k = k + 1 \tag{50}$$

6. GOTO STEP 3.

In practice, the steplength/linesearch parameter κ is typically set to equal one. However, if $(\|\mathbf{f}_{k+1}\|_\infty / \|\mathbf{f}_k\|_\infty) \geq \epsilon$, where ϵ is typically set between $\frac{1}{2}$ and $\frac{4}{5}$, then an attempt is made to improve the update by adjusting κ. Although various techniques for performing the linesearch were attempted, the simple techniques of divide and conquer and sampling proved most efficient due to the fact that $\xi^\alpha \in [0, 1]$. It is noted that each step of the constitutive iteration requires computation of the term $\mathbf{M} = \exp[\Delta t \mathbf{Q}]$. Since this operation has the potential to dominate the cost of the constitutive iteration, particularly upon extension to higher dimensions, means by which this process can be improved and/or minimized are worth investigating. Appendix A discusses steps which can be taken to reduce the cost of the computation of $\mathbf{M}$, while on the constitutive side heuristics can be developed to improve the starting iterate and hence also reduce the number of operations.

As an illustration of the difficulty in solving the constitutive equations, Fig. 2 shows the one norm of the residual function plotted over the domain $\{(\xi^-, \xi^+) \in [0, 1] \times [0, 1] \,|\, \xi^- + \xi^+ \leq 1\}$ for a particularly difficult set of initial conditions. As is apparent, modification of the standard Newton–Raphson scheme is necessary to provide an efficient solution routine. Numerical tests of the proposed integration scheme are provided in Section 4.

4. Numerical simulations

In this section we provide examples which demonstrate the qualitative and quantitative behavior of the numerical implementation. As a baseline, simulations of simple thermal and mechanical cycles are

Table 1
Approximate material properties for TiNi

Reference density	$\rho = 6448.0$ kg/m^3
Young's moduli, Austenite	$\mathbb{C}^a = 67.0 \times 10^9$ Pa
Young's moduli, Martensite	$\mathbb{C}^-, \mathbb{C}^+ = 26.3 \times 10^9$ Pa
Latent heat	$\lambda_T = 14.5 \times 10^3$ J/kg
Heat capacity	$c = 4.00 \times 10^2$ J/(kg K)
Transformation rate	$\omega = 16 \times 10^3$ 1/s
Transformation volume	$\Delta v = 2.71 \times 10^{-27}$ m^3
Bain strain	$\boldsymbol{E}_T^{\alpha} = \pm 0.07$

presented first. Then, as a test of the constitutive algorithm under spatially inhomogeneous loading patterns, a truss system similar to that suggested by Govindjee and Kasper (1999) is run. The final examples are of a shape memory beam undergoing large rotations, a simulation of an SMA mechanism, and a plastically deformed shape memory beam.

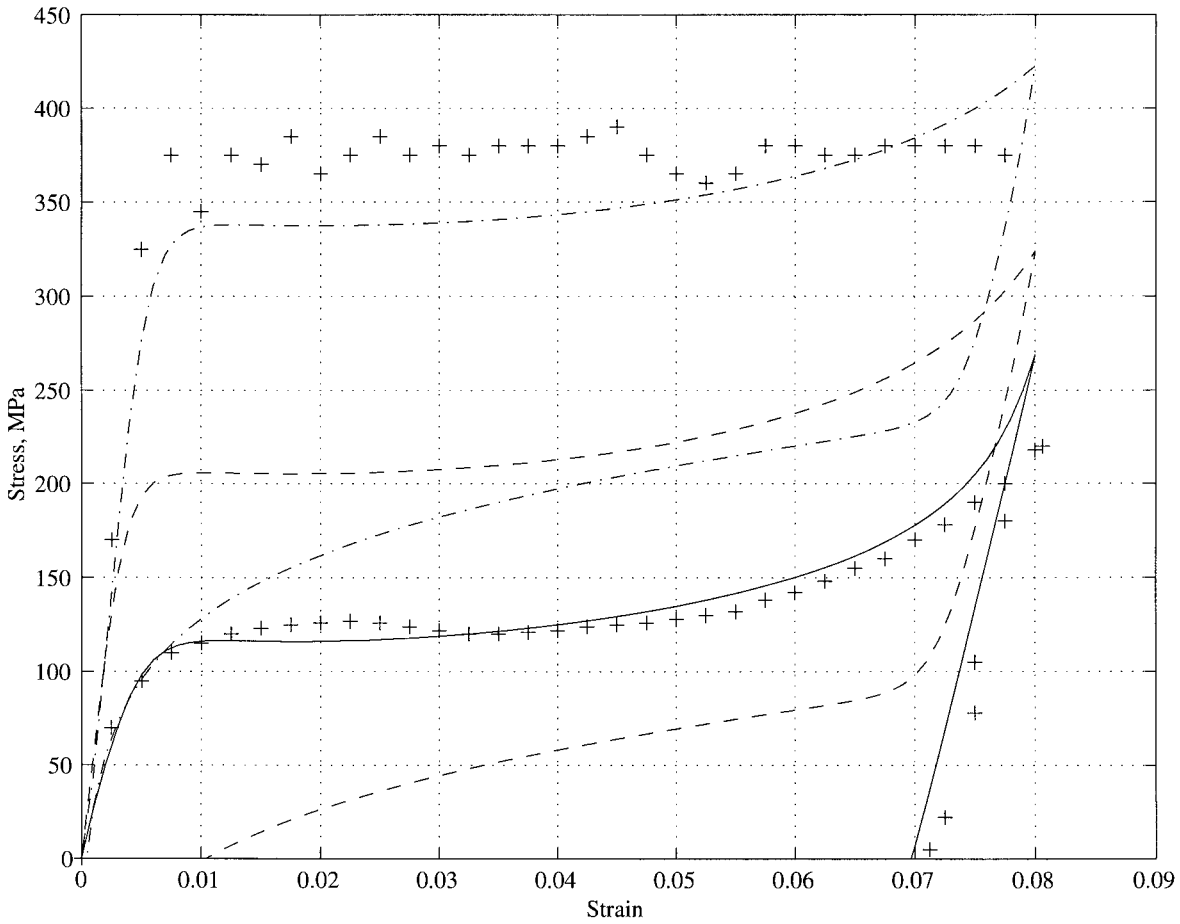


Fig. 3. Mechanical load cycle of TiNi bar under isothermal conditions. The solid, dashed, and dash–dot lines indicate runs at 263, 293, 323 K. The + points indicate experimental data by Liang as per Brinson and Lammering (1993).

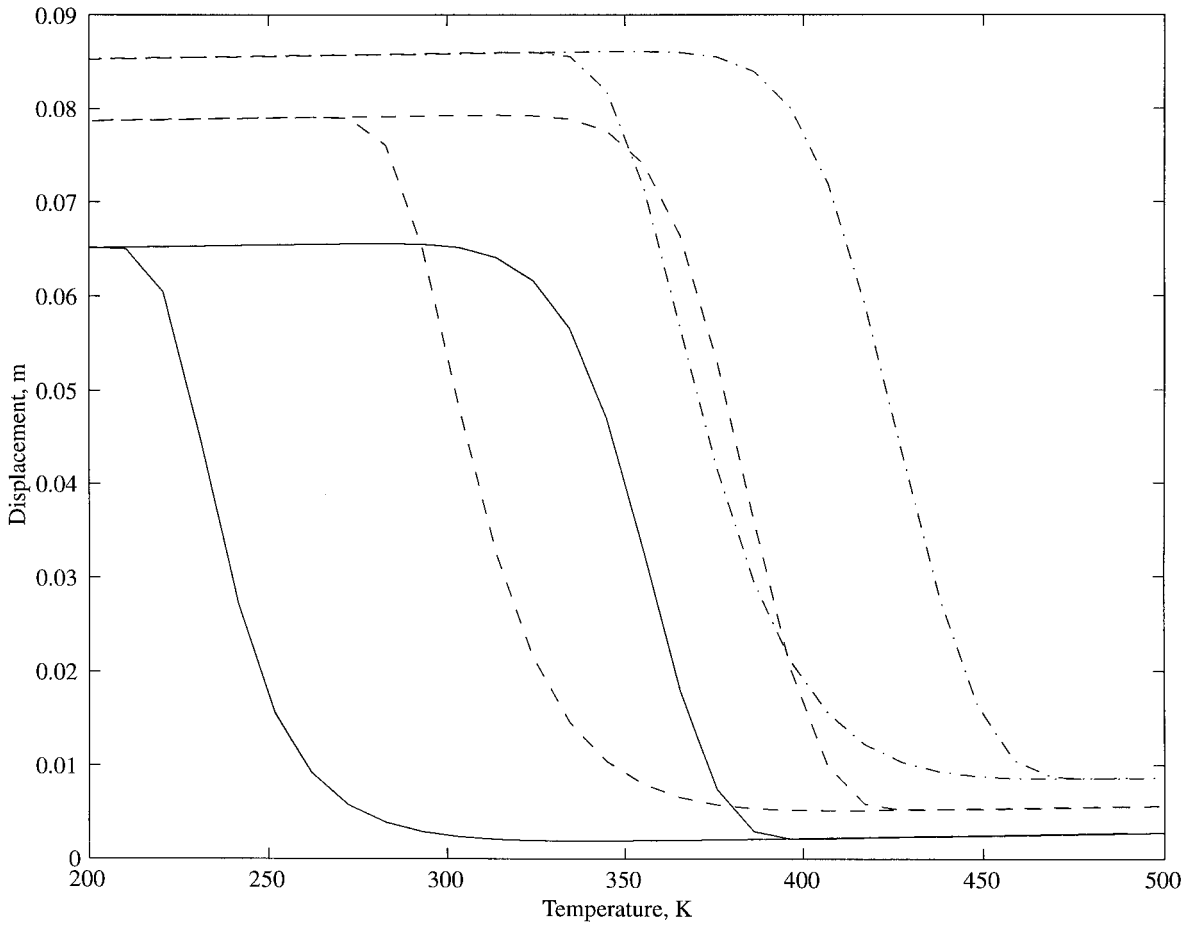


Fig. 4. Thermal load cycle of TiNi bar under isostress conditions. The solid, dashed and dashed–dot lines indicate loads of 100, 300, and 500 MPa, respectively.

4.1. Unidimensional mechanical and thermal cycles

To demonstrate the constitutive response of the model to an applied isothermal mechanical load cycle, we begin by subjecting the model to a strain path at several temperatures. The material under consideration is a titanium nickel alloy (TiNi) for which approximate material parameters are chosen as shown in Table 1. As an order of magnitude check, the test runs at 263 and 323 K are compared to the experimental data of Liang as reported in Brinson and Lammering (1993). The initial conditions for the lowest temperature curve were set as $\xi^{\mathrm{a}}=0$, $\xi^{-}=\xi^{+}=\frac{1}{2}$, so the run represents a martensitic detwinning transformation. The next two curves, run at 293 and 323 K, started in a pure austenitic state ($\xi^{\mathrm{a}}=1$) and exhibited superelasticity to varying degrees in accordance with the temperature. The applied strain rate was quasistatic ($\dot{\boldsymbol{E}}=1\times10^{-4}\mathrm{s}^{-1}$) and the reference temperature $\theta_0=273$ K. The variance of the transformation load and recoverable strain is seen in Fig. 3 to be realistic for the intended class of materials.

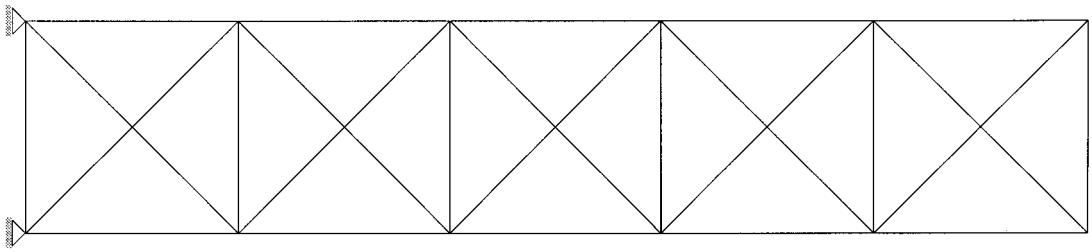


Fig. 5. Truss system configuration. Each bay measures 0.2 m square, with each bar having a cross sectional area of 0.000123 m^2 (0.00625 m radius).

The second series of tests considered an isostress thermal load cycle of a one-dimensional bar under several values of an applied tension. The results of the three runs, loaded at 100, 300 and 500 MPa, respectively, are shown in Fig. 4.

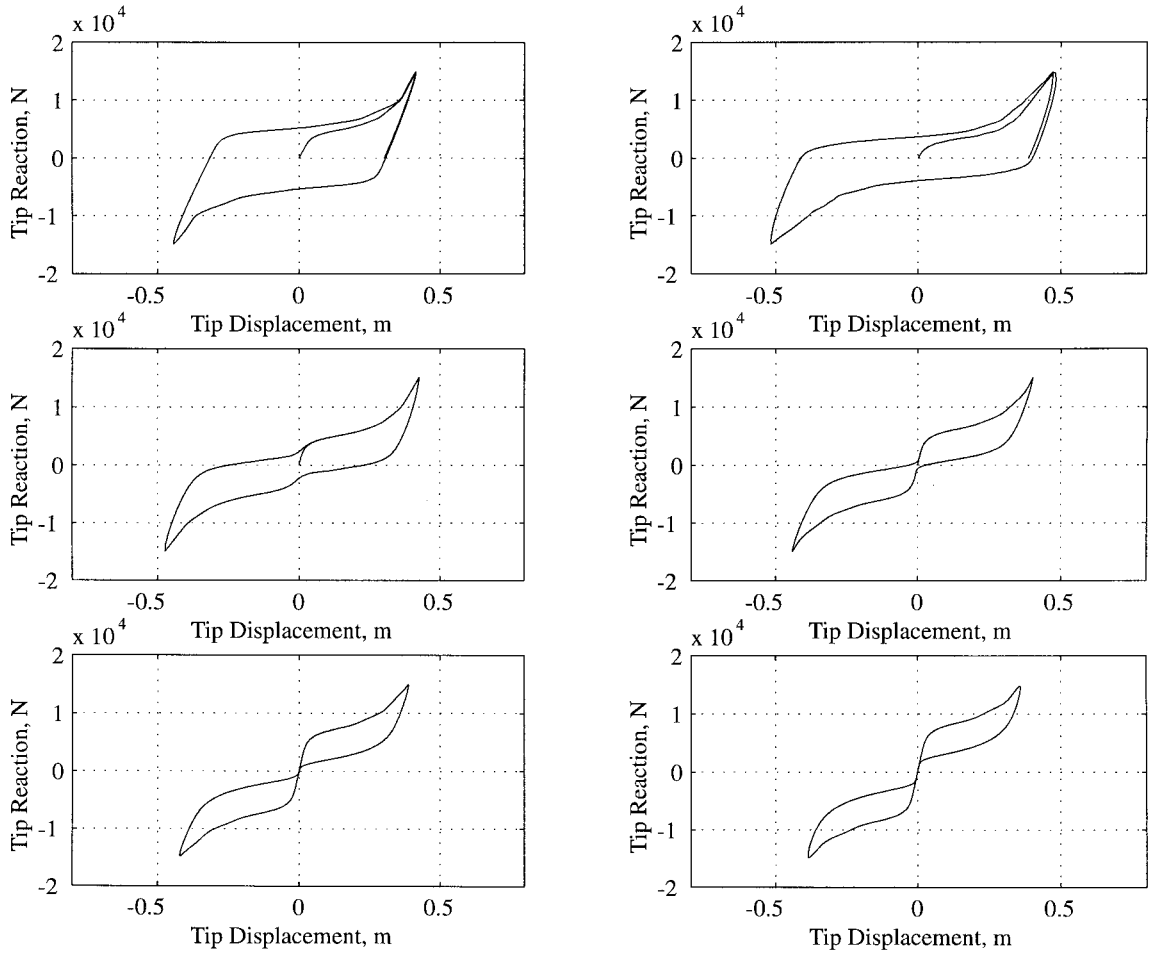


Fig. 6. Simulation of truss system response to mechanical cycling. The progression from left to right and top to bottom shows identical load paths at 253, 283, 313, 333, 353, and 373 K. The structural system is shown in Fig. 5.

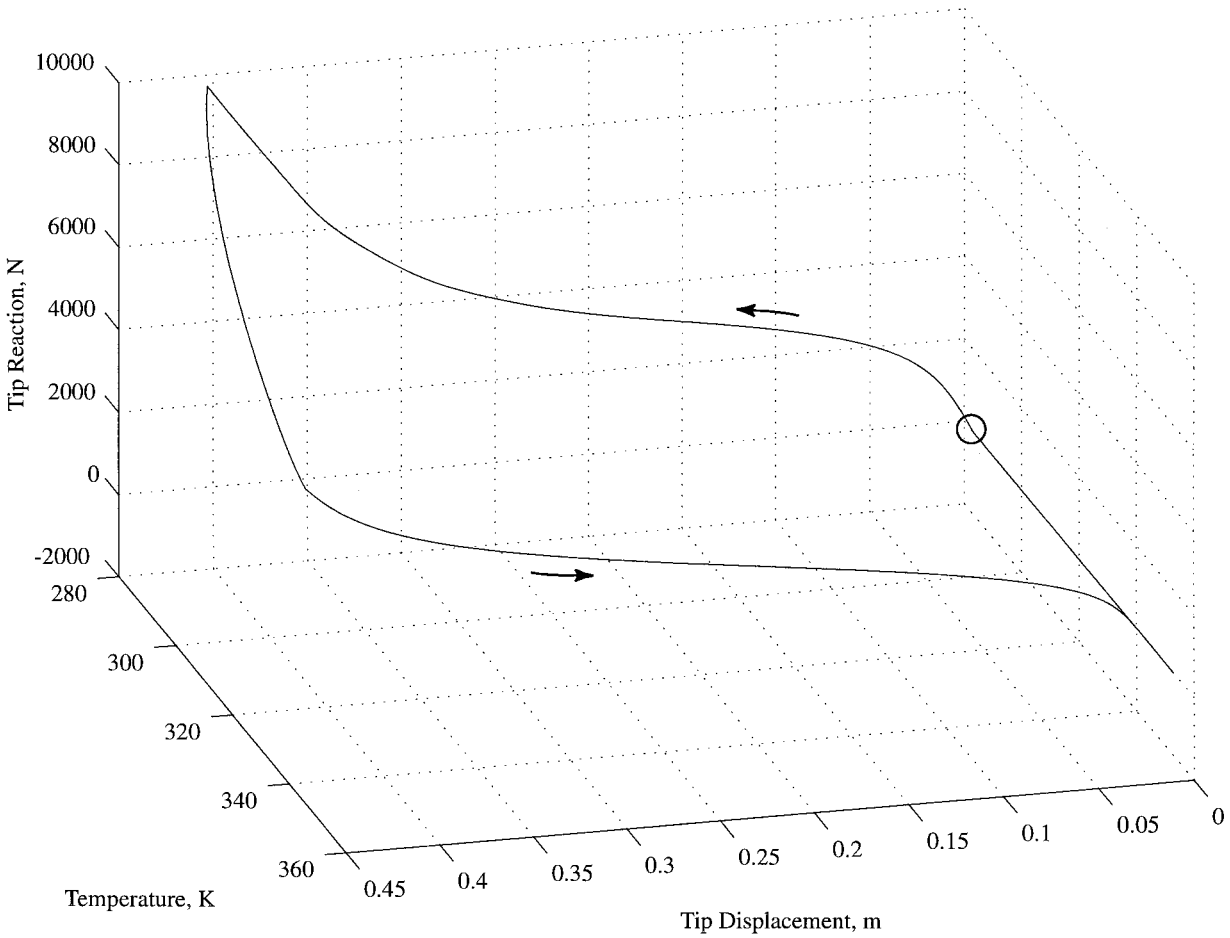


Fig. 7. Simulation of truss system response to a mechanical-thermal closed cycle starting from the state $\xi^- = \xi^+ = \frac{1}{2}$ and a temperature of 283 K. The circle denotes the start and finish points, while the path direction is indicated by the arrows. The structure is that of Fig. 5.

In each case, the material is assumed to start in the austenite phase at the high temperature, and transformation to martensite is induced by lowering the temperature. The simulations indicate that with increasing load, the transition occurs at higher temperature as is observed in TiNi alloys.

4.2. Truss system examples

Fig. 5 shows the physical layout and dimensions of a cantilever truss similar to that which was suggested as an example problem by Govindjee and Kasper (1999). By virtue of the spatially inhomogeneous stress state in the structure, this system provides a test of the constitutive iteration over a wide range of initial conditions. For both the mechanical load cycles and the thermo-mechanical cycles discussed below, the tip loading was applied at a quasi-static rate and the material parameters listed in Table 1 were employed.

The load deflection response under mechanical cycling as measured at the end of the truss is shown in Fig. 6. The six runs, which are ordered left to right and then downward, simulated isothermal

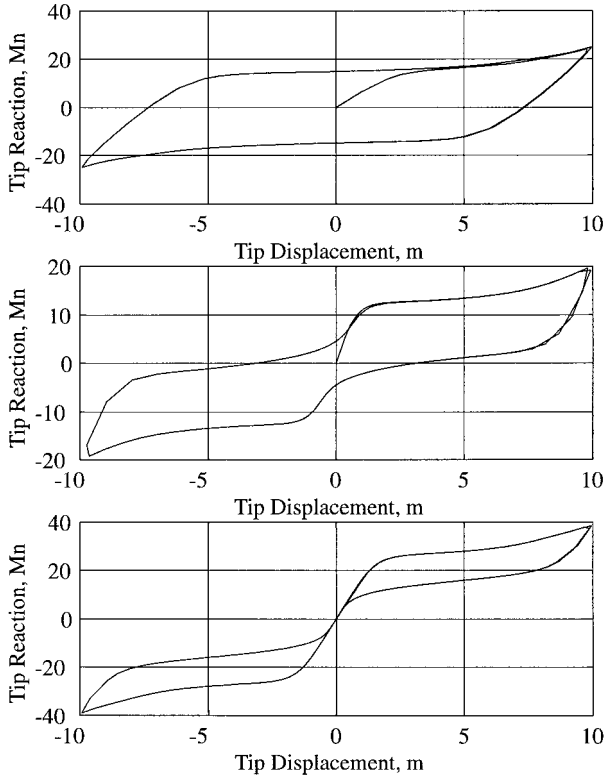


Fig. 8. Beam simulation: mechanical cycles at various temperatures. Initial conditions (from top to bottom) were $\xi^- = \xi^+ = \frac{1}{2}$ at 253 K, $\xi^a = 1$ at 313 K, and $\xi^a = 1$ at 373 K.

conditions at 253, 283, 313, 333, 353, and 373 K, respectively. The initial phase state for temperatures below 293 K were set as $\xi^- = \xi^+ = 0.5$, while the higher temperature runs were initialized with $\xi^a = 1$; here and throughout the remainder the reference temperature $\theta_0 = 293$ K. By running identical load cycles at increasing temperatures, the series demonstrates how the model progresses toward superelastic behavior.

A combined thermal and mechanical cycle was then applied to the system as indicated in Fig. 7. As previously, the plot indicates reaction at the tip of the structure. The elements were initialized in a martensitic state ($\xi^- = \xi^+ = 0.5$) at a temperature of 283 K, and subsequently loaded and unloaded under isothermal conditions. The temperature was then increased under zero load to a maximum of 353 K, by which time the structure had essentially recovered its original configuration. The cycle was then closed by decreasing the temperature to its initial value.

4.3. Beam examples

Beam simulations were performed by incorporating the constitution presented above into the beam formulation of Simo et al. (1984). The cross section of each beam was divided into five layers, with a five point quadrature rule applied in each layer (see Kasper, 1997 for details). The simulated cantilever

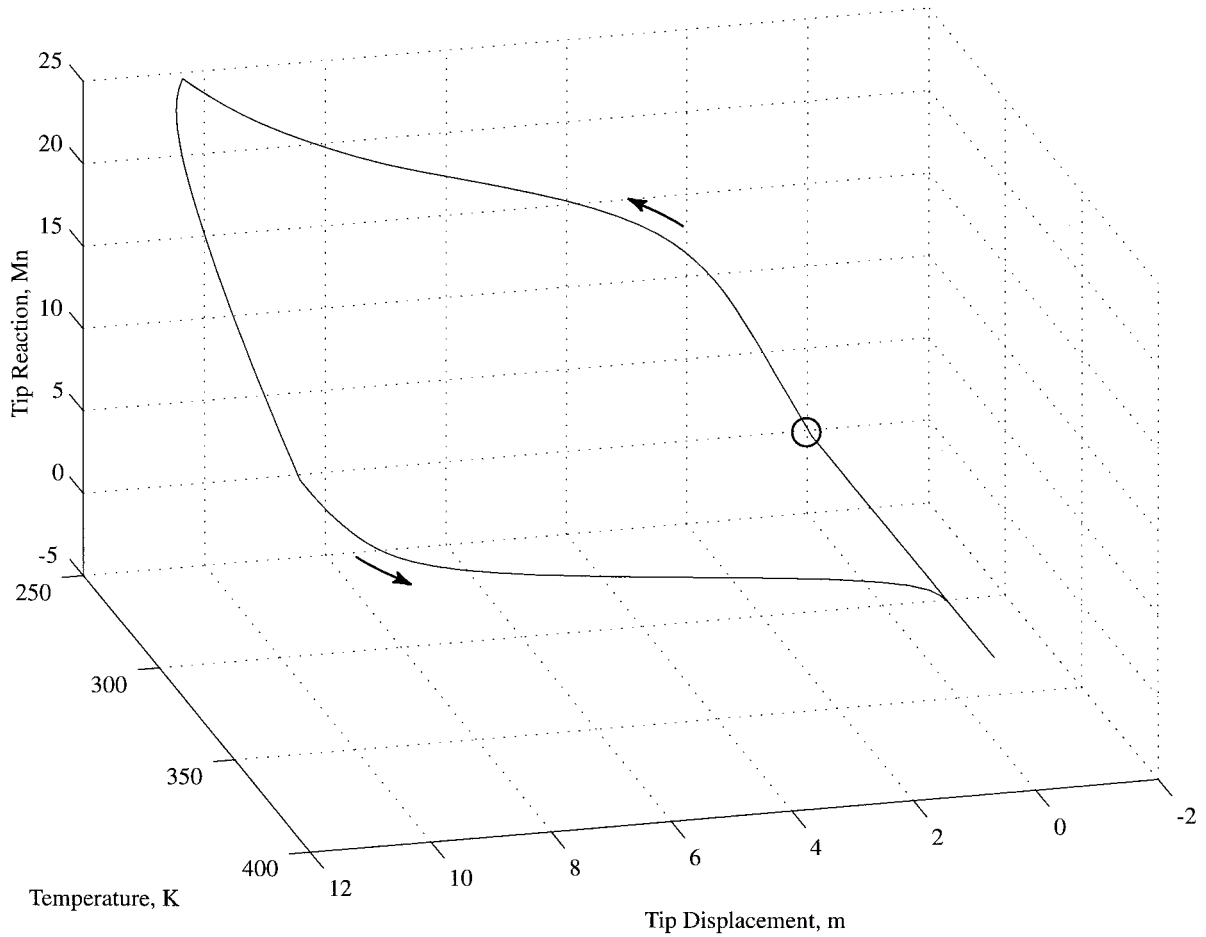
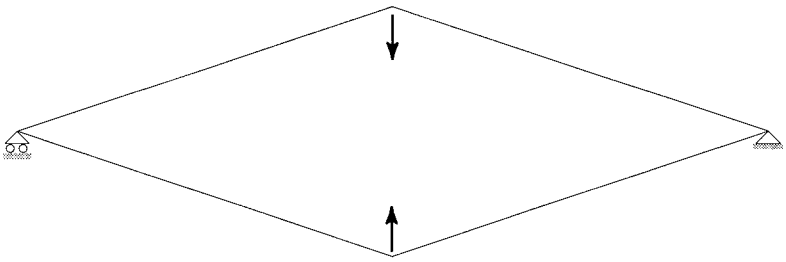


Fig. 9. Beam simulation: thermo-mechanical cycle with initial conditions $\xi^- = \xi^+ = \frac{1}{2}$ at 253 K. The circle marks the start and finish points along the path; the arrows indicate the direction traveled.

beam was then divided into ten elements along its length, and displacement driven at the far end. The length and cross section were set at 20 and 2×1 m, respectively, and the material parameters are those listed in Table 1.

The first set of simulations consists of three isothermal runs at 253, 313 and 373 K. The same displacement path was applied to the end of each beam, with the load displacement results indicated in Fig. 8. The low temperature cycle was initialized to $\xi^- = \xi^+ = 0.5$, whereas the second and third simulations began in the austenite phase. As with the truss example, superelastic behavior is induced with increased temperature.

The next simulation begins with the aforementioned beam in a martensitic state ($\xi^- = \xi^+ = 0.5$) at 253 K. An end load is then applied and removed at this temperature, leaving the beam with a residual deformation due to the martensitic transformation. The temperature is then increased to a maximum of 373 K to induce recovery of the deformation as indicated in Fig. 9. Return to the initial temperature then closes the cycle.



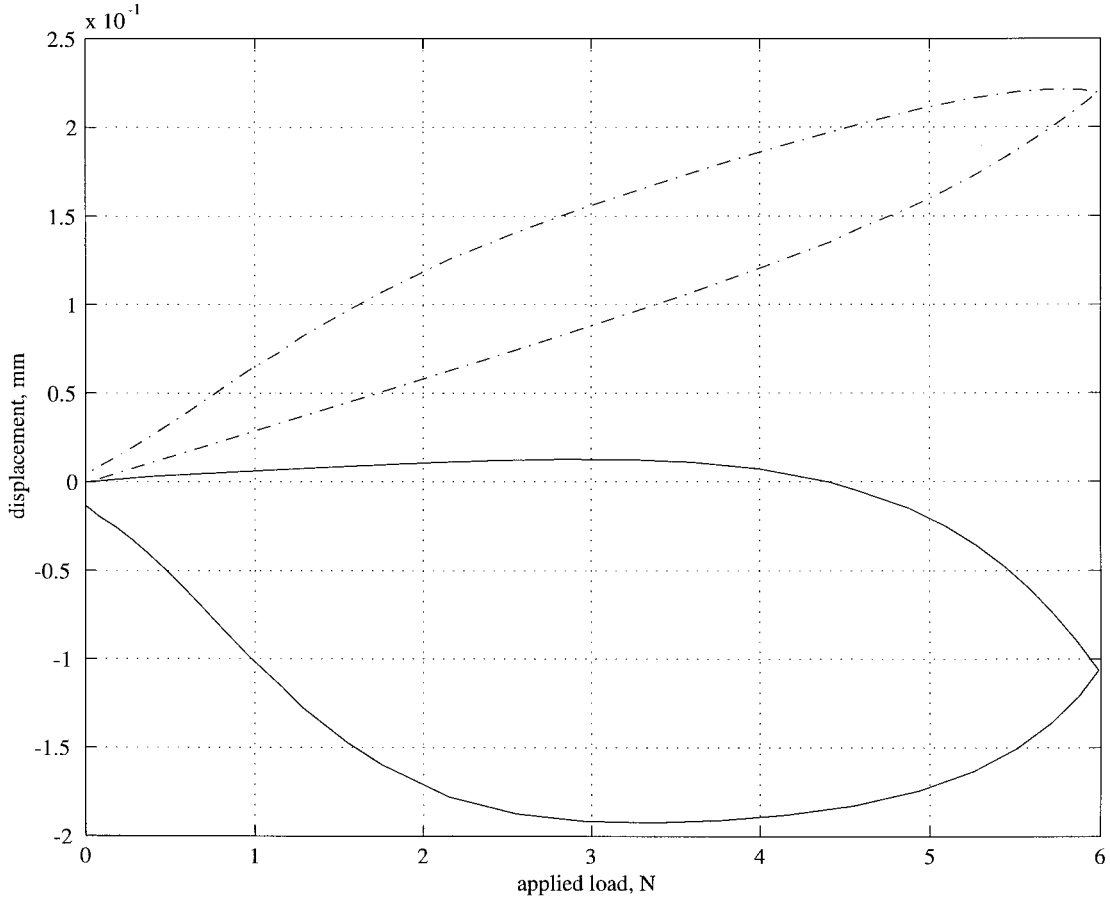


Fig. 10. The top figure shows the layout of the stent mechanism analyzed. The device measures 4.2 mm wide by 1.4 mm tall with a square cross section of 0.0508 mm in each direction. The lower figure shows the total lateral deflection (solid line) and the total vertical height reduction (dash–dot line) under the applied pinching load. The simulation is intended as a rough proof of concept for such devices.

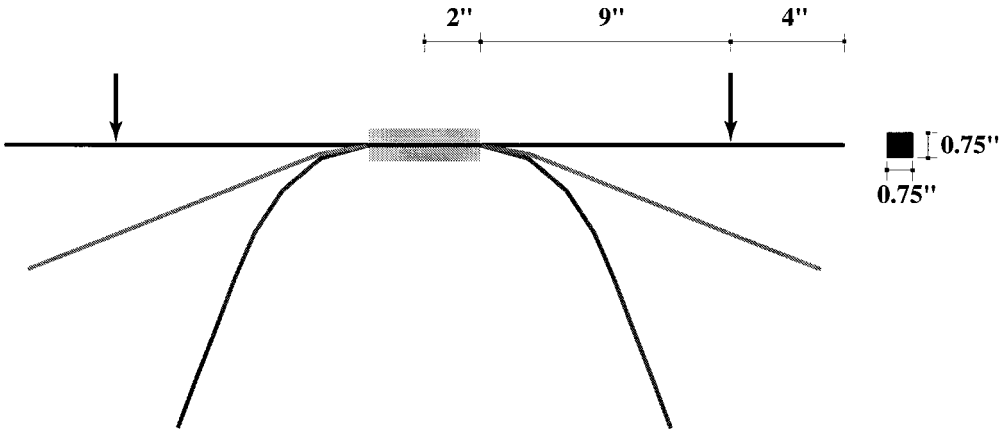


Fig. 11. Shape memory beam simulation incorporating plasticity. The bottom deflected shape shows the residual deformation induced by the mechanical load/unload cycle, whereas the intermediate configuration demonstrates partial recovery following a thermal cycle. The original geometry is given at the top; see Fig. 12 for the response curves. Note that the region of the cantilever beyond the zone of plasticity completely recovers the initial shape following the thermal cycle.

4.4. SMA mechanism

The next simulation is of a structure representing an idealized portion of a SMA stent. The chosen dimensions of the lattice cell are shown in Fig. 10, along with the lateral and vertical displacements plotted against the pinching load applied to crush the stent vertically. The members are beam elements of the type described in the previous simulations, and the temperature was held constant at 330 K throughout the run. At the maximum applied load, the vertical dimension of the stent is reduced by 33%, which is fully recovered upon removal of the load. Note the nonintuitive behavior of the lateral displacement curve caused by the formation of the ξ^- variant under load. While this demonstration does not attempt to model any existing stent mechanism, the simulation does serve as a rough proof of concept for such devices. It is interesting to note that in practice, similar devices are designed in both superelastic (as shown here) and martensitic versions (Russell, 1998).

4.5. Plasticity example

The final example incorporates the effects of plasticity as described in Appendix C. The test geometry, which is provided in Fig. 11, is that of a cantilever beam. The material parameters were chosen to roughly match those reported in Govindjee and Kasper (1997), although no attempt at calibration was made. The simulation consisted of a mechanical cycle at a constant temperature, followed by a thermal cycle in the unloaded state. The configurations and the stress–strain-temperature response curves given in Figs. 11 and 12 indicate the interplay of plasticity and the shape memory effect, particularly near the base of the cantilever. It is interesting to note that even with the significant plastic deformation, over half the mechanically induced residual deformation was recoverable.

5. Discussion

In this paper we have addressed the computational aspects of implementing a rate dependent constitutive model for shape memory alloy materials. This model, which is derived from the referenced

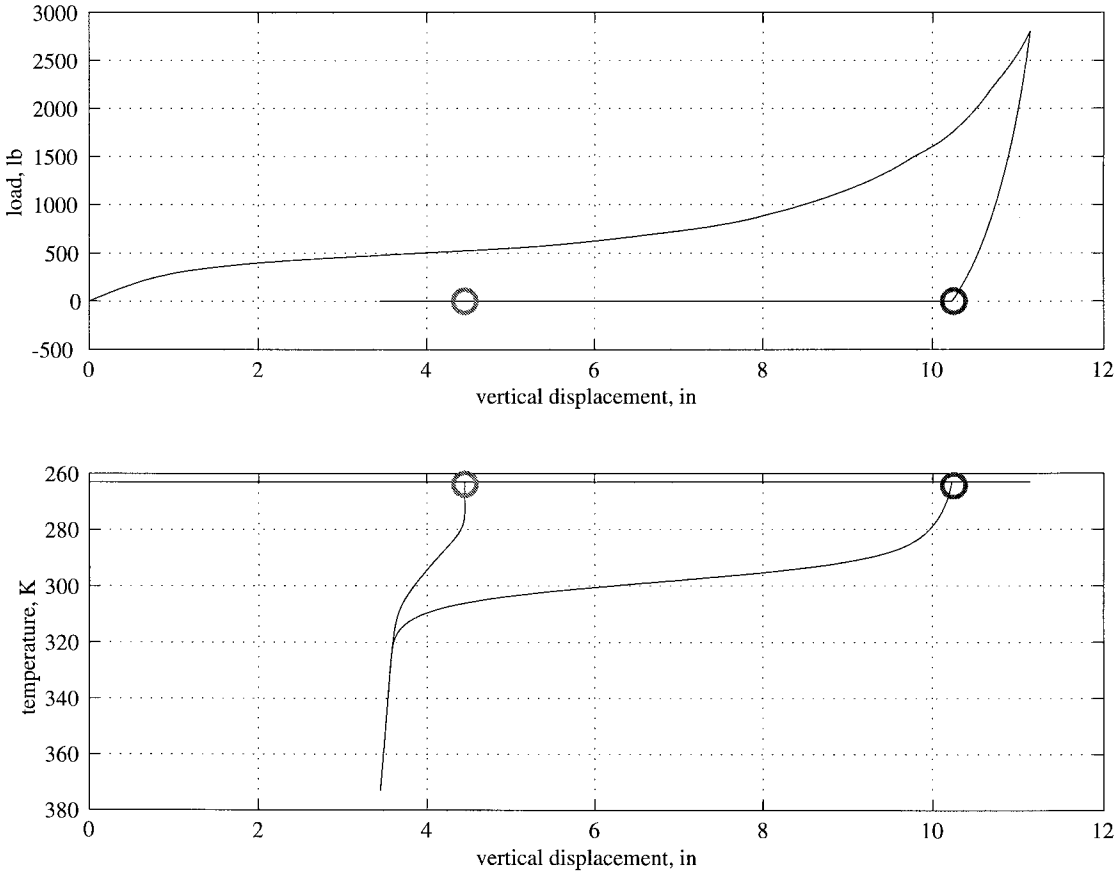


Fig. 12. Thermomechanical cycling of a shape memory beam accounting for plasticity. The circled points coincide with the deflected shapes given in Fig. 11. Despite the large residual deformation following unload, over 50% of the residual deformation is recovered by the thermal cycle.

literature, has been shown to realistically capture the one-dimensional behavior of TiNi alloys under both thermal and mechanical loading. While an explicit proof has not been provided, numerical computation of the model dissipation indicates that the constitution is second law compliant in accordance with the Clausius–Planck inequality. Additionally, the model is extendible to higher spatial dimensions where the number of variants must be increased to account for the observed behavior of shape memory alloys; the primary difficulty in moving to higher dimensions lies in computing the energy barriers.

From a computational viewpoint, it has been found that the proposed integration algorithm provides a robust means for solving the highly nonlinear constitutive equation set which arises from such models, particularly when driven over a wide range of loading paths. On the downside, this approach is not as efficient in terms of flop counts as a phase space approach, although preliminary suggestions have been made for reducing the expense of the current approach.

Items of interest for further investigation include the incorporation of rate dependent plasticity into the constitutive model, extension to higher dimensions, and further reductions in the computational expense required for large scale structural system simulation.

Acknowledgements

The authors gratefully acknowledge the support of the Lawrence Livermore National Laboratory through subcontract B502681 to DOE prime contract W-7405-ENG-48.

Appendix A

Computation of exp[$\tilde{\boldsymbol{Q}}$]

Noting that the individual off-diagonal entries in the transition probability matrix are bounded by $0 \leq P_{ij} \leq 1$, it is apparent that $\tilde{\mathbf{Q}}$ is a non-normal, negative semidefinite matrix with the maximum possible absolute eigenvalue, $|\lambda_{\max}|$, bounded by $|\lambda_{\max}| \leq 2(v-1)\omega\Delta t$. Here, v is taken as the number of variants which throughout this discussion we will assume is small (say $v \leq 24$). Given that zero is also an eigenvalue, the spacing between the maximum and minimum eigenvalues is also bounded by $2(v-1)\omega\Delta t$, and similarly it is noted that the infinity norm of $\tilde{\mathbf{Q}}$ is strictly bounded by $2(v-1)\omega\Delta t$. This measure may be used as one potential indicator of problem 'stiffness' or difficulty in computing the exponential (Stewart, 1994) as discussed next.

For the problem at hand, a number of means of computing $\exp[\tilde{\mathbf{Q}}]$ were tried in a comparative manner. The techniques included the Taylor series method, uniformization, Pade approximation and several ordinary differential equation techniques. Tests were run with both spaced and repeated eigenvalues for problems of increasing stiffness as approximated by the infinity norm of $\mathbf{Q}$. The results indicated that Pade and uniformization with scaling and squaring are most efficient for the typical problems encountered in this application.

The cost of each computation, using Pade as a baseline, is dominated by the expense of the squaring portion of the routine when $\|\tilde{\mathbf{Q}}\|_\infty$ grows. In such cases, the cost of the simplest implementation is roughly $O(2\ \log(\|\tilde{\mathbf{Q}}\|_\infty)v^3)$ where v is again the number of variants. In the following, three potential techniques for reducing costs are mentioned.

Reduction of dimensionality

As indicated by the zero eigenvalue of $\tilde{\mathbf{Q}}$, a linear dependence stemming from the fact that $\Sigma\dot{\xi}^\alpha = 0$ (and $\Sigma\xi^\alpha = 1$), may be utilized to reduce the dimension of the problem. If one ignores the relatively small cost required to perform the reduction, the cost of the integration immediately becomes proportional to $(v-1)^3$ rather than $(v)^3$. Thus, this simple consideration can reduce costs significantly when the exponential routine is called frequently.

Triangular decomposition

Another approach which may prove more efficient than Pade or uniformization is to compute the Schur decomposition of $\tilde{\mathbf{Q}}$:

$$\tilde{\mathbf{Q}} = \mathbf{S}\tilde{\mathbf{T}}\mathbf{S}^{\mathrm{T}}$$

where $\mathbf{S}$ is orthogonal, and $\tilde{\mathbf{T}}$ is upper triangular (for a $\tilde{\mathbf{Q}}$ with real eigenvalues). The problem is then reduced to the exponentiation of the block triangular matrix $\tilde{\mathbf{T}}$:

$$\exp[\tilde{\mathbf{Q}}] = \mathbf{S}\exp[\tilde{\mathbf{T}}]\mathbf{S}^{\mathrm{T}}.$$

As discussed by Moler and Van Loan (1978), the potential difficulty with this technique lies in errors associated with repeated eigenvalues (which may occur frequently in the current application). However,

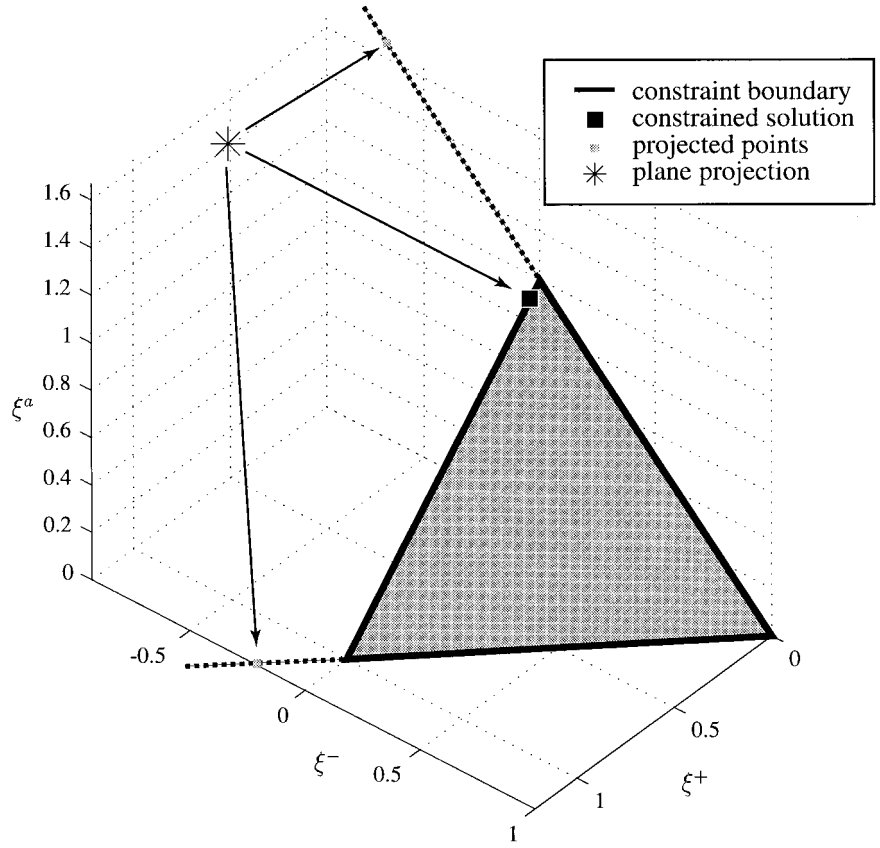


Fig. A1. Geometric interpretation of constraint set. The * locates the projection of a random point into the plane $\xi^a + \xi^- + \xi^+ = 1$, with the bounded triangular area indicating the interior of the constraint set. Once in the plane, a solution satisfying the remaining constraints is found by a closest point projection to the constraint surface.

the initial techniques described by Parlett (1976) for computing functions of a triangular matrix have since been refined to largely overcome such difficulties. While the cost of a Schur decomposition is still conservatively $O(25(v)^3)$ (or $O(25(v-1)^3)$ with reduction in dimensionality) (Demmel, 1997), this coefficient may prove to be smaller than the previous estimate when $\|\tilde{\mathbf{Q}}\|_\infty$ is large.

Constrained integration

Integration of the kinetic equations must be performed in such a manner that the $2v + 1$ constraints

$$\sum_\alpha \xi^\alpha = 1 \tag{A1}$$

$$0 \le \xi^\alpha \le 1 \tag{A2}$$

are satisfied. The tolerance for satisfaction must be strict due to the fact that violations of the constraints leads to nonphysical results which contradict the principle of mass conservation. This implies that for our application the comparative evaluation of integration schemes must take into account not just the norm of error, but the location of that error in volume fraction space (see Fig. A1). Hence, the

preliminary results gathered in choosing an integration technique mentioned above were influenced by the fact that for this class of problems errors in the exponential map still tend to preserve both Eqs. (A1) and (A2), whereas errors in the standard ordinary differential equation solvers tested tended to violate Eqs. (A2). This suggests that the combination of a standard o.d.e. solver with the individual volume fraction constraints may prove an efficient technique. Fig. A1 illustrates how the constraints may be enforced from a simple geometric projection viewpoint. Recently, Govindjee and Hall (1999) have shown this to be the method of choice.

Appendix B

Algorithmic tangent

While the implicit iteration outlined in Section 3 solves the discrete form of the constitutive equations at a local point, it is also necessary to consider the global problem which is arrived at by assembling contributions from each element. In particular, during the global solution it is often advantageous to obtain the tangent to the large scale problem. However, by the nature of the assembly process, the construction of the global tangent may be reduced to the assembly of the local tangents, or rather, the linearization of the local constitutive equations. As was pointed out by Simo and Taylor (1986), the linearization of the constitution must be consistent with the discrete form of the continuum equations, which in general differ from the linearization of the continuum equation set. In this section, the derivation of the consistent (algorithmic) tangent is outlined for the numerical model presented in the body of the paper.

Due to the highly nonlinear nature of the constitution, the consistent tangent is not readily expressible in compact form. As such, only the essential components will be presented. We begin by taking the variation of the discrete one-dimensional stress [see Eq. (33)]

$$\delta \boldsymbol{S}_{n+1} = \frac{\partial \boldsymbol{S}_{n+1}}{\partial \mathbb{C}_{n+1}} \delta \mathbb{C}_{n+1} + \frac{\partial \boldsymbol{S}_{n+1}}{\partial \boldsymbol{E}_{n+1}} \delta \boldsymbol{E}_{n+1} + \frac{\partial \boldsymbol{S}_{n+1}}{\partial \theta_{n+1}} \delta \theta_{n+1} + \frac{\partial \boldsymbol{S}_{n+1}}{\partial \xi^{-}_{n+1}} \delta \xi^{-}_{n+1} + \frac{\partial \boldsymbol{S}_{n+1}}{\partial \xi^{+}_{n+1}} \delta \xi^{+}_{n+1} \tag{B1}$$

where the following terms are easily derived.

$$\frac{\partial \boldsymbol{S}_{n+1}}{\partial \mathbb{C}_{n+1}} = \boldsymbol{E}_{n+1} - \zeta(\theta_{n+1} - \theta_0)$$

$$\frac{\partial \boldsymbol{S}_{n+1}}{\partial \boldsymbol{E}_{n+1}} = \mathbb{C}_{n+1}$$

$$\frac{\partial \boldsymbol{S}_{n+1}}{\partial \theta_{n+1}} = -\zeta \mathbb{C}_{n+1}$$

$$\frac{\partial \boldsymbol{S}_{n+1}}{\partial \xi^{-}_{n+1}} = -\mathbb{C}^{-} \boldsymbol{E}^{-}_{\mathrm{T}}$$

$$\frac{\partial \boldsymbol{S}_{n+1}}{\partial \xi^{+}_{n+1}} = -\mathbb{C}^{+} \boldsymbol{E}^{+}_{\mathrm{T}} \tag{B2}$$

Combining Eqs. (B2), Eq. (B1), and the following expression for the variation of the effective modulus

$$\delta\mathbb{C}_{n+1} = \mathbb{C}^{\mathrm{a}}\delta\xi^{\mathrm{a}}_{n+1} + \mathbb{C}^{-}\delta\xi^{-}_{n+1} + \mathbb{C}^{+}\delta\xi^{+}_{n+1} \tag{B3}$$

it is seen that the primary task is the linearization of the transformation kinetics governing the evolution of the volume fractions. The discrete version of the kinetic relations given in Eq. (36) may be expressed in update form as

$$\boldsymbol{\xi}_{n+1} = \mathbf{M}_{n+1}\boldsymbol{\xi}_n \tag{B4}$$

where the functional relationship $\mathbf{M}_{n+1} = \hat{\mathbf{M}}_{n+1}(\hat{\mathbf{Q}}_{n+1})$ is noted. Then, by temporarily dropping the time subscripts $(n+1)$, denoting $\boldsymbol{\xi}_n$ by $\boldsymbol{\xi}^0$, and adopting indicial notation with summation over repeated indices one finds that

$$\delta\xi_i = \delta M_{ij}\xi^0_j = \frac{\partial M_{ij}}{\partial \tilde{Q}_{kl}}\delta\tilde{Q}_{kl}\xi^0_j = \mathbb{Z}_{ijkl}\delta\tilde{Q}_{kl}\xi^0_j \tag{B5}$$

where $\mathbb{Z}$ is the tangent of the chosen integration method. The evaluation of $\delta\tilde{\mathbf{Q}} = \omega\Delta t\delta\mathbf{Q}$ requires an expression for the variation of each unique term comprising $\mathbf{Q}$. Based on Eq. (29) and the relation $(b_{\alpha\beta})_{n+1} = \hat{b}_{\alpha\beta}(\boldsymbol{S}_{n+1}, \theta_{n+1})$, the symbolic evaluation of a typical term is carried out in the following manner

$$\begin{aligned}
\delta P_{\alpha\beta} &= \exp\left[\frac{-\Delta v b_{\alpha\beta}}{k\theta}\right]\left(\frac{-\Delta v}{k\theta}\delta b_{\alpha\beta} + \frac{\Delta v b_{\alpha\beta}}{k\theta^2}\delta\theta\right)\\
&= P_{\alpha\beta}\left(\frac{-\Delta v}{k\theta}\left(\frac{\partial b_{\alpha\beta}}{\partial\theta}\delta\theta + \frac{\partial b_{\alpha\beta}}{\partial \boldsymbol{S}}\delta\boldsymbol{S}\right) + \frac{\Delta v b_{\alpha\beta}}{k\theta^2}\delta\theta\right)\\
&= P_{\alpha\beta}\left(\frac{-\Delta v}{k\theta}\,\frac{\partial b_{\alpha\beta}}{\partial\theta} + \frac{\Delta v b_{\alpha\beta}}{k\theta^2}\right)\delta\theta + P_{\alpha\beta}\left(\frac{\partial b_{\alpha\beta}}{\partial \boldsymbol{S}}\right)\delta\boldsymbol{S}
\end{aligned} \tag{B6}$$

where once again the time subscripts $(n+1)$ have been neglected in the interest of notational clarity. The remaining terms $\partial b_{\alpha\beta}/\partial\theta$ and $\partial b_{\alpha\beta}/\partial\boldsymbol{S}$ must be worked out individually based on the particular values of α and β indicated in Eq. (30), and as such are not explicitly written here. However, by allowing $\mathbf{H}$ and $\mathbf{L}$ to, respectively, represent the compilation of the first and second terms in Eq. (B6), the variation of the volume fractions is given by

$$\delta\boldsymbol{\xi}_{n+1} = (\omega\Delta t\mathbb{Z}_{ijkl}\mathbf{H}_{kl}\boldsymbol{\xi}^0_j)\delta\theta_{n+1} + (\omega\Delta t\mathbb{Z}_{ijkl}\mathbf{L}_{kl}\boldsymbol{\xi}^0_j)\delta\boldsymbol{S}_{n+1}. \tag{B7}$$

With this result Eq. (B1) may be rewritten accounting for Eqs. (B2), (B3), and (B7). The expression is simplified by adopting the notation

$$\delta\boldsymbol{S}_{n+1} = \tilde{f}_{\mathrm{t}}\delta\theta_{n+1} + \tilde{f}_{\mathrm{s}}\,\delta\boldsymbol{S}_{n+1} + \tilde{f}_{\mathrm{e}}\,\delta\boldsymbol{E}_{n+1} \tag{B8}$$

where the scalars $\tilde{f}_{\mathrm{s}}$, $\tilde{f}_{\mathrm{e}}$, and $\tilde{f}_{\mathrm{t}}$ have the expected meaning. However, since the current implementation treats the temperature as an assigned field, $\delta\theta_{n+1} = 0$, and the tangent required by the code to move the global solution forward is found by isolating the stress term. The result is

$$\delta \boldsymbol{S}_{n+1} = \mathbb{C}^{\text{alg}}_{n+1} \delta \boldsymbol{E}_{n+1} = \left[\frac{\tilde{f}_{\text{e}}}{1 - \tilde{f}_{\text{s}}} \right] \delta \boldsymbol{E}_{n+1} \tag{B9}$$

where $\mathbb{C}^{\text{alg}}_{n+1}$ denotes the algorithmic tangent.

Appendix C

Incorporation of plasticity

In this section the one-dimensional constitution developed thus far is addended with a standard isotropic linear hardening plasticity model. The approach taken is to assume that the shape memory transformations and plastic evolution are uncoupled in the sense that only one will occur at any given time. However, in accordance with experimental observations (Vandermeer et al., 1981), the development of plastic strain is assumed to influence the shape memory effect by hindering the recoverability of the high symmetry phase. Although it is noted that the lack of sufficient experimental evidence suggests that this approach is somewhat speculative, Govindjee and Kasper (1997) have demonstrated that for the data available, the method captures the essential (one dimensional) features in a realistic manner. As such, their proposal is extended to the model presented.

Rather than repeat the well-known continuum plasticity equations, we will work directly with the time integrated equation set presented in Section 3, beginning with the stress computation. The modified form of Eq. (33) is

$$\begin{aligned} \boldsymbol{S}^{\text{tr}}_{n+1} = {} & \hat{\xi}^{-} \mathbb{C}^{-} [\boldsymbol{E}_{n+1} - \boldsymbol{E}^{\text{p}}_{n} - \boldsymbol{E}^{-}_{\text{T}} - \zeta(\theta - \theta_0)] + \hat{\xi}^{+} \mathbb{C}^{+} [\boldsymbol{E}_{n+1} - \boldsymbol{E}^{\text{p}}_{n} - \boldsymbol{E}^{+}_{\text{T}} - \zeta(\theta_{n+1} - \theta_0)] \\ & + \hat{\xi}^{\text{a}} \mathbb{C}^{\text{a}} [\boldsymbol{E}_{n+1} - \boldsymbol{E}^{\text{p}}_{n} - \zeta(\theta_{n+1} - \theta_0)] \end{aligned} \tag{C1}$$

where $\hat{\xi}^{\alpha}$ is defined by $\hat{\xi}^{\alpha} = \xi^{\alpha} + \tilde{\xi}^{\alpha}$, and $\tilde{\xi}^{\alpha}$ represents the volume fraction of the material which is locked in due to the plasticity. This formulation assumes that the remaining portion, ξ^{α}, is free to evolve in the manner presented in Algorithm 1. The evolution of the locked in fractions, $\tilde{\xi}^{\alpha}$, is given by the two parameter function (Govindjee and Kasper, 1997)

$$h(\gamma_{n+1}) = (1 - \vartheta) \exp[-k\gamma_{n+1}] + \vartheta \tag{C2}$$

which gives the recoverable austenite fraction for a given equivalent plastic strain (γ_{n+1}). The plasticity algorithm is thus as given in Algorithm 2.

Algorithm 2 (Plasticity equations)**.**

1. Input at time t^{n+1}:

$$\{\boldsymbol{E}^{\text{p}}_{n}, \mathbb{C}_{n+1}, \boldsymbol{\xi}^{\text{tr}}_{n+1}, \tilde{\boldsymbol{\xi}}_{n}, \gamma_{n}\} \tag{C3}$$

2. Compute trial yield function value:

$$\phi^{\text{tr}} = | \boldsymbol{S}^{\text{tr}}_{n+1} | - (\boldsymbol{S}_{y} + H\gamma_{n}) \tag{C4}$$

3. IF ($\phi_{\rm tr} < 0$) THEN ELASTIC

$$\boldsymbol{S}_{n+1} = \boldsymbol{S}^{\rm tr}_{n+1} \tag{C5}$$

$$\boldsymbol{E}^{\rm p}_{n+1} = \boldsymbol{E}^{\rm p}_{n+1} \tag{C6}$$

$$\gamma_{n+1} = \gamma_n \tag{C7}$$

$$\tilde{\xi}_{n+1} = \tilde{\xi}_n \tag{C8}$$

$$\xi_{n+1} = \xi^{\rm tr}_{n+1} \tag{C9}$$

$$\text{RETURN} \Longrightarrow \text{END} \tag{C10}$$

4. ELSE

$$\Delta\gamma = \phi^{\rm tr}/(\mathbb{C}_{n+1} + H) \tag{C11}$$

$$\boldsymbol{S}_{n+1} = \boldsymbol{S}^{\rm tr}_{n+1} - \Delta\gamma\mathbb{C}_{n+1}\ \text{sign}(\boldsymbol{S}^{\rm tr}_{n+1}) \tag{C12}$$

$$\boldsymbol{E}^{\rm p}_{n+1} = \boldsymbol{E}^{\rm p}_{n} + \Delta\gamma\ \text{sign}(\boldsymbol{S}^{\rm tr}_{n+1}) \tag{C13}$$

$$\gamma_{n+1} = \gamma_n + \Delta\gamma \tag{C14}$$

$$\tilde{\xi}_{n+1} = \tilde{\xi}_n + \left[1 - \frac{h(\gamma_{n+1})}{h(\gamma_n)}\right]\xi^{\rm tr}_{n+1} \tag{C15}$$

$$\xi_{n+1} = \frac{h(\gamma_{n+1})}{h(\gamma_n)}\xi^{\rm tr}_{n} \tag{C16}$$

5. END.

The above algorithm, which is based on the radial return mapping of Simo and Hughes (1998), does not account for simultaneous plastic flow and transformation. Hence, the procedure for computing the composite constitution is as follows. First, Algorithm 1 (with the noted modification to the stress computation) is called to determine $\mathbb{C}_{n+1}$, ξ_{n+1}, and $\boldsymbol{S}_{n+1}$. These values are then entered into Algorithm 2 as $\mathbb{C}_{n+1}$, $\xi^{\rm tr}_{n+1}$, and $\boldsymbol{S}^{\rm tr}_{n+1}$ along with $\tilde{\xi}_n$, γ_n, and $\boldsymbol{E}^{\rm p}_n$. To ensure that the two algorithms do not overlap the conditions $\|\hat{\xi}_{n+1} - \hat{\xi}_n\| < \text{TOL}$ can be used. The algorithmic tangent is set to $\mathbb{C}_{n+1}H/(\mathbb{C}_{n+1}+H)$ if plastic flow has occurred. Note that experimental data suggests that phase transformation and plastic flow regions are well separated as evidenced by the transformation and yield stress values.

References

Abeyaratne, R., Kim, S., 1997. Cyclic effects in shape-memory alloys: A one-dimensional continuum model. International Journal of Solids and Structures 34, 3273–3289.

Abeyaratne, R., Kim, S-J., Knowles, J., 1994. A one-dimensional continuum model for shape-memory alloys. International Journal of Solids and Structures 31, 2229–2249.

Achenbach, M., 1989. A model for an alloy with shape memory. International Journal of Plasticity 5, 371–395.

Achenbach, M., Müller, I., 1985. Simulation of material behavior of alloys with shape memory. Archives of Mechanics 35, 537–585.

Atkin, R., Craine, R., 1976. Continuum theories of mixtures: Basic theory and historical development. The Quarterly Journal of Mechanics and Applied Mathematics 29, 209–244.

Boyd, J., Lagoudas, D., 1996a. A thermodynamic constitutive model for the shape memory materials. Part I. The monolithic shape memory alloy. International Journal of Plasticity 12, 805–842.

Boyd, J., Lagoudas, D., 1996b. A thermodynamic constitutive model for the shape memory materials. Part II. The SMA composite material. International Journal of Plasticity 12, 843–873.

Brinson, L., Bekker, A., 1997. Temperature-induced phase transformation in a shape memory alloy: Phase diagram based kinetics approach. Journal of the Mechanics and Physics of Solids 45, 949–988.

Brinson, L., Lammering, R., 1993. Finite element analysis of the behavior of shape memory alloys and their applications. International Journal of Solids and Structures 30, 3261–3280.

Demmel, J., 1997. Applied Numerical Linear Algebra. Society for Industrial and Applied Mathematics.

Govindjee, S., Hall, G., 1999. Computational aspects of solid–solid phase transformation modeling with a Gibbs function. In: Proceedings Smart Structures and Materials: Mathematics, Modeling and Control, Vol. 3667. The International Society for Optical Engineering, pp. 302–313.

Govindjee, S., Kasper, E., 1997. A shape memory model for Uranium-Niobium accounting for plasticity. Journal of Intelligent Materials, Systems, and Structures 8, 815–823.

Govindjee, S., Kasper, E., 1999. Computational aspects of shape memory alloy modeling with phase diagrams. Computer Methods in Applied Mechanics and Engineering 171, 309–326.

Kasper, E. 1997. Shape memory materials: Constitutive modeling and finite element analysis. Ph.D. thesis, University of California at Berkeley.

Lubliner, J., Auricchio, F., 1996. Generalized plasticity and shape memory alloys. International Journal of Solids and Structures 33, 991–1003.

Moler, C., Van Loan, C., 1978. Nineteen dubious ways to compute the exponential of a matrix. Society for Industrial and Applied Mathematics Review 20, 801–836.

Müller, I., Xu, H., 1991. On the pseudo-elastic hysteresis. Acta Metallurgica 39, 263–271.

Parlett, B., 1976. A recurrence among the elements of functions of triangular matrices. Linear Algebra and its Application 14, 117–121.

Russell, S. 1998. NiTi technical pages. Technical Report, Shape Memory Applications, Inc.

Simo, J., Hughes, T., 1998. Computational Inelasticity. Springer-Verlag, New York.

Simo, J., Taylor, R., 1986. Consistent tangent operators for rate independent elastoplasticity. Computer Methods in Applied Mechanics and Engineering 48, 101–118.

Simo, J., Taylor, R., Hjelmstad, K., 1984. Numerical formulations of elasto-viscoplastic response of beams accounting for the effect of shear. Computer Methods in Applied Mechanics and Engineering 42, 301–330.

Stewart, W., 1994. Introduction to the Numerical Solution of Markov Chains. Princeton University Press.

Vandermeer, R., Ogle, J., Northcutt, W.J., 1981. A phenomenological study of the shape memory effect in polycrystalline Uranium–Niobium alloys. Metallurgical Transactions A, Physical Metallurgy and Materials Science 12A, 733–741.