NUMERICAL MODELING OF NONLINEAR THERMODYNAMICS IN SMA WIRES

DANIEL R. REYNOLDS AND PETR KLOUČEK

Abstract. We present a mathematical model describing the thermodynamic behavior of shape memory alloy wires, as well as a computational technique to solve the resulting system of partial differential equations. The model consists of conservation equations based on a new Helmholtz free energy potential. The computational technique introduces a viscosity-based continuation method, which allows the model to handle dynamic applications where the temporally local behavior of solutions is desired. Computational experiments document that this combination of modeling and solution techniques appropriately predicts the thermally- and stress-induced martensitic phase transitions, as well as the hysteretic behavior and production of latent heat associated with such materials.

1. Introduction

In this paper, we consider a fully-coupled nonlinear model governing the thermodynamic behavior of shape memory alloy (SMA) wires. Shape memory alloys are characterized by their unique ability to undergo a crystal phase transformation, known as the first-order martensitic phase transformation. This transformation induces such materials to achieve up to 8% pseudo-elastic deformation (deformation without material damage) through small changes to their temperature. As a result of this uniquely sensitive responses to small temperature variations, these materials are currently being considered for such multi-ranging tasks as nano-machinery, biomedical implants, self-repairing shielding devices, and active damping applications.

An impediment to the development of mathematical models for shape memory alloys is that their unique behaviors arise from nonlinear thermoelastic material responses. To predict the behavior of such a material whose free-energy potential has multiple local minima, the continuum thermodynamic equations result in non-elliptic, highly degenerate systems of equations. Thus in addition to the nonlinear model, we propose a computational technique to solve this system of equations.

This paper is organized in the following manner. We first discuss the key nonlinear thermodynamic behaviors of shape memory alloys in Section 2. Then Section 3 details our nonlinear thermodynamic model for these materials. We then discuss the computational techniques used to render the model solvable in Section 4, and conclude with computational results verifying the model's validity and predictive properties in Section 5.

2. Relevant Physical Phenomena

Depending on the material temperature, shape memory alloys may exist in one of a variety of different crystal structures. The transition temperatures dividing these phases are contingent on the specific alloy, though

Key words and phrases. Thermodynamic Modeling, Nonconvex Partial Differential Equation Solvers, Shape Memory Alloys.

This work was performed in part under the auspices of the U.S. Department of Energy by University of California Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48. The authors were supported in part by the grant NSF DMS–0107539, by LACSI through LANL contract number 03891–99–23, as part of the prime contract W–7405–ENG-36 between the Department of Energy and the Regents of the University of California, by the NASA grant SECTP NAG5–8136, and by a grant from the Schlumberger Foundation.

UCRL-CONF-201761

Addresses: D.R. Reynolds, Lawrence Livermore National Lab, Box 808, L-551, Livermore, CA 94551

P. Klouček, Dept. of Computational and Applied Math, Rice University, 6100 Main Street, Houston, TX 77005

E-mail: reynoldd@llnl.gov, kloucek@rice.edu

.

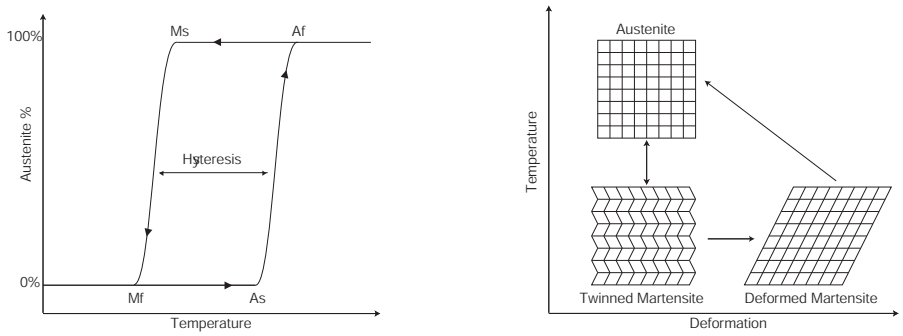


FIGURE 1. Typical hysteresis loop for a martensitic transformation (left): the transformation from austenite to martensite occurs along a different pathway than the reverse.
Twinning and the shape memory effect (right): note the twinning as austenite transforms into martensite. The martensitic lattice is easily deformed under stress. When heated, the deformed martensite resumes the original austenitic configuration (adapted from [24]).

in general these materials exhibit distinct high and low temperature states. The higher-temperature phase is called austenite, and the lower-temperature phase is called martensite. However, there exists a 'middle-range' of temperatures at which either or both phases may be seen, depending on the thermomechanical history of the material. This is indicative of sudden transformations in regions of the crystal lattice between phases as the transformations occur. It is this lack of smoothness in the transition between the two states that distinguishes the first-order phase transformation from other phase transformations, and gives rise to such phenomena as *hysteresis*, *twinning*, the *shape memory effect*, and a release of *latent heat*.

When martensite is heated it begins to transform at the *austenite start temperature* (A_s). This phase change is completed at the *austenite finish temperature* (A_f). Similarly, the reverse transformation exhibits *martensite start* and *martensite finish* temperatures, M_s and M_f, [6]. The transition temperatures (A_s, A_f, M_s, M_f) are typically not the same, in that the heating transformation occurs at a higher temperature than the cooling transformation, see Figure 1. The difference between these temperatures is called the *hysteresis*, and provides a measure of the overall energy absorbed by the material during the phase transformations [13]. Another nonlinear thermal phenomenon unique to the first-order phase transition is the release of *latent heat* [27]. This release of latent heat is associated with an increase in entropy during the transition, due to the conversion between potential and thermal energy [19, 27].

A further behavior unique to shape memory alloys is a phenomenon known as *twinning*. As the alloy is cooled from the austenitic parent phase, martensitic variants are created in a twinned pattern, in which the atoms on opposite sides of *twinning planes* achieve displacements with mirror symmetry [6]. This occurs since the crystal lattice strives to achieve minimal potential energy states for a given temperature [21]. This minimum energy state changes discontinuously as the temperature is lowered between a minimum corresponding to the unique austenitic lattice to minima corresponding to one of many martensitic lattice variants. Thus, during the cooling process, many different martensitic phases are created simultaneously from the original austenitic state, all having equal energy. At this point martensite is easily deformable (Figure 1, right). While most materials deform by slip or dislocation motion, martensite responds to stress by changing the orientation of its crystal lattice through movement of the twin boundaries to the most accomodating variant to the applied stress [24]. However, once the temperature in increased, the crystal lattice returns to the single austenitic orientation, a phenomenon called the *shape memory effect* [24], as seen in Figure 1.

Shape memory alloys permit a second means of phase transformation resulting from stress as opposed to temperature. Here, austenitic shape memory alloys may be transformed into the martensitic phase when loaded with high stresses. As stresses are increased, the crystal lattice transforms into the martensitic variant corresponding to maximum strain and minimum stress, rather than allow plastic effects such as slip or dislocation motion. Upon release of these stresses, the material transforms back into the more stable austenitic phase. This process may be seen in Figure 2.

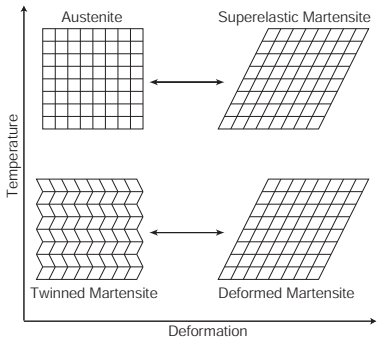


FIGURE 2. Superelasticity: The material begins in the stable austenitic state under zero initial load. As the load increases, the lattice deforms elastically. When the stresses exceed a stress tolerance, the lattice transforms into the metastable martensitic lattice rather than suffer plastic damage.

3. Nonlinear Thermodynamic Model

We model the above behaviors within a continuum thermodynamic framework, in which the material behavior is determined by a *free energy density*. This quantity was first discussed in the framework of general phase transitions by L.D. Landau in the 1940s and A.F. Devonshire in the early 1950s [7, 18]. They postulated that global and local minima of the free energy with respect to the deformation gradient dictate a material's stable and metastable states. The application of these principles to a description of shape memory alloys was first examined by Ericksen in the mid 1980s [8, 9].

Variable	Range	Units	Description
u	$\mathbb{R}^N$	m	deformation
θ	$\mathbb{R}^+ \setminus \{0\}$	K	absolute temperature
γ	$\mathbb{R}^{N\times N}$		deformation gradient, $\gamma = \nabla u$
b	$\mathbb{R}^N$	N kg^{-1}	external body force
r	$\mathbb{R}$	W kg^{-1}	external heat supply
q	$\mathbb{R}^N$	W m^{-2}	referential heat flux
σ	$\mathbb{R}^{N\times N}$	Pa	first Piola-Kirchoff stress tensor
η	$\mathbb{R}$	J kg^{-1} K^{-1}	entropy density
ε	$\mathbb{R}$	J kg^{-1}	internal energy density
Ψ	$\mathbb{R}$	J kg^{-1}	Helmholtz free energy density

TABLE 1. Variables and their units used within the model. All quantities are defined for points (x,t) in the domain $\Omega \times \mathbb{R}^+$.

Remark 3.1. *We make the following remarks on the variables and notation used in the remainder of this paper.*

(1) *The austenitic reference configuration is given by* $u(x) = x$.
(2) *Due to the multiplicity of stable phases, the deformation has limited regularity,* $u \in L^\infty(\mathbb{R}^+; H^1(\Omega))$, *and the deformation gradient* $\gamma(x,t)$ *is only defined* almost everywhere $x \in \mathbb{R}^N$.
(3) *The Helmholtz free energy is defined according to Landau as* $\Psi = \varepsilon - \theta\eta$, [18].
(4) *Time derivatives adhere to the following shorthand,* $\frac{\partial}{\partial t} w(x,t) = \dot{w}(x,t)$.

Our thermodynamic model considers the fully-coupled nonlinear system of equations given by the conservation of mass, conservation of linear momentum, conservation of angular momentum, conservation of energy and the Clausium-Duhem inequality. In the current notation, this is represented by the system

$$\begin{aligned} \rho_0 \ddot{u} &= \mathrm{Div}(\rho_0 \partial_\gamma \Psi + \alpha\dot{\gamma}) + \rho_0 b \\ \rho_0 c_p \dot{\theta} &= \rho_0 \theta \partial^2_{\gamma\theta}\Psi \cdot \dot{\gamma} + \mathrm{Tr}\left(\dot{\gamma}^T \alpha \dot{\gamma}\right) + \mathrm{Div}(\kappa \det(\gamma) \nabla\theta) + \rho_0 r \end{aligned} \tag{3.1}$$

(see [23] for the derivation). Here, ρ_0 is the referential density, κ is the thermal conductivity and c_p is the specific heat capacity. The model assumes that the body is under constant pressure, with possibly changing

volume. A further assumption is that any possible viscous stresses may be incorporated through the term $\alpha\dot{\gamma}$, where $\alpha \in \mathbb{R}^{3\times 3}$ is the symmetric viscous anisotropy matrix having units $\frac{\text{kg}}{\text{m s}}$. Moreover, the model (3.1) uses an inhomogeneous version of the heat flux,

$$q(\gamma, \nabla\theta) = -\kappa \det(\gamma)\nabla\theta, \tag{3.2}$$

where $\kappa \in \mathbb{R}^+ \setminus \{0\}$ is the coefficient of thermal conductivity. This choice incorporates information on the material state (encoded in the deformation gradient γ) into the description of the heat flow through shape memory materials, as suggested by James and Kloučeк [16] and Abeyaratne and Knowles [1].

3.1. Helmholtz Free Energy

We use a one-dimensional form of the Helmholtz free energy, $\Psi(\gamma, \theta)$, which provides the thermoelastic material description within the model (3.1). The construction of this free energy, as well as its relative merits over other forms may be found in [23]. The full form of our Helmholtz free energy is given by

$$\Psi(\gamma,\theta) = C_M(\theta)\mathcal{W}_M(\gamma) + C_C(\theta)\mathcal{W}_C(\gamma) + C_A(\theta)\mathcal{W}_A(\gamma) + c_p(\theta - \theta\ln\theta). \tag{3.3}$$

Here, for martensitic and austenitic elastic moduli, E_m and E_a, the nonconvex martensitic isothermal energy profile $\mathcal{W}_M$ is given by

$$\mathcal{W}_M(\gamma) = \begin{cases} \frac{E_m}{2\rho_0}(\gamma - 1 - \gamma_0)^2, & \gamma < 1-\gamma_0, \\ \frac{E_m}{8\rho_0\gamma_0^2}\left[(\gamma-1)^4 - 2\gamma_0^2(\gamma-1)^2 + \gamma_0^4\right], & \gamma \in [1-\gamma_0, 1+\gamma_0], \\ \frac{E_m}{2\rho_0}(\gamma - 1 + \gamma_0)^2, & \gamma > 1+\gamma_0, \end{cases} \tag{3.4}$$

Similarly, the nonconvex isothermal energy profile at the transition temperature $\mathcal{W}_C$ is given by

$$\mathcal{W}_C(\gamma) = \begin{cases} \frac{E_a+E_m}{4\rho_0}(\gamma - 1 - \gamma_0)^2, & \gamma < 1-\gamma_0, \\ \frac{E_a+E_m}{16\rho_0\gamma_0^4}\left[(\gamma-1)^6 - 2\gamma_0^2(\gamma-1)^4 + \gamma_0^4(\gamma-1)^2\right], & \gamma \in [1-\gamma_0, 1+\gamma_0], \\ \frac{E_a+E_m}{4\rho_0}(\gamma - 1 + \gamma_0)^2, & \gamma > 1+\gamma_0. \end{cases} \tag{3.5}$$

The austenitic isothermal energy profile $\mathcal{W}_A$ is given by

$$\mathcal{W}_A(\gamma) = \frac{E_a}{2\rho_0}(\gamma - 1)^2. \tag{3.6}$$

Lastly, the thermal transition functions $C_M(\theta)$, $C_C(\theta)$ and $C_A(\theta)$ are used to connect the isotherms $\mathcal{W}_M$, $\mathcal{W}_C$ and $\mathcal{W}_A$ together. Denoting the martensite finish temperature as θ_M, the austenite finish temperature as θ_A, and the central temperature where both phases have equal energy as θ_C, these functions are constructed as piecewise cubic splines satisfying the following requirements:

- $C_A(\theta) = 1$ for $\theta \geq \theta_A$, and $C_A(\theta) = 0$ for $\theta \leq \theta_C$,
- $C_M(\theta) = 1$ for $\theta \leq \theta_M$, and $C_M(\theta) = 0$ for $\theta \geq \theta_C$,
- $C_C(\theta) = 1$ for $\theta = \theta_C$, and $C_C(\theta) = 0$ for $\{\theta : \theta \leq \theta_M, \text{or } \theta \geq \theta_A\}$,
- $C_A(\theta) + C_M(\theta) + C_C(\theta) = 1$ for all θ.

The form (3.3) for the free energy has been constructed to reproduce the first-order phase transformation, while remaining rich enough to allow for control over the material constants for different shape memory alloys. Therefore the thermodynamic model (3.1), with the Helmholtz free energy (3.3) at its core, should fully describe the thermodynamic behaviors described in Section 2. Section 5 provides computational results showing that indeed the proposed model fulfills this task. However, we first consider the numerical solution of this system of equations.

4. Solution Technique

4.1. Continuous Space-Time Galerkin

We first consider the temporal and spatial discretizations of the thermodynamic system (3.1), in order to arrive at a finite-dimensional, nonlinear, nonconvex root-finding problem. Our first consideration in this discretization is that the resulting scheme satisfy discrete versions of the original conservation laws. For this purpose we employ a *Continuous Time Galerkin* (CTG) method for the time-discretization of the PDE system. CTG methods were first introduced in the context of ordinary differential equations by Hulme [14], and further expanded by French and Schaeffer [11], and French and Jensen [10]. According to these studies, CTG methods retain continuity in time, so the original conservation principles remain applicable to the resulting discrete systems. Moreover, when combined with finite element approximations of the spatial variables, CTG schemes provide for a unified treatment of both the spatial and temporal discretizations.

To this end, we use piecewise linear approximations for our state variables u and θ in both space and time, as dictated by the underlying regularity of the PDE system. We first consider the following discretizations of our computational domain $[0, L] \times [0, \infty)$,

$$\begin{aligned} &0 = x_0 < x_1 < \ldots < x_N = L, \quad I_j = (x_{j-1}, x_j), \\ &0 = t_0 < t_1 < t_2 < \ldots, \quad J_n = (t_{n-1}, t_n). \end{aligned}$$

Associated with these discretizations, we have the piecewise linear finite element approximation spaces

$$\begin{aligned} \mathcal{S} &= \{\chi \in \mathcal{S} \ : \ \chi|_{I_j} \in P_1(I_j), \ j = 1, \ldots, N\} \subset C([0, L]) \\ \mathcal{V} &= \{\tau \in \mathcal{V} \ : \ \tau|_{J_n} \in P_1(J_n), \ n = 1, \ldots, \infty\} \subset C([0, T]), \end{aligned}$$

and define the overall approximation space $\mathcal{A} = \mathcal{S} \otimes \mathcal{V}$. We also define the inner-product

$$\langle\!\langle u \, , v \rangle\!\rangle_n = \int_0^L \int_{t_{n-1}}^{t_n} u(x,t)\, v(x,t) \, \mathrm{d}t \, \mathrm{d}x.$$

Since the momentum equation from our system (3.1) is a second-order in time ordinary differential equation in u, we introduce the *weak velocity* $v \in \mathcal{A}$ as the weak solution to the equation

$$\langle\!\langle v \, , \varphi \rangle\!\rangle = \langle\!\langle \dot{u} \, , \varphi \rangle\!\rangle, \qquad \forall \varphi \in L^2(J_n; \mathcal{S}).$$

With the above definitions and notation, we find time-level solutions to the model as functions $X(x,t) = [u(x,t), v(x,t), \theta(x,t)] \in \mathcal{A} \times \mathcal{A} \times \mathcal{A}$ providing the weak solution to the first-order in time systems

$$\mathbf{g}^n(X, \varphi) = \begin{pmatrix} \langle\!\langle \dot{u} - v \, , \varphi \rangle\!\rangle_n \\ \left\langle\!\left\langle \rho_0 \dot{v} - \rho_0 \partial^2_{\gamma x} \Psi - \alpha \, \partial^2_{xx} v - \rho_0 b \, , \varphi \right\rangle\!\right\rangle_n \\ \left\langle\!\left\langle \rho_0 c_p \dot{\theta} - \rho_0 \theta \partial^2_{\gamma\theta} \Psi \dot{\gamma} - \alpha \, \dot{\gamma}^2 - \kappa \partial_x \left(\gamma \partial_x \theta\right) - \rho_0 r \, , \varphi \right\rangle\!\right\rangle_n \end{pmatrix} = 0, \tag{4.1}$$

for all test functions $\varphi \in L^2(J_n; \mathcal{S})$, on each time interval J_n.

Due to the piecewise linearity of the approximation space $\mathcal{V}$, the state variables u, v and θ may therefore be represented in the time interval J_n using only the time level states $u(x, t_n)$, $u(x, t_{n-1})$, etc. Furthermore, for computational simplicity we represent the quantities $\gamma(x,t)$, $r(x,t)$, $b(x,t)$, $\partial_\gamma \Psi(x,t)$ and $\partial^2_{\gamma,\theta} \Psi(x,t)$ linearly in time as well. We assume that errors introduced through these linear approximations are of the same magnitude as those introduced through the piecewise linear approximation of u, v and θ. For notational simplification, we denote these time-level states as $u(x, t_n) = u^n(x)$. Lastly, due to the piecewise-linear finite element discretization in space, the nonlinear system (4.1) may be equivalently written for the coefficients $X^n_{\Delta x} = [u^n_1, u^n_2, \ldots, u^n_N, v^n_1, \ldots, v^n_N, \theta^n_1, \ldots, \theta^n_N]$ as the finite-dimensional nonlinear system

$$\mathbf{g}^{n, \Delta x}(X^{n-1}_{\Delta x}, X^n_{\Delta x}, \varphi) = 0, \quad \text{for all} \varphi \in L^2(J_n, \mathcal{S}). \tag{4.2}$$

4.2. Krylov Inexact Newton Solver

In order to solve the finite-dimensional nonlinear system (4.2), we employ a standard Krylov Inexact Newton method (see [5, 15]). We add to this a standard backtracking linesearch method to provide both a robustification strategy as well as diagnostic information on the state of the system (see the next section).

In particular, for the Krylov Inexact Newton system solution, we use a preconditioned GMRES algorithm, with the preconditioner resulting from an incomplete LU factorization of the sparse Newton system matrix. We tried a variety of standard inexactness parameters and linesearch methods, all having relatively similar performance, solving the system (4.2) very quickly for the majority of time steps of each simulation. However, during moments of phase transition, the standard solver proved inadequate. The adjustments to the solution method during these moments of difficulty are discussed in the following section.

4.3. Viscosity-Based Continuation

The central difficulty in solving the system (4.2) occurs at moments when regions of the simulated shape memory wire go through the first-order phase transformation. These numerical difficulties, as well as the transition itself, result from the nonconvexity of the free energy Ψ. Theoretical results have only been able to show existence of weak solutions for similar models incorporating large viscosity coefficients α (see [12, 19]). However, viscous effects are not observed in physical tests on shape memory alloys (see Seelecke et al. [25, 26]). Furthermore, Antman and Seidman have shown that the addition of such linear viscous effects violates basic physical principles such as material frame indifference and observer invariance (see [2, 3]). We are thus faced with the difficulty of balancing solvability of the problem (represented by an inflated viscosity coefficient α) with accuracy in the model solution (represented by the zero-viscosity solution).

In order to balance these contrary demands, we propose the following viscosity-based continuation method. We detect moments of phase transition by monitoring iterates in the linesearch method. Once these iterates show signs of stagnation, and thus phase transition, the viscosity coefficient α is increased to a level proposed by the theoretical results in order to render the model well-posed. The aforementioned inexact Newton method is then applied to solve this perturbed system for a perturbed solution at the current time step. From this point, the viscosity level α is progressively lowered, using the previous viscosity-perturbed solutions to determine the initial guesses for the Inexact Newton solver to better calculate the same time step. For this viscous continuation, we attempted using initial guesses as both the previous viscosity-perturbed solution as well as a linearly extrapolated guess from the previous two viscosity-perturbed solutions. Both worked equally well.

This algorithm is designed to serve two key purposes. First, it assures that the overall time-stepping scheme remains stable by reverting to the solution of a perturbed system during moments of computational difficulty. Second, the continuation method shown here reduces the amount of error in the solution introduced through the artificial dissipation. In computational tests, this method typically achieves solutions using viscosity coefficients at least an order of magnitude smaller than those required without reduction of the viscosity (i.e. [12, 22]).

5. Computational Verification of the Model

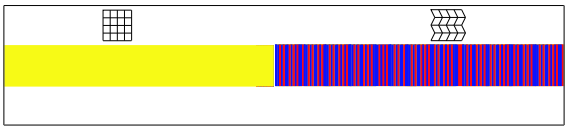
The computational results in this section are designed to provide numerical evidence that the proposed model is indeed able to reproduce the range of nonlinear behaviors unique to shape memory alloys. Thus in the following sections we show results on the thermally- and stress-induced phase transformations; we also provide evidence that these phase changes give rise to hysteretic and latent heat effects.

For the calculations, we consider the wire to have length L with small cross-sectional area β. The heating of the SMA wire is controlled through the term $r(x,t)$. The vibrational forces are implemented by the term $b(x,t)$, and affect the material at the $x = L$ end of the wire. This end of the wire is considered

thermally insulated. The $x = 0$ end of the wire is held fixed in both deformation and temperature. The material constants used here are those of the SMA NiTi: $\rho_0 = 6.45 \times 10^3 \frac{\text{kg}}{\text{m}^3}$, $\kappa = 10 \frac{\text{W}}{\text{m K}}$, $c_p = 322 \frac{\text{J}}{\text{kg K}}$, $E_a = 7.5 \times 10^{10}$Pa, $E_m = 2.8 \times 10^{10}$Pa, $\gamma_0 = 0.02$, $L = 0.05$m, $\beta = 2 \times 10^{-8}\text{m}^2$, $\theta_A = 350$K, $\theta_C = 335$K, and $\theta_M = 320$K.

In the following examples, we make use of color plots to represent the various material phases at different locations along the wire. Since the material phases at equilibrium correspond to $\gamma(x) = 1$ for austenite and $\gamma(x) = 1 \pm \gamma_0$ for martensite, and also since these dynamic experiments allow for deviations away from equilibrium, we may choose between the phases in the following manner at each point x in the wire: if $\gamma(x) \geq 1 + \gamma_0/2$, it is colored as martensite$^+$ (blue or black); if $\gamma(x) \leq 1 - \gamma_0/2$, is is colored as martensite$^-$ (red or dark grey); otherwise it is colored as austenite (yellow or light grey). An example of this technique is given in the following Figure.

FIGURE 3. Phase Visualization Technique: Note the austenitic regions in the left half, and twinned martensitic regions in the right.



5.1. Thermally-Induced Phase Transformations

The first computational experiments that we present exhibit the thermally-induced phase transformation. These numerical experiments were performed under constant heating (or cooling) of the wire through the external heat supply term r. We begin with the wire as relaxed austenite, with velocity set to zero and initial temperature at the transition temperature θ_C. From the beginning of the simulation, the heat supply term is held constant at $r(x,t) = -6$ kJ g^{-1} in order to lower the temperature of the wire into the martensitic temperature range. We present time level states of the temperature distribution, and phase plots in the wire at the times $t = 0.1$ ms, 0.9124 ms, 0.9204 ms, and 1.7288 ms.

As shown in the temperature plots in Figure 4, the temperature decreases from the transition temperature $\theta_C = 335$K down to near the martensite finish temperature at 320K. At this point, however, distributed regions of the wire begin transforming to the martensite phase. As this transformation occurs, the temperature of the transforming regions jumps due to the latent heat effects. Moreover, phase plots in Figure 4 show that the transformation occurs in (distributionally) about half of the wire at first. Then because of the thermal increase, more heat must be removed from the wire in order to finish the transformation to martensite. Finally after the transformation is complete the heat again smoothes out through the length of the wire.

We may similarly consider the phase plots in Figure 4. As the temperature is decreased, we see that as time goes on the austenitic state destabilizes, leaving alternating regions of austenite and martensite variants, until the transformation is complete and the wire resides in the stable martensitic configuration. A key item of note in these plots is that due to the multiplicity of equally-stable martensitic variants, the resulting fully-transformed martensitic state of the wire is not uniformly twinned throughout the length of the wire. These deviations from the perfectly twinned configuration are realistic of nature, where the sudden formation of martensite upon cooling results in mostly twinned structures, along with possibly random variations from the fully-twinned state.

We note that the reverse transformation from the low-temperature martensitic state to the high-temperature austenite state is similar to the above transformation.

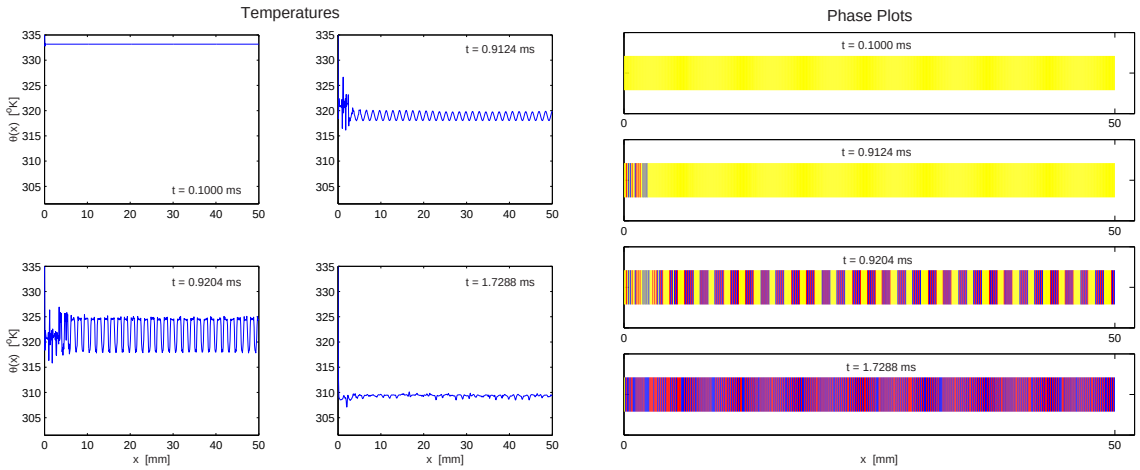


FIGURE 4. Temperatures and Phases for the Thermally-Induced Transformation: Note the non-smooth latent heat effects as each lattice element transforms independently, as well as the eventual smoothing of the temperature in time (left). As the temperature is decreased, the austenite transforms in a distributed fashion into the different martensite variants (right).

5.2. Stress-Induced Phase Transformation

A key reason for the consideration of nonlinear models as opposed to linearized models, is the ability of nonlinear models to handle *both* the thermally- and stress-induced phase transformations [19]. We thus consider the following computational experiment testing the ability of the model to handle the stress-induced phase transformation, as discussed in Section 2. In this experiment, we begin with the wire as relaxed austenite, with zero initial velocity and with initial temperature $\theta_0 = 345K$ (between θ_C and θ_A). The experiment runs for a simulation time of 2 ms. In the first 0.5 ms, body forces linearly increase on the rightmost end of the wire from 0 to $1 \times 10^5 \frac{\mathrm{N}}{\mathrm{kg}}$, providing enough tensile force to induce the transformation to martensite$^+$. Then the body forces decrease linearly to $-1 \times 10^5 \frac{\mathrm{N}}{\mathrm{kg}}$ for 1 ms, first relaxing the wire and then compressing it back through the austenite state and into the martensite$^-$ state. For the final 0.5 ms of the simulation, the body forces then linearly return to the relaxing state of $0 \frac{\mathrm{N}}{\mathrm{kg}}$.

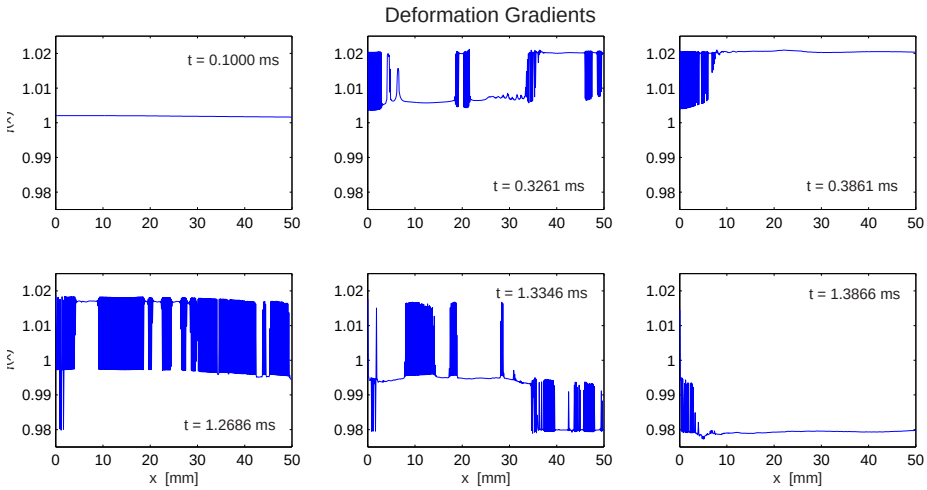


FIGURE 5. Deformation Gradient Plots for the Stress-Induced Phase Transformations: Note the distinct clustering near the various stable states at $\gamma \in \{1, 1+\gamma_0, 1-\gamma_0\}$.

These effects may be seen in the time-level states of the deformation gradients in Figure 5. We plot the states at the time values $t = 0.1$ms, 0.3261ms, 0.3861ms, 1.2686ms, 1.3346ms and 1.3866ms. Examining the deformation gradients in time, we see the effects of increasing tensile load in the first three plots, showing the sudden jumps in various parts of the wire from the austenite phase to the martensite$^+$ phase. As with the temperature-induced transformation, small regions of the wire transform seemingly independently from each other, forming regions of mixed phase in the wire, until the load has increased enough to pull the entire wire into the martensite$^+$ phase. The final three plots show the subsequent compression of the wire, in which lattice elements in the wire first begin to transform back to the austenite state ($t = 1.2686$ms), and then due to the large compressive force they continue the transformation into the martensite$^-$ state.

5.3. Hysteresis

In addition to the thermally- and stress-induced phase transformations, our nonlinear model also accounts for the hysteresis discussed in Section 2. Upon collecting simulation data from repeated thermally-induced phase transformations, we arrive at the phase vs. temperature plot in the following Figure. We used data from 8 experiments, in which the heat supplies varied in the range ± 4 to $\pm 10 \frac{\text{kJ}}{\text{g ms}}$. Each of these simulations calculated 2,000 time steps, resulting in 16,000 data points on the plot representing the percentage of austenite in the wire as a function of the average temperature.

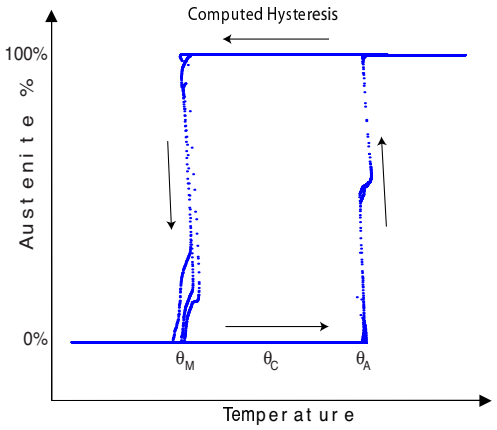


Figure 6. Computed Hysteresis loop. The plot is assembled using a number of simulations. The average state of the wire at each time level is represented by one point. Note the hysteretic gap between transformation temperatures, as predicted in Figure 1.

5.4. Latent Heat

As has been evidenced in the computational experiments of the thermally-induced phase transformation in Section 5.1, the proposed model accounts for the effects of the latent heats of transformation. Moreover, these effects are implicitly accounted for by the nonlinearity of the model, and do not need to be supplied as additional terms, as must be done in many linearized models [4, 17, 20]. The latent heat may be directly observed through changes in the overall enthalpy of the wire in time. At constant pressure, the total enthalpy (in Joules) of the wire at the time t is given by

$$H(t) = \int_\Omega \rho_0 \, \theta(\mathbf{x}, t) \, \eta(\mathbf{x}, t) \, \mathrm{d}\mathbf{x} = \rho_0 \beta \int_0^L \theta(x, t) \, \eta(x, t) \, \mathrm{d}x. \tag{5.1}$$

Figure 7 shows the time histories of the enthalpy from two thermally-induced phase transformations. The plot on the left corresponds to the overall enthalpy resulting from the martensite to austenite transformation, and the plot on the right corresponds to the reverse transformation. The items to note in these plots are the sudden enthalpy jumps, corresponding to the latent heats of transformation. Moreover, published values of latent heat for these materials provide the latent heat coefficient per unit volume of NiTi as $k_l = 1.56 \times 10^8 \frac{\text{kg}}{\text{m s}^{-2}}$. Converting this number into units corresponding to the total change in enthalpy

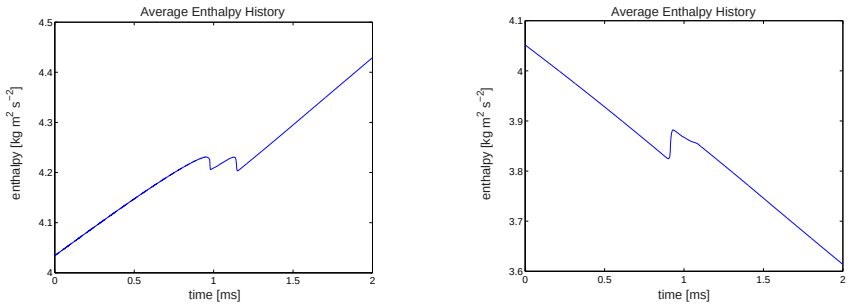


FIGURE 7. Enthalpy Plots Showing Latent Heat Effects: The thermally-induced martensite to austenite transformation is on the left, and the reverse transformation is on the right.

gives an enthalpy change of $0.156\frac{\text{kg m}^2}{\text{s}^2}$. Thus, we see that the predicted latent heats of transformation from Figure 7 are within a factor of two of the measured values.

6. Conclusions

Due to our retention of the full nonlinear model (3.1), as well as the proposed solution method from Section 4, this model accurately describes the nonlinear physics of shape memory alloy wires. Thus it may provide key information predicting the material response expected from various shape memory alloy wires, without the need for possibly detailed and costly experimentation. This information could then be used to aid in the design of macroscale devices using shape memory alloy wires. Additionally, through further research into the behavior of the model, controlling methods may be investigated for taking advantage of the nonlinear behaviors unique to shape memory alloys.

An area for related work lies in the expansion of the current model to higher dimensions. We believe that the general nonlinear thermodynamic model in Section 3, and the overall solution method proposed in Section 4 would work similarly for the higher-dimensional cases of thin films and solids

References

[1] R. Abeyaratne and J. Knowles. On the driving traction acting on a surface of strain discontinuity in a continuum. *Journal of Applied Mechanics*, 38:345–360, 1990.
[2] S.S. Antman. Physically unacceptable viscous stresses. *Z. Angew. Math. Phys.*, 49:980–988, 1998.
[3] S.S. Antman and T.I. Seidman. Quasilinear hyperbolic-parabolic equations of one-dimensional viscoelasticity. *J. Diff. Eqns.*, 124:132–185, 1996.
[4] A. Bekker, L.C. Brinson, and K. Issen. Localized and diffuse thermoinduced phase transformation in 1D shape memory alloys. *J. Intell. Mater. Syst. and Struct.*, 9:355–365, 1998.
[5] P.N. Brown and Y. Saad. Hybrid krylov methods for nonlinear systems of equations. *SIAM J. Scientific and Statistical Computing*, 11:450–481, 1990.
[6] W.J. Buehler and F.E. Wang. A summary of recent research on the nitinol alloys and their potential application in ocean engineering. *Ocean Eng.*, 1:105–120, 1967.
[7] A.F. Devonshire. Theory of ferroelectrics. *Adv. Phys.*, 3:86–130, 1954.
[8] J.L. Ericksen. On the symmetry of deformable crystals. *Arch. Rational Mech. Anal.*, 72:1–13, 1979.
[9] J.L. Ericksen. Constitutive theory for some constrained elastic crystals. *J. Solids and Structures*, 22:951–964, 1986.
[10] D.A. French and S. Jensen. Behaviour in the large of numerical solutions to one-dimensional nonlinear viscoelasticity by continuous time Galerkin methods. *Computer Methods in Applied Mechanics and Engineering*, 86:105–124, 1991.
[11] D.A. French and J.W. Schaeffer. Continuous finite element methods which preserve energy properties for nonlinear problems. *Appl. Math. Comput.*, 39:271–295, 1990.
[12] G. Friesecke and G. Dolzmann. Implicit time discretization and global existence for a quasi-linear evolution equation with nonconvex energy. *SIAM J. Math. Anal.*, 28(2):363–380, March 1997.
[13] H. Funakubo. *Shape Memory Alloys*. Gordon and Breach Science Publishers, New York, 1987.
[14] B.L. Hulme. Discrete Galerkin and related one-step methods for ordinary differential equations. *Math. Comp.*, 26:881–891, 1972.

[15] C.T. Kelley. *Iterative Methods for Optimization*. SIAM, Philadelphia, 1999.
[16] P. Klouček. Toward the computational modeling of nonequilibrium thermodynamics of the martensitic transformations. *IMA Preprint Series*, 1418, July 1996.
[17] D.C. Lagoudas, Z. Bo, and M.A. Qidwai. A unified thermodynamic constitutive model for SMA and finite element analysis of active metal matrix composites. *Mechanics of Composite Materials and Structures*, 3:153–179, 1996.
[18] E.M. Lifshitz and L.P. Pitaevskii. *Statistical Physics by L.D. Landau and E.M. Lifshitz*, volume 5 of *Course of Theoretical Physics, Part I*. Pergamon Press, Oxford, 3rd edition, 1980.
[19] M. Niezgódka and J. Sprekels. Existence of solutions for a mathematical model of structural phase transitions in shape memory alloys. *Math. Meth. Appl. Sci.*, 10:197–223, 1988.
[20] E.R. Oberaigner, K. Tanaka, and F.D. Fischer. Damping of rod vibrations caused by impact. *SPIE*, 3667:566–577, 1999.
[21] K. Otsuka and C.M. Wayman. *Shape Memory Materials*, chapter Mechanism of Shape Memory Effect and Superelasticity, pages 27–48. Cambridge University Press, 1998.
[22] R.L. Pego. Stabilization in a gradient system with a conservation law. *Proc. Amer. Math. Soc.*, 114, 1992.
[23] D.R. Reynolds. *A Nonlinear Thermodynamic Model for Phase Transitions in Shape Memory Alloy Wires*. PhD thesis, Rice University, Houston, Texas, May 2003.
[24] J. Ryhänen. *Biocompatibility Evaluation of Nickel-Titanium Shape Memory Metal Alloy*. PhD thesis, University of Oulu, 1999.
[25] S. Seelecke. Modeling the dynamic behavior of shape memory alloys. *International Journal of Non-Linear Mechanics*, 37:1363–1374, 2002.
[26] S. Seelecke and I. Müller. Shape memory alloy actuators: Modeling and simulation. *ASME Applied Mechanics Reviews*, 56(6), 2003.
[27] P. Wollants, J.R. Roos, and L. Delaey. Thermally- and stress-induced thermoelastic martensitic transformations in the reference frame of equilibrium thermodynamics. *Progress in Materials Science*, 37:227–288, 1993.