One-Dimensional Thermomechanical Constitutive Relations for Shape Memory Materials

C. LIANG AND C. A. ROGERS*
Smart Materials and Structures Laboratory, Mechanical Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

ABSTRACT: The use of the thermoelastic martensitic transformation and its reverse transformation has recently been proposed and demonstrated for several active control applications. However, the present constitutive models have lacked several important fundamental concepts that are essential for many of the proposed intelligent material system applications such as shape memory hybrid composites.

A complete, unified, one-dimensional constitutive model of shape memory materials is developed and presented in this paper. The thermomechanical model formulation herein will investigate important material characteristics involved with the internal phase transformation of shape memory alloys. These characteristics include energy dissipation in the material that governs the damping behavior, stress-strain-temperature relations for pseudoelasticity, and the shape memory effect. Some numerical examples using the model are also presented.

INTRODUCTION

SEVERAL books [1–3] and many technical papers have been devoted to shape memory alloy (SMA) materials. They describe the fundamental concepts associated with the martensitic phase transformation characteristics of Nitinol and other shape memory materials. Because of the particular characteristics of this material, namely the shape memory effects (SME) and pseudoelasticity, many applications have been proposed that can utilize the material's force or stiffness transduction capabilities [2–4]. Rogers [5] has demonstrated that SMA materials can be embedded in composite structures to actively control and modify the dynamic and structural behavior of SMA hybrid composite materials. SMA hybrid composite materials have become one of the important material system configurations of today's intelligent material systems and structures. This concept is, however, limited by our present knowledge and understanding of the behavior and controllability of the shape memory alloy constituent. An extensive and organized data base of the material's behavior and mechanical/physical characteristics is desperately needed. The development of a unified and theoretical constitutive model appears to be a critical path element in the development of SMA based intelligent material systems and structures.

There are two approaches to establish a constitutive relation for any material. One is the macroscopic phenomenological method that requires a significant amount of experimental work; the other is the microscopic physical method that derives the constitutive relation from fundamental physical concepts. The phenomenological approach is often used in engineering practice; however, it can rarely explain the physics behind the material's behavior or character. The microscopic physical method can successfully provide the fundamental explanations to different experimental phenomena, however, its numerical predictions and simulations are often complex and distant from phenomenological observations.

Besides the conventional treatment to the stress-induced martensitic phase transformation [1–3,6], Müller and his colleagues proposed another model based on shape memory effect phenomenology, thermodynamics and statistical physics [7–11]. For a small lattice, the potential energy at a certain temperature is schematically shown in Figure 1. There are two stable minima corresponding to the martensitic twins and a metastable center for the austenitic state as shown in Figure 1. The shear length of each layer of the martensitic lattice is denoted by Δ, and the stable martensitic phases are denoted by M_+ and M_-. If an external force P is applied, the total potential energy is given by $\Phi - P\Delta$, as shown in Figure 2. At a lower temperature, the material particles lie still in their potential well. The yielding from M_- to M_+ will occur when the applied external load is so large that the potential barrier on the left is eliminated. This condition is shown on the top of Figure 2. The bottom of Figure 2 refers to higher temperatures at which particles are fluctuating about their minima with a mean kinetic energy that is proportional to the temperature. The height of the pools of the potential wells of Figure 2 indicates the strength of the fluctuation. The figure indicates that at higher temperatures the yielding from M_- to M_+ will occur at a lower load than that at lower temperatures. If there is such a fluctuation, it may be described using statis-

This article was reprinted from the April 1990 issue of the *Journal of Intelligent Material Systems and Structures*, Vol. 1, No. 2, pp. 207–234.

*Author to whom correspondence should be addressed.

1045-389X/97/04 0285-18 $10.00/0

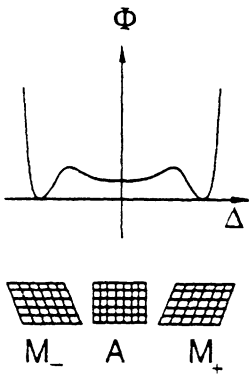

Figure 1. *Potential energy of a lattice vs. shear deformation length [11].*

tical mechanics. Assuming the distribution function of phase k is N_k in a single lattice, the total deformation $D - D_0$ may be written as:

$$D - D_0 = \frac{1}{\sqrt{2}} (N_{M-} \Delta_{M-} + N_{M+} \Delta_{M+} + N_A \Delta_A) \quad (1)$$

The thermodynamics behind the phase transformation can be described as: "the energy E tries to minimize by pulling all particles into the depths of the potential wells and the entropy S attempts to maximize by distributing the particles evenly over the available range of shear lengths. In this competition, it is the free energy,

$$\Psi = E - TS \quad (2)$$

that achieves a minimum" [11]. The free energy and the distribution function N_k can be related using statistical mechanics. The minimization of the free energy may give the explicit expression of $D - D_0$ and martensitic volume fraction ξ in terms of the applied load and temperature.

Tanaka [12] developed a model in 1982 using a concept similar to Müller's, in which the phase transformation is basically governed by the minimization of the free energy. The

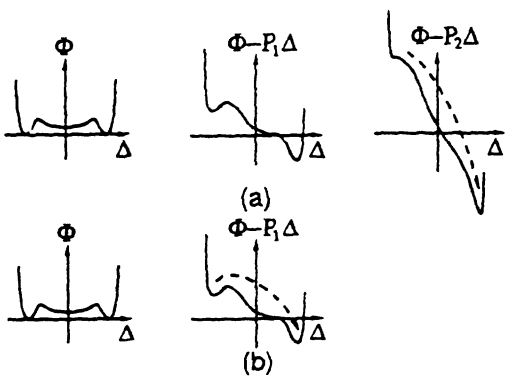

Figure 2. *Influence of external force and temperature on Φ-Δ relation [11].*

energy equation and Clausius-Duhem inequality are used to describe the hysteresis associated with the phase transformation. This model has been used to study superelasticity [13], pseudoelasticity, and SMA materials [14–16], qualitatively.

Tanaka's model considers a one-dimension metallic material of length L that is undergoing either martensitic transformation or its reverse transformation. The energy balance equation and Clausius-Duhem inequality in the reference X-coordinate are expressed as:

$$\begin{aligned} \rho \dot{U} - \sigma L + \frac{\partial q_{sur}}{\partial X} - \rho q &= 0 \\ \rho \dot{S} - \rho \frac{q}{T} + \frac{\partial}{\partial X}\left(\frac{q_{sur}}{T}\right) &\geq 0 \end{aligned} \quad (3)$$

where ρ is the density in the current deformed configurations, σ the Cauchy stress, U and q_{sur} the internal energy density and the heat flux from the surroundings, and T, q, and S are the temperature, internal heat source, and entropy density, respectively.

The state variables for the material are stress, strain, temperature, and extent of phase transformation, ξ, which is defined as the martensite fraction. Since only three of the above variables are independent, the state variable is defined as:

$$\Lambda \equiv (\bar{\varepsilon}, t, \xi) \quad (4)$$

The Helmholtz free energy defined in Equation (2) is a function of the state variable Λ. Equations (2) and (4) can be used to express the Clausius-Duhem inequality in the reference configuration as a function of the state variable. Because the SME involves large deformations, the Green strain, $\bar{\varepsilon}$, and Piola-Kirchhoff stress, $\bar{\sigma}$, are used. The new form of the Clausius-Duhem inequality may be given as:

$$\begin{aligned} &\left(\bar{\sigma} - \rho_0 \frac{\partial \Psi}{\partial \bar{\varepsilon}}\right)\dot{\bar{\varepsilon}} - \left(S + \frac{\partial \Psi}{\partial T}\right)\dot{T} - \frac{\partial \Psi}{\partial \xi}\dot{\xi} \\ &- \frac{1}{\rho_0 T}\frac{\rho}{\rho_0} q f^{-1} \frac{\partial T}{\partial X} \geq 0 \end{aligned} \quad (5)$$

where f is the deformation gradient and ρ_0 is the density in the reference configurations. From the thermodynamics of continuous media, the coefficients of $\dot{\bar{\varepsilon}}$ and $\dot{T}$ should vanish. Therefore, the inequality of Equation (5) provides the mechanical constitutive equation as:

$$\bar{\sigma} = \rho_0 \frac{\partial \Psi}{\partial \bar{\varepsilon}} = \sigma(\bar{\varepsilon}, T, \xi) \quad (6)$$

From the above equation, the rate form of the mechanical constitutive equation is obtained as:

$$\dot{\bar{\sigma}} = \frac{\partial \sigma}{\partial \bar{\varepsilon}} \dot{\bar{\varepsilon}} + \frac{\partial \sigma}{\partial T} \dot{T} + \frac{\partial \sigma}{\partial \xi} \dot{\xi} = D\dot{\bar{\varepsilon}} + \Theta \dot{T} + \Omega \dot{\xi} \quad (7)$$

where D is the Young's modulus, Θ the thermoelastic tensor, and Ω the transformation tensor, a metallurgical quantity. These material properties derived from thermomechanics are given as:

$$D = \rho_0 \frac{\partial \Psi^2}{\partial \bar{\varepsilon}^2} \qquad \Theta = \rho_0 \frac{\partial \Psi^2}{\partial \bar{\varepsilon} \partial T} \qquad \Omega = \rho_0 \frac{\partial \Psi^2}{\partial \bar{\varepsilon} \partial \xi} \quad (8)$$

Assuming an exponential form for the relation of martensite fraction and temperature during the phase transformations, and using the relation given in Equation (7), Tanaka studied the stress-strain behavior of SMA materials qualitatively. The martensite fraction during the phase transformation is of the following form:

$$\begin{aligned} \xi_{M \to A} &= \exp[A_a (T - A_s) + B_a \sigma] \\ \xi_{A \to M} &= 1 - \exp[A_m (T - M_s) + B_m \sigma] \end{aligned} \quad (9)$$

where A_a, A_m, B_a, and B_m are material constants in terms of the transition temperatures, A_s, A_f, M_s, and M_f, etc.

Based on the potential and deflection relations, as shown in Figures 1 and 2, and thermomechanical free energy concepts, Niezgodka, Sprekels, and Hoffman [17–21] suggested a dynamic model for SMA materials. Like Tanaka's model, their model uses the same governing equations including the energy equation and Clausius-Duhem inequality. The only difference is that instead of solving the governing equations, Tanaka gives an explicit equation in terms of stress, strain, temperature, and martensite fraction, whereas Niezgodka, Sprekels, and Hoffmann actually solve the governing equations by assuming an expression of free energy in terms of strain and temperature.

These models essentially explain the physical nature of the shape memory effect associated with the martensitic phase transformations. However, most of these models cannot be applied easily in mechanical engineering design and computations. These models do give a good qualitative assessment but generally require thermomechanical parameters that are either impractical or difficult to measure.

Cory and McNichols have developed a thermomechanical model based on so-called non-equilibrium thermostatics (NET) systems [22–24]. If a thermodynamic system undergoes an equilibrium path, which means that every position of the path must correspond to an equilibrium state, the process is required to take place very slowly (i.e., quasi-static). Such processes are likewise reversible. Since shape memory alloys involve and irreversible phase transformation and have macroscopic hysteric characteristics, their thermodynamics paths are described by a non-equilibrium thermostatic process. For NET systems, instead of using the inequality form of the second law of thermodynamics, a path-dependent heat generation term is introduced and the newly modified second law of thermodynamics is given as:

$$TdS = dQ + Td\bar{S} \quad (10)$$

where dQ is the path-dependent heat into the system, whereas the newly introduced term $d\bar{S}$ is the internally generated entropy change for a differential path element relating directly to the energy dissipation of the hysteresis. Cory and McNichol summarized the following points [20]:

1. All thermodynamic paths of force, length of SMA wire, and temperature (FLT) in its state space are limited within a bounded volume.
2. The surface of the bounded volume is the asymptotic boundary of the paths.
3. All observed paths are one of two types, each type having a characteristic "direction" in state space. The two types are defined as AM (austenite to martensite) and MA (martensite to austenite). Therefore, as shown in Figure 3, the surfaces of the bounded volume are distinguished as AM and MA surfaces.
4. The direction of the path is related to the increase or decrease of a new state variable Z defined especially for the NET system. The "history" dependence of SMA state variables (FLT) is completely determined by the FLT location (called F_0, L_0, and T_0) of the last change in the sign of the differential dZ and is independent of the history of the system prior to that time.
5. The initial slope of each single controlled state variable

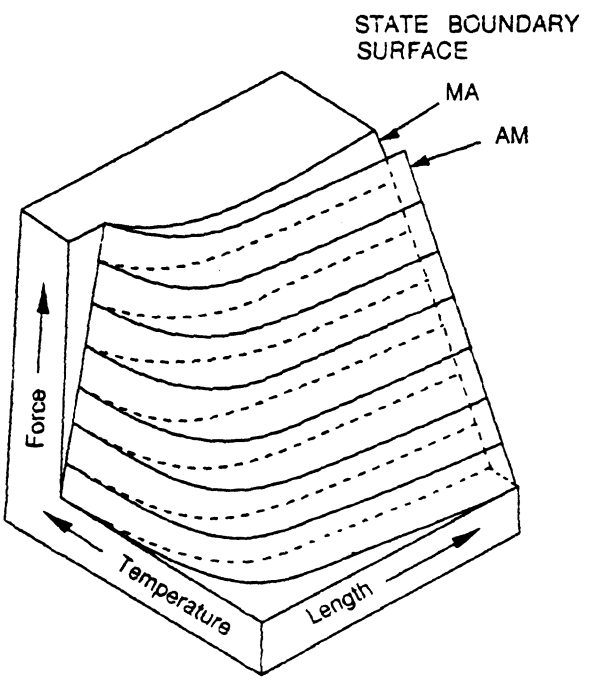

Figure 3. *Schematic diagram of nitinol (helix) state space [24].*

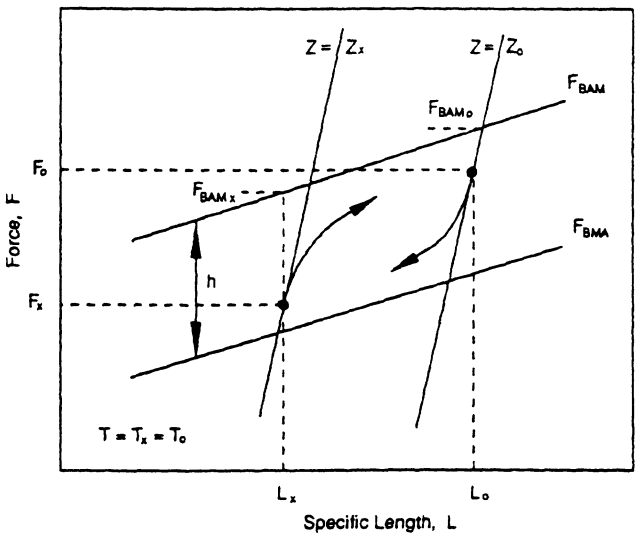

Figure 4. *Schematic isothermal cut of nitinol state space [24].*

path is nearly the same and is tangent to a set of Z surfaces in FLT space.

6. There exist reversible paths in the FLT space that form a surface of constant Z value.

Figure 4 gives a schematic isothermal cut of the SMA state space from which the phenomenological equations of the stress-strain curves can be determined. First, the AM surface is given by:

$$F_{BAM}(L,T) = (a - f)L + (b + g)T - N \qquad (11)$$

where *BAM* represents the "boundary surface of the martensitic phase transformation". The MA surface is given by:

$$F_{BMA}(L,T) = F_{BAM}(L,T) - h \qquad (12)$$

where *BMA* represents "boundary surface of the MA transformation". The Z surface, which is tangent to the initial slope, is described by:

$$Z = aL + bT - F \qquad (13)$$

The stress-strain AM path is phenomenologically described by:

$$F_{BAM} - F = (F_{BAMx} - F_x)\exp\left|\frac{g(T - T_x) - f(L - L_x)}{F_{BAMx} - F_x}\right| \qquad (14)$$

where the history variables (L_x, T_x, F_x), correspond to the minimum Z at $Z_x = Z(L_x, T_x, f_x)$, and the subscript x of F_{BAM} and F indicate their corresponding values at (L_x, T_x). For the stress-strain curve of the MA path, the equation is expressed as:

$$F - F_{BMA} = (F_{BMA0} - F_0)\exp\left|\frac{g(T_0 - T) - f(L_0 - L)}{F_0 - F_{BMA0}}\right| \qquad (15)$$

where (L_0, T_0, F_0), correspond to the maximum Z at $Z_0 = Z(L_0, T_0, F_0)$, and the subscript "0" of F_{BMA} and F indicate their corresponding values at (L_0, T_0). In the above equations, a, b, g, f, h, and N are measurable material constants. Besides the stress-strain relations, this model has been used to predict the thermodynamic path behavior, heat dissipation and other thermodynamic parameters. However, the SME behavior is directly related to the martensitic phase transformation. Cory's model does not consider the phase transformation as the internal driving force that results in the hysteresis. The state variable, Z, also lacks a physical explanation (probably related to the martensite/austenite fraction). Cory's model describes only the one-dimensional stress-strain relation of SMA. The other phase transition characteristics, such as the recovery effects, have not been modeled.

Based on the work of Tanaka, Niezgodka and Sprekel, and others, we believe that the stress, strain, temperature, and martensite fraction provide a complete set of state variables for SMA systems. Tanaka's rate form of the constitutive equation, Equation (7), is of a relatively simple form and yet is derived from a fundamental thermodynamic point of view. This equation will be used as our basic governing equation. In order to develop a time-independent model rate form of the constitutive equation, Equation (7) is integrated with respect to time and yields what is called the unified constitutive relation as:

$$\bar{\sigma} - \bar{\sigma}_0 = D(\bar{\varepsilon} - \bar{\varepsilon}_0) + \Theta(T - T_0) + \Omega(\xi - \xi_0) \tag{16}$$

where variables with subscript "0" correspond to the initial conditions. It is assumed that the material properties are constant here.

EXPERIMENTAL PHENOMENOLOGY

Since the stress-induced martensitic phase transformation is one of the most important characteristics of SMA materials, it is necessary to start our model from the phase transformation relations and the effects of stress on the phase transformations.

As illustrated in Figure 5, the martensite fraction changes as a function of temperature under a free stress condition. The four important temperature parameters in the figure are: martensite finish temperature (M_f), martensite start temperature (M_s), austenite start temperature (A_s), and austenite finish temperature (A_f). There are two types of SMA materials: one has the typical ξ-T relation as shown in Figure 5 in which $A_s > M_s$, while the other type is characterized by $A_s < M_s$. Since most commercially available SMA materials belong to the first category, only the first type of SMA materials are considered in this paper.

Unlike Tanaka's model in which he described the ξ-T relation with an exponential function as expressed in Equation (9), we have selected a cosine function to describe the ξ-T relation. The two equations describing the martensite fraction during the $M \rightarrow A$ and $M \leftarrow A$ transformation under free stress conditions are assumed to have the following forms:

$$\xi = \frac{1}{2}\{\cos[a_A(T - A_s)] + 1\} \tag{17}$$

and for the $M \leftarrow A$ transformation

$$\xi = \frac{1}{2}\{\cos[a_M(T - M_f)] + 1\} \tag{18}$$

where the two material constants a_A and a_M are determined from:

$$a_A = \pi/(A_f - A_s)$$
$$a_M = \pi/(M_s - M_f) \tag{19}$$

If the $M \rightarrow A$ transformation starts from a state that has mixed austenite and martensite phases, denoted by (ξ_M, T_M), it is assumed that there will be no new austenite phase during the heating process until the temperature is higher than A_s. For temperatures above A_s the transformation is described by:

$$\xi = \frac{\xi_M}{2}\{\cos[a_A(T - A_s)] + 1\} \tag{20}$$

For $\xi_M = 1$, Equation (20) becomes simply Equation (17). Similarly, if $M \leftarrow A$ transformation starts from (ξ_A, T_A), there will be no new martensite phase until it is cooled to a temperature lower than M_s, and the transformation from M_s to M_f is described by:

$$\xi = \frac{1 - \xi_A}{2}\cos[a_M(T - M_f)] + \frac{1 + \xi_A}{2} \tag{21}$$

The three transition temperatures, M_f, M_s, A_s, are linearly related to the applied stress as shown in Figure 6. However, the change of A_f is more complicated. To simplify the model, it is assumed that the A_f line is straight and parallel to the others as shown in Figure 6. Two other material constants that indicate the influence of stress on the transition temperatures are given by:

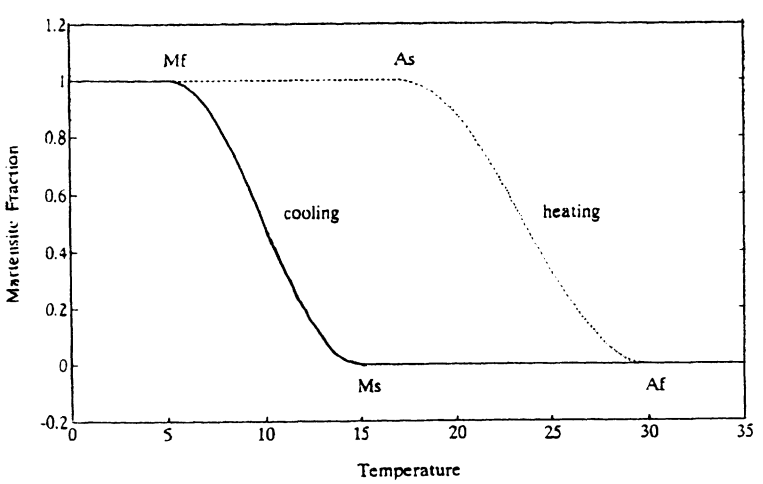

Figure 5. *Martensite fraction vs. temperature.*

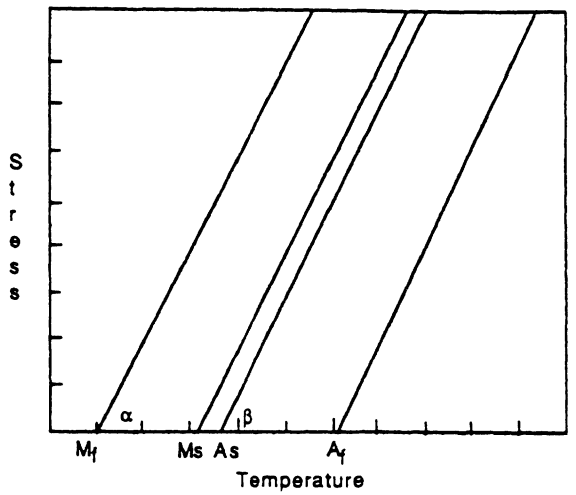

Figure 6. *Transition temperatures vs. stress.*

$$C_M = \tan(\alpha)$$
$$C_A = \tan(\beta) \tag{22}$$

where α and β are shown in Figure 6, and it is observed that $C_A = C_M$. Combining both Figures 5 and 6, it can be seen that the martensite-temperature hysteresis loop will move to the left if external stresses are applied. This movement is reflected in Equations (20) and (21) by adding a stress induced phase transition term, which gives the relation for the $M \rightarrow A$ transformation as:

$$\xi = \frac{\xi_M}{2}\{\cos[a_A(T - A_s) + b_A\sigma] + 1\} \tag{23}$$

and for the $M \leftarrow A$ transformation the expression is:

$$\xi = \frac{1 - \xi_A}{2}\cos[a_M(T - M_f) + b_M\sigma] + \frac{1 + \xi_A}{2} \tag{24}$$

where the two new material constants are:

$$b_A = -a_A/C_A$$
$$b_M = -a_M/C_M \tag{25}$$

Since the variables for the cosine function in the above equations are limited to the range 0 to π, the stress range in which the $M \rightarrow A$ transformation may occur can be derived as:

$$C_A(T - A_s) - \frac{\pi}{|b_A|} \leq \sigma \leq C_A(T - A_s) \tag{26}$$

and the corresponding stress range for the reverse transformation is given by:

$$C_M(T - M_f) - \frac{\pi}{|b_M|} \leq \sigma \leq C_M(T - M_f) \tag{27}$$

CONSTITUTIVE MODELING OF THE SHAPE MEMORY EFFECT

Shape memory alloys may be used in several geometric forms and utilize various transduction principles for active control applications. One approach is to use SMA wire or fiber as distributed force actuators. SMA force actuators are first elongated at a low temperature and then unloaded to generate some martensitic residual strain; upon heating the wire, the martensitic residual strain will be restored. The shape memory recovery generally falls into one of the following categories. First is free recovery in which there is no external load on the wire and, therefore, no work. The second category includes fully restrained "recovery". In this case, the martensitic residual strain of the SMA wire is restrained from being restored to its original length. This restraint causes large internal "recovery" stress to be generated. The third category is called controlled recovery. In this case, some residual martensitic strain is restored (the restored part is called the recovery strain), but the wire is still under some tension that is required to prevent full recovery of the residual strain. The following models may be applied to any of these three categories.

Stress and Strain Relations

The constitutive model presented in this paper considers only one-dimensional system, therefore the material will often be referred to as wire or fiber. To give the constitutive relations for the SME, we must first consider the stress-strain relations that provide the initial conditions for the recovery process.

Now, consider a piece of SMA wire undergoing an isothermal mechanical loading and unloading. It is assumed that the surrounding temperature is below A_s, so there is no pseudoelastic effect. The initial state of a piece of SMA wire can range from all martensite to all austenite, depending on the heat processing and the ambient temperature. Here, we consider only the case that has loading and unloading temperatures between M_s and A_s and 100 percent austenite before loading. To achieve a stress induced martensitic phase, the applied stress must be in the stress range given by Equation (27). In Equation (16), for the isothermal process, there is no T term, and ξ is zero if the stress is less than $C_M(T - M_s)$, as given in Equation (27). After taking the integration with time and applying the initial conditions that are zero stress, strain and full austenite, Equation (16) changes into the following form:

$$\bar{\sigma} = D\bar{\varepsilon} \tag{28}$$

The linear elastic stress limit $\bar{\sigma}_{lin}$ and its corresponding strain $\bar{\varepsilon}_{lin}$ are defined as:

$$\bar{\sigma}_{lin} = C_M(T - M_f) - \frac{\pi}{|b_M|} = C_M(T - M_s)$$
$$\bar{\varepsilon}_{lin} = \bar{\sigma}_{lin}/D \tag{29}$$

The experimental results shown in Figures 7 and 8 seem to verify the relations of Equation (29). Note that the yield stress shown in Figure 8 is 0.2 percent yield stress, which is different from the linear elastic stress limit defined here.

When the applied stress becomes greater than $\bar{\sigma}_{lin}$, there will be stress induced martensite; Equation (16) now becomes:

$$\bar{\sigma} - \bar{\sigma}_0 = D(\bar{\varepsilon} - \bar{\varepsilon}_0) + \Omega(\xi - \xi_0) \tag{30}$$

with the initial conditions, $\bar{\sigma}_0 = \bar{\sigma}_{lin}$, $\bar{\varepsilon}_0 = \bar{\varepsilon}_{lin}$, and $\xi_0 = 0$. Equation (30) then simply becomes:

$$\bar{\sigma} = D\bar{\varepsilon} + \Omega\xi \tag{31}$$

Together Equations (29), (31) and (24) yield the stress-strain curve. To determine the value of Ω from the mechanical engineering point of view, which is concerned with the crystallographic volume change due to transformation and is a well-defined physical quantity in metallurgy, it is necessary to consider the unloading process at $\xi = 1$. Since the ambient temperature is below A_s, there is no pseudoelasticity—it is a linear martensitic elastic unloading. If the loading continues after the martensitic phase transformation finishes (where $\xi = 1$), it is a linear elastic loading defined by the martensitic elastic modulus and it continues until the stress reaches the second yield strength where permanent martensitic slip (plastic strain) occurs, as shown in Figure 9. The constitutive equations given here can describe the stress-strain relations up to the second yield point.

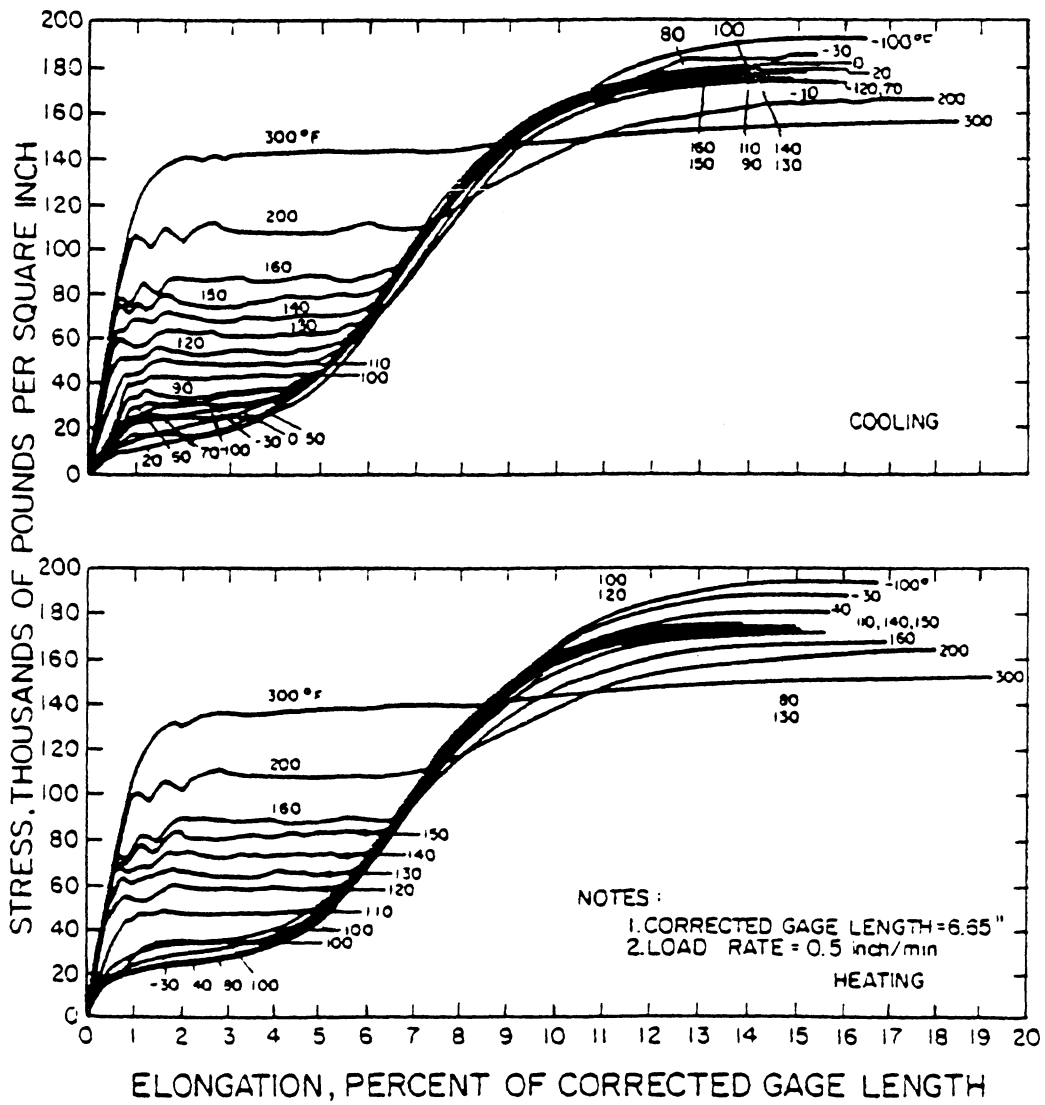

Figure 7. Stress vs. elongation (M_s = 15°C) [3].

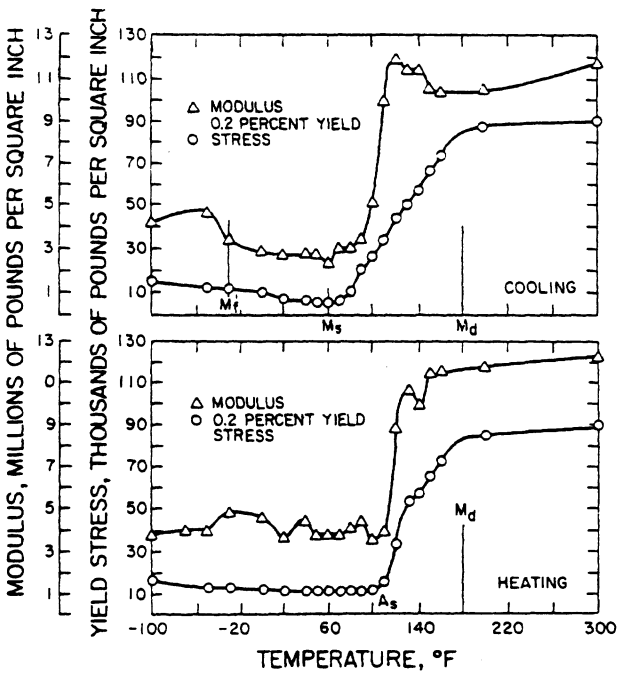

Figure 8. *The effect of temperature on the yield stress and elastic modulus [3].*

After unloading at $\xi = 1$, the martensitic residual strain (the unloaded phase is completely deformed martensite), $\bar{\varepsilon}_{res}$, can be derived as:

$$\bar{\varepsilon}_{res} = -\frac{\Omega}{D} \tag{32}$$

This residual strain, which corresponds to the completely deformed martensitic phase, is the maximum recoverable strain. Experimental results have shown that this maximum recoverable martensitic strain, or recovery strain limit, is almost a constant between M_f and A_f^m, at which point it is impossible to induce martensite by stress. The recovery strain limit, $\bar{\varepsilon}_L$, is therefore considered a temperature-independent material constant; it has the following relation with Ω and D:

$$\bar{\varepsilon}_L = -\frac{\Omega}{D} \tag{33}$$

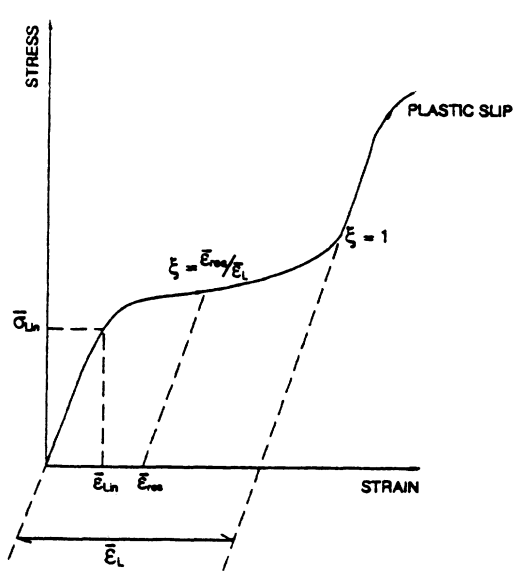

Figure 9. *Typical stress-strain relationship of SMA.*

If the unloading occurs before $\xi = 1$, the SMA material is composed of both a completely deformed martensitic phase and austenite phase. It is assumed here that the austenitic crystallic deformation contributes only to the elastic deformation. Therefore, the martensitic residual strain can be derived as:

$$\bar{\varepsilon}_{res} = -\xi\frac{\Omega}{D} \tag{34}$$

The martensite fraction, ξ, which is the initial condition for the recovery process, can be expressed as:

$$\xi = -\bar{\varepsilon}_{res}\frac{D}{\Omega} \quad \text{or} \quad \xi = \frac{\bar{\varepsilon}_{res}}{\bar{\varepsilon}_L} \tag{35}$$

If the surrounding temperature is below M_s, there is some original martensite. (To distinguish the existing undeformed martensite from the stress induced martensite, the martensite under the free stress state is called the original martensite.) The discussion of original and deformed martensite is more complex. A future paper will be published that will be devoted particularly to this subject.

Constitutive Relation of the SME

FREE RECOVERY

For the free shape recovery, the martensitic residual strain, which is the primary parameter in the determination of recovery strain-temperature relations, will appear only in the constitutive relations as an initial condition. In spite of what happened in the loading and unloading process, the strain/shape recovery depends only on the final state of the material, namely, the deformed martensite fraction. For the free recovery, there is no stress. Equation (16) then becomes:

$$D(\bar{\varepsilon}^r - \bar{\varepsilon}_0) + \Theta(T - T_0) + \Omega(\xi - \xi_0) = 0 \tag{36}$$

where the initial conditions are:

$$\begin{aligned} \bar{\varepsilon}_0 &= \bar{\varepsilon}_{res} \\ T_0 &= A_s \\ \xi_0 &= \xi_M = \frac{\bar{\varepsilon}_{res}}{\bar{\varepsilon}_L} \end{aligned} \tag{37}$$

This equation ignores the thermal expansion effects that are negligible in comparison to the recovery effects. Since free recovery must occur above A_s, the austenite start temperature is set as the initial condition for the temperature. In the above equation, the martensite fraction is governed by Equation (23) in which σ is zero. Note that the recovery strain is still measured in reference to the original coordinate X. Equations (23), (36), and the above initial conditions give the constitutive equation for the free recovery as:

$$\bar{\varepsilon}^r = \bar{\varepsilon}_{res} - \frac{1}{D}\left\{\Theta(T - A_s) + \frac{\Omega}{2} \times \frac{\bar{\varepsilon}_{res}}{\bar{\varepsilon}_L}[\cos a_A(T - A_s) - 1]\right\} \tag{38}$$

The material properties D, Θ, and $\bar{\varepsilon}_L$ can be determined experimentally. The transformation tensor Ω is then calculated using Equation (33). The following discussion about the restrained recovery will give the method to determine the thermoelastic tensor, Θ, experimentally.

RESTRAINED RECOVERY

For the restrained recovery, there is no change in strain. As observed in experiments, the recovery stress, $\bar{\sigma}^r$, is only a function of temperature and martensitic residual strain. The constitutive equation from Equation (16) then becomes:

$$\bar{\sigma}^r - \bar{\sigma}_0^r = \Theta(T - T_0) + \Omega(\xi - \xi_0) \tag{39}$$

The initial state in Figure 5 is (ξ_M, T_M), where ξ_M is ξ_0, and T_M can be found from Equation (18) and rewritten as:

$$T_M = M_f + \frac{\cos^{-1}[2(\xi_M - 0.5)]}{a_M} \tag{40}$$

From T_M to A_s, there is no new austenite; the recovery stress-temperature relation is linear and given by:

$$\bar{\sigma}^r - \bar{\sigma}_0^r = \Theta(T - T_M) \tag{41}$$

The initial stress, $\bar{\sigma}_0^r$, is assumed to be zero at T_M. If there is some stress, the austenite start temperature will be increased as shown in Figure 6. By defining the new austenite start temperature as A_s^m, and combining Equations (41) and (23) at the transition point where $a_A(A_s^m - A_s) + b_A\bar{\sigma}^r$ equals 0, one obtains the new transition temperature as:

$$A_s^m = \frac{C_A A_s - \Theta T_M}{C_A - \Theta} \tag{42}$$

By substituting this temperature into Equation (41) one obtains the corresponding stress $\bar{\sigma}_{As}^r$ as:

$$\bar{\sigma}_{As}^r = \Theta(A_s^m - T_m) \tag{43}$$

If the ambient temperature is higher than A_s^m, then there is new austenite and the constitutive relation has the complete form of Equation (39). The initial conditions are:

$$\bar{\sigma}_0^r = \bar{\sigma}_{As}^r \qquad T_0 = A_s^m \qquad \xi_0 = \xi_M = \bar{\varepsilon}_{res}/\bar{\varepsilon}_L \tag{44}$$

The calculation obtained by replacing $\bar{\sigma}$ with $\bar{\sigma}^r$ in Equation (23), and substituting it and the above initial conditions into Equation (39), yields the final constitutive equation for the restrained recovery as:

$$\bar{\sigma}^r - \bar{\sigma}_{As}^r = \Theta(T - A_s^m) + \frac{\Omega}{2}\frac{\bar{\varepsilon}_{res}}{\bar{\varepsilon}_L}\left\{\cos[a_A(T - A_s) + b_A\bar{\sigma}^r] - 1\right\} \tag{45}$$

The above stress-temperature relation is not explicit; for a given temperature, iteration is needed for convergence.

When ambient temperature reaches a certain level, there will be no stress induced martensite for the $M \rightarrow A$ transformation. This temperature is denoted as A_f^m. In this case, the recovery stress-temperature relation will be linear:

$$\bar{\sigma}^r - \bar{\sigma}_0^r = \Theta(T - T_0) \tag{46}$$

where T_0 is A_f^m and $\bar{\sigma}_0^r$ is the recovery stress at $T = A_f^m$, resulting in

$$\bar{\sigma}^r = \bar{\sigma}_{Af}^r = \Theta(A_f^m - A_s^m) + \Omega\left(0 - \frac{\bar{\varepsilon}_{res}}{\bar{\varepsilon}_L}\right) + \bar{\sigma}_{As}^r \tag{47}$$

A_f^m can be solved by substituting Equation (47) into Equation (23) as:

$$A_f^m = \frac{a_A A_s - b_A\bar{\sigma}_{As}^r + b_A\Omega\xi_M + b_A\Theta A_s^m + \pi}{a_A + b_A\Theta} \tag{48}$$

Equation (46) indicates that the recovery stress has a linear relation with temperature after T higher than A_f^m. This point is observed in the experiment shown in Figure 10. This relation indicates an approach to determine the thermoelastic tensor, Θ, experimentally. Figure 10 shows the restrained recovery stress vs. temperature for different initial strains. It can be seen that the recovery stress is almost a linear relation

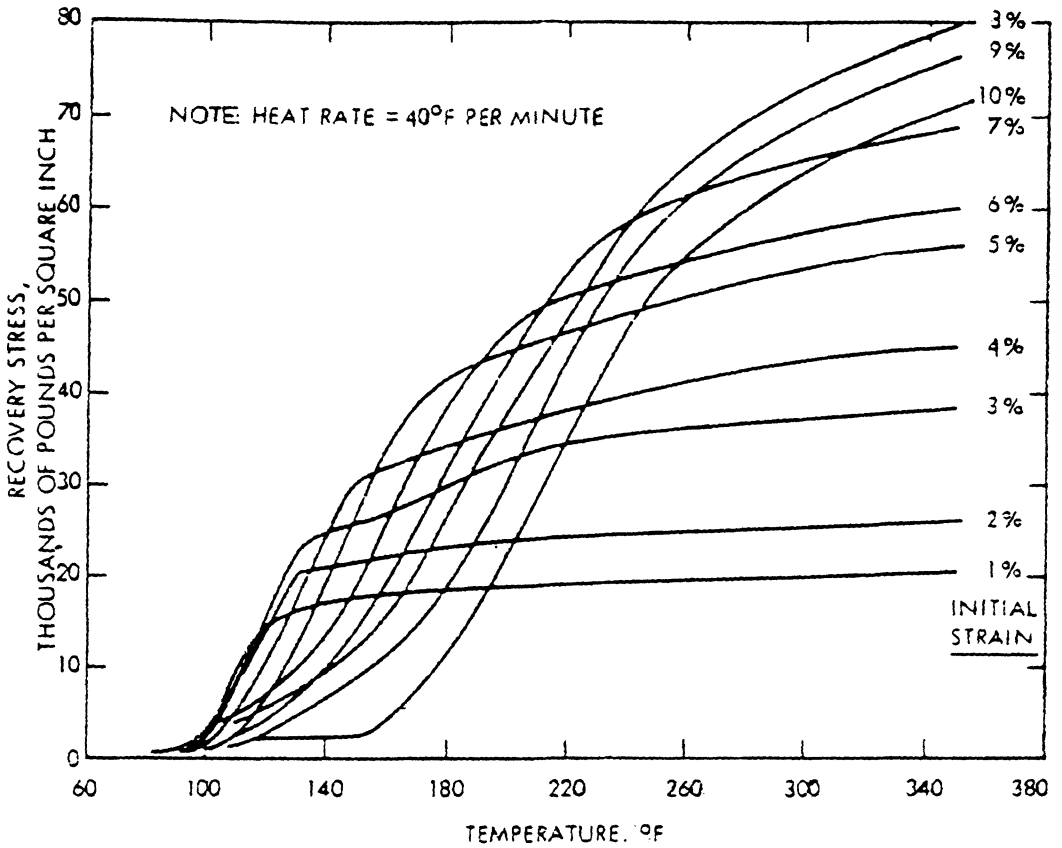

Figure 10. *Recovery stress vs. temperature for different initial strains [3].*

with temperature at higher temperatures after the phase transformation has finished. The tangent of that slope is defined as the thermoelastic tensor, Θ. Compared with other material constants such as D and Ω, Θ is much smaller. We define the recovery stress, $\bar{\sigma}^r_{Af}$, at temperature, A^m_f, as the maximum recovery stress. It can be determined from the expression that the maximum recovery stress is a linear function of the martensitic residual strain or initial strain. Figure 11 shows that the maximum recovery stress vs. martensitic residual strain is, in fact, linear. Note that all the discussion presented in this paper is based on the condition that $\bar{\varepsilon}_{res}$ is less than $\bar{\varepsilon}_L$.

The recovery stress-temperature relations are different for cooling and heating. The preceding discussion was devoted to the heating process. Experiments reveal hysteric behavior for heating and cooling cycles as shown in Figure 12. However, using the above approach, the mathematical simulation will in fact describe and predict the complete recovery stress-temperature hysteresis.

Assume cooling begins when the restrained SMA wire is heated to some temperature. Before the cooling begins, the recovery stress-temperature relations are given by Equations (41), (45) and (46). After cooling starts, which is to say that an $A \rightarrow M$ process is initiated, Equation (24) must be used with Equation (39) to describe the process. In this case, the initial conditions are:

$$\begin{aligned} \bar{\sigma}^r_0 &= \bar{\sigma}^r_c \\ T_0 &= T_c \\ \xi_0 &= \xi_c \end{aligned} \tag{49}$$

where the subscript c indicates the initial cooling state. Considering the effect of stress on M_s, the reverse phase transformation will occur at temperature, M^m_s, which is higher than M_s. Before the temperature reaches this transition point, there is no new martensite. As in solving A^m_s for the heating process, this transition temperature, M^m_s, and its corresponding stress, $\bar{\sigma}^r_{Ms}$, can be determined from:

$$M^m_s = \frac{a_M M_f - b_M \bar{\sigma}^r_c + b_M \Theta T_c + \pi}{a_M + b_M \Theta} \tag{50}$$

and

$$\bar{\sigma}^r_{Ms} = \bar{\sigma}^r_c + \Theta(M^m_s - T_c) \tag{51}$$

The phase transformation occurs for temperatures below M^m_s. Equations (24) and (39) yield the governing equation:

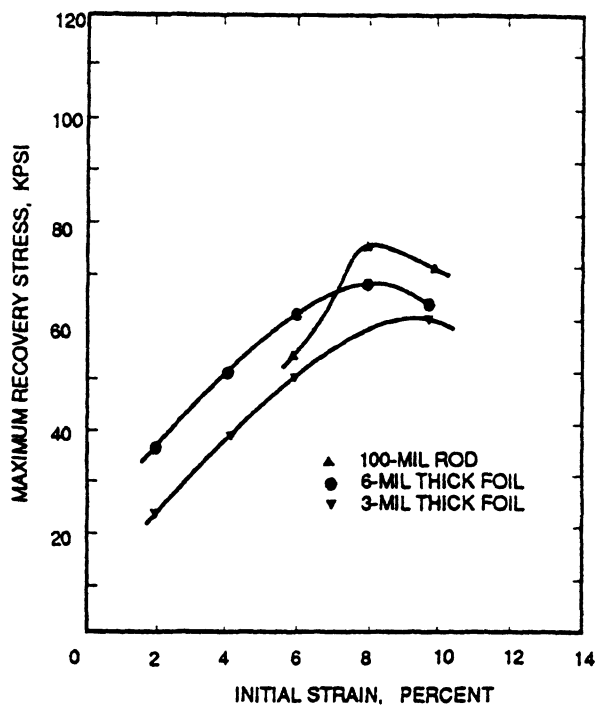

Figure 11. Max. recovery stress vs. initial strain ($\bar{\varepsilon}_L$ = 8%) [3].

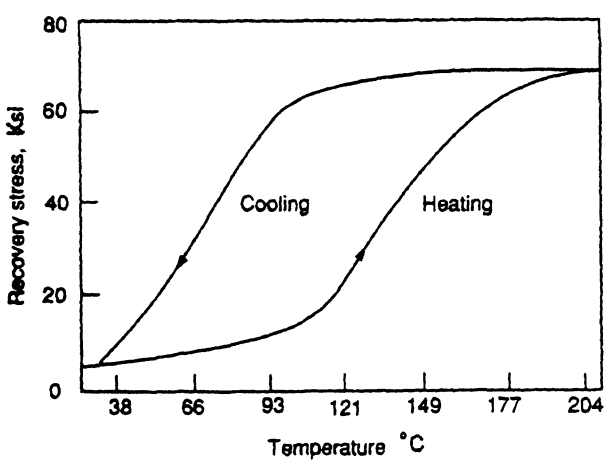

Figure 12. Hysteresis of SMA.

$$\bar{\sigma}^r - \bar{\sigma}^r_{Ms} = \Theta(T - M_s^m) + \Omega\frac{1-\xi_c}{2}\{\cos[a_M(T - M_f) + b_M\bar{\sigma}^r] + 1\} \tag{52}$$

Similarly, if the SMA wire is cooled to M_f^m, which is higher than M_f because of the stress effect, the $M \leftarrow A$ transformation finishes, and the recovery stress and temperature is again a linear relation. M_f^m and its corresponding stress, $\bar{\sigma}^r_{Mf}$, are calculated from:

$$\bar{\sigma}^r_{Mf} = \Theta(M_f^m - M_f) + \Omega(1-\xi_c) + \bar{\sigma}^r_{Ms} \tag{53}$$

and

$$M_f^m = \frac{a_M M_f - b_M\bar{\sigma}^r_{Ms} - b_M\Omega(1-\xi_c) + b_M\Theta M_s^m}{a_M + b_M\Theta} \tag{54}$$

The linear recovery stress-temperature relation after the temperature drops below M_f^m is given by:

$$\bar{\sigma}^r - \bar{\sigma}^r_{Mf} = \Theta(T - M_f^m) + \Omega(1-\xi_c) \tag{55}$$

In this section, some new temperature parameters, namely, A_s^m, A_f^m, M_s^m, and M_f^m are introduced. Like the phase transition temperatures of the SMA materials, A_s, M_s, etc., the parameters are the start and finish temperatures of the phase transformation under stress. Therefore, in this paper they are referred to as the mechanical transition temperatures. The relations to calculate their corresponding stresses, $\bar{\sigma}^r_{As}$, $\bar{\sigma}^r_{Af}$, $\bar{\sigma}^r_{Ms}$, and $\bar{\sigma}^r_{Mf}$ have been derived and explained.

CONTROLLED RECOVERY

Controlled recovery is defined as the shape memory recovery process during which some recovery strain is allowed during the heating and cooling cycle. Since the martensite fraction-temperature relation has been described in a piecewise continuous fashion, the recovery stress-strain temperature should also be studied in a similar manner. For the controlled recovery, Equation (16) can be expressed as:

$$\bar{\sigma}^r - \bar{\sigma}^r_0 = D(\bar{\varepsilon}^r - \bar{\varepsilon}^r_0) + \Theta(T - T_0) + \Omega(\xi - \xi_0) \tag{56}$$

The controlled recovery process can be subdivided into two categories defined by their loading mechanism. The first category assumes that a constant load is applied, e.g., a piece of elongated SMA wire is loaded by a dead weight. In this case, Equation (56) is expressed as:

$$D(\bar{\varepsilon}^r - \bar{\varepsilon}^r_0) + \Theta(T - T_0) + \Omega(\xi - \xi_0) = 0 \tag{57}$$

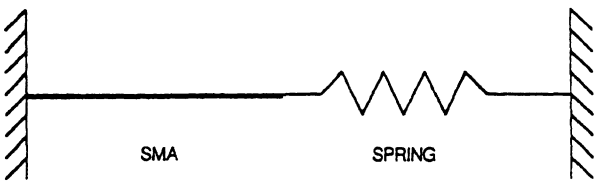

Figure 13. Simplified SMA force actuator.

where the stress caused by the dead weight appears in the martensite fraction term. Transition points can be determined following the reasoning described in the restrained recovery section.

The second loading category assumes that the loading is proportional to the recovery strain and may be modeled as SMA-spring structures. Most SMA force actuators [2] can also be simplified to this model. Consider a spring-SMA system as shown in Figure 13. If the spring constant is k, the length of the SMA wire L, and the cross section area of the SMA wire s, then the stress-strain relation for the spring is described by:

$$\bar{\sigma}^r - \bar{\sigma}_0^r = \frac{kL}{s}(\bar{\varepsilon}_{res} - \bar{\varepsilon}^r) \tag{58}$$

Substituting this equation into Equation (56) yields:

$$\left(1 + s\frac{D}{kL}\right)(\bar{\sigma}^r - \bar{\sigma}_0^r) = \Theta(T - A_0) + \Omega(\xi - \xi_0) \tag{59}$$

This equation can be handled in the same way as the restrained recovery. It should be noted that the stress-temperature relation as given in Equation (59) is path dependent, and a nonlinear spring will yield very different state functions and results. In the above equation, ξ is also governed by Equation (23) or (24) according to whether the SMA wire is being heated or cooled.

Table 1. Material constants for a copper based SMA.

(MPa)		(°C)				(MPa/°C)	
D	Ω	M_s	M_f	A_s	A_f	C_A	C_M
7000	−70	−27	−34	−25	−14	1.5	1.5

NUMERICAL EXAMPLE

Based on data given in Reference [17], some numerical examples will be given here for a copper based alloy. At this time, because of the lack of experimental results, we cannot verify our prediction of stress-strain and recovery stress-strain temperature relations with experimental data. However, a detailed experimental program to verify this model is currently in progress and will be published in the near future. The experimental results that are available do, in fact, show the same relationships and trends as those predicted with the model described in this paper.

Table 1 gives the material constants that are required for our numerical simulations. From the table, the maximum recoverable strain, or recovery strain limit, ε_L, is found as one percent from Equation (33).

Figure 14 shows the stress-strain relations predicted for two temperatures. Great similarities are found when this prediction is compared with the experimental results shown in Figure 7. Of course, in order to make the simulation fit the experimental results completely, more work is necessary. Figure 15 shows the martensite fraction as a function of stress. It is clearly indicated that for stress less than 10 MPa, the martensite fraction is described by aus-

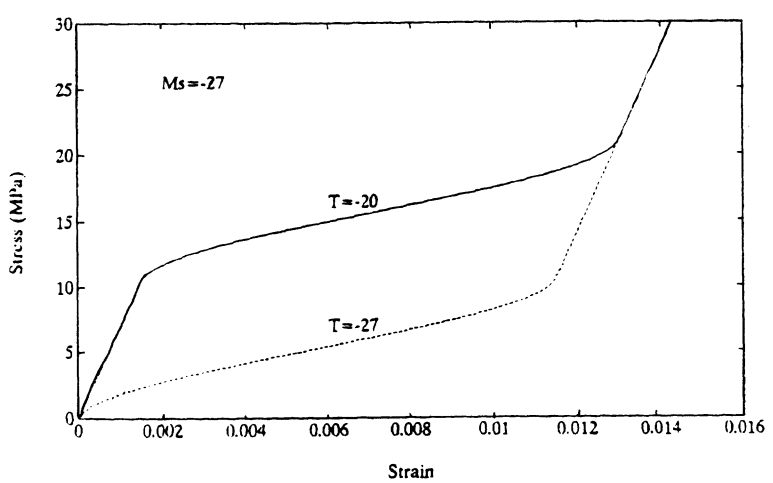

Figure 14. Prediction of stress-strain relation for a copper based SMA.

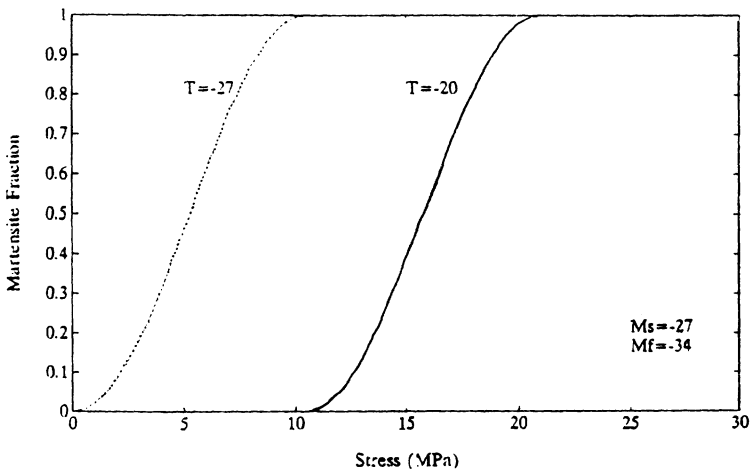

Figure 15. *Martensite fraction vs. stress A → M.*

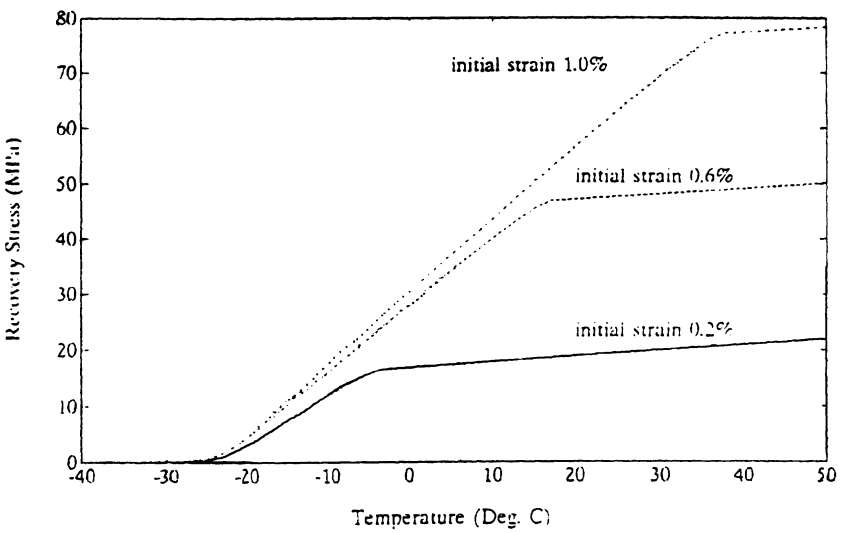

Figure 16. *Recovery stress vs. temperature for restrained SMA wire.*

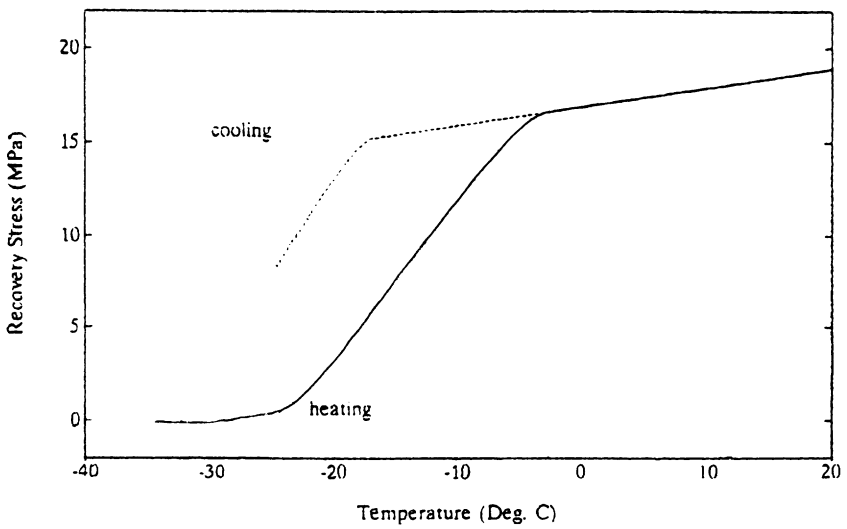

Figure 17. *Predicted hysteresis of a copper based SMA with 0.2% initial strain.*

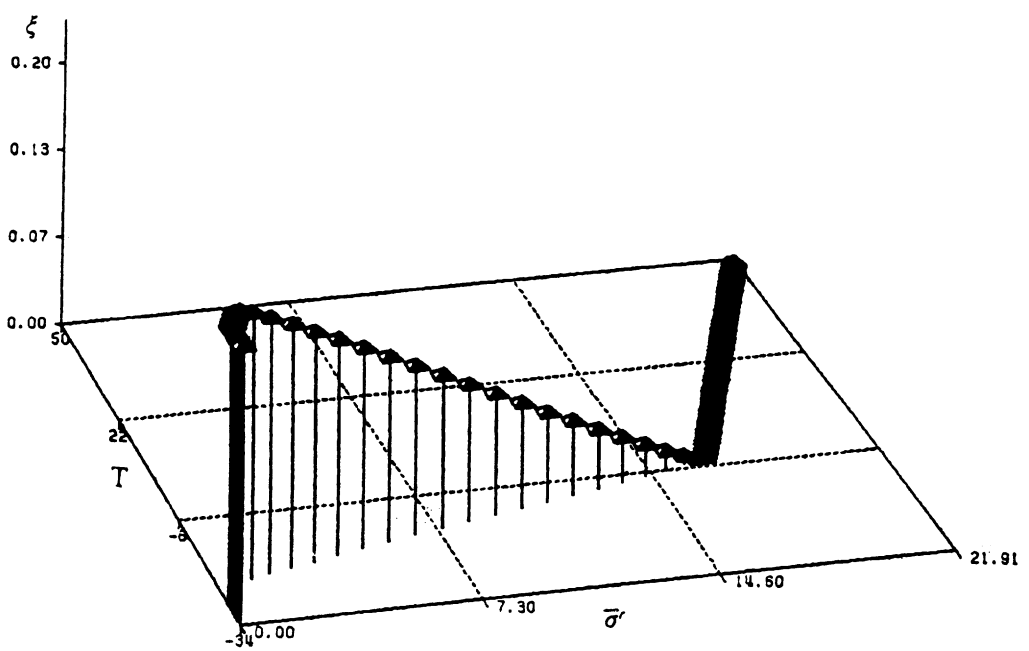

Figure 18. *ξ vs. $\bar{\sigma}^r$ and T for a restrained SMA wire with 0.2% initial strain.*

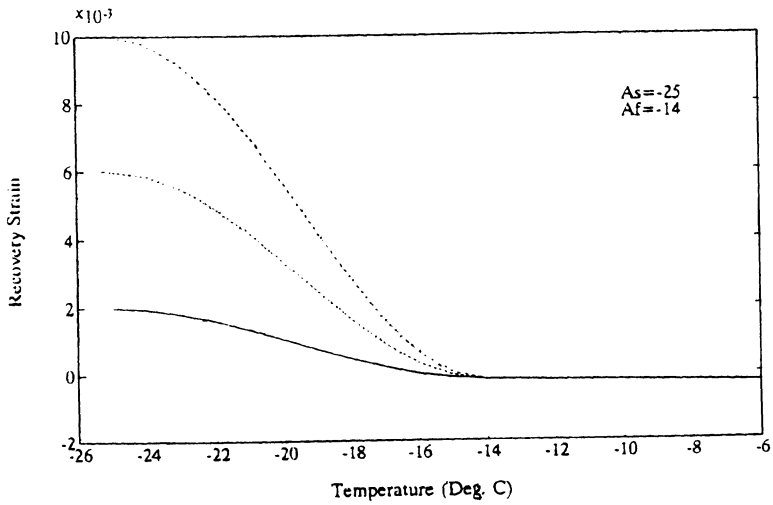

Figure 19. *Recovery strain vs. temperature of free recovery.*

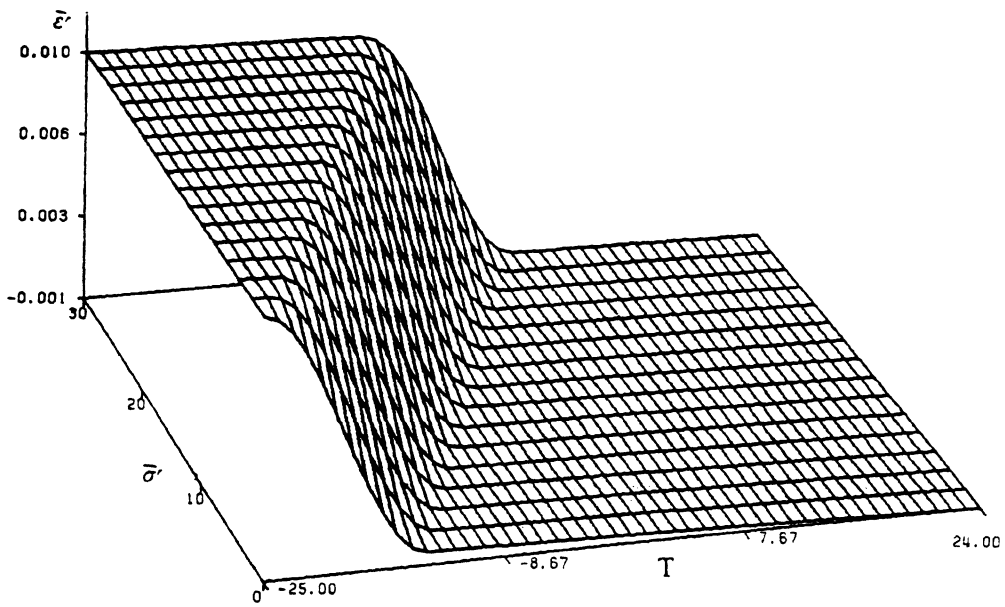

Figure 20. $\bar{\varepsilon}^r$ *vs.* T *and* $\bar{\sigma}^r$ *for constant load SMA wire structures.*

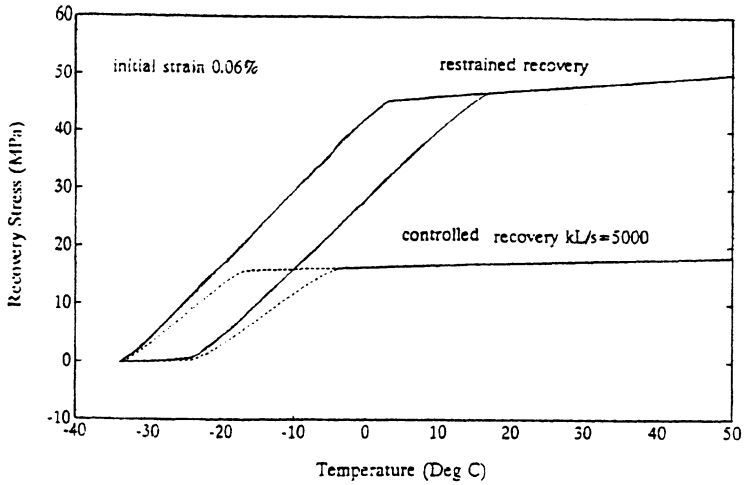

Figure 21. *Comparison of $\bar{\sigma}^r$-T hysteresis between restrained and controlled recovery.*

tenitic elasticity. Martensitic elasticity appears to be dominant for stresses greater than 22 MPa for the testing temperature of –20°C.

The recovery stress vs. temperature of a restrained wire for three different initial strains is shown in Figure 16. Here, Θ is assumed to be 0.1 MPa/°C. This result is very similar to the experimental results shown in Figure 10. Figure 17 shows the calculated SMA hysteresis for the heating and cooling process. Again this demonstrates that our simulation does fit the experimental phenomena. As illustrated in Figure 18, the martensite fraction changes as a function of stress and temperature for a restrained wire with 0.2% initial strain, which corresponds to 20% of initial deformed martensite [see Equation (35)].

Figure 19 illustrates the recovery strain vs. temperature for the free recovery using three different initial strains. Since there is no stress to affect the phase transformation, the free recovery processes start at A_s and end at A_f.

Figure 20 shows the recovery strain vs. temperature for different levels of applied stress. This is a simulation of the constant loading scenario. Once again, the model correlates well with the experimental phenomenological observations. For SMA-spring structures, the recovery stress-temperature hysteresis is shown in Figure 21.

CONCLUSIONS

This paper describes a unified, thermomechanical constitutive model based on Tanaka's early work. A more accurate equation is proposed to fit the martensite fraction and temperature relations in order to quantitatively predict the stress-strain relation and shape memory behavior. This improved model more accurately predicts the experimental observations using the most basic material constants that are required to completely describe the behavior of shape memory alloys. Besides the stress-strain relation, this paper focuses more on the SME and will eventually provide a theoretical guide to the design of SMA based intelligent materials and structures. The theory has demonstrated the ability to model the observed experimental phenomena; the numerical examples presented correlated very well with the experimental observations.

Further work will concentrate on incorporating experimental observations into the theory. Temperature dependent material constants and even the martensite fraction dependent material properties will be considered in order to develop better prediction and analysis capabilities as well as to develop the multi-dimensional constitutive relations for shape memory alloys.

NOMENCLATURE

A	austenite phase
A_f, A_s	austenite finish and start temperatures
A_f^m, A_s^m	mechanical austenite finish and start temperatures
a, b, f, g, h	material constants (Cory's model)
a_A, a_M	material constants associated with the temperature induced phase transformation
b_A, b_M	material constants associated with the stress induced phase transformation
C_A, C_M	material constants related to the influence of stress on the phase transformation
D	elastic modulus
E	energy
F, F_0, F_x	force
k	spring constant
L	length of SMA wire or spring as specified
L_0, L_x	specific lengths
M	martensite phase

M_f, M_s martensite finish and start temperatures
M_f^m, M_s^m mechanical martensite finish and start temperatures
N_{M-}, N_{M+}, N_A distribution function
Q heat
q, q_{sur} internal heat source and heat flux from the surroundings
S entropy density
s cross section area of SMA wire or fiber
T temperature
T_M, T_A initial state temperatures of SMA
U internal energy density
X reference coordinate
Z, Z_x, Z_0 new state variables in Cory's model
Δ shear length of each layer of martensitic lattice
ε engineering strain
$\bar{\varepsilon}$ Green strain
$\bar{\varepsilon}^r$ recovery Green strain
$\bar{\varepsilon}_L$ max. recoverable strain or recovery strain limit
$\bar{\varepsilon}_{lin}$ linear elastic strain limit
$\bar{\varepsilon}_{res}$ martensitic residual strain
Θ thermoelastic tensor
Λ generalized state variable
ξ martensite fraction
ξ_A, ξ_M initial martensite fractions of SMA
ρ mass density in the deformed configuration
ρ_0 mass density in the reference configuration
σ engineering stress
$\bar{\sigma}$ Piola-Kirchhoff stress
$\bar{\sigma}_{lin}$ linear elastic stress limit
$\bar{\sigma}^r$ recovery stress
$\bar{\sigma}^r_{As}$ stress corresponding to A_s^m
$\bar{\sigma}^r_{Af}$ stress corresponding to A_f^m
$\bar{\sigma}^r_{Mf}$ stress corresponding to M_f^m
$\bar{\sigma}^r_{Ms}$ stress corresponding to M_s^m
Φ potential energy of SMA lattice
Ω transformation tensor

Superscripts

r recovery
m mechanics

Subscripts

A austenite
Af austenite finish
As austenite start
c cooling
f finish
L limit
lin linear
M martensite
Mf martensite finish
Ms martensite start
res residual
s start
sur surroundings
0 initial condition

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the Office of Naval Research Young Investigator Program. ONR N00014-88-K-0566.

REFERENCES

1. Perkins, J. 1976. *Shape Memory Effects in Alloys.* N.Y., London: Plenum Press.
2. Funakubo, Hiroyasu (translated by J. B. Kennedy). 1987. *Shape Memory Alloys.* Gordon and Breach Science Publishers.
3. Jackson, C. M., H. J. Wagner and R. J. Wasilewski. 1972. "55-Nitinol—The Alloy with a Memory: Its Physical Metallurgy, Properties, and Applications", NASA-SP-5110, p. 91.
4. Wayman, C. M. 1980. "Some Applications of Shape-Memory Alloys", *Journal of Metals* (June).
5. Rogers, C. A. and H. H. Robertshaw. 1988. "Shape Memory Alloy Reinforced Composites", *Eng. Sci. Preprints,* ESP25.88027 (June 20–22).
6. Warlimont, H., L. K. Deleay, R. V. Krishnan and H. Tas. 1974. "Review: Thermoelasticity, Pseudoelasticity and the Memory Effects Associated with Martensitic Transformations, Part 3: Thermodynamics of Kinetics", *Journal of Material Science,* 9:1545–1555.
7. Müller, I. 1979. "A Model for a Body with Shape Memory", *Arch. Rat. Mech. Anal.,* 70:61–77.
8. Müller, I. and K. Wilmanski. 1980. "A Model for Phase Transformation in Pseudoelastic Bodies", *Il Nuovo Cimento,* 57B:238–318.
9. Müller, I. 1985. "Pseudoelasticity in Shape Memory Alloys—An Extreme Case of Thermoelasticity", IMA Preprint No. 169 (July).
10. Achenbach, M., T. Atanackovic and I. Müller. 1986. "A Model for Memory Alloys in Plane Strain", *Int. J. of Solid Structures,* 22(2): 171–193.
11. Müller, I. "Shape Memory Alloys—Phenomenology and Simulation", *Shape Memory Alloys,* preprints.
12. Tanaka, K. and S. Nagaki. 1982. "A Thermomechanical Description of Materials with Internal Variables in the Process of Phase Transformation", *Ingenieur-Archiv,* 51:287–299.
13. Tanaka, K. and R. Iwasaki. 1985. "A Phenomenological Theory of Transformation Superplasticity", *Engineering Fracture Mechanics,* 21(4):709–720.
14. Tanaka, K. 1986. "A Thermomechanical Sketch of Shape Memory Effect: One-Dimensional Tensile Behavior", *Res. Mechanica,* 18:251–263.
15. Sato, Y., K. Tanaka and S. Kobayshi. 1986. "Pseudoelasticity and Shape Memory Effect Associated with Stress-Induced Martensitic Transformation: A Thermomechanical Approach", *Trans. of Japan Soc. of Aero. & Space Science,* Vol. 28.
16. Sato, Y. and K. Tanaka. 1988. "Estimation of Energy Dissipation in Alloys Due to Stress-Induced Martensitic Transformation", *Res. Mechanica,* 23:381–393, England: Elsevier Applied Science Publishers Ltd.
17. Niezgodka, M. and J. Sprekels. 1986. "On the Dynamics of Structural Phase Transformation in Shape Memory Alloys", Preprint No. 114, Inst. fur Mathe, Uni. Augsburg.

18. Hoffmann, K. H. and J. Sprekels. 1987. "Phase Transformation in Shape Memory Alloys I: Stability and Optimal Control", Preprint No. 136, Inst. fur Mathe, Uni. Augsburg.
19. Niezgodka, M. and J. Sprekels. 1985. "Existence of Solutions for a Mathematical Model of Structural Phase Transitions in Shape Memory Alloys", Preprint No. 89, Inst. fur Mathe, Uni. Augsburg.
20. Niezgodka, M., S. M. Zheng and J. Sprekels. 1986. "Global Solutions to Structural Phase Transitions in Shape Memory Alloys", Print No. 105, Inst. fur Mathe, Uni. Augsburg.
21. Hoffmann, K. H. and S. M. Zheng. 1986. "Uniqueness for Nonlinear Coupled Equations Arising from Alloy Mechanism", Preprint No. 118, Inst. fur Mathe, Uni. Augsburg.
22. Cory, J. S. 1978. "Nitinol Thermodynamic State Surfaces", *Journal of Energy*, 2(5).
23. Cory, J. S. and J. L. McNichols, Jr. 1985. "Nonequilibrium Thermostatics", *Journal of Applied Physics*, 58(9):1.
24. McNichols, J. L., Jr. and J. S. Cory. 1987. "Thermodynamics of Nitinol", *Journal of Applied Physics*, 61(3):1.