Differential Hysteresis Modeling of a Shape Memory Alloy Wire Actuator

Sushant M. Dutta, *Student Member, IEEE* and Fathi H. Ghorbel, *Senior Member, IEEE*

***Abstract*— In this paper, we develop a complete mathematical model of a shape memory alloy (SMA) wire actuated by an electric current and a bias spring. The operation of the SMA actuator involves different physical phenomena, such as heat transfer, phase transformation with temperature hysteresis, stress-strain variations and electrical resistance variation accompanying the phase transformation. We model each of these phenomena in a modular fashion. A key feature of the proposed model is that one or more of its modules can be extended to fit other SMA applications. At the heart of the proposed model is a differential hysteresis model capable of representing minor hysteresis loops. We generate the temperature profile for the hysteresis model using lumped parameter analysis. We extend the variable sublayer model to represent actuator strain and electrical resistance. This model can be used to develop a position control system for the actuator. Simulation results from the model are found to be in good agreement with experimental data.**

***Index Terms*— Hysteresis, modeling, shape memory alloy.**

I. Introduction

SHAPE memory alloys (SMAs) are metallic alloys which exhibit the shape memory effect, whereby, a shape memory alloy deformed at a low temperature will regain its original undeformed state when heated to a higher temperature. This property is a consequence of the reversible crystalline phase transformation that occurs between the low temperature martensite (M) phase and the high temperature austenite (A) phase of the SMA. The two phases have the same chemical composition but different crystallographic structures, and hence, different thermal, mechanical and electrical properties. The austenitic phase is highly symmetric and usually cubic, while the martensitic phase is non-symmetric and usually monoclinic [1]. In the absence of applied stress, the M phase is formed without any observable macroscopic shape change. However, a large macroscopic inelastic strain can be obtained by applying mechanical stress on the SMA, which forces the M phase variants to *detwin*, i.e., reorient into a single variant. When the SMA is heated to a higher temperature, the M phase returns to the highly symmetric A phase and the inelastic strain is, therefore, recovered and a large force is associated with this strain recovery. The shape memory effect discussed above is called *One Way Memory Effect* (OWME). Some SMAs exhibit a *Two Way Memory Effect* (TWME), whereby the A phase transforms directly to detwinned M phase during cooling without the application of mechanical stress. SMAs can be made to acquire TWME by a cyclic loading procedure called *training* [2]. In this paper, we restrict our discussion to OWME only. Apart from M and A phases, the occurrence of another phase, called the *R-phase*, has also been observed in SMAs [3], [4]. However, the SMA wire used in this research is fully annealed near-equiatomic TiNi alloy, and therefore, does not exhibit R-phase [4]. SMAs possess electrical resistance, and it is convenient to heat them using the Joule effect, whereby, an electric current passing through the SMA increases its temperature. Cooling is usually accomplished by natural air convection, and is, therefore, slower. Where possible, the SMA may be cooled faster using water or another fluid.

The combination of their high stiffness, high strength and large recovery strain offers great potential for the use of SMAs as actuators in diverse applications. These include aircraft wing shape control [5], mobile robot actuators [6], microrobot manipulation [7], active endoscopy [8], prosthetic end-effector actuator [9], micro rotary actuator [10], micro fiber switch [11], electrical and mechanical connectors [12], smart structures and composites [13].

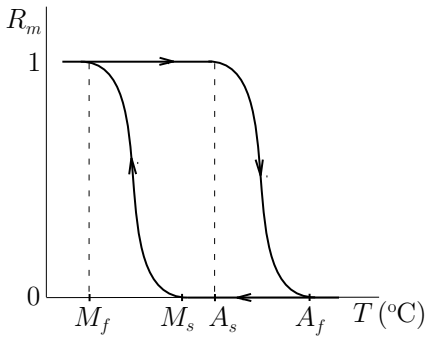

Fig. 1. Schematic of martensite fraction–temperature hysteresis.

The extent of $A - M$ transformation is characterized by the martensite fraction R_m. Martensite fraction is defined as the volume fraction of martensite present in the SMA at any instant. Therefore, $0 \leq R_m \leq 1$. One can also define the austenite fraction $R_a = 1 - R_m$ (where $R_m, R_a \in [0, 1]$) as the volume fraction of austenite present in the SMA. In general, at any given temperature, the phases A and M coexist. In principle, the thermal, mechanical and electrical properties of SMAs can be predicted if R_m and the stress history of the SMA are known. However, it is not possible to measure R_m by non-destructive means [14]. Differential Scanning Calorimetry (DSC) is a well-known method for analyzing phase transformations, and it has been shown that the $A - M$ phase transformation exhibits significant temperature hysteresis [3]. This hysteresis is mainly caused by frictional effects associated with the movement of $A-M$ interfaces and $M-M$ interfaces with different crystallographic orientations [14]. A typical

The authors are with the Department of Mechanical Engineering and Materials Science, Rice University, Houston, TX 77005 USA. E-mail: ghorbel@rice.edu .

martensite fraction–temperature hysteresis schematic is shown in Fig. 1. The phase transformation is characterized by the transformation start and finish temperatures. A_s is the austenite start temperature, A_f is the austenite finish temperature, M_s is the martensite start temperature, and M_f is the martensite finish temperature.

Most of the SMA actuator applications mentioned earlier in this section use SMA wires, because they are easy to cut, to connect, and to activate electrically. Since one way memory effect is exhibited in most cases, it is common to obtain inelastic strain in the martensite phase by virtue of a bias spring connected in series with the SMA wire. Thus the wire contracts when heated, and it expands with the aid of the bias spring when cooled. A schematic of a spring-biased SMA wire actuator is shown in Fig. 2. L_0 is the undeformed length of the SMA wire heated to 100% austenitic state and at zero pretension. ϵ_0 is the strain caused by spring pretension, while the SMA is still heated to 100% austenitic state. ϵ_r is the recoverable strain caused by $A-M$ phase transformation. The total strain is $\epsilon = \epsilon_0 + \epsilon_r$. In this paper, we propose a complete mathematical model of this spring-biased SMA wire actuator.

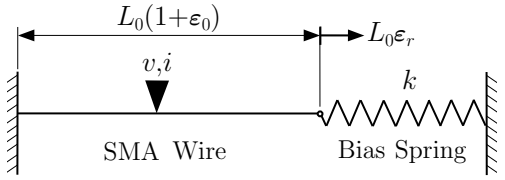

Fig. 2. Schematic of spring-biased SMA wire actuator.

This paper uses experimental results obtained by Ma *et al.* [1] on their experimental setup for spring-biased SMA wire actuator. Here, we give only an outline of the measurement and actuation system of the setup. The details of the experimental setup are given in [1]. A potential difference $v(t)$ is applied across the SMA wire using a programmable power supply. This causes an electric current $i(t)$ to flow through the SMA wire, which gets heated up via the Joule effect. Both $v(t)$ and $i(t)$ are measurable in real-time. Therefore, power vi and electrical resistance $R = v/i$ can also be computed in real-time. The output variable is the transformation strain ϵ_r. The actual strain is measured using a LVDT.

The motivation behind modeling of SMA wire actuator is to be able to control ϵ_r to follow a desired strain $\epsilon_{rd}(t)$. Thus, the SMA wire actuator system can be considered to be any one of the following I/O systems: (a) $v(t) - \epsilon_r(t)$, (b) $i(t) - \epsilon_r(t)$, (c) $R(t) - \epsilon_r(t)$ or (d) $vi - \epsilon_r(t)$. The system dynamics are shown in Fig. 3.

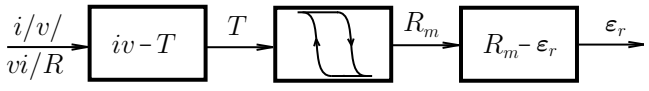

Fig. 3. Block diagram of SMA wire actuator system.

Position control strategies of SMA actuators can be divided into two broad categories. One method uses actual strain measurement and a feedback control loop, whereby the strain error is minimized using any of the several standard closed loop control schemes. However, strain measurement might be expensive [1] or impractical [15]. The other method is to use a mathematical model of the actuator, which predicts the actuator strain $\hat{\epsilon_r}$. This gives us two control options:

- feedback control using computed error $e = \epsilon_{rd} - \hat{\epsilon_r}$, and
- model-based control using inverse dynamics model.

Thus, mathematical modeling of SMA actuators is invaluable to developing control systems for them. While new applications for SMA actuators have been prolifically developed over the years, their mathematical modeling remains an open research problem. Ikuta *et al.* [3] proposed the *variable sublayer model* to derive expressions for both strain and electrical resistance of SMAs. The variable sublayer model hypothesizes that the SMA consists of parallel connected sublayers of different phases with different mechanical characteristics. They also proposed a phenomenological algebraic model for martensite fraction–temperature hysteresis. Their resistance modeling did not produce satisfactory results. It has also been shown [16] that the hysteresis model developed in [3] is not suitable for simulation of a closed-loop control system.

Brailovski *et al.* [17] used a lumped parameter convective heat transfer equation to generate the temperature profile for SMAs, given the voltage and the current. Madill and Wang [16] also used the same equation, assuming constant parameters. They extended the work of [3] to model a SMA wire actuator under constant load and used the model to propose an L2-stable position proportional control system. However, it has been shown that the parameters in the heat transfer equation are temperature-dependent [18], [19]. Recently, Ma *et al.* [1] modeled the system dynamics of Fig. 3 as a black box. They used a neural network model of $R(t) - \epsilon_r(t)$ for position PID control of a spring-biased SMA wire actuator. However, they did not model minor hysteresis loops.

Hysteresis modeling in SMAs has been studied extensively in the literature, and forms a significant part of the SMA actuator modeling literature. Bo and Lagoudas [14] provide an comprehensive review of hysteresis modeling in SMAs. The Preisach model has been widely used to model SMA hysteresis. However, it has been pointed out that the Preisach model is unable to represent dead zones of transformation or drift of hysteresis loops with cyclic loading during partial transformation [20]. Also, computation of the weighting function for every point in the hysteresis region is a tedious process. Many other algebraic models for SMA hysteresis have also been developed [14], [3], [16], [20], [21]. From the control systems point of view, it is advantageous to model the hysteresis using a differential model. Then, the hysteresis model can easily be incorporated into the SMA dynamics state space equations. A differential hysteresis model also facilitates the stability analysis of the system. Yet, there is a lack of differential models of SMA hysteresis in the literature. Likhachev [22] proposed a general differential model for representing SMA hysteresis minor loops, which is probably the only differential model of SMA hysteresis in the literature.

In this paper, we follow a modular approach to modeling each system element (see Fig. 3) of the spring-biased SMA

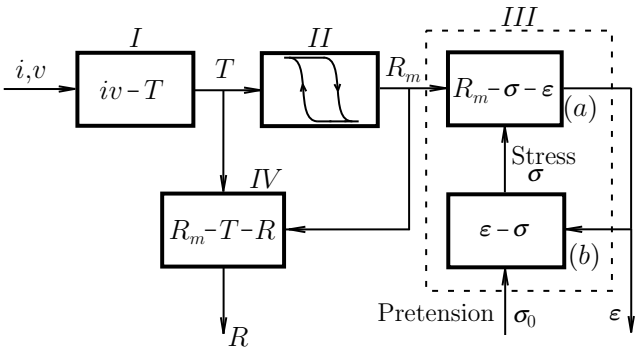

Fig. 4. Proposed modular system block diagram.

wire actuator shown in Fig. 2. Our approach is summarized in Fig. 4. We make four main contributions:

- We use temperature-dependent parameters in the lumped parameter convective heat transfer equation to obtain accurate temperature profile. (Fig. 4, module *I*)
- We propose particular functions for Likhachev's general differential hysteresis model [22] to model the major and minor loops for martensite fraction–temperature hysteresis. (Fig. 4, module *II*)
- We extend Madill and Wang's [16] strain analysis for SMA wire actuator under constant load to spring-biased SMA wire actuator. (Fig. 4, modules *III(a)* and *III(b)*)
- We use the variable sublayer model [3] to model SMA resistance. We propose nonlinear temperature coefficients of resistivity for each phase. Modeling of SMA resistance has significant applications. For example, the electrical resistance of a SMA wire may be used to detect the crack propagation in a SMA composite [15]. (Fig. 4, module *IV*)

The remainder of the paper is organized as follows. Section II describes the heat transfer problem and the computation of the temperature profile. Section III presents the martensite fraction–temperature hysteresis differential model. Section IV establishes the actuator strain as a function of pretension and martensite fraction. Section V describes electrical resistance calculations. Section VI discusses parameter identification and simulation results. In Section VII, we present our conclusions.

II. Heat Transfer and Temperature

In this section, we describe the heat transfer problem of the SMA wire and develop the temperature profile for the SMA wire. This constitutes module *I* of Fig. 4.

Electrical power is used to heat the SMA wire via the Joule effect. Thus, while the system gains heat energy from the electrical power, it loses heat energy to the environment. The balance of the heat energy governs the temperature of the SMA wire. This is essentially a transient heat transfer problem. A straightforward method for solving this heat transfer problem is the lumped parameter analysis, whereby, it is assumed that the internal resistance of the wire to heat conduction is negligible compared to the convective heat transfer with the environment. This assumption is valid for heat transfer in thin metal samples [17]. We also assume that only natural convection occurs. Let

$\rho =$ mass density of SMA wire [kgm^{-3}],
$c =$ specific heat of SMA wire [Jkg^{-1}°C^{-1}],
$L_0 =$ undeformed SMA wire length in 100% A state [m],
$d_0 =$ cross-sectional diameter of undeformed SMA wire [m],
$\epsilon_0 =$ strain caused by pretension load in 100% A state,
$\epsilon_r =$ strain caused by $A - M$ phase transformation,
$\epsilon = \epsilon_0 + \epsilon_r =$ total strain,
$h =$ convection heat transfer coefficient [Wm^{-2}°C^{-1}],
$v =$ voltage across SMA wire [V],
$i =$ current through SMA wire [A], and
$T_{amb} =$ ambient temperature [°C].

Then, the temperature T [°C] of the SMA wire is governed by the following lumped parameter convective heat transfer equation:

$$\rho c \frac{\pi d_0^2 L_0}{4} \frac{dT}{dt} = vi - \pi d_0 L_0 \left(1 + \frac{\epsilon}{2}\right) h(T - T_{amb}).$$

Neglecting $\epsilon/2$, we get

$$\rho c \frac{\pi d_0^2 L_0}{4} \frac{dT}{dt} \approx vi - \pi d_0 L_0 h(T - T_{amb}). \tag{1}$$

Madill and Wang [16] used the above equation assuming that both h and c are constants. However, h is, in general, temperature-dependent [18] and, for SMAs, c is also temperature-dependent [19]. It was found that the following expressions for h and c satisfy experimental results:

$$h = \begin{cases} a_1 - a_2 T, & \dot{T} \geq 0, \\ a_3 + a_4 \operatorname{erf}\left(\frac{T - m_1}{n_1}\right), & \dot{T} < 0, \end{cases} \tag{2}$$

$$c = b_1 + b_2 \operatorname{erf}\left(\frac{T - m_2}{n_2}\right), \tag{3}$$

where a_1, a_2, a_3, a_4, b_1, b_2, m_1, n_1, m_2 and n_2 are constant parameters. Fig. 5 shows a sinusoidal voltage input with decreasing amplitude [1]. The temperature profile for this input is obtained from (1) and is shown in Fig. 6. Strain calculations in Section IV show that for the simulations in this modeling study, the value of ϵ does not exceed 0.105. Thus we are justified in neglecting the $\epsilon/2$ term in (1).

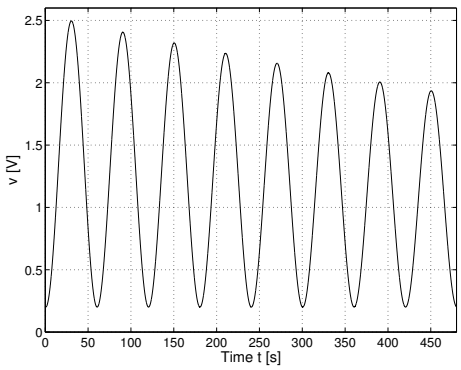

Fig. 5. Input voltage vs time plot.

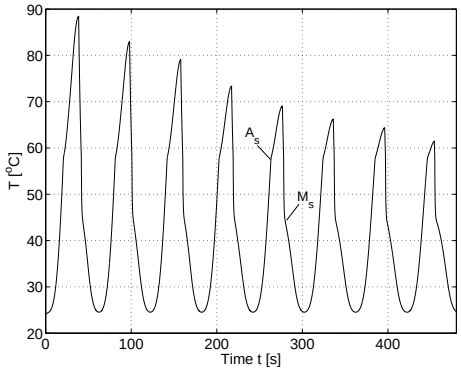

Fig. 6. Simulated temperature vs time plot.

Since the actual temperature was not measured in the experimental setup, the simulated temperature profile could not be compared with the actual temperature profile. Hence, direct validation of the heat transfer model was not possible. It is worth noting that, while (1) has a physical basis, the functions in (2) and (3) are obtained phenomenologically. The choice of these functions and the validity of the heat transfer model can be justified as follows:

- Fig. 7 shows the simulated temperature profile along with the experimentally measured transformation strain. When the strain decreases, the temperature increases. While the strain is constant at its minima, the temperature decreases. When the strain increases, the temperature decreases further. While the strain is constant at its maxima, the temperature increases. This is exactly as observed experimentally [22].
- Transformation temperatures A_s and M_s are shown in Fig. 6. It is known that the $M - A$ transformation is exothermic, while the $A - M$ transformation is endothermic [23]. This observation is also captured by the above heat transfer model. In Fig. 6, the decrease in slope above A_s represents the exothermic transformation, while the decrease in slope below M_s represents the endothermic transformation.

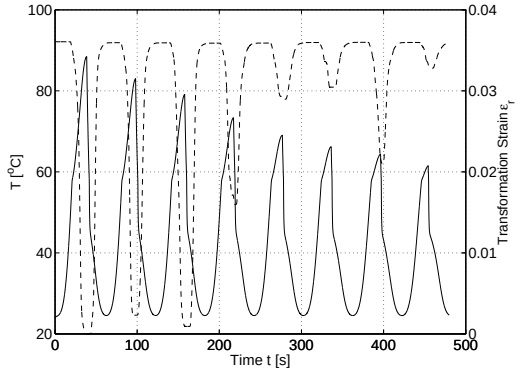

Fig. 7. Simulated temperature and measured transformation strain vs time plot.

In this section, we developed the temperature profile for the SMA wire actuator as a function of the input electrical power and time. As Fig. 4 shows, this is module *I* of the proposed system model, and thus, forms the basis for the remainder of the paper. Some important issues regarding the development of the temperature profile will be discussed in Section VI.

III. Martensite Fraction–Temperature Hysteresis

The crystalline phase transformation between martensite and austenite, and hence the relationship between martensite fraction and temperature, is hysteretic. In this section, we propose special functions for the general differential hysteresis model proposed by Likhachev [22] and delineate the formulation of the martensite fraction–temperature differential hysteresis model, which forms module *II* of Fig. 4.

The typical $R_m - T$ hysteretic relationship for complete $A-M$ transformation was shown schematically in Fig. 1. The hysteresis loop corresponding to complete $A-M$ transformation is called the major hysteresis loop. The area enclosed by the major hysteresis loop is called the hysteresis region. Incomplete $A - M$ transformation yields minor hysteresis loops within the major hysteresis loop, as shown schematically in Fig. 8. The underlying assumption is that the shape and transformation temperatures of minor loops are the same as those of the major loop [3], [16]. Bekker and Brinson [20] have also cited experimental data which corroborates this assumption. The shape of the minor loops is pertinent to hysteresis modeling of the SMA wire actuator because forward and reverse phase transformations can occur at any temperature within the hysteresis region while the actuator is in operation. Another assumption is that the transformation temperatures M_f, M_s, A_s, and A_f are constant throughout. This is not strictly true, since it is known that transformation temperatures of SMAs change with applied stress. A 350 MPa variation of applied stress can change transformation temperatures by upto 100 °C [24]. However, stress calculations in Section IV show that the stress variation throughout the operation of the actuator is merely 4 MPa. Hence, the variation in transformation temperatures is negligible and our assumption is justified.

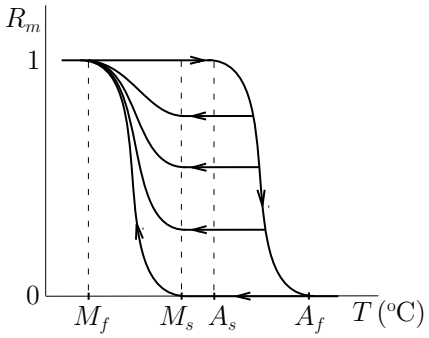

Fig. 8. Schematic of major and minor hysteresis loops.

We follow a phenomenological approach to hysteresis modeling in this section. The proposed hysteresis model is a *Duhem* [25] differential model. Consider an input (independent) variable $u(t)$, an output (dependent) variable $v(t)$, and a hysteresis relationship on the $u(t) - v(t)$ input-output plane. The Duhem model follows intuitively from the fact that if $u(t)$

is an increasing function, then $v(t)$ increases along one path, and if $u(t)$ is a decreasing function, then $v(t)$ decreases along another path, the slopes of the paths being given by functions g_+ and g_-, respectively. Mathematically, the Duhem model is given by

$$\begin{cases} \dot{v} = g_+(u(t), v(t))(\dot{u})^+ - g_-(u(t), v(t))(\dot{u})^-, \\ \\ v(0) = v_0, \end{cases} \tag{4}$$

where $(\dot{u})^\pm = \frac{|\dot{u}| \pm (\dot{u})}{2}$ and $g_+, g_- \in C^0(\mathbb{R}^2)$ [25]. We will refer to g_+ and g_- as *slope functions*.

We propose to use Gaussian probability distribution functions (PDFs) as the slope functions of the major hysteresis loop. A Gaussian PDF is characterized by the mean μ and the variance σ^2. Hence,

$$g_{+/-}(u) = \frac{1}{\sigma_{+/-}\sqrt{2\pi}} \exp\left(-\frac{(u-\mu_{+/-})^2}{2\sigma_{+/-}^2}\right), \tag{5}$$

where subscripts $+$ and $-$ denote increasing and decreasing curves, respectively. Typical plots obtained from (5) are shown in Fig. 9.

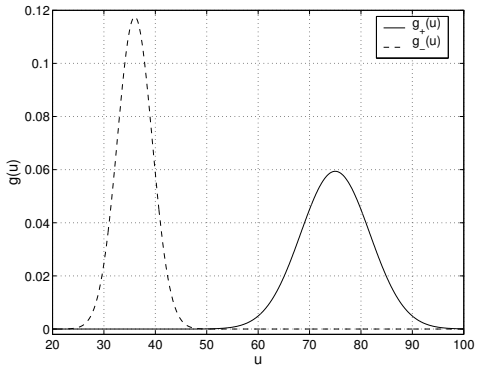

Fig. 9. Gaussian probability distribution functions.

The output v representing the major loop is given by

$$\begin{aligned} v_{+/-} = h_{+/-}(u) &= \int_{-\infty}^{u} g_{+/-}(u')du', \\ &= \frac{1}{2}\left[1 + \text{erf}\left(\frac{u-\mu_{+/-}}{\sigma_{+/-}\sqrt{2}}\right)\right]. \end{aligned} \tag{6}$$

Hence, the major hysteresis loop is given by the differential equation

$$\frac{dv}{du} = \begin{cases} \frac{1}{\sigma_+\sqrt{2\pi}} \exp\left(-\frac{(u-\mu_+)^2}{2\sigma_+^2}\right), & \dot{u} \geq 0, \\ \\ \frac{1}{\sigma_-\sqrt{2\pi}} \exp\left(-\frac{(u-\mu_-)^2}{2\sigma_-^2}\right), & \dot{u} < 0. \end{cases} \tag{7}$$

Typical plots obtained from (6) are shown in Fig. 10. It is evident that the major loop in Fig. 10 would qualitatively match the major loop in Fig. 8 if the y-axis is inverted. This justifies the use of the proposed Gaussian PDFs as slope functions is justified. Now that we have a differential equation for the major loop, we turn our attention to modeling minor

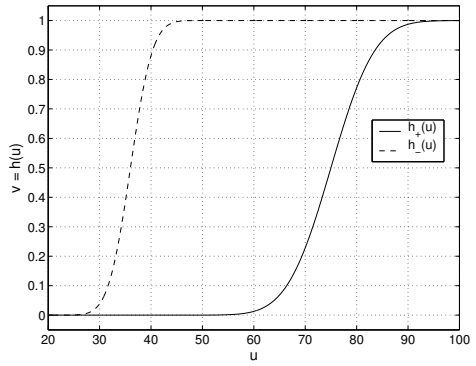

Fig. 10. Gaussian Integral functions represent major loop.

loops. If we multiply the slope functions in (5) by a suitable scaling constant $n_{i_{+/-}} \in [0, 1]$, where $i = 1...N$, we get

$$g_{i_{+/-}}(u) = n_{i_{+/-}} g_{+/-} = \frac{n_{i_{+/-}}}{\sigma_{+/-}\sqrt{2\pi}} \exp\left(-\frac{(u-\mu_{+/-})^2}{2\sigma_{+/-}^2}\right), \tag{8}$$

which is the slope function for the ith minor loop. Typical slope functions for $i = 1, 2, 3$ are shown in Fig. 11. The corresponding minor loop curves are shown in Fig. 12.

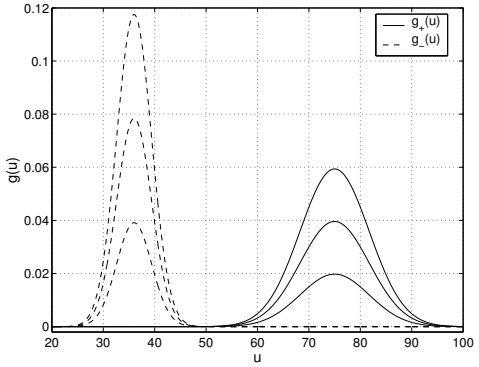

Fig. 11. Slope functions for minor loops.

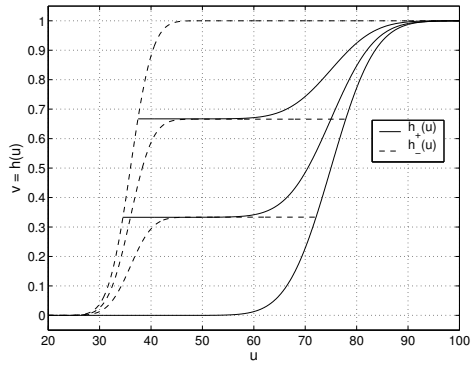

Fig. 12. Minor loop curves.

To obtain a differential model for minor hysteresis loops, it remains to be shown that the scaling constants for each minor curve at any instant t can be expressed in terms of the couple $(u(t), v(t))$. We refer to Fig. 13. $h_{+/-}(u)$ is the

major loop, $h_{1\_}(u)$ and $h_{2\_}(u)$ are minor loop decreasing curves, $g_{+/-}(u)$ is the slope of the major loop and $g_{1\_}(u,v)$ and $g_{2\_}(u,v)$ are slope functions of the minor loop decreasing curves. We observe that scaling constant $n_{1\_}$ can be expressed as:

$$n_{1\_}(u,v) = \frac{b}{a+b} = \frac{h_{1\_}(u) - h_{+}(u)}{h_{-}(u) - h_{+}(u)} = \frac{v - h_{+}(u)}{h_{-}(u) - h_{+}(u)}. \tag{9}$$

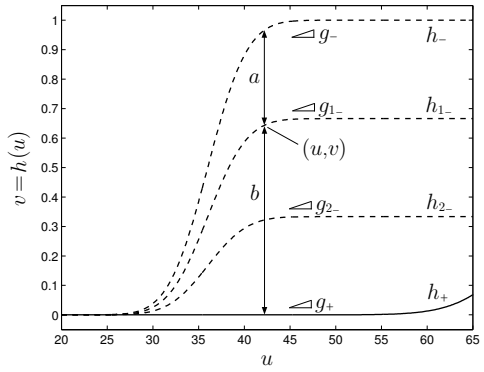

Fig. 13. Method for obtaining differential model for minor loops.

Therefore, the slope function $g_{1\_}$ becomes

$$g_{1\_}(u,v) = \frac{v - h_{+}(u)}{h_{-}(u) - h_{+}(u)} g_{-}(u). \tag{10}$$

It should be noted that for the major loop decreasing curve, $v = h_{-}(u)$, which means that the scaling constant becomes unity. Therefore, the major loop decreasing slope function $g_{-}(u)$ is a special case of (10). A similar analysis can be done for increasing curves. Thus, the complete differential hysteresis model can be expressed as:

$$\frac{dv}{du} = \begin{cases} \frac{h_{-}(u)-v}{h_{-}(u)-h_{+}(u)}\, g_{+}(u), & \dot{u} \geq 0, \\ \\ \frac{v-h_{+}(u)}{h_{-}(u)-h_{+}(u)}\, g_{-}(u), & \dot{u} < 0. \end{cases} \tag{11}$$

We put $u = T$ and $v = R_a$ in (11) to obtain the $R_a - T$ hysteresis model. Putting $1 - v = R_m$ yields the $R_m - T$ hysteresis model:

$$\begin{cases} \frac{dR_m}{dT} = \begin{cases} \frac{h_{-}(T)+R_m-1}{h_{+}(T)-h_{-}(T)}\, g_{+}(T), & \dot{T} \geq 0, \\ \\ \frac{h_{+}(T)+R_m-1}{h_{-}(T)-h_{+}(T)}\, g_{-}(T), & \dot{T} < 0, \end{cases} \\ \\ R_m(0) = 1, \end{cases} \tag{12}$$

where $g_{+/-}$ and $h_{+/-}$ are given by (5) and (6), respectively.

Scaling constants for minor hysteresis loops given in (11) were originally proposed by Likhachev [22]. Fig. 14 shows the hysteresis loops obtained from (12) for the temperature profile shown in Fig. 6. Note that in (12), we have assumed the martensite fraction at time $t = 0$ to be 1. This is a reasonable assumption, because the initial temperature of the actuator is room temperature, which means that the SMA comprises 100% martensite at time $t = 0$. Hence, (12) gives the complete martensite fraction–temperature differential hysteresis model,

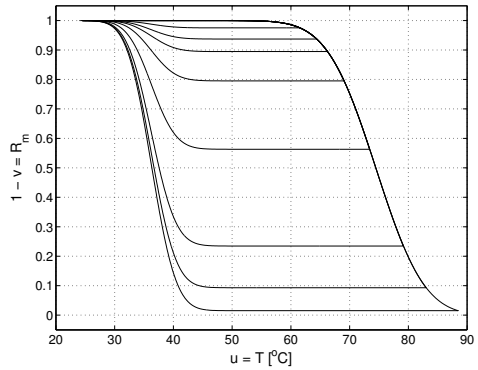

Fig. 14. Simulated martensite fraction vs temperature hysteresis plot.

which forms module *II* of the proposed system model of Fig. 4. Later in this paper, we will use (12) to model the SMA wire strain and electrical resistance as functions of temperature.

IV. Stress and Strain

In this section, we extend the SMA strain results derived by Madill and Wang [16] to the case of the spring-biased SMA wire actuator. This forms module *III* of Fig. 4. Martensite and austenite have very different stress-strain characteristics. They are shown in Fig. 15 and Fig. 16 [16], respectively. Therefore, the SMA wire strain ϵ is a function of both stress σ and martensite fraction R_m. However, stress is caused by the bias spring tension and can be expressed as a function of strain. Therefore, we eliminate σ from the $\epsilon - \sigma - R_m$ relationship and derive ϵ as an explicit function of R_m. Note that $\epsilon = \epsilon_0 + \epsilon_r$. The treatment follows Madill and Wang [16].

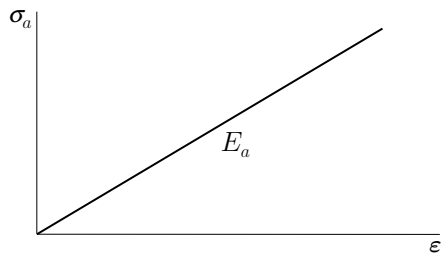

Fig. 15. Schematic stress vs strain curve for austenite.

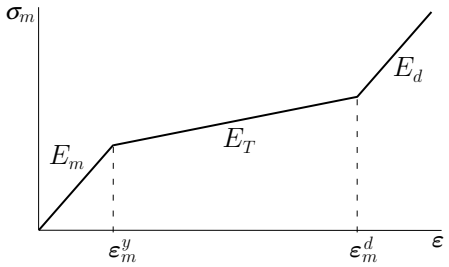

Fig. 16. Schematic stress vs strain curve for martensite.

Let

$\sigma_0 =$ pretension load stress due to bias spring [MPa],
$\Delta =$ spring deformation in 100% austenitic state [m],

σ_a = stress due to 100% austenite [MPa],
σ_m = stress due to 100% martensite [MPa],
E_a = elasticity of austenite [MPa],
E_m = elasticity of fully twinned martensite [MPa],
E_T = elasticity of partly twinned martensite [MPa],
E_d = elasticity of detwinned martensite [MPa],
ϵ_m^y = yield strain of twinned martensite,
ϵ_m^d = minimum strain of detwinned martensite, and
k = bias spring constant [Nm^{-1}].

The pretension load stress σ_0 is given by

$$\sigma_0 = E_a\epsilon_0 = \frac{4k\Delta}{\pi d_0^2(1-\epsilon_0)}. \tag{13}$$

Therefore, the prestrain ϵ_0 is given by

$$\epsilon_0 = \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{16k\Delta}{\pi d_0^2 E_a}}. \tag{14}$$

Also, elongation of the wire causes some spring relaxation, which in turn causes the tensile load on the wire to decrease below σ_0 :

$$\sigma = \frac{4k(\Delta - L_0\epsilon_r)}{\pi d_0^2(1-\epsilon)} = \frac{4k(\Delta - L_0\epsilon_r)}{\pi d_0^2(1-\epsilon_0-\epsilon_r)}. \tag{15}$$

Equation (15) forms module *III(b)* of the proposed system model (Fig. 4). Using the variable sublayer model [3], the tensile load σ gets distributed in the two phases, and is given by

$$\sigma = (1-R_m)\sigma_a + R_m\sigma_m\,. \tag{16}$$

Equation (16), along with Fig. 15 and Fig. 16, forms module *III(a)* of the proposed system model. Substituting for σ from (15), and for σ_a and σ_m from Fig. 15 and Fig. 16, respectively, we obtain the following expression for the total strain $\epsilon = \epsilon_0 + \epsilon_r$:

$$\epsilon = \begin{cases} \epsilon_0 - \frac{d_1}{2c_1} - \frac{\sqrt{d_1^2-4c_1e_1}}{2c_1}, & \text{if } 0 \le \epsilon < \epsilon_m^y, \\ \epsilon_0 - \frac{d_2}{2c_2} - \frac{\sqrt{d_2^2-4c_2e_2}}{2c_2}, & \text{if } \epsilon_m^y \le \epsilon < \epsilon_m^d, \\ \epsilon_0 - \frac{d_3}{2c_3} - \frac{\sqrt{d_3^2-4c_3e_3}}{2c_3}, & \text{if } \epsilon_m^d \le \epsilon, \end{cases} \tag{17}$$

where c_i, d_i and e_i $(i = 1, 2, 3)$ are given in the Appendix. For the simulations in this paper, $\epsilon < \epsilon_m^d$ always. Equation (17) is valid for all values of pretension stress σ_0 because it takes into account the prestrain ϵ_0. However, it is only piecewise differentiable. We approximate the above expression to a polynomial in R_m in order to obtain a continuously differentiable expression for ϵ in the range $\epsilon_m^y \le \epsilon < \epsilon_m^d$:

$$\epsilon = \epsilon_0 + k_1R_m + k_2R_m^2 + k_3R_m^{50}, \tag{18}$$

where k_1, k_2 and k_3 are constant parameters. The above expression is obtained using MATLAB's curve fitting toolbox. For typical parameter values, the strain values obtained from (17) and the approximated strain obtained from (18) are shown in Fig. 17.

In this section, we extended the work of Madill and Wang [16] to obtain the strain in the SMA wire actuator as an explicit function of martensite fraction. As shown in Fig. 4, this result forms module *III* of the proposed system model. It will be combined with the $R_m - T$ hysteresis model developed in the previous section to model the SMA wire strain vs temperature relationship. Simulation results are presented in Section VI.

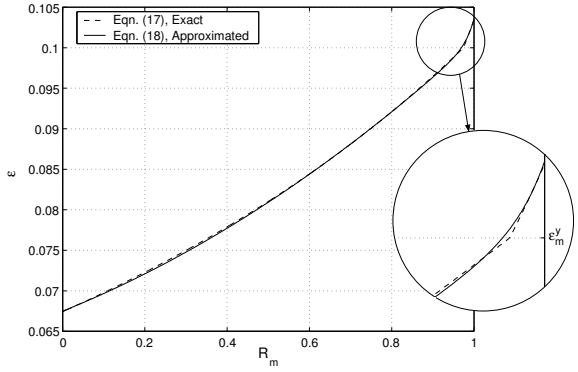

Fig. 17. Simulated strain vs martensite fraction plot.

V. Electrical Resistance

In this section, we model the electrical resistance of the SMA wire as a function of R_m and T, which forms module *IV* of Fig. 4. The electrical resistivities of austenite and martensite are different [3]. Thus the electrical resistance of the SMA wire in any given state depends on the martensite fraction R_m in that state. At the same time, the resistivity of each phase is also a function of the temperature [3]. We use the variable sublayer model [3] to model the electrical resistance of SMA wire actuator. Let

R^a = electrical resistance of 100% austenite [Ω],
R^m = electrical resistance of 100% martensite [Ω],
$\rho_a(T)$ = electrical resistivity of austenite [Ωm], and
$\rho_m(T)$ = electrical resistivity of martensite [Ωm].

The electrical resistance of each phase is given by

$$R^a = \frac{4L_0(1+2\epsilon)}{\pi d_0^2}\rho_a(T), \text{ and} \tag{19}$$

$$R^m = \frac{4L_0(1+2\epsilon)}{\pi d_0^2}\rho_m(T). \tag{20}$$

Using the variable sublayer model [3], the two phases can be considered to be in parallel. Then the SMA wire electrical resistance R is given by

$$\frac{1}{R} = \frac{(1-R_m)}{R^a} + \frac{R_m}{R^m}\,. \tag{21}$$

Substituting for R^a and R^m from (19) and (20), we obtain the following expression for R :

$$\frac{1}{R} = \frac{\pi d_0^2}{4L_0(1+2\epsilon)}\left[\frac{1-R_m}{\rho_a(T)} + \frac{R_m}{\rho_m(T)}\right]. \tag{22}$$

Ikuta *et al.* [3] modeled the resistivities as linear functions of temperature. However, we find that the following nonlinear expressions for the resistivities satisfy experimental results:

$$\rho_a(T) = p_1 + p_2\exp\left(-p_3(T - T_{amb})\right), \text{ and} \tag{23}$$

$$\rho_m(T) = (q_1 - q_2 T)\left[1 + \text{erf}\left(\frac{T - m_3}{n_3}\right)\right] + \sum_{i=0}^{i=9} \alpha_i (T - T_{amb})^i, \quad (24)$$

where p_1, p_2, p_3, q_1, q_2, m_3, n_3 and $\alpha_i (i = 0, ..., 9)$ are constant parameters. The austenite resistivity and martensite resistivity obtained from (23) and (24), respectively, are plotted in Fig. 18. It is pertinent to note that the nonlinear expressions for resistivities in (23) and (24) are obtained phenomenologically. Consider Fig. 14, which shows R_m as a function of T. Suppose T increases from M_f to A_s. We know that $R_m = 1$ in this region. Hence, the electrical resistance R would only depend on ρ_m, and ρ_m should take a form which would yield the shape of the increasing curve in Fig. 20. Similarly, suppose T decreases from A_f to M_s. Then $R_m = 0$, and ρ_a should take a form which would yield the shape of the decreasing curve in Fig. 20. The above reasoning explains qualitatively why the temperature dependence of resistivities of each phase is as shown in Fig. 18. Functions (23) and (24) are obtained using MATLAB's curve fitting toolbox.

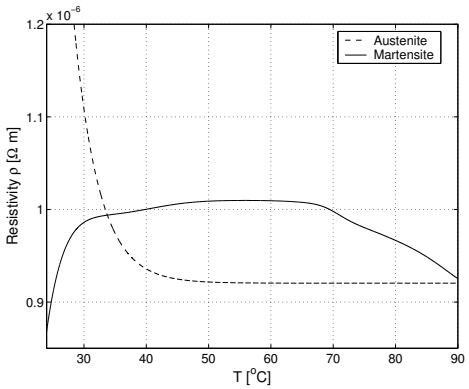

Fig. 18. Simulated electrical resistivity vs temperature plot.

Having obtained R as a function of R_m and T, we will use the hysteresis model developed in Section III to model the SMA wire resistance vs temperature relationship. The next section shows simulation results.

VI. Parameter Identification and Results

Experiments were conducted by Ma *et al.* [1] to study SMA strain during complete and partial phase transformations in a spring-biased SMA wire actuator. We use some of their experimental data to compare and validate the proposed model with. In this section, we present parameter identification for the proposed model and simulation results from the proposed model and their comparison with the experimental data of [1].

Table I shows the parameter values used in this modeling study. Parameters used in (1) and (17) are either obtained empirically, or taken from literature [16]. First estimates for parameters used in (18) were obtained by curve fitting (18) to (17). The remaining parameters were identified using the actual strain and electrical resistance measurements, since the actual temperature measurements were not available, and martensite fraction measurements were not possible. Parameters used in (2), (3), (5), (18), (23), and (24) were obtained by nonlinear least-squares optimization. This was accomplished in MATLAB using the Levenberg-Marquardt algorithm [26] and with strain and electrical resistance errors as the objective functions. The starting values of the parameters for the optimization algorithm were obtained by hand-tuning.

TABLE I
Simulation parameters

ρ	6500 kgm^{-3}	T_{amb}	24 °C
L_0	0.2286 m	d_0	3.81×10^{-4} m
a_1	165 Wm^{-2}°C^{-1}	b_1	1400 Jkg^{-1}°C^{-1}
a_2	0.5 Wm^{-2}°C^{-2}	b_2	1000 Jkg^{-1}°C^{-1}
a_3	300 Wm^{-2}°C^{-1}	m_1	48 °C
a_4	150 Wm^{-2}°C^{-1}	m_2	58 °C
n_1	10 °C	n_2	0.5 °C
μ_+	78.9 °C	σ_+	11.2 °C
μ_-	34 °C	σ_-	5.8 °C
Δ	0.04425 m	k	58.19 Nm^{-1}
E_a	35917 MPa	E_m	20480 MPa
E_T	826 MPa	E_d	16800 MPa
ϵ_m^y	0.1	ϵ_m^d	0.15
k_1	0.0204	k_2	0.01293
k_3	0.0027		
p_1	9.2×10^{-7} Ωm	q_1	3.4×10^{-8} Ωm
p_2	8.4×10^{-7} Ωm	q_2	5.7×10^{-10} Ωm°C^{-1}
p_3	0.2499 °C^{-1}	m_3	70 °C
n_3	3 °C		
α_0	8.7×10^{-7} Ωm	α_1	4.8×10^{-8} Ωm°C^{-1}
α_2	-7.8×10^{-9} Ωm°C^{-2}	α_3	7.0×10^{-10} Ωm°C^{-3}
α_4	-3.7×10^{-11} Ωm°C^{-4}	α_5	1.2×10^{-12} Ωm°C^{-5}
α_6	-2.5×10^{-14} Ωm°C^{-6}	α_7	3.2×10^{-16} Ωm°C^{-7}
α_8	-2.2×10^{-18} Ωm°C^{-8}	α_9	6.7×10^{-21} Ωm°C^{-9}

Fig. 19 shows the SMA transformation strain obtained by applying (18) to the hysteresis model (12). Fig. 20 shows the SMA electrical resistance obtained by applying (22) to the hysteresis model (12). Note that, in both figures, even for the plots titled "experimental", the quantity on the x-axis is the simulated temperature. It has already been pointed out that actual temperature was not measured in the experimental setup. We choose temperature as the x-axis to clearly illustrate the dynamics of strain and electrical resistance as compared to the martensite fraction–temperature hysteresis.

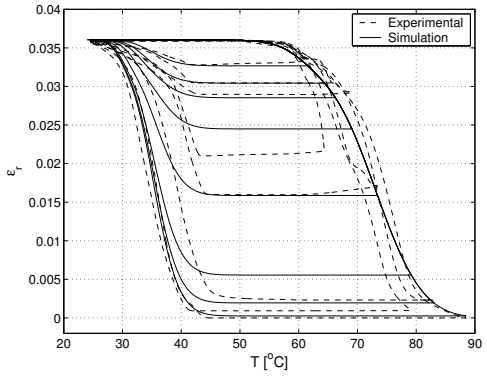

Fig. 19. Strain vs temperature.

Finally, the $\epsilon_r - R$ relation is plotted in Fig. 21. Ma *et al.* [1] used a neural network model of the $\epsilon_r - R$ major loop to achieve position PID control of spring-biased SMA wire actuator. Fig. 21 shows that the proposed model is capable of

modeling $\epsilon_r - R$ minor loops as well, and therefore, may be used to achieve more accurate position control of the actuator.

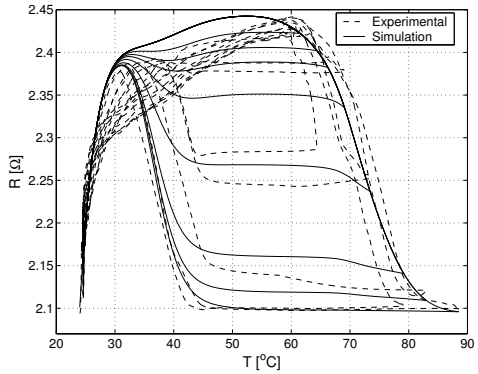

Fig. 20. Electrical resistance vs temperature.

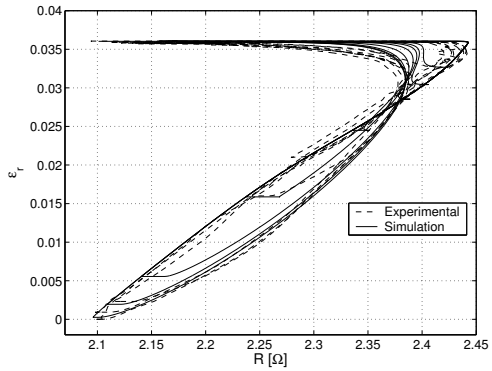

Fig. 21. Strain vs electrical resistance.

It can be seen from Figs. 19, 20 and 21 that the simulation results from the proposed model are in qualitative agreement with the experimental results. The discrepancy between simulated and experimental data can be explained as follows. The measured transformation strain shown in Fig. 7 corresponds to the voltage profile of Fig. 5. Thus, the strain amplitude should gradually decrease with each voltage cycle. However, Fig. 7 shows significant anomalies in the measured strain profile. These can be attributed to the fact that the ambient conditions of the experimental setup are not constant over time and space, since the experimental setup is not contained in a hermetically sealed enclosure and the natural convection assumption made in Section II does not strictly hold. Thus, the actual temperature of the SMA wire deviates indeterministically from what is theoretically calculated. Another source of error in the simulated temperature profile is that the functions in (2) and (3) were obtained empirically from the measured transformation strain, since the actual temperature data was not available. Hence, there is an unavoidable error in the simulated temperature profile, which is carried over to modules *II*, *III*, and *IV*. Additionally, it is common for a film of Titanium Oxide to form at the contacts of the SMA wire over time. This introduces extraneous electrical resistance into the circuit, affecting the experimentally measured electrical quantities. The shorter the length of the SMA wire, the higher the error due to Titanium Oxide resistance. It has also been pointed out by Ma *et al.* [1] that the experiments conducted on the SMA wire do not exhibit accurate repeatability until it has been subjected to sufficient thermal cycling. Hence, the simulation results should be viewed in light of the above issues.

It is evident that the actuator behavior is highly nonlinear due to the complex physics of the SMA. There have been several physics-based thermomechanical models of SMAs (for example [14], [20]), which have not found use in the control of SMA actuators due to their complicated nature. Preisach models have found widespread use in SMA modeling, but their limitations have been mentioned in Section I. The proposed model does not suffer from these limitations. It is computationally simple, enabling inexpensive real-time control of the SMA wire actuator. The above discussion demonstrates the effectiveness of our modeling study.

VII. Conclusions

In this paper, we developed a complete mathematical model for a spring-biased shape memory alloy wire actuator. The model is capable of simulating the temperature, martensite fraction, stress, strain and electrical resistance of the actuator. We used lumped parameter analysis to obtain the temperature profile. We proposed particular functions for Likhachev's general differential hysteresis equations [22] to represent major and minor hysteresis loops. The model also extends the variable sublayer model to obtain strain and electrical resistance of the actuator.

One or more modules of the model can be extended to other applications. For example, the hysteresis and electrical resistance modules may be used to detect the crack propagation in a SMA composite [15]. It should also be possible to eliminate temperature and martensite fraction and to obtain a strain–electrical resistance differential model. This will facilitate resistance feedback position control applications such as [1]. A salient feature of the proposed model is that it is built upon a differential hysteresis model. The differential hysteresis model proposed in Section III uses only four parameters to generate major and minor loops for any temperature profile. This presents a significant advantage over the algebraic model developed in [16], which uses an algorithm to compute new parameter values in each thermal half-cycle. We have also shown that the proposed differential hysteresis model represents the dead zone of phase transformation (Fig. 8, region between M_s and A_s), and is, in that sense, superior to the Preisach model. Furthermore, the proposed differential hysteresis model makes it convenient to obtain a state space model of the SMA actuator. It facilitates the use of the rich theory of ordinary differential equations for stability analysis. It is also feasible to construct an inverse of the differential hysteresis model for inverse compensation of hysteresis.

We also discussed significant issues regarding temperature modeling of SMAs. This discussion yields the direction for enhancement of the proposed model. The proposed model can be refined if actual temperature of the SMA wire is measured during experiments and the heat transfer model is validated using this data. Further, by placing the experimental setup in

a hermetically sealed enclosure, the temperature profile can be much better controlled, which would obviously yield more accurate results.

ACKNOWLEDGMENTS

The authors would like to acknowledge the help of Dr. G. Song and his graduate student N. Ma at the University of Houston for making experimental data available for this modeling study.

APPENDIX

The parametric variables in (17) are given by

$$
\begin{aligned}
c_1 &= \pi d_0^2[E_a + (E_m - E_a)R_m],\\
d_1 &= -4kL_0 - \pi d_0^2[E_a + (E_m - E_a)R_m](1 - 2\epsilon_0),\\
e_1 &= 4k\Delta - \pi d_0^2[E_a + (E_m - E_a)R_m]\epsilon_0(1 - \epsilon_0),
\end{aligned}
$$

$$
\begin{aligned}
c_2 &= \pi d_0^2[E_a + (E_T - E_a)R_m],\\
d_2 &= -4kL_0 - \pi d_0^2[E_a + (E_T - E_a)R_m](1 - 2\epsilon_0)\\
&\quad + \pi d_0^2 R_m \epsilon_m^y (E_m - E_T),\\
e_2 &= 4k\Delta - \pi d_0^2(1 - \epsilon_0)[E_a\epsilon_0 + (E_T - E_a)R_m\epsilon_0\\
&\quad + (E_m - E_T)R_m\epsilon_m^y,
\end{aligned}
$$

$$
\begin{aligned}
c_3 &= \pi d_0^2[E_a + (E_d - E_a)R_m,\\
d_3 &= -4kL_0 - \pi d_0^2[E_a + (E_d - E_a)R_m](1 - 2\epsilon_0)\\
&\quad + \pi d_0^2 R_m[(E_m - E_T)\epsilon_m^y + (E_T - E_d)\epsilon_m^d],\\
e_3 &= 4k\Delta - \pi d_0^2(1 - \epsilon_0)[E_a\epsilon_0 + (E_d - E_a)R_m\epsilon_0\\
&\quad + (E_m - E_T)R_m\epsilon_m^y + (E_T - E_d)R_m\epsilon_m^d].
\end{aligned}
$$

REFERENCES

[1] N. Ma, G. Song, and H.-J. Lee, "Position Control of Shape Memory Alloy Actuators with Internal Electrical Resistance Feedback," in *SPIE - The International Society for Optical Engineering, Smart Structures and Materials: Modeling, Signal Processing, and Control*, vol. 5049, 2003, pp. 46–55.

[2] Z. Bo and D. Lagoudas, "Thermomechanical Modeling of Polycrystalline SMAs under Cyclic Loading, Part I: Theoretical Derivations," *International Journal of Engineering Science*, vol. 37, pp. 1089–1140, 1999.

[3] K. Ikuta, M. Tsukamoto, and S. Hirose, "Mathematical Model and Experimental Verification of Shape Memory Alloy for Designing Micro Actuator," in *IEEE Micro Electro Mechanical Systems*, 1991, pp. 103–108.

[4] T. Saburi, *Shape Memory Materials*. Cambridge: Cambridge University Press, 1998, ch. 3, pp. 49–96.

[5] J. Kudva, B. Sanders, J. Pinkerton-Florance, and E. Garcia, "Overview of the DARPA/AFRL/NASA Smart Wing Phase 2 Program," in *SPIE - The International Society for Optical Engineering, Smart Structures and Materials 2001: Industrial and Commercial Applications of Smart Structures Technologies*, vol. 4332, 2001, pp. 383–389.

[6] M. Hashimoto, M. Takeda, H. Sagawa, I. Chiba, and K. Sat, "Shape Memory Alloy and Robotic Actuators," *Journal of Robotic Systems*, vol. 2, no. 1, pp. 3–25, 1985.

[7] K.-Y. Tu, T.-T. Lee, C.-H. Wang, and C.-A. Chang, "Design of a Fuzzy Walking Pattern (FWP) for a Shape Memory Alloy (SMA) Biped Robot," *Robotica*, vol. 17, no. 4, pp. 373–382, 1999.

[8] K. Ikuta, M. Tsukamoto, and S. Hirose, "Shape Memory Alloy Servo Actuator System with Electric Resistance Feedback and Application for Active Endoscope," in *IEEE International Conference on Robotics and Automation*, vol. 1, 1988, pp. 427–430.

[9] M. Bergamasco, F. Salsedo, and P. Dario, "Shape Memory Alloy Micromotors for Direct-Drive Actuation of Dexterous Artificial Hands," *Sensors and Actuators*, vol. 17, no. 1, pp. 115–119, 1989.

[10] K. Gabriel, W. Trimmer, and J. Walker, "A Micro Rotary Actuator using Shape Memory Alloys," *Sensors and Actuators*, vol. 15, no. 1, pp. 95–102, 1988.

[11] R. Jebens, F. Salsedo, and P. Dario, "Microactuators for Aligning Optical Fibers," *Sensors and Actuators*, vol. 20, no. 1, pp. 65–73, 1989.

[12] J. Harrison and D. Hodgson, "Use of TiNi in Mechanical and Electrical Connectors," in *International Symposium on Shape Memory Effects and Applications*, 1975, pp. 517–523.

[13] V. Michaud, "Can Shape Memory Alloy Composites be Smart?" *Scripta Materialia*, vol. 50, no. 2, pp. 249–253, 2004.

[14] Z. Bo and D. Lagoudas, "Thermomechanical Modeling of Polycrystalline SMAs under Cyclic Loading, Part IV: Modeling of Minor Hysteresis Loops," *International Journal of Engineering Science*, vol. 37, pp. 1205–1249, 1999.

[15] X. Wu, J. Wu, and Z. Wang, "The Variation of Electrical Resistance of Near Stoichiometric NiTi during Thermo-Mechanic Procedures," *Smart Materials and Structures*, vol. 8, pp. 574–578, 1999.

[16] D. Madill and D. Wang, "Modeling and L2-Stability of a Shape Memory Alloy Position Control System," *IEEE Transactions on Control Systems Technology*, vol. 6, no. 4, pp. 473–481, 1998.

[17] V. Brailovski, F. Trochu, and G. Daigneault, "Temporal Characteristics of Shape Memory Linear Actuators and their Application to Circuit Breakers," *Materials and Design*, vol. 17, no. 3, pp. 151–158, 1996.

[18] S. Sukhatme, *A Textbook on Heat Transfer*. Hyderabad: Universities Press, 1996.

[19] J. McNichols and J. Cory, "Thermodynamics of Nitinol," *Journal of Applied Physics*, vol. 61, no. 3, pp. 972–984, 1987.

[20] A. Bekker and L. Brinson, "Phase Diagram based Description of the Hysteresis Behavior of Shape Memory Alloys," *Acta Materialia*, vol. 46, no. 10, pp. 3649–3665, 1998.

[21] Y. Ivshin and T. Pence, "A Constitutive Model for Hysteretic Phase Transition Behavior," *International Journal of Engineering Science*, vol. 32, no. 4, pp. 681–704, 1994.

[22] A. Likhachev, "Differential Equation of Hysteresis: Application to Partial Martensitic Transformation in Shape-Memory Alloys," *Scripta Metallurgica et Materialia*, vol. 32, no. 4, pp. 633–636, 1995.

[23] Y. Liu, "Detwinning Process and its Anisotropy in Shape Memory Alloys," in *SPIE - The International Society for Optical Engineering, Smart Materials*, vol. 4234, 2000, pp. 82–93.

[24] A. Keefe, "Thermo-mechanical Characterization of Shape Memory Alloy Torque Tube Actuators," Master's thesis, University of California, Los Angeles, California, 1999.

[25] A. Visintin, *Differential Models of Hysteresis*. Berlin: Springer-Verlag, 1994.

[26] T. Coleman, M. Branch, and A. Grace, "Optimization Toolbox," User's Guide, The Mathworks Inc., 1999.