Novel Phase Transformation Model for Shape Memory Alloy Actuators

Lawrence Smith[1] | *lsmith@fprin.com*

***Abstract*— A novel phase transformation model for SMA wire actuators is posed, and fit quality to empirical data is measured favorably as compared to other popular models. The phase transformation model is implemented in a full transient model of actuation, accounting for thermal, mechanical, and electrical effects.**

I. MOTIVATION

A. Shape Memory Alloy Basics

Shape Memory Actuator (SMA) wire exhibits two solid phases with distinct crystalline structures, and can generate force or displacement when caused to transition between these phases. Strains introduced when an actuator is at low temperature, and composed of the Martensite phase, are recovered when the actuator is heated to its high-temperature phase, Austenite. This phase transformation is not instantaneous, and occurs smoothly as a function of temperature, stress state, and strain [1]. An effective thermomechanical model of SMA wire requires an accurate, succinct mathematical description of this transformation curve. Several definitions of the transformation curve of various mathematical complexity have been posed, one of the earliest and most widely used by Brinson in 1993 [2]. Though much work has been done towards predicting the behavior of SMA actuators in practical applications, no universally accepted model has emerged[3]. Because SMA actuators are highly attractive as low-complexity, high-cycle, exceedingly compact mechanical power sources, and because of their valuable applications to development of highly reliable devices and products, capturing their behavior in computationally efficient numerical models remains a thoroughly interesting problem.

B. Practical SMA Implementation

When properly implemented, SMA actuators can deliver upwards of 10^6 cycles (transitions between solid phases), with 200 MPa axial stress and 5.0% axial strain available at each cycle. They are also exceedingly compact, with about 1000 times less mass than a solenoid of equivalent mechanical power output. However, special care must be taken to ensure proper function and long life of an SMA actuator, as it may be irreversibly damaged by either high operating temperatures (>200 C) or high stresses (>600 MPa).

In practice, SMA wire can easily be made to undergo transformation by passing current through the wire, which raises the temperature of the wire as a result of Joule heating. The evolution of phase transformation can be monitored by sensing the resistance of the actuator to avoid overheating and permanent damage. In order to prevent the development of excessive stresses, SMA actuators are typically designed in series with a compliant element with an effective stiffness orders of magnitude lower than the axial stiffness of the wire.

C. Shortcomings of Existing Models

SMA wire actuators are highly attractive for use in miniature mechanical devices because of their small size and ability to develop relatively large forces. These actuators can be purchased at diameters that approach $100\,\mu\mathrm{m}$, and manufacturers readily supply typical actuator forces or recoverable strains that SMA wires achieve when *full transformation* - a conversion from nearly fully martensite to nearly fully austenite is achieved. However, with wire of these exceedingly small diameters, several factors challenge full phase transformation:

- surface area to volume ratio increases with decreasing diameter - this creates unfavorable heating conditions because the rate of thermal energy generation (Joule heating) scales with volume and rate of thermal energy loss (convection) scales with surface area
- typical electrical energy sources used at this scale (coin cell batteries) are limited in terms of maximum current draw, voltage stiffness, and stability over the life of the battery
- designs which optimize for miniaturization may leverage actuators that operate near limits: maximum force or displacement available, minimum time or energy required to actuate - and thus device performance may be sensitive to variation in wire properties, load cases, or ambient conditions

These factors mean that full transformation often cannot be assured. Thus, when driving SMA wire actuators at this scale, an accurate description of stress/strain state during partial conversion is *particularly* critical. While many of the existing transformation models are capable of capturing wire behavior in certain distinct regions of the transformation process, none are general enough to be fit effectively throughout the transformation. Additionally, few examples of open-source 1D numerical models of SMA actuators exist. This paper aims to remedy both of these shortcomings.

D. 1D Model Framework

To simulate the performance of a representative SMA actuator in practical application, we will consider the 1D mechanical system depicted in Figure 1. An ideal voltage

*This work was supported by Aproio LLC, a Product Engineering Consulting Company, based in San Jose, California.

[1]L. Smith is pursuing a PhD in Mechanical Engineering at University of Colorado, Boulder. He claims no conflict of interest.

source applies electrical stimulus to the SMA wire of nominal length and diameter, resulting in current flow through the wire. One end of the actuator is fixed and the other is connected in series with a perfect Hookean spring.

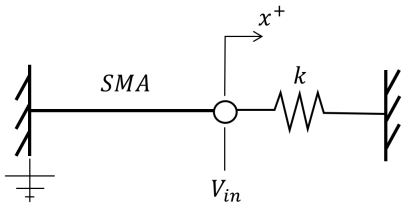

Fig. 1: 1D Mechanical Model of SMA Actuator

The thermomechanical model developed in Section III will represent this simple system by taking the following physical effects into account:

- Thermal effects of Joule heating
- Convective and conductive losses
- Energy storage and absorption by specific heat and latent heat of phase transformation
- Reversible strain recovery by transition between solid phases
- Stress-induced shifts in transformation temperature
- Temperature and Phase dependent material properties like electrical resistance and elastic modulus

II. APR Transformation Curve Definition

A. Phase Transformation Curve Basics

The variable composition of SMA is described by the parameter ξ, the volume fraction martensite - ranging from 1 for a sample that is completely Martensite and 0 for a sample that is completely Austenite. This parameter is a function of the temperature T, stress state σ, and initial volume fraction martensite ξ_0, which is determined by the temperature history of the sample.

B. Asymmetric Piecewise Rational Phase Transformation Description

The transformation curve presented here is a piecewise construction of two rational expressions. For the transition from martensite to austenite, which is induced by heating, we present:

if $T \leq b_A + \sigma/C_A$

$$\xi_{M\to A_1} = \xi_0\big(\frac{a_{A1}}{T - \sigma/C_A - b_A - c_A} + 1 + d_{A1}\big)$$

else:

$$\xi_{M\to A_2} = \xi_0\big(\frac{a_{A2}}{T - \sigma/C_A - b_A + c_A} - d_{A2}\big)$$

The shape of the transformation curve described by these equations is controlled by several independent constants: *b*, the centerline temperature of transformation, which is also the point of coincidence of the pair of curves; and c_{A1}/c_{A2}, parameters that control the curvature of the transformation on either side of the centerpoint. In practice, these are chosen independently and the remaining constants d_{A1},d_{A2},a_{A1}, and a_{A2} are chosen to enforce coincidence and tangency at the point of intersection of the curves.

Similarly for the austenite to martensite transition, which occurs during cooling, the instantaneous volume fraction martensite ξ is given by the pair of curves:

if $T \leq b_M + \sigma/C_M$

$$\xi_{A\to M_1} = (1-\xi_0)\big(\frac{a_{M1}}{T - \sigma/C_A - b_M - c_M} + 1 + d_{M1}\big) + \xi_0$$

else:

$$\xi_{A\to M_2} = (1-\xi_0)\big(\frac{a_{M2}}{T - \sigma/C_A - b_M + c_M} - d_{M2}\big) + \xi_0$$

Note that ξ is a function of both instantaneous temperature T and axial stress σ. The stress-induced shift in phase transformation coefficients C_A and C_M are intrinsic material properties and are determined experimentally, and typically have values in the range of 10 MPa/K. Practically, the addition of stress drives the temperatures required to achieve transformation higher, which can lead to incomplete transformation or lower energy efficiency.

C. Example of Transformation Curves

Representative values for transformation curve input parameters are given in Table I, and produce the example curve in Figure 2. These values will be used in the numerical model introduced in subsequent sections, and demonstrate how the asymmetric piecewise rational (APR) description of SMA phase transformation gives precise control of the mathematical description used in our model.

TABLE I: Representative APR Model Curve Parameters

Parameter	Value	Description
b_A	75	$^\circ C$ Centerpoint Temperature, $M \to A$
c_{A1}	4	$^\circ C$ Curvature Parameter 1, $M \to A$
c_{A2}	2	$^\circ C$ Curvature Parameter 2, $M \to A$
b_M	50	$^\circ C$ Centerpoint Temperature, $A \to M$
c_{M1}	12	$^\circ C$ Curvature Parameter 1, $A \to M$
c_{M2}	4	$^\circ C$ Curvature Parameter 2, $A \to M$

III. Intermediate Results: Comparison to Empirical Data

In order to judge the ability of a model to accurately capture the shape of an observed transformation curve, we can compare the model fit to empirical data. Ideally, we would have readily available ξ vs. T data to compare to the analytical relationship described in the previous section. However, direct empirical measurement of the phase of a sample of wire (volume fraction martensite vs. austenite) is to our knowledge impossible to measure in real time, and must be inferred from macroscale length change or deformation in a SMA sample.

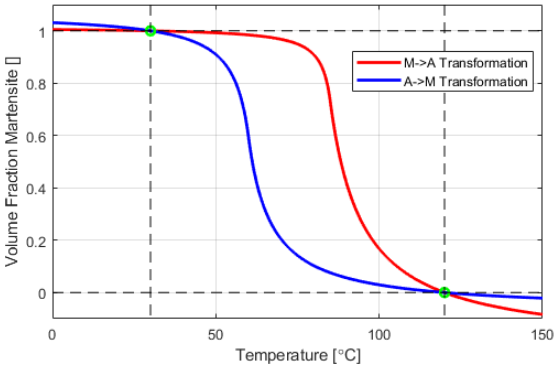

Fig. 2: Representative transformation curves for $M \rightarrow A$ (red) and $A \rightarrow M$ (blue) transitions using the APR model

For our analysis, we use data gathered on Nickel-Titanium SMA wires in uniaxial tension, provided by Hernandez and Hartl from [4]. This data was gathered on a wire placed under constant uniaxial tensile stress equal to 150MPa, which was then heated slowly in a thermal chamber. Under these conditions, the wire transitions slowly from near fully martensite to near fully austenite. The recovered strain in the sample of wire (about 5%) amounts to millimeters of length change, which is easily measurable. We assume that instantaneous volume fraction martensite throughout the full test can be inferred from instantaneous strain:

$$\xi = 1 - \frac{\varepsilon}{\varepsilon_{max}}$$

To make a consistent comparison across models, we compare the fit quality of several models to the same set of empirical data. Best-fit values of the input parameters were found using the standard *Matlab* function *fminsearch*, a derivative-free optimization routine based on the simplex method, to minimize the sum of the squared error between the empirical data and a model fit. Because of the simplicity of this optimization problem, convergence to a global minimum can be assessed by eye, and is ensured when initial guesses which are sufficiently close to the optimum parameter set are supplied.

In Figure 3, note that the APR model obtains a closer fit to the empirical data than the other three models presented, by a factor of about 5. While this is significant, the more critical improvement is apparent when the local error between model predictions and empirical data is analyzed.

In Figure 4, we see that local error stays within $\pm 3\%$ for the APR model, which is a 2x improvement over the other models - *the APR model is more accurate over the full transition region*. The APR model's improvement over other models is particularly striking over the transformation range of $0.7 < \xi < 1.0$. As noted in the introduction to this paper, this region is of particular interest when modeling the actuation of wires in miniature electromechanical devices, which often operate without full actuation being ensured. Thus *the APR model is more accurate in the region of particular interest for small electromechanical devices*.

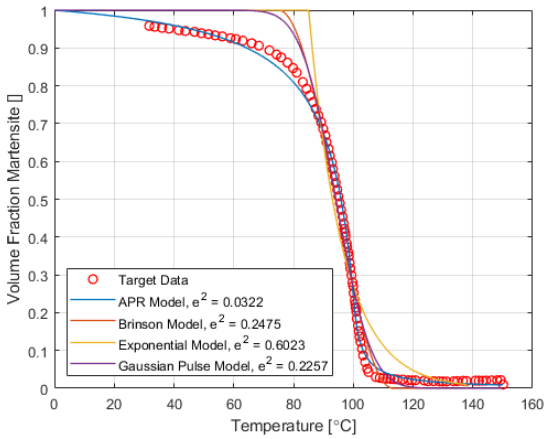

Fig. 3: Visualization of best-fit curves for four models with total error minimized *Data courtesy of D. Hartl and E. Hernandez*

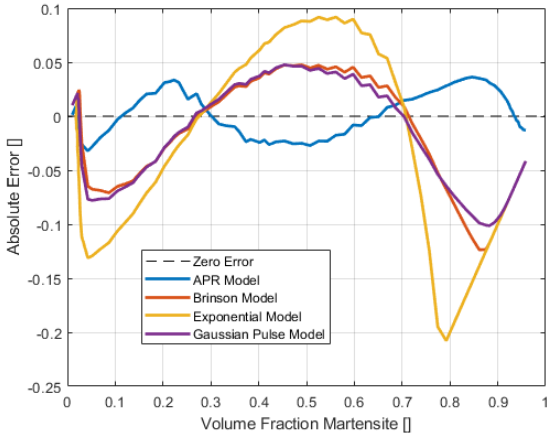

Fig. 4: Local Error between model fit and empirical data for four models, with total error minimized

In the following section, we will develop a full 1D transient electromechanical model of SMA wire around the APR phase transformation model, and observe simulated cycles of actuation

IV. Development of Numerical Model

The development of our full electro-thermo-mechanical model follows closely from [3], [5], and [2] and the reader may refer to these papers for supporting detail.

A. Derivation of Governing Equations

In order to capture the thermal response of the wire, we begin with an incremental form of the first law of thermodynamics:

$$E_{in} = E_{out} + E_{store}$$

We will solve this equation at each timestep in our simulation using the Newton-Raphson method, so the equation

above is rearranged into a residual function that represents equilibrium when f_1 is driven to zero.

$$f_1 = E_{in} - E_{out} - E_{store}$$

To expand this expression, we substitute the following expressions: Joule heating for E_{in}, convective losses for E_{out}, specific heat capacity and latent heat capacity for E_{store}. This leaves the final form of the thermodynamic equation in our model:

$$f_1 = I^2 R\Delta t - hA(T_i - T_a)\Delta t - mc(\Delta T) - mL(\Delta\xi)$$

where Δ denotes the difference in a particular quantity at the current timestep and at the previous timestep:

$$\Delta T = T_i - T_{i-1}$$
$$\Delta\xi = \xi_i - \xi_{i-1}$$
$$\Delta t = t_i - t_{i-1}$$

In our model, the equation above is driven to zero by taking successively better guesses at the temperature T at the given timestep t_i using the Newton-Raphson method:

$$T_{i+1} = T_i - m\frac{f_1}{\frac{\partial f_1}{\partial T}}$$

This guess requires $\frac{\partial f_1}{\partial T}$, which is given in the Appendix and itself contains the piecewise partial derivative of ξ. The relaxation factor m modifies the classical Newton-Raphson method by reducing the magnitude of iteration step sizes, and is discussed in more depth in the following section.

The constitutive law for shape memory actuator wire is widely established for 1D models [3] [5] [2] as:

$$\frac{\partial\sigma}{\partial t} = E\frac{\partial\varepsilon}{\partial t} + \Omega\frac{\partial\xi}{\partial t}$$

where Ω is the phase-dependent stress contribution defined as

$$\Omega = -\varepsilon E(\xi)$$

and $E(\xi)$, the phase-dependent elastic modulus, is defined as

$$E(\xi) = E_A - \xi(E_A - E_M)$$

Converting our governing equation into residual form and discretizing in time, we are left with the final form of the constitutive equation implemented in our model:

$$f_2 = (\sigma - \sigma_0) - E(\varepsilon - \varepsilon_0) - \Omega(\xi - \xi_0)$$

Again, the residual is driven to zero by taking successively better guesses at the axial stress σ at the given timestep t_i

$$\sigma_{i+1} = \sigma_i - \frac{f_2}{\frac{\partial f_2}{\partial T}}$$

The partial derivatives required to evaluate this expression are given in the Appendix.

B. Implementation, Convergence, Error Estimation

In our model, the Newton-Raphson routines which solve our governing equations by driving the residuals f_1 and f_2 are implemented in a pair of nested loops, which iterate towards convergence at each timestep. Secondary outputs (power dissipation, wire length, rate of temperature change, etc.) are calculated after convergence has been reached inside a given timestep.

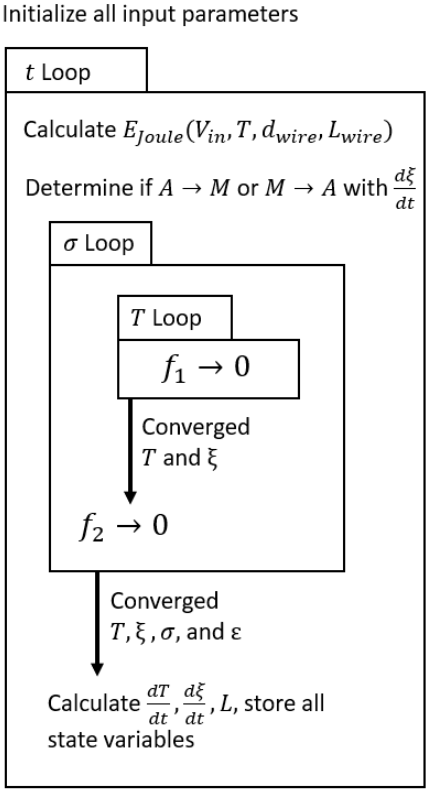

Fig. 5: Diagram of model algorithm

The exit condition for each Newton-Raphson loop involves comparing the current value of f_1 and f_2 to thresholds, which are user-configurable. Because the residual equation f_1 is posed in absolute units of J, the residual value at each timestep gives a measure of how well the system of governing equations is satisfied. We have chosen a residual threshold which keeps the absolute error to around $1e^{-9}J$ for any given timestep.

Several implementation steps have been taken to create a favorable balance between simulation accuracy and execution time. First, a relaxation factor denoted m is applied to the Newton-Raphson steps taken by the inner loop of the solver. This factor was implemented because the residual equation f_1 is less numerically stable when $\frac{\partial\xi}{\partial T}$ is large (e.g. when ξ is around 0.5). Accordingly, the relaxation factor m is calculated at the beginning of each timestep according to the value of $\frac{\partial\xi}{\partial T}$ at the previous timestep:

$$m = 1 - |\frac{\partial\xi}{\partial T}|^{0.1}$$

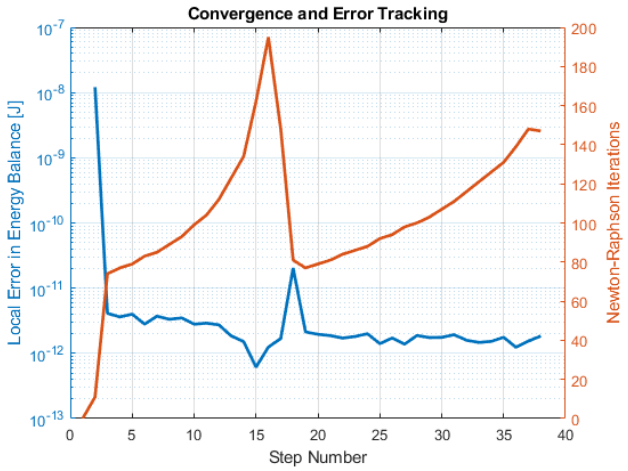

Fig. 6: Absolute convergence error (J) and number of Newton-Raphson iterations executed at each timestep of a simulation of a single actuation cycle

This equation was determined heuristically, and leaves the relaxation factor m very close to 1 during periods of slow transition (increasing NR step size and accelerating convergence) while dropping it during periods of rapid transition (decreasing NR step size and ensuring convergence).

Finally, the size of the timestep at a given iteration is determined adaptively based on the rate that energy is added to the system by Joule heating. This is implemented because at certain values of input parameters (like the inputs used to generate the simulation results in the following section), the time dependent solution can develop up to 1000 times faster during heating than during cooling.

V. Simulation Results

An example of the output of our full system model is given below. For this simulation, representative values for all dimensions and material properties were chosen, and are given in the following tables. Note that phase- or temperature-dependent material properties may be implemented in place of any of the constant values in Table II.

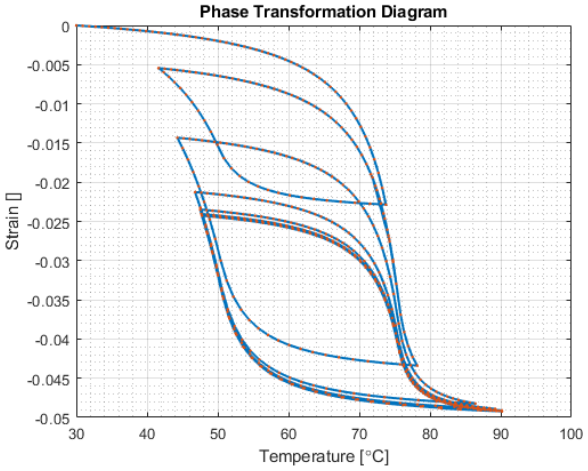

Fig. 7: Wire response to several voltage pulses visualized on a $T-\varepsilon$ phase portrait

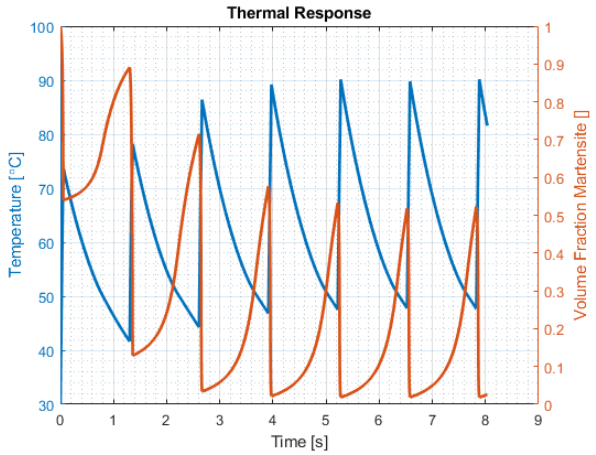

Fig. 8: Wire thermal and mechanical response to several voltage pulses

TABLE II: Shape Memory Alloy Intrinsic Properties

Symbol	Value	Description
k	20	$\frac{W}{mK}$ Thermal conductivity
c	320	$\frac{J}{kgK}$ Specific heat capacity
ρ	6450	$\frac{kh}{m^3}$ Density
h	2400	$\frac{J}{kg}$ Latent Heat of phase transformation
ρ_M	8.2e-7	Ωm Electrical resistivity, M
ρ_A	7.0e-7	Ωm Electrical resistivity, A
ε_{max}	0.055	Maximum recoverable strain
E_M	2.8e10	Pa Young's Modulus, M
E_A	7.5e10	Pa Young's Modulus, A
C_M	8e-6	$\frac{K}{Pa}$ Stress-induced phase shift coefficient, $M \rightarrow A$
C_A	6e-6	$\frac{K}{Pa}$ Stress-induced phase shift coefficient, $A \rightarrow M$

VI. Conclusion

The basics of SMA actuator function are introduced, as are the shortcomings of existing models of phase phase transformation *particularly in regions of interest for miniature mechanical devices*. A novel model of phase transformation is posed, and fit quality to empirical data is examined relative to other models of phase transformation. Finally, implementation of this phase transformation model into a full transient model of actuation, including thermal, electrical, and mechanical effects, is described and sample results are discussed. Full *Matlab* code of this model is available upon request from the author.

TABLE III: Wire Extrinsic Properties, Load Conditions

Symbol	Value	Description
L	$1e$-2	m Wire Length
d	$2e$-4	m Wire Diameter
h	100	$\frac{W}{m^2K}$ Convective coefficient
k_s	100	$\frac{N}{m}$ Spring rate for at non-fixed boundary condition

TABLE IV: Simulation Parameters

Symbol	Value	Description
V_{in}	2.8	V Drive Voltage
t_{on}	0.055	s Duration of voltage pulse with voltage switched on
t_{off}	1.250	s Duration of voltage pulse with voltage switched off
t_{stop}	8	s Duration of simulated time

VII. APPENDIX

A. Partial Derivatives of Residual Expressions

$$\frac{\partial f_1}{\partial T} = -hA\Delta t - mc - mL\frac{\partial \xi}{\partial T}$$

$$\frac{\partial f_2}{\partial \sigma} = 1 - \frac{\partial E}{\partial \sigma}\varepsilon - \frac{\partial \varepsilon}{\partial \sigma}E + \frac{\partial E}{\partial \sigma}\varepsilon_0 + \varepsilon_{max}\left(\frac{\partial E}{\partial \sigma}(\xi - \xi_0) + \frac{\partial \xi}{\partial \sigma}E\right)$$

B. Partial Derivatives of Transformation Equations, Temperature

$$\frac{\partial \xi_{M\to A_1}}{\partial T} = \frac{-\xi_0 a_{A1}}{(T - \sigma/C_A - b_A - c_A)^2}$$

$$\frac{\partial \xi_{M\to A_2}}{\partial T} = \frac{-\xi_0 a_{A2}}{(T - \sigma/C_A - b_A + c_A)^2}$$

$$\frac{\partial \xi_{A\to M_1}}{\partial T} = \frac{-(1-\xi_0) a_{M1}}{(T - \sigma/C_M - b_M - c_M)^2}$$

$$\frac{\partial \xi_{A\to M_2}}{\partial T} = \frac{-(1-\xi_0) a_{M2}}{(T - \sigma/C_M - b_M + c_M)^2}$$

C. Partial Derivatives of Transformation Equations, Stress

$$\frac{\partial \xi_{M\to A_1}}{\partial T} = \frac{\xi_0 a_{A1}}{C_A(T - \sigma/C_A - b_A - c_A)^2}$$

$$\frac{\partial \xi_{M\to A_2}}{\partial T} = \frac{\xi_0 a_{A2}}{C_A(T - \sigma/C_A - b_A + c_A)^2}$$

$$\frac{\partial \xi_{A\to M_1}}{\partial T} = \frac{(1-\xi_0) a_{M1}}{C_M(T - \sigma/C_M - b_M - c_M)^2}$$

$$\frac{\partial \xi_{A\to M_2}}{\partial T} = \frac{(1-\xi_0) a_{M2}}{C_M(T - \sigma/C_M - b_M + c_M)^2}$$

REFERENCES

[1] Cheikh Cisse, Wael Zaki, and Tarak Ben Zineb. "A review of constitutive models and modeling techniques for shape memory alloys". In: *International Journal of Plasticity* 76 (2016), pp. 244–284. ISSN: 07496419. DOI: 10.1016/j.ijplas.2015.08.006. URL: http://dx.doi.org/10.1016/j.ijplas.2015.08.006.

[2] Brinson L.C. "One-Dimensional Constitutive Behavior of shape memory alloys: Thermomechanical Derivation with Non-constant Material Functions and Redefined Martensite Internal Variable". In: *Journal of Intelligent Material Systems and Structures* 4.April (1993), pp. 229–242.

[3] Sonia Marfia and Andrea Vigliotti. *1D SMA Models*. 2014, pp. 99–140. ISBN: 9780080999210. DOI: 10.1016/B978-0-08-099920-3.00005-X.

[4] Edwin A Peraza Hernandez, Darren J Hartl, and Richard J Malak Jr. "DESIGN AND OPTIMIZATION OF AN SMA-BASED SELF-FOLDING STRUCTURAL SHEET WITH SPARSE INSULATING LAYERS". In: *ASME Conference on Smart Materials, Adaptive Structures and Intelligent Systems*. 2014, pp. 1–12.

[5] L Iannucci, P Robinson, and W L H Wan A Hamid. "The Development of a User Defined Material Model for NiTi SMA Wires". In: (2017).