

A reduced-order thermomechanical model and analytical solution for uniaxial shape memory alloy wire actuators

John A Shaw¹ and Christopher B Churchill

Department of Aerospace Engineering, The University of Michigan, 1320 Beal Avenue, Ann Arbor, MI 48109-2140, USA

E-mail: jashaw@umich.edu

Received 17 September 2008, in final form 29 January 2009

Published 8 April 2009

Online at stacks.iop.org/SMS/18/065001

Abstract

A lumped shape memory alloy (SMA) model is derived from the thermodynamic model of Chang *et al* (2006 *Contin. Mech. Thermodyn.* **18** 83–118), using a set of simplifying assumptions, that reduces the system of partial differential equations for an SMA/bias spring actuator to a nonlinear, first-order, ordinary differential equation. Dimensionless state variables and parameters are defined that are useful for characterizing the actuator system and for studying its performance and scaling. A general analytical solution to the nonlinear differential equation governing phase transformation is found in terms of the *Lambert* function for a piecewise constant Joule heating input and a constant temperature convective environment. The analytical solution provides a useful and convenient tool for assessing the time-dependent, hysteretic response of this simple class of SMA actuators, with which design and optimization studies are performed.

(Some figures in this article are in colour only in the electronic version)

Nomenclature

Independent variables:

t	time
τ	dimensionless time
x	reference axial coordinate

SMA phases:

A	Austenite
M	Martensite
M^+, M^-	tensile- M , compressive- M
M^+/M^-	thermal- M

SMA state variables:

ε	nominal axial strain
T	absolute wire temperature
ξ_1, ξ_2, ξ_3	mass fractions M^+, M^-, A

SMA thermodynamic variables:

ϕ	specific Helmholtz free energy
σ	nominal axial stress

s	specific entropy
μ, μ_1	thermo-driving force $A \rightarrow M^+$
μ_2	thermo-driving force $A \rightarrow M^-$
q_x	axial heat flux
A_s, A_f	A start, finish temperatures
M_s, M_f	M^+ start, finish temperatures
<i>SMA physical, mechanical, and electrical constants:</i>	
L, d	reference wire length, diameter
A	reference cross-sectional area
ρ	mass density
E	elastic modulus
β	transformation strain
c_1	phase interaction parameter
ρ_e	electrical resistivity
<i>SMA thermodynamic constants:</i>	
T_R	reference transformation temperature
Δs	equilibrium $A \rightarrow M$ entropy jump
H	enthalpy
c_0	specific heat
u_0, μ_c	kinetic law stiffness, critical driving force

¹ Author to whom any correspondence should be addressed.

σ_M	$M^+/M^- \rightarrow M^+$ reorientation stress
K	reference thermal conductivity

Spring constants:

k_B, k_E	bias, external spring stiffnesses
Δ	bias spring-SMA length mismatch

Actuator, control and environmental variables:

F	SMA axial force
$\delta_0, \varepsilon_0, \sigma_0$	SMA pre-elong., pre-strain, pre-stress
F_B, F_E	bias, external spring forces
δ_B	bias spring displacement
T_a, h	ambient temperature, film coefficient
P_e	electrical power input
t^*	characteristic time
$\mathcal{E}_{in}, \mathcal{E}_{out}$	energy input, external work output

Dimensionless actuator variables:

θ	SMA temperature
θ_a	ambient temperature
θ^{ss}	SMA steady state temperature
$\bar{\sigma}$	SMA axial stress
f	SMA $A \rightarrow M^+$ driving force
\bar{A}_s, \bar{A}_f	A start, finish temperatures
\bar{M}_s, \bar{M}_f	M^+ start, finish temperatures
\bar{P}_e	electrical power
a_n, b_n	time interval constants during transformation
W	Lambert function
$\bar{\mathcal{E}}_{in}, \bar{\mathcal{E}}_{out}$	energy input, external work output

Dimensionless actuator constants:

λ, \bar{c}	SMA latent heat, specific heat
$\bar{\mu}_c$	$A \rightarrow M^+$ critical driving force
$\bar{\sigma}_M$	$M^+/M^- \rightarrow M^+$ reorientation stress
$\bar{\sigma}_0$	SMA pre-stress
$\bar{\sigma}_{max}$	maximum SMA stress
η_B, η_E	bias, external spring stiffnesses
η	spring ratio
$\bar{\Delta}$	bias spring mismatch
δ_E	external spring stroke

1. Introduction

Since their discovery over 50 years ago, there has existed a longstanding interest by engineers to use shape memory alloys (SMAs), such as Nitinol, in various thermomechanical actuator applications. The underlying, reversible, martensitic phase transformations lead to the shape memory effect and superelasticity. Due to their extremely high energy density, NiTi SMAs can be exploited to generate large amounts of mechanical work in light-weight, compact actuators. It is fair to say, however, the design and analysis of actuators with SMA elements has been rather challenging due to the complexity of the thermomechanical behavior of SMAs. They exhibit extremely nonlinear, hysteretic, stress-strain-temperature behavior, which is tightly coupled to ambient heat transfer conditions and imposed boundary conditions. Furthermore, transformation in NiTi SMAs may exhibit instabilities and consequent localization and propagation phenomena in their time-dependent strain and temperature

fields [2, 3]. Fortunately, continuing research in recent years has matured the understanding of SMA behavior and the material is more readily available commercially, improving the potential for developing novel SMA actuators.

The literature is replete with SMA constitutive models of varying complexity (see [1, 4–6] for a few recent examples). The ongoing trend has been to achieve more accuracy and rigor with increasingly complex models. At the same time, the solution of even simple boundary value problems is not a straightforward matter, requiring the solution of coupled, partial differential equations (PDEs) governing equilibrium, heat transfer, and phase transformation kinetics. Usually this must be accomplished numerically to capture the detailed time-dependent nonuniform distributions of the state variables.

For design and optimization studies, a tractable setting is desirable, so our more modest aim in this paper is to demonstrate a simple model of a prototype SMA wire/bias spring actuator that admits analytical solutions for the time-dependent behavior. One of the few examples in the literature coming closest to a closed-form SMA actuator solution was provided in Wu *et al* [7], but unlike our current case, the SMA constitutive model used resulted in a linear PDE that was solved by separation of variables in terms of an infinite series solution in a somewhat different boundary value problem context.

Our paper is organized as follows. Section 2 defines the thermomechanical problem of interest for our prototype actuator system and shows how the full 1D constitutive model of Chang *et al* [1] can be reduced to a ‘lumped’ model, having a minimum set of degrees of freedom, by neglecting the details of nonuniform strain and temperature fields. The resulting model has similarities to many others, including Brinson [8], but with different kinetics. Under our simplifying assumptions, the set of governing PDEs is reduced to a single nonlinear, ordinary differential equation (ODE). We define dimensionless parameters and state variables in order to study the scaling of the solution over a large range of pre-strain, spring properties, environmental conditions, and power inputs. Section 3 shows how the governing ODE during phase transformation can be integrated exactly in terms of a known special function, the *Lambert* function. We believe this is the first such analytical solution for a thermo-mechanically coupled SMA/bias spring model. Section 4 provides a numerical example of the time-dependent behavior of a prototype actuator and a discussion of sensitivities of various parameters. Section 5 discusses a systematic approach to the design and optimization of such an actuator, considering the sizing of springs, actuation stroke, actuation times, minimum power requirements, and energy usage and efficiency. Lastly, summary and conclusions are provided in section 6.

2. Thermomechanical problem of an SMA actuator

As defined in Chang *et al* [1] the thermomechanical problem of interest involves a thin SMA wire of initial length, L , and diameter, d , under uniaxial tensile loading (see figure 1). The reference configuration chosen for the SMA wire is

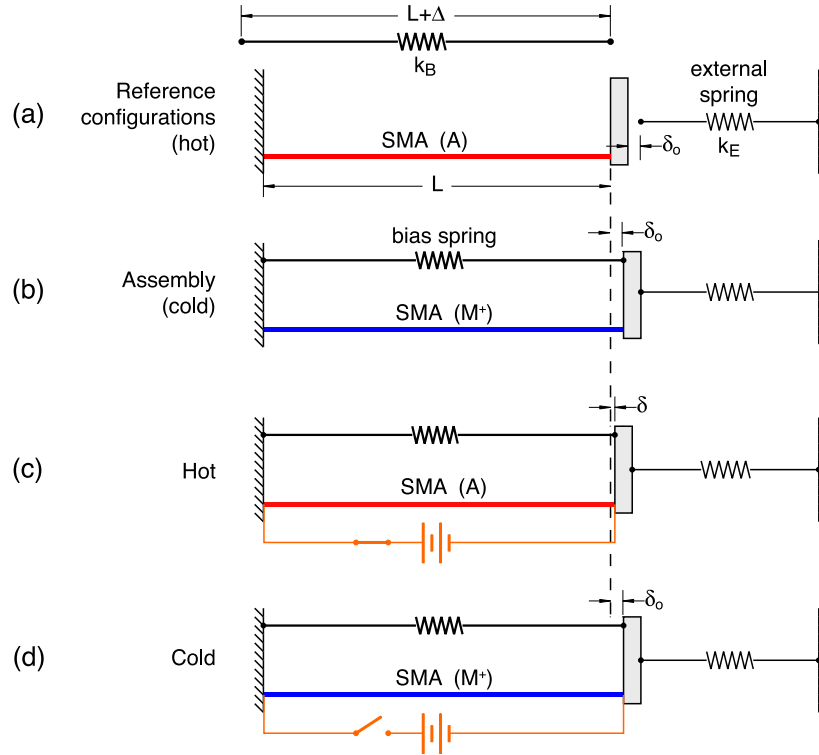


Figure 1. Prototype uniaxial SMA/bias spring actuator working against an external spring as electrical power is applied to the SMA element: (a) reference configurations for SMA element, bias spring, and external spring, (b) after cooling and assembly, (c) contraction of SMA element during Joule heating, (d) reset configuration when cold.

unstressed austenite at a sufficiently high temperature (above the stress-free austenite finish temperature, A_f). We define a reference transformation temperature for the material, T_R , at a lower temperature where martensite and austenite are in thermodynamic equilibrium (near the average of the stress-free martensite start temperature, M_s , and austenite start temperature, A_s). The SMA wire is immersed in a thermal bath (gas or liquid) with convective film coefficient, h , at an ambient temperature, T_a , that is sufficiently below T_R to cause the initial stress-free state of the SMA to be thermal martensite. The left end of the wire in figure 1 is held fixed, and a bias spring of stiffness k_B and initial length mismatch to the SMA element of Δ (figure 1(a)) is force-fit into place. The SMA pre-stress, in mechanical equilibrium with the bias spring force, results in an elongation of the SMA by an amount δ_0 that is assumed to produce tensile martensite (a check for validity of this assumption is provided later). After assembly of the SMA/bias spring actuator, an external spring of stiffness k_E , representing the stiffness of whatever external object is to be actuated against, is attached in its natural (force-free) configuration (figure 1(b)). Electrical power, P_e , can then be applied to resistively heat the SMA wire, causing it to revert to austenite and to contract against the action of both springs to some displacement δ (figure 1(c)). Subsequently, switching off the electrical power causes the SMA temperature to return to ambient temperature and transform back to tensile martensite (figure 1(d)).

The following subsections present (1) a review of the full uniaxial thermodynamic model of Chang *et al* [1], (2) some

simplifying assumptions, and (3) the resulting lumped form of the governing equations.

2.1. The full 1D constitutive model

The constitutive model of Chang *et al* [1] is the starting point for the derivation of the reduced-order, lumped SMA model, so it is briefly reviewed here for completeness. For clarity, we have already omitted strain gradient effects, which are ignored in our current context. The state of each point in the SMA wire is determined by the strain field, $\varepsilon(x, t)$, the temperature field, $T(x, t)$, and the internal phase field vector, $\xi(x, t) \equiv \{\xi_1, \xi_2\}$, for tensile (M^+) and compressive (M^-) variants of martensite, respectively. Constitutive relations were derived from the specific Helmholtz free energy,

$$\begin{aligned} \phi(\varepsilon, T, \xi) = & \frac{E_A + (\xi_1 + \xi_2)\Delta E}{2\rho} [\varepsilon - (\xi_1 - \xi_2)\beta]^2 \\ & - (T - T_R)(\xi_1 + \xi_2)\Delta s \\ & + c_1(1 - \xi_1 - \xi_2)(\xi_1 + \xi_2) \\ & + (c_0 - s_0)(T - T_R) - c_0 T \ln(T/T_R), \end{aligned} \quad (2.1)$$

where the separate lines on the right-hand side are respective contributions from elastic energy, phase-dependent entropy, energy of mixing, and phase-independent thermal energy. The mass density is ρ , the elastic modulus of pure austenite is E_A , the difference between the effective martensite and austenite elastic moduli is $\Delta E \equiv E_M - E_A$ (usually a negative constant), and the stress-free transformation strain is β (a positive material constant). Thermal expansion is neglected

in the elastic energy, since its effect is usually small compared to the transformation strain. The specific entropy change from austenite to martensite is $\Delta s \equiv s_M - s_A$ (a negative constant), which is related to the latent heat of transformation. The free energy of mixing has a material constant, c_1 , that affects the slope of the pseudoelastic transformation path. Lastly, s_0 and c_0 are material constants representing the specific entropy and specific heat, respectively.

The non-negativity requirement of the local entropy production leads to the Gibb's relations

$$s = -\frac{\partial \phi}{\partial T} = s_0 + (\xi_1 + \xi_2) \Delta s + c_0 \ln(T/T_R), \quad (2.2)$$

$$\sigma = \rho \frac{\partial \phi}{\partial \varepsilon} = [E_A + (\xi_1 + \xi_2) \Delta E] [\varepsilon - (\xi_1 - \xi_2) \beta],$$

where σ is the nominal stress. Entropy production arises from the heat flux and phase transformation terms, and sufficient conditions to satisfy the second law of thermodynamics are

$$-q_x \frac{\partial T}{\partial x} \geq 0, \quad (2.3)$$

$$\boldsymbol{\mu} \cdot \frac{\partial \boldsymbol{\xi}}{\partial t} \geq 0, \quad (2.4)$$

where q_x is the axial heat flux, and $\boldsymbol{\mu} \equiv -\phi_{,\boldsymbol{\xi}}$ is the thermodynamic driving force for phase transformation conjugate to the rate of change of $\boldsymbol{\xi} \equiv (\xi_1, \xi_2)$. The first inequality in (2.4) is easily satisfied by assuming the conventional axial heat conduction law $q_x = -K \partial T / \partial x$, where K is a positive thermal conductivity for the material. Written explicitly, the thermodynamic driving forces are

$$\mu_1 = -\frac{\partial \phi}{\partial \xi_1} = \frac{\beta \sigma}{\rho} - \frac{\Delta E}{2\rho} \left[\frac{\sigma}{E_A + (\xi_1 + \xi_2) \Delta E} \right]^2 + (T - T_R) \Delta s - c_1 (1 - 2\xi_1 - 2\xi_2), \quad (2.5)$$

$$\mu_2 = -\frac{\partial \phi}{\partial \xi_2} = -\frac{\beta \sigma}{\rho} - \frac{\Delta E}{2\rho} \left[\frac{\sigma}{E_A + (\xi_1 + \xi_2) \Delta E} \right]^2 + (T - T_R) \Delta s - c_1 (1 - 2\xi_1 - 2\xi_2). \quad (2.6)$$

The second inequality in (2.4) is satisfied if the 'projection' of the driving force to its conjugate phase fraction rate is non-negative. A simple, piecewise linear, kinetic relation is chosen, according to

$$\frac{\partial \boldsymbol{\xi}}{\partial t} = v_0 \langle \boldsymbol{\mu} \cdot \mathbf{m} - \mu_c \rangle \mathbf{m} \quad (2.7)$$

where v_0 is the 'stiffness' coefficient of the kinetic law, $\langle x \rangle$ is the Macaulay discontinuity function ($\langle x \rangle = x$ for $x > 0$ and $\langle x \rangle = 0$ for $x \leq 0$), and \mathbf{m} is the unit vector in the direction of phase transformation in (ξ_1, ξ_2) space. The parameter, μ_c , is a positive constant that represents the critical thermodynamic driving force necessary for hysteretic phase transformation. Within the phase fraction space described above, the direction of phase transformation is assumed to be a unit vector collinear with the thermodynamic driving force ($\mathbf{m} = \boldsymbol{\mu} / \sqrt{\boldsymbol{\mu} \cdot \boldsymbol{\mu}}$), except at a boundary of the admissible phase region ($\xi_1 = 0$, $\xi_2 = 0$, or $\xi_1 + \xi_2 = 1$) where it is a unit vector tangent to the boundary to keep the phase fractions within the admissible region (see figure 2).

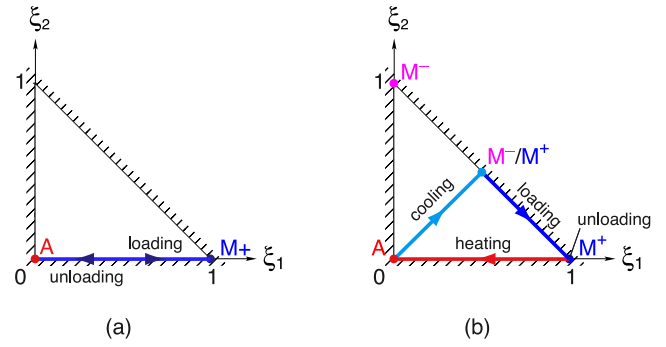


Figure 2. Admissible phase fraction space: (a) superelastic path in tension, (b) shape memory effect path.

Assuming traction-free lateral boundary conditions and no body forces, mechanical equilibrium requires that the axial force be constant along the length, or

$$\frac{\partial F}{\partial x} = 0, \quad (2.8)$$

where the axial tensile force is simply $F \equiv \sigma A$. The heat equation, accounting for sensible and latent heat, axial heat conduction, lateral convection, and Joule heating, is

$$\rho A \left[c_0 \frac{\partial T}{\partial t} - (\boldsymbol{\mu} - T \boldsymbol{\mu}_{,T}) \cdot \frac{\partial \boldsymbol{\xi}}{\partial t} \right] = \frac{\partial}{\partial x} \left(K A \frac{\partial T}{\partial x} \right) - h \pi d (T - T_a) + \frac{P_e}{L}, \quad (2.9)$$

where h is the Fourier law coefficient (or convective film coefficient) for lateral heat transfer from the wire. The governing equations for the thermomechanical boundary value problem are the field equations (2.8) and (2.9), subject to the kinetic law (2.7) at each point $x \in (0, L)$, and the appropriate mechanical and thermal boundary conditions at $x = 0, L$.

2.2. Simplified model

In addition to the assumptions already inherent in Chang *et al* [1] (no inertia effects, no R-phase, no thermal expansion, and no plasticity or shakedown effects), the following assumptions are made to simplify and reduce the system of partial differential equations to ordinary differential equations.

- (1) Spatial gradients of strain (ε), temperature (T), and phase fraction ($\boldsymbol{\xi}$) are neglected. This means that transformation is assumed to be uniform (no transformation fronts or boundary effects), and axial heat conduction is ignored. It is a reasonable assumption for wire with constant cross-sectional area (A) that has already been conditioned (trained) to have repeatable cyclic behavior, and for SMA wire that is thermally insulated at its ends. This is verified in section 4.2 by comparison of the simplified model to a finite element simulation.
- (2) The SMA element is always under sufficient tension (after assembly) to avoid thermal martensite. In other words, $\mu_2 < 0$, is sufficiently negative to cause $\xi_2 = 0$ for all time $t > 0$. This assumption is valid if the bias spring constant and pre-strain are sufficiently large. A condition will be given later to check the validity of this assumption.

- (3) The elastic moduli of austenite and martensite are the same ($\Delta E = 0$). This assumption poses no significant restriction if unloading of the martensite are avoided during assembly and operation.
- (4) The mixing energy between austenite and martensite is neglected ($c_1 = 0$). This assumption neglects any hardening or softening during superelastic transformation, which may be reasonable, or not, depending on the desired accuracy of the simulation.
- (5) The characteristic speed of phase transformation is fast compared to heat transfer rate and mechanical loading rate (i.e., we take the limiting case $v_0 \rightarrow \infty$). This assumption is reasonable for slow to moderate loading rates where displacement rates are small compared to the inherent velocity of martensitic transformations. Roughly speaking, inherent martensitic transformation propagation velocities are only an order of magnitude slower than elastic wave velocities, as measured in high rate impact experiments [9]. For Nitinol this should not be a serious restriction for moderate displacement rates, say $\dot{\delta} < 10^2 \text{ m s}^{-1}$.
- (6) The electrical power P_e and ambient temperature T_a are piecewise constant functions of time.

The state of the wire is now minimally defined by three time-dependent variables, the strain $\varepsilon(t)$, the temperature $T(t)$, and the single phase fraction $\xi(t)$ for tensile martensite ($\xi \in [0, 1]$). Under these assumptions, the specific Helmholtz free energy of the SMA material reduces to

$$\phi(\varepsilon, T, \xi) = \frac{E}{2\rho} [\varepsilon - \xi\beta]^2 - (T - T_R)\xi\Delta s + (c_0 - s_0)(T - T_R) - c_0 T \ln(T/T_R), \quad (2.10)$$

and the following simplified constitutive relations for nominal stress (σ), entropy (s), and transformation driving force ($\mu = \mu_1$) are obtained.

$$\sigma = E[\varepsilon - \xi\beta] \quad (2.11)$$

$$s = s_0 + \xi\Delta s + c_0 \ln(T/T_R) \quad (2.12)$$

$$\mu = \frac{\beta\sigma}{\rho} + (T - T_R)\Delta s. \quad (2.13)$$

Figure 3(a) shows isothermal constitutive responses in stress–strain–temperature space predicted by the simplified model using typical SMA properties of table 1, which were calibrated to the room temperature superelastic wire used in Chang *et al* [1] (guideBB-30 from Memry Corp.). This simplified model has constant stress transformation paths (which could also have been derived from a simple Gibb's mixture rule of the phases.) Figure 3(b) shows a quasi-phase diagram in stress–temperature space, and figure 3(c) shows a stress-free shape memory cycle along with an isothermal superelastic response. In fact, figure 3(c) shows a two-way shape memory effect, since the simplified two-phase (A and M^+) model cannot capture a zero strain martensite that would be necessary to capture a one-way shape memory effect. The behavior of the simplified model, consequently, is more representative of a highly conditioned SMA that has two-way shape memory. If the actual SMA being used has a one-way shape memory, the model should not be used at certain

low stress/low temperature regimes. A suitable check will be provided later in this paper to ensure the model's validity in the latter case.

2.3. Simplified governing equations

For this simplified model, the governing equations for the actuator's operation are algebraic equations for equilibrium and transformation kinetics and an ordinary differential equation in time for heat transfer.

Neglecting inertial effects, equilibrium requires the axial force in the SMA wire to be uniform. It is balanced by the forces in the bias spring and external spring

$$F(t) = F_B(t) + F_E(t). \quad (2.14)$$

Accounting for the natural configurations of the springs and the SMA constitutive equation for stress (2.11) leads to the equilibrium equation for $t \geq 0$ (after assembly),

$$EA[\varepsilon - \beta\xi] = k_B[\Delta - L\varepsilon] + k_E L[\varepsilon_0 - \varepsilon], \quad (2.15)$$

where $\varepsilon_0 \equiv \delta_0/L$ is the pre-strain in the SMA wire caused during cold assembly. Equation (2.15) is an equation relating the unknowns $\varepsilon(t)$ and $\xi(t)$, where the explicit time dependence has been omitted for brevity. At $t = 0$, just after assembly, the external spring is not yet loaded and the initial condition is defined by

$$EA[\varepsilon_0 - \xi_0\beta] = k_B[\Delta - \varepsilon_0 L], \quad (2.16)$$

from which the pre-strain, ε_0 , can be found.

From an analysis of martensite reorientation, using the full model [1], the stress required for reorientation from thermal martensite to tensile martensite is $\sigma_M = \rho\mu_c/(\beta\sqrt{2})$. Using the initial condition for equilibrium (2.16) leads to the following condition that must be satisfied for the bias spring mismatch and stiffness to cause full reorientation of martensite ($\xi_0 = 1$),

$$\frac{\Delta}{L} \geq \beta + \frac{\sigma_M}{E} \left[1 + \frac{EA}{k_B L} \right]. \quad (2.17)$$

Considering now kinetics, forward transformation ($A \rightarrow M^+$) occurs when the thermodynamic driving force, μ , reaches the critical value, μ_c , and the amount of tensile martensite has not yet reached saturation ($0 \leq \xi < 1$). Conversely, reverse transformation ($M^+ \rightarrow A$) occurs when the thermodynamic driving force reaches the critical value, $-\mu_c$, and the amount of austenite has not yet reached saturation ($0 < \xi \leq 1$). Otherwise, no phase transformation occurs ($\dot{\xi} = 0$, where a dot above a variable denotes a derivative with respect to time, t), and the kinetics equation is not active. During phase transformation, however, the SMA stress and temperature are coupled through one of the following conditions

$$\begin{aligned} & \frac{\beta\sigma}{\rho} + (T - T_R)\Delta s \\ & = \begin{cases} \mu_c & \text{if } \xi \in [0, 1) \quad \text{and} \quad \dot{\xi} > 0 (A \rightarrow M^+), \\ -\mu_c & \text{if } \xi \in (0, 1] \quad \text{and} \quad \dot{\xi} < 0 (M^+ \rightarrow A). \end{cases} \end{aligned} \quad (2.18)$$

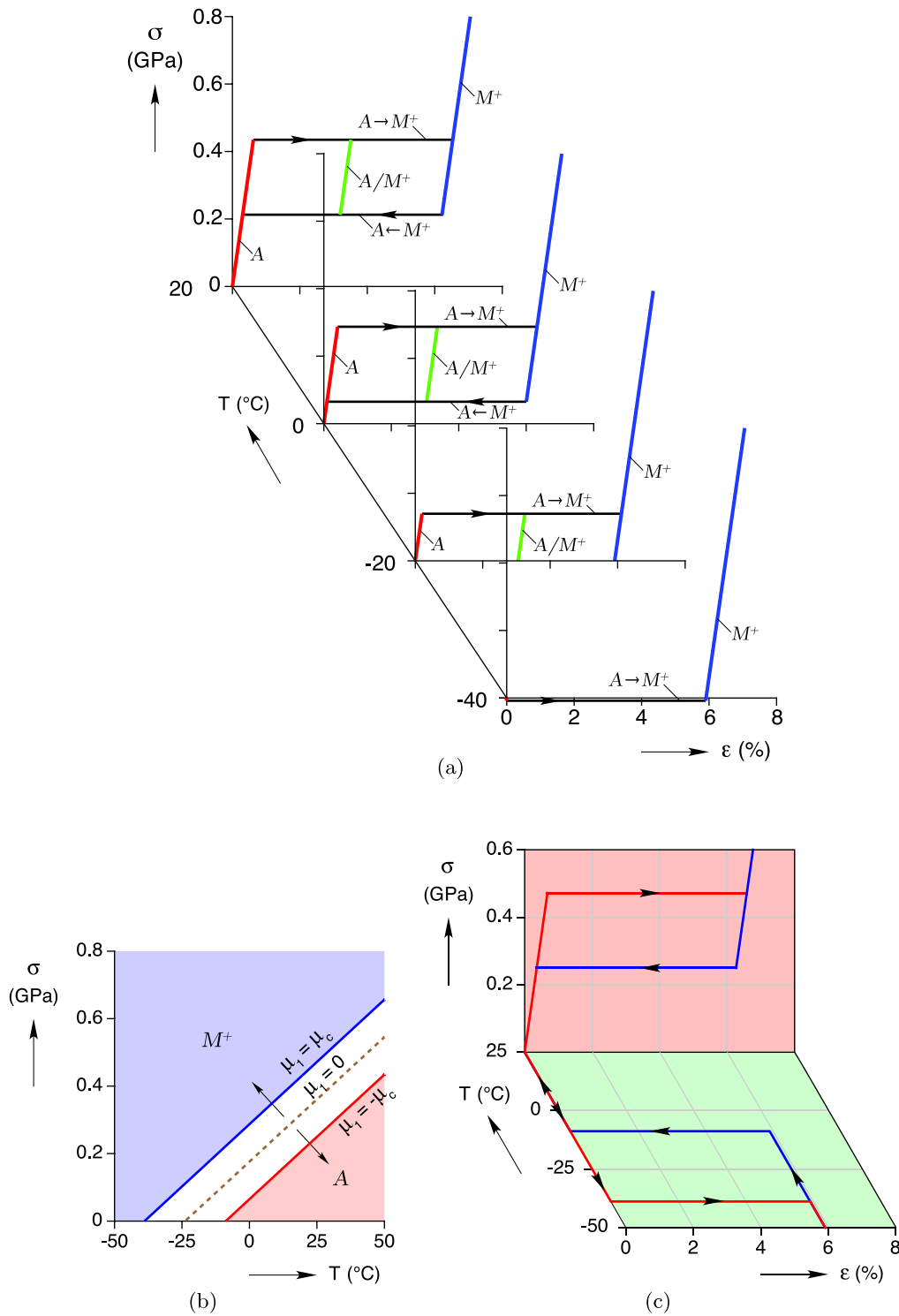


Figure 3. The simplified constitutive model: (a) isothermal responses, (b) transformation stresses, (c) shape memory cycle (ϵ - T plane) and superelastic response (σ - ϵ plane).

It is apparent from these equations that forward and reverse isothermal superelastic transformations are predicted to occur along stress plateaus with stress hysteresis $2\rho\mu_c/\beta$ centered on the stress $(T - T_R)(-\rho\Delta s)/\beta$. These equations represent implicit relations between the state variables $\epsilon(t)$, $T(t)$ and $\xi(t)$, which can be rewritten using the constitutive equation for

stress (2.11) as

$$E(\epsilon - \xi\beta) + \frac{\rho}{\beta}(T - T_R)\Delta s = \begin{cases} \frac{\rho\mu_c}{\beta} & \text{if } \xi \in [0, 1) \text{ and } \dot{\xi} > 0, \\ -\frac{\rho\mu_c}{\beta} & \text{if } \xi \in (0, 1] \text{ and } \dot{\xi} < 0. \end{cases} \quad (2.19)$$

Table 1. SMA parameters.

Mechanical	Thermal	Physical, kinetic, electrical
$L = 60 \text{ mm}$	$T_R = 249.3 \text{ K}$	$\rho = 6.5 \times 10^6 \text{ g m}^{-3}$
$d = 0.762 \text{ mm}$	$T_R \Delta s = -16.74 \text{ J g}^{-1}$	$\mu_c = 1.01 \text{ J g}^{-1}$
$\beta = 0.0591$	$c_0 = 0.5 \text{ J g}^{-1} \text{ K}^{-1}$	$\rho_e = 1 \times 10^{-6} \Omega \text{ m}$
$E = 70 \text{ GPa}$		

Finally, the heat equation (2.9) now simplifies to

$$\rho A L c_0 \dot{T} = \rho A L [\text{sgn}(\dot{\xi}) \mu_c - T \Delta s] \dot{\xi} - h \pi d L (T - T_a) + P_e, \quad (2.20)$$

where the left-hand side is the rate of energy associated with sensible heat, the first term on the right-hand side is the latent heat source rate (sgn is the *signum* function), the second term is the heat loss rate to the environment, and P_e is the applied Joule heat rate.

2.4. Dimensionless parameters and equations

To simplify the equations and clarify the scaling of various quantities we define the following dimensionless variables for time, stress, and temperature

$$\tau \equiv \frac{t}{t^*}, \quad (2.21)$$

$$\bar{\sigma} \equiv \frac{\sigma}{E}, \quad (2.22)$$

$$\theta \equiv \left(\frac{T}{T_R} - 1 \right) \lambda, \quad (2.23)$$

where

$$t^* \equiv \frac{\rho c_0 d}{4h} \quad (2.24)$$

is a characteristic time for the rate of sensible heat change, and

$$\lambda \equiv -\frac{\rho \Delta s T_R}{\beta E} \quad (2.25)$$

is a dimensionless (positive) reference latent heat (associated with the reversible case $\mu_c = 0$), which incidentally, is also related to the Clausius–Clapeyron slope of stress-induced transformation. The following dimensionless parameters (constants) are also helpful to define,

$$\theta_a \equiv \left(\frac{T_a}{T_R} - 1 \right) \lambda, \quad (2.26)$$

$$\bar{c} \equiv -\frac{c_0}{\Delta s}, \quad (2.27)$$

$$\bar{\mu}_c \equiv \frac{\rho \mu_c}{\beta E}, \quad (2.28)$$

$$\eta_B \equiv \frac{k_B L}{E A}, \quad (2.29)$$

$$\eta_E \equiv \frac{k_E L}{E A}, \quad (2.30)$$

$$\eta \equiv \frac{\eta_B + \eta_E}{1 + \eta_B + \eta_E}, \quad (2.31)$$

$$\bar{\Delta} \equiv \frac{\Delta}{L}, \quad (2.32)$$

$$\bar{P}_e \equiv \frac{P_e \lambda}{\pi d L T_R h}, \quad (2.33)$$

where θ_a is a dimensionless ambient temperature, \bar{c} is a dimensionless specific heat, $\bar{\mu}_c$ is a dimensionless critical driving force, η_B is a dimensionless bias spring constant, η_E is a dimensionless external spring constant, $\eta \in [0, 1)$ is a dimensionless ratio of spring constants, $\bar{\Delta}$ is a dimensionless bias spring/SMA length mismatch, and \bar{P}_e is a dimensionless electrical power.

Time derivatives in the heat equation (2.20) can be converted to dimensionless time derivatives, denoted by $(\cdot)' \equiv \partial(\cdot)/\partial\tau$, by the chain rule $\partial(\cdot)/\partial t = \partial(\cdot)/\partial\tau/t^*$. Now the respective governing equations for equilibrium, kinetics, and heat can be written in dimensionless form as

$$\varepsilon(\tau) - \varepsilon_0 + \beta(1 - \eta)[1 - \xi(\tau)] = 0, \quad (2.34)$$

$$\varepsilon(\tau) - \beta \xi(\tau) - \theta(\tau) - \text{sgn}(\xi') \bar{\mu}_c = 0, \quad (2.35)$$

$$\bar{c} [\theta(\tau) + \theta'(\tau) - \theta_a - \bar{P}_e] - [\varepsilon(\tau) - \beta \xi(\tau) + \lambda] \xi'(\tau) = 0. \quad (2.36)$$

Equation (2.34) assumes the condition for full initial reorientation of martensite (2.17) is satisfied (discussed further in section 5), rewritten in dimensionless form as

$$\bar{\Delta} \geq \beta + \bar{\sigma}_M \left[1 + \frac{1}{\eta_B} \right], \quad (2.37)$$

where $\bar{\sigma}_M = \bar{\mu}_c/\sqrt{2}$ is the dimensionless stress to orient Martensite ($M^+/M^- \rightarrow M^+$). The respective pre-strain and dimensionless pre-stress in the SMA element (related by $\varepsilon_0 = \beta + \bar{\sigma}_0$) are

$$\varepsilon_0 = \frac{\beta + \eta_B \bar{\Delta}}{1 + \eta_B}, \quad (2.38)$$

$$\bar{\sigma}_0 = \frac{(\bar{\Delta} - \beta) \eta_B}{1 + \eta_B}. \quad (2.39)$$

According to equations (2.34) and (2.35), the dimensionless start and finish temperatures for the actuator are dependent on the pre-stress and spring stiffnesses, given by

$$\bar{A}_s = \bar{\sigma}_0 + \bar{\mu}_c, \quad \bar{A}_f = \bar{\sigma}_0 + \bar{\mu}_c + \beta \eta, \quad (2.40)$$

$$\bar{M}_s = \bar{\sigma}_0 - \bar{\mu}_c + \beta \eta, \quad \bar{M}_f = \bar{\sigma}_0 - \bar{\mu}_c.$$

A dimensionless steady state temperature, based on the current power level and ambient conditions, is defined as

$$\theta^{\text{ss}} \equiv \theta_a + \bar{P}_e, \quad (2.41)$$

which is shown later to be the asymptotic temperature for long times ($\tau \rightarrow \infty$). A comparison between it and the appropriate dimensionless transformation temperature of equation (2.40) can quickly assess whether transformation will start and/or complete. Furthermore, it is convenient to define a normalized, time-dependent, driving force function (for strictly $\mu_c > 0$) as

$$f(\tau) \equiv \frac{\mu(\varepsilon(\tau), \theta(\tau), \xi(\tau))}{\mu_c} = \frac{\varepsilon(\tau) - \beta \xi(\tau) - \theta(\tau)}{\bar{\mu}_c}, \quad (2.42)$$

which can be monitored to determine whether or not phase transformation occurs. This allows equation (2.35), when active, to be written simply as

$$f = \text{sgn}(\xi'). \quad (2.43)$$

3. Analytical solution

In this section analytical solutions to the system of equations (2.34)–(2.36) are derived for $\varepsilon(\tau)$, $\theta(\tau)$, and $\xi(\tau)$ for piecewise constant electrical power input. Assuming that condition (2.37) is satisfied, the initial conditions at $\tau = 0$ are $\varepsilon(0) = \varepsilon_0$, $\theta(0) = \theta_a$, and $\xi(0) = 1$. Presuming sufficient initial pre-stress exists to create M^+ and sufficient electrical power is supplied to completely transform the SMA wire to Austenite, the stress–strain path is uniquely determined by the equilibrium equation (2.34).

3.1. Mechanical equilibrium paths

An example actuator response, based on equilibrium considerations alone, is shown in figure 4. The initial equilibrium point between M^+ (dotted line) and the bias spring is shown by the solid circle. Open circles show the end points on the pure Austenite elastic curve (dotted line). Note that, using the equilibrium equation (2.34) and the constitutive equation (2.11), one can show that the slopes of the respective dimensionless paths are

$$\frac{d\bar{\sigma}}{d\varepsilon} = \begin{cases} -\eta_B, & \text{bias spring only} \\ -(\eta_B + \eta_E), & \text{bias plus external springs.} \end{cases} \quad (3.1)$$

Our immediate concern is to quantify the transient response of the actuator. The character of the time-dependent solution depends on whether or not phase transformation occurs during the time interval, so the general solution can be derived piecewise in time starting from an initial condition for each time interval $\tau_n \leq \tau \leq \tau_{n+1}$, where $n = 0, 1, 2, \dots$. The power input and ambient temperature for each time interval are then denoted as $\bar{P}_{e,n}$ and $\theta_{a,n}$, respectively. Each τ_n for $n = 1, 2, \dots$ is a time instant that defines either the onset or termination of phase transformation or a change in power level or environment. The normalized driving force function, $f(\tau)$ of equation (2.42), is monitored during the evolution to determine whether or not phase transformation occurs and to determine the time, τ_n , for the onset of transformation.

3.2. Solution without phase transformation

If no phase transformation occurs during the interval $\tau_n \leq \tau \leq \tau_{n+1}$, the phase fraction remains constant at its initial value, $\xi(\tau) = \xi(\tau_n) \equiv \xi_n$. By equilibrium (2.34) the strain $\varepsilon(\tau) = \varepsilon(\tau_n) \equiv \varepsilon_n$ also remains constant at its initial value according to

$$\varepsilon_n = \varepsilon_0 - \beta(1 - \eta)(1 - \xi_n). \quad (3.2)$$

Temperature, $\theta(\tau)$, is the only variable that evolves with time. Substituting $\xi'(\tau) = 0$ into (2.36), results in the classical

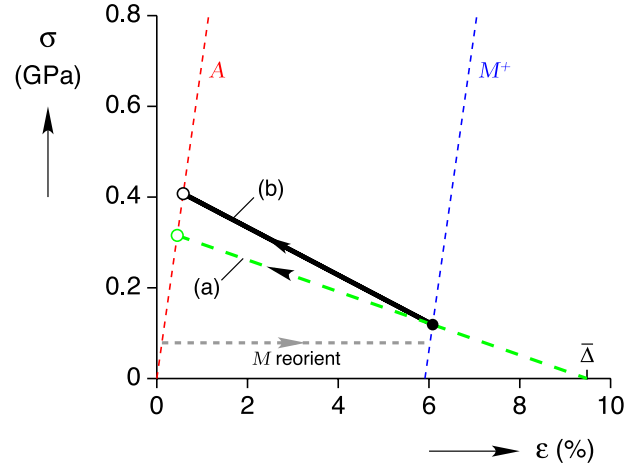


Figure 4. Example actuator response during heating in SMA stress versus strain space: (a) response with bias spring only (dashed line), (b) response with the addition of the external spring (bold line).

sensible heat equation

$$\theta'(\tau) + \theta(\tau) = \theta_{a,n} + \bar{P}_{e,n}, \quad (3.3)$$

which has the general solution

$$\theta(\tau) = \theta_n^{ss} + (\theta_n - \theta_n^{ss})e^{-(\tau - \tau_n)}, \quad \tau \in [\tau_n, \tau_{n+1}] \quad (3.4)$$

where $\theta_n \equiv \theta(\tau_n)$ is the initial temperature for the time interval, and

$$\theta_n^{ss} \equiv \theta_{a,n} + \bar{P}_{e,n} \quad (3.5)$$

is the dimensionless steady state temperature for long times ($\tau \rightarrow \infty$) for the current power and environment.

This solution is valid until phase transformation is detected at time τ_{n+1} , for which

$$\text{Start} \begin{cases} A \rightarrow M^+: & f(\tau_{n+1}) = 1, \\ & \text{and } f'(\tau_{n+1}) > 0, \text{ and } \xi < 1, \\ \text{or} \\ M^+ \rightarrow A: & f(\tau_{n+1}) = -1, \\ & \text{and } f'(\tau_{n+1}) < 0, \text{ and } \xi > 0. \end{cases} \quad (3.6)$$

Since temperature is the only evolving variable, the direction of transformation can be determined equivalently by

$$f'(\tau_{n+1}) > 0 \implies \theta'(\tau_{n+1}) < 0, \quad (3.7)$$

and $f'(\tau_{n+1}) < 0 \implies \theta'(\tau_{n+1}) > 0.$

Defining the starting stress for the interval as $\bar{\sigma}_n = \varepsilon_n - \beta\xi_n$, onset of transformation is detected when

$$\tau_{n+1} = \tau_n + \ln \left(\frac{\theta_n^{ss} - \theta_n}{\theta_n^{ss} - [\bar{\sigma}_n + \text{sgn}(\theta')\bar{\mu}_c]} \right), \quad (3.8)$$

provided, of course, that the logarithm is positive.

3.3. Solution during phase transformation

If, on the other hand, transformation occurs during a time interval $\tau_n \leq \tau \leq \tau_{n+1}$, all three state variables, $\varepsilon(\tau)$, $\theta(\tau)$, and $\xi(\tau)$ evolve from their starting values when phase transformation begins at time τ_n . During the time interval, $\varepsilon(\tau)$ and $\theta(\tau)$ can be solved in terms of $\xi(\tau)$ using equations (2.34) and (2.35), which when substituted into (2.36) gives the following nonlinear, first-order ODE for $\xi(\tau)$.

$$\bar{c} [\beta\eta + \bar{\sigma}_0 - \text{sgn}(\xi')\bar{\mu}_c - \theta_n^{\text{ss}}] - [(1 + \bar{c})\beta\eta + \bar{\sigma}_0 + \lambda]\xi'(\tau) + \beta\eta\xi(\tau) [\xi'(\tau) - \bar{c}] = 0. \quad (3.9)$$

Using the initial condition $\xi(\tau_n) = \xi_n$ for the interval, the solution is found to be

$$\xi(\tau) = \xi_n + \frac{1}{b_n} [a_n - W(a_n e^{[a_n - b_n \bar{c}(\tau - \tau_n)]})], \quad (3.10)$$

$$\tau \in [\tau_n, \tau_{n+1}]$$

where the following dimensionless constants are defined for the time interval

$$a_n \equiv \frac{\beta\eta(1 - \xi_n) + \bar{\sigma}_0 - [\theta_n^{\text{ss}} + \text{sgn}(\xi')\bar{\mu}_c]}{\beta\eta\bar{c} + \lambda + \theta_n^{\text{ss}} + \text{sgn}(\xi')\bar{\mu}_c}, \quad (3.11)$$

$$b_n \equiv \frac{\beta\eta}{\beta\eta\bar{c} + \lambda + \theta_n^{\text{ss}} + \text{sgn}(\xi')\bar{\mu}_c},$$

and where $W(z)$ is the Lambert function [10], also known as the *Product-Logarithm*. This is a special function defined as the implicit solution to $We^W = z$. It is available as a built-in function in symbolic manipulation software, such as Mathematica[®], or can be calculated iteratively as shown in Boyd [11]. The Lambert function $W(z)$ is real valued for $z \geq -e^{-1}$, below which a branch cut exists. It has unique real values for $z \geq 0$ and has two real values for $-e^{-1} < z < 0$ as shown in figure 5. In our case, only the principle branch, $W(z) > -1$, is relevant, so W can be alternately defined by $W(y e^y) = y$ for $y \geq -1$. From this definition, we see that

$$\lim_{\tau \rightarrow \tau_n} W(a_n e^{[a_n - b_n \bar{c}(\tau - \tau_n)]}) = a_n, \quad (3.12)$$

making it apparent that the initial condition is recovered in equation (3.10). While equation (3.9) is nonlinear, it is only first order, and the solution (3.10) was discovered with the help of Mathematica[®] once suitable simplifying constants were defined by equation (3.11). The reader can readily verify the solution by substituting equation (3.10) into (3.9).

The remaining state variables follow from the solution above. Using equations (3.10), (2.34), and (2.35), it can be shown that

$$\theta(\tau) = \theta_n - \frac{\beta\eta}{b_n} [a_n - W(a_n e^{[a_n - b_n \bar{c}(\tau - \tau_n)]})], \quad (3.13)$$

$$\varepsilon(\tau) = \varepsilon_n + \frac{\beta(1 - \eta)}{b_n} [a_n - W(a_n e^{[a_n - b_n \bar{c}(\tau - \tau_n)]})], \quad (3.14)$$

where the initial values are uniquely identified with ξ_n by

$$\theta_n = \beta\eta(1 - \xi_n) + \bar{\sigma}_0 - \text{sgn}(\xi')\bar{\mu}_c \quad (3.15)$$

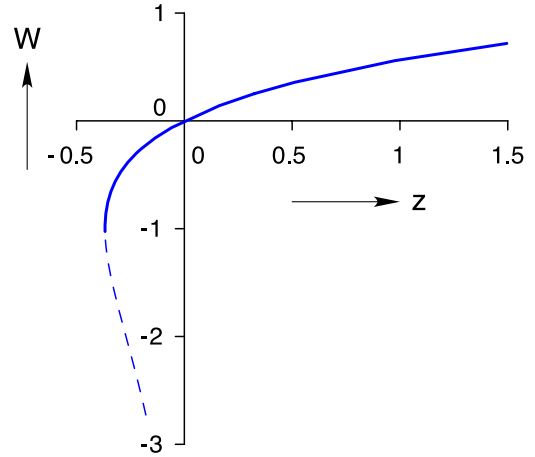


Figure 5. The real-valued part of the Lambert $W(z)$ function.

$$\varepsilon_n = \beta[\eta + (1 - \eta)\xi_n] + \bar{\sigma}_0. \quad (3.16)$$

From (2.11) the evolution of the dimensionless stress is then

$$\bar{\sigma}(\tau) = \sigma_n - \frac{\beta\eta}{b_n} [a_n - W(a_n e^{[a_n - b_n \bar{c}(\tau - \tau_n)]})]. \quad (3.17)$$

Phase transformation terminates when either the driving force is no longer sufficient to sustain the transformation or when the phase fraction saturates, i.e.,

$$\text{End} \begin{cases} A \rightarrow M^+ : & f(\tau_{n+1}) = 1, \\ & \text{and } (f'(\tau_{n+1}) < 0 \text{ or } \xi = 1), \\ \text{or} \\ A \leftarrow M^+ : & f(\tau_{n+1}) = -1, \\ & \text{and } (f'(\tau_{n+1}) > 0 \text{ or } \xi = 0). \end{cases} \quad (3.18)$$

Alternatively, the termination condition can be found as follows. It can be shown according to equation (3.13), that the long time, asymptotic temperature is still θ_n^{ss} as defined in equation (3.5). Note that the temperature evolution is always monotonic within each time interval. If θ_n^{ss} does not exceed \bar{A}_f during heating, or drop below \bar{M}_f during cooling, phase transformation does not reach saturation and the given solutions are valid until the next power change or ambient condition change. In this case, the steady state value of the phase fraction for long times, $\xi^{\text{ss}} \equiv \xi(\infty)$, is found from equation (3.9) by setting $\xi'(\infty) = 0$ and solving as

$$\xi^{\text{ss}} = \frac{\bar{\sigma}_0 - \text{sgn}(\xi')\bar{\mu}_c - \theta_n^{\text{ss}}}{\beta\eta}. \quad (3.19)$$

If, conversely, one of the threshold temperatures is predicted to be exceeded by θ_n^{ss} the transformation saturates at one of following final phase fraction values

$$\xi_f = \begin{cases} 0, & \xi_n' < 0, \\ 1, & \xi_n' > 0. \end{cases} \quad (3.20)$$

In this case, the time interval is to be subdivided, and a new time (τ_{n+1}) is calculated for the termination of transformation.

Table 2. Spring parameters, environment constants, and electrical power.

Springs	Ambient	Power
$k_B = 26.60 \text{ N mm}^{-1}$	$T_a = 223.15 \text{ K}$	$P_e = 1.301 \text{ W}$
$k_E = 13.30 \text{ N mm}^{-1}$	$h = 90 \text{ W m}^{-2} \text{ K}^{-1}$	
$\Delta = 5.688 \text{ mm}$		

The time (τ_{n+1}) at which this happens is found by solving equation (3.10) for $\xi(\tau_{n+1}) = \xi_f$. While at first glance it might appear that an iterative numerical solution is necessary, it can be solved analytically as follows. First, define the intermediate variable

$$y_n \equiv W(a_n e^{[a_n - b_n \bar{c}(\tau_{n+1} - \tau_n)]}), \quad (3.21)$$

which must satisfy

$$y_n = a_n + b_n (\xi_n - \xi_f). \quad (3.22)$$

By the property $y = W(ye^y)$ equality between the two can be rewritten as

$$y_n e^{y_n} = a_n e^{[a_n - b_n \bar{c}(\tau_{n+1} - \tau_n)]}, \quad (3.23)$$

which can then be solved for τ_{n+1} (with y_n replaced) as

$$\tau_{n+1} = \tau_n + \frac{1}{\bar{c}} \left[\xi_f - \xi_n - \frac{1}{b_n} \ln \left(1 - \frac{b_n}{a_n} (\xi_f - \xi_n) \right) \right]. \quad (3.24)$$

In fact, this is just a special instance of the implicit form of equation (3.10) where time can be considered a function of the phase fraction

$$\tau(\xi) = \tau_n + \frac{1}{\bar{c}} \left[\xi - \xi_n - \frac{1}{b_n} \ln \left(1 - \frac{b_n}{a_n} (\xi - \xi_n) \right) \right]. \quad (3.25)$$

4. Numerical example

As a numerical example, the simplified model is compared to the finite element calculations of Chang *et al* [1]. The finite element model was implemented using an in-house research code. Weak forms of the equilibrium and heat equations were solved through staggered Newton–Raphson iteration between thermal and mechanical degrees of freedom. The kinetic law was implemented using an explicit 4th-order Runge–Kutta algorithm. The results presented here used 200 elements along the specimen length, with displacement and its gradient given by Hermite cubic shape functions and temperature by linear shape functions.

The SMA material parameters are taken from table 1, while table 2 lists the spring parameters, ambient environment constants (-50°C stagnant air), and electrical power history similar to that used in [1]. The finite element simulation used thermally insulated ends, for which the temperature field was relatively uniform (as opposed to constant temperature boundary conditions), to provide a fair comparison with our lumped model simulation. The full set of dimensionless constants is provided in table 3. (The subset of mechanical parameters from this list were used to construct figure 4.) These parameters result in a steady state temperature rise of $\Delta T_\infty = 100.6^\circ\text{C}$ ($\Delta\theta^{\text{ss}} = \bar{P}_e = 0.01061$) above the

Table 3. Dimensionless parameters.

SMA	Springs	Ambient	Power
$\beta = 0.0591$	$\eta = 0.06977$	$\theta_a = -0.002758$	$\bar{P}_e = 0.01061$
$\lambda = 0.02630$	$\eta_B = 0.05$		
$\bar{c} = 7.446$	$\eta_E = 0.025$		
$\bar{\mu}_c = 0.001587$	$\bar{\Delta} = 0.0948$		

Table 4. Dimensionless actuator transformation temperatures.

\bar{M}_f	\bar{M}_s	\bar{A}_s	\bar{A}_f
0.00011	0.00424	0.00329	0.00741

ambient temperature when power is applied, a characteristic time of $t^* = 6.88 \text{ s}$ for sensible heating, and an initial pre-stress of $\sigma_0 = 119 \text{ MPa}$ ($\bar{\sigma}_0 = 0.00170$). The specific dimensionless (stress-dependent) temperatures at which transformation occurs in the actuator during heat and cooling are provided in table 4.

4.1. Time-dependent response

Simulation results for a complete heating/cooling cycle are shown graphically in figures 6 for dimensionless quantities. A column of plots are shown against a common time axis for (a) the applied power (\bar{P}_e), (b) evolution of temperature (θ), stress ($\bar{\sigma}$), and strain (ϵ), and (c) M^+ phase fraction (ξ) and normalized driving force (f). Electrical power (1.3 W) is applied for about 7.268 dimensionless time units (50 s), and then turned off for the same time period. Initially, the phase fraction, strain, and stress remain constant, and the temperature rises rapidly according to an exponential curve during sensible heating. This is interrupted once $M^+ \rightarrow A$ phase transformation starts, at which point the phase fraction and strain decrease while the stress increases. The temperature continues to rise by another exponential-like curve, but at a slower rate, since both sensible heat and latent heat must be supplied to transform the SMA material. Once the transformation is complete ($\xi = 0$), the phase fraction, strain, and stress again remain constant while the temperature rises at a more rapid rate according to sensible heating only.

When the power is turned off, a similar, but reverse direction behavior, occurs. The phase fraction, strain, and stress stay constant initially while the temperature decreases rapidly according to sensible cooling. At the onset of $A \rightarrow M^+$ transformation the temperature decrease is arrested somewhat as latent heat is extracted from the surroundings. Once the transformation saturates ($\xi = 1$), the actuator has reset to its initial phase fraction, strain, and stress, and the temperature decreases asymptotically toward the ambient temperature according to sensible cooling again. Overall, the start and termination of phase transformation are accompanied by noticeable discontinuities in the rates of change of all state variables. Finally, the response path taken for the entire cycle is plotted in stress–strain–temperature space in figure 7, showing the complete hysteresis loop.

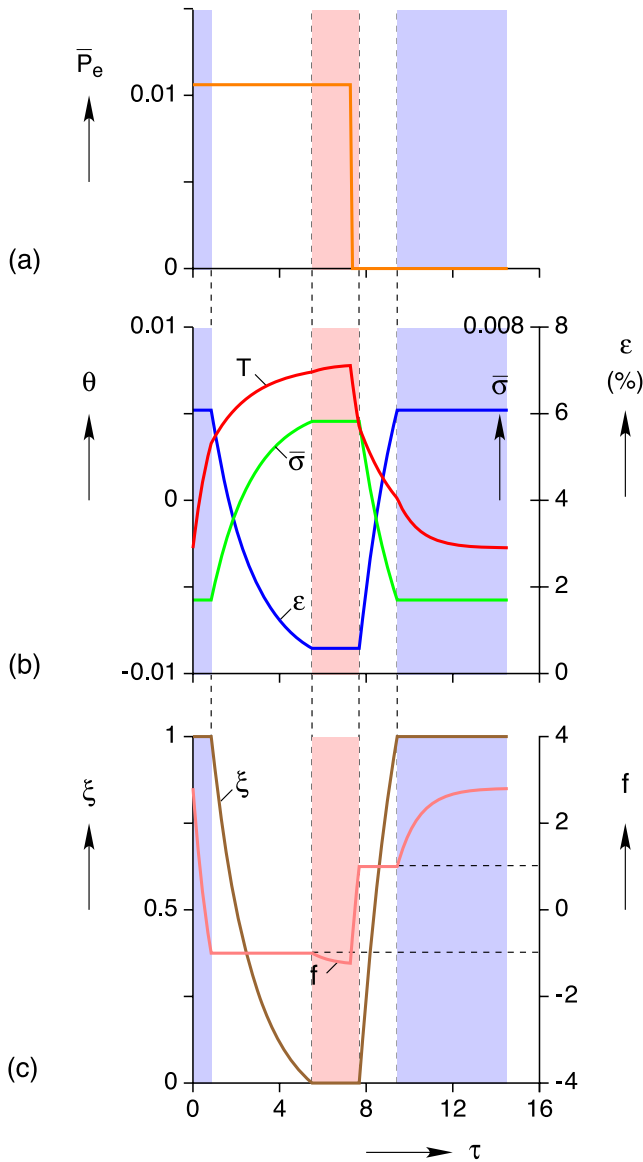


Figure 6. Dimensionless SMA actuator response: (a) the applied power history, (b) temperature, stress, and strain, and (c) M^+ phase fraction and normalized driving force. Light and dark shaded areas indicate time regimes for pure M^+ and A phases, respectively.

4.2. Comparison to finite element simulation

To explore the reasonableness of the lumped model solution we now compare the above simulation to that of the previous finite element numerical simulation (see figures 27 and 28 of [1]). Figure 8(a) shows a comparison of the average strain histories. Figure 8(b) shows a comparison of the temperature histories (mid-span temperature for the FE case). The open and closed circles show the onset and termination of transformations of the lumped model. The slight delay in the onset of transformation of the FE case was due to the input power history that was slightly different, i.e., ramped up linearly over 2.5 s, rather than instantaneously, to the final constant value (and then ramped down over the same time interval when switched off). Figure 8(c) shows force histories in the external spring. Overall, the responses shown in figure 8 are reasonably similar,

Table 5. Numerical results: values at significant times.

n	τ_n	t_n (s)	θ_n	T_n ($^{\circ}\text{C}$)	ϵ_n (%)	ξ_n	$\bar{\sigma}_n$	σ_n (GPa)	Event
0	0	0	-0.00276	-50	6.081	1	0.0017	0.119	heat
1	0.843	5.8	0.00329	7.3	6.081	1	0.0017	0.119	A_s
2	5.504	37.9	0.00741	46.4	0.582	0	0.00582	0.408	A_f
3	7.268	50	0.00778	49.9	0.582	0	0.00582	0.408	cool
4	7.678	52.8	0.00424	16.3	0.582	0	0.00582	0.408	M_s
5	9.431	64.9	0.00011	-22.8	6.081	1	0.0017	0.119	M_f
6	14.537	100	-0.00274	-49.8	6.081	1	0.0017	0.119	—

and other minor differences, such as the jaggedness in the FE temperature history in figure 8(b) can be attributed to the spatial non-uniformities due to mechanical instabilities and localization effects (a focus of that study) captured by the finite element simulation but not by our lumped model. As a convenient design-level simulation model, we consider the results of the lumped model to be quite satisfactory.

4.3. Actuation times

Of particular interest is the duration of time needed to complete such an actuation cycle. Table 5 provides both dimensionless and dimensional values at significant times during the cycle. In this example, sensible heating takes about 5.8 s before transformation starts, $M^+ \rightarrow A$ transformation takes about 32.1 s, and post transformation sensible heating occurs for another 12 s, or so (arbitrary, depending on the duration of heating). Sensible cooling takes about 2.8 s, starting from about 50°C , and $A \rightarrow M^+$ transformation takes about 12.1 s. In this case, the time for transformation during heating is longer than that during cooling due to the heat losses to the ambient environment, but in general that need not be the case if the electrical power input during heating is large enough.

It is also interesting to study the times necessary to start and finish transformation as a function of different heat transfer environments and electrical power inputs. Setting $\theta(\tau_1) = \bar{A}_s$ in equation (3.8) with the appropriate initial conditions, the dimensionless time needed to start $M^+ \rightarrow A$ transformation is

$$\tau_1 = \ln \left(\frac{\theta^{ss} - \theta_a}{\theta^{ss} - \bar{A}_s} \right). \quad (4.1)$$

The dimensionless time needed to finish $M^+ \rightarrow A$ transformation can be found by setting $\xi(\tau_2) = 0$ in equation (3.10) with the appropriate starting conditions.

$$\tau_2 = \tau_1 + \left(1 + \frac{\theta^{ss} + \lambda - \bar{\mu}_c}{\beta \eta \bar{c}} \right) \ln \left(\frac{\theta^{ss} - \bar{A}_s}{\theta^{ss} - \bar{A}_f} \right) - \frac{1}{\bar{c}}. \quad (4.2)$$

These two are converted to dimensional times (t_1 and t_2) and are plotted in figure 9 as a function of convective film coefficient (h) and electrical power input (P_e), respectively, while holding all other parameters fixed. One can see in figure 9(a) that the chosen electrical power ($P_e = 1.3$ W) is insufficient to start the transformation ($t_1 \rightarrow \infty$) for convective film coefficients larger than about 158 $\text{W m}^{-2} \text{K}^{-1}$, and it is insufficient to finish the transformation ($t_2 \rightarrow \infty$) for values larger than 93.9 $\text{W m}^{-2} \text{K}^{-1}$. Figure 9(b) shows that for a

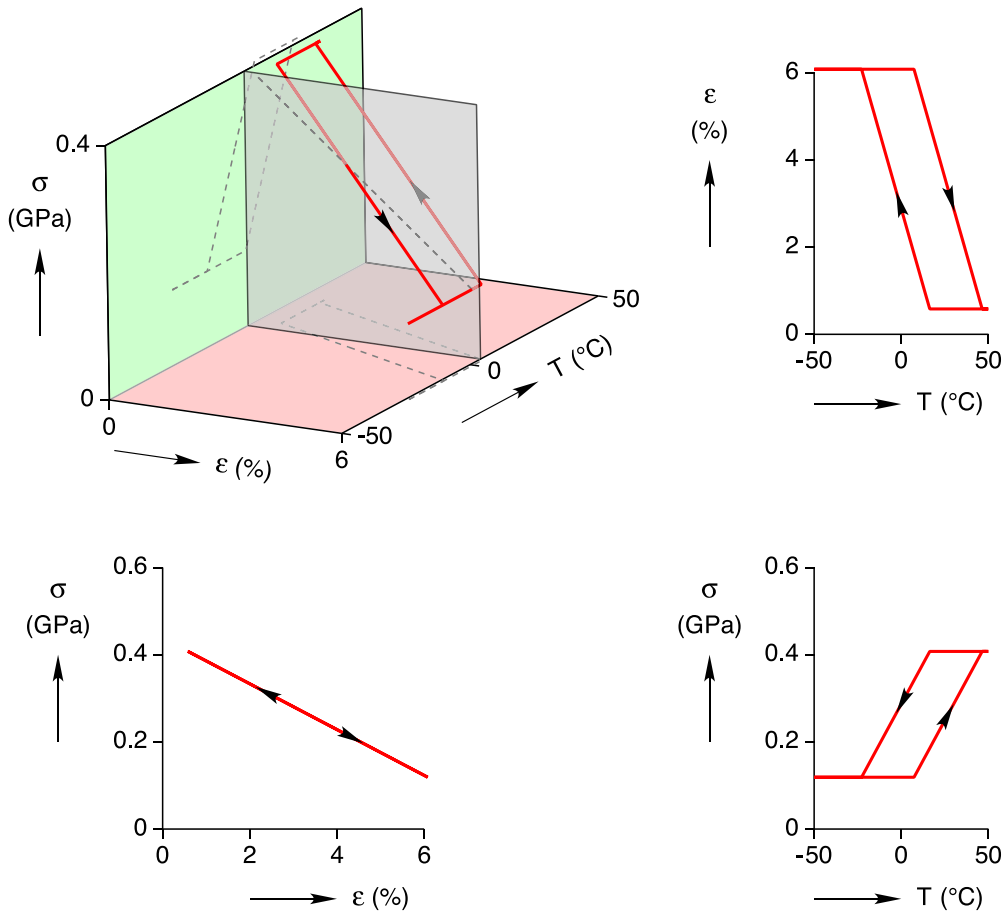


Figure 7. Dimensional SMA actuator response in stress–strain–temperature space along with planar projections.

fixed convection coefficient ($h = 90 \text{ W m}^{-2} \text{ K}^{-1}$) the times to start and finish transformation are strongly dependent on the electrical power input. A modest increase in power input will dramatically decrease the time needed. For low power input they asymptotically approach infinity at particular power levels, 0.741 W to start and 1.25 W to finish the $M^+ \rightarrow A$ transformation. Accordingly, our chosen power, 1.3 W, is barely sufficient for our example actuator system.

5. Design aspects

In this section we consider some issues related to actuator design based on our simple actuator model. The numerical example of section 4 was an ad hoc design, and no attempt was made to optimize it. Here, we discuss a systematic design approach by considering minimum requirements for feasible actuation. We also show how the design could be optimized given SMA stress limits and constraints on power and energy. We start with the parameters of table 3 and then consider modifications of parameters one at a time. Thus, parameters of table 3 are assumed unless stated otherwise throughout this section. The following considers actuation stroke, sizing of springs, actuation time, minimum power requirements, and energy usage and energy efficiency. The section concludes with a discussion of potential refinements to the model to improve its accuracy and generality for design.

5.1. Actuation stroke

The maximum stress clearly occurs when the SMA element is fully A phase. From equations (2.34) and (2.39) the maximum stress in general is

$$\bar{\sigma}_{\max} = \bar{\sigma}_0 + \beta\eta. \tag{5.1}$$

This can be combined with eqn. (2.38) to obtain the actuation stroke of the external spring, $\delta_E = \Delta\varepsilon L$, which is calculated using

$$\Delta\varepsilon = \varepsilon_0 - \bar{\sigma}_{\max}, \tag{5.2}$$

since the dimensionless stress, $\bar{\sigma}_{\max}$ is the numerical equivalent of the minimum SMA strain during actuation. This simplifies to

$$\Delta\varepsilon = \beta(1 - \eta) = \frac{\beta}{1 + \eta_B + \eta_E}. \tag{5.3}$$

According this expression, the best case is $(\delta_E/L)_{\max} = \beta$, which corresponds to complete dead loading ($\eta_B = \eta_E = 0$). It predicts that the stroke is independent of the starting (ambient) temperature and the initial mismatch ($\bar{\Delta}$), and that larger values of η_B and η_E reduce the actuator stroke somewhat. However, one should be wary of this conclusion, since the elastic modulus for Martensite was assumed to be the same as that of Austenite in our simplified SMA model. In reality, the Martensite mechanical response is quite

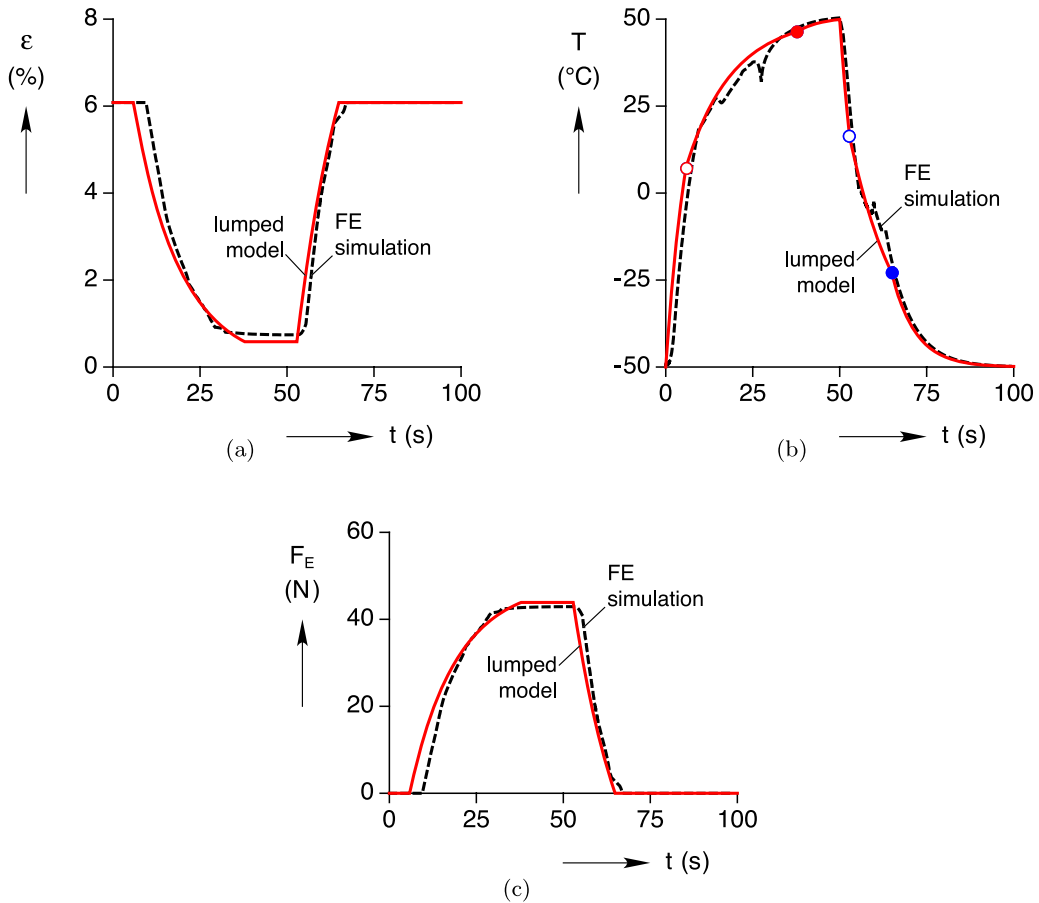


Figure 8. Comparison of simulation histories between lumped actuator model (solid line) and finite element (FE) simulation (dashed line) of Chang *et al* [1]: (a) SMA strain (global average for FE case), (b) temperature (mid-length for FE case), (c) external spring force.

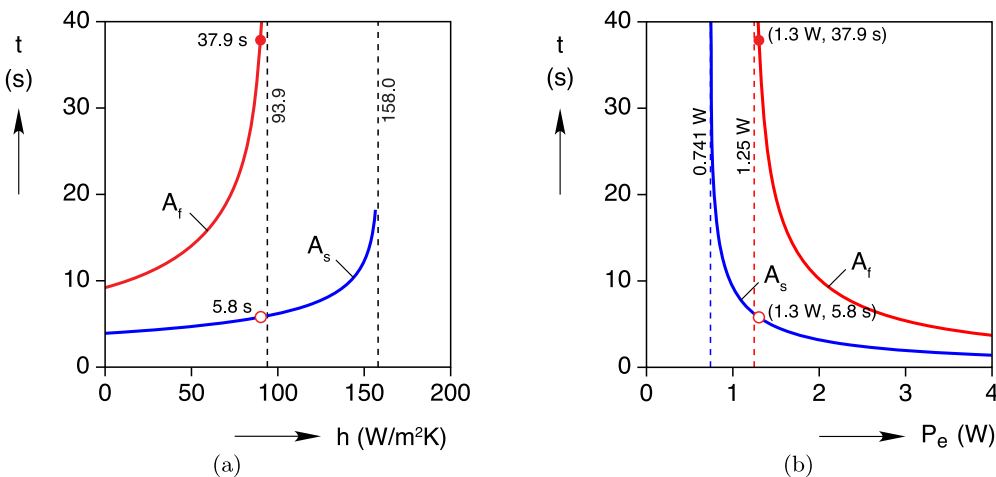


Figure 9. Times to start, t_1 (A_s), and finish, t_2 (A_f), $M^+ \rightarrow A$ transformation: (a) as a function of convective film coefficient, h , (b) as a function of electrical power input, P_e . All other parameters are fixed as given in table 2. Open and closed circles correspond to the numerical example.

nonlinear at large strains, having a tangent modulus much lower than the Austenite modulus. We consider this to be the most severe limitation of our model, since β is treated as a material constant. Changing the ambient temperature θ_a and/or the pre-stress $\bar{\sigma}_0$ may give less accurate stroke predictions.

A possible refinement for generalizing the model to more accurately predict the actuator stroke under general conditions is discussed later in section 5.7. Nevertheless, the actuator model as presented is simple and convenient, so we continue to explore its design implications in the next several sections.

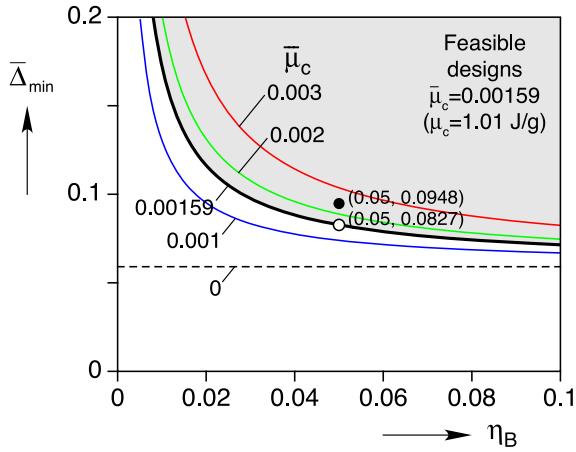


Figure 10. Minimum bias spring mismatch ($\bar{\Delta}_{\min}$) as a function of dimensionless bias spring stiffness (η_B) to ensure initial complete reorientation of Martensite (M^+) during initial actuator assembly (cold). The gray region indicates feasible designs for values of β and $\bar{\mu}_c$ of table 3.

5.2. Sizing of springs

As discussed in section 2.4, equation (2.37) is the design condition to ensure that full initial tensile Martensite ($\xi_0 = 1$) exists upon assembly of the actuator. The lower bound of feasible designs is

$$\bar{\Delta}_{\min} = \beta + \bar{\sigma}_M \left[1 + \frac{1}{\eta_B} \right], \quad (5.4)$$

recalling that $\bar{\sigma}_M \equiv \bar{\mu}_c / \sqrt{2}$ is the dimensionless stress needed to ensure M^+ exists when cold. This is plotted in figure 10 (solid black line). In our design example this corresponds to the minimum pre-stress condition $\sigma_0 \geq \sigma_M (= 78.5 \text{ MPa})$. The filled black circle in the figure is our chosen design ($\bar{\Delta} = 0.0948$), while the open circle is the minimum feasible value

($\bar{\Delta}_{\min} = 0.0827$) for the same bias spring stiffness. This latter case can be considered an ‘optimal’ choice, since it minimizes the stress in the SMA element. For reference, the other lines in the figure indicate lower bounds to other feasible designs for SMA wire with other values of hysteresis ($\bar{\mu}_c$).

Throughout this study we have assumed that no cyclic shakedown, or ratcheting, of the SMA element occurs during operation of the actuator. This is true only if the SMA element has been conditioned (or trained) to achieve repeatable cyclic response, or if the stress level is kept sufficiently low. In either case the maximum SMA stress must be maintained acceptably low during operation. With the maximum stress from equation (5.1), the maximum stress for the ‘optimal design’ is

$$\bar{\sigma}_{\max}^{\text{opt}} = \bar{\sigma}_M + \beta \eta. \quad (5.5)$$

These constant maximum stress lines are plotted, respectively, in figures 11(a) and (b) in the space of bias spring and external spring stiffnesses. The space in figure 11(a) is truncated by a dotted line, which is the minimum bias spring stiffness ($\eta_{B,\min} = 0.0325$) for the chosen bias spring mismatch ($\bar{\Delta} = 0.0948$) according to equation (2.37). The solid black circle indicates the maximum stress of 408 MPa for the chosen values of η_B and η_E in table 3. Figure 11(b) has no such restriction, since $\bar{\Delta} = \bar{\Delta}_{\min}$ has been substituted into equation (5.5). In this case, the open circle indicates a lower maximum stress of 367 MPa, since the optimal bias spring mismatch has been used. The shaded regions of figure 11 show regimes of η_B and η_E to keep the maximum stress under a hypothetical design value of 400 MPa.

Another possible step to minimize the worst case stress is to use a very compliant bias spring with a suitable value of $\bar{\Delta}$ to maintain $\bar{\sigma}_0 = \bar{\sigma}_M$. The designer should use as compliant a bias spring as is reasonable to meet actuator size requirements. In the limit as $\eta_B \rightarrow 0$ the bias spring mismatch gets very large ($\bar{\Delta} \rightarrow \infty$), which corresponds essentially to dead loading by the bias spring. While this may be difficult to achieve in

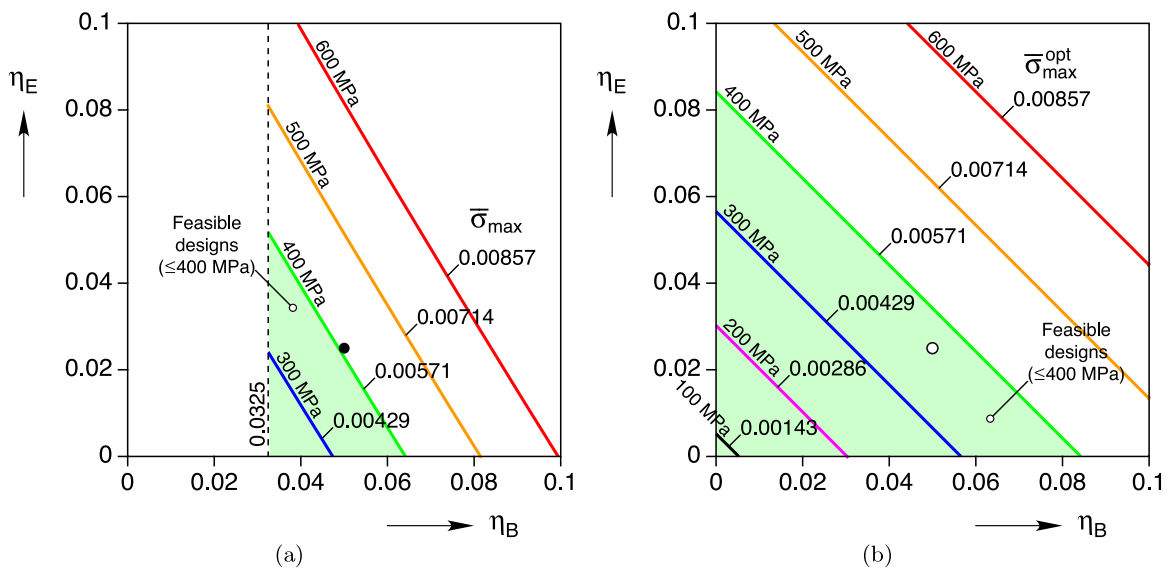


Figure 11. Design stresses as a function of bias spring (η_B) and external spring (η_E) stiffnesses: (a) ‘non-optimal design’ using $\bar{\Delta} = 0.0948$ (solid circle is the example of section 4), (b) ‘optimal design’ using $\bar{\Delta}_{\min}$.

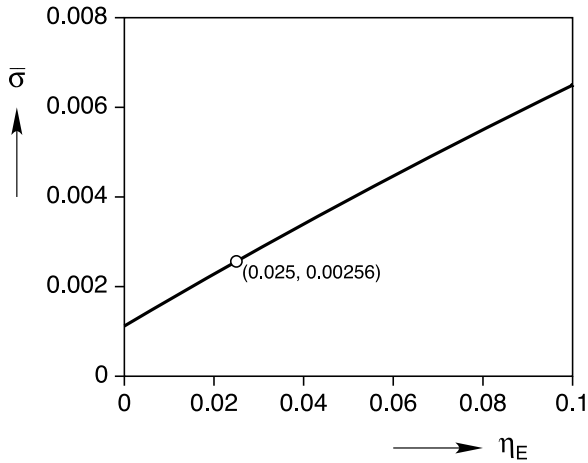


Figure 12. Dimensionless maximum stress ($\bar{\sigma}_{\max}^{\text{best}}$) as a function of external spring stiffness (η_E) for dead loading bias ($\eta_B = 0$, $\bar{\sigma}_0 = \bar{\sigma}_M$).

practice, we consider it as a limiting ideal case. Substituting $\eta_B = 0$ in equation (5.5) gives a best case maximum stress of

$$\bar{\sigma}_{\max}^{\text{best}} = \bar{\sigma}_M + \frac{\beta\eta_E}{1 + \eta_E}, \quad (5.6)$$

which is plotted in figure 12. For the chosen value of $\eta_E = 0.025$ this gives a maximum stress of only 0.002 56 (179 MPa) as shown by the open circle in the figure.

5.3. Actuation times

Here we revisit the issue of actuation times to start and finish $M^+ \rightarrow A$ transformation. Equations (4.1) and (4.2) can be rewritten explicitly in terms of \bar{P}_e as

$$\tau_1 = \ln\left(\frac{\bar{P}_e}{\bar{P}_e + \theta_a - \bar{A}_s}\right), \quad (5.7)$$

$$\tau_2 = \tau_1 + \left(1 + \frac{\bar{P}_e + \theta_a + \lambda - \bar{\mu}_c}{\beta\eta\bar{c}}\right) \ln\left(\frac{\bar{P}_e + \theta_a - \bar{A}_s}{\bar{P}_e + \theta_a - \bar{A}_f}\right) - \frac{1}{\bar{c}}. \quad (5.8)$$

Figure 13(a) shows how figures 9(a) and (b) collapse to a single plot on a dimensionless basis, since the definition of \bar{P}_e includes both the input power and the film coefficient. Other parameters in table 3 are held fixed. Open and closed circles correspond to the example of section 4. Dotted lines indicate minimum dimensionless powers to start (0.006 04) and finish (0.0102) the transformation, i.e., where $\tau_1 \rightarrow \infty$ and $\tau_2 \rightarrow \infty$, respectively. The figure also shows the time interval ($\Delta\tau_{12} = \tau_2 - \tau_1$) during which transformation occurs.

5.4. Minimum power requirements

Minimum dimensionless powers to start and finish the $M^+ \rightarrow A$ transformation are the singular points of equations (5.7) and (5.8), given by

$$(\bar{P}_e)_{\min} = \begin{cases} \bar{A}_s - \theta_a, & \text{start} \\ \bar{A}_f - \theta_a, & \text{finish.} \end{cases} \quad (5.9)$$

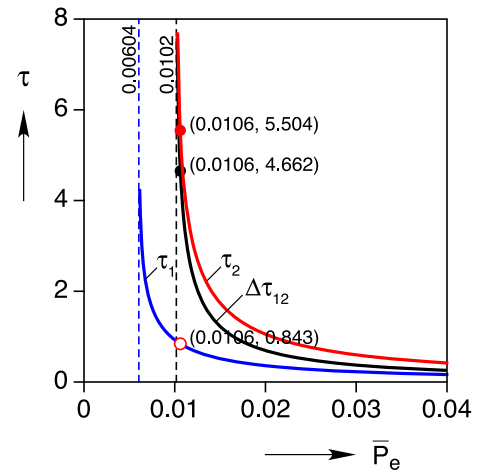


Figure 13. Dimensionless times (τ) during $M^+ \rightarrow A$ transformation as a function of dimensionless power (\bar{P}_e): start time, τ_1 (to reach A_s), time period during $M^+ \rightarrow A$ transformation, $\Delta\tau_{12}$, and finish time, τ_2 (to reach A_f).

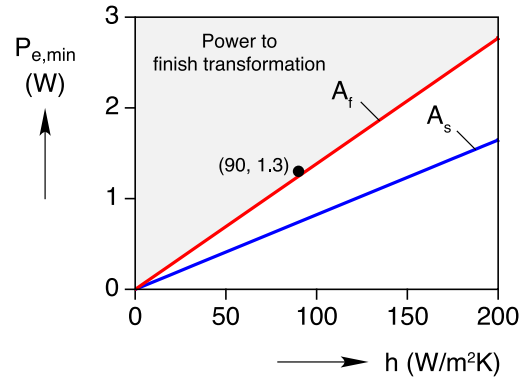


Figure 14. Minimum power requirements to start (A_s) and finish (A_f) $M^+ \rightarrow A$ transformation as a function of film coefficient.

These can be converted to dimensional power requirements using definition (2.33) as simple linear functions of the film coefficient as plotted in figure 14, keeping all other parameters fixed in table 3. The gray region shows feasible powers to achieve complete actuation. The solid circle is the example of section 4.

Again using equation (5.9) we now consider relaxing other parameters to study the minimum power requirements for alternate designs. Figure 15 shows lines of constant dimensionless power needed to achieve A_s in the design space of dimensionless bias and external spring constant (η_B) and bias spring mismatch ($\bar{\Delta}$). The lower bound (dotted black line) is the minimum design curve consistent with the solid black line of figure 10, which requires a minimum power of $\bar{P}_e = 0.005 47$. The bold dotted black line corresponds to the power ($\bar{P}_e = 0.0106$) used in the example of section 4. The gray region is the feasible design space for ($\eta_B, \bar{\Delta}$) for this power level. The solid black circle is the corresponding coordinates ($\eta_B, \bar{\Delta}$) of the example, and the open circles indicate the minimum and maximum values of $\bar{\Delta}$ for the same η_B .

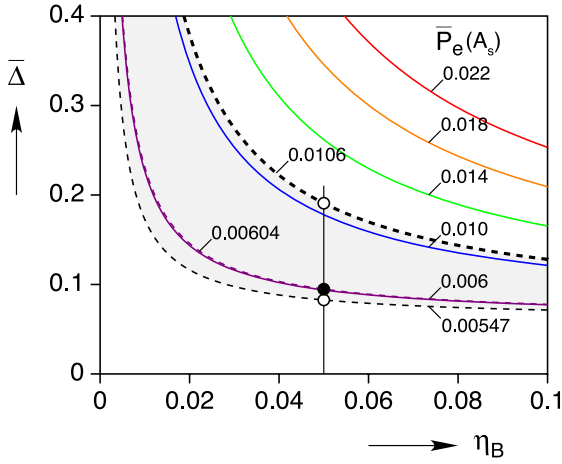


Figure 15. Minimum dimensionless power curves to start $M^+ \rightarrow A$ transformation, $\bar{P}_e(A_s)$, in the design space of dimensionless bias (η_B) and external ($\bar{\Delta}$) spring constants.

Figure 16(a) then shows lines of constant dimensionless power needed to achieve A_f in the design space of dimensionless bias spring and external spring stiffnesses (η_B, η_E) while holding $\bar{\Delta} = 0.0948$ of the example. Clearly, larger values of η_B and η_E require larger power levels, since they raise $\bar{\sigma}_{max}$ and thus raise the corresponding transformation temperature, A_f . The gray region is the feasible design space for (η_B, η_E) for the power level $\bar{P}_e = 0.0106$. The solid black circle is the corresponding coordinates (η_B, η_E) = (0.05, 0.025) of the example, and the open circles indicate the minimum and maximum values of η_E for the same η_B . Figure 16(b) shows a similar design plot where an optimal value of $\bar{\Delta}_{min}$ is chosen according to equation (5.4). This better choice expands the design space to larger permissible values of η_E for a given power level, or allows a lower power level to be used for a given (η_B, η_E) design.

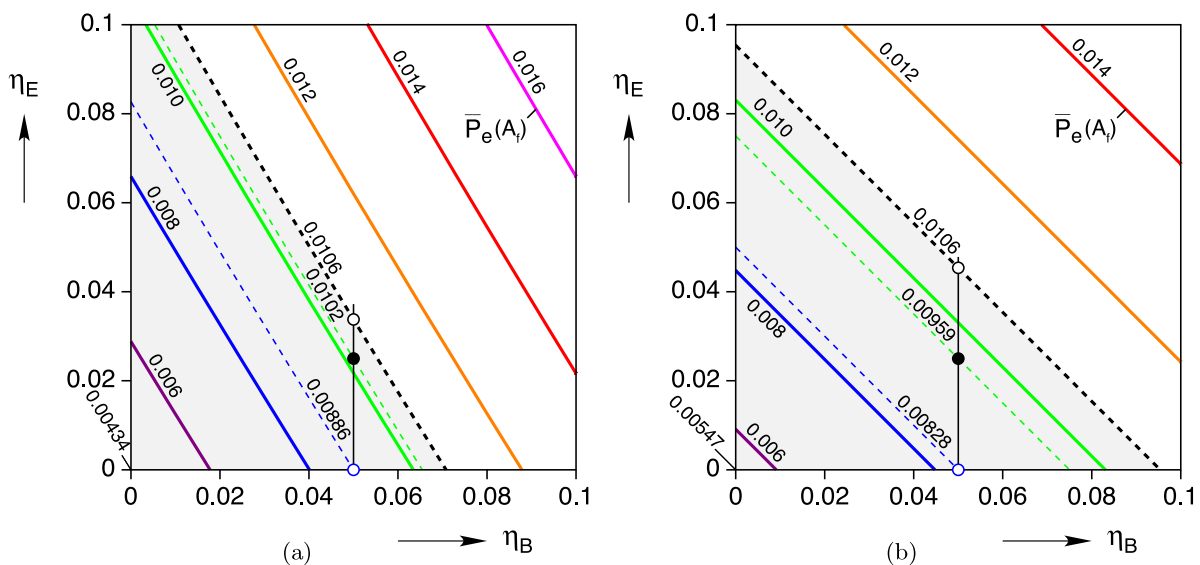


Figure 16. Minimum dimensionless power curves to finish $M^+ \rightarrow A$ transformation, $\bar{P}_e(A_f)$, in the design space of dimensionless bias (η_B) and external (η_E) spring stiffnesses: (a) with fixed $\bar{\Delta} = 0.0948$, (b) with $\bar{\Delta} = \bar{\Delta}_{min}$ from equation (5.4).

5.5. Energy usage

Another consideration beside power requirements is the amount of input energy necessary to start and finish actuation. This could be important if the power source has a finite energy storage, such as a battery or capacitor. Since the applied power is assumed constant during the actuation portion of the cycle, the input energy is simply $\mathcal{E}_{in} = P_e \Delta t$. Defining a dimensionless input energy as

$$\bar{\mathcal{E}}_{in} \equiv \frac{\mathcal{E}_{in}}{\mathcal{E}^*}, \tag{5.10}$$

where a characteristic energy has been defined as $\mathcal{E}^* \equiv EAL\beta$, the dimensionless input energy necessary to reach A_f from T_a is

$$\bar{\mathcal{E}}_{in} = \bar{c} \bar{P}_e \tau_2, \tag{5.11}$$

which, using equations (5.7) and (5.8), works out to be

$$\bar{\mathcal{E}}_{in} = \bar{P}_e \left[\bar{c} \ln \left(\frac{\bar{P}_e}{\bar{P}_e + \theta_a - \bar{A}_f} \right) + \left(\frac{\bar{P}_e + \theta_a + \lambda - \bar{\mu}_c}{\beta \eta} \right) \times \ln \left(\frac{\bar{P}_e + \theta_a - \bar{A}_s}{\bar{P}_e + \theta_a - \bar{A}_f} \right) - 1 \right]. \tag{5.12}$$

The total input energy can be partitioned as

$$\bar{\mathcal{E}}_{in,S} = \bar{c} \bar{P}_e \ln \left(\frac{\bar{P}_e}{\bar{P}_e + \theta_a - \bar{A}_f} \right), \tag{5.13}$$

$$\bar{\mathcal{E}}_{in,T} = \bar{P}_e \left[\left(\frac{\bar{P}_e + \theta_a + \lambda - \bar{\mu}_c}{\beta \eta} \right) \ln \left(\frac{\bar{P}_e + \theta_a - \bar{A}_s}{\bar{P}_e + \theta_a - \bar{A}_f} \right) - 1 \right], \tag{5.14}$$

where $\bar{\mathcal{E}}_{in,S}$ is the total sensible heating contribution to reach A_f , and $\bar{\mathcal{E}}_{in,T}$ is the latent heat and spring energy during transformation from A_s to A_f . Equations (5.13) and (5.12) are plotted against \bar{P}_e in figure 17(a) using the remaining parameters of table 3. One can see that the required input energy drops dramatically as the power is increased, since

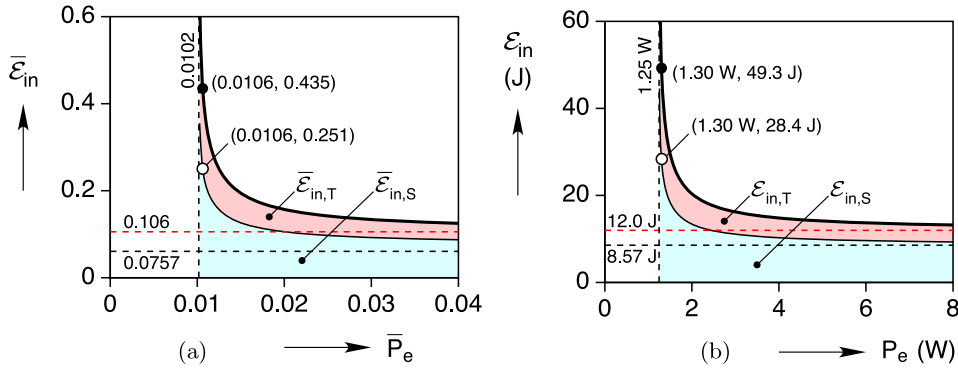


Figure 17. Input energies as a function of applied power to reach A_f with shaded regions indicated sensible heat and latent heat/spring energy portions. Numerical example values are shown by open and closed circles. (a) Dimensionless plot, holding all SMA parameters and ambient temperature constant, (b) dimensional plot, holding all parameter values except input power constant.

the time for ambient heat loss is reduced. The shaded areas show the individual contributions of the sensible heat and latent heat/spring energy contributions. Figure 17(b) also shows a corresponding dimensional plot, holding all parameters fixed except the input power, to give a sense of the energy and power magnitudes for the numerical example.

In the limit as $\bar{P}_e \rightarrow \infty$ the actuation is adiabatic (no ambient heat loss) and the input energy reaches a finite lower bound. The infinite limit can be found by a change of variables to a zero limit and then an application of L'Hôpital's rule as

$$\begin{aligned} \lim_{x \rightarrow \infty} x \ln \left(\frac{x+a}{x+b} \right) &= \lim_{y \rightarrow 0} \frac{\ln \left(\frac{1+ay}{1+by} \right)}{y} \\ &= \lim_{y \rightarrow 0} \frac{a-b}{(1+ay)(1+by)} = a-b. \end{aligned} \quad (5.15)$$

Accordingly, equations (5.13), (5.14), (5.12) become

$$\bar{\mathcal{E}}_{in,S}^{\infty} = \bar{c} (\bar{A}_f - \theta_a), \quad (5.16)$$

$$\bar{\mathcal{E}}_{in,T}^{\infty} = \frac{1}{2} \beta \eta + \lambda + \bar{\sigma}_0, \quad (5.17)$$

$$\bar{\mathcal{E}}_{in}^{\infty} = \frac{1}{2} \beta \eta + \lambda + \bar{\sigma}_0 + \bar{c} (\bar{A}_f - \theta_a), \quad (5.18)$$

as $\bar{P}_e \rightarrow \infty$. The lower bound values $\bar{\mathcal{E}}_{in,S}^{\infty}$ and $\bar{\mathcal{E}}_{in}^{\infty}$ are shown by dashed lines in figure 17(a).

5.6. Energy efficiency

The numerical example of section 4 applied power longer than was necessary to achieve full actuation, just for illustration purposes. In practice, a lower power could be used to maintain the actuator position, or better yet, a latching mechanism could be employed. Inducing temperatures above A_f does no additional work against the external spring, so is wasted energy. In this section, we consider the efficiency of the actuator on an energy basis. We compare the input energy (now denoted \mathcal{E}_{in} , dropping the subscript 2) to the output energy (\mathcal{E}_{out}) of the actuator. The output energy is simply the work done against the external spring, which is

$$\bar{\mathcal{E}}_{out} = \frac{\beta \eta_E}{2(1 + \eta_B + \eta_E)^2}, \quad (5.19)$$

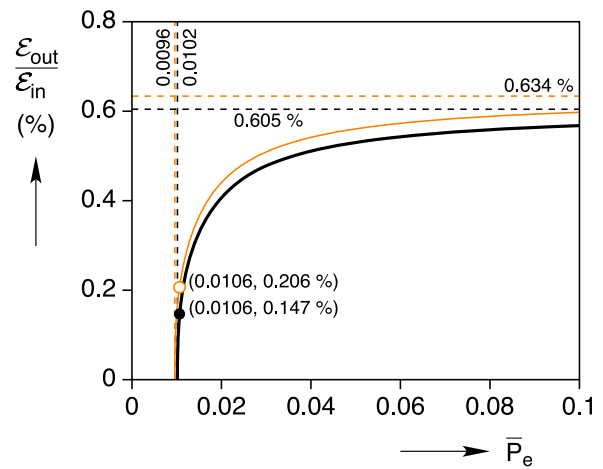


Figure 18. Energy efficiency ($\mathcal{E}_{out}/\mathcal{E}_{in}$) as a function of dimensionless power (\bar{P}_e) for the design of table 3 (bold line) and with minimum bias spring mismatch (thin line).

using a similar dimensionless energy ($\bar{\mathcal{E}}_{out} = \mathcal{E}_{out}/\mathcal{E}^*$). The energy efficiency is the ratio $\mathcal{E}_{out}/\mathcal{E}_{in}$ (or $\bar{\mathcal{E}}_{out}/\bar{\mathcal{E}}_{in}$),

$$\begin{aligned} \frac{\mathcal{E}_{out}}{\mathcal{E}_{in}} &= \frac{\beta \eta_E}{2(1 + \eta_B + \eta_E)^2} \\ &\times \left\{ (\theta^{ss} - \theta_a) \left[\bar{c} \ln \left(\frac{\theta^{ss} - \theta_a}{\theta^{ss} - \bar{A}_f} \right) \right. \right. \\ &\left. \left. + \left(\frac{\theta^{ss} + \lambda - \bar{\mu}_c}{\beta \eta} \right) \ln \left(\frac{\theta^{ss} - \bar{A}_s}{\theta^{ss} - \bar{A}_f} \right) - 1 \right] \right\}^{-1}. \end{aligned} \quad (5.20)$$

The best efficiency in the limit of infinite power input is

$$\begin{aligned} \left(\frac{\mathcal{E}_{out}}{\mathcal{E}_{in}} \right)^{\infty} &= \frac{\beta \eta_E}{(1 + \eta_B + \eta_E)^2} \\ &\times \left[\frac{1}{\beta \eta + 2(\lambda + \bar{\sigma}_0 + \bar{c} [\bar{A}_f - \theta_a])} \right]. \end{aligned} \quad (5.21)$$

The efficiency curve, equation (5.20), is plotted in figure 18 (bold line) against the dimensionless power input using $\theta^{ss} \equiv \bar{P}_e + \theta_a$ with the remaining parameters taken from table 3. The efficiency of our numerical example is quite low (0.147%), while the best case efficiency is 0.605% for infinite power

input. The case of optimal bias spring mismatch ($\bar{\Delta}_{\min}$) is also shown for comparison (thin line), having slightly better efficiency (0.206%) at the chosen input power and upper bound efficiency (0.634%). While the energy density of SMA's is several orders of magnitude larger than other typical adaptive materials [12], the energy usage can be quite lossy for a given work output, which is a well-known issue for typical SMA actuators. As a final case, it is also interesting to study the best possible energy efficiency of dead loading ($\eta = 0$). For this case, the constants in equation (3.11) simplify to

$$a_1 = \frac{\bar{\sigma}_0 - (\bar{P}_e + \theta_a + \bar{\mu}_c)}{\lambda + \bar{P}_e + \theta_a + \bar{\mu}_c}, \quad (5.22)$$

$$b_1 = 0,$$

where $\bar{\sigma} = \bar{\sigma}_0$ is the dead load stress. The governing equation (3.9), simplifies to

$$\bar{c} a_1 - (1 + a_1) \xi'(\tau) = 0, \quad (5.23)$$

which has the simple linear solution in time

$$\xi(\tau) = (\tau - \tau_1) \frac{\bar{c} a_1}{1 + a_1}. \quad (5.24)$$

Proceeding as before, accounting for the time to reach A_s and the time to transform the SMA element, the dead load energy efficiency is

$$\frac{\mathcal{E}_{\text{out}}}{\mathcal{E}_{\text{in}}} = \bar{\sigma}_0 \left\{ \bar{c} \bar{P}_e \ln \left[\frac{\bar{P}_e}{\bar{P}_e + \theta_a - (\bar{\mu}_c + \bar{\sigma}_0)} \right] + \frac{\bar{P}_e (\lambda + \bar{\sigma}_0)}{\bar{P}_e + \theta_a - (\bar{\mu}_c + \bar{\sigma}_0)} \right\}^{-1}. \quad (5.25)$$

Note that $\bar{A}_s = \bar{A}_f = \bar{\mu}_c + \bar{\sigma}_0$ in this case. The limiting case of $\bar{P}_e \rightarrow \infty$ is

$$\left(\frac{\mathcal{E}_{\text{out}}}{\mathcal{E}_{\text{in}}} \right)^{\infty} = \frac{\bar{\sigma}_0}{\bar{\sigma}_0 + \lambda + \bar{c} [\bar{\mu}_c + \bar{\sigma}_0 - \theta_a]}. \quad (5.26)$$

Figure 19 shows contours of energy efficiency, equation (5.25), in the space of $\bar{\sigma}_0$ and \bar{P}_e , leaving the SMA parameters and ambient temperature fixed according to table 3. The efficiency generally improves as $\bar{\sigma}_0$ increases, due to the greater output work done, except near minimum powers (slanted dashed line) where time for actuation becomes long. Here, one can see that energy efficiencies of a few per cent can be achieved for moderate dimensionless stress and power levels.

5.7. Potential model refinements

The actuator model presented herein was developed primarily with simplicity and ease of use in mind. Once the parameters ($\beta, \eta_B, \bar{\Delta}$) have been calibrated to give an accurate pre-stress $\bar{\sigma}_0$, we expect the model to give reasonably accurate performance predictions. However, as noted in section 5.1, allowing η_B and/or $\bar{\Delta}$ to change for fixed β would produced only rough approximations of the pre-stress, since the low temperature Martensite isothermal responses are actually quite nonlinear and temperature dependent.

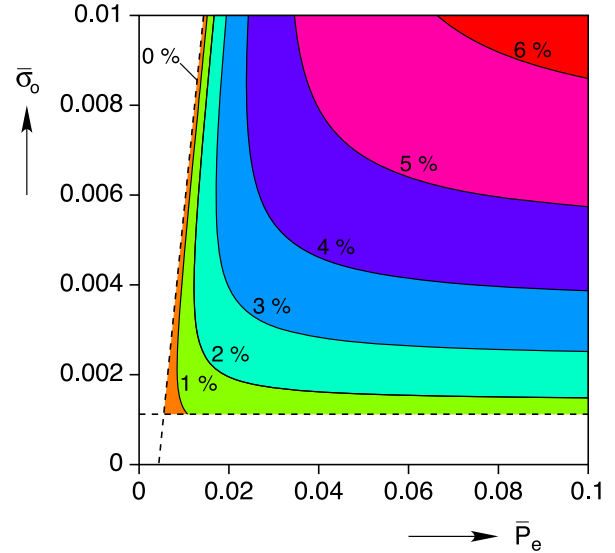


Figure 19. Contours of energy efficiency ($\mathcal{E}_{\text{out}}/\mathcal{E}_{\text{in}}$) as a function of dimensionless power (\bar{P}_e) and stress ($\bar{\sigma}_0$) for a dead load actuator.

Here, we discuss how the model could be generalized somewhat for design studies, focusing on improving the predictions of $\bar{\sigma}_0$, thereby giving more accurate calculations of the actuator stroke over the design space of ($\eta_B, \bar{\Delta}$). There are many ways this could be accomplished, for example using a more sophisticated SMA constitutive model, but instead we outline here an incremental extension of the current model that seems to be a pragmatic alternative that retains the useful analytical solutions. For design purposes, the limitation of the current actuator model stems from treating β as a material constant, independent of the actuator design and ambient temperature. We now relax this interpretation and use it as a fitting parameter for a range of different bias spring designs in order to more accurately capture the initial stress, $\bar{\sigma}_0$.

Figure 20(a) shows eight isothermal, displacement-controlled, load-unload experiments on NiTi wire specimens (as used in Chang *et al* [1]) at relatively low temperatures between -60°C and 10°C . Each as-received specimen was cooled monotonically from room temperature before the experiment, so many started in the R-phase initially and were transformed to tensile Martensite during the loading process, hence the appearance of a double ‘knee’ at low strains for the lowest temperature experiments. (Note that the appearance of the R-phase has no bearing on the quality of the actuator predictions, since the SMA operates at much higher stresses and temperatures at low strains, where the R-phase never appears.) The experiments at the highest temperatures actually show the superelastic behavior of the material. Points at the termination of the loading plateaus are shown by large dots. The overall envelope of these points was fitted by a three-parameter curve (M^+ -fit) of the form (see figure 20(b))

$$\bar{\sigma}(\varepsilon_0) = a + b e^{c \varepsilon_0}, \quad (5.27)$$

with fitting constants

$$\{a, b, c\} = \{7.9561 \times 10^{-4}, 1.7118 \times 10^{-5}, 80.622\}. \quad (5.28)$$

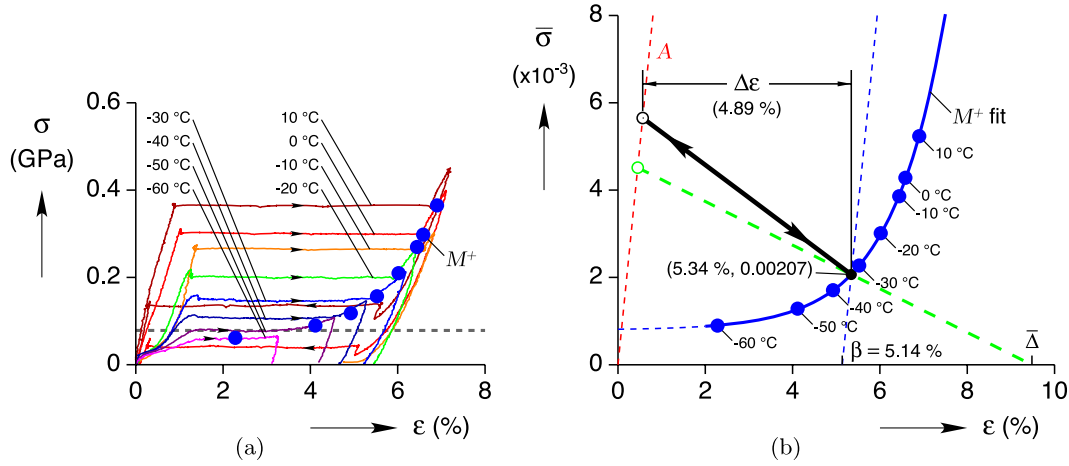


Figure 20. (a) Low temperature isothermal experiments on NiTi wire (same material as used in Chang *et al* [1]) with large solid dots showing the termination points of the loading plateaus. The dotted gray line shows the predicted value for Martensite reorientation from the constitutive model ($\sigma_M = 78.5$ MPa), which agrees with the plateau stress of the experiment at -50°C . (b) Dimensionless stress-strain paths of the actuator, $(\bar{\Delta}, \eta_B) = (0.0948, 0.05)$, for bias spring only (thick dashed line) and with external spring, $\eta_E = 0.025$ (thick solid line). The M^+ -fit curve is a fit of the loading plateau termination points, according to equation (5.27), which represents the pre-stress $\bar{\sigma}_0$ of the actuator at ‘full’ M^+ .

The curve represents the loci of the initial states of the actuation cycle for the whole class of designs. Again, we must check that the assembly takes the SMA to a point somewhere beyond the plateau of the isothermal (cold) response, i.e. pre-stress must intersect the M^+ -fit curve above the ambient temperature indicated on the curve. (This criterion is satisfied, for example, by the intersection point shown in figure 20(b), which corresponds to the plateau stress for an isothermal response near -33°C , well above the point corresponding to the ambient temperature -50°C .) Each design, i.e. combination of $(\bar{\Delta}, \eta_B)$, can be calibrated to a particular value of β , that once fixed, is associated with M^+ ($\xi_0 = 1$). For our previous values $(\bar{\Delta}, \eta_B) = (0.0948, 0.05)$ a new value of transformation strain was calculated, $\beta = 0.0514$ (as shown in figure 20(b)), somewhat smaller than used before (0.0591), but probably more accurate in light of these particular experiments. The actuator path in SMA stress-strain space is shown in figure 20(b) for the cases with a bias spring only (dashed line) and with an external spring (thick solid line) for which the predicted dimensionless stroke is 4.89%.

The calculations of fitted transformation strain (β) and the cold assembly strain (ε_0) are accomplished as follows. The intersection point (small solid dot) in figure 20(b) is defined by equating the dimensionless stresses

$$\varepsilon_0 - \beta = a + b e^{c\varepsilon_0}, \quad (5.29)$$

but now ε_0 is a function of β and $\bar{\Delta}$ according to equation (2.38). The solution is

$$\varepsilon_0 = 1 - \frac{1}{\bar{\Delta}} \left[\frac{a}{\eta_B} + \frac{1}{c} W \left(\frac{bc}{\eta_B} e^{c \left(\bar{\Delta} - \frac{a}{\eta_B} \right)} \right) \right], \quad (5.30)$$

where W is again the Lambert function, and $\{a, b, c\}$ are the fitting constants of the M^+ envelope. Accordingly, the transformation strain is found to be

$$\beta = 1 - \frac{1 + \eta_B}{\bar{\Delta}} \left[\frac{a}{\eta_B} + \frac{1}{c} W \left(\frac{bc}{\eta_B} e^{c \left(\bar{\Delta} - \frac{a}{\eta_B} \right)} \right) \right]. \quad (5.31)$$

Contours of constant ε_0 and β in the design space of $(\eta_B, \bar{\Delta})$ are shown in figures 21(a) and (b), respectively. The open circles in the figure are the values at $(\eta_B, \bar{\Delta}) = (0.05, 0.0948)$ associated with $(\varepsilon_0, \beta) = (5.34\%, 5.14\%)$. These show monotonically increasing values of ε_0 and β as either value of $(\eta_B, \bar{\Delta})$ is increased.

The dimensionless actuator stroke can be calculated from equations (5.2), (2.34), and (5.30) as

$$\Delta\varepsilon = \frac{1}{1 + \eta_B + \eta_E} \left\{ \bar{\Delta} - (1 + \eta_B) \times \left[\frac{a}{\eta_B} + \frac{1}{c} W \left(\frac{bc}{\eta_B} e^{c \left(\bar{\Delta} - \frac{a}{\eta_B} \right)} \right) \right] \right\}. \quad (5.32)$$

A contour plot of constant $\Delta\varepsilon$ in the space of (η_B, η_E) for $\bar{\Delta} = 0.0948$ is shown in figure 22. One can see that for constant η_B , increasing η_E , the actuator stroke is monotonically reduced, yet for constant η_E , increasing η_B , the actuator stroke increases then decreases. For our chosen parameters, $(\eta_B, \eta_E) = (0.05, 0.025)$, the actuator stroke is 4.78% as shown by the open circle in the plot. The stroke without an external spring is slightly larger, 4.89%. Additionally, the dotted line in the plot shows that the stroke changes in a non-monotonic manner (increasing then decreasing) as $\eta_B + \eta_E$ is increased through this point (or most others for that matter), contrary to the overly simplistic equation (5.3).

As a final comment, the original actuator model was simple, making it especially convenient for first-order design calculations. The purpose of this section was to suggest a possible refinement, reusing as much of the previous actuator model as possible, yet this was achieved at the expense of introducing some additional complexity. We have not yet investigated all the implications on the design guidelines presented in earlier parts of this section using this more refined model. That is left for future work, where we also intend to perform relevant experiments on SMA bias spring actuators using the current SMA material for direct comparison, in order

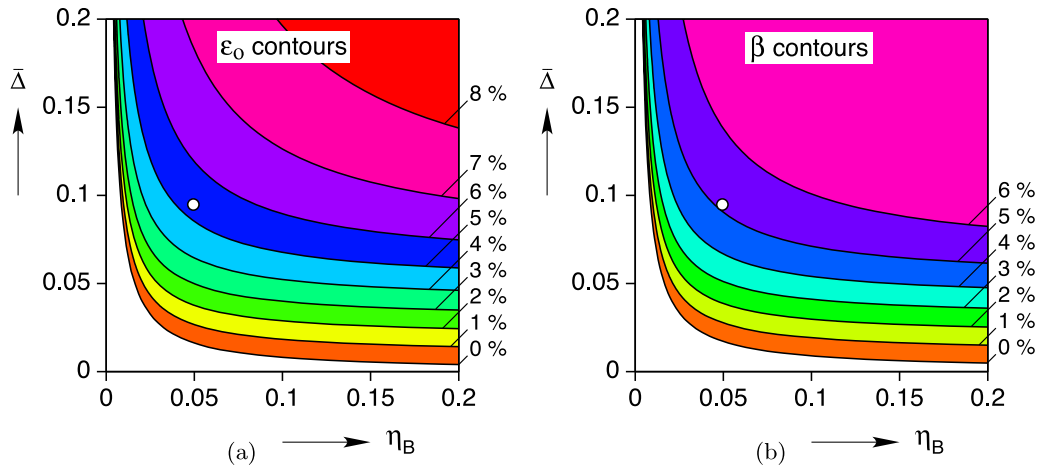


Figure 21. Contours of constant ε_0 (a) and β (b) as a function of $(\eta_B, \bar{\Delta})$. The open circles show the points $(\eta_B, \bar{\Delta}) = (0.05, 0.0948)$.

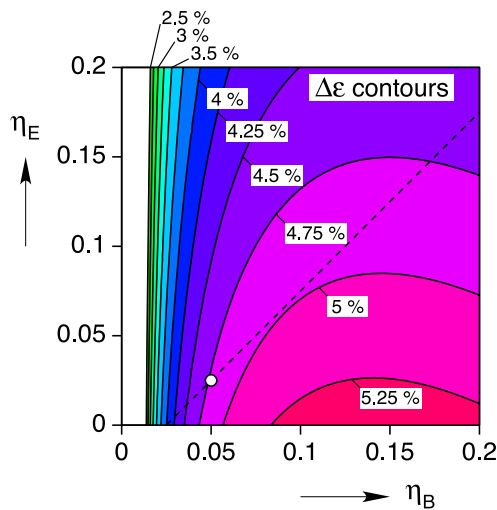


Figure 22. Contours of constant dimensionless actuator stroke, $\Delta\varepsilon$ in the space of dimensionless spring constants (η_B, η_E) for $\bar{\Delta} = 0.0948$. The open circle shows the point $(\eta_B, \eta_E) = (0.05, 0.025)$.

to validate predictions over a wide range of conditions and design parameters.

6. Summary and conclusions

A reduced-order thermodynamic model was presented for a prototype uniaxial SMA/bias spring actuator immersed in a thermal bath and actuated by Joule heating. The constitutive model, a derivative of a more complex, experimentally validated model Chang *et al* [1], was used to solve mechanical equilibrium, heat transfer, and kinetic laws, but in lumped form. A dimensionless set of equations was derived, resulting in a minimum set of dimensionless parameters governing the transient actuator response.

The actuator behavior was governed by a first-order, nonlinear ordinary differential equation (ODE) during phase transformation, rather than the usual set of partial differential equations (PDEs). Analytical solutions were then derived for

the time-dependent actuator response, during both sensible (only) heating/cooling and the combination of sensible/latent heating/cooling for martensitic transformation. To our knowledge, the latter is the first such analytical solution available for an SMA actuator system. During transformation, the evolution of strain, stress, temperature, and tensile martensite phase fraction were found in terms of the special Lambert function, given a prescribed ambient temperature and input power. The analytical solution compared favorably to a finite element simulation using the parent model [1].

The analytical nature of the transient solution to the governing ODE allowed us to study several important actuator performance indicators in terms of the governing non-dimensional parameters. Actuation time, spring size, maximum stress, minimum power, and energy efficiency were derived in closed form. By contrast, countless simulations using traditional numerical techniques to solve PDEs would be required to encompass the entire range of potential actuators. The simpler approach herein gives the designer a quick and convenient model to assess the various measures of actuator performance.

While the model presented gives a good first approximation of actuator performance, its most limiting assumptions are (1) uniform temperature field, (2) uniform strain field, and (3) the rather crude fit of the low temperature martensite reorientation response by the simple constitutive model. Assumption (1) is reasonably satisfied if the ends of the SMA wire are thermally insulated. Assumption (2) is reasonably satisfied if the SMA wire has been conditioned (pre-cycled) to eliminate the occurrence of strain localization associated with transformation fronts, yet the simple actuator model still gave satisfying results compared to the finite element simulation (that included transformation fronts), at least in terms of the gross behavior of the actuator. Assumption (3) was not an issue, provided the transformation strain parameter (β) gave a good estimate of the pre-stress in the SMA element. This pre-stress also needed to be sufficiently large to ensure tensile martensite existed in the actuator's cold state. To achieve accurate stroke predictions for a range of design parameters and ambient conditions it was necessary to reinterpret β as a design parameter, rather

than a material constant, by fitting it to the post-plateau tensile martensite responses.

We expect this simplified setting will be a useful first-order calculational tool for sizing SMA actuators, predicting their performance, and for use in design optimization software to effectively integrate such actuators into complex systems. We hope it will be a convenient tool for engineers to perform initial actuator design studies in a systematic way.

Acknowledgments

The financial support from the General Motors/University of Michigan Collaborative Research Laboratory (CRL) for Smart Materials and Structures is acknowledged with gratitude. The inspiration for this work was the result of helpful discussions and ongoing collaboration with many members of the GM/UM CRL, particularly Diann Brei (UM), Jon Luntz (UM), and Nilesh Mankame (GM). The low temperature experiments on NiTi wire shown were performed by Mark Iadicola while a doctoral student at UM (currently employed at the National Institute for Standards and Technology, NIST).

References

- [1] Chang B-C, Shaw J A and Iadicola M A 2006 Thermodynamics of shape memory alloy wire: modeling, experiments, and application *Contin. Mech. Thermodyn.* **18** 83–118
- [2] Shaw J A and Kyriakides S 1995 Thermomechanical aspects of NiTi *J. Mech. Phys. Solids* **43** 1243–81
- [3] Shaw J A and Kyriakides S 1997 On the nucleation and propagation of phase transformation fronts in a NiTi alloy *Acta Mater.* **45** 683–700
- [4] Papadopoulos P, Jung Y and Ritchie R O 2004 Constitutive modeling and numerical simulation of multivariant phase transformation in superelastic shape memory alloys *Int. J. Numer. Methods Eng.* **60** 429–60
- [5] Popov P and Lagoudas D C 2007 A 3d constitutive model for shape memory alloys incorporating pseudoelasticity and detwinning of self-accommodated martensite *Int. J. Plasticity* **23** 1679–720
- [6] Heintze O and Seelecke S 2008 A coupled thermomechanical model for shape memory alloys—from single crystal to polycrystal *Mater. Sci. Eng. A* **481/482** 389–94
- [7] Wu W, Gordaninejad F and Wirtz R A 1996 Modeling and analysis of a shape memory alloy–elastomer composite actuator *J. Intell. Mater. Syst. Struct.* **7** 441–7
- [8] Brinson L C 1993 One-dimensional constitutive behavior of shape memory alloys: thermomechanical derivation with non-constant material functions *J. Intell. Mater. Syst. Struct.* **4** 229–42
- [9] Escobar J C, Clifton R J and Yang S-Y 1999 Stress-wave-induced martensitic phase transformations in NiTi *Shock Compression of Condensed Matter; AIP Conf. Proc.* **505** 267–70
- [10] Corless R M, Gonnet G H, Hare D E G, Jeffrey D J and Knuth D E 1996 On the Lambert W function *Adv. Comput. Math.* **5** 329–59
- [11] Boyd J P 1998 Global approximations to the principal real-valued branch of the Lambert W -function *Appl. Math. Lett.* **11** 27–31
- [12] Huber J E, Fleck N A and Ashby M F 1997 The selection of mechanical actuators based on performance indices *Proc. R. Soc. A* **453** 2185–205