A three-dimensional constitutive model for the martensitic transformation in polycrystalline shape memory alloys under large deformation

L Xu1, T Baxevanis2 and D C Lagoudas1,3

1 Department of Aerospace Engineering, Texas A&M University, College Station, TX 77843, United States of America
2 Department of Mechanical Engineering, University of Houston, Houston, TX 77204, United States of America
3 Department of Materials Science & Engineering, Texas A&M University, College Station, TX 77843, United States of America

E-mail: sdf007xulei@tamu.edu

Received 4 November 2018, revised 6 March 2019
Accepted for publication 18 April 2019
Published 6 June 2019

Abstract
This work presents a three-dimensional constitutive model for the martensitic transformation in polycrystalline shape memory alloys (SMAs) under large deformation. By utilizing the logarithmic strain and rate, the model is able to account for large strains and rotations that SMA-based components may undertake, but also resolves the stress errors caused by the non-integrable objective rates that are widely used in current commercial finite element software. The model is developed through classical thermodynamic laws combined with the standard Coleman–Noll procedure. The scalar martensitic volume fraction and the second-order transformation strain tensor are chosen as the internal state variables to capture the material response exhibited by polycrystalline SMAs. A detailed implementation procedure of the proposed model is described through a user-defined material subroutine. Numerical experiments considering SMA components including a bar, a beam, a torque tube and a solid flexible structure under stress/thermally-induced phase transformations are investigated via the proposed model, and the results under cyclic loading are compared against the predictions provided by the Abaqus nonlinear solver. The development framework of the proposed model and its implementation procedure can be extended to incorporate other nonlinear phenomena exhibited by SMAs, such as transformation-induced plasticity, viscoplasticity, and damage under large deformation.

Keywords: shape memory alloys, constitutive model, large deformation, logarithmic strain, stress errors

(Some figures may appear in colour only in the online journal)

1. Introduction

Shape memory alloys (SMAs) belong to a specialized subgroup of multifunctional materials known as active materials, and are capable of recovering their pre-defined geometry when subjected to a thermal stimulus above certain temperatures. This unique property of SMAs is achieved through a solid-to-solid state diffusionless phase transformation between the high-symmetry, high-temperature austenitic phase and the low-symmetry, low-temperature martensitic phase [1]. Since the discovery of shape memory effect, SMAs have been extensively investigated as sensors and actuators towards building smart systems integrated with adaptive and morphing features [2, 3]. Recently, aerospace researchers have considered to use...
the SMA-based actuators to reconfigure the shape of a supersonic aircraft to meet the noise and efficiency requirements in response to the real-time changing ambient environment, which has the great potential to realize a commercially viable overland civil supersonic flight in the near future [4, 5].

A substantial number of constitutive theories for SMAs have been proposed so far with the majority of them based on the infinitesimal strain theory for small deformation analysis. Thorough reviews can be found from [6–19]. In general, constitutive models for SMAs can be approximately categorized into three different types: phase-field theory based models, crystal-plasticity theory based models, and classical 

\[ J_2 \]-flow theory based models. The phase-field models, in which order parameters are utilized to differentiate austenitic and martensitic phases, can track microstructure evolution, such as phases front movement, during the phase transformation process [11, 12, 20–24]. Therefore, phase-field theory based models are well suited to investigate the dynamic nucleation and phase morphology growth for martensitic phase transformation in SMAs. The enormous computational time, however, needed to solve the phase kinetic partial differential equations hinders its popularity to the extent that macroscopic structural response is concerned. In regard to crystal-plasticity theory based models [25–27], with the consideration of the effect of material microstructure (e.g. crystal orientation, texture, etc), these models are able to capture the anisotropy in material response (e.g. tension-compression asymmetry) exhibited by textured polycrystalline SMAs. Similar to phase-field methods, the complex implementation procedure of these models to incorporate the microstructure information makes them computational costly. Following the legacy of \( J_2 \)-flow theory, phenomenological SMA models have attracted attention in the engineering community where repetitive designs and optimization procedures on SMA components are needed to find target shapes. By introducing a set of internal state variables (such as volume fraction and transformation strain tensor), \( J_2 \) theory-based SMA models are able to simulate the macroscopic response of an SMA component in an efficient way. The simplicity of this model type and its well-established implementation procedure have allowed it to be widely used among real engineering applications [14, 28–34].

Constitutive models based on the infinitesimal strain theory are able to predict SMA response accurately under small deformation situations. However, it has been reported that SMAs can reversibly deform to a relatively large strain regime up to 8% [35, 36]. Also for specific boundary value problem such as fracture in SMAs, the strain levels close to the crack tip are well beyond 10% [37, 38] within the finite strain regime. In addition to such relatively large strain, SMA-based actuators (e.g. spring, beam, torque tube) may also undergo large rotation during their deployment. For example, an SMA beam component is used as a bending actuator in order to realize a morphing engine shape, where the SMA experienced large bending induced rotation [3]. Another example is the SMA tube component utilized as a torsional actuator to repeatedly rotate a deployable and retractable solar panel [39], where the SMA tube is subjected to cyclic large rotation. Combining the aforementioned two facts that SMAs may undergo large strain and rotation, it is necessary to develop a constitutive model based on a finite deformation framework to provide an accurate prediction for the response of SMAs.

Two kinematic assumptions are often employed in the finite deformation theory, i.e. the multiplicative decomposition of the deformation gradient and the additive decomposition of the strain rate tensor. In the multiplicative approach, the deformation gradient is usually decomposed into an elastic part multiplied with an inelastic part. Finite strain SMA models based on the multiplicative decomposition can be obtained from literature [34, 40–45], among which some advanced capabilities are considered. For examples, Wang and coworkers [43] presented a finite strain SMA model with the fully thermomechanically coupled feature, the consideration of coexistence of different martensitic variants, and accounting for temperature effect on the hysteresis size. In the work of Stupkiewicz and Petryk [44], they proposed a finite strain SMA model to capture the tension-compression asymmetry phenomenon. Damanpack and coworkers [45] also developed an SMA model that considers anisotropic behaviors and reorientation in SMAs at finite deformation. However, it is known that finite strain model based on additive decomposition significantly reduces the complexities of model structure compared to multiplicative models, which in return facilitates the computational efficiency of the finite strain model as a 3D design tool. Therefore, they are widely used in current available finite element (FE) softwares (e.g. Abaqus, ANSYS). However, to satisfy the principle of objectivity, additive models are required to use an objective rate in its rate form hypoelastic constitutive relation. A number of objective rates (such as Zaremba–Jaumann–Noll rate, Green–Naghdi–Dienes rate, and Truesdell rate) have been proposed to meet this goal. However, those objective rates are not essentially ‘objective’ because of their failure to integrate the rate form hypoelastic relation to yield a free energy based hyperelastic stress–strain relation [46]. As a result, spurious phenomena (e.g. shear stress oscillation, artificial stress residuals, etc) are often observed in the predicted response even for simple elastic materials through these objective rates. More details regarding this issue are addressed in appendix B.

It was not until the logarithmic rate was proposed in the literature [47–53] that the previously mentioned self-inconsistency issue related to non-integrable objective rates was resolved. As was proved in the work [47], the logarithmic rate of Eulerian logarithmic strain \( \mathbf{b} \) is exactly identical to the strain rate tensor \( \mathbf{D} \), and the logarithmic strain is the only one among all other strain measures enjoying this important property. Therefore, the finite strain models using logarithmic strain and rate are not only able to capture large strain and large rotation but also are capable of resolving the aforementioned spurious phenomena. This new development in finite deformation theory not only provides solutions to classical finite elastoplastic problems for conventional metallic materials [54, 55], but also sheds light on the finite strain model development for active materials such as SMAs. Few SMA models using an additive approach can be found from [56–62], but some of the very important SMA phase...
transformation characteristics have not been addressed among them, such as the smooth transition during the phase transformation, the stress dependent transformation strain to account for the coexistence of oriented/self-accommodated martensitic variants, and a stress dependent critical driving force to consider the effect of applied stress levels on the size of hysteretic loop. To this end, this work presents a finite strain SMA model formulation based on the additive decomposition using the logarithmic strain and rate. As a continuous development from the infinitesimal SMA model [29], the proposed model has a simple model structure and considers three very important characteristics for SMA response as its infinitesimal counterpart does. These developments combined result in improved computational efficiency and robustness for the proposed finite strain model to predict the SMA response at large deformation, without introducing additional intermediate state variables, such as Mandel stress, that are utilized in the multiplicative models. It is noted that the primary focus of this work is mainly on the formulation of a finite strain SMA model, rather than the development of a constitutive model that can capture the full complexities of the SMA thermomechanical deformation. Thus, tension-compression, latent-heat effects, reorientation between oriented and self-accommodated martensitic variants, cyclic evolution features (transformation-induced plasticity, two-way shape memory effect at stress-free conditions) are not included here for simplicity. Moreover, this work carefully examines the artifical stress errors caused by using other non-integrable objective rates in current commercial FE packages. The capability of the proposed model to eliminate such stress errors shows significant importance for the analysis of SMA-based actuators, e.g. SMA beam and SMA torque tube subjected to cyclic large deformation.

This paper is organized as follows. Section 2 presents the kinematic preliminaries. Section 3 concentrates on the model development based on the logarithmic strain and logarithmic rate. The derivation of the consistent tangent stiffness matrix and the consistent thermal matrix are also provided. In section 4, the detailed implementation procedure for the proposed model is described by using a user-defined material subroutine (UMAT) through the FE software Abaqus. Numerical examples are studied to demonstrate the capability of the proposed model in section 5. Conclusions are presented in section 6. A detailed calibration procedure for the material parameters used in this model is also provided in the appendix C.

2. Preliminaries

2.1. Kinematics

Let material point \( P \) from body \( B \) be defined by a position vector \( \mathbf{X} \) in the reference (undeformed) configuration at time \( t_0 \), and let vector \( \mathbf{x} \) represent the position vector of that material point in the current (deformed) configuration at time \( t \). Therefore, the deformation process of point \( P \) between the reference configuration and the current configuration can be defined through the well-known deformation gradient tensor \( \mathbf{F}(\mathbf{x}, t) \):

\[
\mathbf{F}(\mathbf{x}, t) = \frac{\partial \mathbf{x}}{\partial \mathbf{X}}
\]

(1)

and the velocity field \( \mathbf{v} \) can be defined as

\[
\mathbf{v} = \frac{d\mathbf{x}}{dt} = \mathbf{x}
\]

(2)

based on the velocity field \( \mathbf{v} \), the velocity gradient \( \mathbf{L} \) can be derived as

\[
\mathbf{L} = \frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \mathbf{F} \mathbf{F}^{-1}
\]

(3)

the following polar decomposition equation for deformation gradient \( \mathbf{F} \) is well known

\[
\mathbf{F} = \mathbf{RU} = \mathbf{VR}
\]

(4)

where \( \mathbf{R} \) is the rotation tensor, \( \mathbf{U} \) and \( \mathbf{V} \) are the right (or Lagrangian) and the left (or Eulerian) stretch tensors, respectively, by which the right Cauchy–Green tensor \( \mathbf{C} \) and the left Cauchy–Green tensor \( \mathbf{B} \) can be obtained, as follows

\[
\mathbf{C} = \mathbf{F}^T \mathbf{F} = \mathbf{U}^2,
\]

(5)

\[
\mathbf{B} = \mathbf{F}^{-T} \mathbf{F} = \mathbf{V}^2,
\]

(6)

where \( \mathbf{I} \) is the second order identity tensor. The logarithmic strain (also called Hencky or true strain) of Lagrangian type \( \mathbf{H} \) and Eulerian type \( \mathbf{h} \) can thus be defined as

\[
\mathbf{H} = \frac{1}{2} \ln(\mathbf{C}) = \ln(\mathbf{U}),
\]

(7)

\[
\mathbf{h} = \frac{1}{2} \ln(\mathbf{B}) = \ln(\mathbf{V}).
\]

(8)

It is also well known that the velocity gradient \( \mathbf{L} \) can be additively decomposed into a symmetric part, the strain rate tensor \( \mathbf{D} \), and an anti-symmetric part, the spin tensor \( \mathbf{W} \)

\[
\mathbf{L} = \mathbf{D} + \mathbf{W}, \quad \mathbf{D} = \frac{1}{2} (\mathbf{L} + \mathbf{L}^T), \quad \mathbf{W} = \frac{1}{2} (\mathbf{L} - \mathbf{L}^T).
\]

(9)

2.2. Logarithmic strain, logarithmic rate and logarithmic spin

As was mentioned in section 1, two widely accepted kinematic assumptions, i.e. the multiplicative decomposition of deformation gradient \( \mathbf{F} \) and the additive decomposition of the strain rate tensor \( \mathbf{D} \), are usually considered in finite deformation theory. The multiplicative models use a hyperelastic constitutive relation while a rate form hypoelastic constitutive equation is usually adopted for additive models. The rate form hypoelastic constitutive theory using objective rates has been criticized for its non-integrability because it can not well define an essential elastic material behavior [63], this includes many well known objective rates such as Zaremba–Jaumann rate, Green–Naghdi rate, Truesdell rate, etc [46].
The aforementioned problems about objective rates were solved in the work by Xiao et al. [46–48], Bruhns et al. [49–51] and Meyers et al. [52, 53], where they proved that the logarithmic rate of the Eulerian logarithmic strain \( \mathbf{h} \) is identical with the strain rate tensor \( \mathbf{D} \), by which a hypoelastic model can be exactly integrated to a hyperelastic finite strain model [47]. This unique relationship between logarithmic strain \( \mathbf{h} \) and the strain rate tensor \( \mathbf{D} \) is expressed as follows

\[
\mathbf{h}^{\log} = \dot{\mathbf{h}} + \mathbf{h} \mathbf{Q}^{\log} - \mathbf{Q}^{\log} \mathbf{h} = \mathbf{D},
\]

(10)

where \( \mathbf{Q}^{\log} \) is called the logarithmic spin introduced by [47] defined as

\[
\mathbf{Q}^{\log} = \mathbf{W} + \sum_{i=1}^{n} \left( \frac{1}{1 - (\lambda_i/\lambda_j)} + \frac{2}{\ln(\lambda_i/\lambda_j)} \right) \mathbf{b}_i \mathbf{b}_j
\]

(11)

in which \( \lambda_i, \lambda_j (i, j = 1, 2, 3) \) are the eigenvalues of left Cauchy–Green tensor \( \mathbf{B} \) and \( \mathbf{b}_i, \mathbf{b}_j \) are the corresponding subordinate eigenprojections. As long as the logarithmic spin tensor \( \mathbf{Q}^{\log} \) is defined, the second order rotation tensor \( \mathbf{R}^{\log} \), associated with \( \mathbf{Q}^{\log} \), can be determined through the following differential equation (12). In general cases, the initial condition is assumed as \( \mathbf{R}^{\log}{|}_{t=0} = \mathbf{I} \)

\[
\mathbf{Q}^{\log} = \mathbf{R}^{\log}(\mathbf{R}^{\log})^T
\]

follow the corotational integration definition from [64], and assume the initial conditions \( \mathbf{h}{|}_{t=0} = \mathbf{0} \), equation (10) yields the total logarithmic strain \( \mathbf{h} \) after the logarithmic corotational integration

\[
\mathbf{h} = \int_{t_{\text{corot}}} \mathbf{D} \, dt = \left( \mathbf{R}^{\log} \right)^T \left( \int_{0}^{t} \mathbf{R}^{\log} \dot{\mathbf{D}} \left( \mathbf{R}^{\log} \right)^T \, dt' \right) \mathbf{R}^{\log}.
\]

(13)

2.3. Additive decomposition of logarithmic strain

The kinematic assumption starts with the additive decomposition of the strain rate tensor \( \mathbf{D} \) into an elastic part \( \mathbf{D}^e \) plus a transformation part \( \mathbf{D}^\tau \)

\[
\mathbf{D} = \mathbf{D}^e + \mathbf{D}^\tau.
\]

(14)

The elastic strain rate part \( \mathbf{D}^e \) and the transformation strain rate part \( \mathbf{D}^\tau \) in equation (14) can be rewritten as \( \mathbf{h}^{e,\log} \) and \( \mathbf{h}^{\tau,\log} \) by virtue of the relation in equation (10) respectively

\[
\mathbf{h}^{e,\log} = \mathbf{D}^e; \quad \mathbf{h}^{\tau,\log} = \mathbf{D}^\tau.
\]

(15)

By combining equations (10), (14) and (15), the following equation can be obtained

\[
\mathbf{h}^{\log} = \mathbf{h}^{e,\log} + \mathbf{h}^{\tau,\log}.
\]

(16)

Similar to the results obtained from equation (13), equation (16) can yield the following relation after applying the logarithmic corotational integration

\[
\mathbf{h}^e = \int_{t_{\text{corot}}} \mathbf{D}^e \, dt = \left( \mathbf{R}^{\log} \right)^T \left( \int_{0}^{t} \mathbf{R}^{\log} \dot{\mathbf{D}}^e \left( \mathbf{R}^{\log} \right)^T \, dt' \right) \mathbf{R}^{\log},
\]

(17a)

\[
\mathbf{h}^\tau = \int_{t_{\text{corot}}} \mathbf{D}^\tau \, dt = \left( \mathbf{R}^{\log} \right)^T \left( \int_{0}^{t} \mathbf{R}^{\log} \dot{\mathbf{D}}^\tau \left( \mathbf{R}^{\log} \right)^T \, dt' \right) \mathbf{R}^{\log}.
\]

(17b)

Based on the additive decomposition on the strain rate tensor, combing equations (14), (16), (17a) and (17b), the following additive decomposition on the total logarithmic strain \( \mathbf{h} \) can be achieved, i.e. the total logarithmic strain \( \mathbf{h} \) can be additively split into an elastic strain like part \( \mathbf{h}^e \) plus a transformation strain like part \( \mathbf{h}^\tau \)

\[
\mathbf{h} = \mathbf{h}^e + \mathbf{h}^\tau.
\]

(18)

3. Model formulation

3.1. Thermodynamic framework

The Gibbs free energy potential \( G \) is defined to be a continuous function dependent on Kirchhoff stress tensor \( \mathbf{\tau} \), Eulerian logarithmic strain \( \mathbf{h} \), temperature \( T \) and a set of internal state variables \( \mathbf{\Upsilon} \)

\[
G(\mathbf{\tau}, \mathbf{h}, T, \mathbf{\Upsilon}) = u - \frac{1}{\rho_0} \mathbf{\tau} : \mathbf{h} - s T,
\]

(19)

where \( \rho_0 \) is the density in the reference configuration, \( s \) and \( u \) are the specific entropy and internal energy respectively. From the 2nd law of thermodynamics, the dissipation energy \( \mathcal{D} \) can be written in the form of Clausius–Duhem inequality

\[
\mathcal{D} = \mathbf{\tau} : \mathbf{D} - \rho_0 (u - s T) \geq 0.
\]

(20)

The logarithmic rate of the Gibbs free energy is taken in equation (19). Note that a scalar subjected to an objective rate equals to its conventional time rate, the following equation is derived. An circle hat denotes the logarithmic rate in the following text for brevity

\[
\dot{G}^{\log} = \dot{u} - \frac{1}{\rho_0} \dot{\mathbf{\tau}}^{\log}; \quad \dot{\mathbf{h}}^{\log} = \frac{1}{\rho_0} \dot{\mathbf{\tau}}^{\log} - s \dot{T} - s T
\]

(21)

the following equation is obtained after rearrangement of equation (21)

\[
\dot{u} - s \dot{T} = \dot{G} + \frac{1}{\rho_0} \dot{\mathbf{\tau}}^{\log}: \mathbf{h} + \frac{1}{\rho_0} \dot{\mathbf{\tau}}: \mathbf{h} + s \dot{T}.
\]

(22)

\footnote{The relationship between Kirchhoff stress \( \mathbf{\tau} \) and Cauchy stress \( \mathbf{\sigma} \) is \( \mathbf{\tau} = J \mathbf{\sigma}, \) where \( J = \det(\mathbf{F}) \). Assuming phase transformation to be volume preserving, \( J \) is approximately equivalent to 1, so \( \mathbf{\tau} \approx \mathbf{\sigma} \). Kirchhoff stress \( \mathbf{\tau} \) and Eulerian logarithmic strain \( \mathbf{h} \), called an energetic conjugate pair [65], are usually paired up in the formation of free energy potentials.}
Substitute equation (22) into Clausius–Duhem inequality (20), the dissipation energy is rearranged as the following

\[ \mathcal{D} = -\rho_0 G - \rho_0 \delta \dot{T} - \hat{\mathbf{r}} \cdot \mathbf{h} \geq 0. \]  

(23)

Recall that the Gibbs free energy \( G(\mathbf{r}, T, \mathbf{Y}) \) is a continuous function, chain rule differentiation can be applied on the Gibbs free energy with respect to its independent state variables (i.e. Kirchhoff stress \( \mathbf{r} \), temperature \( T \) and internal state variables \( \mathbf{Y} \)), which gives

\[ \dot{G} = \frac{\partial G}{\partial \mathbf{r}} \cdot \dot{\mathbf{r}} + \frac{\partial G}{\partial T} \cdot \dot{T} + \frac{\partial G}{\partial \mathbf{Y}} \cdot \dot{\mathbf{Y}}. \]  

(24)

Substitute equation (24) into equation (23), the following expression for the dissipation energy \( \mathcal{D} \) is acquired

\[ \mathcal{D} = -\left( \rho_0 \frac{\partial G}{\partial \mathbf{r}} + \mathbf{h} \right) \cdot \dot{\mathbf{r}} - \left( \rho_0 \frac{\partial G}{\partial T} + s \right) \cdot \dot{T} - \rho_0 \frac{\partial G}{\partial \mathbf{Y}} \cdot \dot{\mathbf{Y}} \geq 0. \]  

(25)

Following the standard Coleman–Noll procedure, all admissible values for \( \mathbf{r}, T \) and \( \mathbf{Y} \) have to comply with the dissipation inequality (25) regardless of thermodynamic paths, thereby the constitutive relationships between stress and strain, entropy and temperature can be inferred as

\[ \mathbf{h} = -\rho_0 \frac{\partial G}{\partial \mathbf{r}}, \]  

(26)

\[ s = -\rho_0 \frac{\partial G}{\partial T}. \]  

(27)

Substitute equations (26) and (27) into equation (25), the following reduced form of the dissipation inequality is acquired

\[ -\rho_0 \frac{\partial G}{\partial \mathbf{Y}} \cdot \dot{\mathbf{Y}} \geq 0. \]  

(28)

3.2. Constitutive modeling for SMAs

3.2.1. Thermodynamic potential. The formulation of the proposed model is based on the thermodynamic framework presented in section 3 and the early SMA model developed by Lagoudas and coworkers \([7, 29]\) for small deformation analysis. The model is able to predict the pseudoelastic (isothermal) and actuation (isobaric) response under large deformation including both large strain and large rotation. A quadratic Gibbs free energy potential \( G \) is introduced in equation (29), in which Kirchhoff stress tensor \( \mathbf{r} \) and temperature \( T \) are chosen as the independent state variables. The martensitic volume fraction \( \xi \) and the second order transformation strain tensor \( \mathbf{h}^{\alpha} \) are chosen as internal state variables \( \mathbf{Y} = \{ \xi, \mathbf{h}^{\alpha} \} \) to capture the material response exhibited by polycrystalline SMAs. The Gibbs free energy potential \( G \) is employed as follows

\[ G = -\frac{1}{2\rho_0} \mathbf{r} : \mathbf{S} \mathbf{r} - \frac{1}{\rho_0} \dot{\mathbf{r}} : \mathbf{h}^{\alpha} + c \left[ (T - T_0) - T \ln \left( \frac{T}{T_0} \right) \right] + s_0 (T - T_0) + u_0 + \frac{1}{\rho_0} f(\xi), \]  

(29)

where \( S \) is the effective compliance tensor calculated by equation (30), \( \mathbf{S}^A \) is the compliance tensor for austenitic phase while \( \mathbf{S}^M \) is for martensitic phase, and \( \Delta S \) is the phase difference for the compliance tensor. The effective stiffness tensor \( C \) can be gained by taking the inverse of the effective compliance tensor, i.e. \( C = \mathbf{S}^{-1} \). \( \alpha \) is the second order thermoelastic expansion tensor, \( c \) is the effective specific heat, \( s_0 \) and \( u_0 \) are the effective specific entropy and effective specific internal energy at the reference state. All the aforementioned variables are determined from equation (31)–(34). \( T \) represents the temperature at current state while \( T_0 \) is the temperature at reference state

\[ S(\xi) = S^A + \xi (S^M - S^A) = S^A + \xi \Delta S, \]  

(30)

\[ \alpha(\xi) = \alpha^A + \xi (\alpha^M - \alpha^A) = \alpha^A + \xi \Delta \alpha, \]  

(31)

\[ c(\xi) = c^A + \xi (c^M - c^A) = c^A + \xi \Delta c, \]  

(32)

\[ s_0(\xi) = s_0^A + \xi (s_0^M - s_0^A) = s_0^A + \xi \Delta s_0, \]  

(33)

\[ u_0(\xi) = u_0^A + \xi (u_0^M - u_0^A) = u_0^A + \xi \Delta u_0. \]  

(34)

A smooth hardening function \( f(\xi) \) is proposed in equation (35) to account for the hardening effects in polycrystalline SMAs, such as the plastic strain accumulation after the training procedure, imperfections located at the grain boundary, and nano-precipitates hardening effects, etc \([1]\), where three additional intermediate material parameters \( a_1, a_2, a_3 \) and four curve fitting parameters \( n_1, n_2, n_3, n_4 \) are introduced to better treat the smooth transition behaviors at the initiation and completion of phase transformation

\[ f(\xi) = \begin{cases} \frac{1}{2} a_1 (\xi + \xi_{\alpha+1}^{\alpha+1} (1 - \xi_{\alpha+1}^{\alpha+1})/n_{\alpha+1}) + a_3 \xi, & \xi > 0, \\ \
\frac{1}{2} a_1 (\xi + \xi_{\alpha}^{\alpha} (1 - \xi_{\alpha}^{\alpha})/n_{\alpha}) - a_3 \xi, & \xi < 0. \end{cases} \]  

(35)

Following the standard Coleman–Noll procedure described in section 3, the explicit form for constitutive relation (26) between stress and strain is derived as follows. Note that the nonlinearity in this constitutive relation is implied by the transformation strain \( \mathbf{h}^{\alpha} \)

\[ \mathbf{h} = -\rho_0 \frac{\partial G}{\partial \mathbf{r}} = \mathbf{S} \mathbf{r} + \alpha (T - T_0) + \mathbf{h}^{\alpha} \]  

(36)

the explicit form for constitutive relation (27) between entropy \( s \) and temperature \( T \) can also be derived as

\[ s = -\rho_0 \frac{\partial G}{\partial T} = \frac{1}{\rho_0} \mathbf{r} : \alpha + c \ln \left( \frac{T}{T_0} \right) + s_0 \]  

(37)
the reduced form of the dissipation inequality (28) can be rewritten in terms of the chosen internal state variables \( \mathbf{Y} = \{ \xi, h^r \} \) as
\[
-\rho_0 \frac{\partial G}{\partial h^r} \dot{h}^r - \rho_0 \frac{\partial G}{\partial \xi} \dot{\xi} \geq 0. \tag{38}
\]

### 3.2.2. Evolution equation of internal state variables

The evolution equation for the internal state variables \( \mathbf{Y} = \{ \xi, h^r \} \) is presented here. It is proposed that the logarithmic rate of the transformation strain \( h^r \) is proportional to the rate change of the martensitic volume fraction \( \xi \). This proportional evolution rule is adopted by following the principle of maximum dissipation such that among all the admissible thermodynamic paths, the one dissipating the most energy is chosen during the SMAs phase transformation process [66]. The idea of maximum dissipation for inelastic materials is not new, it was also widely employed for plastic deformed materials to derive the associated flow rule [67]. It is worth pointing out that the rate applied on the transformation strain is the logarithmic rate rather than the conventional time rate. The explicit evolution rule is as follows
\[
\dot{h}^r = \Lambda \dot{\xi}, \quad \Lambda = \begin{cases} \Lambda^\text{fwd}, & \dot{\xi} > 0, \\ \Lambda^\text{rev}, & \dot{\xi} < 0, \end{cases} \tag{39}
\]

where \( \Lambda^\text{fwd} \) is called the forward transformation direction tensor and \( \Lambda^\text{rev} \) is called the reverse transformation direction tensor. They are defined as
\[
\Lambda^\text{fwd} = \frac{3}{2} H^{\text{cur}} \frac{\tau'\tau'}{\tau}, \quad \Lambda^\text{rev} = \frac{h^{r-r}}{\xi} \tag{40}
\]

in which, \( \tau' \) is the deviatoric part of Kirchhoff stress tensor calculated by \( \tau' = \tau - \frac{1}{3} tr(\tau) \mathbf{1} \), \( \mathbf{1} \) is the second order identity tensor. The effective Mises equivalent stress is given by \( \dot{\tau} = \sqrt{2/3} \tau' \). \( h^{r-r} \) and \( \xi' \) represent the transformation strain value and martensitic volume fraction at the reverse transformation starting point. It is common among available SMA models that the magnitude of the inelastic recoverable transformation strain is the same for full transformation under any applied stress levels. This is true when the stress levels is high enough to generate maximum oriented martensitic variants. However, if the applied stress level is not sufficient high, self-accommodated martensitic variants will be generated. This renders the value of transformation strain less than it is in the high stress level case (i.e. the stress dependency of the magnitude of the transformation strain). Therefore, an exponential function \( H^{\text{cur}} \) dependent on current stress levels is introduced to calculate the current transformation strain as shown in equation (41), where \( H^{\text{max}} \) is the maximum (or saturated) transformation strain and \( k_i \) is a curve fitting material parameter
\[
H^{\text{cur}}(\tau) = H^{\text{max}}(1 - e^{-k_i \tau}). \tag{41}
\]

### 3.2.3. Transformation function

The objective in this part is to define a proper transformation criterion to determine the occurrence of the phase transformation. Recall the reduced form for dissipation energy is given by inequality (38) and the relation between \( h^r \) and \( \xi \) is defined through evolution equation (39). Substituting evolution equation (39) into reduced form dissipation inequality (38), the following equation is obtained
\[
(\tau: \Lambda - \rho_0 \frac{\partial G}{\partial \xi}) \dot{\xi} = \pi \dot{\xi} \geq 0 \tag{42}
\]

the above equation implies that all the dissipation energy is directly a result of the change in the martensitic volume fraction. Based upon this, a scalar variable \( \pi \), called the thermodynamic driving force conjugated to the martensitic volume fraction \( \xi \), can thus be defined. Substitution of Gibbs free energy potential \( G \) in equation (29) into equation (42) yields the explicit expression for \( \pi \) as follows
\[
\pi(\tau, T, T, \xi) = \tau: \Lambda + \frac{1}{2} \tau T \Delta S_T + \tau: \Delta \alpha(T - T_0) + \rho_0 \Delta T_0 \tau - \rho_0 \Delta u_0 - \frac{\partial f(\xi)}{\partial \xi}. \tag{43}
\]

where \( \Delta S_T, \Delta \alpha, \Delta T_0, \) and \( \Delta u_0 \) are the phase differences on compliance tensor, thermal expansion tensor, specific heat, reference entropy and reference internal energy, respectively. It can be observed that the thermodynamic driving force \( \pi \) is a function of Kirchhoff stress \( \tau \), temperature \( T \) and martensitic volume fraction \( \xi \). This indicates that the phase transformation process can be activated by two independent sources, namely either the stress or temperature, which correlates quite well with the experimentally observed stress-induced and thermally-induced phase transformations in SMAs. To proceed to the goal of defining a transformation criteria, it is assumed that whenever the thermodynamic driving force \( \pi \) reaches a critical value \( Y(\sim Y) \), the forward (reverse) phase transformation takes place. Therefore a transformation function \( \Phi \) can be defined as the transformation criteria to determine the transformation occurrence as follows
\[
\Phi = \begin{cases} \pi - Y, & \dot{\xi} > 0, \\ -\pi - Y, & \dot{\xi} < 0. \end{cases} \tag{44}
\]

In the infinitesimal strain theory based SMA model [29], a reference critical value \( Y_0 \) and an additional parameter \( D \) were introduced into \( Y \), through which the thermodynamical critical value \( Y \) becomes a function dependent on applied stress levels, see in equation (45). Such treatment let the model consider the effect of applied stress levels on the size of hysteresis loop. This capability is provided through capturing the different slopes \( C_A, C_M \) in the effective stress-temperature phase diagram. The explicit derivation is provided from equation (B.7) to equation (B.11) at the model
3.3. Consistent tangent stiffness and thermal matrix

In this section, a detailed derivation of the consistent tangent stiffness matrix and the thermal matrix is provided to complete the proposed model. For most typical displacement-based FE softwares, such as Abaqus, the consistent tangent matrices are often required to be provided in the UMAT so that the FE solver can achieve a fast convergence for the global equilibrium equations. Normally, consistent tangent matrices can be expressed in the rate form shown in equation (47), where \( L \) is called the consistent tangent stiffness matrix and \( \Theta \) is the consistent thermal matrix

\[
\dot{\tau} = L \dot{h} + \Theta \dot{T}
\]

(47)

applying the logarithmic rate on constitutive equation (36) yields

\[
\dot{\tau} = C [\dot{h} - \alpha \dot{T} - (\Delta \alpha (T - T_0) + \Lambda) \dot{\xi}]
\]

(48)

taking chain rule differentiation on the transformation function equation (44) gives

\[
\dot{\Phi} = \partial_\tau \Phi : \dot{\tau} + \partial_\rho \Phi : \dot{\rho} = 0
\]

(49)

substituting equation (48) back into equation (49) to eliminate \( \dot{\tau} \) and solving it for \( \dot{\xi} \), the following expression for \( \dot{\xi} \) can be obtained

\[
\dot{\xi} = - \frac{\partial_\rho \Phi : \dot{\rho} - \partial_\tau \Phi : \dot{\tau} : C \alpha}{\partial_\rho \phi - \partial_\rho \Phi : C (\Delta \alpha (T - T_0))}
\]

(50)

substituting equation (50) back into the rate form constitutive equation (48) to eliminate \( \dot{\xi} \), considering the phase difference of thermal expansion coefficients can be ignored for martensite and austenite phase, the final explicit expression corresponding to equation (47) can be obtained as follows

\[
\dot{\tau} = \left[ C + \left( \frac{\partial_\rho \Phi - \partial_\tau \Phi : C (\Delta \alpha (T - T_0))}{\partial_\rho \phi - \partial_\rho \Phi : C (\Delta \alpha (T - T_0))} \right) \dot{h} + \left( -C \alpha + \frac{\partial_\rho \Phi - \partial_\tau \Phi : C (\Delta \alpha (T - T_0))}{\partial_\rho \phi - \partial_\rho \Phi : C (\Delta \alpha (T - T_0))} \right) \dot{T} \right]
\]

(51)

in which the consistent tangent stiffness matrix \( L \) is

\[
L = C + \frac{\left( C (\Delta \alpha (T - T_0) + \Lambda) \right) \otimes \left( \frac{\partial_\rho \Phi - \partial_\tau \Phi}{\partial_\rho \phi - \partial_\rho \Phi : C (\Delta \alpha (T - T_0))} \right)}{\partial_\rho \phi - \partial_\rho \Phi : C (\Delta \alpha (T - T_0) + \Lambda)}
\]

(52)

and the consistent thermal matrix \( \Theta \) is

\[
\Theta = -C \alpha + \frac{C (\Delta \alpha (T - T_0) + \Lambda) \otimes \left( \frac{\partial_\rho \Phi - \partial_\tau \Phi}{\partial_\rho \phi - \partial_\rho \Phi : C (\Delta \alpha (T - T_0))} \right)}{\partial_\rho \phi - \partial_\rho \Phi : C (\Delta \alpha (T - T_0) + \Lambda)}
\]

(53)

In order to fully determine the explicit values for \( L \) and \( \Theta \) during the implementation section for the proposed model, the explicit expressions of the following terms \( \partial_\rho \Phi, \partial_\rho \Phi, \partial_\tau \Phi \) used in above equations are derived in appendix E.

4. Numerical implementation

This section focuses on the implementation of the proposed model within FE solvers to solve boundary value problems (BVPs). The implementation flowchart is shown in figure 1. While typically stress and strain information are provided from FE solver, the initial input information used in this model are only the temperatures \( T_n, \Delta T_n \) and deformation gradients at current step \( \mathbf{F}_n \) and next step \( \mathbf{F}_{n+1} \). The reason for using only these information is that other tensorial variables have been rotated by the FE solver before they are used as inputs, in which the rotation tensor is calculated based on the other non-integrable objective rates. This consequently leads to the artificial stress errors described in section 1. During the implementation for the proposed model, a pre-calculation and a rotation procedure are employed before calling the main UMAT subroutine. In the pre-calculation procedure, the logarithmic strain at current step \( \mathbf{h}_n \) and next step \( \mathbf{h}_{n+1} \) are calculated based on \( \mathbf{F}_n \) and \( \mathbf{F}_{n+1} \). The incremental rotation tensor \( \Delta \mathbf{R}^{\text{log}}_n \) based on the logarithmic rate can be calculated by using the exponential map scheme [54, 63, 68]. In the rotation procedure, tensorial variables including \( \mathbf{h}_n, \mathbf{h}^{\text{rot}}_n \) and \( \Lambda_n \) are rotated from step \( n \) configuration to the new configuration at step \( n+1 \) by using the obtained \( \Delta \mathbf{R}^{\text{log}}_n \), thus, the so-called principle of objectivity is preserved.

The rest of implementation procedure consists of two steps, the first step is called the thermoelastic predictor and the second step is called the transformation corrector. During the initialization of thermoelastic step, the total strain \( \mathbf{h}_{n+1} \) and current temperature \( T_{n+1} = T_n + \Delta T_n \) are provided. The initial internal state variables \( \mathbf{Y}^{(0)}_{n+1} \) are assumed to be the same as \( \mathbf{Y}_n \) for the initial Kuhn–Tucker consistency checking, i.e. \( \Phi^{(0)}_{n+1} \leq 0 \). If the initial value of \( \Phi^{(0)}_{n+1} \) satisfies the consistency checking, the \( n + 1 \) step is detected as a thermoelastic response and the UMAT returns to global FE solver for next increment. In case of consistency condition violated, the second step called the transformation corrector is activated to find the updated internal state variables \( \mathbf{Y}^{(1)}_{n+1} \) in order to regain the Kuhn–Tucker consistency. A detailed
summary for the implementation procedure is listed in the table 1.

### 4.1. Thermoelastic prediction

Take the \((n+1)\)th step as an example to go through the thermoelastic prediction process. The total strain tensor \(h_{n+1}\) and the temperature \(T_{n+1}\) are provided from Pre-calculation procedure, and the initial internal state variables \(\gamma_{n+1}^{(0)}\) are assumed the same as \(\gamma_n\):

\[
h_{n+1}^{tr(0)} = h_{n+1}^{tr}; \xi_{n+1}^{(0)} = \xi_{n}^{tr}
\]  

(54)

Based on equation (54), the initial guess for Kirchhoff stress \(\tau_{n+1}^{(0)}\) can be calculated through the constitutive equation (55). Here the integer in the upper parenthesis represents how many iterations have been performed during the transformation correction procedure, and integer zero means that this step is just an initial guess in the thermoelastic procedure. The initial calculation for stress \(\tau_{n+1}^{(0)}\) can be obtained

\[
\tau_{n+1}^{(0)} = C_n\left[h_{n+1}^{tr} - h_{n+1}^{tr(0)} - \alpha_n(T_{n+1} - T_0)\right].
\]  

(55)

After the calculation of \(\tau_{n+1}^{(0)}\), the value of transformation function \(\Phi_{n+1}^{(0)}\) can be evaluated based on equations (43) and (44) for the initial Kuhn–Tucker consistency checking

\[
\Phi_{n+1}^{(0)} = \Phi(\tau_{n+1}^{(0)}; T_{n+1}; \gamma_{n+1}^{(0)}).
\]  

(56)

### 4.1. Thermoelastic prediction

The implementation procedure for the proposed finite strain SMA model.

1. **Initialization**
   - Conduct pre-calculation and rotation procedures.
   - \(k = 0; \xi_{n+1}^{(0)} = \xi_n; h_{n+1}^{tr(0)} = h_{n+1}^{tr}\).

2. **Thermoelastic Predictor**
   - \(\tau_{n+1} = C_n\left[h_{n+1}^{tr} - h_{n+1}^{tr(0)} - \alpha_n(T_{n+1} - T_0)\right]\)
   - Calculate \(\Phi_{n+1}^{(0)}\).
   - IF \(\Phi_{n+1}^{(0)} \leq tol\), GOTO 4 (thermoelastic response).
   - IF \(\Phi_{n+1}^{(0)} > tol\), GOTO 3 (transformation happens).

3. **Transformation Corrector**
   - Calculate residual matrix \(R_{n+1}^{(k)} = h_{n+1}^{tr(1)} - h_{n+1}^{tr(0)} + \Lambda_{n+1}^{(k)}(\xi_{n+1}^{(k)} - \xi_{n}^{tr})\)
   - Calculate consistent tangent stiffness and thermal matrix \(\Theta_{n+1}\).
   - Perform the Newton-Raphson iterations in equation (60).
   - Update variables \(\xi_{n+1}^{(k+1)}; h_{n+1}^{tr(1)}; S_{n+1}^{(k+1)}\)

4. Calculate consistent stiffness matrix \(\mathcal{L}\) and thermal matrix \(\Theta\).
   - \(\mathcal{L} = \mathcal{C} + \left[\mathcal{C}(\Delta S\mathcal{T} + \mathcal{A})\right] \otimes \left[\partial_{\xi^n} \Phi - \partial_{h^n} \Phi \cdot \mathcal{C}(\Delta S\mathcal{T} + \mathcal{A})\right]\)
   - \(\Theta = -\partial_{h^n} \Phi \cdot \mathcal{C}(\Delta S\mathcal{T} + \mathcal{A})\)

5. Exit UMAT and proceed to the global FE solver for the next increment.
If the calculated value of transformation function \( \Phi_{n+1}^{(0)} \) remains under the transformation surface (i.e. \( \Phi_{n+1}^{(0)} \leq \text{tol} \), ‘tol’ is usually set to be \( 10^{-6} \)), step \( n + 1 \) is detected as a thermoelastic response. Therefore the values of current state variables \( \tau_{n+1}^{(0)} \) and \( \Sigma_{n+1}^{(0)} \) are accepted as correct and the UMAT proceeds to the global FE solver for the next increment. In case the transformation surface is violated (i.e. \( \Phi_{n+1}^{(0)} > \text{tol} \)), the transformation corrector step is activated to find the updated state variables until the consistency equation (46) is preserved.

4.2. Transformation correction

This part addresses the iterative procedures required for the transformation corrector to restore the Kuhn–Tucker consistency. In general, the transformation corrector is nothing but a set of Newton–Raphson iterations on equations (57) and (58) to find the updated internal state variables. Take the \( k \)th local iteration for example, the corrector is activated to find a set of \( \Sigma_{n+1}^{(k)} \) which makes the residual terms \( R_{n+1}^{(k)} \) in equation (58) less than ‘tol’

\[
R_{n+1}^{(k)} = -h_{n+1}^{\tau} + h_{n+1}^{\tilde{\tau}} + \Delta \tilde{\tau}_{n+1}^{(k)} (\Phi_{n+1}^{(k)} - \zeta_{n+1}) \leq \text{tol}.
\]

This objective is equivalent to the following convergence conditions

\[
|\xi_{n+1}^{(k+1)} - \xi_{n+1}^{(k)}| \leq \text{tol} ; \ |h_{n+1}^{\tau} - h_{n+1}^{\tilde{\tau}}| \leq \text{tol}.
\]

Use the standard Newton–Raphson procedure\(^5\) to solve equations (57) and (58)

\[
\begin{bmatrix}
\Delta \xi_{n+1}^{(k+1)} \\
\Delta h_{n+1}^{\tau(k+1)}
\end{bmatrix} = -
\begin{bmatrix}
\frac{\partial h_{n+1}^{\tau}}{\partial \xi_{n+1}} & \frac{\partial h_{n+1}^{\tau}}{\partial \Phi_{n+1}} \\
\frac{\partial h_{n+1}^{\tau}}{\partial \Phi_{n+1}} & \frac{\partial h_{n+1}^{\tau}}{\partial \Phi_{n+1}}
\end{bmatrix}
^{-1}
\begin{bmatrix}
\Phi_{n+1}^{(k)} \\
R_{n+1}^{(k)}
\end{bmatrix}.
\]

\(^5\) The explicit expression for the Jacobian matrix during this Newton–Raphson iteration in equation (60) is quite complicated. The symbolic calculation tool in MATLAB is used here to find the Jacobian matrix, and the authors suggest interested readers to utilize this method to perform the tedious calculation. 

| Table 2. A set of representative material parameters used for the parametric numerical study | \( n \) | Value | \( n \) | Value |
|-----------------------------------|---|---|---|---|
| Type                              | Parameter | Value | Parameter | Value |
| Key material parameters           | \( E_A \) | 60 (GPa) | \( C_A \) | 8 (MPa K\(^{-1}\)) |
|                                  | \( E_M \) | 40 (GPa) | \( C_M \) | 6 (MPa K\(^{-1}\)) |
| 12                                | \( \nu_A = \nu_M \) | 0.3 | \( M_f \) | 333 (K) |
|                                  | \( \alpha_A = \alpha_M \) | 1.0 \( \times \) \( 10^{-5} \) (K\(^{-1}\)) | \( M_f \) | 220 (K) |
|                                  | \( H_{max} \) | 3%, 5%, 8% | \( A_f \) | 274 (K) |
|                                  | \( k_i \) | 0.02 | \( A_f \) | 370 (K) |
| Smooth hardening parameters       | \( n_1 \) | 0.5 | \( n_3 \) | 0.5 |
|                                  | \( n_2 \) | 0.5 | \( n_4 \) | 0.5 |

Figure 2. Three sets of load–displacement curves are interpreted into the engineering scale stress–strain curve for the calibration of infinitesimal model, and the true stress–strain curve for the calibration of proposed model. (Engineering scale is denoted by Eng. and true scale is denoted by True.)
6 As the nonlinear solver is activated for implicit analysis (i.e. select NLGEOM on), Abaqus automatically use the logarithmic strain as its strain measure, and the Jaumann rate is the utilized objective rate to account for the large rotation [69].

Table 3. Elastic modulus and transformation strain calibrated based on engineering and true scale stress–strain curves.

|                  | Engineering scale | Case 1 | Case 2 | Case 3 | True scale | Case 1 | Case 2 | Case 3 |
|------------------|-------------------|--------|--------|--------|------------|--------|--------|--------|
| $E_A$ (GPa)      | 60.00             | 60.00  | 60.00  | 60.00  | $E_A$ (GPa) | 60.00  | 60.00  | 60.00  |
| $E_M$ (GPa)      | 35.48             | 34.06  | 32.05  | 32.05  | $E_M$ (GPa) | 40.00  | 40.00  | 40.00  |
| $H_{max}^{true}$| 3.04%             | 5.13%  | 8.33%  |        | $H_{max}$  | 3.00%  | 5.00%  | 8.00%  |

The following results on internal state variables at $(k+1)$th iteration can be obtained

$$
\begin{bmatrix}
\frac{\epsilon_f^{(k+1)}}{\gamma_{n+1}} \\
\frac{h_n^{(k+1)}}{h_{n+1}}
\end{bmatrix} = \begin{bmatrix}
\frac{\epsilon_f^{(k)}}{\gamma_{n+1}} \\
\frac{h_n^{(k)}}{h_{n+1}}
\end{bmatrix} + \begin{bmatrix}
\Delta \frac{\epsilon_f^{(k+1)}}{\gamma_{n+1}} \\
\Delta h_n^{(k+1)}
\end{bmatrix}.
$$

Once the converged values of $\{h_n^{(k+1)}, \frac{\epsilon_f^{(k+1)}}{\gamma_{n+1}}\}$ are found, the current transformation corrector step is labeled as finished and the UMAT proceeds to the next increment. Otherwise, the Newton-Raphson procedure exits at this step after a certain number of iterations and the current FE increment step stops.

5. Numerical results

In this section, the proposed model is used to predict the stress/thermally-induced phase transformations in SMAs subjected to general three-dimensional thermo-mechanical loading. Several numerical examples are presented here to test the capabilities of this model to account for large strains and rotations, and also to resolve the artificial stress errors issue. First, a parametric analysis on a uniaxial SMA bar is studied to show that the proposed model is able to consider the geometry nonlinearity induced by large strains. Second, two BVPs, i.e. an SMA beam and an SMA torque tube subjected to stress-induced phase transformations, are tested as large rotation cases. To show the model is able to resolve the artificial stress errors issue, the cyclic response of the beam and the torque tube are obtained via the proposed model, and the results are compared against the predictions obtained by the Abaqus nonlinear solver. Next, an isobaric BVP of an SMA torque tube subjected to varying thermal loading is investigated to predict the thermally-induced phase transformation. In the end, to show the capability of the model to capture the non-proportional local stress and strain evolutions, a 3D solid flexible structure undergoing a self-expanding process is studied. The proposed model is anticipated to be further validated against experimental data of NiTi and NiTiHf SMAs under uniaxial and other non-uniform loading conditions. The ultimate objective is to validate the capability of the proposed model to predict the response of SMA-based actuators, such as SMA beams and torque tubes, which are intended to be integrated with the future supersonic transport aircraft to realize the morphing capabilities to reduce the sonic boom noise.

5.1. SMA bar under isothermal loading

To test the capability of the proposed model to account for the effects of large strain, an SMA prismatic bar is studied under uniaxial isothermal loading condition. A parametric study is performed with the maximum transformation strain $H_{max} = 3\%, 5\%, 8\%$ to represent three different loading cases. A group of representative material parameters (two material parameter groups combined) used in this example are listed in table 2 referenced from [1, 29]. The SMA prismatic bar has a length $L = 100$ (mm) and an square cross section with an edge length $a = 10$ (mm). It is subjected to a proportional force loading up to 120 (kN) then unloading to 0 (kN), the temperature is kept constant at 380 K throughout the process. Generally, the load–displacement curves provided from such uniaxial experiments are interpreted into the engineering scale stress–strain curves to facilitate the model calibration. However, when the materials experience a strain that is no longer considered small, the geometry nonlinearity due to such strain has to be taken into consideration. To demonstrate that the proposed model accounts for this, the calibrated values of elastic modulus $(E_A, E_M)$ and maximum transformation strain $(H_{max}^{true})$ based on the true stress–strain curve are compared against the values from its infinitesimal counterpart. Three sets of load–displacement curves are generated shown in figure 2(a). They are interpreted into stress–strain curves in two scales, i.e. the true stress (Cauchy stress) versus the true strain (logarithmic strain) curve and the engineering stress (nominal stress) versus engineering strain (infinitesimal strain) curve. By using the calibration procedure described in appendix C, the calibrated values of $E_A, E_M$ and $H_{max}^{true}$ summarized in table 3.

Table 3 shows that the values of $E_A$ are identical in both the two scales. However, the values of $E_M$ in engineering scale change from 35.48 to 32.05 GPa, which indicates a
Actually, such material softening is not real. Instead, it is the effect of disregarding the geometric nonlinearity induced by large strain as described previously. In this case, the geometric nonlinearity means that the bar needs to contract its cross section to compensate for its elongation to preserve the volume conservation. Disregarding the change of cross section results in an unreal decreasing on the values of $E_M$. By doing the calibration based on the true stress–strain curve instead of the engineering one, the proposed model is able to exclude the geometry nonlinearity induced by large strain, so that the calibrated values of $E_M$ remain the same in the three loading cases from true scale. Besides, the values of $H_{\text{max}}$ are also worth to be noted. Although $H_{\text{max}}$ shows different values in the two scales, a relationship exists between the true scale $H_{\text{max}}$ and the engineering scale $H_{\text{eng}}$, i.e. $H_{\text{max}} = \ln(1 + H_{\text{eng}})$. Based on the results from this parametric study, it is shown that the infinitesimal strain assumption may no longer be considered as an accurate approximation when the strain regime is beyond 3%. In order to account for the effects caused by large strain, a finite strain model to consider the geometry nonlinearity is required even in a uniaxial case.

5.2. SMA beam under isothermal loading

The second BVP considered here is an SMA beam subjected to isothermal loading shown in figure 3. The SMA beam component has been investigated as bending actuators in [3] to realize a morphing variable-geometry chevron in order to change the outer engine shell shape to achieve specific aerodynamic characteristics. While only one loading cycle was considered in the previous study, this example examines the cyclic material and structural response. The studied beam has the same geometry as the SMA bar in section 5.1. Refer to
Figure 3, the beam is simply supported with one node being fixed to suppress the rigid body motion, and the upper face is subjected to a traction that ramps up to 24 (MPa) then decreases to 0 (MPa). Temperature is kept constant at 380 K throughout the whole numerical experiment. Material parameters used in this simulation are summarized in table 4. The cyclic material and structural response are obtained by the proposed model for a material point \( p \) (in figure 3) located at the middle bottom position, and are compared against the results obtained from the Abaqus nonlinear solver. Figure 4 shows the obtained cyclic longitudinal stress–strain curve for material point \( p \) under tension. As shown in figure 4(a), the proposed model provided a stable material response, while the Abaqus nonlinear solver predicted a shifting, instead of stable, response shown in figure 4(b). The observation from these results indicate that the spurious material response is obtained due to the usage of non-integrable objective rates in Abaqus as discussed in the introduction. Although the initial several loading cycles are almost the same in the results provided by Abaqus nonlinear solver, the accumulation of artificially introduced stress errors, around −2 MPa for each cycle, gradually drifts the material response left downwards throughout the 100 loading cycles. In total, −200 MPa stress residuals together with −0.6% remnant strains are observed at the end. Such stress errors consist of almost 18% of the maximum stress levels experienced by material point \( p \). As a comparison, figure 5 shows 100 stress–strain curves for another material point subjected to compression at the middle of beam upper surface. The result shows an opposite shifting trend in contrast to the results of point \( p \). Again, a stable compressive stress–strain curve are predicted by the proposed model while the Abaqus nonlinear solver predicts a shifting one. In addition, figure 6 shows the obtained cyclic load–displacement curves for point \( p \). It can be seen that the proposed model predicted a stable structural response while the Abaqus nonlinear solver predicted a shifting structural response. Based on these results, it is demonstrated that the Abaqus nonlinear solver can no longer produce reliable results for the SMA beam subjected to cyclic loading.

5.3. SMA tube under isothermal loading

In this section, the BVP of a three-dimensional SMA torque tube under torsion loading is studied. Refer to figure 7(a), the

| Table 4. Calibrated values of material parameters for equiatomic NiTi [29]. |
|---|---|---|---|
| Type | Parameter | Value | Parameter | Value |
| | \( E_A \) | 60 (GPa) | \( C_A \) | 7.8 (MPa K\(^{-1}\)) |
| | \( E_M \) | 60 (GPa) | \( C_M \) | 7.3 (MPa K\(^{-1}\)) |
| Key material parameters | \( \nu_A = \nu_M \) | 0.3 | \( M_s \) | 333 (K) |
| 12 | \( \alpha_A = \alpha_M \) | \( 1.0 \times 10^{-5} \) (K\(^{-1}\)) | \( M_f \) | 220 (K) |
| | \( H^{\text{max}} \) | 4.7% | \( A_s \) | 274 (K) |
| | \( k_f \) | 0.021 | \( A_f \) | 370 (K) |
| Smooth hardening parameters | \( n_1 \) | 0.5 | \( n_3 \) | 0.5 |
| 4 | \( n_2 \) | 0.5 | \( n_4 \) | 0.5 |
Figure 7. Schematic for a three-dimensional cylindrical SMA torque tube subjected to isothermal torsion loading.

(a) Full size torque tube geometry  (b) Reduced size torque tube geometry

Figure 8. Mises stress residuals accumulated after one loading cycle for the torque tube predicted by the proposed model and the Abaqus nonlinear solver.

(a) Proposed model  (b) Abaqus nonlinear solver

Figure 9. The cyclic stress–strain response for an SMA tube under isothermal loading condition.

(a) Shear stress-strain curve predicted by proposed model  (b) Shear stress-strain curve predicted by Abaqus nonlinear solver
A tube has an inner radius $r = 3.0$ (mm) and thickness $t/r = 0.1$. In order to reduce the computational cost, a representative tube segment $L/r = 2/3$ is analyzed here. Boundary conditions are depicted in figure 7(b), the tube left face is fixed and the right face is subjected to a torsion loading. The torque proportionally increases to 25 (Nm) then unloads to 0 (Nm), the temperature is kept constant at 380 K. The torque tube undergoes a fully forward phase transformation from austenitic phase to martensitic phase followed by a reverse phase transformation from martensitic phase to austenitic phase. The material parameters used in this simulation are from table 4.

The cyclic shear stress–strain response of a material point from the tube outer surface at the right end is predicted by the proposed model and Abaqus nonlinear solver. As shown in figure 9(b), similar to the results observed in the SMA beam case, a shifting response is predicted by the Abaqus nonlinear solver due to the accumulation of shear stress errors. In contrast, a stable response is predicted by the proposed model shown in figure 9(a). More specifically, figure 8 shows the magnitude of stress residual accumulated after one loading cycle. The value of Mises stress residual predicted by the proposed model is almost zero compared to a value around 4 MPa predicted by the Abaqus nonlinear solver. As a result of the accumulation of such shear stress errors, the shear stress levels required to start the forward phase transformation spuriously decreases in the case of Abaqus nonlinear solver shown in figure 9(b). Besides, the maximum shear stress levels at the end of forward transformation increases, and the shape of hysteresis loop also changes. The cyclic structural response of the torque tube is also provided in figure 10 by plotting the applied torque versus the twist angle $\theta_z$. It can be seen that a stable structural response is predicted by the proposed model shown in figure 10(a) compared to a shifting structural response predicted by the Abaqus nonlinear solver shown in figure 10(b). From the observation on these results,
it is seen that the Abaqus nonlinear solver is not able to predict reliable results for the SMA torque tube subjected to 100 shearing cycles any more. Thus, the proposed model that can resolve the shear stress errors is required for the SMA torque tube subjected to cyclic torsion loading.

5.4. SMA tube under isobaric loading

In order to test the capability of the proposed model to predict the thermally-induced phase transformation in SMAs, a three-dimensional cylindrical SMA tube is studied under cyclic isobaric torsional loading, i.e. subjected to a constant torsion load with temperature variation cycles. The SMA torque tubes has been investigated as rotational actuators to realize a morphing wing during the plane take-off and cruise regime [70–73]. The design and optimization of such SMA-based morphing structure requires a thorough understanding on the response of SMA torque tubes subjected to cyclic isobaric loading. To that end, the SMA tube component is analyzed under cyclic isobaric loading conditions. The model has the same geometry and material information as the tube simulation in section 5.3. The loading condition is as follows, a 3 N m torque load is applied to the tube right end and the temperature varies between 250 and 390 K for 100 cycles. Cyclic shear strain-temperature and \( \theta_z \)-temperature curves are obtained via the proposed model and the Abaqus nonlinear solver.

As it is shown in figure 11(a), a stable cyclic shear strain-temperature response for the tube is predicted by the proposed model. In contrast, a shifting cyclic strain-temperature response is predicted by using the Abaqus nonlinear solver shown in figure 11(b). More specifically, it can be seen that the isobaric response drifts downwards at \( T = 240 \text{ K} \) and is lifted up at \( T = 390 \text{ K} \) from cycle to cycle due to the stress error accumulation. Similar to the observation on the strain-temperature response, the cyclic \( \theta_z \)-temperature response is stable in the case of proposed model, and it is a shifting response predicted by the Abaqus nonlinear solver. The comparison on \( \theta_z \)-temperature response is plotted in figure 12. Based upon the analysis of an SMA tube subjected to thermal loading cycles, it is shown that the accumulated stress errors from Abaqus nonlinear solver result in a shifting cyclic isobaric response, and such artificial stress errors can be eliminated by using the proposed model.

5.5. 3D analysis of a flexible SMA structure

In this section, a 3D solid SMA flexible structure is studied to demonstrate that the proposed model can be used as a 3D structural design tool. This SMA flexible structure can be used as a stent to provide a less invasive method for administering support to diseased arteries, veins or other vessels in the human body. They are crimped into a smaller shape outside the body then inserted into the diseased artery. After being delivered into the desired position, the flexible structure expands automatically by using the phase transformation of SMAs [74]. The loading path of this self-expanding structure can be described with three steps, i.e. (1) the flexure is firstly crimped outside the body by external constraints and attached to a constraint container device called catheter or cannula. (2) The flexure is inserted into the body while the temperature increases from the room temperature to the body temperature. (3) The structure recovers its original shape when the constraint is removed [75–77]. Although there are analysis for similar type of flexible structure by other researchers [78, 79], among which only a small unit cell of the structure is analyzed. Here a full-scale 3D SMA flexible structure designed for repairing aortic dissection [74, 80] is studied to provide the global response of the structure during the self-expanding process. The structure is 32 (mm) long, 25.4 (mm) in outer diameter with 4 struts and 0.5 (mm) thick in the radial direction. Due to the curvature of the structure strut, the SMA flexure experiences large rotation and stress concentrations around the hinge part, which in return results in a complex local non-proportional stress and strain evolutions at the hinge. The loading path of this analysis is indicated by the red curve in figure 13(b). The material parameters used in this simulation are from table 5.
Figure 13. The expanded and crimped shapes, loading path, von Mises stress distribution and martensitic volume fraction contour of the SMA flexible structure during the self-expanding analysis.

### Table 5. Calibrated values of material parameters for NiTi (50.8% at Ni) [29].

| Type                  | Parameters | Value     | Parameters | Value     |
|-----------------------|------------|-----------|------------|-----------|
|                        | $E_A$      | 32.5 (GPa) | $C_A$      | 3.5 (MPa K$^{-1}$) |
|                        | $E_M$      | 23.0 (GPa) | $C_M$      | 3.5 (MPa K$^{-1}$) |
| Key material parameters| $\nu_A = \nu_M$ | 0.3       | $M_s$      | 264 (K)   |
| 12                    | $\alpha_A = \alpha_M$ | $2.2 \times 10^{-5}$ (K$^{-1}$) | $M_f$      | 160 (K)   |
|                       | $H^{\text{max}}$ | 3.3%      | $A_s$      | 217 (K)   |
|                       | $k_i$      | N/A       | $A_f$      | 290 (K)   |
| Smooth hardening parameters | $n_1$      | 0.17      | $n_3$      | 0.25      |
|                        | $n_2$      | 0.27      | $n_4$      | 0.35      |

* The values of transformation temperatures ($M_s, M_f, A_s, A_f$) are referenced from [81] in order to realize the self-expanding process within human body environment, the rest of values of the material parameters are taken from [29] as they are not provided from [81].
The expanded and crimped shapes of the SMA flexure during the analysis are shown in figure 13(a). As illustrated in figure 13(c), stress concentration due to the strut curved part is observed. The von Mises contour indicates that a local non-proportional stress field is evolved at the hinge location during the crimping process. As shown in figure 13(d), while the straight strut part is still in the austenitic phase, the stress-induced martensitic phase transformation is activated by the stress concentration at the hinge location subjected to bending. The martensitic volume fraction contour for the stress concentration due to the strut curved part is provided in terms of a 3D stress-temperature-displacement curve in figure 13(b), in which the red curve indicates the actual response while the blue curve is the projection of the blue curve on the stress-radial reduction ratio plane. This well captured non-proportional stress evolution and martensitic phase transformation in this example demonstrates that the proposed model can be used as an efficient tool for the 3D analysis and design of complex SMA-based structures.

6. Concluding remarks

Based on the SMA model proposed by Lagoudas and co-workers for small deformation analysis, a three-dimensional constitutive model for martensitic transformation in polycrystalline SMAs accounting for large deformation has been proposed in this work. Three important characteristics in SMA response are considered, i.e. the smooth transition during the phase transformation, the stress-dependent transformation strain to account for the coexistence of oriented/self-accommodated martensitic variants, and a stress-dependent critical driving force to consider the effect of applied stress levels on the size of the hysteresis loop. The proposed model is formulated based on the finite deformation framework that utilizes logarithmic strain and rate such that it not only accounts for the large strains and rotations that an SMA component may undertake, but also resolves the artificial stress errors that are caused by using other non-integrable objective rates. The proposed model is able to predict the stress-induced and thermally induced phase transformations in SMAs under general three-dimensional thermo-mechanical loading. In particular, it was shown in the example of an SMA bar that the proposed model accounts for the geometry nonlinearity induced by large strains, so that it corrects the spurious material softening in the results from its infinitesimal counterpart. In the numerical examples of an SMA beam and an SMA torque tube, it was demonstrated that the proposed model captures the large rotations that SMA-based components may undertake. By comparing the predicted cyclic response to the results obtained through the Abaqus nonlinear solver, the proposed model demonstrated that it effectively resolves the artificial stress errors. In the end, a 3D solid flexible structure experiencing local, non-proportional stress and strain evolution was analyzed by the proposed model, which shows the proposed model can be used as an efficient tool for the 3D analysis and design of complex SMA-based structures. The detailed formulation of the proposed model and its implementation procedures make it readily used by other researchers. The model can be further extended to incorporate additional nonlinear phenomena exhibited by SMAs, such as transformation-induced plasticity, viscoplasticity, and damage evolution.

Acknowledgments

The authors would like to acknowledge the financial support provided by the Qatar National Research Fund (QNRF) under grant number: NPRP 7-032-2-016, and the NASA University Leadership Initiative (ULI) project under the grant number: NNX17AJ96A. The conclusions in this work are solely made by the authors and do not necessarily represent the perspectives of QNRF and NASA.

Data availability

The user-defined material subroutine implemented in commercial FE software Abaqus for the proposed finite strain SMA model is available to download from the GitHub: https://github.com/Aero-tomato/SMA-UMAT.

Appendix A. Artificial stress residuals due to the non-integrable objective rates

In this appendix section, BVP is investigated to study the artificial stress residuals caused by using other non-integrable objective rates. As it is discussed in the introduction, the rate form hypoelastic constitutive theory has been criticized for its inconsistent choices on objective rates [63], this includes the well-known objective rates (e.g. Zaremba–Jaumann rate, Green–
Naghdi rate, Truesdell rate, etc). In other words, the rate form hypoelastic constitutive equation fails to be integrated to deliver an algebraic hyperelastic constitutive equation via the so-called objective rates, because of which spurious phenomena (e.g. shear stress oscillation, artificial stress residuals, etc) are often observed in a simple elastic deformation [46]. It was not until recently such self-inconsistent issues about hypoelastic constitutive models has been resolved by the logarithmic rate proposed by [46–53]. As their work proved that the logarithmic rate of the logarithmic strain $\mathbf{h}$ of its Eulerian type is equivalent to the strain rate $\mathbf{D}$, by which a grade-zero hypoelastic model can be exactly integrated into a finite strain elastic model based on logarithmic strain.

A.1. The cyclic response of an elastic square

Refer to figure A1 for the BVP schematics, a two-dimensional elastic square with length $H$ is under a closed path cyclic loading. The upper line of the square is subjected to a displacement control circular deformation, the deformation over geometry ratio is as $r/H = 0.2$ to induce large deformation strain. The stress components are examined by the hypoelastic equation (A.1) for 10 loading cycles. In equation (A.1), $\mathbf{C}$ is the stiffness tensor and a circle over $\tau$ means the different objective rates adopted

$$\tau = \mathbf{C} \cdot \mathbf{D}. \tag{A.1}$$

Based on the results from [47], hypoelastic equation (A.1) can be self-consistently integrated to a hyperelastic constitutive equation (A.2) based on logarithmic strain through the logarithmic corotational integration [64]

$$\tau = \mathbf{C} \cdot \mathbf{h}. \tag{A.2}$$

Kirchhoff stress components are obtained by equation (A.1) with three different objective rates, i.e. Jaumann rate, Green–Naghdi rate and Logarithmic rate. The predicted results are presented in figures A2–A4. The stress results are normalized by the material Young’s modulus $E$.
The stress residuals are examined at the end of the loading cycle and summarized in table A1. First, the stress components in all the three cases showed periodic oscillation. Since the material is confined to behave elastically, the deformation should be indissipative to anticipate that all the stress components should return to zero value in the end. However, the predicted stress components in the cases of Jaumman rate and Green–Naghdi rate showed artificial stress residuals are introduced. Refer to figure A2 for the case of Jaumman rate, stress residuals $\tau_{11}$ is $0.0164$, $\tau_{22}$ is $-0.0164$ and $\tau_{12}$ is $0.0411$ after the 10 loading cycles. In the case of Green–Naghdi rate, although there are inconsiderable stress residuals for $\tau_{11}$ and $\tau_{22}$ components, the shear residuals $\tau_{12}$ is $-0.023$. In contrast, all the stress residuals are almost negligible in the case of Logarithmic rate, which demonstrates that the hypoelastic constitutive equation utilizing logarithmic rate can be self-consistently integrated to deliver a hyperelastic equation based on the logarithmic strain. Interested readers are encouraged to further read [52, 82].

### Appendix B. Calibration of the material parameters

In this section, the material parameters utilized in the proposed model are identified from a set of one-dimensional experimental data. Note that the strain measure used here should be in the true (or logarithmic) scale rather than the engineering (or infinitesimal) scale. Material parameters used in the proposed model can be categorized into three groups, i.e. the key material parameters, smooth hardening parameters and intermediate parameters. First, the material constants such as elastic modulus $E_A$, $E_M$, Poisson’s ratios $\nu_A$ and $\nu_B$, the thermal expansion tensors $\alpha_A$ and $\alpha_M$, stress influenced coefficients $C_A$ and $C_M$ from the phase diagram (or called clausius clapeyron coefficient), critical phase transformation temperatures $A_s$, $A_f$, $M_s$, $M_f$ at stress free state are determined. Secondly, the hardening parameters describing the smooth transition feature are discussed. Finally, the intermediate parameters are derived based on the aforementioned two parameter groups. All the material parameters used in this model are summarized in table 2.

Because the data is provided in a one-dimensional case, all tensorial variables of the proposed model have to be reduced into a 1D scalar value. For example, the stress tensor is reduced as $\tau_{11}$, the logarithmic strain tensor is reduced as $\varepsilon_{11}$, etc. Constitutive equation (36) can be rewritten as one-dimensional form as follows

$$\tau = E[h - \alpha(T - T_0) - h^T],$$

where the effective elastic modulus $E$ is calculated by using the rule of mixture as follows

$$E = [1/E_A + \xi(1/E_M - 1/E_A)]$$

The evolution equation (39) is also reduced in one-dimensional form as

$$\Lambda = \Lambda_{11} = \begin{cases} H^{\text{wt}}(\tau) \sgn(\tau) & \xi > 0, \\ \frac{\theta - \varepsilon}{\varepsilon} & \xi < 0, \end{cases}$$

the thermodynamic driving force $\pi$ in the one-dimensional case can thus be obtained

$$\pi = \tau \Lambda + \frac{1}{2} \Delta \gamma + \tau \Delta \alpha (T - T_0) + \rho_0 \Delta \xi T$$

$$- \rho_0 \Delta \varepsilon \left[T - T_0 - T \ln \left(\frac{T}{T_0}\right)\right] - \rho \Delta u_0 - \frac{df}{d\xi}$$
the transformation function in one-dimensional form can be calculated based on equation (B.4). Considering the phase difference for the thermal expansion $\Delta \alpha$ and specific heat $\Delta c$ are small enough to be ignored, the following transformation functions for the forward case and the reverse case can be obtained respectively

$$\Phi_{fwd}(\tau, T, \xi) = \left[ \tau \Lambda + \frac{1}{2} \Delta S \tau^2 + r_0 \Delta s_0 T \right]$$

$$- r_0 \Delta u_0 - \frac{\partial f}{\partial \xi} - Y = 0.$$  \hspace{1cm} (B.5)

$$\Phi_{rev}(\tau, T, \xi) = \left[ \tau \Lambda + \frac{1}{2} \Delta S \tau^2 + r_0 \Delta s_0 T \right]$$

$$- r_0 \Delta u_0 - \frac{\partial f}{\partial \xi} - Y = 0.$$  \hspace{1cm} (B.6)

As described in the first paragraph of this section, there are three sets of material parameters that need to be identified. First, let’s consider the material constants $(E_A, E_M, v_A, v_M, \alpha_A, \alpha_M)$. Elastic modulus $E_A$, $E_M$ can be determined through a pseudoelastic stress and strain curve by calculating the slopes at martensitic phase and austenite phase. Poisson’s ratio is attained using a widely accepted value of $\nu_A = \nu_M = 0.33$ found in [29]. The thermal expansion coefficients are usually considered as $\alpha_A = \alpha_M$, which can be calibrated through an isobaric actuation experiment. The maximum transformation strain $H_{max}$ can be determined from the pseudoelastic experimental and the value of parameter $k_t$ is chosen to best fit the $H^{car}$ curve. The stress influence coefficients and the critical phase transformation temperatures $(C_A, C_M, M_r, M_s, A_s, A_f)$ can be calibrated through the phase diagram. Second, the material parameters related to the smooth hardening features are discussed. The coefficients $n_1$, $n_2$, $n_3$, $n_4$ without specific physical meanings are determined to best match the smoothness in corners of material response. Lastly, there are seven intermediate material parameters $(\rho_0 \Delta s_0, \rho_0 \Delta u_0, a_1, a_2, a_3, Y_0$ and $D)$ that need to be calculated to complete the model. Determination of such intermediate parameters requires a set of seven algebraic equations. The needed four equations come from transformation constraints as the Kuhn–Tucker condition equation (46) (i.e. $\Phi_{rev}(\tau, T, \xi) = 0$). The fifth equation comes from the continuity of Gibbs free energy at the end of the forward transformation ($\xi = 1$). The needed five algebraic equations are summarized as follows

1. Start of the forward transformation at zero stress ($\tau = 0; T = M_s; \xi = 0$). $\Phi_{fwd}(0, M_s, 0) = \rho_0 \Delta s_0 M_s - r_0 \Delta u_0 - a_3 - Y_0 = 0$

2. Finish of the forward transformation at zero stress ($\tau = 0; T = M_f; \xi = 1$). $\Phi_{fwd}(0, M_f, 1) = \rho_0 \Delta s_0 M_f - r_0 \Delta u_0 - a_1 - a_3 - Y_0 = 0$

3. Start of reverse transformation under zero stress ($\tau = 0; T = A_s; \xi = 1$). $\Phi_{rev}(0, A_s, 1) = -\rho_0 \Delta s_0 A_s + r_0 \Delta u_0 + a_2 - a_3 - Y_0 = 0$

4. Finish of the reverse transformation under zero stress ($\tau = 0; T = A_f; \xi = 0$). $\Phi_{rev}(0, A_f, 0) = -\rho_0 \Delta s_0 A_f + r_0 \Delta u_0 - a_3 - Y_0 = 0$

5. The gibbs free energy continuity at the end of forward transformation ($\xi = 1$). $f(\xi = 1)|_{\xi \geq 0} = f(\xi = 1)|_{\xi \leq 0}$

The above five algebraic equations yield the following expression for the five out of seven intermediate model parameters

$$a_1 = \rho_0 \Delta s_0 (M_f - M_s); \quad a_2 = \rho_0 \Delta s_0 (A_s - A_f)$$

$$a_3 = \frac{1}{4} a_2 \left( 1 + \frac{1}{n_3 + 1} \right) - \frac{1}{4} a_1 \left( 1 + \frac{1}{n_1 + 1} \right)$$

$$\rho_0 \Delta u_0 = \frac{1}{2} \rho_0 \Delta s_0 (M_f + A_f)$$

$$Y_0 = \frac{1}{2} \rho_0 \Delta s_0 (M_s + A_f) - a_3.$$  \hspace{1cm} (B.7)

Another two equations are derived from the Kuhn–Tucker condition in order to complete the calculation. For a one-dimensional uniaxial experiment, the Kuhn–Tucker condition (46) requires equation (B.8) to hold true at any specific stress level $\tau^*$

$$d\Phi = \frac{\partial \tau}{\partial \tau^*} d\tau + \frac{\partial \tau}{\partial T} dT + \frac{\partial \tau}{\partial \xi} d\xi = 0.$$  \hspace{1cm} (B.8)

Evaluate $d\Phi$ at the start point of the forward phase transformation (i.e. $\xi = 0$), and at the finish point of the forward phase transformation (i.e. $\xi = 1$), the incremental part of martensitic volume fraction should be zero (i.e. $d\xi = 0$) in both of the aforementioned cases. Therefore, the relationships between the stress temperature coefficients $C_M, C_A$ and the stress temperature slopes $\frac{d\tau}{dT}$ can be obtained.

For the forward transformation case, $\xi > 0$

$$C_M = \frac{d\tau}{dT} \bigg|_{\tau^*, \xi > 0} = \frac{-\rho_0 \Delta s_0}{\Lambda + \tau: \partial_\tau \Lambda + \Delta S \tau - \partial_\xi Y}.$$  \hspace{1cm} (B.9)

For the reverse transformation case, $\xi < 0$

$$C_A = \frac{d\tau}{dT} \bigg|_{\tau^*, \xi < 0} = \frac{-\rho_0 \Delta s_0}{\Lambda + \tau: \partial_\tau \Lambda + \Delta S \tau - \partial_\xi Y}.$$  \hspace{1cm} (B.10)

Using the equations (B.9) and (B.10), the rest two intermediate material parameters $\rho_0 \Delta s_0$ and $D$ can thus be expressed as follows

$$D = \frac{(C_M - C_A)(H^{car} + \tau \partial_\tau H^{car} + \tau \Delta S)}{(C_M + C_A)(H^{car} + \tau \partial_\tau H^{car})}.$$  \hspace{1cm} (B.11)

$$\rho_0 \Delta s_0 = \frac{-2 C_M C_A (H^{car} + \tau \partial_\tau H^{car} + \tau \Delta S)}{C_M + C_A}.$$  \hspace{1cm} (B.12)

**Appendix C. Model validation**

In order to validate the proposed model against experimental results, a uniaxial pseudoelastic tensile test is performed on a NiTi SMA [1]. In the experiment, a NiTi SMA strip is loaded at a constant temperature of 320 K, which is larger than the austenite finish temperature $A_s$. The strip is subjected to traction up to 600 MPa then unloaded to 0 MPa. The material
parameters listed in Table C1 are calibrated based on the experimental result. Figure C1 shows the comparison between the experimental results and the simulations from the proposed model, it clearly demonstrates that the proposed model predicts the stress-strain response of the NiTi strip quite well, including the phase transformation starting and finishing points, and the size of the hysteresis loop. As the experimentally tested specimen is an untrained SMA, there is an amount of transformation-induced plastic strain remained at the end of the loading. This phenomenon can be captured by extending the current model with the consideration of the transformation-induced plasticity.

Appendix D. Supplementary calculation for consistent tangent stiffness and thermal matrix

The consistent tangent stiffness and thermal matrix are derived in section 3.3. In order to determine the explicit values for $\zeta$ and $\Theta$ during the implementation of the proposed model, the explicit expressions of the following terms $\partial_\tau \Phi$, $\partial_\zeta \Phi$, $\partial_\Theta \Phi$ used in equation (51) are needed. First, the partial derivative of transformation function $\Phi$ with respect to stress $\tau$ can be obtained through differentiating equation (44) by $\tau$. Utilize the expression for $\pi$ in equation (43), it obtains

$$\partial_\tau \Phi = \begin{cases} \partial_\tau \pi - \partial_\tau Y, & \dot{\xi} > 0, \\ -\partial_\tau \pi - \partial_\tau Y, & \dot{\xi} < 0, \end{cases} \quad (D.1)$$

where the partial derivative of the thermodynamic driving force $\pi$ with respect to stress $\tau$ is

$$\partial_\tau \pi = \Lambda + (\partial_\tau \Lambda) \tau + \Delta \tau + \Delta \alpha(T - T_0) \quad (D.2)$$

and the partial derivative of critical driving force value $Y$ with respect to stress $\tau$ is

$$\partial_\tau Y = D[\Lambda + (\partial_\tau \Lambda) \tau] \quad (D.3)$$

based on the expression for the transformation direction tensor in equation (40), the partial derivative of $\Lambda$ with respect to stress $\tau$ are provided for the forward and reverse transformation cases as follows

$$\partial_\tau \Lambda = \begin{cases} \frac{3}{2} \partial_\tau H_{\text{cur}} \otimes \frac{\tau'}{\tau} + \frac{3}{2} H_{\text{cur}} \partial_\tau (\frac{\tau'}{\tau}), & \dot{\xi} > 0, \\ 0, & \dot{\xi} < 0 \end{cases} \quad (D.4)$$

where the partial derivative of the term $\left(\frac{\tau'}{\tau}\right)$ with respect to stress $\tau$ is provided in the following equation, in which $I$ is the fourth order identity tensor and $I$ is the second order identity tensor. It can be observed that $\partial_\tau \Lambda$ only has value for the forward transformation case while it has value zero for the reverse transformation case

$$\partial_\tau \left(\frac{\tau'}{\tau}\right) = \frac{1}{\tau} \left(1 - \frac{1}{3} I \otimes I - \frac{2}{3} \frac{\tau'}{\tau} \otimes \frac{\tau'}{\tau} \right) \quad (D.5)$$

to calculate the partial derivative of the current maximum transformation strain $H_{\text{cur}}$ with respect to stress $\tau$, the following result can be obtained based on equation (41)

$$\partial_\tau H_{\text{cur}} = \frac{3}{2} H_{\text{max}} \partial_\tau \left(\frac{\tau'}{\tau}\right) \quad (D.6)$$
Follow the similar procedure to obtain $\partial_t \Phi$, the partial derivative of the transformation function $\Phi$ with respect to martensitic volume fraction $\xi$, and the partial derivative of the transformation function $\Phi$ with respect to temperature $T$ can be calculated as follows

$$
\partial_t \Phi = \left\{ \begin{array}{ll}
\frac{1}{2} a_1 [n_1 \xi^{n_1} + n_2 (1 - \xi) \xi^{n_1}], & \xi > 0,
\frac{1}{2} a_2 [n_3 \xi^{n_3} + n_4 (1 - \xi) \xi^{n_3}], & \xi < 0,
\end{array} \right.
$$

(D.7)

$$
\partial_T \Phi = \left\{ \begin{array}{ll}
\tau : \Delta \alpha + \rho_0 \Delta c \ln \left( \frac{T}{T_0} \right) + \rho_0 \Delta s_0, & \xi > 0,
- \tau : \Delta \alpha + \rho_0 \Delta c \ln \left( \frac{T}{T_0} \right) + \rho_0 \Delta s_0, & \xi < 0.
\end{array} \right.
$$

(D.8)

ORCID iDs

L Xu @ https://orcid.org/0000-0002-1952-3835

References

[1] Lagoudas D C 2008 Shape Memory Alloys: Modeling and Engineering Applications (Berlin: Springer)
[2] Peraza-Hernandez E A, Hartl D J, Malak R J Jr and Lagoudas D C 2014 Origami-inspired active structures: a synthesis and review Smart Mater. Struct. 23 094001
[3] Hartl D J and Lagoudas D C 2007 Aerospace applications of shape memory alloys Proc. Inst. Mech. Eng. G 221 535–52
[4] Lazzara D S, Magee T, Shen H and Mabe J H 2019 Off-design sonic boom performance for low-boom aircraft AIAA Scitech 2019 Forum p 0606
[5] Xu L, Solomou A, Baxevanis T and Lagoudas D C 2019 A three-dimensional constitutive modeling for shape memory alloys considering two-way shape memory effect and transformation-induced plasticity AIAA Scitech 2019 Forum p 1195
[6] Raniecki B, Lexact and Tanaka K 1992 Thermodynamic models of pseudoelastic behaviour of shape memory alloys Arch. Mech., Arch. Mech. Stosow. 44 261–84
[7] Boyd J G and Lagoudas D C 1996 A thermodynamical constitutive model for shape memory materials. part: I. The monolithic shape memory alloy Int. J. Plast. 12 805–42
[8] Birman V 1997 Review of mechanics of shape memory alloy structures Appl. Mech. Rev. 50 629–45
[9] Raniecki B and Lexact C 1998 Thermodynamics of isotropic pseudoelasticity in shape memory alloys Eur. J. Mech. A 17 185–205
[10] Patoor E, Eberhardt A and Berveiller M 1996 Micromechanical modelling of superelasticity in shape memory alloys Le J. Phys. IV 6 C1–277
[11] Levitas V I 1998 Thermomechanical theory of martensitic phase transformations in inelastic materials Int. J. Solids Struct. 35 889–940
[12] Levitas V I and Preston D L 2002 Three-dimensional landau theory for multivariant stress-induced martensitic phase transformations: I. Austenite martensite Phys. Rev. B 66 134206
[13] Patoor E, Lagoudas D C, Entchev P B, Brinson L C and Gao X 2006 Shape memory alloys. part: I. General properties and modeling of single crystals Mech. Mater. 38 391–429
[14] Zaki W and Moumni Z 2007 A three-dimensional model of the thermomechanical behavior of shape memory alloys J. Mech. Phys. Solids 55 2455–90
[15] Saint-Sulpice L, Chirani S A and Calloch S 2009 A 3d super-elastic model for shape memory alloys taking into account progressive strain under cyclic loadings Mech. Mater. 41 12–26
[16] Hackl K and Heinen R 2008 A micromechanical model for pretextured polycrystalline shape-memory alloys including elastic anisotropy Cont. Mech. Thermodyn. 19 499–510
[17] Chemisky Y, Duval A, Patoor E and Zineb T B 2011 Constitutive model for shape memory alloys including phase transformation, martensitic reorientation and twins accommodation Mech. Mater. 43 361–76
[18] Sedlak P, Frost M, Benesová B, Zineb T B and Stittner P 2012 Thermomechanical model for niti-based shape memory alloys including r-phase and material anisotropy under multi-axial loadings Int. J. Plast. 39 132–51
[19] Cisse C, Zaki W and Zineb T B 2016 A review of constitutive models and modeling techniques for shape memory alloys Int. J. Plast. 76 244–84
[20] Chen L-Q 2002 Phase-field models for microstructure evolution Annu. Rev. Mater. Res. 32 113–40
[21] Steinbach I and Pezzolla F 1999 A generalized field method for multiphase transformations using interface fields Physica. D 134 885–93
[22] Steinbach I and Apel M 2006 Multi phase field model for solid state transformation with elastic strain Physica D 217 153–60
[23] Mamivand M, Zaeem M A and El Kaidir H 2013 A review on phase field modeling of martensitic phase transformation Comput. Mater. Sci. 77 304–11
[24] Zhong Y and Zhu T 2014 Phase-field modeling of martensitic microstructure in niti shape memory alloys Acta Mater. 75 337–47
[25] Thamburaja P and Anand L 2001 Polycrystalline shape-memory materials: effect of crystallographic texture J. Mech. Phys. Solids 49 709–37
[26] Wang X, Xu B and Yue Z 2008 Micromechanical modelling of the effect of plastic deformation on the mechanical behaviour in pseudoelastic shape memory alloys Int. J. Plast. 24 1307–32
[27] Yu C, Kang G, Kan Q and Song D 2013 A micromechanical constitutive model based on crystal plasticity for thermo-mechanical cyclic deformation of niti shape memory alloys Int. J. Plast. 44 161–91
[28] Auricchio F and Taylor R L 1997 Shape-memory alloys: modelling and numerical simulations of the finite-strain superelastic behavior Comput. Methods Appl. Mech. Eng. 143 175–94
[29] Lagoudas D, Hartl D, Chemisky Y, Machado L and Popov P 2012 Constitutive model for the numerical analysis of phase transformation in polycrystalline shape memory alloys Int. J. Plast. 32 155–83
[30] Brinson L and Lammering R 1993 Finite element analysis of the behavior of shape memory alloys and their applications Int. J. Solids Struct. 30 5261–80
[31] Brinson L C 1993 One-dimensional constitutive behavior of shape memory alloys: thermomechanical derivation with non-constant material functions and redefined martensite internal variable J. Intell. Mater. Syst. Struct. 4 229–42
[32] Leclercq S and Lexact C 1996 A general macroscopic description of the thermomechanical behavior of shape memory alloys J. Mech. Phys. Solids 44 953–80
[33] Lexact C 2013 Shape-Memory Alloys Handbook (New York: Wiley)
[34] Reese S and Christ D 2008 Finite deformation pseudo-elasticity of shape memory alloys-constitutive modelling and finite element implementation Int. J. Plast. 24 455–82
[35] Shaw J A 2000 Simulations of localized thermo-mechanical behavior in a nita shape memory alloy Int. J. Plast. 16 541–62
[36] Jani J M, Leary M, Subic A and Gibson M A 2014 A review of shape memory alloy research, applications and opportunities Mater. Des. 65 575–83
[37] Haghgooyan B, Shafagh N, Aydiner C C and Anlas G 2016 Experimental and computational investigation of the effect of phase transformation on fracture parameters of an sma Smart Mater. Struct. 25 075010
[38] Haghgooyan B, Hayrettin C, Baxevanis T, Karaman I and Lagoudas D C 2019 Fracture toughness of niti-towards establishing standard test methods for phase transforming materials Acta Mater. 162 226–38
[39] Wheeler R, Saunders R, Pickett K, Eckert C, Stroud H, Fink T, Gakhar K, Boyd J and Lagoudas D 2015 Design of a reconfigurable sma-based solar array deployment mechanism ASME 2015 Conf. on Smart Materials, Adaptive Structures and Intelligent Systems (American Society of Mechanical Engineers) V001T02A010
[40] Ziolkowski A 2007 Three-dimensional phenomenological thermodynamic model of pseudoelasticity of shape memory alloys at finite strains Contin. Mech. Thermodyn. 19 379–98
[41] Evangelista V, Marfia S and Sacco E 2010 A 3d sma constitutive model in the framework of finite strain Int. J. Numer. Methods Eng. 81 761–85
[42] Wang J, Moumni Z, Zhang W, Xu Y and Zaki W 2017 A 3d finite-strain-based constitutive model for shape memory alloys accounting for thermomechanical coupling and martensite reorientation Smart Mater. Struct. 26 065006
[43] Wang J, Moumni Z, Zhang W and Zaki W 2017 A thermomechanically coupled finite deformation constitutive model for shape memory alloys based on hencky strain Int. J. Eng. Sci. 117 51–77
[44] Stupkiewicz S and Petryk H 2013 A robust model of pseudoelasticity in shape memory alloys Int. J. Numer. Methods Eng. 93 747–69
[45] Damanpack A, Bodaghi M and Liao W 2017 A finite-strain constitutive model for anisotropic shape memory alloys Mech. Mater. 112 129–42
[46] Xiao H, Bruhns O and Meyers A 2006 Elastoplasticity beyond small deformations Acta Mech. 182 31–111
[47] Xiao H, Bruhns I O and Meyers I A 1997 Logarithmic strain, logarithmic spin and logarithmic rate Acta Mech. 124 89–105
[48] Xiao H, Bruhns O and Meyers A 1997 Hypo-elasticity model based upon the logarithmic stress rate J. Elast. 47 51–68
[49] Bruhns O, Xiao H and Meyers A 1999 Self-consistent eulerian rate type elasto-plasticity models based upon the logarithmic stress rate Int. J. Plast. 15 479–520
[50] Bruhns O, Xiao H and Meyers A 2001 Large simple shear and torsion problems in kinematic hardening elasto-plasticity with logarithmic rate Int. J. Solids Struct. 38 8701–22
[51] Bruhns O, Xiao H and Meyers A 2001 A self-consistent eulerian rate type model for finite deformation elastoplasticity with isotropic damage Int. J. Solids Struct. 38 657–83
[52] Meyers A, Xiao H and Bruhns O 2003 Elastic stress ratcheting and corotational stress rates Tech. Mech. 23 92–102
[53] Meyers A, Xiao H and Bruhns O 2006 Choice of objective rate in single parameter hypoelastic deformation cycles Comput. Struct. 84 1134–40
[54] Zhu Y, Kang G, Kan Q and Bruhns O T 2014 Logarithmic stress rate based constitutive model for cyclic loading in finite plasticity Int. J. Plast. 54 34–55
[55] Zhu Y, Kang G, Kan Q, Bruhns O T and Liu Y 2016 Thermo-mechanically coupled cyclic elasto-viscoplastic constitutive model of metals: theory and application Int. J. Plast. 79 111–52
[56] Müller C and Bruhns O 2006 A thermodynamic finite-strain model for pseudoelastic shape memory alloys Int. J. Plast. 22 1658–82
[57] Teerioho J P 2013 An extension of a shape memory alloy model for large deformations based on an exactly integrable eulerian rate formulation with changing elastic properties Int. J. Plast. 43 153–76
[58] Xiao H 2014 An explicit, straightforward approach to modeling sma pseudoelastic hysteresis Int. J. Plast. 53 228–40
[59] Yu C, Kang G, Kan Q and Zhu Y 2015 Rate-dependent cyclic deformation of super-elastic niti shape memory alloy: thermo-mechanical coupled and physical mechanism-based constitutive model Int. J. Plast. 72 60–90
[60] Xu L, Baxevanis T and Lagoudas D C 2017 A finite strain constitutive model for martensitic transformation in shape memory alloys based on logarithmic strain 25th AIAA/AHS Adaptive Structures Conf. p 0731
[61] Xu L, Baxevanis T and Lagoudas D 2017 A finite strain constitutive model considering transformation induced plasticity for shape memory alloys under cyclic loading VIII ECCOMAS Thematic Conf. on Smart Structures and Materials pp 1645–1477
[62] Xu L, Baxevanis T and Lagoudas D 2018 A three-dimensional constitutive model for polycrystalline shape memory alloys under large strains combined with large rotations ASME 2018 Conf. on Smart Materials, Adaptive Structures and Intelligent Systems V002T02A007
[63] Simo J C and Hughes T J 2006 Computational Inelasticity vol 7 (New York: Springer)
[64] Khan A S and Huang S 1995 Continuum Theory of Plasticity (New York: Wiley)
[65] Xiao H, Bruhns O and Meyers A 1998 Objective corotational rates and unified work-conjugacy relation between eulerian and lagrangean strain and stress measures Arch. Mech. 50 1015–45
[66] Qidwai M and Lagoudas D 2000 On thermomechanics and transformation surfaces of polycrystalline niti shape memory alloy material Int. J. Plast. 16 1309–43
[67] Hill R 1948 A variational principle of maximum plastic work in classical plasticity Q. J. Mech. Appl. Math. 1 18–28
[68] MIEHE C 1996 Exponential map algorithm for stress updates in anisotropic multiplicative elastoplasticity for single crystals Int. J. Numer. Methods Eng. 39 3367–90
[69] Abaqus 2014 6.14 documentation Dassault Systems Simulia Corporation, Providence, RI
[70] Saunders R, Hartl D, Herrington J, Hodge L and Mabe J 2014 Optimization of a composite morphing wing with shape memory alloy torsional actuators ASME 2014 Conf. on Smart Materials, Adaptive Structures and Intelligent Systems (American Society of Mechanical Engineers) V002T02A014
[71] Mabe J, Brown J and Calkins F 2014 Flight test of a shape memory alloy actuated adaptive trailing edge flap, part 1 Proc. SMST 2014 the Int. Conf. on Shape Memory and Superelastic Technologies
[72] Calkins F and Mabe J 2016 Flight test of a shape memory alloy actuated adaptive trailing edge flap ASME 2016 Conf. on Smart Materials, Adaptive Structures and Intelligent Systems (American Society of Mechanical Engineers) V001T04A007
[73] Icardi U and Ferrero L 2009 Preliminary study of an adaptive Structures and Intelligent Systems V002T02A007
[74] Jayendran R, Nour B and Ruim A 2018 Fluid-structure interaction (fsi) analysis of stent-graft for aortic
endovascular aneurysm repair (evar): material and structural considerations J. Mech. Behav. Biomed. Mater. 87 95–110

[75] Stoeckel D, Pelton A and Duerig T 2004 Self-expanding nitinol stents: material and design considerations Eur. Radiol. 14 292–301

[76] Kapoor D 2017 Nitinol for medical applications: a brief introduction to the properties and processing of nickel titanium shape memory alloys and their use in stents Johnson Matthey Technol. Rev. 61 66–76

[77] Liu X, Wang Y, Yang D and Qi M 2008 The effect of ageing treatment on shape-setting and superelasticity of a nitinol stent Mater. Charact. 59 402–6

[78] Arghavani J, Auricchio F and Naghdabadi R 2011 A finite strain kinematic hardening constitutive model based on hencky strain: general framework, solution algorithm and application to shape memory alloys Int. J. Plast. 27 940–61

[79] Hartl D J, Chatzigeorgiou G and Lagoudas D C 2010 Three-dimensional modeling and numerical analysis of rate-dependent irrecoverable deformation in shape memory alloys Int. J. Plast. 26 1485–507

[80] Jayendiran R, Nour B and Ruimi A 2018 Computational fluid-structure interaction analysis of blood flow on patient-specific reconstructed aortic anatomy and aneurysm treatment with dacron graft J. Fluids Struct. 81 693–711

[81] Kapoor D 2017 Nitinol for medical applications: A brief introduction to the properties and processing of nickel titanium shape memory alloys and their use in stents Johnson Matthey Technol. Rev. 61 66–7

[82] Xiao H, Bruhns O and Meyers A 1999 Existence and uniqueness of the integrable-exactly hypoelastic equation and its significance to finite inelasticity Acta Mech. 138 31–50