Shape-memory alloys: macromodelling and numerical simulations of the superelastic behavior

Ferdinando Auricchio\textsuperscript{a,}* , Robert L. Taylor\textsuperscript{b} , Jacob Lubliner\textsuperscript{c}

\textsuperscript{a}Universita' di Roma 'Tor Vergata', Dipartimento di Ingegneria Civile, Via della Ricerca Scientifica, 00133 Rome, Italy
\textsuperscript{b}University of California at Berkeley, Berkeley, CA, USA
\textsuperscript{c}University of California at Berkeley, Berkeley, CA, USA

Received 11 July 1995

Abstract

Shape-memory alloys show features not present in materials traditionally used in engineering; as a consequence, they are the basis for innovative applications.

A review of the available literature shows a dearth of computational tools to support the design process of shape-memory-alloy devices. A major reason is that conventional inelastic models do not provide an adequate framework for representing the unusual macrobehavior of shape-memory materials.

The present work focuses on a new family of inelastic models, based on an internal-variable formalism and known as generalized plasticity. Generalized plasticity is adopted herein as framework for the development of one- and three-dimensional constitutive models for shape-memory materials.

The proposed constitutive models reproduce some of the basic features of shape-memory alloys, such as superelasticity, different material behavior in tension and compression, and the single-variant-martensite reorientation process.

For isothermal conditions the implementation of the model in a finite-element scheme and the form of the algorithmically consistent tangent are discussed in detail.

Numerical simulations of typical tests performed on shape-memory materials (e.g. uniaxial loading, four-point bending and three-point bending tests) are presented and compared with available experimental data.

Based on the overall developments, it appears that the proposed approach is a viable basis for the development of an effective computational tool to be used in the simulation of shape-memory-alloy devices.

1. Introduction

In the 1960s, Buehler and Wiley [7] developed a series of nickel–titanium alloys, with a composition of 53 to 57% nickel by weight, that exhibited an unusual effect: severely deformed specimens of the alloys, with residual strains of 8–15\%, regained their original shape after a thermal cycle. This effect became known as the shape-memory effect (Fig. 1), and the alloys exhibiting it were named shape-memory alloys.

It was later found that not only do other materials have the shape-memory property, but that at sufficiently high temperatures such materials also possess the property of superelasticity, that is, the recovery of large deformations during mechanical loading–unloading cycles performed at constant temperature (Fig. 2).

* Corresponding author. e-mail: auricchio@tovx1.ccd.utovrm.it.

0045-7825/97/017.00 © 1997 Elsevier Science S.A. All rights reserved

PII S0045-7825(96)01232-7
Fig. 1. Shape-memory effect. At the end of a mechanical loading-unloading path (ABC) performed at constant temperature, the material presents residual deformation (AC). The residual strain may be recovered through a thermal cycle (CDA).

Fig. 2. Superelasticity. At constant temperature the material is first loaded (ABC), showing a nonlinear behavior. When unloaded (CDA), the reverse transformation occurs, with zero final permanent strain. Note the hysteresis.

As a consequence of superelastic and shape-memory behavior, shape-memory alloys lend themselves to innovative applications. Recently proposed designs based on such materials range from self-expanding micro-structures for the treatment of body vessel occlusions to devices for the deployment and control of space structures such as antennas and satellites.

A review of the available literature and personal contacts in the industry, however, show a dearth of computational tools to support the design process. A major reason is the fact that conventional models of inelastic behavior, such as classical plasticity, do not provide an adequate framework for representing superelastic and shape-memory behavior.

Based on some pioneering work by Phillips and collaborators [11,37], over the past two decades a new family of inelastic models has been developed [20] that allows the description of features not representable by classical plasticity, though it includes classical plasticity as a special case; it has accordingly been called generalized plasticity [21]. As has recently been discussed [3,4,23], the numerical implementation of models belonging to this family is straightforward. Consequently generalized plasticity appears to be a viable and flexible environment for the development of constitutive materials with complex behavior.

The chief objective of this work is to propose a plausible initial development of computational tools for design with shape-memory alloys through an exploration of the applicability of generalized plasticity to the representation of the superelastic behavior, and specifically (1) the development of constitutive models that reproduce the superelastic behavior, (2) the numerical implementation of such models in a finite-element setting, and (3) the simulation of applications to show the viability of the proposed approach as an effective computational tool for the design of devices based on shape-memory alloys.

The work is organized as follows. Section 2 explains the reason for the increasing interest on shape-memory materials through a brief survey of SMA-based applications exploiting the superelastic behavior. Section 3 overviews the SMA micromechanics and introduces the terminology used in the study. The survey of the phase transitions occurring in shape-memory alloys allows the interpretation of the macroscopic material behavior and prepares the ground for the development of a continuum material model. Section 4 describes a major simplifying assumption (isotropic behavior) and present a one-dimensional and a three-dimensional constitutive model, which reproduce the superelastic behavior. Section 5 describes a time-discrete isothermal version of the three-dimensional model; its algorithmic implementation within a finite-element framework is carefully addressed. Numerical tests performed to check the algorithm are presented. The model ability of reproducing macroscopic effects, such as Luders band, is discussed. Section 6 assesses the ability to perform simulation of typical

---

1 In the following the following abbreviations are used: SMA for shape-memory alloy and SE for superelasticity.
SMA-based applications and show the potentialities of the proposed approach. Finally, Section 7 closes and comments on the work presented. Extensions and future developments are also discussed.

2. Superelastic applications

Shape-memory alloys have unique properties which are not present in many materials traditionally used in engineering applications. Accordingly, their use introduces new design capabilities, which make it possible to improve device performance as well as to propose innovative solutions.

The present section reviews applications exploiting the superelastic behavior; the goal is to use practical examples to show the SMA range of potentialities and to explain the reasons for the growing interest in SMA-based devices. In general, superelasticity-based applications take advantage of one of the following features: (1) the possibility of recovering large deformations (up to strain of 8–15%); (2) the existence of a transformation stress plateau, which guarantees constant stress over non-negligible strain intervals.

2.1. Medical guidewires

A guidewire is a long, thin, metallic wire, passed into the body through a natural opening or a small incision. It serves as a guide for the safe introduction of various therapeutic and diagnostic devices. The use of superelastic alloys may: (a) reduce the complication of the guidewire taking a permanent kink, which may be difficult to remove from the patient without injury; (b) increase steerability, that is, the ability to transmit a twist at one end of the guidewire into a rotation of identical degree at the other end [43].

2.2. Stents

Stent is the technical word indicating self-expanding micro-structures, which are currently investigated for the treatment of hollow-organ or duct-system occlusions. The stent is initially stretched out to reach a small profile, which facilitates a safe, traumatic insertion of the stent. After being released from the delivery system, the stent self-expands to over twice its compressed diameter and exerts a nearly constant, gentle, radial force on the vessel wall [32].

2.3. Orthodontics

During orthodontic therapy tooth movement is obtained through a bone remodeling process, resulting from forces applied to the dentition. The optimal tooth movement is achieved by applying forces that are low in magnitude and continuous in time. Recently, coil springs using superelastic materials have been designed. As proved experimentally [40], they produce excellent results due to the constant stress that shape-memory alloys are able to exert during a substantial part of the transformation.

2.4. The Homer Mammalok

The Homer Mammalok is a needle wire localizer designed to meet an important and demanding need in the medical field. It is often recommended that any woman over the age of forty has mammograms every other year. This procedure frequently identifies small lesions which must be removed and examined microscopically. The procedure for removing the lesion is made difficult because there are no landmarks to guide the surgeon and because the lesion cannot be easily distinguished in any obvious way from the surrounding tissue. At the same time, as little as possible excess tissue should be removed. The method currently used is to place a needle in the breast at the site of the lesion as a guide to the surgeon during the removal. To prevent accidental advancement or withdrawal during the transit to the operating room, the end of the needle has a barb; due to the sharp barb shape, tissue damages
are possible and the repositioning the needle is not possible. Superelastic alloys can be used to have self-deploying, rounded hooks, which do not move, due to their shape, and which can be easily repositioned without causing tissue damages [36].

2.5. Eyeglass frames

By using the superelasticity, eyeglass frame components can withstand extreme deformations and spring back completely [9].

2.6. Conclusions

The survey presented above clearly shows that shape-memory materials have a wide range of applications. Recently, international conferences on superelastic and shape-memory technologies have also been organized, testifying to an increasing interest on the subject. However, a review of the available literature and personal contacts in industry shows a lack of computational design tools. Accordingly, the aim of the following sections is to propose a plausible initial development of such design tools.

To reach this goal, we first review the phase transformations occurring in shape-memory materials. This allows an interpretation of the superelastic effect from a micromechanical point of view. A thermomechanical framework already presented in literature is then adopted for the development of SMA constitutive models. Finally, after discussing a finite-element implementation of the models, some sample problems are studied.

3. Essential features of the shape-memory alloy micromechanics

The present section presents a survey of the shape-memory alloy micromechanics. It then focuses on Nitinol, which is probably the shape-memory material most commonly used in applications. The main objectives are to help the understanding of the material macroscopic behavior and to prepare the ground for the development of constitutive models.

3.1. Phase transformations

Shape-memory alloys belong to a class of materials that can undergo reversible changes in the crystallographic symmetry-point-group. Such changes can be interpreted as martensitic phase transformations, that is, as solid–solid displacive-diffusionless (or lattice-distortive) phase transformations between a crystallographically more-ordered phase, called austenite or parent phase, and a crystallographically less-ordered phase, called martensite. Typically, the austenite is stable at high temperatures and low values of the stress, while the martensite is stable at low temperatures and high values of the stress. For a stress-free state we indicate with $T_{s}^{A}$ the temperature above which only the austenite is stable and with $T_{s}^{M}$ the temperature below which only the martensite is stable.

The phase transformations between austenite and martensite are the key to explain the superelasticity effect. For the simple case of uniaxial tensile stress a brief explanation follows (Fig. 2). Consider a specimen in the austenitic state and at a temperature greater than $T_{s}^{A}$; accordingly, at zero stress only the austenite is stable. If the specimen is loaded, while keeping the temperature constant, the material presents a nonlinear behavior $(ABC)$ due to a stress-induced conversion of austenite into martensite. Upon unloading, while again keeping the temperature constant, a reverse transformation from martensite to austenite occurs $(CDA)$ as a result of the instability of the martensite at zero stress. At the end of the loading–unloading process no permanent strains are present and the stress–strain path is a closed hysteresis loop.

The study of martensitic transformations can be approached from different points of view. In the following we briefly review some aspects of the kinematics and the kinetics of the martensitic phase transformations occurring in shape-memory alloys.
3.1.1. Kinematics

By the kinematics of a martensitic phase transformation we mean the changes of atomic structures occurring during the phase transformation.

3.1.1.1. Formation of a single martensite crystal. In general, single martensite crystals have a platelet shape. From experimental observation occurring during the formation of a martensite single crystal from a parent single crystal specimen, it can be deduced that the plate formation is accomplished by a macroscopic deformation, called *shape change* or *shape deformation*. Moreover, a well-defined interface or contact plane between the austenite and the martensite, the so-called *habit plane*, is observed.

From a crystallographic point of view, we just recall that in general SMA-parent phases have super-lattice BCC structures and are classified as \(\beta\)-phase alloys. The martensite crystals have *periodic stacking order structures*. Since in the martensite atoms of different radii are packed without any symmetry, the super-lattice structure tends to deform slightly, resulting in a typical monoclinic configuration.

3.1.1.2. Self-accommodation and martensite variants. As was recognized by Bain [5], the formation of martensite plates cannot simply follow the single-crystal mechanism discussed above. In fact, this mechanism introduces a major accommodation problem, due to the misfit between the martensite and the surrounding austenite.

If there is no preferred direction for the occurrence of the transformation, the martensite takes advantage of the existence of different possible habit planes, forming a series of crystallographically equivalent variants. The product phase is then termed *multiple-variant martensite* and is characterized by a *twinned* structure, which minimizes the misfit between the martensite and the surrounding austenite (Fig. 3).

On the other hand, if there is a preferred direction for the occurrence of the transformation (often associated with a state of stress), all the martensite crystals tend to be formed on the most favorable habit plane. The product phase is then termed *single-variant martensite* and is characterized by a *detwinned* structure, which again minimizes the misfit between the martensite and the surrounding austenite (Fig. 4).

According to the existence of different types of single-variant martensitic species, the conversion of each single-variant martensite into different single variants is possible. Such process, known as *reorientation process*, can be interpreted as a family of martensitic phase transformations and is associated with changes in the parameters governing the single-variant martensite production (hence, it is often associated to non-proportional changes of stresses).

3.1.1.3. Martensite–martensite transformations. Depending on the alloy, different martensitic struc-
tured may progressively be formed from the austenite during the cooling process. They all have stacking-order structures with the same basal plane and differ just in the stacking sequences.

The process of production of these different martensitic structures can also be interpreted as martensitic phase transformations.

3.1.2. Kinetics

By the kinetics of a phase transformation we mean the phase transformation evolution process. The martensitic transformations occur by nucleation and growth.

Depending on the accommodation response between the parent and the product phases (either plastic or elastic), the martensitic phase transformations are classified as non-thermoelastic (such as in Fe-Ni) or thermoelastic (in typical β-phase alloys)\(^2\) [46]. For thermoelastic transformations the accommodation of the martensite plates within the parent phase is essentially elastic, with no dislocation generation; the interface remains glissile, capable of backward movements, and the reverse transformation proceeds through shrinkage of martensite plates. The magnitude of the hysteresis, which is related to the transformation driving force, is in general small. Shape-memory alloys usually present thermoelastic martensitic transformations, guaranteeing the reversibility of the transformations.

Once the conditions for the nucleation are satisfied, depending on the alloy, the martensitic transformation may proceed in three different modes: the athermal mode, the isothermal mode and the burst mode [38]. In general, shape-memory alloys undergo the athermal (or achronic) mode, that is, the amount of martensite formed is a function only of the temperature and not of the length of time at which the alloy is held at that temperature. Athermal transformations start at well-defined temperatures, which are usually insensitive to rate-effects.

The phase-transformation kinetics is in general strongly influenced by a complex combination of internal and external parameters. Internal parameters are the alloy system, the composition and the lattice structure including defects. External parameters are the thermomechanical treatment and the training. Therefore, for each alloy a careful macroscopic characterization of the material kinetics through experimental investigation is needed.

3.2. An example of shape-memory material: Nitinol

Having laid down the general features of the martensitic transformations, we now specialize the discussion to nickel-titanium alloys, which are probably the shape-memory materials most frequently used in commercial applications.

Shape-memory properties for nickel (Ni) titanium (Ti) alloys were discovered in the 1960s, at the Naval Ordnance Laboratory (NOL) [7]; hence, the acronym NiTi NOL or Nitinol, which is commonly used when referring to Ni-Ti based shape-memory alloys. Starting from the 1970s, Ni-Ti has been widely investigated due to its frequent use in applications [10,16,48].

3.2.1. Metallurgical properties

A detailed discussion of crystallographic aspects, including possible self-accommodation mechanisms, can be found in [6,24–27,33,34].

The phase transformations occur in temperature ranges which strongly depend on the material composition. To obtain a desired phase-transformation temperature range, the required composition accuracy is mostly higher than the one obtainable by usual error in chemical analysis. Accordingly, in most cases the transformation temperatures are directly measured as quality control, rather than the chemical composition. Depending on the composition, \( T^{AM} \) ranges from \(-100^\circ C\) to \(100^\circ C\).

\(^2\) The non-thermoelastic phase transformations are also called strong, the thermoelastic phase transformations are also called weak.
3.2.2. Thermomechanical properties

The amount of heat recoverable memory strain and the size of the hysteresis loop strongly depend on alloy composition, thermomechanical processing, testing direction and deformation mode (that is, if the material is in simple tension, simple compression or shear) [1,31]. If strains larger than the recoverable strain are induced, then the reversible martensitic processes and the dislocations resulting from plastic flow interact, resulting in a reduction of the memory strain. For the full austenite–martensite phase transformation the recoverable memory strain is of the order of 8%, while the hysteresis width is typically of 30–50°C.

For uniaxial states of stress and in the usual range of applications the stress-temperature regions in which the phase transformations may occur are delimited with good approximation by straight lines with slopes ranging from 2.5 MPa/°C to over 15 MPa/°C. Experimental evidence shows that:

• Phase transformations do not exhibit pressure dependence in the case of long-aged Ni–Ti; for short-aged Ni–Ti the R-phase (B2-R) transition is unaffected by pressure, while the martensitic (R-B19') transformation is pressure dependent [17].

• Phase transformations are insensitive to temperature rates and to stress rates [8].

3.3. Toward a continuum mechanics model

The presented survey on the SMA micromechanics and in particular the discussion on the Nitinol highlight the variety and the complexity of the phenomena occurring at the material microscale. Accordingly, the construction of a shape-memory-alloy constitutive model based directly on the phase-transition micromechanics seems to be an overwhelming task. Thus, we resort to a thermomechanical continuum theory based on an internal-variable formalism. The specific material micromechanics is then tied in through the choice of appropriate simplifying assumptions, internal variables, and evolution equations for the internal variables.

From the previously discussed kinematics of the SMA martensitic phase transformations, we may deduce some indications on the structure of the internal variables. In fact, depending on the level of description to be reached, the internal variables should describe the presence and the features of the different phases; thus, they should be related to the phase fractions and orientations.

On the other hand, from the previously discussed kinetics of the SMA martensitic phase transformations, we may deduce some general features to take into account in the constitutive model through the choice of the internal-variable evolutionary equations. For example, the model should be able to take into consideration:

• Different initial and final conditions for each phase transformation to reproduce the hysteretic behavior.

• Dependence of the phase-transition initial conditions on stress and temperature.

• Pressure dependence of some phase transition in terms of initial conditions and evolutionary equations.

• Possibility of completely reversing the phase-transformation effects through appropriate thermomechanical cycles.

Finally, since experimental evidence shows that the phase transformations occurring in several shape-memory alloys are rate-independent, the goal of the following sections is to develop inelastic rate-independent models able to reproduce the superelastic behavior.

4. Constitutive models for shape-memory alloys

A general inelastic theory well suited for the development of constitutive models for materials undergoing solid–solid phase transitions has been presented in [7,22]. Such a theory is specialized herein to a specific class of materials undergoing solid–solid phase transitions, the shape-memory alloys. After discussing a major simplifying assumption, a one-dimensional model (1D-1 model) and a three-dimensional model (3D-1 model) are presented. The 1D-1 model has already been presented in
literature [22]; however, due to the novelty of the inelastic frame adopted, it is reviewed here with a didactic purpose. Moreover, since for proportional loading the 3D-1 model reduces to the 1D-1 model, the latter can also be interpreted as the basis for the development of the more complex three-dimensional model.

4.1. A simplifying assumption: isotropy

The survey on the kinematics of the SMA martensitic phase transformations presented in Section 3.1.1 pointed out the existence of several types of martensite, each one present with its own family of variants. To limit the dimension of the problem to handle, in the following we consider only one type of martensite. Moreover, we assume the material to be isotropic; accordingly, we do not distinguish between the different single-variant species. Hence, we propose to work with only two phases: the austenite \( A \) and the single-variant martensite \( S \).

Taking into account only two phases, the number of possible phase transformations drastically decreases. However, to properly model the three-dimensional macrobehavior of shape-memory materials, we need to take into account the single-variant martensite reorientation process. Therefore, in the most general case we consider three phase transformations:

- conversion of austenite into single-variant martensite \( A \rightarrow S \)
- conversion of single-variant martensite into austenite \( S \rightarrow A \)
- single-variant martensite reorientation \( S \rightarrow S \).

Finally, we recall that all the proposed models are rate-independent; this assumption, with good approximation, is appropriate for applications using Nitinol (refer to Section 3.2) as well as other shape-memory alloys frequently used.

4.2. 1D-1 model

The 1D-1 model reproduces the superelastic effect for one-dimensional states of stress (or equivalently, for proportional loading conditions).

4.2.1. Control and internal variables

As control variables, we assume the uniaxial stress, \( \sigma \), and the relative temperature, \( T \), subsequently referred to as the temperature. As internal variables, we may choose either the single-variant martensite fraction, \( \xi_S \), or the austenite fraction, \( \xi_A \), which are assumed to satisfy the following relation at each time:

\[
\xi_S + \xi_A = 1 \tag{1}
\]

Accordingly, there is only one independent internal variable, chosen herein to be \( \xi_S \). From Eq. (1) we may also get a relation between the rates of the fractions:

\[
\dot{\xi}_S + \dot{\xi}_A = 0 \tag{2}
\]

For simplicity, during the construction of the model we prefer to deal with both parameters, ensuring that the fraction evolutions always satisfy Eq. (2); as a consequence, at each time Eq. (1) is also satisfied. Finally, for each phase transformation we find it more convenient to establish first the evolutionary equation associated with the reducing fraction, deriving the evolution of the other fraction by enforcement of Eq. (2).

---

3 Several constitutive models for shape-memory alloys have been proposed and are available in the literature. Without even attempting to give an exhaustive list, few examples can be found in [12,13,35,39,44,45,47].

4 By convention, the capital letters \( S \) and \( A \) used as indices refer to specific fractions \( (S = \text{single-variant martensite}, \ A = \text{austenite}) \). Moreover, \( \xi_i = 0 \) \( (i = S, A) \) indicates the absence of the corresponding phase in the material, while \( \xi_i = 1 \) indicates that the material is completely in such a phase.
4.2.2. Phase transformations and activation conditions

To reproduce the superelastic behavior in a uniaxial state of stress, only two phase transformations can be considered: the conversion of austenite into single-variant martensite \((A \rightarrow S)\) and the conversion of single-variant martensite into austenite \((S \rightarrow A)\).

As showed experimentally, for uniaxial stress-temperature changes and in the usual range of applications, the regions in which phase transformations may occur are delimited with good approximation by straight lines [14,31]. Assigning a fraction change to each process we set:

\[
\begin{align*}
\dot{\xi}_s &= \dot{\xi}_s^{AS} + \dot{\xi}_s^{SA} \\
\dot{\xi}_A &= \dot{\xi}_A^{AS} + \dot{\xi}_A^{SA}
\end{align*}
\]

(3)

(4)

We now discuss the two production processes in more detail.

4.2.2.1. Single-variant martensite production. To describe the region corresponding to the single-variant martensite production process, we introduce the functions

\[
\begin{align*}
F^{AS} &= \sigma - C^{AS}T \\
F^{AS}_s &= F^{AS} - R^{AS}_s \\
F^{AS}_f &= F^{AS} - R^{AS}_f
\end{align*}
\]

(5)

(6)

(7)

with

\[
\begin{align*}
R^{AS}_s &= C^{AS}T^{AS}_s \\
R^{AS}_f &= C^{AS}T^{AS}_f
\end{align*}
\]

(8)

(9)

where \(C^{AS}, T^{AS}_s\) and \(T^{AS}_f\) are material parameters, with the latter two representing initial and final temperatures at which the transformation may occur at zero stress (Fig. 5). The region in which the transformation may take place is described by

\[
F^{AS}_s > 0, \quad F^{AS}_f < 0 \Rightarrow F^{AS}_s F^{AS}_f < 0
\]

(10)

Moreover, to activate the production of single-variant martensite, a stress increase, a temperature decrease or a proper combination of these actions should occur. Hence, we require

Fig. 5 1D-1 model: production of single-variant martensite. The shaded area represents states where the phase transformation can occur. The arrow indicates the direction of activation.

\[^3\text{By convention, the superscripts refer to specific evolution processes; accordingly, the superscript AS refers to the conversion of austenite into single-variant martensite, while the superscript SA refers to the conversion of single-variant martensite into austenite.}\]
Taking into account Eqs. (10) and (11), we express the fraction evolutionary equations relative to the $A \rightarrow S$ phase transformation in the form

$$
\frac{dF^A}{dt} = K^A \left( -F^A F^A \right) \\
\frac{dF^S}{dt} = -\frac{dF^A}{dt}
$$

where $K^A$ is a scalar function of the state variables (control and internal variables) and $\langle \cdot \rangle$ is the Macaulay bracket, defined as $\langle x \rangle = (x + |x|)/2$ where $|\cdot|$ is the absolute value.

4.2.2.2. Austenite production. Similarly to the above, we introduce the functions

$$
F^SA = T^SA - C^SA T
$$

where $C^SA$, $T^SA$ and $T_{f}^{SA}$ are material parameters, with the latter two representing initial and the final temperatures at which the transformation may occur at zero stress (Fig. 6). The region in which the transformation may take place is described by

$$
F^SA < 0, \quad F_{f}^{SA} > 0 \Rightarrow F^SA F_{f}^{SA} < 0
$$

Moreover, to activate the production of austenite, a stress decrease, a temperature increase or a proper combination of these actions should occur. Hence, we require

$$
\tilde{F}^SA < 0
$$

Taking into account Eqs. (17) and (18), we express the fraction evolutionary equations relative to the $S \rightarrow A$ phase transformation in the form

$$
\frac{dF^S}{dt} = K_1^S \left( -F^S F^A \right) \langle F^A \rangle \\
\frac{dF^A}{dt} = -\frac{dF^S}{dt}
$$

Fig. 6. 1D-1 model: production of austenite. The shaded area represents states where the phase transformation can occur. The arrow indicates the direction of activation.
where $K_{j}^{SA}$ is a scalar function of the state variables.

**REMARK 4.1.** Because of the general framework in which the model is developed, there is no limitation to the relative position of the two phase-transition zones; hence, they may intersect or they may be disjoint, since neither case would be problematic for the constitutive model.

### 4.2.3. Flow rule

To complete the specification of the rate equations, we need to choose $K_{j}^{AS}$ and $K_{j}^{SA}$. We restrict attention to very simple first-order forms. A possible choice is an evolution of exponential type, such as

$$
\dot{\xi}^{AS} = -\dot{\xi}^{SA} = -\beta^{AS} \xi^{A} \left( \frac{-F_{s}^{AS}F_{f}^{AS}}{|F_{s}^{AS}F_{f}^{AS}|} \right) \left( \frac{F_{s}^{AS}}{F_{f}^{AS}} \right)
$$

(19)

$$
\dot{\xi}^{SA} = -\dot{\xi}^{SA} = -\beta^{SA} \xi^{A} \left( \frac{-F_{s}^{SA}F_{f}^{SA}}{|F_{s}^{SA}F_{f}^{SA}|} \right) \left( \frac{-F_{s}^{SA}}{F_{f}^{SA}} \right)
$$

(20)

where $\beta^{AS}$ and $\beta^{SA}$ are scalar constants, measuring the rates at which the transformations proceed. A second possible choice is an evolution of power type, such as:

$$
\dot{\xi}^{AS} = -\dot{\xi}^{AS} = +\pi^{AS} \xi^{A} \left( \frac{-F_{s}^{AS}F_{f}^{AS}}{|F_{s}^{AS}F_{f}^{AS}|} \right) \frac{F_{s}^{AS}}{F_{f}^{AS}}
$$

(21)

$$
\dot{\xi}^{SA} = -\dot{\xi}^{SA} = -\pi^{SA} \xi^{A} \left( \frac{-F_{s}^{SA}F_{f}^{SA}}{|F_{s}^{SA}F_{f}^{SA}|} \right) \frac{-F_{s}^{SA}}{F_{f}^{SA}}
$$

(22)

In the following we limit the discussion of the power-type evolution to $\pi^{AS} = \pi^{SA} = 1$; hence, the evolution equation is termed as linear type. A discussion of the proposed evolutionary equations can be found in [2]. Clearly, other choices for $K_{j}^{AS}$ and $K_{j}^{SA}$ are possible [29]; however, from experimental results a first-order kinetic rule seems to be adequate [18].

Since the model has only one independent internal variable (chosen to be $\xi^{A}$), we need only one evolutionary equation. Recalling Eqs. (1)--(3), we have

$$
\dot{\xi} = \dot{\xi}^{AS} + \dot{\xi}^{SA} = \beta^{AS} \left( 1 - \xi^{A} \right) \left( \frac{-F_{s}^{AS}F_{f}^{AS}}{|F_{s}^{AS}F_{f}^{AS}|} \right) \frac{F_{s}^{AS}}{F_{f}^{AS}} - \beta^{SA} \xi^{A} \left( \frac{-F_{s}^{SA}F_{f}^{SA}}{|F_{s}^{SA}F_{f}^{SA}|} \right) \frac{-F_{s}^{SA}}{F_{f}^{SA}}
$$

(23)

for the exponential model and

$$
\dot{\xi} = \dot{\xi}^{AS} + \dot{\xi}^{SA} = -\left( 1 - \xi^{A} \right) \left( \frac{-F_{s}^{AS}F_{f}^{AS}}{|F_{s}^{AS}F_{f}^{AS}|} \right) \frac{F_{s}^{AS}}{F_{f}^{AS}} - \xi^{A} \left( \frac{-F_{s}^{SA}F_{f}^{SA}}{|F_{s}^{SA}F_{f}^{SA}|} \right) \frac{-F_{s}^{SA}}{F_{f}^{SA}}
$$

(24)

for the linear model. The Macaulay brackets manage the choice of the active evolution. We may express the flow rules also in the following equivalent forms:

$$
\dot{\xi} = H^{AS} \beta^{AS} \left( 1 - \xi^{A} \right) \frac{F_{s}^{AS}}{(F_{f}^{AS})^2} + H^{SA} \beta^{SA} \xi^{A} \frac{F_{s}^{SA}}{(F_{f}^{SA})^2}
$$

(25)

for the exponential model and

$$
\dot{\xi} = -H^{AS} \xi^{A} \frac{F_{s}^{AS}}{(F_{f}^{AS})^2} + H^{SA} \xi^{A} \frac{F_{s}^{SA}}{(F_{f}^{SA})^2}
$$

(26)

for the linear model. The scalar parameters $H^{AS}$ and $H^{SA}$ embed the phase-transformation activation conditions and are defined by the relations
\[
H^{AS} = \begin{cases} 
1 & \text{if } F_s^{AS} > 0, \quad F_f^{AS} < 0, \quad \dot{F}^{AS} > 0 \\
0 & \text{otherwise} 
\end{cases} \quad (27)
\]

\[
H^{SA} = \begin{cases} 
1 & \text{if } F_s^{SA} < 0, \quad F_f^{SA} > 0, \quad \dot{F}^{SA} < 0 \\
0 & \text{otherwise} 
\end{cases} \quad (28)
\]

**Remark 4.2.** The flow equations (23) and (24) can be integrated in closed form, as described in [2].

### 4.2.4. Test examples

We now test the ability of the model to reproduce the superelastic behavior under multiple stress cycles. Restricting the discussion to a small-deformation regime, we assume an additive decomposition of the total strain \( \varepsilon \)

\[
\varepsilon = \varepsilon^e + \varepsilon_L \xi_S
\]

where \( \varepsilon^e \) is the elastic strain and \( \varepsilon_L \) is the maximum residual strain. The elastic strain is assumed to be linearly related to the stress

\[
\sigma = E \varepsilon^e
\]

with \( E \) the elastic modulus. We limit the test example to the exponential flow rule and choose the following material parameters

\[
E = 1000 \text{ MPa}, \quad C^{AS} = C^{SA} = 1 \text{ MPa/°C} \\
\varepsilon_L = 10, \quad \beta^{AS} = \beta^{SA} = 10 \text{ MPa} \\
T_f^{AS} = 10 \text{ °C}, \quad T_f^{SA} = 130 \text{ °C}
\]

We assume to start always with a specimen in the austenitic phase (\( \xi_S = 0 \)) and to keep the temperature constant (\( T = 160 \text{ °C} > T_f^{SA} \)). In the first simulation (Fig. 7) we induce a complete set of transformations (that is, a complete \( A \rightarrow S \) and a complete \( S \rightarrow A \) transformation), followed by partial unloading.

---

The maximum residual strain \( \varepsilon_L \), regarded as a material constant, is a measure of the maximum deformation obtainable by multiple-variant martensite detwinning, hence, it is a measure of the maximum deformation obtainable aligning all the single-variant martensites in one direction. Within a small-deformation regime the assumed additive strain decomposition is consistent with the experimental observation in which for zero stress the residual strain is associated only with the amount of single-variant martensite.

Fig. 7. 1D-1 model. Multiple stress cycles at constant temperature: stress vs. strain. A complete transformation path is followed by partial loading and partial unloading.

Fig. 8. 1D-1 model. Multiple stress cycles at constant temperature: stress vs. strain. A complete transformation path is followed by partial loading and complete unloading.
reloading cycles. A partial reloading implies an incomplete direct transformation \((A \rightarrow S)\), while a partial unloading implies an incomplete reverse transformation \((S \rightarrow A)\). Note that the model describes a series of loops, which are internal to the complete loading–unloading cycle; such internal loops present a ratcheting effect which stabilizes after a few cycles. As test examples, we also consider the case of partial loading with complete unloading, the case of partial unloading with complete loading and the case of partial loading and partial unloading (Figs. 8–10). Comparing these results with experimental data available in the literature \([19,30,35]\), it is possible to conclude that the proposed model has an appropriate qualitative behavior.

4.3. 3D-1 model

The 3D-1 model reproduces the superelastic behavior for three-dimensional states of stresses; hence, it is a three-dimensional generalization of the 1D-1 model.

4.3.1 Control and internal variables

As control variables, we assume the stress, \(\tau\), and the temperature, \(T\). To capture the phase transformation effects, we introduce the transformation strain \(\epsilon'\), given by

\[
\epsilon' = \epsilon, u
\]

where \(u\) is the scaled transformation strain and \(\epsilon_s\) is a scalar parameter representing the maximum deformation obtainable only by detwinning of the multiple-variant martensite. Accordingly, as internal variables, we assume \(u\) and the single-variant martensite fraction, \(\xi_S\); thus, the model has one tensorial and one scalar internal variable.

4.3.2. Phase transformations and activation conditions

We consider three phase transformations:
- conversion of austenite into single-variant martensite \((A \rightarrow S)\),
- conversion of single-variant martensite into austenite \((S \rightarrow A)\),
- reorientation of the single-variant martensite \((S \rightarrow S)\).

Assigning a change of \(u\) and \(\xi\) to each processes, we set

\[
\begin{align*}
\dot{u} &= \dot{u}^{AS} + \dot{u}^{SA} + \dot{u}^{SS} \\
\dot{\xi} &= \dot{\xi}^{AS} + \dot{\xi}^{SA} \\
\end{align*}
\]

Fig. 9. 1D-1 model. Multiple stress cycles at constant temperature: stress vs. strain. A complete transformation path is followed by complete loading and partial unloading.

Fig. 10. 1D-1 model. Multiple stress cycles at constant temperature: stress vs. strain. A complete transformation path is followed by a series of partial loading and a series of partial unloading.
Note that the reorientation process occurs at constant value of the martensite fraction, consequently, there is no corresponding contribution in the evolution of the single-variant martensite fraction, that is, \( \dot{\xi}_s^{ss} = 0 \).

Each of the three phase transformations is assumed to occur in a specific region of the control-variable hyperplane \( \tau - T \), as discussed in the following.

4.3.2.1. Conversion of austenite into single-variant martensite \( (A \rightarrow S) \). As mentioned in Section 3, some phase transformations are pressure-dependent. To model such an effect, we introduce a Drucker–Prager-type loading function

\[
F^{AS}(\tau, T) = \|t\| + 3\alpha p - C^{AS}T
\]

where \( t \) is the deviatoric part of the stress (defined as: \( t = \tau - \text{tr}(\tau)1/3 \)), \( p \) is the pressure, \( C^{AS} \) and \( \alpha \) are material parameters and \( \| \cdot \| \) indicates the Euclidean norm. Generalizing the one-dimensional model (Eqs. (5)–(9)), the initial and final transformation function can be expressed as

\[
\begin{align*}
F_s^{AS} &= F^{AS} - R_s^{AS} \\
F_f^{AS} &= F^{AS} - R_f^{AS}
\end{align*}
\]

with

\[
\begin{align*}
R_s^{AS} &= \left[ \sigma_s^{AS} \left( \sqrt{2/3} + \alpha \right) - C^{AS} T_s^{AS} \right] \\
R_f^{AS} &= \left[ \sigma_f^{AS} \left( \sqrt{2/3} + \alpha \right) - C^{AS} T_f^{AS} \right]
\end{align*}
\]

where \( \sigma_s^{AS}, \sigma_f^{AS}, T_s^{AS} \) and \( T_f^{AS} \) are all material parameters. We assume that the conditions for the conversion of austenite into single-variant martensite are

\[
F_s^{AS} > 0, \quad F_f^{AS} < 0, \quad \dot{F}^{AS} > 0 \tag{34}
\]

and that the scaled transformation strain corresponding to the \( A \rightarrow S \) phase transformation evolves as follows

\[
\dot{u}^{AS} = \dot{\xi}_s^{AS} N^{AS} \tag{35}
\]

where

\[
N^{AS} = \frac{M^{AS}}{\|M^{AS}\|}, \quad M^{AS} = \frac{\partial F^{AS}}{\partial \tau} \tag{36}
\]

Recalling again the discussion on the 1D-1 model, for the evolution of the single-variant martensite fraction we may assume either an exponential form:

\[
\dot{\xi}_s^{AS} = H^{AS} \beta^{AS} (1 - \xi_s) \frac{\dot{F}^{AS}}{(F_f^{AS})^2}
\]

with \( \beta^{AS} \) a material parameter, measuring the speed of the transformation, or a linear form:

\[
\dot{\xi}_s^{AS} = -H^{AS} (1 - \xi_s) \frac{\dot{F}^{AS}}{F_f^{AS}}
\]

The scalar parameter \( H^{AS} \) embeds the conditions for the activation of the phase transformation (Eq. (34)) and it is defined by the relation

\[
H^{AS} = \begin{cases} 
1 & \text{if } F_s^{AS} > 0, \; F_f^{AS} < 0, \; \dot{F}^{AS} > 0 \\
0 & \text{otherwise}
\end{cases} \tag{37}
\]
4.3.2.2. Conversion of single-variant martensite into austenite ($S \rightarrow A$). Similarly to the above, we introduce a Drucker–Prager-type loading function

$$F^SA(\tau, T) = \|\tau\| + 3\alpha p - C^SA T$$

where $C^SA$ is a material parameter. Generalizing the 1D-1 model (Eqs. (12)–(16)), the initial and final transformation function can be expressed as

$$F^SA_s = F^SA - R^SA_s$$
$$F^SA_f = F^SA - R^SA_f$$

with

$$R^SA_s = \left[ \sigma^SA_s \left( \sqrt{\frac{2}{3}} + \alpha \right) - C^SA T^SA_s \right]$$
$$R^SA_f = \left[ \sigma^SA_f \left( \sqrt{\frac{2}{3}} + \alpha \right) - C^SA T^SA_f \right]$$

where $\sigma^SA_s$, $\sigma^SA_f$, $T^SA_s$, $T^SA_f$ are all material parameters. We assume that the conditions for the conversion of single-variant martensite into austenite are

$$F^SA_s < 0, \quad F^SA_f > 0, \quad F^SA < 0$$

(38)

and that the scaled transformation strain corresponding to the $S \rightarrow A$ phase transformation evolves as follows

$$u^SA = \dot{\xi}^SA N^SA$$

(39)

where

$$N^SA = \frac{u}{\|u\|}$$

(40)

Recalling again the discussion on the 1D-1 model, for the evolution of the single-variant martensite fraction we may assume either an exponential form:

$$\dot{\xi}^SA_s = H^SA \beta^SA \xi^SA \frac{F^SA}{(F^SA_f)^i}$$

(41)

with $\beta^SA$ a material parameter, measuring the speed of the transformation, or a linear form:

$$\dot{\xi}^SA_s = H^SA \xi^SA \frac{F^SA}{F^SA_f}$$

(42)

The scalar parameter $H^SA$ embeds the conditions for the phase transformation (Eq. (38)) and it is defined by the relation

$$H^SA = \begin{cases} 1 & \text{if } F^SA_s < 0, \quad F^SA_f > 0, \quad F^SA < 0 \\ 0 & \text{otherwise} \end{cases}$$

(43)

For later developments, we note that the conversion of martensite into austenite is associated only with a rescaling of the transformation strain $u$.

4.3.2.3. Single-variant martensite reorientation process ($S \rightarrow S$). To model the single-variant martensite reorientation process for non-proportional changes of stress (changes of direction or rotation), we set

$$F^SS = \|\tau\| + 3\alpha p - C^SS T$$
$$F^SS_s = F^SS - R^SS_s$$

with
\[ R_{s}^{ss} = \left[ \sigma^{ss} \left( \sqrt{\frac{2}{3}} + \alpha \right) - C^{ss} T^{ss} \right] \]  

(44)

where \( C^{ss}, \sigma^{ss}, \) and \( T^{ss} \) are material parameters. We assume that the condition for the activation of the reorientation process are

\[ F^{ss} > 0, \quad N^{ss} \neq 0 \]  

(45)

where

\[ N^{ss} = \frac{M^{ss}}{\| M^{ss} \|}, \quad M^{ss} = \frac{\partial F^{ss}}{\partial \tau} \]

and that the scaled transformation strain corresponding to the \( S \rightarrow S \) phase transformation evolves as follows

\[ u^{ss} = H^{ss} \xi_{s} \tilde{N}^{ss} \]  

(46)

where

\[ H^{ss} = \begin{cases} 1 & \text{if } F^{ss} > 0 \\ 0 & \text{otherwise} \end{cases} \]  

(47)

4.3.2.4. Additional considerations. From a physical point of view it seems plausible that whenever the material has enough energy to induce a conversion of austenite into single-variant martensite, then it has also enough energy to reorient the martensite fraction already resent. Hence, we set \( C^{as} = C^{ss} \) and \( R^{as} = R^{ss} \), such that \( F^{as} = F^{ss} \). Moreover, recalling the expression for \( F^{as} \) and \( F^{ss} \), we have

\[ M^{as} = M^{ss} = \frac{t}{\| t \|} + \alpha 1 = M \]  

(48)

\[ N^{as} = N^{ss} = \frac{1}{1 + 3\alpha} \left( \frac{t}{\| t \|} + \alpha 1 \right) = N \]  

(49)

We can now integrate the rate equations for the scaled transformation strain, obtaining the following equation

\[ u = \xi_{s} \tilde{N} \]  

(50)

where \( \tilde{N} \) is the current value of \( N \) if \( H^{ss} = 1 \), otherwise it is equal to the last value of \( N \) attained when \( H^{ss} \) was equal to 1. Consistency of Eq. (50) with the rate form can be checked computing its time derivative and considering the two cases \( H^{ss} = 1 \) and \( H^{ss} = 0 \):

- If \( H^{ss} = 1 \), then \( \tilde{N} = N \) and, by Eq. (50), \( u \) is in the direction of \( N \). From Eq. (40) we may conclude that \( N^{as} = N \). Hence, the rate form of Eq. (50) returns

\[ \dot{u} = \dot{\xi}_{s} N + \xi_{s} \dot{N} \]

which is exactly equal to the sum of the three transformation strain rate Eqs. (35), (39) and (46).

- If \( H^{ss} = 0 \), then \( H^{as} = 0 \) and only \( H^{as} \) can be non-zero. Moreover, since \( H^{ss} = 0 \), \( \tilde{N} \) is fixed. Hence, the rate form of Eq. (50) returns

\[ \dot{u} = \dot{\xi}_{s}^{as} \tilde{N} \]

Recalling the definition of \( \tilde{N} \) and that the \( S \rightarrow A \) phase transformation is a rescaling of \( u \), we have that \( \tilde{N} = u / \| u \| \), from which we obtain Eq. (39).

**Remark 4.3.** Due to the particular form of \( N \), the condition \( \tilde{N} \neq 0 \) required for the reorientation process to be active (Eq. (46)) reduces to
where \( n = t/\|t\| \) and \( I \) is the fourth-order identity tensor. Hence, the reorientation process occurs only for non-radial changes of the deviatoric stress.

4.3.3. Model review

Due to the Drucker–Prager-type loading functions, we split the transformation strain \( u \) into the volumetric and the deviatoric component

\[
u = v + \left( \frac{w}{3} \right) I, \quad w = \text{tr}(u)
\]

with \( \text{tr}(\cdot) \) being the trace operator. The constitutive Eq. (50) for the transformation strain \( u \) can also be split into volumetric and deviatoric parts:

\[
w = 3\alpha \xi \quad \text{(51)}
\]

\[
v = \xi \chi
\]

where \( \chi \) is the deviatoric part of \( \bar{N} \). The martensite fraction evolution represents the only part of the model still in rate form

\[
\xi = \dot{\xi} + \xi
\]

For the exponential case we have

\[
\dot{\xi} = \frac{\dot{F} - F}{F} \quad \text{(55)}
\]

\[
\xi = \frac{\xi}{F} \quad \text{(56)}
\]

while for the linear case we have

\[
\dot{\xi} = \frac{\dot{F}}{F} \quad \text{(57)}
\]

\[
\xi = \frac{\xi}{F} \quad \text{(58)}
\]

with the usual definitions for \( H^{AS} \) and \( H^{SA} \) (Eqs. (37) and (43)).

4.3.4. Material-parameter specialization to uniaxial states

The model presented so far is in a three-dimensional context. Accordingly, the material parameters involved in the constitutive equations are relative to a three-dimensional setting.

If the model is specialized to the case of uniaxial states of stress and strain, then the constitutive equations should involve the corresponding one-dimensional quantities, which are indicated herein with a superposed tilde and which may be easily computed from the corresponding three-dimensional parameters through the relations:

\[
B^{AS} = \tilde{B}^{AS} \left( \frac{2}{3} + \alpha \right)
\]

\[
B^{SA} = \tilde{B}^{SA} \left( \frac{2}{3} + \alpha \right)
\]

\[
\nu = \tilde{\nu} / \left( \frac{2}{3} + \alpha \right)
\]
5. Time-discrete isothermal models and algorithmic implementations

This section describes a time-discrete isothermal version of the three-dimensional 3D-1 model presented in Section 4.3. The algorithmic implementation of the model within a finite-element framework and the form of the consistent tangent tensor are discussed in detail. The discussion is limited only to a small deformation regime.

While during the development of the time-continuous model (Section 4.3) we assumed stress and temperature as control variables, during the development of the isothermal time-discrete models we assume the strain as the only control variable. This choice is consistent with the fact that, from the point of view of integration scheme, the time-discrete problem is considered strain-driven. Accordingly, we consider two time values, say \( t_n \) and \( t_{n+1} > t_n \), such that \( t_{n+1} \) is the first time value of interest after \( t_n \). Then, knowing the strain at time \( t_{n+1} \) and the solution at time \( t_n \), we should compute the new solution at time \( t_{n+1} \). To minimize the appearance of subscripts (and to make the equations more readable), we introduce the convention

\[ a_n = a(t_n), \quad a = a(t_{n+1}) \]

where \( a \) is any generic quantity. Therefore, the subscript \( n \) indicates a quantity evaluated at time \( t_n \), while no subscript indicates a quantity evaluated at time \( t_{n+1} \). We now consider a time-discrete isothermal version of the 3D-1 model under the assumption of small deformations.

5.1. Kinematics and elastic constitutive equation

Limiting the discussion to a small deformation setting, we additively decompose the total strain \( \varepsilon \) into an elastic part, \( \varepsilon^e \), and a transformation part, \( \varepsilon^{tr} \)

\[ \varepsilon = \varepsilon^e + \varepsilon^{tr} \]  \hspace{1cm} (59)

where, recalling Eq. (31)

\[ \varepsilon^{tr} = - \varepsilon_L u \]  \hspace{1cm} (60)

We choose a free-energy function quadratic in the elastic strain

\[ \psi = \psi(\varepsilon^e) = \frac{1}{2} \varepsilon^e D^e \varepsilon^e \]  \hspace{1cm} (61)

where \( D^e \) is the fourth-order rank elastic modulus tensor. Accordingly, the stress \( \tau \) is given by

\[ \tau = \frac{\partial \psi}{\partial \varepsilon} = D^e \varepsilon^e = D^e (\varepsilon - \varepsilon_L u) \]  \hspace{1cm} (62)

Recalling the Drucker–Prager form of the loading functions and the hypothesis of isotropy (see Section 4.1 for a discussion) and assuming the elastic tensor \( D \) to be constant and isotropic, it is convenient to split Eqs. (59), (60) and (62) into the volumetric and the deviatoric components. Using the notation of Section 4.3.3, we have

\[ \varepsilon = \varepsilon - \frac{\theta}{3} \mathbf{1} \]  \hspace{1cm} (63)

\[ \varepsilon^e = \varepsilon^e - \frac{\theta^e}{3} \mathbf{1} \]  \hspace{1cm} (64)

\[ \varepsilon = \varepsilon^e + \varepsilon_L \xi s m \]  \hspace{1cm} (65)

\[ \theta = \theta^e + 3 \alpha \varepsilon_L \xi s \]  \hspace{1cm} (66)
\[ p = K \theta^e = K(\theta - \epsilon_L w) \]  
\[ t = 2G\epsilon^e = 2G(e - \epsilon_s v) \]

where \( K \) is the bulk modulus and \( G \) is the shear modulus.

5.2. Time-discrete model

Using a backward-Euler integration formula, the discrete forms of Eqs. (52)-(54) are given by

\[ w = 3\alpha \xi_S \]  
\[ v = \xi_S m \]  
\[ \xi_S = \xi_{S,n} + (\lambda_S^{AS} + \lambda_S^{SA}) \]

where

\[ \lambda_S^{AS} = \int_{t_n}^{t_{n+1}} \xi_S^{AS} \, dt \]  
\[ \lambda_S^{SA} = \int_{t_n}^{t_{n+1}} \xi_S^{SA} \, dt \]

Similarly, using a backward-Euler scheme to integrate the time-continuous evolutionary equations yields the corresponding time-discrete evolutionary equations. Written in residual form and after clearing fractions, the time-discrete evolutionary equations specialize to

\[ \mathcal{R}^{AS} - (F^{AS})^2 \lambda_S^{AS} - H^{AS} \beta^{AS} (1 - \xi_S)(F^{AS} - F^{AS}_n) = 0 \]  
\[ \mathcal{R}^{SA} = (F^{SA})^2 \lambda_S^{SA} - H^{SA} \beta^{SA} \xi_S (F^{SA} - F^{SA}_n) = 0 \]

for the exponential case (Eqs. (55)-(56)), while for the linear case (Eqs. (57)-(58)) they specialize to

\[ \mathcal{R}^{AS} = F^{AS}_f \lambda_S^{AS} + H^{AS} (1 - \xi_S)(F^{AS} - F^{AS}_n) = 0 \]  
\[ \mathcal{R}^{SA} = F^{SA}_f \lambda_S^{SA} - H^{SA} \xi_S (F^{SA} - F^{SA}_n) = 0 \]

The quantities \( \lambda_S^{AS} \) and \( \lambda_S^{SA} \) can be computed expressing \( F^{AS} \) and \( F^{SA} \) as functions of \( \lambda_S^{AS} \) and \( \lambda_S^{SA} \) and requiring the satisfaction of \( \mathcal{R}^{AS} \) and \( \mathcal{R}^{SA} \).

5.3. Integration algorithm for the time-discrete model

A return-map algorithm is used as the integration scheme for the time-discrete model. Initially suggested by Maenchen and Sack [28] and Wilkins [50] for the solution of plasticity formulations, the return map provides an efficient and robust integration scheme based on a discrete enforcement of the evolutionary equations [41,42]. It belongs to the family of elastic-predictor/inelastic-corrector algorithms and, hence, is a two-part algorithm. In the first part, a purely elastic trial state is computed; in the second part, if the trial state violates the constitutive model, an inelastic correction is computed using the trial state as an initial condition. The details of the algorithm for the time-discrete model proposed here are:

(1) Trial state

Assume that no phase transformations occur (i.e. \( w = w_n, v = v_n, \xi_S = \xi_{S,n}, \lambda_S^{AS} = \lambda_S^{SA} = 0 \)). Accordingly, compute the trial pressure and the trial deviatoric part of the stress
\[ p^{TR} = K(\theta - \epsilon_L w_n) \]
\[ t^{TR} = 2G(e - \epsilon_L u_n) \]

(2) Check reorientation process \((S \rightarrow S)\)
Compute \(F^{SS}\)
If \(F^{SS}_s > 0\) then
set \(H^{SS} = 1\)
update \(w, m\) and \(v\)
recompute \(p^{TR}\) and \(t^{TR}\)
else
set \(H^{SS} = 0\)
end if

(3) Check solutions with \(\xi_S = 0\) and \(\xi_S = 1\)
Compute \(F^{AS}|_{\xi_S = 1}\) and \(F^{AS}|_{\xi_S = 0}\)
If \(F^{AS}|_{\xi_S = 1} > R^A_f\) then
\(\xi_S = 1\) is the appropriate solution
solution found
set \(H^{AS} = 1\)
else
\(\xi_S = 1\) is not the appropriate solution
end if
If \(F^{SA}|_{\xi_S = 0} < R^S_f\) then
\(\xi_S = 0\) is the appropriate solution
solution found
else
\(\xi_S = 0\) is not the appropriate solution
end if
If solution found then skip to 7

(4) Check \(A \rightarrow S\) transformation
Compute \(F^{AS}, F^{AS}_s\) and \(F^{AS}_n\)
If \(F^{AS}_s > 0, F^{AS} > F^{AS}_s\) and \(\xi_{S,n} < 1\) then
set \(H^{AS} = 1\)
else
set \(H^{AS} = 0\)
end if

(5) Check \(S \rightarrow A\) transformation
Compute \(F^{SA}, F^{SA}_s\) and \(F^{SA}_n\)
If \(F^{SA}_s < 0, F^{SA} < F^{SA}_s\) and \(\xi_{S,n} > 0\) then
set \(H^{SA} = 1\)
else
set \(H^{SA} = 0\)
end if

(6) Compute martensite evolution
If \(H^{AS} = 1\) or \(H^{SA} = 1\) then
compute \(\lambda^{AS}_S\) and \(\lambda^{SA}_S\)
update \(\xi_S, v, w, p\) and \(t\)
end if

(7) Compute algorithmic tangent
If \(H^{AS} = 1\) or \(H^{SA} = 1\) or \(H^{SS} = 1\) then
compute algorithmic inelastic tangent
else
  compute elastic tangent
end if

**REMARK 5.1.** Step 6 is performed solving Eqs. (74) and (75) or Eqs. (76) and (77) using a Newton-type iterative algorithm

\[
\begin{bmatrix}
  (\lambda_S^{AS})^{k+1} \\
  (\lambda_S^{SA})^{k+1}
\end{bmatrix}
= \begin{bmatrix}
  (\lambda_S^{AS})^k \\
  (\lambda_S^{SA})^k
\end{bmatrix}
- \begin{bmatrix}
  \partial R^{AS} \\
  \partial \lambda_S^{AS} \\
  \partial R^{SA} \\
  \partial \lambda_S^{SA}
\end{bmatrix}^{-1}
\begin{bmatrix}
  (\partial R^{AS})^k \\
  (\partial R^{SA})^k
\end{bmatrix}
\]

(78)

where

\[
R^{-1} = \begin{bmatrix}
  a & b \\
  c & d
\end{bmatrix}^{-1} = \begin{bmatrix}
  B & C \\
  D & E
\end{bmatrix}
\]

and the superscript \( k \) indicates the iteration index.

**5.4. Algorithmic tangent**

In the following we address the construction of the tangent matrix consistent with the time-discrete model. The use of a consistent tangent tensor preserves the quadratic convergence of the Newton method, which we adopt for the incremental solution of the momentum equations from a finite-element scheme.

Linearizing Eqs. (67)–(71), we get

\[
\frac{dp}{dt} = K(d\theta - \epsilon_L d\omega)
\]

\[
\frac{dt}{dt} = 2G(de - \epsilon_L dw)
\]

\[
\frac{dw}{dt} = 3\alpha(d\lambda_S^{AS} + \lambda_S^{SA})
\]

\[
\frac{dv}{dt} = (d\lambda_S^{AS} + d\lambda_S^{SA})m + H^{SS} \xi_S dn
\]

We need to compute \( dn \) only for the case \( H^{SS} = 1 \). When \( H^{SS} = 1 \), \( t, e \) and \( v \) are all in the same direction (Eq. (68)); hence

\[
n = \frac{t}{\|t\|} = \frac{e}{\|e\|}
\]

In [2] we show that

\[
dn = d\left(\frac{t}{\|t\|}\right) = d\left(\frac{e}{\|e\|}\right) = \frac{1}{\|e\|} [I - n \otimes n] \frac{de}{\|e\|}
\]

where

\[\|e\| = \|t\| / (2G) + \epsilon_L \xi_S\]

and \( I \) is the fourth-order identity tensor. Using these results in the linearized elastic relations, we get

\[
\frac{dp}{dt} = K d\theta - 3K \alpha \epsilon_L (d\lambda_S^{AS} + d\lambda_S^{SA})
\]

\[
\frac{dt}{dt} = 2G\left[I - H^{SS} \epsilon_L \xi_S \frac{1}{\|e\|} (I - n \otimes n)\right] \frac{de}{\|e\|} - 2G \epsilon_L (d\lambda_S^{AS} + d\lambda_S^{SA}) m
\]

The scalar quantities \( d\lambda_S^{AS} \) and \( d\lambda_S^{SA} \) are computed from the linearization of the discrete-time evolutionary equations (Eqs. (74) and (75) for the exponential model and Eqs. (76) and (77) for the linear model) as function of \( e, \lambda_S^{AS} \) and \( \lambda_S^{SA} \). Accordingly
\[
\begin{align*}
\frac{dR^{AS}}{dA^S} &= \frac{\partial R^{AS}}{\partial A^S} dA^S + \frac{\partial R^{AS}}{\partial \varepsilon} d\varepsilon = 0 \\
\frac{dR^{SA}}{dA^S} &= \frac{\partial R^{SA}}{\partial A^S} dA^S + \frac{\partial R^{SA}}{\partial \varepsilon} d\varepsilon = 0
\end{align*}
\]

which can be written as
\[
\begin{bmatrix}
\frac{dA^S}{d\varepsilon} \\
\frac{dA^S}{d\varepsilon}
\end{bmatrix} = -
\begin{bmatrix}
\frac{\partial R^{AS}}{\partial \varepsilon} \\
\frac{\partial R^{SA}}{\partial \varepsilon}
\end{bmatrix} d\varepsilon
\]

For the particular models presented here, we have
\[
\begin{align*}
\frac{\partial R^{AS}}{\partial \varepsilon} &= \frac{\partial R^{AS}}{\partial F^{AS}} \frac{\partial F^{AS}}{\partial \varepsilon} \\
\frac{\partial R^{SA}}{\partial \varepsilon} &= \frac{\partial R^{SA}}{\partial F^{SA}} \frac{\partial F^{SA}}{\partial \varepsilon}
\end{align*}
\]

and
\[
\begin{align*}
\frac{\partial F^{AS}}{\partial \varepsilon} &= 2Gn + 3ak1 \\
\frac{\partial F^{SA}}{\partial \varepsilon} &= 2Gn + 3ak1
\end{align*}
\]

Thus, using the expression for \( R^{-1} \) from Eq. (79), we can solve the system 80 in terms of \( dA^S \) and \( dA^S \)
\[
\begin{align*}
\frac{dA^S}{d\varepsilon} &= (T^{AS}_1 n + T^{AS}_2 1): d\varepsilon \\
\frac{dA^S}{d\varepsilon} &= (T^{SA}_1 n + T^{SA}_2 1): d\varepsilon
\end{align*}
\]

where
\[
\begin{align*}
T^{AS}_1 &= 2G(BA^{AS} + CA^{SA}) \\
T^{AS}_2 &= 3k\alpha(BA^{AS} + CA^{SA}) \\
T^{SA}_1 &= 2G(DA^{AS} + EA^{SA}) \\
T^{SA}_2 &= 3k\alpha(DA^{AS} + EA^{SA})
\end{align*}
\]

and
\[
A^{AS} = -\frac{\partial R^{AS}}{\partial F^{AS}} \\
A^{SA} = -\frac{\partial R^{SA}}{\partial F^{SA}}
\]

Finally, we obtain
\[
d\tau = D d\varepsilon
\]

where
\[
D = [K^n(1 \otimes 1) + 2G^*I_{dev} + M^n_1(n \otimes n) - M^n_2(m \otimes n) - M^n_3(m \otimes 1) - M^n_4(1 \otimes n)]
\]

with
\[ K^* = K \{ 1 - 3\varepsilon_1 \alpha (T_1^{AS} + T_2^{SA}) \} \]

\[ 2G^* = 2G \left\{ 1 - H^{SS} \varepsilon_1 \varepsilon_2 \frac{\|\varepsilon\|}{\|\varepsilon\|} \right\} \]

\[ M_1^* = H^{SS} 2G \varepsilon_1 \varepsilon_2 \frac{\|\varepsilon\|}{\|\varepsilon\|} \]

\[ M_2^* = 2G \varepsilon_1 (T_1^{AS} + T_1^{SA}) \]

\[ M_3^* = 2G \varepsilon_1 (T_2^{AS} + T_2^{SA}) \]

\[ M_4^* = 3K \alpha_1 \varepsilon_1 (T_1^{AS} + T_1^{SA}) \]

5.5. Specialization to linear and exponential flow rules

We now consider how some of the quantities previously introduced specialize for the case of linear and exponential flow rules. First of all, we note that

\[ \frac{\partial F^{AS}}{\partial \lambda^{AS}} = \frac{\partial F^{AS}}{\partial \lambda^{SA}} = [2G(n : m) + 9K \alpha^2] \varepsilon_1 = G_1 \]

Accordingly, for the linear flow rule, we have

\[ a = \frac{\partial R^{AS}}{\partial \lambda^{AS}} = -G_1 \lambda^{AS} - [(F^{AS} - F'^{AS}) + (1 - \xi) G_1] + F'^{AS} \]

\[ b = \frac{\partial R^{AS}}{\partial \lambda^{SA}} = -G_1 \lambda^{SA} - [(F^{AS} - F'^{AS}) + (1 - \xi) G_1] \]

\[ c = \frac{\partial R^{SA}}{\partial \lambda^{SA}} = -G_1 \lambda^{SA} - [(F^{SA} - F'^{SA}) - \xi G_1] \]

\[ d = \frac{\partial R^{SA}}{\partial \lambda^{SA}} = -G_1 \lambda^{SA} - [(F^{SA} - F'^{SA}) - \xi G_1] + F'^{SA} \]

and

\[ A^{AS} = -\frac{\partial R^{AS}}{\partial F^{AS}} = -\lambda^{AS} - (1 - \xi) \]

\[ A^{SA} = -\frac{\partial R^{AS}}{\partial F^{SA}} = -\lambda^{SA} + \xi \]

For the exponential flow rule, we have

\[ a = \frac{\partial R^{AS}}{\partial \lambda^{AS}} = \beta^{AS}(F^{AS} - F'^{AS}) + G_1[\beta^{AS}(1 - \xi) - 2\lambda^{AS} F'^{AS}] + (F'^{AS})^2 \]

\[ b = \frac{\partial R^{AS}}{\partial \lambda^{SA}} = \beta^{AS}(F^{AS} - F'^{AS}) + G_1[\beta^{AS}(1 - \xi) - 2\lambda^{AS} F'^{AS}] \]

\[ c = \frac{\partial R^{SA}}{\partial \lambda^{SA}} = -\beta^{SA}(F^{SA} - F'^{SA}) + G_1[\beta^{SA} \xi - 2\lambda^{SA} F'^{SA}] \]

\[ d = \frac{\partial R^{SA}}{\partial \lambda^{SA}} = -\beta^{SA}(F^{SA} - F'^{SA}) + G_1[\beta^{SA} \xi - 2\lambda^{SA} F'^{SA}] + (F'^{SA})^2 \]
and

\[ A^{AS} = -\frac{\partial R^{AS}}{\partial F^{AS}} = -2\lambda^{AS} F^{AS}_f + \beta^{AS}(1 - \xi_s) \]

\[ A^{SA} = -\frac{\partial R^{AS}}{\partial F^{AS}} = -2\lambda^{SA} F^{SA}_f + \beta^{SA} \xi_s \]

5.6. Test examples

We now present some simple numerical examples to show the performance of the discrete model and of the algorithmic implementation discussed above. All the simulations presented are obtained by employing the exponential flow rule. We choose the following material parameters:

\[ E = 1000 \text{ MPa} \quad \text{C}^{AS} = C^{SA} = 1 \text{ MPa/°C} \]

\[ \epsilon_l = 10\% \quad \beta^{AS} = \beta^{SA} = 10 \text{ MPa} \]

\[ T^{AS}_f = 10\^\circ \text{C} \quad T^{AS}_s = 70\^\circ \text{C} \quad T^{SA}_f = 90\^\circ \text{C} \quad T^{SA}_s = 130\^\circ \text{C} \]

The parameter \( \alpha \) is set equal to zero unless otherwise stated.

5.6.1. Uniaxial (proportional) stress cycles

We consider a specimen undergoing a uniaxial (proportional) loading-unloading history. Recalling the discussion on the 1D-1 model, for such a loading history it is possible to compute a closed-form solution which can be useful to evaluate the accuracy of the discrete solution obtained by the finite-element scheme. In Fig. 11 the stress-strain response is plotted for different values of the integration time step. Each integration time step corresponds to a different but constant strain step since the specimen is loaded controlling the displacements. Fig. 11 shows that the algorithmic response converges to a stable solution as the integration time step reduces and that an acceptable solution can be obtained also for large time steps.

The response of the model to uniaxial tension-compression tests is presented in Fig. 12 for different values of the parameter \( \alpha \). The parameter \( \alpha \) is computed from the initial value of the \( AS \) phase transformation in tension, \( (\sigma_1^{AS})_T \), and compression, \( (\sigma_1^{AS})_C \), from the relation

\[ \alpha = \frac{1}{3} \left( \frac{(\sigma_1^{AS})_C - (\sigma_1^{AS})_T}{(\sigma_1^{AS})_C + (\sigma_1^{AS})_T} \right) \quad (82) \]

Fig. 11. 3D-1 model. Uniaxial stress cycle. Comparison between the closed form solutions and the finite-element solution for different values of the integration time step. The specimen is loaded controlling the displacements.

Fig. 12. 3D-1 model. Tension–compression cycles. Stress–strain responses.
5.6.2. Biaxial (non-proportional) stress cycles

To test the model during a non-proportional stress cycle, we consider a biaxial loading condition. The specimen is constrained to obtain zero strain along one direction (x₁) and is loaded controlling the strain along a second direction (x₂, orthogonal to x₁). Along a third direction (x₃, orthogonal to the two previous ones) the specimen is free; hence, the stress is zero. This represents the behavior of a flat sheet (Fig. 13) free in the thickness direction (x₃) and restrained along the boundaries parallel to one of the in-plane directions (x₁). During the loading-unloading history the direction of the stress tensor varies, reproducing a non-proportional loading; hence, the reorientation process of the single-variant martensite is active.

The stress responses along the x₁ and the x₂ directions versus the strain along the x₁ direction is plotted in Fig. 14. The ability of the model to reproduce the superelastic behavior also for multiaxial non-proportional loading history can be easily asserted.

5.6.3. Luders bands

During the uniaxial testing of a shape-memory alloy bar, the presence of Luders bands is commonly noticed. The phase transformation is, in fact, observed to start in a localized zone of the specimen (in general one of the two ends) and to progressively propagate within the specimen, while maintaining a well-defined interface between the austenite and the martensite. We believe that the Luders-band effect is initiated by small variations in the homogeneity of the bar.

To model this effect, we consider a bar of length L = 100 mm and normal area 1 mm². The bar is loaded uniaxially as described in Fig. 15. To represent a geometric nonhomogeneity, the bar is assumed to be slightly tapered so as to initiate the transformation at the loaded end [we set: (Aₐ, A₁)/A₀ = 10⁻⁴, where A₀ is the area at the built-in end and A₁ is the area at the loaded end]. In Fig. 16 we plot the displacement of several points along the bar versus time. The presence of a moving boundary can be clearly detected from the difference of the responses; hence, the proposed model is able to reproduce the Luders-band effect.

6. Finite-element simulation of shape-memory-alloy devices

We now want to assess our ability to perform simulations for typical SMA-based devices exploiting the superelastic behavior.

The experimental data as well as the specimen specifications were provided by Dr. A.R. Pelton and
Dr. T.W. Duerig of Nitinol Device & Components, Inc.⁷. The material adopted is a standard binary Ni–Ti alloy with composition:

<table>
<thead>
<tr>
<th>Element</th>
<th>Ni</th>
<th>Ti</th>
<th>O</th>
<th>C</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight %</td>
<td>55.5</td>
<td>44.4</td>
<td>0.026</td>
<td>0.053</td>
<td>0.000016</td>
</tr>
<tr>
<td>atomic %</td>
<td>50.8</td>
<td>49.2</td>
<td>0.087</td>
<td>0.237</td>
<td>0.0085</td>
</tr>
</tbody>
</table>

A detailed description of the material and its manufacturing aspects can be found in [49].

6.1. Material-parameter characterization

The first step is to properly identify the parameters to use in the material constitutive model. To do so, we consider the results from uniaxial tension tests performed on a circular wire of diameter $D = 1.49$ mm. The test procedure is described in [49].

Two uniaxial tension tests were performed controlling the displacements. In the first test, a simple loading-unloading history was applied to the specimen reaching a maximum 8% strain. The stress-strain response is reported in Fig. 17. The test was designed to give some general information on the regions and the kinetics of the phase transformations. In the second test, a multiple loading-unloading history was applied to the specimen reaching a maximum 2% strain. The stress-strain response is reported in Fig. 18. The test was designed to give more information on the elastic properties of the parent phase (austenite) and on the starting value of the forward phase transformation (conversion of austenite into single-variant martensite).

From the inspection of the experimental stress-strain curves the following material parameters are chosen for both the exponential and the linear model:

\[
E = 60 \, 000 \, \text{MPa}, \quad \nu = 0.3, \quad \xi_L = 7.5\% \\
\theta_I^{\text{AS}} = \theta_f^{\text{AS}} = 0^\circ \text{C}, \quad C^{\text{AS}} = 0 \, \text{MPa/}^\circ\text{C} \\
\theta_I^{\text{AS}} = \theta_f^{\text{AS}} = 0^\circ \text{C}, \quad C^{\text{AS}} = 0 \, \text{MPa/}^\circ\text{C} \\
\]

For the exponential model we also set

⁷Nitinol Device & Components, Inc., 48501 Warm Spring Blvd., Fremont, 94539 CA, USA.
The stress–strain responses simulated using the exponential and the linear models are reported respectively with dash–dot and dash–dash lines in Figs. 19 and 20; a good match with the experimental data can be observed, especially for the exponential model. A better match could be obtained modeling the difference in elastic properties between the austenite and the martensite as well as through a more careful choice of the material parameters (for example, using a least-square algorithm minimizing the difference between the experimental data and the numerical solutions). These aspects will be addressed in future work.
Due to the limited number of tension tests available, it is not possible to accurately assess the maximum deformation obtainable by detwinning of the multiple-variant martensite ($\tilde{\varepsilon}_t$), the starting and final values of the reverse transformation as well as the difference between the elastic properties of the austenite and the martensite. Accordingly, the values chosen for these parameters are quite arbitrary.

Due to the limitation in the experimental facilities, it was also not possible to obtain experimental data for tests different than uniaxial tension or non-proportional loading tests. Thus, it is not possible to determine the difference in material response between tension and compression from experimental data. From consideration relative to a Ni-Ti-10\% Cu alloy [22] the parameter $\alpha$ is set equal to 0.15.

6.2. Four-point bending test

A four-point bending test was performed on a wire of circular cross section with diameter $D = 1.49$ mm. The test procedure is described in [49].

The setting of the test is idealized in Fig. 21. The beam is simply supported on rollers, sitting at a fixed distance $L = 20$ mm. The output parameters recorded during the experiments are the load $F$ and the bottom fiber deflection of the beam middle-span section.

Due to symmetry conditions only half of the half-span beam needs to be studied. For the numerical simulations we use the mesh presented in Fig. 22. In Fig. 23 the plotted load–deflection curves are computed using both the exponential and the linear model. The experimental load–deflection curve is also reported. From the inspection of the figures it is possible to conclude that:

![Fig. 21. Four-point bending test: test problem.](image1)

![Fig. 22. Four-point bending test: finite-element mesh.](image2)

![Fig. 23. Four-point bending test. Load–deflection curve. Finite element solution obtained using the exponential model (dash–dot line) and the linear model (continuous line). The experimental curve is also reported with a dash–dash line.](image3)
The predicted load-deflection curves match quite well the experimental data, except in the part of the curve corresponding to the final unloading portion. The size of the predicted hysteresis loop also slightly differs from the one found experimentally.

The exponential and the linear models predict similar results, except in the part of the curve corresponding to the initial unloading portion.

6.3. Three-point bending test

A three-point bending test was performed on the same wire and with the same test procedure used for the four-point bending test [49]; the setting of the test is idealized in Fig. 24. The wire has again a circular cross section with diameter $D = 1.49$ mm; the beam is simply supported on rollers, sitting at a fixed distance of length $L = 20$ mm. The output parameters recorded during the experiments are the load $F$ and the bottom fiber deflection of the beam middle-span section.

Due to symmetry conditions only half of the half-span beam needs to be studied. For the numerical simulations we use the mesh presented in Fig. 22. In Fig. 25 the plotted load-deflection curves are computed using both the exponential and the linear model. The experimental load-deflection curve is also reported. From the inspection of the figures it is possible to conclude that:

- The predicted load-deflection curves match quite well the experimental data, except in the part of the curve corresponding to the final loading portion (peak load).
- It is significantly different from the four-point bending test the fact that we now predict well the part of the curve corresponding to the final unloading and the size of the hysteresis loop.
- The exponential and the linear models predict again similar results.

6.4. Considerations on the simulations

From all the analyses presented we may conclude that:

- The different material behavior in tension and compression has to be taken into account during the modeling of shape-memory materials.
- Both the exponential and the linear models have a good ability to properly reproduce uniaxial tension tests, the former showing better matches with the available experimental data.
- The exponential and the linear models give comparable results in terms of global response parameters (for example, in terms of load-displacements curves).
- Both the exponential and the linear models produces results which are in good agreement with available experimental data for the case of four- and three-point bending tests.

Fig. 24. Three-point bending test: test problem.

Fig. 25 Three-point bending test. Load–deflection curve. Finite element solution obtained using the exponential model (dash–dot line) and the linear model (continuous line). The experimental curve is also reported with a dash–dash line.
• The discussed framework can be used to study the global and the local response of structures with complex geometry and it can be used to obtain valuable design information.

Based on the overall simulation presented, we may conclude that the proposed approach is a viable basis for an effective computational tool to be used in the simulation of shape-memory-alloy devices.

7. Closure and future research directions

The present work proposes an initial plausible development of a computation tool to be used during the design of SMA-based devices. To reach this goal, one- and three-dimensional thermomechanical constitutive models are developed, using an internal-variable inelastic framework able to describe solid–solid phase transformations.

The internal variables are chosen between a scalar variable (the martensitic fraction) and one second-order tensor variable (the transformation strain). A linear-type and an exponential flow rules for the internal variables are discussed.

The proposed constitutive models reproduce some of the basic features of shape-memory alloys, such as superelasticity, different material behavior in tension and compression, and the single-variant-martensite reorientation process.

For an isothermal case we discuss in detail the numerical implementation within a finite-element scheme for both the case of small- and large-deformation regimes.

To assess the viability of the proposed approach we simulate the response of some simple SMA typical structures (uniaxial test, four-point bending test, three-point bending test) for which experimental data are available. From the simulations it is possible to evince that the proposed framework:

• produces results which are in good agreement with the experimental data,
• can be used to study devices with different geometry and/or material properties,
• is able to give valuable local information such as displacement and/or stress–strain histories.

The match between the experimental and the predicted data are good, especially considering that the parameters used in the constitutive model have been extrapolated on the basis of only two tension tests.

The overall discussion leads us to conclude that the proposed approach is a plausible initial development of an effective computational tool for the simulation of SMA-based devices. However, additional research and developments are needed in order to obtain a fully general and reliable tool.

From the modeling point of view the required extensions and improvements are:

• extension of the isothermal time-discrete formulation to a large deformation regime,
• implementation of the complete thermomechanical model in a finite-element framework,
• more accurate assessment and modeling of the reorientation process for the single-variant martensites
• modeling of the different elastic properties of the austenite and the martensite.

All the proposed developments have to be followed by an experimental campaign, in order to better understand the material macroscopic behavior as well as the possible interaction between the different phase transformations.

References