Nečas Center for Mathematical Modeling

A microscopically motivated constitutive model for shape memory alloys: formulation, analysis and computations

M. Frost, B. Benešová, P. Sedlák

Preprint no. 2013-017

http://ncmm.karlin.mff.cuni.cz/

A microscopically motivated constitutive model for shape memory alloys: formulation, analysis and computations

Miroslav Frost[1,2], Barbora Benešová[3], Petr Sedlák[4,5]

[1] Institute of Thermomechanics ASCR, Prague, Czech Republic
[2] Mathematical Institute, Charles University in Prague, Prague, Czech Republic
[3] Department of Mathematics I, RWTH Aachen University, Germany
[4] New Technologies Research Center, University of West Bohemia, Plzeň, Czech Republic
[5] Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Prague, Czech Republic

Abstract

We present a 3D constitutive model for NiTi polycrystalline shape memory alloys exhibiting transformations between three solid phases (austenite, R-phase, martensite). The "full modeling sequence" comprising of formulation of modeling assumptions, construction of the model, mathematical analysis and numerical implementation and validation is presented.

Namely, by formulating micromechanics-inspired modeling assumptions we concentrate on describing the dissipation mechanism – a refined form of this description makes our model especially useful for complex loading paths. We then embed the model into the so-called energetic framework – extended to our case – while taking advantage of describing the dissipation mechanism through the so-called dissipation distance. We prove existence of energetic solutions to our model by a backward Euler scheme. This is then implemented into a finite element software and numerical simulations compared with experiments are also presented.

1 Introduction

Shape memory alloys (SMA) are metallic materials exhibiting remarkable properties like being able to sustain and recover large strains or to "remember" the initial configuration and return to it with temperature change. These properties arise from a reversible rearrangement of the crystal lattice associated with the so-called martensitic phase transformation. This is a transition from the (typically cubic) parent phase, referred to as austenite, to the product phase, called martensite and having a less symmetric crystal lattice. The transformation can be induced by thermal and mechanical loads as wells as by a combination of these.

The martensitic phase is then characterized by an internal microstructure of different strain states corresponding to different variants of martensite. This microstructure may rearrange depending on the current loading conditions; this will be referred to as *reorientation* hereinafter. It is this reorientation that is at the heart of the unique response of SMA (cf. e.g. [1]).

Clearly, the specific thermomechanical properties of SMA make them attractive from the point of view of application. Many products made out of SMA polycrystals are already being used or developed: in medicine, in automotive industry, in textile industry, etc. [2, 3, 4].

Due to the multiscale character of SMA and due to different objectives in modeling, the variety of models that have been proposed in literature is very large ranging from atomistic to macroscopical, from purely static to fully (thermomechanically) evolutionary

ones, from those which focus only on particular phenomena to more general ones (which may, however, disregard some peculiarities); cf. e.g. [5, 6, 7, 8] for surveys.

From the point of view of applications, however, *macroscopic*, phenomenological models for polycrystalline SMA form a powerful tool due to their easy implementation, less time-consuming calculations and the possibility to be adjusted for the particular material easily. Hence, a large number of such models has been proposed to date, e.g. [9, 10, 11, 12, 13, 14, 15, 16].

Performance of several of them was compared within a unique activity called "Round-robin SMA modeling" when experimentally measured behavior of a NiTi wire subjected to complex thermomechanical loading conditions has been compared to predictions of the individual models; see [17] for details. Assessment of the results revealed that especially in the case when multiple deformation mechanisms were coming into play, i.e. in *general (non-proportional) thermomechanical loadings*, the performance of the currently available models is still not completely satisfactory. Note that the chosen material, a NiTi alloy, is of particular importance since it is the commercially most successful SMA.

The difficulty in modeling the response due to non-proportional loadings stems from an interactions of multiple deformation modes (elasticity, two-stage phase transformation, reorientation) coming into play and the response becomes very sensitive to changes in loading conditions [18].

In this work, we derive a macroscopic, phenomenological model that is based on a simple modeling idea on the interaction of (some of) these processes. This allows us, in particular, to describe the rate-independent[1] dissipation mechanism related to (simultaneous) phase transformation and reorientation of martensite. Further, we embed the model into the energetic framework for rate-independent processes to prove existence of solutions and to propose a suitable numerical approximation. This is then implemented for verification and prediction.

Therefore, this paper includes a presentation of the "full modeling sequence", representing an *interplay and synergy* of mechanics, applied mathematical analysis and numerics, consisting of the following main steps:

1. *Formulation of simplifying modeling assumptions based on experimental results*: Based on a number of experiments available on polycrystalline NiTi samples and motivated by microscopic behavior of SMA on the level of the single crystal, we propose simplifying modeling assumptions that form the heart of the proposed model within this work. The microscopic motivation is especially important to us when we propose that a "zero-macroscopic-strain" type of martensite forms at the martensite-austenite interface. Microscopically this notion reminds of twins (laminates) well know from single crystal studies. Let us stress that we are aware that the situation in the polycrystal is more complex and this assumption is very simplifying and might not even be always satisfied. Nevertheless, the assumption is simple enough to allow us to formulate a well-performing model and still transfers some of the essence of the martensitic transition at the microscopic level to the macroscopic one.

2. *Formulation of a thermodynamically consistent model*: Following our modeling assumptions we formulate our model in the framework of generalized standard solids (GSM). In contrast to the original works due to [21], we, however, work with the so-called dissipation distance (cf. [22]) instead of the dissipation potential.

[1] In the range of mechanical loading frequencies where assumption of isothermality is satisfied, which will be the regime in which we work here, SMA behave in a rate-independent manner [19, 20].

3. *Mathematical analysis of the model*: We embed our model into the so-called energetic framework due to [23] which allows us to prove existence of solutions to the model by a constructive method that simultaneously provides a conceptual numerical algorithm. Unlike in the previous works concerning the energetic framework, we have to cope here with a dissipation distance *dependent on temperature*; we, hence, extend the framework also to this case.

4. *Numerical implementation*: Based on the mathematical analysis, we develop a suitable numerical approximation of the model, which is used for implementation into the finite element software Abaqus.

5. *Validation of the model by comparison of simulations with experiments*: Implementation of the model is employed in simulations of a superelastic helical spring loaded in tension. We confirm plausibility of the model by comparing numerical results with experiments on a real spring. Let us note that information from such a numerical analysis helped to shed light on the problem of fatigue lifetime of SMA braided stents in [24].

Although most published SMA models concentrate on only some of the aforementioned steps, we present the *full* modeling sequence to assure that the models satisfies the requirements of *thermodynamical consistency* and *mathematical rigor* as well as quantitative prediction of experimental results within a good accuracy.

Let us also note that although we believe that especially the microscopically-motivated modeling ansatz is more general, we adapt the model here to NiTi samples similar to those used in the Roundrobin activity. Therefore, other troublesome phenomena specific to these samples, namely the two-step transformation through intermediate R-phase and material anisotropy, are also incorporated into the model by the approach proposed in [25].

The article is organized as follows. In section 2, we will briefly introduce the internal variables, derive the specific form of the dissipation mechanism and construct the constitutive model. In section 3, we will express the general problem of a quasistatic mechanical loading of a NiTi-based SMA body with prescribed temperature evolution as a mathematical time-evolutionary problem and establish properties of its solutions utilizing an elegant method, namely the notion of energetic solutions, recently developed for rate-independent processes. In section 4, we will briefly summarize the numerical implementation and demonstrate the predictive capabilities of the model.

2 Model Derivation

For model formulation, we adopt the formalism of so-called *generalized standard models* (also called generalized standard materials) (GSM) developed by [21]; so, we have to specify two functions (potentials) to construct a constitutive model. The first one is the energy function, f, expressing the energy stored in the material, the other one is the dissipation function, d, describing the energy transformed into heat by the material during the loading process.

However, ideas presented further on in this paper can be advantageously expressed in the form of so-called *dissipation distance*, δ, due to [23, 26]. This dissipation distance gives the *minimal possible dissipation* when the system transits between two arbitrary states $\boldsymbol{\alpha}_A$ and $\boldsymbol{\alpha}_B$; i.e.

$$\delta(\boldsymbol{\alpha}_A, \boldsymbol{\alpha}_B) := \inf \Big\{ \int_{\boldsymbol{\alpha}_A \to \boldsymbol{\alpha}_B} d(\boldsymbol{\alpha}, \dot{\boldsymbol{\alpha}}) \, \mathrm{d}\Gamma; \text{ over all smooth paths from } \boldsymbol{\alpha}_A \text{ to } \boldsymbol{\alpha}_B \Big\} \qquad (1)$$

Moreover, if the system transits from $\boldsymbol{\alpha}_A$ and $\boldsymbol{\alpha}_B$ in an infinitesimally short time-increment the dissipation distance approximately gives the total dissipated energy on transition between the two states (similar concepts have also been used in e.g. [27, 28]).

Remark 2.1 (From dissipation distance to dissipation function)**.** Clearly, obtaining the dissipation distance in a closed form via formula (1) can be a very hard task. Nevertheless, as in this paper, it can be advantageous to derive (a candidate) dissipation distance by observing a transition between two states in a very short time increment. However, one may ask which requirements such a candidate has to satisfy in order to guarantee that it can be derived from dissipation potential via (1).

It is easily seen that *necessary* requirements for this to hold true are that $\delta(\boldsymbol{\alpha}_A, \boldsymbol{\alpha}_B) \geq 0$ for all admissible states of the system (as the dissipation function is non-negative) and that δ satisfies the so-called *triangle inequality*

$$\delta(\boldsymbol{\alpha}_A, \boldsymbol{\alpha}_B) \leq \delta(\boldsymbol{\alpha}_A, \boldsymbol{\alpha}_C) + \delta(\boldsymbol{\alpha}_C, \boldsymbol{\alpha}_B)$$

for all admissible states of the system $\boldsymbol{\alpha}_A$, $\boldsymbol{\alpha}_B$, $\boldsymbol{\alpha}_C$. As the dissipation potential has to be convex in the second variable for thermodynamic consistency, we also need to demand that

$$\delta(\boldsymbol{\alpha}_A, \boldsymbol{\alpha}_B) \text{ is convex in the second variable.}$$

Moreover, if we *knew* that δ was obtained from some dissipation function through (1) the corresponding dissipation function could be found through

$$d(\boldsymbol{\alpha}, \dot{\boldsymbol{\alpha}}) = \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \delta(\boldsymbol{\alpha}, \boldsymbol{\alpha} + \epsilon \dot{\boldsymbol{\alpha}}), \tag{2}$$

respectively (see [22] for details). However, to the authors' knowledge, it is not known whether these requirements are truly *sufficient*. Yet, working with a dissipation distance satisfying these requirements (together with some coercivity) allows us to work within a *thermodynamically and mathematically consistent framework*; so, we will impose only these in what follows without verifying whether the dissipation distance really stems from a dissipation function.

Within this section, we first introduce internal variables of the model. Then, we present the concept of *transformation favourable martensite*, which is central for ensuing derivation of the proposed dissipation distance. Finally, we briefly describe the form of the free energy.

2.1 Internal variables

As the primary state variables, we choose the total strain tensor, $\boldsymbol{\varepsilon}$, and temperature, T. Following a common approach in GSM [29] that has also been widely used in macroscopic SMA modeling, cf. [12, 14, 30], we introduce two internal variables for description of the state of material. The first, scalar one is the volume fraction of martensite, $\xi \in [0, 1]$. The other one is a tensor variable, $\boldsymbol{\varepsilon}^{\text{in}}$, coming from the conventional small strain decomposition:

$$\boldsymbol{\varepsilon}^{\text{in}} = \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{\text{el}}, \tag{3}$$

where $\boldsymbol{\varepsilon}^{\text{el}}$ represents the common thermoelastic strain tensor. Hence, $\boldsymbol{\varepsilon}^{\text{in}}$ is the only macroscopic variable used for storing information about microscopic internal structure of

martensite.[2] It shall be noted that we do not consider other inelastic strain contributions as plasticity, retained martensite accumulation, etc. since they are beyond the scope of the presented model.

Remark 2.2. One could formally decompose inelastic strain into the following product: $\boldsymbol{\varepsilon}^{\mathrm{in}} = \xi\boldsymbol{\varepsilon}^{\mathrm{tr}}$, where $\boldsymbol{\varepsilon}^{\mathrm{tr}}$ represents a "homogenized" mean transformation strain of martensite and was used as a primary variable (together with ξ) in e.g. [12, 14]. However, as it will turn out later, the choice of the pair $\xi, \boldsymbol{\varepsilon}^{\mathrm{in}}$ is suitable for derivations in Subsection 2.2 and advantageous for mathematical treatment in Section 3.

Since the relative volume change associated with martensitic transformation is very small [33], it is (similarly as in works [10, 12, 14]) neglected. In addition, crystallographic considerations show [34] that there exists a maximum value of strain that is attainable due to phase transformation; so, the inelastic strain is considered to lie in a bounded convex set. Moreover, the area of this set has to shrink linearly if the volume fraction of martensite shrinks.[3] Thus, we further restrict $\boldsymbol{\varepsilon}^{\mathrm{in}}$ as follows:

$$\boldsymbol{\varepsilon}^{\mathrm{in}} \in \left\{\boldsymbol{x} \in \mathbb{R}^{3\times 3} : \boldsymbol{x} \text{ is symmetric, } \mathrm{tr}(\boldsymbol{x}) = 0,\ \langle \boldsymbol{x} \rangle \leq \xi\right\}, \tag{4}$$

where $\mathrm{tr}(\boldsymbol{x})$ denotes the trace of a tensor $\boldsymbol{x}$ and $\langle\cdot\rangle : \mathbb{R}^{3\times 3} \to \mathbb{R}^{+}$ is a suitable positively 1-homogeneous convex function; by a particular form of this function, tension-compression asymmetry and anisotropic material behavior is captured in the model.

As the presented model is adapted to NiTi samples, it is important to capture also the effect of another martensitic-type transformation specific to this alloy: the R-phase transformation [18]. Here, we concentrate only on two important phenomena associated with this transformation: namely, the dramatic change of elastic behavior and the distinct entropy change. To this end, we introduce the volume fraction of R-phase, $\eta \in [0, 1]$, as an additional internal variable[4]; then, since an arbitrary mixture of the three phases (austenite, R-phase, martensite) can be found in the material, the set of natural constraints on ξ, η reads as:

$$0 \leq \eta \leq 1 - \xi \leq 1. \tag{5}$$

2.2 Dissipation

During general thermal and mechanical loading of SMA, mutual interaction between two important dissipative processes, namely phase transformation and reorientation, occurs. To effectively capture this in the model we first formulate a set of assumptions, (A1)-(A7), which actually *build up* the presented model. The first three of them are based on a simplified modeling notion of evolution of the internal structure of martensite during the phase transformation. This notion is inspired by microstructural observations of SMA on the single-crystalline level while formulation of the other assumptions is rather influenced by macroscopic experimental findings. The assumptions are then used to propose a suitable dissipation distance.

[2] This is in contrast to common treatment in micromechanics-based models, where such an information is often stored in a vector of volume fractions and/or transformation strains of all microstructural variants, e.g. [31, 32].

[3] As the aforementioned crystallographic considerations force indeed rather the transformation strain to be uniformly bounded, the upper bound on magnitude of the mean transformation strain must be, with respect to the decomposition in Remark 2.2, multiplied by ξ to obtain the upper bound on $\boldsymbol{\varepsilon}^{\mathrm{in}}$.

[4] Our approach neglects the transformation strain related to transformation between austenite and R-phase. Justification of such a substantial simplification is further discussed in [25].

Note that we concentrate mostly on the dissipation mechanism since, as far as the free energy is concerned, most available macroscopic SMA models consider similar *dominant* energy contributions [5]. The dissipation mechanism is much less explored, though, and so we concentrate on describing implications of the modeling ideas (A1)-(A7) to it.

Remark 2.3 (Neglecting dissipation due to R-phase transformation)**.** The transformation between austenite and R-phase is also of the martensitic type [35]; however, we apply the following simplifications. Since the phase transformation between austenite and R-phase exhibits a very narrow hysteresis in experimental studies [36], we *neglect* the associated dissipation. Accordingly, we suppose that the same amount of energy is dissipated when austenite transforms to martensite through R-phase as when it transforms directly.[5] This could be justified by a substantially smaller change in crystal lattice associated with the austenite-to-R-phase transformation than for the austenite-to-martensite one [35].

Let us stress, at this point, that our derivation is *local* in the sense that our considerations apply to any single arbitrary material point.

2.2.1 Assumptions of a two-step martensitic transformation

During transformation from austenite to martensite a specific form of structure of martensite compatible within the structure of austenite is formed [1]. Thus, requirements of crystallographic coherence between the parent phase and the product phase impose some conditions on the twinned structure of the formed martensite and application of an external load may lead to a further (usually complex) evolution of this structure [37, 38]. Extending this observation for any material point of the polycrystalline sample, we formulate the first assumption as follows:

(A1) *Austenite transforms to martensite always through formation of a particular martensitic structure. This structure will be called transformation favorable martensite (TFM) hereinafter and it is characterized by zero net macroscopic strain at the moment of creation. Immediate subsequent reorientation of TFM is possible.*

Let us stress that formation of TMF is presumed not only in purely temperature-induced transformation, but also in every transformation from austenite to martensite induced by any combination of stress and temperature changes.

In [37, 39] it was observed that compatible interfacial structure forms and develops when martensite, regardless whether twinned or detwinned, transforms back to austenite. Motivated by such observations, the idea of TFM is now extended to the reverse transformation, too; i.e., we suppose that TFM is always present also during disappearance of martensite:

(A2) *A necessary condition for initialisation of a reverse transformation in any amount of martensite is formation of TFM within it.*

The formulation emphasizes that if the actual martensite is not the TFM, then the structure of the disappearing martensite is *adjusted first* to TFM; only then the reverse transformation may proceed.

Due to the previous assumptions, we can assign two additive contributions to dissipation during thermomechanical loading of SMAs. The first one is connected with phase transformation between austenite and TFM and we shall call it *transformation part of*

[5]The same assumptions has recently been used in the micromechanical model by [32].

dissipation distance, δ^{tr}, in the following. The second contribution is connected to structure evolution of martensite and it is called *reorientation part of dissipation distance*, δ^{reo}. Thus, we arrive at a quite natural assumption:

(A3) *The contributions to dissipation associated to formation/disappearence of TMF and reorientation of martensite are additive.*

Hence, $\delta = \delta^{\mathrm{tr}} + \delta^{\mathrm{reo}}$.

Remark 2.4. Distinction between dissipation related to transformation and reorientation can be found in several other recent SMA models as in those of [14] or [12]. In our model it emerges inherently due to the notion expressed by (A1) and (A2).

Remark 2.5. Making profit from experimental observations, we attempted to identify those characteristics of the phase transformation at the single crystalline level which are, in our opinion, *of key importance* for description of dissipation at the polycrystalline level. Of course, the real situation in polycrystals is much more complex than in single crystals. Nevertheless, we believe that albeit simplified, our assumptions capture the essential ingredient of the evolution of martensitic structure in polycrystals.

2.2.2 Derivation of the transformation part of dissipation

If austenite is cooled under stress-free conditions, martensite phase transformation starts at a temperature denoted usually M_{s} and finishes at temperature M_{f} (i.e. $M_{\mathrm{f}} \leq M_{\mathrm{s}}$). Similarly, the reverse transformation occurs in a temperature range from A_{s} to A_{f} under stress-free conditions (i.e. $A_{\mathrm{s}} \leq A_{\mathrm{f}}$) forming together a hysteresis loop in a closed transformation cycle. Such conditions correspond well to a situation where only energy related to the formation of TFM is dissipated. Thus, the dissipation distance should depend, in such a case, *only in the volume fraction of martensite in the initial and finite state*. Since more energy is dissipated when a larger volume fraction of the material transforms and *due to the rate-independence* we assume the following:

(A4) *The transformation part of dissipation distance is proportional to the absolute value of the difference between its initial and final volume fraction of martensite. The positive proportionality factor may depend (linearly) on the volume fraction.*

For simplicity, we assume that the dependence of the proportionality factor in (A4) on the volume fraction is linear. This corresponds to the situation sketched in Fig. 1 and the corresponding form reads, then, as (we again distinguish initial and terminal state by subscripts A and B, respectively):

$$\xi_B \geq \xi_A : \quad \delta^{\mathrm{tr}}(\xi_A, \xi_B) = \Delta s^{AM} \left[(T_0 - M_{\mathrm{s}}) + \frac{\xi_A + \xi_B}{2} (M_{\mathrm{s}} - M_{\mathrm{f}}) \right] |\xi_B - \xi_A|, \tag{6}$$

$$\xi_B < \xi_A : \quad \delta^{\mathrm{tr}}(\xi_A, \xi_B) = \Delta s^{AM} \left[(A_{\mathrm{f}} - T_0) + \frac{\xi_A + \xi_B}{2} (A_{\mathrm{s}} - A_{\mathrm{f}}) \right] |\xi_B - \xi_A|. \tag{7}$$

where Δs^{AM} and T_0 adjust the transformation temperatures with respect to the particular form of f, see (12).

Remark 2.6. Let us note that corresponding type of dissipation function was analytically studied by [40] in model denoted "M2" in that work. Choice $M_{\mathrm{s}} = M_{\mathrm{f}} < T_0$ and $A_{\mathrm{s}} = A_{\mathrm{f}} > T_0$ leads to the situation denoted "M1" in that work and corresponds to the simplest non-trivial dissipative mechanism.

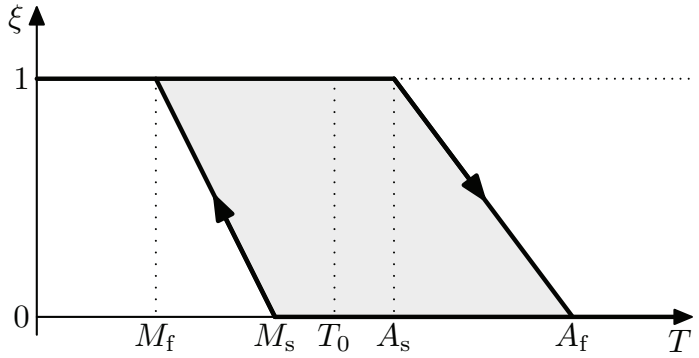

Figure 1: The evolution of volume fraction of martensite with temperature under stress-free condition in case only the transformation dissipation is considered in the model.

Remark 2.7. One may ask whether the form of the proportionality factor as given in, say, (6) is the only possible as given by thermodynamic limitations. This is not entirely so; however, when assuming linearity in the volume fraction, several restrictions can be derived only from these consideration. To illustrate some of them, take the most general form of δ^{tr} which reads as

$$\big(C + A\xi_A + B\xi_B\big)|\xi_B - \xi_A|, \tag{8}$$

with some A, B, C. For definiteness, let us concentrate only the case when $\xi_B \geq \xi_A$; then, as the proportionality factor has to be positive, we deduce that $C > 0$. Moreover, as the dissipation distance needs to be convex in the second variable, also $B \geq 0$. As the dissipation distance needs to fulfil the triangle inequality we obtain that $A \geq B$; this can be seen by taking $\xi_B = 1$, $\xi_A = 0$ and the intermediate step $\xi_C = 1/2$. So, in particular, the proportionality factor could not depend, e.g. *only* on the last step. We also have that $A < B + C$. To see, this take $\xi_B = 1$ and $\xi_A = \epsilon$ for some small ϵ small. Now as in this case a smaller amount of martensite transforms than in the case when the starting state was zero. So,

$$\big(C + A\epsilon + B)(1-\epsilon) < B + C \quad \Rightarrow (A - C - B)\epsilon - A\epsilon^2 < 0$$

for all $\epsilon > 0$, which can only be satisfied if the above condition holds. Even more conditions could be derived when testing the triangle inequality jointly in the case when $\xi_B \geq \xi_A$ and $\xi_B < \xi_A$.

Note that the constants chosen in (6)-(7) indeed satisfy all the above constrains.

Remark 2.8. The value of reference temperature T_0 determines the ratio between energy dissipated during forward and reverse transformation.

2.2.3 Derivation of the reorientation part of dissipation

Microstructure observations reveal that several types of structure rearrangement may be involved in evolution of martensite due to applied stress [38, 41]. All such processes are related to existence of twin boundaries within the material and their movement contributes to change of the inelastic strain of the specimen. In our model, description of complex internal structure is simplified to one tensorial variable. Thus, we neglect possible dependence of amount of dissipated energy on type of internal microstructure rearrangement and make following simple assumption:

(A5) *If martensite is subject to reorientation, the corresponding dissipation distance is proportional to the norm of the difference between its final and initial inelastic strain.*

A proportionality coefficient in units of stress must appear in formula for the dissipation distance so that the resulting dissipation distance keeps the physical dimension of energy. It seems natural to relate this coefficient to stress needed for initialization of martensite reorientation at temperatures below M_s. Since it is often observed [42, 43] that this stress decreases with increasing temperature,[6] we, finally, make an assumption on the temperature dependence of dissipation related to reorientation as follows:

(A6) *Dissipation accompanying reorientation processes is proportional to temperature.*

We denote the proportionality coefficient $\sigma^{\mathrm{reo}}(T)$.

Now, we are going to formulate the reorientation dissipation distance explicitly. When austenite to martensite transformation occurs (i.e. $\xi_B > \xi_A$), we directly follow assumptions (A5) and (A6) and obtain

$$\xi_B > \xi_A : \quad \delta^{\mathrm{reo}}(\boldsymbol{\varepsilon}_A^{\mathrm{in}}, \xi_A, T_B, \boldsymbol{\varepsilon}_B^{\mathrm{in}}, \xi_B) = \sigma^{\mathrm{reo}}(T_B)\|\boldsymbol{\varepsilon}_B^{\mathrm{in}} - \boldsymbol{\varepsilon}_A^{\mathrm{in}}\|. \tag{9}$$

An analogous consideration leads to extending the definition also for pure reorientation, i.e. $\xi_B = \xi_A$.

However, the situation is less clear when some martensite disappears (i.e. $\xi_A > \xi_B$). The disappearing martensite and remaining martensite shall be treated separately, since they differ in the final state of inelastic strain. In line with assumption (A2), the disappearing martensite forms TFM, which is characterized by zero inelastic strain. Yet, by this assumption it is still not evident how much of the inelastic strain $\boldsymbol{\varepsilon}_A^{\mathrm{in}}$ is lost by transformation of the disappearing martensite to TFM. A rather sophisticated approach to this issue based on substitution of internal variable $\boldsymbol{\varepsilon}^{\mathrm{in}}$ by an internal function was proposed in [44], where, however, only the one-dimensional situation is tackled which cannot be straightforwardly adapted to the 3D case. So, we confine ourselves to the following simple assumption:

(A7) *The inelastic strain is distributed homogeneously within martensite.*

Hence, the initial value of the inelastic strain corresponding to the disappearing martensite is then proportional the ratio $(\xi_A - \xi_B)/\xi_A$. Considering the temperature dependence as in (9), the dissipation distance reads

$$\delta_1^{\mathrm{reo}}(\boldsymbol{\varepsilon}_A^{\mathrm{in}}, \xi_A, T_B, \boldsymbol{\varepsilon}_B^{\mathrm{in}}, \xi_B) = \sigma^{\mathrm{reo}}(T_B)\Big\|\frac{\xi_A - \xi_B}{\xi_A}\boldsymbol{\varepsilon}_A^{\mathrm{in}} - \mathbf{0}\Big\| = \sigma^{\mathrm{reo}}(T_B)\Big\|\boldsymbol{\varepsilon}_A^{\mathrm{in}} - \frac{\xi_B}{\xi_A}\boldsymbol{\varepsilon}_A^{\mathrm{in}}\Big\|.$$

Moreover, the remaining martensite, whose initial inelastic strain is then proportional to the ratio ξ_B/ξ_A, can reorient, which adds the second term

$$\delta_2^{\mathrm{reo}}(\boldsymbol{\varepsilon}_A^{\mathrm{in}}, \xi_A, T_B, \boldsymbol{\varepsilon}_B^{\mathrm{in}}, \xi_B) = \sigma^{\mathrm{reo}}(T_B)\Big\|\boldsymbol{\varepsilon}_B^{\mathrm{in}} - \frac{\xi_B}{\xi_A}\boldsymbol{\varepsilon}_A^{\mathrm{in}}\Big\|. \tag{10}$$

The sum $\delta_1^{\mathrm{reo}} + \delta_2^{\mathrm{reo}}$ gives the the reorientation dissipation distance for martensite transforming to austenite

$$\xi_B < \xi_A : \quad \delta^{\mathrm{reo}}(\boldsymbol{\varepsilon}_A^{\mathrm{in}}, \xi_A, T_B, \boldsymbol{\varepsilon}_B^{\mathrm{in}}, \xi_B) = \sigma^{\mathrm{reo}}(T_B)\left[\Big\|\frac{\xi_B}{\xi_A}\boldsymbol{\varepsilon}_A^{\mathrm{in}} - \boldsymbol{\varepsilon}_A^{\mathrm{in}}\Big\| + \Big\|\boldsymbol{\varepsilon}_B^{\mathrm{in}} - \frac{\xi_B}{\xi_A}\boldsymbol{\varepsilon}_A^{\mathrm{in}}\Big\|\right]. \tag{11}$$

[6] This could be associated with increasing mobility of twin boundaries with increasing temperature.

Remark 2.9. To the authors' knowledge there is no systematic experimental study which would provide deeper insight on situation discussed in the previous paragraph. Nevertheless, since the constitutive model is intended for a finite element method calculations, a "homogeneity" approximation like (A7) is reasonable (cf. [12].)

2.3 Free energy

We consider the free energy consisting only of two dominant contributions, namely the elastic and chemical energy, which are pivotal in SMA models, e.g. [10, 15, 45, 46]. The rule of mixtures is employed to determine the specific free energy of the multiphase material and the Reuss model for strain decomposition is adopted. We omit the details in this work and refer the reader to [25], where the procedure yielding the Helmholtz free energy

$$\begin{aligned} f(T,\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}^{\mathrm{tr}},\xi,\eta) &= \frac{1}{2}K\mathrm{tr}(\boldsymbol{\varepsilon})^2 + G(\xi,\eta)\|\mathrm{dev}(\boldsymbol{\varepsilon}) - \boldsymbol{\varepsilon}^{\mathrm{in}}\|^2 \\ &\quad + \Delta s^{AM}(T-T_0)\xi + \Delta s^{AR}(T - T_0^R(\xi,\eta))\eta \\ &\quad + u_0^A - s_0^A T + c^A\left[(T-T_0) - T\ln\left(\frac{T}{T_0}\right)\right]. \end{aligned} \tag{12}$$

is described. Note that the superscripts ^{A},^{R} and ^{M} distinguish entities related to austenite, R-phase and martensite, respectively. The constant K denotes the bulk modulus (considered independent on the phase composition), G the shear modulus, Δs the difference of specific entropy per unit volume. u^A, s^A and c^A are the specific internal energy, entropy and heat capacity of austenite at a reference temperature T_0.[7] $\mathrm{dev}(\boldsymbol{x})$ denotes the deviatoric part of a tensor $\boldsymbol{x}$.

Although the transformation between austenite and R-phase is considered non-dissipative, the form of the free energy allows to capture two important effects related that transformation. First, due to considered different values of shear modulus of the phases, the substantial change of the elastic behavior observed e.g. in [47] is covered. Second, the term proportional to Δs^{AR} reflects the distinct change of entropy driving the transformation process. Moreover, to account for a possibly wide temperature range in which the austenite-to-R-phase transformation takes place, we define T_0^R as

$$T_0^R(\xi,\eta) := R_{\mathrm{s}} + \frac{\eta}{1-\xi}\frac{R_{\mathrm{f}} - R_{\mathrm{s}}}{2}, \tag{13}$$

where R_{s} and R_{f} denote the initial and final temperature of the transformation from austenite to R-phase, i.e. $R_{\mathrm{s}} > R_{\mathrm{f}}$. See [25] for details.

To sum up the final form of free energy and dissipation distance (recall (A3)) reads as

[7] With respect to their physical meaning, we assume that all these constants are positive in Section 3.

follows:

$$
\begin{aligned}
f(T,\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}^{\mathrm{in}},\xi,\eta) &= \frac{1}{2}K\mathrm{tr}(\boldsymbol{\varepsilon})^2 + G(\xi,\eta)\|\mathrm{dev}(\boldsymbol{\varepsilon})-\boldsymbol{\varepsilon}^{\mathrm{in}}\|^2 \\
&\quad + \Delta s^{AM}(T-T_0)\xi + \Delta s^{AR}(T-T_0^R(\xi,\eta))\eta \\
&\quad + u_0^A - s_0^A T + c^A\left[(T-T_0) - T\ln\left(\frac{T}{T_0}\right)\right], \qquad (14)
\end{aligned}
$$

$$
\delta(T,\boldsymbol{\varepsilon}_A^{\mathrm{in}},\xi_A,\boldsymbol{\varepsilon}_B^{\mathrm{in}},\xi_B) = \begin{cases} \Delta s^{AM}\left[T_0 - M_{\mathrm{s}} + \frac{\xi_A+\xi_B}{2}(M_{\mathrm{s}}-M_{\mathrm{f}})\right]|\xi_B-\xi_A| & \\ \quad + \sigma^{\mathrm{reo}}(T_B)\|\boldsymbol{\varepsilon}_B^{\mathrm{in}}-\boldsymbol{\varepsilon}_A^{\mathrm{in}}\| & \text{if } \xi_B \geq \xi_A, \\ \Delta s^{AM}\left[A_{\mathrm{f}} - T_0 + \frac{\xi_A+\xi_B}{2}(A_{\mathrm{s}}-A_{\mathrm{f}})\right]|\xi_B-\xi_A| & \qquad (15) \\ \quad + \sigma^{\mathrm{reo}}(T_B)\left[\|\frac{\xi_B}{\xi_A}\boldsymbol{\varepsilon}_A^{\mathrm{in}}-\boldsymbol{\varepsilon}_A^{\mathrm{in}}\| + \|\boldsymbol{\varepsilon}_B^{\mathrm{in}}-\frac{\xi_B}{\xi_A}\boldsymbol{\varepsilon}_A^{\mathrm{in}}\|\right] & \text{if } \xi_B < \xi_A. \end{cases}
$$

Remark 2.10. By a straightforward calculation (cf. [48]) it can be shown that dissipation distance proposed here satisfies the necessary conditions from Remark 2.1; namely, it is non-negative, convex in the second variable and, importantly, satisfies the triangle inequality.

Remark 2.11. As outlined in Remark 2.1, if the dissipation function were derived from some dissipation potential, this had to be determined by (2). It is worth mentioning that if this (candidate) dissipation function is derived, it corresponds to the one heuristically obtained in [25] by careful analysis of experimental phase diagrams. Utilizing this dissipation function, we could formulate the standard governing equations/inclusions within the GSM framework to obtain time-evolutionary problem of quasistatic mechanical loading of a NiTi SMA body with prescribed temperature evolution, see [25]. It is also worth mentioning, that, even though we shall not formulate these equations/inclusions and rather give an energetic formulation of the problem, after performing time-discretization the *same* problem as in [25] is obtained.

Remark 2.12. It is apparent by comparing formulas in (15) that the dissipation distance is continuous for $\xi_B = \xi_A$.

Remark 2.13. In accord with experimental observations, we assume that described processes are strongly dissipative, i.e. $\boldsymbol{\alpha}_A \neq \boldsymbol{\alpha}_B \Rightarrow \delta(\boldsymbol{\alpha}_A,\boldsymbol{\alpha}_B) \neq 0$. With respect to this, in our case it is sufficient to put

$$
M_{\mathrm{s}} < T_0 < A_{\mathrm{s}} \quad \text{and} \quad \sigma^{\mathrm{reo}}(T) > 0 \qquad (16)
$$

for any T considered.

3 Mathematical Analysis

In this section, we examine the mathematical properties of the proposed model and embed it into the energetic framework due to Mielke and collaborators (cf. e.g. the works [26, 49, 50, 51]) which we generalize for dissipation distances dependent on time through the external loading by temperature. In particular, we also propose a conceptual numerical scheme that is further used in Section 4 and prove convergence to so-called *energetic solutions*. This analysis shows that our model is well-behaving in the sense that we can prove that solutions to it exist in a natural sense;[8] moreover, the numerical approximation

[8] Importance of such an analysis for SMA models was highlighted in [52] for instance.

is, due to the presented analysis, expected to be well-behaving, for example to be *mesh independent*.

Let us start by presenting a *global formulation* of our model.

3.1 Global formulation

The constitutively defined Helmholtz energy and dissipation distance (14)–(15) are the starting point for the energetic formulation, as outlined in e.g. [26, 51]. However, instead of the Helmholtz free energy, we shall rather include the force terms into the energy and work in terms of the *global* Gibbs free energy $\mathcal{E}$ and the *global* dissipation distance $\mathcal{D}$.

In the following, we denote by Ω the *reference configuration* (usually the stress-free austenitic state of the specimen) and $\partial\Omega$ denotes its boundary. We assume the following splitting $\partial\Omega = \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}} \cup N$ with Γ_{D}, Γ_{N} open, disjoint and the two-dimensional measure of N being zero. On the part Γ_{N} the surface force F_{surf} is acting on the specimen while on the part Γ_{D} Dirichlet boundary conditions for the displacement are prescribed. For simplicity, we restrict ourselves here to *zero Dirichlet boundary conditions*; i.e., $u(x) = 0$ on Γ_{D}. Let us note that, by shifting, considering any other constant-in-time Dirichlet boundary condition is immediate; sufficiently smooth time-dependent Dirichlet boundary conditions could be, again by shifting, incorporated , too – this would, however, make the analysis presented bellow more technical.

To shorten the notation, let us introduce a vector of dissipative variables

$$\alpha := (\boldsymbol{\varepsilon}^{\mathrm{in}}, \xi).$$

With this notation, we introduce the *global*, i.e. integrated over the reference configuration, dissipation distance (recall (15)):

$$\mathcal{D}(t, \alpha, \tilde{\alpha}) := \mathcal{D}^{\mathrm{tr}}(t, \alpha, \tilde{\alpha}) + \mathcal{D}^{\mathrm{reo}}(t, \alpha, \tilde{\alpha}) = \int_\Omega \delta^{\mathrm{tr}}(t, \alpha, \tilde{\alpha})\,\mathrm{d}x + \int_\Omega \delta^{\mathrm{reo}}(t, \alpha, \tilde{\alpha})\,\mathrm{d}x. \tag{17}$$

Remark 3.1. Due to the triangle inequality mentioned in Remark 2.10, the global dissipation satisfies that

$$\mathcal{D}(t, \alpha_A, \alpha_C) \leq \mathcal{D}(t, \alpha_A, \alpha_B) + \mathcal{D}(t, \alpha_B, \alpha_C), \tag{18}$$

for any fixed time t. Moreover, the following coercivity property

$$\mathcal{D}(t, \alpha_A, \alpha_B) \geq c^* \|\alpha_A - \alpha_B\|_{L^1(\Omega)}. \tag{19}$$

with some $c^* > 0$ reflects Remark 2.13.

We have the following form of the global Gibbs free energy

$$\mathcal{E}(t, u, \alpha, \eta) := \int_\Omega f(t, \varepsilon(u), \alpha, \eta) + \nu\|\nabla\alpha(t)\|^2 - F_{\mathrm{vol}}(t) \cdot u\,\mathrm{d}x - \int_{\Gamma_{\mathrm{N}}} F_{\mathrm{surf}}(t) \cdot u\,\mathrm{d}S, \tag{20}$$

where F_{vol} is the prescribed volume force acting on the specimen and $\nu > 0$ is an arbitrary (small) constant. The term $\nu\|\nabla\alpha(t)\|^2$ is then a regularization term as it penalizes fast spatial changes of the internal variables. This term can be linked to capillarity effects [53]; however, we include it in order to make the analysis tractable – it could be avoided in some special cases, see e.g. [51], which we do not consider here. Note that, apart from the regularization term, (20) represents the standard global form of the Gibbs free energy corresponding to the Helmholtz free energy $f(t, \varepsilon(u), \alpha, \eta)$.

In accordance with the concept of generalized standard materials (cf. Section 2), the two functionals $\mathcal{E}$, $\mathcal{D}$ determine the *rate-independent evolution* of the state and internal variables of the specimen; however the values of these variables are limited by natural constrains expressed in (4) and (5). Thus, we shall define the set of admissible states $\mathcal{Q}$ through

$$\mathcal{Q} := \left\{(u,\alpha,\eta) \in \mathcal{U}\times\mathcal{V}\times\mathcal{Z} \ : \eta(x) \le 1-\xi(x) \text{ for a.a. } x\in\Omega \text{ with } \alpha = (\boldsymbol{\varepsilon}^{\mathrm{in}},\xi)\right\} \tag{21}$$

where

$$\mathcal{U} := \{u \in W^{1,2}(\Omega,\mathbb{R}^3) : u = 0 \text{ on } \Gamma_{\mathrm{D}}\} \tag{22}$$

$$\begin{aligned}\mathcal{V} := \{(\boldsymbol{\varepsilon}^{\mathrm{in}},\xi) \in W^{1,2}(\Omega,\mathbb{R}^{3\times 3})\times W^{1,2}(\Omega) \ : \ &\boldsymbol{\varepsilon}^{\mathrm{in}} \text{ is a traceless, symmetric matrix,}\\ &\langle\boldsymbol{\varepsilon}^{\mathrm{in}}(x)\rangle \le \xi(x) \text{ for a.a.} x\in\Omega \text{ and } 0\le\xi(x)\le 1 \text{ for a.a. } x\in\Omega\}\end{aligned} \tag{23}$$

$$\mathcal{Z} := \{\eta \in L^\infty(\Omega,\mathbb{R})) \ : \ 0\le\eta(x)\le 1 \text{ for a.a. } x\in\Omega\}. \tag{24}$$

Note that, we also included the Dirichlet boundary condition into the space $\mathcal{U}$.

The triple $(\mathcal{Q},\mathcal{E},\mathcal{D})$ then defines a *rate-independent system* the solutions to which describe, in our modeling approach, the evolution of a polycrystalline SMA-specimen.

Finally, let us introduce the overall dissipated energy over the path $(u,\alpha,\eta) : [0,\mathcal{T}] \mapsto \mathcal{Q}$ (when actually only the path in α's plays a role since those are the dissipative variables):

$$\mathrm{Diss}_{\mathcal{D}}(\alpha;[0,\mathcal{T}]) := \mathrm{Diss}^{\mathrm{tr}}(\alpha;[0,\mathcal{T}]) + \mathrm{Diss}^{\mathrm{reo}}(\alpha;[0,\mathcal{T}]). \tag{25}$$

First, we define that the overall energy loss due to transformation can be computed as follows:

$$\mathrm{Diss}^{\mathrm{tr}}(\alpha;[0,\mathcal{T}]) := \sup\Big\{\sum_{i=1}^{N}\mathcal{D}^{\mathrm{tr}}(\alpha(t_{i-1}),\,\alpha(t_i)) \ : \ \text{a.p.p. } \ 0=t_0\le t_1\le\ldots\le t_N=\mathcal{T}\Big\}, \tag{26}$$

where we abbreviated "over all possible partitions" by "a.p.p.". This definition follows the works [26, 51] and is motivated by the idea that if the dissipative variables change smoothly then this formula integrates up all the losses due to this process (cf. Remark 3.2). However, it can incorporate also more general situations; for example, it can be easily seen that if the values of the internal dissipative variables jump in finitely many points of the considered time-interval, then (26) counts exactly the energy dissipated during these jumps.

As far as the reorientation losses are concerned, we know that the reorientation part of the dissipation depends on the temperature which, roughly speaking, represents a "weight" on the energy dissipated due to the change on the dissipative variables. To capture this phenomenon we define

$$\mathrm{Diss}^{\mathrm{reo}}(\alpha,[0,t]) := \int_0^t \sigma^{\mathrm{reo}}(T(s))\,\mathrm{d}\mu_\alpha(s). \tag{27}$$

Here we denoted by μ_α the Radon measure defined on $[0,\mathcal{T}]$ by prescribing its values on every closed set $S=[s,z]\subset[0,\mathcal{T}]$, as

$$\mu_\alpha(S) := \sup\left\{\sum_{i=1}^{N}\overline{\mathcal{D}}^{\mathrm{reo}}(\alpha(t_{i-1}),\alpha(t_i)) \ : \ \text{a.p.p. } \ s=t_0\le t_1\le\ldots\le t_N=z\right\}, \tag{28}$$

and thus for any Borel subset $A \subset [0, \mathcal{T}]$

$$\mu_\alpha(A) = \sup\left\{\sum_i \mu_j(V_i) : \bigcup_i V_i \subset A \text{ with } V_i \text{ closed intervals from } [0,\mathcal{T}]\right\}, \tag{29}$$

where

$$\begin{aligned}\overline{\mathcal{D}}^{\rm reo}(\alpha(t_{i-1}),\alpha(t_i)) &:= \int_\Omega \overline{\delta}^{\rm reo}(\alpha(t_{i-1}),\alpha(t_i))\,{\rm d}x\\ &:= \int_\Omega \delta^{\rm reo}(t,\alpha(t_{i-1}),\alpha(t_i))\,{\rm d}x \text{ with } \sigma^{\rm reo}(T(t)) \equiv 1;\end{aligned} \tag{30}$$

hence, the measure itself *does not depend* on the prescribed function $\sigma^{\rm reo}(T(t))$.

The principle behind this definition is similar as above; the usage of the measure allows us to incorporate the dependence of the dissipated energy on the temperature. Again, it is instructive to observe an evolution of the internal variable α that exhibits only a finite number of jumps. In such a case (27) together with (28) counts exactly the dissipated energy at the jumps weighted by the actual temperature at the jumps.

Remark 3.2. It follows from the theory of functions of bounded variation (see e.g. [54]) that the definition of the total dissipated energy over the prescribed path is designed in such a way that, for sufficiently smooth processes,

$$\mathrm{Diss}_{\mathcal{D}}(\alpha;[0,\mathcal{T}]) = \int_0^{\mathcal{T}}\int_\Omega d(t,\alpha(t),\dot\alpha(t))\,{\rm d}x\,{\rm d}t,$$

if δ is such that it is associated to the dissipation function d, see Remark 2.1.

Remark 3.3 (Convexity of the free energy)**.** The free energy $f(t,\varepsilon(v(t)),\alpha(t),\eta(t))$ defined through (14) can be shown to be convex by calculating the Hessian of f which turns out to be positive semi-definite if the constrains $0 \le \xi \le 1$, $0 \le \eta \le 1-\xi$ hold. Actually, one relies here on the fact that the functions $h(x,y) = \frac{x^2}{y}$ and $\tilde h(x,y) = \frac{x^2}{1-y}$ are convex, provided $0 \le y \le 1$.

Let us note that this property is available *because* we work with the inelastic strain and not the transformation strain mentioned in Remark 2.2.

3.2 Definition of energetic solutions, data qualifications

Following [26, 49, 50, 51], we define:

Definition 3.4 (Energetic solution)**.** Let $\mathcal{E},\mathcal{D}$ be given by definitions (17) and (20) and let $(u_0,\alpha_0,\eta_0) \in \mathcal{Q}$. Then triplet $(u(t),\alpha(t),\eta(t)) : [0,\mathcal{T}] \mapsto \mathcal{Q}$ is called an *energetic solution* to the rate-independent system $(\mathcal{Q},\mathcal{E},\mathcal{D})$ if $\partial_t\mathcal{E}(t,u(t)) \in L^1([0,\mathcal{T}])$ and if the following conditions are satisfied:

- **Stability condition:**

$$\mathcal{E}(t,u(t),\alpha(t),\eta(t)) \le \mathcal{E}(t,\tilde u,\tilde\alpha,\tilde\eta) + \mathcal{D}(t,\alpha(t),\tilde\alpha) \quad \forall\,(\tilde u,\tilde\alpha,\tilde\eta) \in \mathcal{Q} \tag{31}$$

- **Energy balance:**

$$\mathcal{E}(t,u(t),\alpha(t),\eta(t)) + \mathrm{Diss}_{\mathcal{D}}(\alpha,[0,\mathcal{T}]) = \mathcal{E}(0,u(0),\alpha(0),\eta(0)) + \int_0^{\mathcal{T}} \partial_t\mathcal{E}(s,u(s))\,{\rm d}s. \tag{32}$$

- **Initial condition:** $(u(0), \alpha(0), \eta(0)) = (u_0, \alpha_0, \eta_0)$ a.e. in Ω.

Remark 3.5 (Justification of the definition from a mechanical point of view)**.** In [26] the definition of energetic solutions has been linked to the so-called *realizability* principle which postulates that whenever a process *is admissible* from the point of view of thermodynamics it *will happen immediately*. This principle then justifies the stability condition which essentially states that a state can be *only* stable if a process leading to a lower energy state is thermodynamically not admissible. It shall be noted that this approach does not lead to the same results as other ones, e.g. vanishing viscosity limits [55, 56, 57]. Nevertheless, we use it since it adapts perfectly to our modeling setup.

Remark 3.6 (Relation to classical variational inequalities in the framework of generalized standard materials)**.** It has been proved in [58] that if specific convexity/continuity conditions on $\mathcal{E}$ are satisfied and also conditions on the dissipation function corresponding to the dissipation distance $\mathcal{D}$ are posed, then there exists a *unique* energetic solution which moreover satisfies the weak formulation of the classical doubly non-linear inclusion appearing in the framework of generalized standard solids (see e.g. [59]). However, since the convexity condition demanded in [58] is stricter than strict convexity (which is not satisfied in our case) the results are not applicable to our model. One could try to improve the situation by introducing regularizing terms which, in some cases, can be even linked to physically relevant phenomena like hardening (see [25]); but this is beyond the scope of the present discussion.

Finally in this subsection, we summarize the needed **data qualifications**.

(D1) *Domain of the specimen*: Ω is a regular, bounded, Lipschitz domain; $\Gamma_{\rm D} \subset \partial\Omega$ is such that its two-dimensional measure is positive.

(D2) *Domain of the inelastic strain*: $\langle\cdot\rangle$ is convex, positively 1-homogenous and it holds $\langle \boldsymbol{\varepsilon}^{\rm in}\rangle = 0 \Rightarrow \boldsymbol{\varepsilon}^{\rm in} = 0$,

(D3) *Forces*: Let $F_{\rm vol} \in C^1([0,\mathcal{T}], L^2(\Omega;\mathbb{R}^3))$ and $F_{\rm surf} \in C^1([0,\mathcal{T}], L^2(\Gamma_{\rm N};\mathbb{R}^3))$.

(D4) *Prescribed temperature*: Let $T \in C^1([0,\mathcal{T}])$.

Remark 3.7. Since we are considering sufficiently small specimen (e.g. wires) heated uniformly on the boundary, we can assume that the temperature within the specimen is constant and fully determined by the reservoir into which the specimen is placed; hence (similarly as in [60, 61]) we can impose the regularity (D4). Let us note, however, that in case of a larger specimen when the heat equation is included into the governing system, such a regularity is usually not obtained (cf. e.g. [62, 63]).

3.3 Existence theory

In this subsection, we follow the works [23, 51] and also [60, 61] to prove existence of energetic solutions to the rate-independent system $(\mathcal{Q}, \mathcal{E}, \mathcal{D})$. Let us just note that in addition to these works (and in particular also to the latter ones), we cope here with a dissipation distance dependent on temperature which calls for a modification of the available proofs. This is outlined, more specifically, in *Step 6* and *Step 7* of the proof of Theorem 3.8 below.

Theorem 3.8 (Existence of an energetic solution)**.** *Let $\mathcal{E}, \mathcal{D}$ be defined by* (20) *and* (17)*, let (D1)-(D4) hold true and let initial conditions $(u_0, \alpha_0, \eta_0) \in \mathcal{Q}$ satisfy the stability*

condition (31). *Then there exists an energetic solution of the rate-independent system* $(\mathcal{Q}, \mathcal{E}, \mathcal{D})$.

Sketch of proof. For lucidity, we divide the proof into 7 steps. Let us stress, at this point, that hereinafter C will denote a positive *generic* constant that may vary from expression to expression.

Step 1 (Time-incremental problem):
Let $0 = t_0 \leq t_1 \leq \dots t_{N(\tau)} = \mathcal{T}$ be a partition of the time-interval $[0, \mathcal{T}]$ such that $\max_i(t_{i+1} - t_i) \leq \tau$. In the spirit of [23, 51] we design a time-discretization of (31)–(32) via the backward Euler method. To be more specific, we call the triple $(u_\tau^k, \alpha_\tau^k, \eta_\tau^k) \in \mathcal{Q}$ a *discrete energetic solution* to rate-independent system $(\mathcal{Q}, \mathcal{E}, \mathcal{D})$ at time-level $k = 1, \dots, N(\tau)$ if it solves

$$\begin{aligned} &\text{Minimize } \mathcal{E}(t_k, u, \alpha, \eta) + \mathcal{D}(t_k, \alpha_\tau^{k-1}, \alpha) \\ &\text{subject to } (u, \alpha, \eta) \in \mathcal{Q} \end{aligned} \tag{TIP}$$

with $(u_\tau^0, \alpha_\tau^0, \eta_\tau^0) = (u_0, \alpha_0, \eta_0) \in \mathcal{Q}$ defined through the initial condition.

The existence of solutions to (TIP) is shown by the *direct method* (cf. [64]); to this end we mostly rely on the fact that the cost function in (TIP) is *convex* (cf. Remark 3.3) and the minimization is performed on a closed convex set. In addition, the cost function in (TIP) is coercive in the sense that $\mathcal{E}(t_k, u, \alpha, \eta) \geq C(\|u\|_{W^{1,2}(\Omega;\mathbb{R}^3)} + \|\alpha\|_{W^{1,2}(\Omega;\mathbb{R}^{3\times 3}_{\times}\mathbb{R})})$ whenever $(u, \alpha, \eta) \in \mathcal{Q}$.

As to the coercivity, we use the regularization terms and the uniform bounds on α imposed through the set $\mathcal{V}$ to get the coercivity in α. To show the coercivity in u, we exploit the elastic part of the Helmholtz free energy combined with Korn's inequality and assumption (D4) that allows us to estimate the force terms.

Thus, we take an infimizing sequence of the cost functional in (TIP) $([u^k]_j, [\alpha^k]_j, [\eta^k]_j) \in \mathcal{Q}$ and, due to the coercivity, we can extract a weakly* converging subsequence in $W^{1,2}(\Omega;\mathbb{R}^3) \times W^{1,2}(\Omega;\mathbb{R}\times\mathbb{R}^3) \times L^\infty(\Omega)$ (not relabelled); the weak* limit is denoted $(u^k, \alpha^k, \eta^k) \in W^{1,2}(\Omega;\mathbb{R}^3) \times W^{1,2}(\Omega;\mathbb{R}\times\mathbb{R}^3) \times L^\infty(\Omega)$. Because closed convex sets are weakly* closed we actually have that $(u^k, \alpha^k, \eta^k) \in \mathcal{Q}$. Finally, due to convexity of the the cost function, (u^k, α^k, η^k) is indeed a solution to (TIP).

Step 2 (Discrete stability condition and two-sided energy inequality):
Exactly the same way as in [51, 53] (see also [48] for more details), we can show that the discrete energetic solutions to the rate-independent system $(\mathcal{Q}, \mathcal{E}, \mathcal{D})$ satisfy the *discrete stability condition*

$$\mathcal{E}(t_k, u_k, \alpha_k, \eta_k) \leq \mathcal{E}(t_k, \tilde{u}, \tilde{\alpha}, \tilde{\eta}) + \mathcal{D}(t_k, \alpha_k, \tilde{\alpha}) \quad \forall\, (\tilde{u}, \tilde{\alpha}, \tilde{\eta}) \in \mathcal{Q} \tag{33}$$

and the energy inequalities

$$\begin{aligned} \int_{t_{k-1}}^{t_k} \partial\mathcal{E}_t(s, u_k(s))\,\mathrm{d}s &\leq \mathcal{E}(t_k, u_k, \alpha_k, \eta_k) + \mathcal{D}(t_k, \alpha_{k-1}, \alpha_k) - \mathcal{E}(t_{k-1}, u_{k-1}, \alpha_{k-1}, \eta_{k-1}) \\ &\leq \int_{t_{k-1}}^{t_k} \partial_t\mathcal{E}(s, u_{k-1}(s))\,\mathrm{d}s. \end{aligned} \tag{34}$$

Let us stress, however, that in order to derive (34) as in [51, 53], we need to exploit that the dissipation distance satisfies the triangle inequality (18) as noted in Remark 3.1.

Step 3 (Defining interpolants and a-priori estimates):
Based on solutions of (TIP) established in *Step 1*, we introduce their *piece-wise constant* interpolants through

$$\begin{aligned}[u_\tau(t),\alpha_\tau(t),\eta_\tau(t)] &:= [u_\tau^k,\alpha_\tau^k,\eta_\tau^k] \quad \text{if } t\in[t_k,t_{k+1}),\ k=0,\dots,N-1 \text{ and}\\ [u_\tau(\mathcal{T}),\alpha_\tau(\mathcal{T}),\eta_\tau(\mathcal{T})] &:= [u_\tau^N,\alpha_\tau^N,\eta_\tau^N] \end{aligned} \tag{35}$$

It follows from (34) that the interpolants satisfy the following bounds:

$$\|u_\tau\|_{L^\infty([0,\mathcal{T}];W^{1,2}(\Omega;\mathbb{R}^3))}\le C, \tag{36}$$

$$\|\eta_\tau\|_{L^\infty([0,\mathcal{T}],L^\infty(\Omega))}\le 1, \tag{37}$$

$$\|\alpha_\tau\|_{L^\infty([0,\mathcal{T}],W^{1,2}(\Omega;\mathbb{R}^{3\times3}\times\mathbb{R}))}\le C, \tag{38}$$

$$\mathrm{Var}_{L^1(\Omega;\mathbb{R}^{3\times3}\times\mathbb{R})}(\alpha_\tau;[0,\mathcal{T}])\le C, \tag{39}$$

where the variation is defined through

$$\mathrm{Var}_{L^1(\Omega;\mathbb{R}^{3\times3}\times\mathbb{R})}(\alpha_\tau;[0,\mathcal{T}]) := \sup\Big\{\sum_{i=1}^N \|\alpha(t_{i-1})-\alpha(t_i)\|_{L^1(\Omega;\mathbb{R}^{3\times3}\times\mathbb{R})}\ : \tag{40}$$

$$\text{a.p.p. } 0=t_0\le t_1\le\ldots\le t_N=t\Big\}. \tag{41}$$

Indeed, by summing the upper inequality up to some arbitrary $l\in 1,\dots,N(\tau)$ in (34), we can estimate

$$\mathcal{E}(t_l,u_\tau^l,\alpha_\tau^l,\eta_\tau^l)+\sum_{i=1}^l \mathcal{D}(t_i,\alpha_\tau(t_{i-1}),\alpha_\tau(t_i))\le \mathcal{E}(0,u_0,\alpha_0,\eta_0)+\int_0^{t_l}|\partial_t\mathcal{E}(s,u_\tau(s))|\,\mathrm{d}s. \tag{42}$$

Then, (36) and (38) are obtained from the coercivity of $\mathcal{E}$, which was already pointed out in Step 1, combined with discrete Gronwall's lemma and assumption (D3) – this assumption is exploited to estimate $\int_0^{t_l}|\partial_t\mathcal{E}(s,u_\tau(s))|\,\mathrm{d}s$ by a combination of Hölder's and Young's inequalities. Having (36) already at our disposal, (D3) readily gives

$$\int_0^{\mathcal{T}}|\partial_t\mathcal{E}(s,u_\tau(s))|\,\mathrm{d}s\le C. \tag{43}$$

The estimate (37) is induced by the fact that $\eta_\tau(t)\in\mathcal{Z}$ for all $t\in[0,\mathcal{T}]$. As far as (39), we obtain it from (42) when putting $l=N(\tau)$:

$$\begin{aligned}\mathrm{Var}_{L^1(\Omega;\mathbb{R}^{3\times3}\times\mathbb{R})}(\alpha_\tau;[0,\mathcal{T}]) &= \sum_{k=1}^{N(\tau)}\|\alpha_\tau^k-\alpha_\tau^{k-1}\|_{L^1(\Omega;\mathbb{R}^{3\times3}\times\mathbb{R})}\\ &\le \sum_{k=1}^{N(\tau)}\frac{1}{c^*}\big(\mathcal{D}(t_k,\alpha_\tau^{k-1},\alpha_\tau^k)+\mathcal{E}(\mathcal{T},u_\tau(\mathcal{T}),\alpha_\tau(\mathcal{T}),\eta_\tau(\mathcal{T}))\big)\\ &\le \frac{1}{c^*}\Big(\mathcal{E}(0,u_0,\alpha_0,\eta_0)+\int_0^{\mathcal{T}}|\partial_t\mathcal{E}(s,u_\tau(s))|\,\mathrm{d}s\Big)\le C.\end{aligned}$$

For more details on this step, the reader is referred to, e.g., [48].

Step 4 (Selection of subsequences):
Let us now find a cluster point of the sequence $(u_\tau,\alpha_\tau,\eta_\tau)$ for $\tau\to0$ which we shall, in the

following, show to be an energetic solution to the rate-independent system $(\mathcal{Q}, \mathcal{E}, \mathcal{D})$. By an application of a suitable version of Helly's selection principle [see Theorem 1.126 and Remark 1.127 in 65], [23], we deduce from (39) that there exists an $\alpha \in L^\infty([0,\mathcal{T}], W^{1,2}(\Omega)) \cap BV([0,\mathcal{T}], \mathrm{L}^1(\Omega))$ and a subsequence of α_τ's (not relabelled) so that

$$\forall\, t \in [0,\mathcal{T}] \;:\; \alpha_\tau(t) \rightharpoonup \alpha(t) \text{ in } W^{1,2}(\Omega;\mathbb{R}^{3\times 3}\times\mathbb{R}) \Rightarrow \alpha_\tau(t) \to \alpha(t) \text{ in } \mathrm{L}^2(\Omega;\mathbb{R}^{3\times 3}\times\mathbb{R}). \tag{44}$$

Further, we define $\theta(t) := \limsup_{\tau\to 0} \partial_t\mathcal{E}(t, u_\tau(t))$; note that $\theta(t)$ is measurable on $[0,\mathcal{T}]$.

Let us *fix* $t \in [0,\mathcal{T}]$ and let us define a *time-dependent subsequence* of $\tau's$ labelled $\tau_{k(t)}$ such that $\theta(t) = \lim_{\tau_{k(t)}\to 0} \partial_t\mathcal{E}(t, u_\tau(t))$. With t still fixed, we choose, if necessary, a further subsequence of $\tau_{k(t)}$ (not relabelled) such that

$$u_{\tau_{k(t)}} \rightharpoonup u(t) \quad \text{for } l\to\infty \text{ in } W^{1,2}(\Omega), \tag{45}$$

$$\eta_{\tau_{k(t)}} \overset{*}{\rightharpoonup} \eta(t) \quad \text{for } l\to\infty \text{ in } L^\infty(\Omega), \tag{46}$$

owing to (36)-(38). Thus, by (D2) and the particular form of the Gibbs free energy (20), $\theta(t) = \partial_t\mathcal{E}(t,u(t))$ and so, by an application of the Fatou's lemma,

$$\limsup_{\tau\to 0} \int_0^{\mathcal{T}} \partial_t\mathcal{E}(s,u_\tau(s))\,\mathrm{d}s \le \int_0^{\mathcal{T}} \theta(s)\,\mathrm{d}s = \int_0^{\mathcal{T}} \partial_t\mathcal{E}(s,u(s))\,\mathrm{d}s \tag{47}$$

Finally, we realize that due to the mentioned weak* closedness of convex closed sets, $(u(t),\alpha(t),\eta(t)) \in \mathcal{Q}$ for all $t\in[0,\mathcal{T}]$.

Step 5 (Stability of the limit function):
Let us show that $(u,\alpha,\eta) \in \mathcal{Q}$ found in the previous step satisfies (31). Due to the convexity of the Helmholtz free energy f we have that $\mathcal{E} : \mathbb{R}\times\mathcal{Q} \mapsto \mathbb{R}$ is weakly lower semicontinuous; hence, for some $t\in[0,\mathcal{T}]$ fixed, we have that

$$\begin{aligned}\mathcal{E}(t,u,\alpha,\eta) &\le \liminf_{\tau_{k(t)}\to 0} \mathcal{E}(t_{\tau_{k(t)}}, u_{\tau_{k(t)}}, \alpha_{\tau_{k(t)}}, \eta_{\tau_{k(t)}}) \\ &\le \limsup_{\tau_{k(t)}\to 0} \Big(\mathcal{E}(t_{\tau_{k(t)}}\tilde{u},\tilde{\alpha},\tilde{\eta}) + \mathcal{D}(t_{\tau_{k(t)}}, \alpha_{\tau_{k(t)}}(t),\tilde{\alpha})\Big)\end{aligned} \tag{48}$$

with $t_{\tau_{k(t)}} := \inf\{t_l \,:\, t\le t_l;\, l=1,\dots,N(\tau_{k(t)})\}$ and for any $(\tilde{u},\tilde{\alpha},\tilde{\eta}) \in \mathcal{Q}$; we also exploited the discrete stability (33).

On the right hand side, we use that $\mathcal{E}(t,\cdot,\cdot,\cdot)$ is continuous due to (D3) and (D4) and $\lim_{\tau_{k(t)}\to 0} \mathcal{D}(\tau_{k(t)}, \alpha_{\tau_{k(t)}}(t),\tilde{\alpha}) = \mathcal{D}(t,\alpha(t),\tilde{\alpha})$. To see the latter, we employ the Nemytskiĭ continuity of $\mathcal{D}(\cdot,\alpha,\cdot) : L^2(\Omega;\mathbb{R}^{3\times 3}\times\mathbb{R}) \mapsto \mathbb{R}$ (note that, thanks to (44), $\alpha_{\tau_{k(t)}} \to \alpha$ *strongly* in $L^2(\Omega;\mathbb{R}^{3\times 3}\times\mathbb{R})$) and the continuity of $\mathcal{D}(t,\cdot,\cdot) : \mathbb{R}\mapsto\mathbb{R}$. While the second property is a direct consequence of (D4), the Nemytskiĭ continuity is guaranteed by the linear growth of $\mathcal{D}(\cdot,\alpha,\cdot)$ (with respect to α) and the fact that $\delta(\cdot,\alpha,\cdot)$ is continuous.

Thus,

$$\begin{aligned}\mathcal{E}(t,u,\alpha,\eta) &\le \liminf_{\tau_{k(t)}\to 0} \mathcal{E}(t_{\tau_{k(t)}}, u_{\tau_{k(t)}}, \alpha_{\tau_{k(t)}}, \eta_{\tau_{k(t)}}) \\ &\le \limsup_{\tau_{k(t)}\to 0} \Big(\mathcal{E}(t_{\tau_{k(t)}},\tilde{u},\tilde{\alpha},\tilde{\eta}) + \mathcal{D}(\tau_{k(t)}, \alpha_{\tau_{k(t)}}(t),\tilde{\alpha})\Big) \\ &\le \mathcal{E}(t,\tilde{u},\tilde{\alpha},\tilde{\eta}) + \mathcal{D}(t,\alpha(t),\tilde{\alpha})\end{aligned}$$

for any $(\tilde{u},\tilde{\alpha},\tilde{\eta})\in\mathcal{Q}$, i.e. we verified (31). In particular, when taking $(\tilde{u},\tilde{\alpha},\tilde{\eta})=(u,\alpha,\eta)$ in the above inequality, we also get that

$$\mathcal{E}(t,u,\alpha,\eta)=\lim_{\tau_{k(t)}\to 0}\mathcal{E}(t_{\tau_{k(t)}},u_{\tau_{k(t)}},\alpha_{\tau_{k(t)}},\eta_{\tau_{k(t)}}). \tag{49}$$

Step 6 (Upper energy inequality):
In order to prove the upper energy inequality

$$\mathcal{E}(\mathcal{T},u(\mathcal{T}),\alpha(\mathcal{T}),\eta(\mathcal{T}))+\mathrm{Diss}_{\mathcal{D}}(\alpha,[0,\mathcal{T}])\le\mathcal{E}(0,u(0),\alpha(0),\eta(0))+\int_0^{\mathcal{T}}\partial_t\mathcal{E}(s,u(s))\,\mathrm{d}s,$$

we first need to show the weak lower semicontinuity of $\mathrm{Diss}_{\mathcal{D}}(\alpha,[0,\mathcal{T}])$ along the sequence α_τ selected in 4. We will establish it separately for the transformation and reorientation part:

$$\mathrm{Diss}^{\mathrm{tr}}(\alpha,[0,\mathcal{T}])\le\liminf_{\tau\to 0}\sum_{k=1}^{N(\tau)}\mathcal{D}^{\mathrm{tr}}(\alpha_\tau(t_{k-1}),\alpha_\tau(t_k)) \tag{50}$$

and

$$\mathrm{Diss}^{\mathrm{reo}}(\alpha,[0,\mathcal{T}])\le\liminf_{\tau\to 0}\sum_{k=1}^{N(\tau)}\mathcal{D}^{\mathrm{reo}}(t_k,\alpha_\tau(t_{k-1}),\alpha_\tau(t_k)). \tag{51}$$

In fact, only the reorientation part needs our attention since, for any partition of the interval $[0,\mathcal{T}]$ $0\le t^1\le t^2\le\ldots\le t^K=\mathcal{T}$, we have that $\mathcal{D}^{\mathrm{tr}}(\alpha_\tau(t^{k-1}),\alpha_\tau(t^k))\to\mathcal{D}^{\mathrm{tr}}(\alpha(t^{k-1}),\alpha(t^k))$ whenever $\alpha_\tau\rightharpoonup\alpha$ in $\mathcal{V}$; this already implies (50).

Concentrating thus on the reorientation part, we take again some partition of the interval $[0,\mathcal{T}]$ $0\le t^1\le t^2\le\ldots\le t^K=\mathcal{T}$. Since $\alpha_\tau(t)\to\alpha(t)$ strongly in $L^2(\Omega;\mathbb{R}^{3\times 3}\times\mathbb{R})$ for all $t\in[0,\mathcal{T}]$ we have that $\alpha_\tau(t,x)\to\alpha(t,x)$ for a.a. $x\in\Omega$ and all $t\in[0,\mathcal{T}]$ (at least in terms of a not-relabelled subsequence). Because all α_τ's are in $\mathcal{V}$, we know that (for the components of this particular α_τ) $\langle\boldsymbol{\varepsilon}^{\mathrm{in}}_\tau(x)\rangle\le\xi_\tau(x)$. This property assures continuity of $\overline{\delta}^{\mathrm{reo}}$ in the sense $\overline{\delta}^{\mathrm{reo}}(\alpha_\tau(t^{k-1},x),\alpha_\tau(t^k,x))\to\overline{\delta}^{\mathrm{reo}}(\alpha(t^{k-1},x),\alpha(t^k,x))$ for a.a. $x\in\Omega$. Thus, $\overline{\delta}^{\mathrm{reo}}(\alpha(t^{k-1},x),\alpha(t^k,x))$ is measurable and exploiting the growth of $\overline{\delta}^{\mathrm{reo}}$ as well as the Lebesgue dominated convergence theorem we get that

$$\lim_{\tau\to 0}\overline{\mathcal{D}}^{\mathrm{reo}}(\alpha_\tau(\tau^{k-1}),\alpha_\tau(\tau^k))=\overline{\mathcal{D}}^{\mathrm{reo}}(\alpha(\tau^{k-1}),\alpha(\tau^k))$$

which standardly yields that

$$\mu_\alpha([s,z])\le\liminf_{k\to\infty}\mu_{\alpha_\tau}([s,z]). \tag{52}$$

for any subinterval $[s,z]\subset[0,\mathcal{T}]$; notice that μ_{α_τ} is defined through (28) with α_τ in place of α. Moreover, (52) can be extended to also to all Borel sets A in $[0,\mathcal{T}]$ (in particular also to half-open intervals) through (29); i.e.

$$\mu_\alpha(A)\le\liminf_{k\to\infty}\mu_{\alpha_\tau}(A). \tag{53}$$

Since $\sigma^{\mathrm{reo}}(T(t))$ is positive and continuous, we may approximate it by piece-wise constant, positive functions on half-open intervals to get from (53) that

$$\int_0^{\mathcal{T}}\sigma^{\mathrm{reo}}(T(t))\,\mathrm{d}\mu_\alpha(t)\le\liminf_{k\to\infty}\int_0^{\mathcal{T}}\sigma^{\mathrm{reo}}(T(t))\,\mathrm{d}\mu_{\alpha_\tau}(t). \tag{54}$$

Indeed, in more detail, we take an equidistant partition $0 = t_0 \le t_1 \le \ldots \le t_n = \mathcal{T}$ of the interval $[0, \mathcal{T}]$ with $t_i = t_0 + ih$ for $i \in \{0, \ldots, n\}$ and $h := \mathcal{T}/n$ (note that this partition does *not* need to coincide with any of the previously chosen ones) and define a sequence of functions $\sigma_n^{\rm reo}(t) : [0, \mathcal{T}] \to \mathbb{R}$ as

$$\begin{aligned} \sigma_n^{\rm reo}(t) &:= \sigma^{\rm reo}(T(t_{i-1})) \quad \text{if} \quad t \in [t_{i-1}, t_i),\ i = 1, \ldots, N \quad \text{and} \\ \sigma_n^{\rm reo}(t_n) &:= \sigma^{\rm reo}(T(\mathcal{T})) \end{aligned} \tag{55}$$

Due to (D4), we have that $\sigma_n^{\rm reo}(t) \to \sigma^{\rm reo}(T(t))$ pointwise. Clearly, from (53),

$$\int_0^{\mathcal{T}} \sigma_n^{\rm reo}(t)\,{\rm d}\mu_\alpha(t) \le \liminf_{k\to\infty} \int_0^{\mathcal{T}} \sigma_n^{\rm reo}(t)\,{\rm d}\mu_{\alpha_\tau}(t);$$

(54) is then obtained by Lebegue's dominated convergence theorem owing to (D4) providing a uniform bound on $\sigma_n^{\rm reo}(t)$ and (39).

Therefore, we can exploit (51) and the weak*-lower semicontinuity of the Gibbs free energy as well as (34) (with $l = N(\tau)$) to get

$$\begin{aligned} &\mathcal{E}(\mathcal{T},u(\mathcal{T}),\alpha(\mathcal{T}),\eta(\mathcal{T})) + {\rm Diss}(\alpha,[0,t]) \\ &\quad\le \liminf_{\tau\to 0} \Big(\mathcal{E}(\mathcal{T},u_\tau(\mathcal{T}),\alpha_\tau(\mathcal{T}),\eta_\tau(\mathcal{T})) + \sum_{k=1}^{N(\tau)} \mathcal{D}^{\rm reo}(t_k,\alpha_\tau(t_{k-1}),\alpha_\tau(t_k)) + \mathcal{D}^{\rm tr}(\alpha_\tau(t_{k-1}),\alpha_\tau(t_k))\Big) \\ &\quad\le \limsup_{\tau\to 0} \Big(\mathcal{E}(0,u_0,\alpha_0,\eta_0) + \int_0^{\mathcal{T}} \partial_t\mathcal{E}(s,u_\tau(s))\,{\rm d}s\Big) \le \mathcal{E}(0,u_0,\alpha_0,\eta_0) + \int_0^{\mathcal{T}} \partial_t\mathcal{E}(s,u(s))\,{\rm d}s, \end{aligned}$$

where the last inequality is due to (47).

Step 7 (Lower energy inequality):
Finally, we prove that

$$\mathcal{E}(\mathcal{T},u(\mathcal{T}),\alpha(\mathcal{T}),\eta(\mathcal{T})) + {\rm Diss}_{\mathcal{D}}(\alpha,[0,\mathcal{T}]) \ge \mathcal{E}(0,u(0),\alpha(0),\eta(0)) + \int_0^{\mathcal{T}} \partial_t\mathcal{E}(s,u(s))\,{\rm d}s;$$

actually, as was observed in e.g. [49], this is indeed a consequence of the stability condition (31). To see this, let $0 = t_0 \le t_1 \le \ldots \le t_K = \mathcal{T}$, $t \in [0, \mathcal{T}]$ be a sequence of some arbitrary partitions of $[0, \mathcal{T}]$ with $\Delta_K := \max\{t_j - t_{j-1} \ :\ j = 1, \ldots, K\}$.

Let us define the piece-wise constant interpolant $\sigma_K^{\rm reo}(T(t))$ of $\sigma^{\rm reo}$ as follows:

$$\begin{aligned} \sigma_K^{\rm reo}(T(t)) &:= \sigma^{\rm reo}(T(t_{j-1})) \quad \text{if} \quad t \in [t_{j-1}, t_j),\ j = 1, \ldots, K \quad \text{and} \\ \sigma_K^{\rm reo}(T(\mathcal{T})) &:= \sigma^{\rm reo}(T(\mathcal{T})). \end{aligned} \tag{56}$$

Due to the continuity of $\sigma^{\rm reo}(T(t))$ (and thus, in particular, uniform boundeness) on $[0, \mathcal{T}]$ (cf. (D4)), we may employ Lebesgue's dominated convergence theorem to get

$$\lim_{K\to\infty} \int_0^{\mathcal{T}} \sigma_K^{\rm reo}(T(t))\,{\rm d}\mu_\alpha(t) = \int_0^{\mathcal{T}} \sigma^{\rm reo}(T(t))\,{\rm d}\mu_\alpha(t). \tag{57}$$

Moreover, exploiting the definition of the interpolant $\sigma_K^{\rm reo}$ and positivity of $\sigma^{\rm reo}$ in (16), we obtain the following inequality:

$$\begin{aligned} \int_0^{\mathcal{T}} \sigma_K^{\rm reo}(T(t))\,{\rm d}\mu_\alpha(t) &= \sum_{j=1}^{K} \int_{t_{j-1}}^{t_j} \sigma_K^{\rm reo}(T(t_{j-1}))\,{\rm d}\mu_\alpha(t) \\ &\ge \sum_{j=1}^{K} \sigma_K^{\rm reo}(T(t_{j-1})) \int_{t_{j-1}}^{t_j} {\rm d}\mu_\alpha^*(t) = \sum_{j=1}^{K} \mathcal{D}^{\rm reo}(t_{j-1},\alpha(t_{j-1}),\alpha(t_j)), \end{aligned} \tag{58}$$

where the Radon measure μ^*_α is defined for every closed set $S := [s, z] \subset [0, \mathcal{T}]$ through $\mu^*_\alpha(S) := \overline{\mathcal{D}}^{\rm reo}(\alpha(s), \alpha(z))$. Notice that in the inequality on the second line we actually exploited the definition (28).

Now, testing (31) in the point $(u(t_{j-1}), \alpha(t_{j-1}), \eta(t_{j-1}))$ by $(u(t_j), \alpha(t_j), \eta(t_j))$ gives (after summation over $j = 1, \dots, K$)

$$\begin{aligned}
&\mathcal{E}(t, u(t), \alpha(t), \eta(t)) + {\rm Diss}^{\rm tr}(\alpha, [0,t]) + \int_0^{\mathcal{T}} \sigma^{\rm reo}_K(T(t))\,{\rm d}\mu_\alpha(t) - \mathcal{E}(0, u(0), \alpha(0), \eta(0)) \\
&\quad \geq \mathcal{E}(t, u(t), \alpha(t), \eta(t)) + \sum_{j=1}^{K} \mathcal{D}^{\rm tr}(\alpha(t_{j-1}), \alpha(t_j)) \\
&\qquad\qquad + \sum_{j=1}^{K} \mathcal{D}^{\rm reo}(t_{j-1}, \alpha(t_{j-1}), \alpha(t_j)) - \mathcal{E}(0, u(0), \alpha(0), \eta(0)) \\
&\quad \geq \sum_{j=1}^{K} \int_{t_{j-1}}^{t_j} \partial_t \mathcal{E}(s, u(t_j), \alpha(t_j), \eta(t_j))\,{\rm d}s \\
&\quad = \underbrace{\sum_{j=1}^{K} \partial_t \mathcal{E}(t_j, u(t_j), \alpha(t_j), \eta(t_j))|t_j - t_{j-1}|}_{\rm (I)} - \underbrace{\sum_{j=1}^{K} \varrho_j |t_j - t_{j-1}|}_{\rm (II)},
\end{aligned} \tag{59}$$

where we defined

$$\varrho_j := \frac{1}{|t_j - t_{j-1}|} \int_{t_{j-1}}^{t_j} \left[\partial_t \mathcal{E}(s, u(t_j), \alpha(t_j), \eta(t_j)) - \partial_t \mathcal{E}(t_j, u(t_j), \alpha(t_j), \eta(t_j))\right] {\rm d}s. \tag{60}$$

As $K \to \infty$ in (59) term (I) converges to $\int_0^{\mathcal{T}} \partial_t \mathcal{E}(s, u(s))\,{\rm d}s$ by [23, 66] which allows to approximate the Lebesgue integral by Riemann sums, while term (II) converges to 0 by the uniform continuity of $\partial_t \mathcal{E}(t, \cdot)$ on the compact set $[0, \mathcal{T}]$ (cf. [23] for details). □

4 Numerical implementation and simulations

Now we proceed with a brief description of the way the numerical implementation of the proposed model is treated. We specify forms of the material functions and values of material parameters and present a validating simulation.

4.1 Concept of numerical implementation

The analysis of existence of energetic solutions provides grounds for numerical implementation of the model; indeed (TIP) provides the starting point for our implementation. Let us just note that, in order to simplify the implementation, we omit the regularization term in the energy – and denote the modified energy as $\tilde{\mathcal{E}}$ – since ν is presumed to be very small.[9]

The global minimization problem in a time-point t_k $(k = 1, \dots, N)$ of a time-discretization $0 = t_0 \leq t_1 \leq \dots \leq t_N = \mathcal{T}$ then takes the form

$$\begin{aligned}
&\text{Minimize } \tilde{\mathcal{E}}(t_k, u, \varepsilon^{\rm in}, \xi, \eta) + \mathcal{D}(t_k, \varepsilon^{\rm in}_{k-1}, \xi_{k-1}, \varepsilon^{\rm in}, \xi) \\
&\text{subject to } (u, \varepsilon^{\rm in}, \xi, \eta) \in \mathcal{Q};
\end{aligned} \tag{61}$$

[9] Note that this term is not relevant when proving existence of discrete energetic solutions.

where $(u_0, \varepsilon_0^{\text{in}}, \xi_0, \eta_0) := (u(0), \varepsilon^{\text{in}}(0), \xi(0), \eta(0)) \in \mathcal{Q}$ defined through initial conditions. A possible effective method of solution of (61) is to split it into two complementary subproblems, which are to be solved consecutively.[10] To this end, let us note that $\mathcal{D}$ is *independent* on u, hence, minimization in (61) with respect to u – other variables fixed and considered as parameters for the moment – can be viewed as minimization of a potential energy with respect to displacement, which is a rather standard problem solvable by finite element method.

Hence, the first subproblem corresponds to *minimization with respect to displacement* on the level of domain Ω, i.e. finding global mechanical balance, and it can be iteratively solved by employing a finite element software. At p-th iteration ($p \in \mathbb{N}$), the problem takes the form:

$$\begin{aligned} &\text{Minimize } \tilde{\mathcal{E}}(T_k, u, \varepsilon_{k,p-1}^{\text{in}}, \xi_{k,p-1}, \eta_{k,p-1}) \\ &\text{subject to } u \in \mathcal{U}^{\text{disc}}; \end{aligned} \tag{62}$$

where $(\varepsilon_{k,p-1}^{\text{in}}, \xi_{k,p-1}, \eta_{k,p-1})$ now represent a set of parameters resulting from the second subproblem and $\mathcal{U}^{\text{disc}}$ is a suitable approximation of $\mathcal{U}$. We put $(\varepsilon_{k,0}^{\text{in}}, \xi_{k,0}, \eta_{k,0}) := (\varepsilon_{k-1}^{\text{in}}, \xi_{k-1}, \eta_{k-1})$ and the set is upgraded after every iteration in (62). Let us note the number of parameters in this set depends on spatial discretization of Ω.

The second subproblem corresponds to *minimization only with respect to internal variables*, i.e. local thermodynamical balance. It can be solved in *each material point independently* and represents the constitutive response:

$$\begin{aligned} &\text{Minimize } f(T_k, \boldsymbol{\varepsilon}_{k,p}, \boldsymbol{\varepsilon}^{\text{in}}, \xi, \eta) + \delta(T_k, \xi_{k-1}, \boldsymbol{\varepsilon}_{k-1}^{\text{in}}, \xi, \boldsymbol{\varepsilon}^{\text{in}}) \\ &\text{subject to } (\boldsymbol{\varepsilon}^{\text{in}}, \xi, \eta) \text{ satisfying constraints in (5) and (4).} \end{aligned} \tag{63}$$

Here, in each material point, strain $\varepsilon_{k,p}$ is derived from displacement $u_{k,p}$ generated in actual iteration of the minimization problem (62). Since the cost function in *convex* the consolidated result of the iteration process corresponds (with respect to numerical error) to the solution of (61) or, in other words the solutions of the iteration process converge to solutions of (61) [68, 69].

Constraints given by inequalities in (5) and (4) can be imposed by inclusion of appropriate form of indicator function to the minimized (cost) function in (63). Due to physical reasons [25], it may be advantageous to use a regularized form of the indicator function, denoted $r(\boldsymbol{\varepsilon}^{\text{tr}}, \xi, \eta)$ in that work, which assures smooth behavior when internal variables approach limiting values.

In the present work, the minimization problem (63) was solved by the method introduced by [70]. It is an iterative, derivative-free optimization algorithm suitable for non-smooth functions. The corresponding subroutine was then implemented into Abaqus finite element analysis software. Details including the particular form of the regularized indicator function can be found in [25].

4.2 Specification of material functions

To capture the transformation strain anisotropy of the material and the well-known tension-compression asymmetry (e.g. [71]), we specify the function $\langle \cdot \rangle$ as follows:

$$\langle \boldsymbol{\varepsilon}^{\text{in}} \rangle = \frac{I_2(\mathbb{D}\boldsymbol{\varepsilon}^{\text{in}})}{k} \frac{\cos\left(\frac{1}{3}\arccos(1 - a(I_3(\mathbb{D}\boldsymbol{\varepsilon}^{\text{in}}) + 1))\right)}{\cos\left(\frac{1}{3}\arccos(1 - 2a)\right)}, \tag{64}$$

[10] Such an approach is sometimes termed *combined minimization principle*, see e.g. [67].

k [1]	E^A [GPa]	E^M [GPa]	E^R [GPa]	G^A [GPA]	G^M [GPa]	G^R [GPa]
0.07	71	41	43	25	22	15
a [1]	A [1]	B[1]	φ [rad]	L [1]	M [1]	N [1]
0.8	1	1	0	1	1.44	1.44
A_s [°C]	A_f [°C]	M_s [°C]	M_f [°C]	R_s [°C]	R_f [°C]	T_0 [°C]
-17	-13	-22	-24	38	1	-19
Δs^{AM} [J/°C]	Δs^{AR} [J/°C]	σ_0^{reo} [MPa]	Σ^{reo} [MPa/°C]	c^{reg} [MPa]	E^{int} [MPa]	
0.364	0.121	130	-0.9	10	30	

Table 1: Table of material parameters used in simulations. Parameters A, B, φ, L, M, N specify the linear mapping $\mathbb{D}$, parameters $c^{\text{reg}}, E^{\text{int}}$ determine the regularized indicator function r, see [25] for details.

where $I_2(\boldsymbol{x}) = \sqrt{\frac{2}{3}}\|\boldsymbol{x}\|$ and $I_3(\boldsymbol{x}) = 4\frac{\det(\boldsymbol{x})}{I_2(\boldsymbol{x})^3}$. The material parameter k represents the maximum transformation strain in tension, whereas parameter a characterizes the tension-compression asymmetry and ranges between 0 (no tension-compression asymmetry) and 1 (transformation strain in compression is a half of the strain in tension). The *linear mapping* $\mathbb{D} : \mathbb{R}^{3\times3} \to \mathbb{R}^{3\times3}$ is chosen so that the model can be adjusted to material symmetry of macroscopic inelastic strain observed in experiments. The particular form of $\mathbb{D}$ used in our simulations can be found in [25] and it corresponds to transversal isotropy usually observed in NiTi wires.

As mentioned in Subsection 2.2.3, it is often experimentally observed that the stress needed for initialization of martensite reorientation decreases with increasing temperature. This motivates the particular form of the material function σ^{reo} to be a decreasing linear function of temperature in the temperature range of interest, i.e.

$$\sigma^{\text{reo}}(T) = \sigma_0^{\text{reo}} + \Sigma^{\text{reo}} \cdot (T - T_0), \tag{65}$$

where σ_0^{reo} and Σ^{reo} are material parameters.

The above forms are chosen to provide enough freedom for adaptation of the model to material properties which can differ due to variation in composition or processing, albeit trying to keep the number of input parameters low to facilitate the experimental effort prior to simulations.

4.3 Numerical simulation of a NiTi helical spring

In this subsection we present a simulation of loading of a NiTi superelastic helical spring at constant temperature. Such a spring was used in an investigation of fatigue lifetime of NiTi wire in combined loading in [24].

Figure 2 presents two-coil model of the spring in finite element model software Abaqus (7680 eight-node hexaedral elements C3D8R used). The outer diameter of a coil is 3.14 mm, its length is 1.05 mm, the coil wire has 0.19 mm in diameter. The material parameters of the wire are summarized in Table 1.[11] Let us note that elastic properties in tension and torsion were considered as independent (not related through the bulk modulus) because of the strong texture in the real wire [25]. The symmetry axis of transverse isotropy of the material is supposed to follow the axis of the wire forming the spring. The corresponding uniaxial and shear response of a material point at temperature at which the spring was loaded, $T_{\text{sim}} = 20°\text{C}$, are presented in Fig. 3.

[11] Note that restrictions (16) are satisfied for a reasonable range of working temperature. Moreover, due to [72], requirements imposed on function $\langle\cdot\rangle$ in (D2) can be also verified.

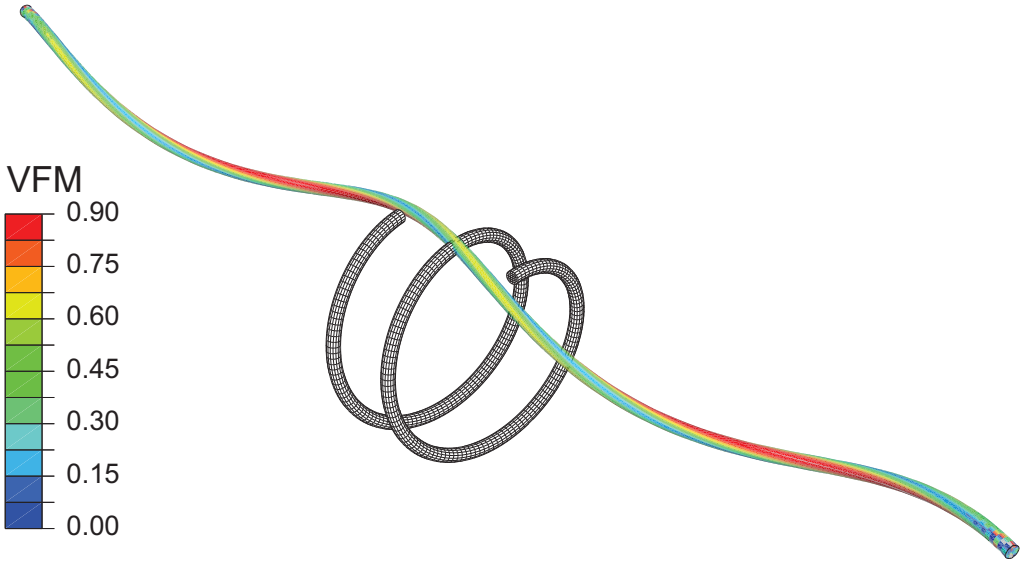

Figure 2: A helical spring model: undeformed shape in grey (used meshing shown) and deformed shape at 19.3 mm stroke in colors; color map shows distribution of volume fraction of martensite.

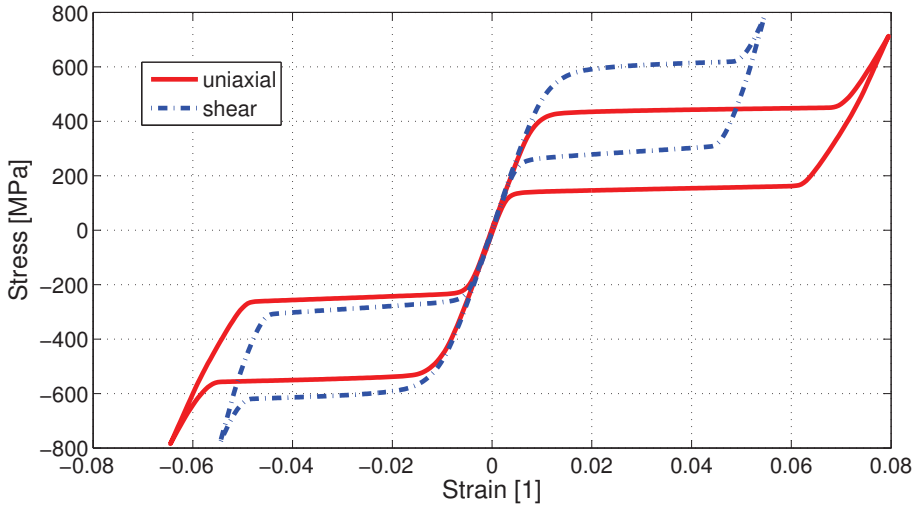

Figure 3: Uniaxial and pure shear response of a material element with parameters according to Table 1. Red (solid) line corresponds to uniaxial loading, blue (dot-dashed) line to pure shear loading. In shear, values of equivalent shear strain ($\gamma/\sqrt{3}$) and shear stress ($\sqrt{3}\tau$) are reported. Note pronounced tension-compression asymmetry in uniaxial loading (since $a \neq 0$) and reduced transformation strain in shear (since $N \neq 1$).

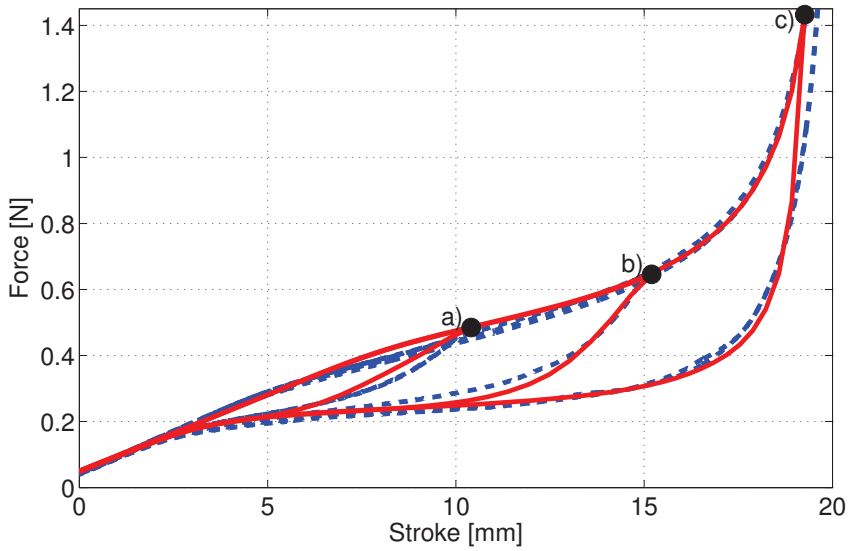

Figure 4: Experimental (dashed blue line) and simulated (solid red line) force-stroke response of a helical spring for three loading-unloading cycles with maximum stroke of 10.4 mm, 15.2 mm and 19.3 mm.

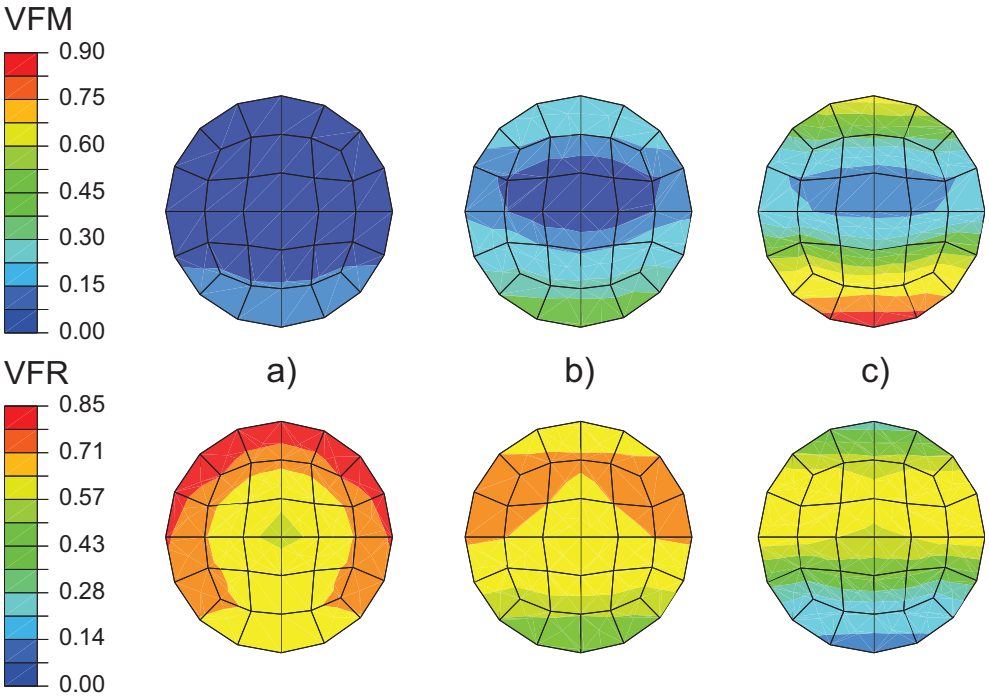

Figure 5: Distribution of volume fraction of martensite (VFM) and R-phase (VFR) within a cross section of the wire for stroke of a) 10.4 mm, b) 15.2 mm, c) 19.3 mm in forward loading. The lowest point of the cross-section is the closest one to the spring axis.

Figure 2 also shows the deformed shape of the spring at stroke 19.3 mm. The contour plot provides basic information about distribution of volume fraction of martensite on the surface of the spring.

In Fig. 4, simulated force-stroke response for three loading-unloading cycles is compared to experiments. Even though clamping of the real spring may influence the response (especially at the highest stroke when the spring faces uncoiling), a rather good correspondence between experiments and simulations is reached in general.

Distribution of volume fraction of martensite and R-phase within the cross section of the wire far from the ends at three strokes during loading (denoted by black circles in Fig. 4) are presented in Fig. 5. In stress free state, approx. 49% of R-phase and 51% of austenite is present in the material ($R_f < T_{sim} < R_s$). When loaded the wire is deformed both in torsion and bending. First, shear induces increase of R-phase concentration near the surface of the spring. Moreover, the inner part of the surface which is closest to the spring axis is deformed in tension, which promotes further transformation to R-phase and leads to early occurrence of martensite in this region, see a). Due to tension/compression asymmetry (cf. Fig. 3), the same process induced by compression on the opposite side of the surface is retarded and transformation from R-phase to martensite starts at higher loading, see b). The tension/compression asymmetry also causes that the neutral axis in loading is shifted from the center of the cross section of the wire towards the outer surface of the spring. Hence, the minimum of volume fraction of martensite is closer to the outer surface, as readily seen in c).

Acknowledgements

We would like to thank dr. P. Šittner for providing experimental measurements used in Subsection 4.3 and prof. T. Ben Zineb for providing access to LEM3 computation cluster (Metz, France).

Funding

This work was conducted within the institutional project CEZ:AV0Z20760514 of IT ASCR, v.v.i., and within the CENTEM project [CZ.1.05/2.1.00/03.0088]. The research has also been supported by the Grant Agency of the Czech Republic through Grants [13-13616S, P201/10/0357] and the Grant Agency of the Charles University in Prague [41110].

References

[1] Bhattacharya K. Microstructure of martensite. Why it forms and how give rise to the shape-memory effect. Oxford: Oxford University Press; 2003.

[2] Machado LG, Savi MA. Medical applications of shape memory alloys. *Braz J Med Biol Res.* 2003; 36:683–691.

[3] Williams E, Elahinia MH. An Automotive SMA Mirror Actuator: Modeling, Design and Experimental Evaluation. *J Intell Mater Systems Struct.* 2008; 19:1425–1434.

[4] Vokoun D, Sedlák P, Frost M, Pilch J, Majtás D, Šittner P. Velcro-like fasteners based on NiTi micro-hook arrays. *Smart Mater Struct.* 2011; 20:085027.

[5] Lagoudas DC, Entchev PB, Popov P, Patoor E, Brinson LC, Gao X. Shape memory alloys, Part II: Modeling of polycrystals. *Mech Mater*. 2006; 38:430–62.

[6] Patoor E, Lagoudas DC, Entchev PB, Brinson LC, Gao X. Shape memory alloys, Part I: General properties and modeling of single crystals. *Mech Mater*. 2006; 38:391–429.

[7] Khandelwal A, Buravalla V. Models for Shape Memory Alloy Behavior: An overview of modeling approaches. *Int J Struct Changes*. 2009; 1:111–148.

[8] Roubíček T. Models of microstructure evolution in shape memory materials. In: Ponte Castaneda Pea, editor. Nonlinear Homogenization and its Application to Composites, Polycrystals and Smart Materials. NATO Sci. Ser. II/170. Kluwer, Dordrecht; 2004. p. 269–304.

[9] Auricchio F, Petrini L. A three-dimensional model describing stress-temperature induced solid phase transformations: solution algorithm and boundary value problems. *Int J Numer Meth Engng*. 2004; 61:807–836.

[10] Arghavani J, Auricchio F, Naghdabadi R, Reali A, Sohrabpour S. A 3-D phenomenological constitutive model for shape memory alloys under multiaxial loadings. *Int J Plast*. 2010; 26:976–991.

[11] Bouvet C, Calloch S, Lexcellent C. A Phenomenological Model for Pseudoelasticity of Shape Memory Alloys Under Multiaxial Proportional and Nonproportional Loading. *Eur J Mech A*. 2004; 23:37–61.

[12] Chemisky Y, Duval A, Patoor E, Ben Zineb T. Constitutive model for shape memory alloys including phase transformation, martensitic reorientation and twins accommodation. *Mech Mater*. 2011; 43:361–376.

[13] Peultier B, Ben Zineb T, Patoor E. Macroscopic constitutive law for SMA: Application to structure analysis by FEM. *Mech Mater*. 2006; 38:510–524.

[14] Zaki W, Moumni Z. A three-dimensional model of the thermomechanical behavior of shape memory alloys. *J Mech Phys Solids*. 2007; 55:2455–2490.

[15] Panico M, Brinson LC. A three-dimensional phenomenological model for martensite reorientation in shape memory alloys. *J Mech Phys Solids*. 2007; 55:2491–2511.

[16] Lagoudas DC, Hartl DJ, Chemisky Y, Machado LG, Popov P. Constitutive model for the numerical analysis of phase transformation in polycrystalline shape memory alloys. *Int J Plast*. 2012; 32–33:155–183.

[17] Pilch J, Heller L, Šittner P. Roundrobin web page; 2009. Available from: `http://ofm.fzu.cz/ofm/index.php/en/roundrobin`.

[18] Šittner P, Heller L, Pilch J, Sedlák P, Frost M, Chemisky Y, et al. Roundrobin SMA modeling. In: Šittner P, Heller L, Paidar V, editors. ESOMAT 2009 - The 8th European Symposium on Martensitic Transformations. EDP Sciences; 2009. p. 08001.

[19] Sadjadpour A, Bhattacharya K. A micromechanics-inspired constitutive model for shape-memory alloys. *Smart Mater Struct*. 2007; 16:1751–1765.

[20] Grabe C, Bruhns OT. On the viscous and strain rate dependent behavior of polycrystalline NiTi. *Int J Solids Struct*. 2008; 45:1876–1895.

[21] Halphen B, Nguyen QS. Sur les matériaux standard généralisés. *J Mecanique*. 1975; 14:39–63.

[22] Mielke A. A mathematical framework for standard generalized materials in the rate-independent case. In: R Helmig BW A Mielke, editor. Multifield problems in Fluid and Solid Mechanics. vol. 28. lecture notes in applied and computational mechanics ed. Springer, Berlin; 2006. .

[23] Francfort G, Mielke A. Existence results for a class of rate-independent material models with nonconvex elastic energies. *J Reine Angew Math*. 2006; 595:55–91.

[24] Frost M, Sedlák P, Benešová B, Ben Zineb T, Šittner P. Macroscopic Thermomechanical Model Suitable for Simulations of Anisotropic NiTi Shape Memory Alloys with R-Phase. In: Proceedings of The International Conference on Shape Memory and Superelastic Technologies, Prague; 2013. p. 300–301.

[25] Sedlák P, Frost M, Benešová B, Šittner P, Ben Zineb T. Thermomechanical model for NiTi-based shape memory alloys including R-phase and material anisotropy under multi-axial loadings. *Int J Plast*. 2012; 39:132–151.

[26] Mielke A, Theil F, Levitas VI. A variational formulation of rate-independent phase transformations using an extremum principle. *Arch Ration Mech An*. 2002; 162(2):137–177.

[27] Petryk H, Stupkiewicz S. Interfacial energy and dissipation in martensitic phase transformations. Part I: Theory. *Journal of the Mechanics and Physics of Solids*. 2010; 58:390–408.

[28] Petryk H. Incremental energy minimization in dissipatiove solids. *R C Mecanique*. 2003; 331:469–474.

[29] Maugin GA. The Thrermomechanics of Plasticity and Fracture. Cambridge University Press; 1992.

[30] Luig P, Bruhns OT. On the modeling of shape memory alloys using tensorial internal variables. *Materials Science and Engineering A*. 2008; 481–482:379–383.

[31] Bartel T, Hackl K. A micromechanical model for martensitic phase transformations in shape memory alloys based on energy relaxation. *ZAMM, Z Angew Math Me*. 2009; 89(10):792–809.

[32] Heinen R, Miro S. Micromechanical modeling of NiTi shape memory alloys including austenite, R-phase, and martensite. *Comput Methods Appl Mech Engrg*. 2012; 229-232:44–55.

[33] Jacobus K, Sehitoglu H, Balzer M. Effect of stress state on the stress-induced martensitic transformation in polycrystalline Ni-Ti alloy. *Metall*. 1996; 27:3066–3073.

[34] Otsuka K, Wayman CM. Shape Memory Materials. Cambridge University Press; 1998.

[35] Otsuka K, Ren X. Physical metallurgy of Ti-Ni-based shape memory alloys. *Prog Mater Sci*. 2005; 50:511–678.

[36] Olbricht J, Yawny A, Pelegrina JL, Dlouhý A, Eggeler G. On the Stress-Induced Formation of R-Phase in Ultra-Fine-Grained Ni-Rich NiTi Shape Memory Alloys. *Metall Mater Trans A*. 2011; 42A:2556–2574.

[37] Seiner H, Landa M. Non-classical austenite-martensite interfaces observed in single crystals of Cu-Al-Ni. *Phase Transitions*. 2009; 82:793–807.

[38] Liu Y, Xie Z, Van Humbeeck J, Delaey L. Deformation of shape memory alloys associated with twinned domain re-configurations. *Mat Sci Eng A*. 1999; 273–275:679–684.

[39] Seiner H, Sedlák P, Landa M. Shape recovery mechanism observed in single crystals of Cu-Al-Ni shape memory alloy. *Phase Transitions*. 2008; 81:537–551.

[40] Bernardini D, Pence TJ. Models for one-variant shape memory materials based on dissipation functions. *Int J NonLin Mech*. 2002; 37:1299–1317.

[41] Zhang JX, Sato M, Ishida A. Deformation mechanism of martensite in Ti-rich TiNi shape memory alloy thin films. *Acta Mater*. 2006; 54:1185–1198.

[42] Kockar B, Karaman I, Kim JI, Chumlyakov YI, Sharp J, Yu CJ. Thermomechanical cyclic response of an ultrafine-grained NiTi shape memory alloy. *Acta Mater*. 2008; 56:3630–3646.

[43] Urbina C, De la Flor S, Ferrando F. Effect of thermal cycling on the thermomechanical behaviour of NiTi shape memory alloys. *Mat Sci Eng A*. 2009; 501:197–206.

[44] Frost M, Sedlák P, Sippola M, Šittner P. Thermomechanical model for NiTi shape memory wires. *Smart Mater Struct*. 2010; 19:094010.

[45] Bonetti E, Frémond M, Lexcellent C. Global Existence and Uniqueness for a Thermomechanical Model for Shape Memory Alloys with Partition of the Strain. *Math Mech Solids*. 2006; 11:251–275.

[46] Stupkiewicz S, Petryk H. A robust model of pseudoelasticity in shape memory alloys. *Int J Numer Meth Engng*. 2013; 93:747–769.

[47] Šittner P, Landa M, Lukáš P, Novák V. R-phase transformation phenomena in thermomechanically loaded NiTi polycrystals. *Mech Mater*. 2006; 38:475–492.

[48] Frost M. Modeling of phase transformations in shape memory materials. Charles University in Prague, Faculty of Mathematics and Physics; 2012.

[49] Mainik A, Mielke A. Existence results for energetic models for rate-independent systems. *Calc Var Partial Differ Equ*. 2005; 22(1):73–99.

[50] Mielke A. Evolution of rate-independent systems. In: C Dafermos EF, editor. Handbook of Differential Equations: Evolutionary Diff. Eqs. Amsterdam: Elsevier/North-Holland; 2005. p. 561–559.

[51] Mielke A, Theil F. On rate-independent hysteresis models. *NODEA-Nonlinear Diff*. 2004; 11(2):151–189.

[52] Krejčí P, Stefanelli U. Existence and non-existence for the full thermomechanical Souza–Auricchio model of shape memory wires. *Math Mech Solids*. 2011; 16:349–365.

[53] Kružík M, Mielke A, Roubíček T. Modelling of microstructure and its evolution in shape-memory-alloy single-crystals, in particular in CuAlNi. *Meccanica*. 2005; 40(5-6):389–418.

[54] Ambrosio L, Fusco N, Pallara D. Functions of Bounded Variation and Free Discontinuity Problems. Oxford Mathematical Monographs. Clarendon Press; 2000.

[55] Dal Maso G, DeSimone A, Mora M, Morini M. A vanishing viscosity approach to quasistatic evolution in plasticity with softening. *Arch Ration Mech An*. 2008; 189(3):469–544.

[56] Efendiev M, Mielke A. On the rate-independent limit of systems with dry friction and small viscosity. *J Conv Anal*. 2006; 13(1):151.

[57] Knees D, Mielke A, Zanini C. On the inviscid limit of a model for crack propagation. *Math Models Methods Appl Sci*. 2008; 18(09):1529–1569.

[58] Mielke A, Rossi R. Existence and uniqueness results for a class of rate-independent hysteresis problems. *Math Models Methods Appl Sci*. 2007; 17(1):81–123.

[59] Nguyen QS. Stability and Nonlinear Solid Mechanics. John Wiley; 2000.

[60] Mielke A, Petrov A. Thermally driven phase transformation in shape-memory alloys. *Adv Math Sci Appl*. 2007; 17(2):667–685.

[61] Mielke A, Paoli L, Petrov A. On existence and approximation for a 3D model of thermally induced phase transformations in shape-memory alloys. *SIAM J Math Anal*. 2009; 41(4):1388–1414.

[62] Roubíček T. Thermodynamics of rate-independent processes in viscous solids at small strains. *SIAM J Math Anal*. 2010; 42(1):256–297.

[63] Benešová B, Roubíček T. Micro-to-meso scale limit for shape-memory-alloy models with thermal coupling. *Multiscale Model Simul*. 2012; 10(3):1059–1089.

[64] Dacorogna B. Direct Methods in the Calculus of Variations. vol. 78 of *Applied Mathematical Sciences*. 2nd ed. Springer-Verlag, New York; 2008.

[65] Barbu V, Precupanu T. Convexity and Optimization in Banach Spaces. Springer, New York; 2012.

[66] Dal Maso G, Francfort GA, Toader R. Quasistatic crack growth in nonlinear elasticity. *Arch Ration Mech An*. 2005; 176(2):165–225.

[67] Hackl K, Fischer FD. On the relation between the principle of maximum dissipation and inelastic evolution given by dissipation potentials. *Proc R Soc A*. 2008; 464:117–132.

[68] Bourdin B. Numerical implementation of the variational formulation for quasi-static brittle fracture. *Interface Free Bound*. 2007; 9(3):411–430.

[69] Bourdin B, Francfort GA, Marigo JJ. The variational approach to fracture. *J Elasticity*. 2008; 91(1-3):5–148.

[70] Nelder JA, Mead R. A simplex method for function minimization. *Comput J*. 1965; 7:308–313.

[71] Gall K, Sehitoglu H, Chumlyakov YI, Kireeva IV. Tension-compression asymmetry of the stress-strain response in aged single crystal and polycrystalline NiTi. *Acta Mater.* 1999; 43:1203–17.

[72] Laydi MR, Lexcellent C. Yield Criteria for Shape Memory Materials: Convexity Conditions and Surface Transport. *Math Mech Solids.* 2010; 15:165–208.