Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org

pp. **X–XX**

THERMAL CONTROL OF THE SOUZA-AURICCHIO MODEL FOR SHAPE MEMORY ALLOYS

Michela Eleuteri

Università degli Studi di Verona,
Strada le Grazie 15, I-37134 Verona, Italy.

Luca Lussardi

Università Cattolica del Sacro Cuore,
via dei Musei 41, I-25121 Brescia, Italy.

Ulisse Stefanelli

Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes" - CNR,
via Ferrata 1, I-27100 Pavia, Italy and
Weierstrass-Institut für Angewandte Analysis und Stochastik,
Mohrenstrasse 39, D-10117 Berlin, Germany.

(Communicated by the associate editor name)

Abstract. We address the thermal control of the quasi-static evolution of a polycrystalline shape memory alloy specimen. The thermomechanical evolution of the body is described by means of the phenomenological Souza-Auricchio model [6, 53]. By assuming to be able to control the temperature of the body in time we determine the corresponding quasi-static evolution in the *energetic* sense. By recovering in this context a result by Rindler [49, 50] we prove the existence of optimal controls for a suitably large class of cost functionals and comment on their possible approximation.

Dedicated to Professor Michel Frémond on the occasion of his 70th birthday.

1. **Introduction.** Shape memory alloys (SMAs) are examples of *active materials*: comparably large strains can be induced (activated) by means of external mechanical, thermal, or magnetic stimuli. At suitably high temperatures SMAs completely recover strains as large as 8% during loading-unloading cycles (note that conventional steels plasticize around 1% strains). This is the so-called *super-elastic* SMA behavior. At lower temperatures permanent deformations remain under unloading. Still, the specimen can be forced to recover its original shape by heating: this is the so called *shape-memory* effect. Finally, some specific SMAs are *ferro-magnetic*: completely recoverable strains can be induced by the action of an external magnetic field. This amazing macroscopic behavior is the effect of an abrupt and diffusionless solid-solid phase transformation between different crystallographic configurations (phases): the *austenite* (mostly cubic, predominant at high temperature and low stresses) and the *martensites* (lower symmetry variants, favored at low temperature or high stresses) [21, 22, 23].

2000 *Mathematics Subject Classification.* 74C05; 49J20.
Key words and phrases. Shape Memory Alloys, optimal control.

The remarkable thermo-magneto-mechanical behavior of SMAs is at the basis of a variety of innovative applications ranging from sensors and actuators, to Aerospace, Biomedical, and Seismic Engineering [16], just to mention a few hot topics. Correspondingly, the interest for the efficient modeling, analysis, and control of SMAs behavior has trimmed an intense research activity in the last decades and a whole menagerie of models has been proposed by addressing different alloys (NiTi, CuAlNi, Ni_2MnGa, among many others) at different scales (atomistic, microscopic with micro-structures, mesoscopic with volume fractions, macroscopic) and emphasizing different principles (minimization of stored energy vs. maximization of dissipation, phenomenology vs. rational crystallography and Thermodynamics) and different structures (single crystals vs. polycrystalline aggregates, possibly including intragranular interaction) [51]. Without any claim of completeness, we shall refer to [4, 19, 21, 25, 26, 31, 32, 33, 43, 47, 48, 55] for macroscopic SMA modeling results.

Our focus here is on a phenomenological, internal-variable-type model for polycrystalline materials which is able of describing both the shape-memory and the super-elastic effect. The model has been originally advanced in the small-strain regime by Souza, Mamiya, & Zouain [53] and then combined with finite elements by Auricchio & Petrini [6, 7, 8]. We hence resort in referring to it as the *Souza-Auricchio* model in the following.

The motivation our the interest of Souza-Auricchio model is twofold. Firstly, the model is extremely *simple*: in the three-dimensional situation, the constitutive behavior of the specimen is determined by the knowledge of just 8 material parameters (note that linearized thermo-plasticity with linear hardening already requires 5 material parameters) which can be easily fitted from real experimental data. Secondly, the Souza-Auricchio model has a remarkable *variational structure* being indeed formulated within the by-now classical frame of generalized (thermo-)plasticity models. This in particular entails both the robustness of the Souza-Auricchio model with respect to approximations and discretizations and its efficiency in accommodating modifications and extensions to more general situations. In particular, the Souza-Auricchio model has been extended in order to include permanent deformation effects [9, 10], asymmetric material behavior [1], ferromagnetic effects [2, 3], and finite strains [17, 18].

From the mathematical viewpoint, existence and approximation of solutions of the Souza-Auricchio model in three-dimensional isothermal quasi-static evolution problem has been obtained [5] and convergence rates for space-time discretization of the problem are derived in [38, 39]. The analysis of the extension of the Souza-Auricchio model including permanent deformations is detailed in [34], the ferromagnetic model is discussed in [11, 12, 54], and the analysis of the finite strains situation is in [24]. Results in the direction of including temperature changes in the Souza-Auricchio model have been obtained by Mielke, Paoli, & Petrov [37, 40] (given temperature), [44] (unknown temperature but viscous), and in [29, 30] (unknown temperature, 1D).

The focus of this paper is on the thermal control of a SMA specimen under the Souza-Auricchio model. The control of SMA devices is obviously of a paramount importance with respect to applications. As such, it represents a clear emerging front in the vast Engineering SMA literature. The corresponding mathematical literature is comparably much less developed. The Reader shall however be referred to [14, 15, 27, 28, 45, 52] for a selection of results. Here, we assume to be able to control the temperature of the specimen in time. This is particularly the case when a SMA body is relatively *thin* in at least one direction and undergoes relatively *low-frequency* loading-unloading cycles. In this case the heat produced via deformation and phase-change can be assumed to be (almost) instantaneously dissipated in the environment (in contrast, note that *high-frequency* mechanical probing on *wires* reveals that the heat production due to the dissipative phase-transformation is not at all negligible [46]).

A second *Ansatz* of our analysis is that the temperature is *space-homogeneous* in the bulk. We concentrate on this setting as we assume that the temperature of the specimen is here the control variable and we see little motivation (if any) in considering a control device able to prescribe a non-homogeneous temperature in a metallic bulk. On the other hand, in case of a fully coupled thermomechanical system (a case which is not under consideration here as the temperature is given) an inhomogeneous temperature evolution is generically to be expected, even in the simple one-dimensional geometry of a wire. Although we stick here to the homogeneous situation, we have however to stress that the non-homogeneous temperature case would also be amenable from the mathematical viewpoint under minor modifications.

The novelty of our contribution is twofold. At first, we establish a novel existence result for the so-called *state problem*, namely, given the temperature, to determine the quasi-static mechanical evolution of the SMA specimen. As we comment below, our existence result is somewhat stronger than the available ones from [37, 40, 44]. In particular, in contrast to the above-mentioned papers, the original non-regularized formulation of the Souza-Auricchio model can be directly accommodated in our setting. Moreover, less regularity is required on the given temperature. This last aspect turns out to be crucial with respect to optimal control as we aim at considering the largest possible set of control temperatures.

Our second novel point is that of proving the existence of an optimal control for a suitably large class of cost functionals depending on both mechanics and temperature. The applicative interest in this perspective resides in the possibility of efficiently activating SMA devices by controlling the temperature of the specimen via Joule's heating. This is one of the basic technological activation mechanisms currently exploited in real applications [16]. This is to our knowledge the first control result in the specific setting of the Souza-Auricchio model and, more generally, of (thermo-)generalized plasticity. Our argument is basically the concrete application of the abstract theory developed by Rindler [49] on existence of optimal controls in the frame of rate-independent systems. Note that, as the state problem is non-smooth and, in particular, not even known to admit a *unique* solutions, the quest for necessary optimality conditions seems presently out of reach. We collect some comments on possible future developments of this investigation in Subsection 4.1 below.

This is the plan of the paper. In Section 2 we recall the basic features of the Souza-Auricchio model, collect our assumptions, and state the main results. Section 3 reports on the analysis of the state problem whereas the existence proof for optimal controls is developed in Section 4 together with some comments for possible future research.

2. Main results.

2.1. Mechanical problem formulation.

We shall start by recalling some ingredients of the Souza-Auricchio model. The Reader is however referred to the original contributions in [6, 7, 8, 53] for further motivation, detail, numerical illustration, and validation.

We denote by $\mathbb{R}^{d\times d}_{\rm sym}$ $(d=2,3)$ the space of symmetric 2-tensors in $\mathbb{R}^d$ endowed with the usual scalar (contraction) product $a{:}b = {\rm tr}(ab) := a_{ij}b_{ij}$ (summation convention) and the corresponding norm $|a| = \sqrt{a{:}a}$. The space $\mathbb{R}^{d\times d}_{\rm sym}$ is orthogonally decomposed as $\mathbb{R}^{d\times d}_{\rm sym} = \mathbb{R}^{d\times d}_{\rm dev} \oplus \mathbb{R}1_2$, where $\mathbb{R}1_2$ is the subspace spanned by the identity 2-tensor 1_2, while $\mathbb{R}^{d\times d}_{\rm dev}$ is the subspace of all *deviatoric* symmetric tensors.

The non-empty, connected, bounded, and open subset $\Omega \subset \mathbb{R}^d$ represents the reference configuration of the body. Given the displacement $u : \Omega \to \mathbb{R}^d$ from the reference configuration with $u \in H^1(\Omega;\mathbb{R}^d)$ we denote the symmetric gradient $(\nabla u + \nabla u^\top)/2$ of u by $\varepsilon(u)$. In particular, throughout the paper we will make tacit use of the well-known Korn inequality

$$c_{\rm Korn}||u||^2_{H^1(\Omega;\mathbb{R}^d)} \le ||u||^2_{L^2(\Gamma_{\rm Dir};\mathbb{R}^d)} + ||\varepsilon(u)||^2_{L^2(\Omega;\mathbb{R}^{d\times d}_{\rm sym})}$$

for any $u \in H^1(\Omega;\mathbb{R}^d)$ and some constant $c_{\rm Korn} > 0$ depending just on Ω.

Within the small-strain regime, the linearized strain $\varepsilon(u)$ is additively decomposed into an *elastic part* $\varepsilon_{\rm el}$ and an *inelastic* (or transformation) part z as

$$\varepsilon(u) = \varepsilon_{\rm el} + z.$$

The elastic part of the linearized strain is directly related to the stress exerted on the body. In particular, we assume the elastic material response $\sigma = \mathbb{C}\varepsilon_{\rm el}$ where σ stands for the stress and $\mathbb{C}$ is the positive-definite and symmetric elasticity 4-tensor.

The internal variable $z \in \mathbb{R}^{d\times d}_{\rm dev}$ will be regarded as the tensorial descriptor of the internal phase distribution in the material. The deviatoric nature of z reflects the fact that martensitic transitions are assumed to be volume preserving. The norm $|z|$ represents a measure of product phase (detwinned martensite) vs. parent phase (austenite and twinned martensite). The direction $z/|z|$ instead describes the predominant orientation of the product phase.

In order to better illustrate the role of the internal variable z let us momentarily leave our polycrystalline framework and consider the single-crystal case of a cubic-tetragonal martensitic system (as such of NiMgGa, FePd, and FePt alloys, among many others). By fixing a reference frame (e_1, e_2, e_3) aligned with the edges of the cubic austenitic crystal, the corresponding inelastic strain tensors z_i relative to the i-th tetragonal martensitic phase $(i = 1, 2, 3)$ read

$$z_i = \frac{\varepsilon_L}{\sqrt{6}}(1_2 - 3e_i \otimes e_i) \quad \forall i = 1, 2, 3.$$

Here, $\varepsilon_L > 0$ denotes the maximal strain which is obtainable via reorientation of martensitic variants. In particular, $|z_i| = \varepsilon_L$. By indicating with λ_i the local phase proportion of the i-th martensitic phase ($0 \le \lambda_i \le 1$ and $\lambda_1+\lambda_2+\lambda_3 \le 1$), one gets the overall inelastic strain to be of the form $z = \lambda_1 z_1 + \lambda_2 z_2 + \lambda_3 z_3$ so that

$$|z| = \varepsilon_L(\lambda_1^2+\lambda_2^2+\lambda_3^2) \quad \text{and} \quad \frac{z}{|z|} = \frac{\lambda_1 z_1 + \lambda_2 z_2 + \lambda_3 z_3}{\varepsilon_L(\lambda_1^2+\lambda_2^2+\lambda_3^2)}.$$

In particular, the first equality above qualifies $|z|$ as an indicator of the total proportion of (detwinned) martensitic phase in the material whereas the director $z/|z|$

distinguishes among the different martensites as we have that

$$\frac{z}{|z|} = \frac{z_i}{\varepsilon_L} \quad \Longleftrightarrow \quad \lambda_i = 1.$$

We assume the boundary $\partial\Omega$ to be Lipschitz and we let $\Gamma_{\rm Dir} \cup \Gamma_{\rm tr} = \partial\Omega$ where $\Gamma_{\rm Dir} \cap \Gamma_{\rm tr} = \emptyset$ and $\mathcal{H}^{d-1}(\Gamma_{\rm Dir}) > 0$. The body will be subject to a given surface traction on the part $\Gamma_{\rm tr}$ of the boundary. On the other hand, non-homogeneous Dirichlet conditions for the displacement will be prescribed on $\Gamma_{\rm Dir}$. More precisely, by letting

$$u^{\rm Dir} \in W^{1,1}(0,T;H^1(\Omega;\mathbb{R}^d)) \tag{1}$$

be given, the trace of $u^{\rm Dir}$ on $\Gamma_{\rm Dir}$ plays the role of the prescribed boundary value for the displacement u. In particular, for all given times $t \in [0,T]$, the set of admissible states $(u(t), z(t))$ is given by $\mathcal{Y}(u^{\rm Dir}(t))$ where

$$\bar{u} \in H^1(\Omega;\mathbb{R}^d) \mapsto \mathcal{Y}(\bar{u}) = \Big\{(u,z) \in H^1(\Omega;\mathbb{R}^d) \times L^2(\Omega;\mathbb{R}^{d\times d}_{\rm dev}) \ : \ u = \bar{u} \text{ on } \ \Gamma_{\rm Dir}\Big\}.$$

We assume each phase to be isotropic and described by the same elasticity tensor. This is of course a crude simplification but still does not jeopardize the overall output of the model [6, 53]. In particular, we denote the *elastic energy* functional $\mathcal{C} : H^1(\Omega;\mathbb{R}^{d\times d}_{\rm sym}) \to [0,\infty)$ as

$$\mathcal{C}(a) := \frac{1}{2}\int_\Omega a{:}\mathbb{C}a \,{\rm d}x.$$

Hence, the elastic contribution to the stored energy of the material is simply given by $\mathcal{C}(\varepsilon_{\rm el}) = \mathcal{C}(\varepsilon(u){-}z)$.

Following the original modelization from [6, 53], the inelastic part of the stored energy of the material is a function of the inelastic strain z and of the temperature θ only. In particular, the *inelastic energy* density takes the form

$$(\theta, z) \mapsto \beta(\theta)|z| + \frac{c_{\rm h}}{2}|z|^2 + I_K(z). \tag{2}$$

Here, β is a given Lipschitz continuous nonnegative function describing the temperature dependence of the inelastic response of the medium. In particular, $\beta(\theta)$ corresponds to the austenite-martensite transition critical stress at temperature $\theta > 0$. The original choice $\beta(\theta) = b(\theta{-}\theta_{\rm m})^+$ of the Souza-Auricchio model ($b > 0$ and $\theta_{\rm m}$ being a critical temperature for the martensite-austenite equilibrium in the stress-free configuration) is included in our frame.

The constant $c_{\rm h} > 0$ is a classical linearized kinematic hardening coefficient and I_K stands for the indicator function of the closed, convex set $K := \{z \in \mathbb{R}^{d\times d}_{\rm dev} \ : \ |z| \leq \varepsilon_L\}$. In particular, $I_K(z) = 0$ if $z \in K$ and $I_K = \infty$ elsewhere. In passing, one shall note that the existence and optimal control issues discussed here do not rely on the particular form of the inelastic energy and could be adapted to much more general situations. We however prefer to stick to the original modeling choice in (2) for the sake of definiteness.

A last term has to be introduced in the overall stored energy of the system in order to penalize martensite-martensite interfaces. Indeed, we include the *interfacial energy* term

$$z \mapsto \nu \int_\Omega |\nabla z|$$

where $\nu > 0$ is a scale parameter. In particular, $1/\nu$ represents the overall length of martensite-martensite interfaces within some reference domain. Note that the latter integral bears the meaning of a total variation and, as such, bears also a crucial compactifying effect. The occurrence of this interfacial term however does not prevent z from possibly exhibiting jumps. This is a particularly desirable feature in connection with shape memory alloys where sharp (often just few atomic layers thick) phase boundaries are usually observed.

By assuming the temperature of the body $\theta \in W^{1,1}(0,T)$ to be (spatially homogeneous and) prescribed, the stored-energy functional $\mathcal{W}(\cdot,\cdot;\theta(t)) : \mathcal{Y}(u^{\rm Dir}(t)) \to [0,+\infty]$ for the body at time $t \in [0,T]$ will be hence given by

$$\begin{aligned}\mathcal{W}(u,z;\theta(t)) &= \mathcal{C}(\varepsilon(u){-}z) + \int_\Omega \Big(\frac{c_{\rm h}}{2}|z|^2{+}I(z)\Big)\,{\rm d}x + \nu\int_\Omega |\nabla z| + \int_\Omega \beta(\theta(t))|z|\,{\rm d}x\\ &=: \mathcal{E}(u,z) + \mathcal{F}(\theta(t),z)\end{aligned}$$

In particular, the functional $\mathcal{E}(u,z)$ collects all terms above which are independent of time (i.e., of the temperature θ) whereas $\mathcal{F}(\theta(t),z)$ contains the only temperature-driven term. Note that, along with the above provisions, the stored-energy functional $\mathcal{W}(\cdot,\cdot;\theta(t))$ turns out to be uniformly convex in $H^1(\Omega;\mathbb{R}^d)\times L^2(\Omega;\mathbb{R}^{d\times d}_{\rm dev})$ albeit non-smooth. In particular, its domain is the convex and closed subset of $\mathcal{Y}(u^{\rm Dir}(t))$ given by $\{(u,z)\in\mathcal{Y}(u^{\rm Dir}(t)) \ : \ z\in K \text{ a.e. in } \Omega\}$.

In addition to the above-prescribed boundary displacement conditions on $\Gamma_{\rm Dir}$, we shall consider some imposed body force f and surface traction g, as well. We assume to be given

$$f\in W^{1,1}(0,T;L^2(\Omega;\mathbb{R}^d)), \quad g\in W^{1,1}(0,T;L^2(\Gamma_{\rm tr};\mathbb{R}^d)) \tag{3}$$

and define the *total load* $\ell \in W^{1,1}(0,T;(H^1(\Omega;\mathbb{R}^d))')$ as

$$\langle \ell(t),u\rangle := \int_\Omega f(t){\cdot}u\,{\rm d}x + \int_{\Gamma_{\rm tr}} g(t){\cdot}u\,{\rm d}\mathcal{H}^{d-1} \quad \forall u\in H^1(\Omega;\mathbb{R}^d),\ t\in[0,T].$$

Starting from a suitable initial state (u^0,z^0), the dissipative evolution of the quasi-static system is governed by the occurrence of a *dissipation (pseudo-)potential* $\mathcal{D}\colon L^1(\Omega;\mathbb{R}^{d\times d}_{\rm dev})\to[0,\infty)$ given by

$$\mathcal{D}(\dot z) := R\int_\Omega |\dot z|\,{\rm d}x$$

for some activation radius $R>0$. Note that the dissipation potential acts on the rate $\dot z$ for it encodes the action of dissipative forces. Moreover, we let the *total dissipation* of the process on the time interval $[s,t]\subseteq[0,T]$ be given by

$${\rm Diss}_{\mathcal{D}}(z,[s,t]) := \sup\Big\{\sum_{i=1}^N \mathcal{D}(z(t_i){-}z(t_{i-1})) : \{s=t_0<t_1<\dots<t_N=t\}\Big\},$$

the sup being taken among all partitions of the interval $[s,t]$.

For the sake of later purposes, we shall define the set of *stable states* $\mathcal{S}(t,\theta)$ at time t and for the temperature θ as

$$\begin{aligned}\mathcal{S}(t,\theta) := \Big\{&(u,z)\in\mathcal{Y}(u^{\rm Dir}(t)) : \mathcal{E}(u,z)<\infty \ \text{ and } \ \forall(\bar u,\bar z)\in\mathcal{Y}(u^{\rm Dir}(t)),\\ &\mathcal{E}(u,z)+\mathcal{F}(\theta,z)-\langle\ell(t),u\rangle \le \mathcal{E}(\bar u,\bar z)+\mathcal{F}(\theta,\bar z)-\langle\ell(t),\bar u\rangle+\mathcal{D}(z-\bar z)\Big\}.\end{aligned} \tag{4}$$

The relevance of this notion resides on the fact that, given a stable state (u,z), no competitor state $(\bar u,\bar z)$ can be preferred in terms of balance between total energy and dissipation. This global minimality requirement, sometimes disputable in other situations, is here fully motivated by the convexity of the total energy of the body.

2.2. **Energetic solvability of the state problem.** Assume that we are given the space-homogeneous temperature $\theta \in W^{1,1}(0,T)$. We shall be interested in establishing the existence of a suitably weak solution to the quasi-static *state problem*

$$
\begin{aligned}
\mathbb{C}(\varepsilon(u){-}z) = \sigma \quad & \text{in } \Omega \times (0,T),\\
\nabla{\cdot}\sigma + f = 0 \quad & \text{in } \Omega \times (0,T),\\
u = u^{\mathrm{Dir}} \quad & \text{on } \Gamma_{\mathrm{Dir}} \times (0,T),\\
\sigma n = g \quad & \text{on } \Gamma_{\mathrm{tr}} \times (0,T),\\
\partial\mathcal{D}(\dot z(t)) + \partial_z \mathcal{W}(u(t),z(t);\theta(t)) \ni 0 \quad & \text{in } L^2(\Omega;\mathbb{R}^{d\times d}_{\mathrm{dev}}), \forall t \in (0,T),\\
u(0) = u^0, \quad z(0) = z^0 \quad & \text{in } \Omega
\end{aligned}
$$

where σ stands for the stress, n is the outward normal to Γ_{tr}, and ∂ is the (possibly partial) subdifferential in $L^2(\Omega;\mathbb{R}^{d\times d}_{\mathrm{dev}})$ in the sense of convex analysis [13]. In particular, we focus on *energetic solutions* à la Mielke [36, 42], namely trajectories $t \mapsto (u(t),z(t)) \in \mathcal{Y}(u^{\mathrm{Dir}}(t))$ such that $(u(0),z(0)) = (u^0,z^0)$, $t \mapsto \langle \dot\ell(t), u(t)\rangle$ and $t \mapsto \beta'(\theta(t))\dot\theta(t)|z(t)|$ are integrable, and, for all $t \in [0,T]$, we have the two conditions:

Stability:

$$(u(t),z(t)) \in \mathcal{S}(t,\theta(t)). \tag{5}$$

Energy balance:

$$
\begin{aligned}
&\mathcal{E}(u(t),z(t)) + \mathcal{F}(\theta(t),z(t)) - \langle \ell(t),u(t)\rangle + \mathrm{Diss}_{\mathcal{D}}(z,[0,t])\\
&= \mathcal{E}(u(0),z(0)) + \mathcal{F}(\theta(0),z(0))\\
&- \langle \ell(0),u(0)\rangle + \int_0^t \int_\Omega \beta'(\theta(s))\dot\theta(s)|z| \,\mathrm{d}x\,\mathrm{d}s - \int_0^t \langle \dot\ell(s),u(s)\rangle \,\mathrm{d}s.
\end{aligned} \tag{6}
$$

Our result on the state problem reads as follows.

Theorem 2.1 (Existence for the state problem)**.** *Assume* (1) *and* (3)*. Given* $\theta \in W^{1,1}(0,T)$ *and an initial value* $(u^0,z^0) \in \mathcal{S}(0,\theta(0))$ *there exists an* energetic solution (u,z) *of the state problem in the sense of* (5)-(6)*. Moreover, all energetic solutions belong to* $W^{1,1}(0,T;H^1(\Omega;\mathbb{R}^d){\times}L^2(\Omega;\mathbb{R}^{d\times d}_{\mathrm{dev}}))$*.*

The proof of this result follows via time discretization by applying the general strategy for existence of energetic solutions of rate-independent systems [36]. In particular, we shall implement in our concrete case the improvements introduced by Rindler [49] to the by-now classical existence proof by Francfort & Mielke [20]. In particular, we cannot rely here on the absolute continuity of the power of external actions and shall hence provide a refined discussion of the relative terms into the energy balance equation. A sketch of this argument is given in Section 3 below.

Our result shall be compared with the former analysis by Mielke, Petrov, & Paoli in [37, 40] where existence theories for the Souza-Auricchio model in the non-isothermal case have been firstly obtained. These results differ from ours under three aspects. At first, in [37, 40] a regularization (smoothing) of the inelastic energy density (both in θ and z) is needed. This in particular prevents the Authors from considering in [37, 40] the original inelastic energy density of Souza-Auricchio in (2) which is instead included in our analysis. On the other hand, one shall note that the smooth setting of [37, 40] allows for a uniqueness theory which is here out of reach. Secondly, the given temperature in [37, 40] was taken to be C^1-regular in time whereas we just require absolute continuity. This issue is indeed crucial here as we are aiming at optimal controls and compactness of temperatures (controls) will clearly play a role. Finally, we include here the total variation of z into the stored energy instead of its squared gradient as in [37, 40]. This has the advantage of allowing discontinuous in space inelastic strains z and of requiring no extra (and somewhat disputable) boundary conditions.

2.3. **Optimal control.** We shall now come to the controllability statement. As we have already mentioned, we are here focusing on the situation of a SMA specimen which deforms under given mechanical loading under the influence of a controlled space-homogeneous time-dependent temperature θ. In particular, given the temperature $\theta \in W^{1,1}(0,T)$ let us denote by

$$\mathrm{Sol}(\theta) \subset W^{1,1}(0,T; H^1(\Omega;\mathbb{R}^d) \times L^2(\Omega;\mathbb{R}^{d\times d}_{\mathrm{dev}}))$$

the set of energetic solutions of the state problem from Theorem 2.1. Let the set of admissible temperatures (controls) be denoted by $\Theta \subset W^{1,1}(0,T)$. Then, the optimal control problem consists in the minimization of a given cost functional

$$\mathcal{J} : W^{1,1}(0,T; H^1(\Omega;\mathbb{R}^d) \times L^2(\Omega;\mathbb{R}^{d\times d}_{\mathrm{dev}})) \times \Theta \to (-\infty, \infty]$$

which is depending on both the energetic solution and the control. Our problem is to find an *optimal control* $\theta_* \in \Theta$ and a corresponding *optimal energetic solution* $(u_*, z_*) \in \mathrm{Sol}(\theta_*)$ such that

$$(u_*, z_*) \in \mathrm{Arg\,Min}\,\{\mathcal{J}(u,z,\theta) \mid (u,z) \in \mathrm{Sol}(\theta),\ \theta \in \Theta\}\,.$$

In order to possibly find optimal controls we shall consider the following standard requirements.

Compatibility of the initial value and the controls:

$$(u^0, z^0) \in \mathcal{S}(0, \theta(0)) \quad \forall \theta \in \Theta. \tag{7}$$

Compactness of controls:

$$\Theta \text{ is weakly compact in } W^{1,r}(0,T) \text{ for some } r > 1. \tag{8}$$

Lower semicontinuity of the cost functional:

$$\left.\begin{array}{l} \theta_n \to \theta \text{ weakly in } W^{1,r}(0,T) \\ (u_n, z_n) \in \mathrm{Sol}(\theta_n), \\ (u_n, z_n) \to (u,z) \text{ weakly-star in} \\ L^\infty(0,T; H^1(\Omega;\mathbb{R}^d) \times L^2(\Omega;\mathbb{R}^{d\times d}_{\mathrm{dev}})) \end{array}\right\} \Rightarrow \ \mathcal{J}(u,z,\theta) \le \liminf_{n\to\infty} \mathcal{J}(u_n, z_n, \theta_n). \tag{9}$$

The compatibility condition in (7) was already presented in [49] and is just intended to ensure that the initial values are stable regardless of the choice of the control. In case all $\theta \in \Theta$ share the same initial value (which is somehow natural in applications where $\theta(0)$ is usually room temperature) the compatibility condition (7) can be easily checked. On the other hand, note that if $\ell(0) = 0$ the choice

$z = 0$ (which corresponds to pure austenite for high temperatures and specific accommodated martensites at low temperatures) fulfills the compatibility (7) under no restrictions on Θ.

The compactness of Θ from (8) is here chosen just for the sake of simplicity. In particular it can be relaxed by asking for some extra coercivity with respect to θ on the functional $\mathcal{J}$. We stick to assumption (8) for the sake of definiteness only.

The lower semicontinuity requirement in (9) is standard. For the sake of illustration, let us remark that a possible *quadratic* cost functional covered by our theory is

$$\mathcal{J}(u,z,\theta) = \int_0^T\int_\Omega |u-u_\mathrm{d}|^2 \mathrm{d}x\,\mathrm{d}t + \int_0^T\int_\Omega |z-z_\mathrm{d}|^2 \mathrm{d}x\,\mathrm{d}t + \int_\Omega |u(T)-u_\mathrm{f}|^2\,\mathrm{d}x + \int_\Omega |z(T)-z_\mathrm{f}|^2\,\mathrm{d}x$$

where $u_\mathrm{d} : [0,T] \to L^2(\Omega;\mathbb{R}^d)$, $z_\mathrm{d} : [0,T] \to L^2(\Omega;\mathbb{R}^{d\times d})$ are given desired displacement and inelastic strain profiles whereas $u_\mathrm{f} \in L^2(\Omega;\mathbb{R}^d)$ and $z_\mathrm{f} \in L^2(\Omega;\mathbb{R}^{d\times d})$ are given target states. Note that the latter functional is not lower semicontinuous with respect to the weak-star topology of

$$L^\infty(0,T;H^1(\Omega;\mathbb{R}^d)\times L^2(\Omega;\mathbb{R}^{d\times d}_\mathrm{dev})).$$

Still, it fulfills (9) as an effect of the requirement $(u_n, z_n) \in \mathrm{Sol}(\theta_n)$ which indeed provides extra compactness.

Our main result is the following.

Theorem 2.2 (Existence of optimal controls)**.** *Under assumptions* (1)*,* (3)*, and* (7)*-*(9) *there exists an optimal control* θ_* *and a corresponding optimal energetic solution* $(u_*, z_*) \in \mathrm{Sol}(\theta_*)$*.*

We shall give a direct proof of this theorem in Section 4 for the sake of completeness. Still, let us mention that the result may be equivalently obtained by applying in the present situation the abstract theory developed by Rindler [49].

3. **State problem.** This section provides a proof of Theorem 2.1. As already mentioned, the argument follows the by-now classical proof of convergence of time discretization for rate-independent evolution problems [36]. Still, exactly in the same spirit of [49], some specific care is devoted to ascertain the convergence of the work of external actions. In particular, these work read here

$$\int_0^t\int_\Omega \beta'(\theta(s))\dot\theta(s)|z|\,\mathrm{d}x\,\mathrm{d}s - \int_0^t \langle\dot\ell(s), u(s)\rangle\,\mathrm{d}s \qquad (10)$$

which fails to fulfill the classical absolute continuity requirement [36, Assumption (A5)] of the general energetic solvability theory as θ and ℓ are just absolutely continuous here. Note that the former existence results for given temperature in [37, 40] by-pass this problem by requiring the temperature to be C^1. This would however be not satisfactory here as we are indeed interested in optimal control via θ. Namely, we shall better consider the least possible time-regularity for the temperature θ entailing the solvability of the state problem. We circumvent here the lack of absolute continuity of the above work by directly considering the concrete from (10). For the sake of completeness we aim at providing here a sketch of the proof.

Let us start by changing variables in order to reduce to a time-independent state space. In particular, we let $v = u - u^{\rm Dir}$ and focus on the pair (v,z) taking values in the space $\mathcal{Y}_0 := \mathcal{Y}(0)$. We easily compute that

$$\mathcal{E}(u,z) = \mathcal{E}(v,z) + \int_\Omega \varepsilon(u^{\rm Dir}){:}\mathbb{C}(\varepsilon(v){-}z)\,{\rm d}x + \mathcal{C}(\varepsilon(u^{\rm Dir}))$$

and

$$\langle \ell(t), u\rangle = \langle \ell(t), v\rangle + \langle \ell(t), u^{\rm Dir}\rangle.$$

Let now $L : [0,T] \to \mathcal{Y}_0'$ be given by

$$\langle L(t),(v,z)\rangle := -\int_\Omega \varepsilon(u^{\rm Dir}(t)){:}\mathbb{C}(\varepsilon(v){-}z)\,{\rm d}x + \langle \ell(t), v\rangle \quad \forall (v,z)\in\mathcal{Y}_0,\ t\in[0,T]$$

and notice that $L \in W^{1,1}(0,T;\mathcal{Y}_0')$. Eventually, (u,z) is an energetic solution of the quasi-static evolution problem (5)-(6) if and only if $(v,z) : t \mapsto \mathcal{Y}_0$ is such that $(v(0),z(0)) = (v^0,z^0) := (u^0{-}u^{\rm Dir}(0), z^0)$, $t \mapsto \langle \dot\ell(t), v(t)\rangle$ and $t \mapsto \beta'(\theta(t))\dot\theta(t)|z(t)|$ are integrable, and we have, for all $t\in[0,T]$,

Stability:

$$\begin{aligned}&(v(t),z(t)) \in \mathcal{S}'(t,\theta(t)) := \{(v,z)\in\mathcal{Y}_0 :\ \forall\,(\bar v,\bar z)\in\mathcal{Y}_0,\\ &\mathcal{E}(v,z) + \mathcal{F}(\theta(t),z) - \langle L(t),(v,z)\rangle\\ &\ \le\ \mathcal{E}(\bar v,\bar z) + \mathcal{F}(\theta(t),\bar z) - \langle L(t),(\bar v,\bar z)\rangle + \mathcal{D}(z{-}\bar z)\},\end{aligned} \tag{11}$$

Energy balance:

$$\begin{aligned}&\mathcal{E}(v(t),z(t)) + \mathcal{F}(\theta(t),z(t)) - \langle L(t),(v(t),z(t))\rangle + {\rm Diss}_{\mathcal{D}}(z,[0,t])\\ &\ = \mathcal{E}(v(0),z(0)) + \mathcal{F}(\theta(0),z(0)) - \langle L(0),(v(0),z(0))\rangle\\ &\ + \int_0^t\int_\Omega \beta'(\theta)\dot\theta|z|\,{\rm d}x\,{\rm d}s - \int_0^t \langle\dot\ell, v\rangle\,{\rm d}s.\end{aligned} \tag{12}$$

We shall prove the existence of such (v,z) via time discretization.

Time discretization. Assume to be given a sequence of partitions $\{0 = t_0^n < t_1^n < \cdots < t^n_{N^n-1} < t^n_{N^n} = T\}$ with diameter $\tau^n = \max_{i=1,\ldots,N^n}(t_i^n{-}t_{i-1}^n)$ going to 0 as $n\to\infty$. We inductively define (a sequence of) unique solutions $\{(v_i^n,z_i^n)\}_{i=0}^{N^n}$ of the incremental problems

$$\begin{aligned}(v_i^n,z_i^n) = {\rm Arg\,Min}_{(v,z)\in\mathcal{Y}_0}\Big(&\mathcal{E}(v,z) + \mathcal{F}(\theta(t_i^n),z)\\ &- \langle L(t_i^n),(v,z)\rangle + \mathcal{D}(z{-}z_{i-1}^n)\Big)\end{aligned} \tag{13}$$

for $i = 1,\ldots,N^n$ with $(v_0^n,z_0^n) = (v^0,z^0)$. To this aim it suffices to observe that the map $(u,z)\mapsto \mathcal{E}(u,z) + \mathcal{F}(\theta(t),z) - \langle L(t),(u,z)\rangle + \mathcal{D}(z{-}\bar z)$ is uniformly convex and lower semicontinuous in $\mathcal{Y}_0$ for any given $\theta(t)\in\mathbb{R}$ and $\bar z \in L^1(\Omega;\mathbb{R}^{d\times d}_{\rm dev})$.

Next, we denote by (v_n,z_n) the right-continuous and piecewise constant interpolant of the values $\{(v_i^n,z_i^n)\}_{i=0}^{N^n}$ on the partition. Moreover, we let $t_n\colon [0,T]\to[0,T]$ be given by $t_n(t) = t_{i-1}^n$ for $t\in[t_{i-1}^n, t_i^n)$ for $i\in 1,\ldots,N^n$.

Stability at the discrete level. The minimality in (13) entails that (v_i^n, z_i^n) is stable at time t_i^n, that is $(v_i^n, z_i^n) \in \mathcal{S}'(t_i^n, \theta(t_i^n))$, for all $i = 1, \dots, N$. Indeed, for any $(\bar{v}, \bar{z}) \in \mathcal{Y}_0$, we get

$$\begin{aligned}
&\mathcal{E}(v_i^n, z_i^n) + \mathcal{F}(\theta(t_i^n), z_i^n) - \langle L(t_i^n), (v_i^n, z_i^n)\rangle + \mathcal{D}(z_i^n - z_{i-1}^n) \\
&\le \mathcal{E}(\bar{v}, \bar{z}) + \mathcal{F}(\theta(t_i^n), \bar{z}) - \langle L(t_i^n), (\bar{v}, \bar{z})\rangle + \mathcal{D}(\bar{z} - z_{i-1}^n) \\
&\le \mathcal{E}(\bar{v}, \bar{z}) + \mathcal{F}(\theta(t_i^n), \bar{z}) - \langle L(t_i^n), (\bar{v}, \bar{z})\rangle + \mathcal{D}(\bar{z} - z_i^n) + \mathcal{D}(z_i^n - z_{i-1}^n)
\end{aligned}$$

and the term $\mathcal{D}(z_i^n - z_{i-1}^n)$ cancels out.

Convergence to a time-continuous evolution. Taking into account the minimality (13) of (v_i^n, z_i^n) we deduce that

$$\begin{aligned}
&\mathcal{E}(v_i^n, z_i^n) - \mathcal{E}(v_{i-1}^n, z_{i-1}^n) + \mathcal{F}(\theta(t_i^n), z_i^n) - \mathcal{F}(\theta(t_{i-1}^n), z_{i-1}^n) \\
&- \langle L(t_i^n), (v_i^n, z_i^n)\rangle + \langle L(t_{i-1}^n), (v_{i-1}^n, z_{i-1}^n)\rangle + \mathcal{D}(z_i^n - z_{i-1}^n) \\
&\le \mathcal{F}(\theta(t_i^n), z_{i-1}^n) - \mathcal{F}(\theta(t_{i-1}^n), z_{i-1}^n) - \langle L(t_i^n) - L(t_{i-1}^n), v_{i-1}^n\rangle.
\end{aligned}$$

Summing up for i from 1 to $m \le N^n$, we get

$$\begin{aligned}
&\mathcal{E}(v_m^n, z_m^n) - \mathcal{E}(v^0, z^0) + \mathcal{F}(\theta(t_m^n), z_m^n) - \mathcal{F}(\theta(0), z^0) \\
&- \langle L(t_m^n), (v_m^n, z_m^n)\rangle + \langle L(0), (v^0, z^0)\rangle + \sum_{i=1}^m \mathcal{D}(z_i^n - z_{i-1}^n) \\
&\le \int_0^{t_m^n} \int_\Omega \beta'(\theta)\dot{\theta}|z_n| \,\mathrm{d}x\,\mathrm{d}s - \int_0^{t_m^n} \langle \dot{\ell}, v_n\rangle \,\mathrm{d}s. \qquad (14)
\end{aligned}$$

By exploiting the discrete Gronwall Lemma and the coercivity of $\mathcal{E}$ we deduce that

$$\sup_{t\in[0,T]} \big(\mathcal{E}(v_n(t), z_n(t)) + \mathcal{F}(\theta(t), z_n(t))\big) \ \text{ and } \ \mathrm{Diss}_{\mathcal{D}}(z_n, [0,T])$$
$$\text{are bounded independently of } n. \qquad (15)$$

We make use of the generalization of Helly's selection principle from [35], and find a (not relabeled) subsequence and a non-decreasing function $\phi : [0,T] \to [0,\infty)$ such that

$$\begin{aligned}
&z_n(t) \to z(t) \ \text{ weakly-star in } \ BV(\Omega; \mathbb{R}^{d\times d}_{\mathrm{dev}}) \ \ \forall t \in [0,T] \\
&\mathrm{Diss}_{\mathcal{D}}(z_n, [0,t]) \to \phi(t) \ \ \forall t \in [0,T] \\
&\mathrm{Diss}_{\mathcal{D}}(z, [s,t]) \le \phi(t) - \phi(s) \ \ \forall [s,t] \subset [0,T]. \qquad (16)
\end{aligned}$$

Moreover, by exploiting the quadratic character of $\mathcal{E}(\cdot, z)$ we readily have from minimality (13) that $v_n = \mathcal{L}(z_n, L\circ t_n)$ where $\mathcal{L} : L^1(\Omega; \mathbb{R}^{d\times d}_{\mathrm{dev}}) \times \mathcal{Y}_0' \to H^1(\Omega; \mathbb{R}^d)$ is a linear and continuous operator. In particular $v_n = \mathcal{L}(z_n, L\circ t_n) \to \mathcal{L}(z, L) =: v$.

Stability of the limit trajectory. We now prove that the set

$$\mathbb{S} := \bigcup_{t\in[0,T]} (t, \mathcal{S}'(t, \theta(t)))$$

is closed with respect to the weak topology of $\mathbb{R}\times\mathcal{Y}_0$. Let $(t_k, v_k, z_k) \in \mathbb{S}$ with $t_k \to t$ and $(v_k, z_k) \to (v, z)$ weakly in $\mathcal{Y}_0$. By the lower semicontinuity of $\mathcal{E}$ and $\mathcal{F}$ and taking into account the strong continuity of $\mathcal{D}$ in $L^1(\Omega; \mathbb{R}^{d\times d}_{\mathrm{dev}})$ and the continuity of

θ and L we get

$$\begin{aligned}
&\mathcal{E}(v,z)+\mathcal{F}(\theta(t),z)-\langle L(t),(v,z)\rangle\\
&\leq \liminf_{k\to+\infty}\big(\mathcal{E}(v_k,z_k)+\mathcal{F}(\theta(t_k),z_k)-\langle L(t_k),(v_k,z_k)\rangle\big)\\
&\leq \liminf_{k\to+\infty}\big(\mathcal{E}(\bar v,\bar z)+\mathcal{F}(\theta(t_k),\bar z)-\langle L(t_k),(\bar v,\bar z)\rangle+\mathcal{D}(z_k-\bar z)\big)\\
&=\mathcal{E}(\bar v,\bar z)+\mathcal{F}(\theta(t),\bar z)-\langle L(t),(\bar v,\bar z)\rangle+\mathcal{D}(z-\bar z),
\end{aligned}$$

for any $(\bar v,\bar z)\in\mathcal{Y}_0$. Then $(t,v,z)\in\mathbb{S}$.

The aim is now to exploit the latter closure property in order to prove that $(v(t),z(t))$ is a stable state, i.e. (11) holds. Note that $t\mapsto t_n(t)$ converges uniformly to the identity and $(v_n(t),z_n(t))=(v_n(t_n(t)),z_n(t_n(t)))$ converges to $(v(t),z(t))$. Hence, the stability (11) follows since we have

$$(t_n(t),v_n(t_n(t)),z_n(t_n(t)))\in\mathbb{S}$$

and the latter is closed.

Upper energy estimate. We can rewrite the inequality (14) in the following way

$$\begin{aligned}
&\mathcal{E}(v_n(t),z_n(t))+\mathcal{F}(\theta(t_n(t)),z_n(t))-\langle L(t_n(t)),(v_n(t),z_n(t))\rangle+\mathrm{Diss}_{\mathcal{D}}(z_n,[0,t_n(t)])\\
&\leq \mathcal{E}(v^0,z^0)+\mathcal{F}(\theta(0),z^0)-\langle L(0),(v^0,z^0)\rangle\\
&+\int_0^{t_n(t)}\int_\Omega \beta'(\theta)\dot\theta|z_n|\,\mathrm{d}x\,\mathrm{d}s-\int_0^{t_n(t)}\langle\dot\ell,v_n\rangle\,\mathrm{d}s.
\end{aligned}$$

We now pass to the lim inf in the latter relation by exploiting the lower semicontinuity of $\mathcal{E}$ and $\mathcal{F}$, the integrability of $\dot\ell$ and of $(\beta\circ\theta)\dot{}$, and the boundedness of v_n from (15)-(16). By the Lebesgue Dominated Convergence Theorem, we deduce that

$$\begin{aligned}
&\mathcal{E}(v(t),z(t))+\mathcal{F}(\theta(t),z(t))-\langle L(t),(v(t),z(t))\rangle+\mathrm{Diss}_{\mathcal{D}}(z,[0,t])\\
&\leq \mathcal{E}(v^0,z^0)+\mathcal{F}(\theta(0),z^0)-\langle L(0),(v^0,z^0)\rangle\\
&+\int_0^{t}\int_\Omega \beta'(\theta)\dot\theta|z|\,\mathrm{d}x\,\mathrm{d}s-\int_0^{t}\langle\dot\ell,v\rangle\,\mathrm{d}s,
\end{aligned}\tag{17}$$

for all $t\in[0,T]$, i.e., the upper energy estimate.

Lower energy estimate. Let us now check the converse inequality with respect to (17). Fix $t\in[0,T]$ and assume to be given a sequence of partitions $\{0=s_0^m< s_1^m<\cdots<s_{M^m-1}^m<s_{M^m}^m=t\}$ such that $\max_{j=1,\dots,M^m}(s_j^m-s_{j-1}^m)\to 0$. We shall let $s_m(s):=s_j^m$ for $s\in(s_{j-1}^m,s_j^m]$, $j=1,\dots,M^m$, $v_m:=v\circ s_m$, and $z_m:=z\circ s_m$. From the stability condition $(v(s_{j-1}^m),z(s_{j-1}^m))\in\mathcal{S}'(s_{j-1}^m,\theta(s_{j-1}^m))$ we have

$$\begin{aligned}
&\mathcal{E}(v(s_{j-1}^m),z(s_{j-1}^m))+\mathcal{F}(\theta(s_{j-1}^m),z(s_{j-1}^m))-\langle L(s_{j-1}^m),(v(s_{j-1}^m),z(s_{j-1}^m)\rangle\\
&\leq \mathcal{E}(v(s_j^m),z(s_j^m))+\mathcal{F}(\theta(s_{j-1}^m),z(s_j^m))-\langle L(s_{j-1}^m),(v(s_j^m),z(s_j^m))\rangle+\mathcal{D}(z_j^m-z_{j-1}^m).
\end{aligned}$$

We now add $\mathcal{F}(\theta(s_j^m),z_j^m)-\mathcal{F}(\theta(s_{j-1}^m),z_j^m)-\langle L(s_j^m)-L(s_{j-1}^m),(v(s_j^m),z(s_j^m))\rangle$ to both sides and rearrange the terms in order to obtain

$$\begin{aligned}
&\mathcal{E}(v(s_j^m),z(s_j^m))+\mathcal{F}(\theta(s_j^m),z(s_j^m))-\langle L(s_j^m),(v(s_j^m),z(s_j^m))\rangle+\mathcal{D}(z(s_{j-1}^m)-z(s_j^m))\\
&\geq \mathcal{E}(v(s_{j-1}^m),z(s_{j-1}^m))+\mathcal{F}(\theta(s_{j-1}^m),z(s_{j-1}^m))-\langle L(s_{j-1}^m),(v(s_{j-1}^m),z(s_{j-1}^m))\rangle\\
&+\mathcal{F}(\theta(s_j^m),z(s_j^m))-\mathcal{F}(\theta(s_{j-1}^m),z(s_j^m))-\langle L(s_j^m)-L(s_{j-1}^m),(v(s_j^m),z(s_j^m))\rangle.
\end{aligned}$$

Summing up for $j = 0, \dots, M^m$ we deduce that

$$\begin{aligned}
&\mathcal{E}(v(t), z(t)) + \mathcal{F}(\theta(t), z(t)) - \langle L(t), (v(t), z(t))\rangle + \mathrm{Diss}_{\mathcal{D}}(z, [0,t]) \\
&\geq \mathcal{E}(v^0, z^0) + \mathcal{F}(\theta(0), z(0)) - \langle L(0), (v^0, z^0)\rangle \\
&+ \sum_{j=1}^{M^m} \int_\Omega \big(\beta(\theta(s_j^m)) - \beta(\theta(s_{j-1}^m))\big)|z(s_j^m)| \,\mathrm{d}x - \int_0^t \langle \dot{\ell}(s), v_m(s)\rangle \,\mathrm{d}s. \qquad (18)
\end{aligned}$$

We can handle the first term in the last line of (18) as follows

$$\sum_{j=1}^{M^m} \int_\Omega \big(\beta(\theta(s_j^m)) - \beta(\theta(s_{j-1}^m))\big)|z(s_j^m)| \,\mathrm{d}x = \int_0^t \int_\Omega \left(⨍_m \frac{\mathrm{d}}{\mathrm{d}t}(\beta\circ\theta)\right) |z \circ s_m| \,\mathrm{d}x \,\mathrm{d}s$$

where we used some obvious notation for the piecewise mean on the partition. As $\mathrm{Diss}_{\mathcal{D}}(z, [0,t]) < \infty$ we have that z is continuous in $L^1(\Omega, \mathbb{R}^{d\times d}_{\mathrm{dev}})$ with the exception of at most a countable number of times. This in particular entails that $z_m \to z$ pointwise almost everywhere in $[0,t]$. Moreover, $\beta\circ\theta \in W^{1,1}(0,T)$ and one has that

$$⨍_m \frac{\mathrm{d}}{\mathrm{d}t}(\beta\circ\theta) \to \beta'(\theta)\dot{\theta} \quad \text{a.e. in } [0,t].$$

Hence, by Dominated Convegence we can conclude that

$$\sum_{j=1}^{M^m} \int_\Omega \big(\beta(\theta(s_j^m)) - \beta(\theta(s_{j-1}^m))\big)|z(s_j^m)| \,\mathrm{d}x \to \int_0^t \int_\Omega \beta'(\theta)\dot{\theta}|z| \,\mathrm{d}x \,\mathrm{d}s.$$

From the fact that $(v(s), z(s)) \in \mathcal{S}'(s, \theta(s))$ for all $s \in [0,t]$ we readily deduce that $v = \mathcal{L}(z, L)$. In particular, v has at most a countable number of discontinuity points in time and $v_m \to v$ pointwise almost everywhere. Eventually, we can pass to the limit into inequality (18) and conclude that

$$\begin{aligned}
&\mathcal{E}(v(t), z(t)) + \mathcal{F}(\theta(t), z(t)) - \langle L(t), (v(t), z(t))\rangle + \mathrm{Diss}_{\mathcal{D}}(z, [0,t]) \\
&\geq \mathcal{E}(v^0, z^0) + \mathcal{F}(\theta(0), z(0)) - \langle L(0), (v^0, z^0)\rangle \\
&+ \int_0^t \int_\Omega \beta'(\theta)\dot{\theta}|z| \,\mathrm{d}x \,\mathrm{d}s - \int_0^t \langle \dot{\ell}, v\rangle \,\mathrm{d}s.
\end{aligned}$$

Absolute continuity of the evolution. Let us now prove that indeed $t \mapsto (v(t), z(t))$ is absolutely continuous in $L^2(\Omega;\mathbb{R}^d)\times H^1(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$. In order to do so, it is notationally convenient to switch back to the original variables $(u,z) = (v + u^{\mathrm{Dir}}, z)$. From the stability (5) at time s and the uniform convexity of $\mathcal{E} + \mathcal{F}$

of constant $\alpha > 0$ we get that

$$\begin{aligned}
&\alpha\|u(t)-u(s)\|^2_{H^1(\Omega;\mathbb{R}^d)} + \alpha\|z(t)-z(s)\|^2_{L^2(\Omega;\mathbb{R}^{d\times d}_{\rm dev})}\\
&\le \mathcal{E}(u(t),z(t)) + \mathcal{F}(\theta(s),z(t)) - \langle \ell(s),u(t)\rangle - \mathcal{E}(u(s),z(s))\\
&- \mathcal{F}(\theta(s),z(s)) + \langle \ell(s),u(s)\rangle + \mathcal{D}(z(t)-z(s))\\
&\le \mathcal{E}(u(t),z(t)) + \mathcal{F}(\theta(t),z(t)) - \langle \ell(t),u(t)\rangle - \mathcal{E}(u(s),z(s)) - \mathcal{F}(\theta(s),z(s))\\
&+ \langle \ell(s),u(s)\rangle + {\rm Diss}_{\mathcal{D}}(z,[s,t]) - \mathcal{F}(\theta(t),z(t)) + \mathcal{F}(\theta(s),z(t)) + \langle \ell(t)-\ell(s),u(t)\rangle\\
&\overset{(6)}{=} \int_s^t\int_\Omega \beta'(\theta)\dot\theta|z|\,{\rm d}x\,{\rm d}s - \int_\Omega \big(\beta(\theta(t)) - \beta(\theta(s))\big)|z(t)|\,{\rm d}x\\
&- \int_s^t \langle \dot\ell,u\rangle\,{\rm d}s + \langle \ell(t)-\ell(s),u(t)\rangle\\
&= \int_s^t\int_\Omega \beta'(\theta(r))\dot\theta(r)\big(|z(r)|-|z(t)|\big)\,{\rm d}x\,{\rm d}r - \int_s^t \langle \dot\ell(r),u(r)-u(t)\rangle\,{\rm d}r\\
&\le \int_s^t\int_\Omega |\beta'(\theta(r))\dot\theta(r)|\,|z(r)-z(t)|\,{\rm d}x\,{\rm d}r\\
&+ \int_s^t \|\dot\ell(r)\|_{(H^1(\Omega;\mathbb{R}^d))'}\|u(r)-u(t)\|_{H^1(\Omega;\mathbb{R}^d)}{\rm d}r.
\end{aligned}$$

Hence, by means of Gronwall's Lemma one checks that

$$\|u(t)-u(s)\|_{H^1(\Omega;\mathbb{R}^d)} + \|z(t)-z(s)\|_{L^2(\Omega;\mathbb{R}^{d\times d}_{\rm dev})} \le c\int_s^t\int_\Omega \big(|\beta'(\theta)\dot\theta|+|\dot\ell|\big){\rm d}x\,{\rm d}s$$

for some suitable constant c depending just on α and T. In particular, the absolute continuity of $t\mapsto(v(t),z(t))$ ensues.

4. **Existence of an optimal control.** We shall finally turn to the proof of Theorem 2.2. Let (u_n,z_n,θ_n) with $(u_n,z_n)\in{\rm Sol}(\theta_n)$ be a minimizing sequence for $\mathcal{J}$, namely

$$\mathcal{J}(u_n,z_n,\theta_n) \to \inf\big\{\mathcal{J}(u,z,\theta)\ :\ (u,z)\in{\rm Sol}(\theta),\ \theta\in\Theta\big\}.$$

Owing to the compactness (8), we can extract a not relabeled subsequence in such a way that both $\theta_n\to\theta$ and $\beta\circ\theta_n\to\beta\circ\theta$ weakly in $W^{1,r}(0,T)$ and uniformly. By exploiting the energy balance (6) we readily get that

$$\sup_{t\in[0,T]}\big(\mathcal{E}(u_n(t),z_n(t))+\mathcal{F}(\theta_n(t),z_n(t))\big)\ \text{ and }\ {\rm Diss}_{\mathcal{D}}(z_n,[0,T])$$
$$\text{are bounded independently of } n.$$

We can hence extract again (still not relabeling) in order to get that $z_n\to z$ pointwise in $BV(\Omega;\mathbb{R}^{d\times d}_{\rm dev})$, $u_n\to u$ pointwise in $H^1(\Omega;\mathbb{R}^d)$ (by linearity), and $(u_n,z_n)\to(u,z)$ weakly-star in $L^\infty(0,T;H^1(\Omega;\mathbb{R}^d)\times L^2(\Omega;\mathbb{R}^{d\times d}_{\rm dev}))$. The proof of Theorem 2.1 can be adapted to the present situation in order to ensure that $(u,z)\in{\rm Sol}(\theta)$. To this aim, the differences arise solely in the treatment of those terms containing θ_n. In particular, the above-mentioned convergences of θ_n and z_n entail directly the convergence

$$\int_\Omega \beta(\theta_n(t))|z_n(t)|\,{\rm d}x \to \int_\Omega \beta(\theta(t))|z(t)|\,{\rm d}x \quad \forall t\in[0,T].$$

Moreover, one can also check for the limit

$$\int_0^t \int_\Omega \beta'(\theta_n)\dot{\theta}_n |z_n| \, \mathrm{d}x \, \mathrm{d}s \to \int_0^t \int_\Omega \beta'(\theta)\dot{\theta}|z| \, \mathrm{d}x \, \mathrm{d}s$$

as we have that $\beta'(\theta_n)\dot{\theta}_n \to \beta'(\theta)\dot{\theta}$ weakly in $L^r(0,T)$ and (by possibly extracting again) $z_n \to z$ strongly in $L^p(\Omega\times[0,T];\mathbb{R}^{d\times d}_{\mathrm{dev}})$ for all $p \in [1,\infty)$ (recall that z_n are uniformly bounded in $\mathbb{R}^{d\times d}_{\mathrm{dev}}$). As we now have that $(u,z) \in \mathrm{Sol}(\theta)$, the assertion follows directly from the lower semicontinuity assumption (9).

4.1. **Some comments on the approximation of optimal controls.** The present existence result represents just a first step in the direction of optimally control the complex thermomechanical behavior of SMA via the Souza-Auricchio model. Indeed, one would also be interested in taking this investigation further into possibly *computing* optimal controls. However, this seems to be a quite delicate task due to the crucially non-smooth nature of the Souza-Auricchio model.

As already mentioned, energetic solutions to the Souza-Auricchio model are not known to be unique. This would require to establish necessary optimality conditions in the frame of set-valued state-to-control mappings. One possible way to restore uniqueness at the state-problem level seems that of regularizing the Souza-Auricchio model by smoothing the inelastic energy density in (2). This approach has been followed by Mielke, Paoli, & Petrov [37, 40] for existence (but note that here the temperature is continuously differentiable with respect to time) and, in the isothermal case, regularizations have been proved to converge to the original non-smooth setting in [5].

Hence, a possible strategy in order to compute optimal controls could be that of considering this regularized situation, optimally controlling the regularized problem, and showing the approximation of the limiting non-smooth case via regularized optimal controls. This kind of argument would quite naturally fall within the general frame introduced by Rindler [50]. However, the specific form of the work of external actions in (10) once again prevents from applying the general results form [50] and some extension of the latter seems rather necessary. Moreover, the possibility of obtaining necessary conditions even at the regularized case seems presently demanding. In this respect, the reader can consider the recent and quite technical analysis by Wachsmuth [56, 57] on the comparably simpler linearized-kinematic-hardening elastoplasticity case.

From the mechanical viewpoint the above-mentioned regularization is rather arguable for it crucially affects the predictive capabilities of the model. Indeed, as an effect of smoothing we loose the possibility of predicting both initiation of the parent-to-product phase transformation under tension and the saturation effect on inelastic evolution. These two are indeed crucial features as they allow at the algorithmic level to efficiently distinguish between purely elastic and inelastic evolution. This distinction turns out to contribute to the remarkable robustness of the Souza-Auricchio model with respect to approximations. We hence prefer to stick here to the original non-smooth situation in order to maximally preserve these distinctive traits of the Souza-Auricchio model.

A second possible strategy for computing optimal control would be that of considering some time or even space-time discretized (approximate) optimal control procedure in the spirit of [50]. This development also appears not at all trivial. On the one hand, one should note that the derivation of necessary optimality conditions at the discrete level appears to be possibly amenable under some regularization only (see above). Moreover, the already mentioned obstruction in applying the available general results due to the specific form of the work of external actions would demand for some significant adaptation of the theory from [41, 49, 50]. We shall hence resort in reconsidering these opportunities elsewhere.

Acknowledgements. Ulisse Stefanelli acknowledges partial support from 7FP-IDEAS-ERC-StG no. 200497 *(BioSMA)*, the CNR-AVČR joint project *SmartMath*, and the Alexander von Humboldt Foundation. The authors gratefully indebted to the Referees for their careful reading of the manuscript and valuable comments.

REFERENCES

[1] F. Auricchio, A. Reali, and U. Stefanelli. A macroscopic 1D model for shape memory alloys including asymmetric behaviors and transformation-dependent elastic properties. *Comput. Methods Appl. Mech. Engrg.*, 198 (2009), 17-20:1631–1637.

[2] F. Auricchio, A.-L. Bessoud, A. Reali, and U. Stefanelli. A phenomenological model for ferromagnetism in shape-memory materials. In preparation, (2011).

[3] F. Auricchio, A.-L. Bessoud, A. Reali, and U. Stefanelli. A three-dimensional phenomenological models for magnetic shape memory alloys. *GAMM-Mitt.*, 34 (2011), 1:90–96.

[4] F. Auricchio and J. Lubliner. A uniaxial model for shape-memory alloys. *Internat. J. Solids Structures*, 34 (1997), 3601–3618.

[5] F. Auricchio, A. Mielke, and U. Stefanelli. A rate-independent model for the isothermal quasi-static evolution of shape-memory materials. *Math. Models Meth. Appl. Sci.*, 18 (2008), 1:125–164.

[6] F. Auricchio and L. Petrini. Improvements and algorithmical considerations on a recent three-dimensional model describing stress-induced solid phase transformations. *Internat. J. Numer. Methods Engrg.*, 55 (2002), 11:1255–1284,

[7] F. Auricchio and L. Petrini. A three-dimensional model describing stress-temperature induced solid phase transformations. Part I: solution algorithm and boundary value problems. *Internat. J. Numer. Meth. Engrg.*, 61 (2004), 807–836.

[8] F. Auricchio and L. Petrini. A three-dimensional model describing stress-temperature induced solid phase transformations. Part II: thermomechanical coupling and hybrid composite applications. *Internat. J. Numer. Meth. Engrg.*, 61 (2004), 716–737.

[9] F. Auricchio, A. Reali, and U. Stefanelli. A phenomenological 3D model describing stress-induced solid phase transformations with permanent inelasticity. In *Topics on mathematics for smart systems*, pages 1–14. World Sci. Publ., Hackensack, NJ, 2007.

[10] F. Auricchio, A. Reali, and U. Stefanelli. A three-dimensional model describing stress-induces solid phase transformation with residual plasticity. *Int. J. Plasticity*, 23 (2007), 2:207–226.

[11] A.-L. Bessoud and U. Stefanelli. Magnetic shape memory alloys: three-dimensional modeling and analysis. *Math. Models Meth. Appl. Sci.*, to appear, 2011.

[12] A.-L. Bessoud, M. Kružík, and U. Stefanelli. A macroscopic model for magnetic shape memory alloys. *Preprint IMATI-CNR*, 23PV10/21/0, 2010.

[13] H. Brézis. *Operateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, Math Studies, Vol.5, North-Holland, Amsterdam/New York, 1973.

[14] M. Brokate and J. Sprekels. Optimal control of shape memory alloys with solid-solid phase transitions. In *Emerging applications in free boundary problems (Montreal, 1990)*, volume 280 of *Pitman Res. Notes Math. Ser.*, pages 208–214. Longman Sci. Tech., Harlow, 1993.

[15] N. Bubner, J. Sokołowski, and J. Sprekels. Optimal boundary control problems for shape memory alloys under state constraints for stress and temperature. *Numer. Funct. Anal. Optim.*, 19 (1998), 5-6:489–498.

[16] T.W. Duerig and A.R. Pelton editors. *SMST-2003 Proceedings of the International Conference on Shape Memory and Superelastic Technology Conference*, ASM International, 2003.

[17] V. Evangelista, S. Marfia, and E. Sacco. Phenomenological 3D and 1D consistent models for shape-memory alloy materials. *Comput. Mech.*, 44 (2009), 3:405–421.

[18] V. Evangelista, S. Marfia, and E. Sacco. A 3D SMA constitutive model in the framework of finite strain. *Internat. J. Numer. Methods Engrg.*, 81 (2010), 6:761–785.

[19] F. Falk. Model free energy, mechanics and thermodynamics of shape memory alloys. *Acta Metallurgica*, 28 (1990), 1773–1780.

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

[21] M. Frémond. Matériaux à mémoire de forme. *C. R. Acad. Sci. Paris Sér. II Méc. Phys. Chim. Sci. Univers Sci. Terre*, 304 (1987), 239–244.

[22] M. Frémond, *Non-smooth Thermomechanics*, Springer-Verlag, Berlin, 2002.

[23] M. Frémond and S. Miyazaki, *Shape Memory Alloys*, CISM Courses and Lectures, vol. 351, Springer-Verlag, 1996.

[24] S. Frigeri and U. Stefanelli. Existence and time-discretization for the finite-strain Souza-Auricchio constitutive model for shape-memory alloys. Contin. Mech. Thermodyn., to appear, 2012.

[25] S. Govindjee and C. Miehe. A multi-variant martensitic phase transformation model: formulation and numerical implementation. *Comput. Methods Appl. Mech. Engrg.*, 191 (2001), 215–238.

[26] D. Helm and P. Haupt. Shape memory behaviour: modelling within continuum thermomechanics. *Intern. J. Solids Struct.*, 40 (2003), 827–849.

[27] K.-H. Hoffmann and D. Tiba. Control of a plate with nonlinear shape memory alloy reinforcements. *Adv. Math. Sci. Appl.*, 7 (1997), 1:427–436.

[28] K.-H. Hoffmann and A. Żochowski. Control of the thermoelastic model of a plate activated by shape memory alloy reinforcements. *Math. Methods Appl. Sci.*, 21 (1998), 7:589–603.

[29] P. Krejčí and U. Stefanelli. Existence and nonexistence for the full thermomechanical Souza-Auricchio model of shape memory wires. *Math. Mech. Solids*, 16 (2011), 4:349–365.

[30] P. Krejčí and U. Stefanelli. Well-posedness of a thermo-mechanical model for shape memory alloys under tension. *M2AN Math. Model. Numer. Anal.*, 44 (2010), 6:1239–1253.

[31] D.C. Lagoudas, P.B. Entchev, P. Popov, E. Patoor, L.C. Brinson, and X. Gao. Shape memory alloys, Part II: Modeling of polycrystals. *Mech. Materials*, 38 (2006), 391–429.

[32] P. Popov and D.C. Lagoudas. A 3-D constitutive model for shape memory alloys incorporating pseudoelasticity and detwinning of self-accommodated martensite. *Int. J. Plasticity*, 23 (2007), 1679–1720.

[33] V.I. Levitas. Thermomechanical theory of martensitic phase transformations in inelastic materials. *Intern. J. Solids Struct.*, 35 (1998), 889–940.

[34] M. Eleuteri, L. Lussardi, and U. Stefanelli. A rate-independent model for permanent inelastic effects in shape memory materials. *Netw. Heterog. Media*, 6 (2011), 1:145-165.

[35] A. Mainik and A. Mielke. Existence results for energetic models for rate-independent systems. *Calc. Var. Partial Differential Equations*, 22(1):73–99, 2005.

[36] A. Mielke. Evolution of rate-independent systems. In C. Dafermos and E. Feireisl, editors, *Handbook of Differential Equations, evolutionary equations*, volume 2, pages 461–559. Elsevier, 2005.

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

[38] A. Mielke, L. Paoli, A. Petrov, and U. Stefanelli. Error estimates for space-time discretizations of a rate-independent variational inequality. *SIAM J. Numer. Anal.*, 48 (2010), 5: 1625–1646.

[39] A. Mielke, L. Paoli, A. Petrov, and U. Stefanelli. Error bounds for space-time discretizations of a 3d model for shape-memory materials. In K. Hackl, editor, *IUTAM Symposium on Variational Concepts with Applications to the Mechanics of Materials*, pages 185–197. Springer, 2010. Proceedings of the IUTAM Symposium on Variational Concepts, Bochum, Germany, Sept. 22–26, 2008.

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

[41] A. Mielke and F. Rindler. Reverse approximation of energetic solutions to rate-independent processes. *NoDEA Nonlinear Differential Equations Appl.*, 16 (2009) 1:17–40.

[42] A. Mielke and F. Theil. On rate-independent hysteresis models. *NoDEA, Nonlinear Diff. Equations Applications*, 11 (2004), 151–189.

[43] B. Peultier, T. Ben Zineb, and E. Patoor. Macroscopic constitutive law for SMA: Application to structure analysis by FEM. *Materials Sci. Engrg.: A*, 438-440 (2006), 454–458.

[44] L. Paoli and A. Petrov. Global existence result for phase transformations with heat transfer in shape memory alloys. *Preprint WIAS*, 1608 (2011).

[45] I. Pawłow and A. Żochowski. A control problem for a thermoelastic system in shape memory materials. *Sūrikaisekikenkyūsho Kōkyūroku*, 1210 (2011), 8–23, Free boundary problems (Japanese) (Kyoto, 2000).

[46] R. Peyroux, A. Chrysochoos, Ch. Licht, and M. Löbel. Phenomenological constitutive equations for numerical simulations of SMA's structures. Effect of thermomechanical couplings. *J. Phys. C4 Suppl.*, 6 (1996), 347–356.

[47] B. Raniecki and Ch. Lexcellent. R_L models of pseudoelasticity and their specification for some shape-memory solids. *Eur. J. Mech. A Solids*, 13 (1994), 21–50.

[48] S. Reese and D. Christ. Finite deformation pseudo-elasticity of shape memory alloys – Constitutive modelling and finite element implementation. *Int. J. Plasticity*, 24 (2008), 3: 455-482.

[49] F. Rindler. Optimal control for nonconvex rate-independent evolution processes. *SIAM J. Control Optim.*, 47 (2008), 6:2773–2794.

[50] F. Rindler. Approximation of rate-independent optimal control problems. *SIAM J. Numer. Anal.*, 47 (2009), 5:3884–3909.

[51] T.Roubíček. Models of microstructure evolution in shape memory alloys, in *Nonlinear Homogenization and its Appl.to Composites, Polycrystals and Smart Materials*, P. Ponte Castaneda, J.J. Telega, B. Gambin eds., NATO Sci. Series II/170, Kluwer, Dordrecht, 2004, pp.269–304.

[52] J. Sokołowski and J. Sprekels. Control problems with state constraints for shape memory alloys. *Math. Methods Appl. Sci.*, 17 (1994), 12:943–952.

[53] A. C. Souza, E. N. Mamiya, and N. Zouain. Three-dimensional model for solids undergoing stress-induced tranformations. *Eur. J. Mech. A Solids*, 17 (1998), 789–806.

[54] U. Stefanelli. Magnetic control of magnetic shape-memory single crystals. *Phys. B*, to appear (2011).

[55] P. Thamburaja and L. Anand. Polycrystalline shape-memory materials: effect of crystallographic texture. *J. Mech. Phys. Solids*, 49 (2001), 709-737.

[56] G. Wachsmuth. Optimal Control of Quasistatic Plasticity with Linear Kinematic Hardening, Part I: Existence and Discretization in Time. Preprint SPP1253-118, 2011.

[57] G. Wachsmuth. Optimal Control of Quasistatic Plasticity with Linear Kinematic Hardening, Part II: Regularization and Optimality Conditions. Preprint SPP1253-119, 2011.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: michela.eleuteri@univr.it
E-mail address: l.lussardi@dmf.unicatt.it
E-mail address: ulisse.stefanelli@imati.cnr.it