ERROR ESTIMATES FOR SPACE-TIME DISCRETIZATIONS OF A RATE-INDEPENDENT VARIATIONAL INEQUALITY

ALEXANDER MIELKE*, LAETITIA PAOLI†, ADRIEN PETROV‡, AND ULISSE STEFANELLI§

Abstract. This paper deals with error estimates for space-time discretizations in the context of evolutionary variational inequalities of rate-independent type. After introducing a general abstract evolution problem, we address a fully-discrete approximation and provide a priori error estimates. The application of the abstract theory to a semilinear case is detailed. In particular, we provide explicit space-time convergence rates for the isothermal Souza-Auricchio model for shape-memory alloys.

Key words. Shape-memory materials, evolutionary variational inequalities, rate-independent processes, space-time discretization, error estimates.

AMS subject classifications. 49J40, 74C05, 74N30

1. Introduction.

The present analysis is concerned with error estimates for space-time discretizations in the context evolutionary variational inequalities of rate-independent type. More precisely, let $\mathcal{Q}$ be a Hilbert space, $\mathcal{E} : [0,T] \times \mathcal{Q} \to \mathbb{R}$ with $T > 0$ and $\Psi : \mathcal{Q} \to [0,\infty)$ be the energy and dissipation functionals, respectively. We assume that $\mathcal{E}(t,\cdot)$ and Ψ are continuous and convex. Moreover, as is common in modeling hysteresis effect in mechanics, we assume that the system is rate-independent which amounts in asking that Ψ is positively homogeneous of degree 1, i.e., $\Psi(\gamma v) = \gamma\Psi(v)$ for all $\gamma \geq 0$.

The aim of this work is to show that the solutions $q : [0,T] \to \mathcal{Q}$ of the non-smooth differential inclusion

$$0 \in \partial\Psi(\dot{q}(t)) + \mathrm{D}_q\mathcal{E}(t,q(t)) \quad \text{a.e. in} \quad (0,T) \tag{1.1}$$

can be well-approximated by spatially discretized time-incremental minimization problems. The difficulty here is the non-smoothness of the subdifferential operator $\partial\Psi(\cdot)$ as well as the nonlinearity of the map $q \mapsto \mathrm{D}_q\mathcal{E}(t,q)$. In the linear case this would reduce to classical evolutionary variational inequalities for which the numerics is well studied, see e.g. [HaR99, ACZ99, AlC00, COV06, CKO06, Car99, LiB96, LiB97, OrP04].

In particular, we are here specifically interested in a semi-linear case where the potential energy has the following form

$$\forall \widehat{q} \in \mathcal{Q}: \; \mathcal{E}(t,\widehat{q}) \stackrel{\text{def}}{=} \frac{1}{2}\langle \mathbf{A}\widehat{q}, \widehat{q}\rangle_{\mathcal{Q}} + \mathcal{H}(\widehat{q}) - \langle \boldsymbol{\ell}(t), \widehat{q}\rangle_{\mathcal{Q}}. \tag{1.2}$$

Here $\mathbf{A}$ is a symmetric positive definite operator, $\mathcal{H}$ is a differentiable and convex functional and $\boldsymbol{\ell} \in \mathrm{C}^1([0,T],\mathcal{Q}')$ is the external loading. This setting is closely related to the isothermal Souza-Auricchio model for shape-memory alloys (SMA). The latter are

* Institut für Mathematik, Humboldt-Universität zu Berlin, Rudower Chaussee 25, 12489 Berlin, Germany (mielke@wias-berlin.de) §

†LaMUSE (Laboratoire de Mathématiques de l'Université de Saint-Etienne), 23 rue Paul Michelon, 42023 Saint-Etienne Cedex 02, France (laetitia.paoli@univ-st-etienne.fr)

‡Weierstraß-Institut für Angewandte Analysis und Stochastik, Mohrenstraße 39, 10117 Berlin, Germany (petrov@wias-berlin.de)

§Istituto di Matematica Applicata e Tecnologie Informatiche - CNR, via Ferrata 1, 27100 Pavia, Italy (ulisse.stefanelli@imati.cnr.it)

metallic alloys showing some surprising thermo-mechanical behavior, namely, strongly deformed specimens regain their original shape after a thermal cycle (*shape-memory effect*). Moreover, within some specific (suitably high) temperature range, they are *super-elastic*, meaning that they fully recover comparably large deformations. These features are not present (at least to this extent) in most materials traditionally used in Engineering and, thus, are at the basis of innovative and commercially valuable applications. Nowadays, shape-memory alloys are successfully used in many applications among which biomedical devices (vascular stents, archwires, endo-guidewires) and MEMS (actuators, valves, mini-grippers and positioners). The Souza-Auricchio model here considered is a phenomenological, small-deformation model describing both the shape memory and the superelastic effect (although in the present isothermal reduction no shape memory effect is actually reproduced). The reader is referred to [SMZ98, AuS01, AuP04, ARS09] for the derivation and the mechanics and [AuS04, AuS05, MiP07, AMS08] for the mathematical analysis.

The paper is organized as follows. After introducing more precisely in Section 2 our assumptions, we recall a well-posedness result from [MiT04]. Then an error estimate for space-time discretizations is derived. To do so, we choose a sequence of partitions $\{ 0 = t_0^\tau < t_1^\tau < \cdots < t_{k^\tau}^\tau = T \}$ of the time interval $[0, T]$ with $\max\{ t_k^\tau - t_{k-1}^\tau \ : \ k = 1, ..., k^\tau \} \leq \tau$ and a sequence $(\mathcal{Q}_h)_{h>0}$ of finite-dimensional spaces exhausting $\mathcal{Q}$. Then, the space-time discretized incremental minimization problem

$$q_{\tau,h}^k \overset{\text{def}}{=} \operatorname{Argmin}\left\{ \mathcal{E}(t_k^\tau, \widehat{q}_h) + \Psi(\widehat{q}_h - q_{\tau,h}^{k-1}) \,\middle|\, \widehat{q}_h \in \mathcal{Q}_h \right\}$$

has a unique solution by uniform convexity. Thus, it is possible to define the piecewise affine interpolants $q_{\tau,h} : [0, T] \to \mathcal{Q}_h$.

Our error estimates rely on an abstract *approximation condition*. We refer to (2.10) for its most general version and give here, for brevity, a slightly strengthened form:

$$\begin{aligned} &\exists C > 0 \ \forall h \in (0, 1] \ \forall (t, q_h, w) \in [0, T] \times \mathcal{Q}_h \times \mathcal{Q} \ \exists v_h \in \mathcal{Q}_h : \\ &\langle \mathrm{D}_q \mathcal{E}(t, q_h), v_h - w \rangle_{\mathcal{Q}} + \Psi(v_h - w) \leq C h^\beta \big(1 + \|q_h\|_{\mathcal{Q}}^2\big) \|w\|_{\mathcal{Q}}. \end{aligned} \tag{1.3}$$

Under suitable additional assumption we construct a constant C such that

$$\|q_{\tau,h}(t) - q(t)\|_{\mathcal{Q}} \leq C\big(h^{\beta/2} + \sqrt{\tau} + \|q_{\tau,h}(0) - q(0)\|_{\mathcal{Q}}\big). \tag{1.4}$$

In Section 3 we show that condition (1.3) can be established by assuming that $\mathcal{H}$ and Ψ are lower order, if compared with $\mathbf{A}$. This means there exists a bigger space $\mathcal{X}$ with $\mathcal{Q} \subset \mathcal{X}$ and $\mathcal{X}' \subset \mathcal{Q}'$ such that $\Psi : \mathcal{X} \to [0, \infty)$ is continuous and that $\mathrm{D}_q \mathcal{H} \in \mathrm{C}^{1,\mathrm{Lip}}(\mathcal{Q}, \mathcal{X}')$. The power β then relates to an interpolation estimate. Moreover, for any suitable initial condition $q(0)$, we can find $q_h(0)$ such that $\|q_h(0) - q(0)\|_{\mathcal{Q}} = \mathcal{O}(h^{\beta/2})$, which provides the desired convergence of space-time discretizations. We emphasize that our convergence rates are obtained without any further assumptions on the smoothness of the solutions to be approximated. This is particularly remarkable in connection with linearized elastoplasticity. Indeed, up to now, convergence rates for linearized elastoplasticity have been obtained in [AlC00] (classical theory) and [DE*07] (strain-gradient theory), by assuming higher smoothness-in-time on the solutions. Here instead the convergence analysis follows under natural regularity conditions. Note however that our overall assumptions will correspond to the occurrence of gradient terms and, in particular, classical linearized elastoplasticity cannot be directly accommodated in our setting.

Eventually, we show in Section 4 that the abstract result obtained for semi-linear problems remains also valid for the isothermal Souza-Auricchio model. Related convergence results for models of phase transformations in shape-memory alloys were obtained in [KMR05, MiR06, MPP08], however, there no convergence rates were obtained. In fact, for the relevant models the uniqueness of solutions is not known and hence, only convergence of suitable subsequences has been established.

2. An abstract approximation result. We consider a Hilbert space $\mathcal{Q}$ with dual $\mathcal{Q}'$. The norm of $\mathcal{Q}$ and the duality product between $\mathcal{Q}'$ and $\mathcal{Q}$ are denoted by $\|\cdot\|_{\mathcal{Q}}$ and $\langle\cdot,\cdot\rangle_{\mathcal{Q}}$, respectively. For some reference time $T > 0$ we are given an energy functional $\mathcal{E} : [0,T] \times \mathcal{Q} \to \mathbb{R}$ and a dissipation potential $\Psi : \mathcal{Q} \to [0,\infty)$. We assume that Ψ is positively homogeneous of degree 1, which makes the system rate-independent. Moreover, Ψ will be assumed to be bounded on bounded sets and to satisfy the triangle inequality. Hence, we have that

$$\forall \gamma > 0\ \forall q \in \mathcal{Q} :\ \Psi(\gamma q) = \gamma \Psi(q), \tag{2.1a}$$

$$\exists c^{\Psi} > 0\ \forall q \in \mathcal{Q} :\ \Psi(q) \leq c^{\Psi} \|q\|_{\mathcal{Q}}, \tag{2.1b}$$

$$\forall q_1, q_2 \in \mathcal{Q} :\ \Psi(q_1 + q_2) \leq \Psi(q_1) + \Psi(q_2). \tag{2.1c}$$

Notice that (2.1a) and (2.1c) imply that Ψ is convex.

In this abstract section we pose quite general conditions on $\mathcal{E}$ that will be specified to the semilinear case in the following section. Finally, in Section 4, we will show that these conditions are satisfied for the Souza-Auricchio model for phase transformations in SMA, see [MiP07, AMS08]. To simplify the presentation we give slightly stronger conditions than those that are really needed. We use the convention that a function $f \in C^k(\mathcal{Q}, Y)$ is k times Fréchet differentiable such that the kth derivative is still continuous and bounded on bounded sets. We let

$$\mathcal{E} \in \mathrm{C}^3([0,T] \times \mathcal{Q}, \mathbb{R}), \tag{2.2a}$$

$$\exists \kappa > 0\ :\ \mathcal{E}(t,\cdot) \text{ is } \kappa\text{-uniformly convex, i.e, } \mathrm{D}_q^2 \mathcal{E}(t,q) \geq \kappa I, \tag{2.2b}$$

We consider the following doubly nonlinear evolution equation

$$0 \in \partial\Psi(\dot{q}(t)) + \mathrm{D}_q \mathcal{E}(t, q(t)) \text{ a.e. in } (0,T). \tag{2.3}$$

As usual, $\dot{(\ )}$ denotes the time derivative $\frac{\mathrm{d}}{\mathrm{d}t}$. We say that q is a solution of the rate-independent system $(\mathcal{Q}, \mathcal{E}, \Psi)$ if $q \in W^{1,1}([0,T], \mathcal{Q})$ and (2.3) holds. We say that q solves the initial-value problem $(\mathcal{Q}, \mathcal{E}, \Psi, q^0)$ if additionally $q(0) = q^0$ holds.

Using the definition of the subdifferential $\partial\Psi(\dot{q})$, relation (2.3) turns out to be equivalent to the *variational inequality*

$$\forall v \in \mathcal{Q} :\ \langle \mathrm{D}_q \mathcal{E}(t, q(t)), v - \dot{q}(t) \rangle_{\mathcal{Q}} + \Psi(v) - \Psi(\dot{q}(t)) \geq 0. \tag{2.4}$$

We define the *set of stable states* at time t via

$$\mathcal{S}(t) \stackrel{\text{def}}{=} \big\{ q \in \mathcal{Q} \,\big|\, \forall \widehat{q} \in \mathcal{Q} :\ \mathcal{E}(t,q) \leq \mathcal{E}(t, \widehat{q}) + \Psi(\widehat{q} - q) \big\}. \tag{2.5}$$

Since 1-homogeneity of Ψ implies $\partial\Psi(\dot{q}) \subset \partial\Psi(0)$ we see that (2.3) implies $q(t) \in \mathcal{S}(t)$ a.e. in $(0,T)$. This can be seen as a static stability condition, which has to hold for all $t \in [0,T]$ by continuity of $\mathrm{D}_q\mathcal{E}$ and the closedness of $\partial\Psi(0)$, entailing the natural restriction $q^0 \in \mathcal{S}(0)$ for the initial datum. The following results provide useful a priori estimtates.

Proposition 2.1. *Assume that* (2.1) *and* (2.2) *hold.*

(a) *Then, for all* $t \in [0,T]$ *we have*

$$q \in \mathcal{S}(t) \quad \Longleftrightarrow \quad -\mathrm{D}_q\mathcal{E}(t,q) \in \partial\Psi(0). \tag{2.6}$$

(b) *There is a constant* $C_0^R > 0$ *such that*

$$q \in \mathcal{S}(t) \implies \|q\|_{\mathcal{Q}} \le C_0^R, \ \|\mathrm{D}_q\mathcal{E}(t,q)\|_{\mathcal{Q}'} \le c^\Psi \quad \text{and} \tag{2.7a}$$

$$\forall \widehat{q} \in \mathcal{Q}: \ \mathcal{E}(t,q) + \frac{\kappa}{2}\|\widehat{q}-q\|_{\mathcal{Q}}^2 \le \mathcal{E}(t,\widehat{q}) + \Psi(\widehat{q}-q). \tag{2.7b}$$

(c) *If* $(t,q^0) \in [0,T]\times\mathcal{Q}$ *and* q_* *minimizes* $q \mapsto \mathcal{E}(t,q)+\Psi(q-q^0)$*, then* $q_* \in \mathcal{S}(t)$*.*

Proof. Part (a) follows from the very definition of subdifferential, for more details, the reader is referred to [MiT04] . Moreover, (2.7b) is an immediate consequence of the fact that $q \in \mathcal{S}(t)$ is the unique minimizer of the functional $\widehat{q} \mapsto \mathcal{E}(t,\widehat{q}) + \Psi(\widehat{q}-q)$, which is still κ-uniformly convex (cf. [MPP08, Theorem 4.1]).

To establish (2.7a) we first observe that $\eta \in \partial\Psi(0)$ implies $\|\eta\|_{\mathcal{Q}'} \le c^\Psi$ because of (2.1b). Now let $\Lambda = \sup_{t\in[0,T]} \|\mathrm{D}_q\mathcal{E}(t,0)\|_{\mathcal{Q}'}$ and estimate

$$\begin{aligned}\kappa\|q\|_{\mathcal{Q}}^2 = \kappa\|q-0\|_{\mathcal{Q}}^2 &\le \langle \mathrm{D}_q\mathcal{E}(t,q)-\mathrm{D}_q\mathcal{E}(t,0), q-0\rangle \\ &\le \big(\|\mathrm{D}_q\mathcal{E}(t,q)\|_{\mathcal{Q}'}+\|\mathrm{D}_q\mathcal{E}(t,0)\|_{\mathcal{Q}'}\big)\|q\|_{\mathcal{Q}} \le \big(c^\Psi+\Lambda\big)\|q\|_{\mathcal{Q}},\end{aligned}$$

which implies that (2.7a) holds with $C_0^R = \big(c^\Psi+\Lambda\big)/\kappa$. This proves Part (b).

Part (c) follows easliy form Part (a), since the minimizer satisfies $-\mathrm{D}_q\mathcal{E}(t,q_*) \in \partial\Psi(q_*-q^0) \subset \partial\Psi(0)$. ☐

We treat now the question of the error estimate of space-time discretizations. Let us choose a set of parameters $h \in (0,1]$ (mesh sizes) having in mind the limit $h \to 0$ and let $\mathcal{Q}_h$ be closed subspaces of $\mathcal{Q}$. Typically, $\mathcal{Q}_h$ is a finite-dimensional subspace of $\mathcal{Q}$, like a finite-element space. By convention, let $\mathcal{Q}_0 \overset{\text{def}}{=} \mathcal{Q}$ to include the full case via $h = 0$.

It is convenient to introduce the *set of stable states* $\mathcal{S}_h(t)$ for any $t \in [0,T]$ by simply replacing $\mathcal{Q}$ by $\mathcal{Q}_h$ in (2.5).

We recall now that for all $h \in [0,1]$ the rate-independent variational inequality (2.4) restricted to $\mathcal{Q}_h$ admits a unique solution $q_h : [0,T] \to \mathcal{Q}_h$ for any given stable initial data q_h^0, i.e, $q_h^0 \in \mathcal{S}_h(0)$. This existence theory has been developed in [MiT04] and is based on the construction of a sequence of incremental minimization problems. The theory avoids any compactness arguments and uses smoothness to obtain strong convergence. More precisely, we consider a second approximation parameter $\tau \in (0,T]$ (time step) and a partition $\Pi^\tau = \{0 = t_0^\tau < t_1^\tau < \ldots < t_{k^\tau}^\tau = T\}$ with

$$t_k^\tau - t_{k-1}^\tau \le \tau \quad \text{for} \quad k = 1,\ldots,k^\tau.$$

We let $q_{\tau,h}^0 \overset{\text{def}}{=} q_h^0$ and we consider the following incremental problems:

$$(\mathrm{IP})^{\tau,h} \quad \begin{cases} \text{for } k = 1,\ldots,k^\tau \text{ find} \\ q_{\tau,h}^k \in \operatorname{Argmin}\big\{ \mathcal{E}(t_k^\tau,\widehat{q}_h)+\Psi(\widehat{q}_h-q_{\tau,h}^{k-1}) \,\big|\, \widehat{q}_h \in \mathcal{Q}_h \big\}. \end{cases}$$

By uniform convexity and continuity, which implies weak lower semicontinuity, the solutions $q_{\tau,h}^k$ exist and are uniquely determined. We define an approximate solution $q_{\tau,h} : [0,T] \to \mathcal{Q}_h$ as the piecewise affine interpolants given by

$$q_{\tau,h}(t) \overset{\text{def}}{=} \tfrac{t_k^\tau - t}{t_k^\tau - t_{k-1}^\tau}\, q_{\tau,h}^{k-1} + \tfrac{t - t_{k-1}^\tau}{t_k^\tau - t_{k-1}^\tau}\, q_{\tau,h}^k \ \text{for } t \in [t_{k-1}^\tau, t_k^\tau], \ k = 1,\ldots,k^\tau, \tag{2.8}$$

where $q^k_{\tau,h}$ solves $(\mathrm{IP})^{\tau,h}$.

Then, for each fixed $h \in [0,1]$, we show that a subsequence of $q_{\tau,h}$ has a limit as τ tends to 0 and this limit function $q_h : [0,T] \to \mathcal{Q}_h$ satisfies (2.4), where $\mathcal{Q}$ is replaced by $\mathcal{Q}_h$.

In rate-independent problems uniqueness results and Lipschitz-continuous dependence on the initial data are rather exceptional, as usually strong assumptions on the nonlinearities are needed, see [MiT04, MiR07]. In the present case these assumptions hold and we are able to conclude for the convergence of the whole sequence $q_{\tau,h}$ to the unique solution of $(\mathcal{Q}_h, \mathcal{E}, \Psi, q^0_h)$. Let us summarize this discussion in the following statement, which is a slight generalization of Theorem 7.1 in [MiT04], in particular since we state uniformity in $h \geq 0$.

THEOREM 2.2. *Assume* (2.1) *and* (2.2). *Then, for all $h \in [0,1]$ and all $q^0_h \in \mathcal{S}_h(0)$, there exists a unique solution $q_h \in \mathrm{C}^{\mathrm{Lip}}([0,T], \mathcal{Q}_h)$ of the initial-value problem $(\mathcal{Q}, \mathcal{E}, \Psi, q^0_h)$. Moreover, there exist positive constants C^R_0, C^R_1 and $\bar{C}$ such that, for all $h \in [0,1]$ and all partitions Π^τ, we have*

$$\|q_{\tau,h}(t)\|_{\mathcal{Q}} \leq C^R_0,\ \|q_h(t)\|_{\mathcal{Q}} \leq C^R_0 \text{ for all } t \in [0,T]; \tag{2.9a}$$

$$\|\dot{q}_{\tau,h}(t)\|_{\mathcal{Q}} \leq C^R_1,\ \|\dot{q}_h(t)\|_{\mathcal{Q}} \leq C^R_1 \text{ for a.a. } t \in [0,T]; \tag{2.9b}$$

$$\|q_{\tau,h}(t) - q_h(t)\|_{\mathcal{Q}} \leq \bar{C}\sqrt{\tau} \text{ for all } t \in [0,T]. \tag{2.9c}$$

The important fact is that estimate (2.9c) for the time approximation is uniform in h. The reader is referred to the Appendix for the detailed proof of (2.9c) which is a crucial ingredient to obtain the error estimate of space-time discretizations. Condition (2.9a) follows from Propostion 2.1 by combining parts (b) and (c). Concerning (2.9b), we leave the verification to the reader since it suffices to follow the ideas developed in [MiT04].

We are now addressing the question of the limit $h \to 0$. For this, we have to impose suitable conditions that allow us to approximate elements in $\mathcal{Q}$ via elements of $\mathcal{Q}_h$. Again we will use smoothness and uniform convexity in the spirit of Section 7.2 in [MiT04]. The *approximation condition* for our error bounds involves additional symmetric operators $\mathbf{B}_h \in \mathrm{Lin}(\mathcal{Q}, \mathcal{Q}')$ and reads as follows:

$$\exists C^{\mathbf{A}}, C^{\mathbf{B}} > 0\ \forall h \in (0,1]\ \forall t \in [0,T],\ q_h \in \mathcal{S}_h(t),\ q \in \mathcal{S}(t),\ w \in \mathcal{Q}\ \exists v_h \in \mathcal{Q}_h :$$

$$|\langle \mathbf{B}_h q, q\rangle_{\mathcal{Q}}| \leq C^{\mathbf{B}} h^\beta, \tag{2.10a}$$

$$\langle \mathrm{D}_q\mathcal{E}(t,q_h), v_h - w\rangle_{\mathcal{Q}} + \Psi(v_h - w) \leq C^{\mathbf{A}}\big(h^\beta + \|q_h - q\|^2_{\mathcal{Q}}\big)\|w\|_{\mathcal{Q}} + \langle \mathbf{B}_h q, w\rangle_{\mathcal{Q}}. \tag{2.10b}$$

This condition is formulated in such a way that we still see the interplay between the potential forces $\mathrm{D}_q\mathcal{E}(t,q)$ and the dissipation Ψ, because of the definition of $\mathcal{S}$ and $\mathcal{S}_h$. Moreover, the stability sets are usually much smaller to turn (2.10) into weaker statements than those obtained by considering q_h and q in large balls. Clearly, condition (1.3) implies (2.10) with $\mathbf{B}_h \equiv 0$.

THEOREM 2.3. *Assume that $\mathcal{Q}$, $\mathcal{Q}_h$, $\mathcal{E}$ and Ψ satisfy* (2.1), (2.2), *and that* (2.10) *hold. Then, there exists a constant $C_* > 0$ such that, for any $h \in (0,1]$, $q^0_h \in \mathcal{S}_h(0)$, any partition Π^τ, and any $q^0 \in \mathcal{S}(0)$, the unique solution q of the initial-value problem $(\mathcal{Q}, \mathcal{E}, \Psi, q^0)$ satisfies the estimate*

$$\|q_{\tau,h}(t) - q(t)\|_{\mathcal{Q}} \leq C_*\big(h^{\beta/2} + \sqrt{\tau} + \|q^0_h - q^0\|_{\mathcal{Q}}\big) \text{ for all } t \in [0,T], \tag{2.11}$$

where $q_{\tau,h} : [0,T] \to \mathcal{Q}_h$ is defined via (2.8) *with $q^0_{\tau,h} = q^0_h$.*

There are two possible strategies to establish the desired result. For each fixed $h \in (0,1]$ we may discretize in time and show that the error between the time-discrete $q_{\tau,h}$ and time-continuous solutions q_h can be controlled by $\sqrt{\tau}$, uniformly in h. Then, we can use variational inequalities on the time-continuous level to estimate $\|q_h(t)-q(t)\|^2_{\mathcal{Q}}$. This is the approach of the proof given below. Another alternative would be to consider a fixed time-discretization and to estimate $\|q^k_{\tau,h}-q^k_\tau\|^2_{\mathcal{Q}}$ uniformly with respect to τ and $k=1,\dots,k_\tau$ (cf. [AMS08]).

In the following, the notations for the constants introduced in the proofs are valid only in the proof.

Proof. Since the first term in the right-hand side of

$$\|q_{\tau,h}(t)-q(t)\|_{\mathcal{Q}} \le \|q_{\tau,h}(t)-q_h(t)\|_{\mathcal{Q}} + \|q_h(t)-q(t)\|_{\mathcal{Q}}, \tag{2.12}$$

is already estimated in (2.9c) it remains to estimate the second one. Since q_h solves $(\mathcal{Q}_h,\mathcal{E},\Psi,q^0_h)$ and q solves $(\mathcal{Q},\mathcal{E},\Psi,q^0)$ we have the two variational inequalities

$$\forall v_h \in \mathcal{Q}_h : \langle \mathrm{D}_q\mathcal{E}(t,q_h(t)), v_h-\dot q_h(t)\rangle_{\mathcal{Q}} + \Psi(v_h) - \Psi(\dot q_h(t)) \ge 0, \tag{2.13}$$

$$\forall v \in \mathcal{Q} : \langle \mathrm{D}_q\mathcal{E}(t,q(t)), v-\dot q(t)\rangle_{\mathcal{Q}} + \Psi(v) - \Psi(\dot q(t)) \ge 0, \tag{2.14}$$

which hold a.e. in $(0,T)$. We may choose $v=\dot q_h(t)$ in (2.14) and add it to (2.13), obtaining

$$\langle \mathrm{D}_q\mathcal{E}(t,q_h(t)), v_h-\dot q_h(t)\rangle_{\mathcal{Q}} + \langle \mathrm{D}_q\mathcal{E}(t,q(t)), \dot q_h(t)-\dot q(t)\rangle_{\mathcal{Q}} + \Psi(v_h) - \Psi(\dot q(t)) \ge 0.$$

Employing the triangle inequality (2.1c) we find

$$\langle \mathrm{D}_q\mathcal{E}(t,q_h(t))-\mathrm{D}_q\mathcal{E}(t,q(t)), \dot q_h(t)-\dot q(t)\rangle_{\mathcal{Q}} \le \langle \mathrm{D}_q\mathcal{E}(t,q_h(t)), v_h-\dot q(t)\rangle_{\mathcal{Q}} + \Psi(v_h-\dot q(t)).$$

Since $q_h(t)\in\mathcal{S}_h(t)$ and $q(t)\in\mathcal{S}(t)$ we can use (2.10b) and find

$$\begin{aligned}&\langle \mathrm{D}_q\mathcal{E}(t,q_h(t))-\mathrm{D}_q\mathcal{E}(t,q(t)), \dot q_h(t)-\dot q(t)\rangle_{\mathcal{Q}}\\ &\le C^{\mathbf{A}}\big(h^\beta+\|q_h(t)-q(t)\|^2_{\mathcal{Q}}\big)\|\dot q(t)\|_{\mathcal{Q}} + \langle \mathbf{B}_h q(t), \dot q(t)\rangle_{\mathcal{Q}},\end{aligned} \tag{2.15}$$

where we took advantage from the fact that v_h in (2.13) was arbitrary. Define now

$$\gamma(t) \overset{\text{def}}{=} \langle \mathrm{D}_q\mathcal{E}(t,q_h(t))-\mathrm{D}_q\mathcal{E}(t,q(t)), q_h(t)-q(t)\rangle_{\mathcal{Q}} \ge \kappa\|q_h(t)-q(t)\|^2_{\mathcal{Q}}, \tag{2.16}$$

where we used the κ-uniform convexity of $\mathcal{E}$. We have

$$\begin{aligned}\dot\gamma = {}&\langle \partial_t\mathrm{D}_q\mathcal{E}(t,q_h)-\partial_t\mathrm{D}_q\mathcal{E}(t,q), q_h-q\rangle_{\mathcal{Q}} + 2\langle \mathrm{D}_q\mathcal{E}(t,q_h)-\mathrm{D}_q\mathcal{E}(t,q), \dot q_h-\dot q\rangle_{\mathcal{Q}}\\ &+ \langle \mathrm{D}_q\mathcal{E}(t,q)-\mathrm{D}_q\mathcal{E}(t,q_h)+\mathrm{D}^2_q\mathcal{E}(t,q_h)[q_h-q], \dot q_h\rangle_{\mathcal{Q}}\\ &+ \langle \mathrm{D}_q\mathcal{E}(t,q_h)-\mathrm{D}_q\mathcal{E}(t,q)+\mathrm{D}^2_q\mathcal{E}(t,q)[q-q_h], \dot q\rangle_{\mathcal{Q}}.\end{aligned}$$

Using the smoothness of $\mathcal{E}$, cf. (2.2), (2.15) implies that there exists $C_1>0$ (independent of h) such that

$$\dot\gamma \le 0 + 2C^{\mathbf{A}}C^R_1 h^\beta + 2\langle \mathbf{B}_h q, \dot q\rangle_{\mathcal{Q}} + C_1\big(\|\dot q\|_{\mathcal{Q}}+\|\dot q_h\|_{\mathcal{Q}}\big)\|q_h-q\|^2_{\mathcal{Q}}.$$

Owing to Theorem 2.2, (2.16), and the notation $\widehat C \overset{\text{def}}{=} 2C^R_1\max\big(C^{\mathbf{A}}, C_1\big)$, we deduce that

$$\dot\gamma \le \widehat C\Big(h^\beta+\frac{\gamma}{\kappa}\Big) + 2\langle \mathbf{B}_h q, \dot q\rangle_{\mathcal{Q}}.$$

Multiplication by $\mathrm{e}^{-\widehat{C}t/\kappa}$ and integration over $(0,t)$ results in

$$\gamma(t)\mathrm{e}^{-\widehat{C}t/\kappa} \leq \gamma(0) + \kappa\big(1-\mathrm{e}^{-\widehat{C}t/\kappa}\big)h^\beta + \int_0^t 2\mathrm{e}^{-\widehat{C}s/\kappa}\langle \mathbf{B}_h q(s), \dot{q}(s)\rangle_{\mathcal{Q}}\,\mathrm{d}s.$$

We multiply now by $\mathrm{e}^{\widehat{C}t/\kappa}$ and integrate by parts the last term on the right-hand side. Since $\mathbf{B}_h$ is a symmetric operator, we obtain

$$\begin{aligned}\gamma(t) \leq \gamma(0)\mathrm{e}^{\widehat{C}t/\kappa} + \kappa\big(\mathrm{e}^{\widehat{C}t/\kappa}-1\big)h^\beta + \big[\mathrm{e}^{\widehat{C}(t-s)/\kappa}\langle \mathbf{B}_h q(s), q(s)\rangle_{\mathcal{Q}}\big]_0^t \\ + \tfrac{\widehat{C}}{\kappa}\int_0^t \mathrm{e}^{\widehat{C}(t-s)/\kappa}\langle \mathbf{B}_h q(s), q(s)\rangle_{\mathcal{Q}}\,\mathrm{d}s.\end{aligned}$$

Since $q(t) \in \mathcal{S}(t)$, condition (2.10a) allows us to estimate the last two terms on the right-hand side, and we obtain

$$\kappa\|q_h(t) - q(t)\|^2_{\mathcal{Q}} \leq \gamma(t) \leq \mathrm{e}^{\widehat{C}t/\kappa}\big(\gamma(0) + \widehat{c}\,h^\beta\big) \quad \text{where } \widehat{c} = \kappa + 2C^{\mathbf{B}}. \tag{2.17}$$

Note that $q(0)$ and $q_h(0)$ are bounded, uniformly with respect to h. Hence we conclude that there exists $C_2 > 0$ (independent of h) such that $\gamma(0) \leq C_2\|q_h^0-q^0\|^2_{\mathcal{Q}}$. This implies that the solutions $q : [0,T] \to \mathcal{Q}$ and $q_h : [0,T] \to \mathcal{Q}_h$ of the rate-independent systems $(\mathcal{Q}, \mathcal{E}, \Psi, q^0)$ and $(\mathcal{Q}_h, \mathcal{E}, \Psi, q_h^0)$, respectively, satisfy

$$\|q_h(t)-q(t)\|^2_{\mathcal{Q}} \leq \frac{1}{\kappa}\mathrm{e}^{\widehat{C}T/\kappa}\big(C_2\|q_h^0-q^0\|^2_{\mathcal{Q}} + \widehat{c}\,h^\beta\big).$$

Together with (2.12) this completes the proof. ◻

REMARK 2.4. *From the proof it is clear that we may replace the symmetric linear operators* $\mathbf{B}_h$ *in* (2.10) *by more general functionals* $\mathcal{B}_h \in \mathrm{C}^1(\mathcal{Q},\mathbb{R})$*. Instead of* (2.10a) *one requires* $|\mathcal{B}_h(q)| \leq C^{\mathbf{B}}h^\beta$*, and* $\langle \mathbf{B}_h q, w\rangle$ *is replaced by* $\langle \mathrm{D}_q\mathcal{B}_h(q), w\rangle$ *in* (2.10b) *and elsewhere.*

3. Specification to the semi-linear case. In this section, we apply the abstract theory developed above to the case where the energy has a leading-order quadratic part and a lower-order nonlinear part $\mathcal{H}$ that is still convex. Moreover, the dissipation potential will also be of lower-order. Then, we will be able to exploit the situation where the approximation of points $q \in \mathcal{Q}$ via points $q_h \in \mathcal{Q}_h$ has a order of convergence in the weaker norm $\|\cdot\|_{\mathcal{X}}$, where $\mathcal{X}$ is a Banach space such that $\mathcal{Q} \subset \mathcal{X}$ densely and continuously and $\mathcal{X}' \subset \mathcal{Q}'$. We will use the symbol $\langle\cdot,\cdot\rangle_{\mathcal{X}}$ for the duality pairing between $\mathcal{X}'$ and $\mathcal{X}$. Recall that we have that

$$\forall x' \in \mathcal{X}'\ \forall q \in \mathcal{Q}\ :\ \langle x', q\rangle_{\mathcal{X}} = \langle x', q\rangle_{\mathcal{Q}}.$$

More precisely, the energy functional has the following form:

$$\forall t \in [0,T]\ \forall q \in \mathcal{Q}:\ \mathcal{E}(t,q) \overset{\mathrm{def}}{=} \tfrac{1}{2}\langle \mathbf{A}q, q\rangle_{\mathcal{Q}} + \mathcal{H}(q) - \langle \boldsymbol{\ell}(t), q\rangle_{\mathcal{Q}}, \tag{3.1a}$$

where

$$\mathbf{A} \in \mathrm{Lin}(\mathcal{Q},\mathcal{Q}'),\ \mathbf{A} = \mathbf{A}^*,\ \text{and } \exists\kappa > 0\ \forall \widehat{q} \in \mathcal{Q}:\ \langle \mathbf{A}\widehat{q}, \widehat{q}\rangle_{\mathcal{Q}} \geq \kappa\|\widehat{q}\|^2_{\mathcal{Q}}, \tag{3.1b}$$

$$\mathcal{H} \in \mathrm{C}^3(\mathcal{Q};\mathbb{R}),\ \mathcal{H} : \mathcal{Q} \to \mathbb{R} \text{ convex},\ \text{and } \mathrm{D}_q\mathcal{H} \in \mathrm{C}^0(\mathcal{Q};\mathcal{X}'), \tag{3.1c}$$

$$\boldsymbol{\ell} \in \mathrm{C}^3([0,T];\mathcal{X}'). \tag{3.1d}$$

We call $\boldsymbol{\ell}$ the external loading and $\mathcal{H}$ the hardening potential. Clearly, (3.1) implies that $\mathcal{E}$ satisfies assumptions (2.2) and that the derivative is semilinear, namely $\mathrm{D}_q\mathcal{E}(t,q) = \mathbf{A}q + \mathrm{D}_q\mathcal{H}(q) - \boldsymbol{\ell}(t)$.

For the dissipation functional Ψ we strengthen the condition (2.1) as follows:

$$\Psi : \mathcal{Q} \to [0,\infty) \text{ satisfies (2.1) and } \exists C^\Psi > 0\ \forall q \in \mathcal{X} :\ \Psi(q) \le C^\Psi \|q\|_{\mathcal{X}}. \tag{3.2}$$

The next result establishes a new a priori estimate for solutions, or more generally for stable states. Taking advantage of the semilinear structure we obtain a bound for $\|\mathbf{A}q\|_{\mathcal{X}'}$, which is crucial to establish the approximation condition (2.10b). For this result, we introduce the notations $C_1^{\mathcal{H}} \stackrel{\text{def}}{=} \sup_{\|q\|_{\mathcal{Q}} \le C_0^R} \|\mathrm{D}_q\mathcal{H}(q)\|_{\mathcal{X}'}$ and $C_0^{\boldsymbol{\ell}} \stackrel{\text{def}}{=} \sup_{t\in[0,T]} \|\boldsymbol{\ell}(t)\|_{\mathcal{X}'}$.

Proposition 3.1. *Assume that* (3.1) *and* (3.2) *hold. Then, there exists a constant* $C^{\mathcal{X}}$ *such that for all* (t,q) *with* $q \in \mathcal{S}(t)$ *we have* $\mathrm{D}_q\mathcal{E}(t,q), \mathbf{A}q \in \mathcal{X}'$, $\|\mathrm{D}_q\mathcal{E}(t,q)\|_{\mathcal{X}'} \le C^\Psi$, *and* $\|\mathbf{A}q\|_{\mathcal{X}'} \le C^{\mathcal{X}}$.

Proof. By Proposition 2.1 there exists $C_0^R > 0$ such that $\|q\|_{\mathcal{Q}} \le C_0^R$ and $-\mathrm{D}_q\mathcal{E}(t,q) \in \partial\Psi(0)$ for all $q \in \mathcal{S}(t)$. The second condition in (3.2) implies that every $\eta \in \partial\Psi(0) \subset \mathcal{Q}'$ satisfies $|\langle \eta, v\rangle| \le C^\Psi \|v\|_{\mathcal{X}}$. Thus, we have $\eta \in \mathcal{X}' \subset \mathcal{Q}'$ and $\|\eta\|_{\mathcal{X}'} \le C^\Psi$ for every $\eta \in \partial\Psi(0)$. We find $\mathbf{A}q = \mathrm{D}_q\mathcal{E}(t,q) - \mathrm{D}_q\mathcal{H}(q) + \boldsymbol{\ell}(t) = -\eta - \mathrm{D}_q\mathcal{H}(q) + \boldsymbol{\ell}(t) \in \mathcal{X}'$ with the bound

$$\|\mathbf{A}q\|_{\mathcal{X}'} \le \|\eta - \mathrm{D}_q\mathcal{H}(q) + \boldsymbol{\ell}(t)\|_{\mathcal{X}'} \le C^\Psi + C_1^{\mathcal{H}} + C_0^{\boldsymbol{\ell}}.$$

Thus, the assertion holds with $C^{\mathcal{X}} \stackrel{\text{def}}{=} C^\Psi + C_1^{\mathcal{H}} + C_0^{\boldsymbol{\ell}}$. ◻

As a corollary, every solution of $(\mathcal{Q},\mathcal{E},\Psi)$ satisfies $\|\mathbf{A}q(t)\|_{\mathcal{X}'} \le C^{\mathcal{X}}$ for all $t \in [0,T]$.

To satisfy the approximation condition we have to find vectors $v_h \in \mathcal{Q}_h$ approximating a given $w \in \mathcal{Q}$ in a suitable way. For this we assume the existence of linear operators $\mathbf{P}_h : \mathcal{Q} \to \mathcal{Q}_h$ with the following properties. There exist positive constants $C_0^{\mathbf{P}}$ and $C_i^{\mathbf{P}}$ and positive exponents α_i for $i = 1,2,3$, such that for all $h \in (0,1]$, $v \in \mathcal{Q}$, and $v_h \in \mathcal{Q}_h$ we have

$$\|\mathbf{P}_h v\|_{\mathcal{Q}} \le C_0^{\mathbf{P}} \|v\|_{\mathcal{Q}}, \tag{3.3a}$$

$$\|(\mathbf{P}_h - \mathbf{I})v\|_{\mathcal{X}} \le C_1^{\mathbf{P}} h^{\alpha_1} \|v\|_{\mathcal{Q}}, \tag{3.3b}$$

$$\|(\mathbf{P}_h^*\mathbf{A} - \mathbf{A}\mathbf{P}_h)v\|_{\mathcal{Q}'} \le C_2^{\mathbf{P}} h^{\alpha_2} \|v\|_{\mathcal{Q}}, \tag{3.3c}$$

$$\|(\mathbf{P}_h - \mathbf{I})v_h\|_{\mathcal{Q}} \le C_3^{\mathbf{P}} h^{\alpha_3} \|v_h\|_{\mathcal{Q}}, \tag{3.3d}$$

where $\mathbf{I}$ denotes the identity on $\mathcal{Q}$.

In Subsection 4.2 we will see that the above convergence rates can be easily realized in practice. In particular, if $\mathbf{P}_h$ is a projection, then (3.3d) holds with $C_3^{\mathbf{P}} = 0$ and any $\alpha_3 > 0$. Moreover, if $\mathbf{P}_h$ commutes with $\mathbf{A}$ like Galerkin projections, then (3.3c) holds with $C_2^{\mathbf{P}} = 0$ and any $\alpha_2 > 0$. We keep the more general setting here, since in general cases the Galerkin projection may not work well with the functionals $\mathcal{H}$ and Ψ, see e.g. [AMS08, MiR06].

In light of the above proposition the following result will be useful in the sequel. It provides an approximation result for $q \in \mathcal{Q}$ in the $\mathcal{Q}$-norm under the additional assumption of higher regularity, i.e., $\mathbf{A}q \in \mathcal{X}'$.

Lemma 3.2. *Assume* (3.1b) *and* (3.3). *Then, there exists* $C_4^{\mathbf{P}} > 0$ *such that for each* $h \in (0,1]$ *and* $q \in \mathcal{Q}$ *with* $\mathbf{A}q \in \mathcal{X}'$ *we have the estimate*

$$\|(\mathbf{P}_h - \mathbf{I})q\|_{\mathcal{Q}} \le C_4^{\mathbf{P}} \max\big\{ \big(h^{\alpha_1}\|q\|_{\mathcal{Q}}\|\mathbf{A}q\|_{\mathcal{X}'}\big)^{1/2}, h^{\alpha_2}\|q\|_{\mathcal{Q}}, h^{\alpha_3/2}\|q\|_{\mathcal{Q}} \big\}. \tag{3.4}$$

Proof. To estimate $\eta_h \stackrel{\text{def}}{=} \|(\mathbf{P}_h{-}\mathbf{I})q\|_{\mathcal{Q}}$ we employ $\mathbf{A}$ via (3.1b) and (3.3). Using the abbreviation $R \stackrel{\text{def}}{=} \|q\|_{\mathcal{Q}}$ we obtain

$$\begin{aligned}
\kappa\eta_h^2 &\leq \langle \mathbf{A}(\mathbf{P}_h{-}\mathbf{I})q, (\mathbf{P}_h{-}\mathbf{I})q\rangle_{\mathcal{Q}} \\
&= \langle (\mathbf{P}_h^*\mathbf{A}{-}\mathbf{A}\mathbf{P}_h)(\mathbf{P}_h{-}\mathbf{I})q, q\rangle_{\mathcal{Q}} + \langle \mathbf{A}(\mathbf{P}_h{-}\mathbf{I})\mathbf{P}_h q, q\rangle_{\mathcal{Q}} - \langle \mathbf{A}(\mathbf{P}_h{-}\mathbf{I})q, q\rangle_{\mathcal{Q}} \\
&\leq \eta_h C_2^{\mathbf{P}} h^{\alpha_2} R + C_3^{\mathbf{P}} h^{\alpha_3} \|\mathbf{P}_h q\|_{\mathcal{Q}} \|\mathbf{A}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')} R + \|\mathbf{A}q\|_{\mathcal{X}'} \|(\mathbf{P}_h{-}\mathbf{I})q\|_{\mathcal{X}} \\
&\leq \eta_h C_2^{\mathbf{P}} h^{\alpha_2} R + C_3^{\mathbf{P}} C_0^{\mathbf{P}} h^{\alpha_3} \|\mathbf{A}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')} R^2 + \|\mathbf{A}q\|_{\mathcal{X}'} C_1^{\mathbf{P}} h^{\alpha_1} R \\
&\leq \frac{\kappa}{2}\eta_h^2 + \frac{1}{2\kappa}\big(C_2^{\mathbf{P}}\big)^2 h^{2\alpha_2} R^2 + C_3^{\mathbf{P}} C_0^{\mathbf{P}} h^{\alpha_3} \|\mathbf{A}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')} R^2 + \|\mathbf{A}q\|_{\mathcal{X}'} C_1^{\mathbf{P}} h^{\alpha_1} R,
\end{aligned}$$

where we used $y_1 y_2 \leq \frac{\kappa}{2} y_1^2 + \frac{1}{2\kappa} y_2^2$ in the last passage. Canceling the first term on the right-hand side we have the desired estimate. ∎

Before formulating the main theorem we give the typical situation we have in mind. Note that in the examples 3.3 and 3.7, the derivative with respect to x is denoted by $(\cdot)'$.

EXAMPLE 3.3. *Consider* $\Omega = (0,1)$, $\mathcal{Q} = \mathrm{H}_0^1(\Omega)$, $\|q\|_{\mathcal{Q}}^2 = \int_0^1 (q'(x))^2\,\mathrm{d}x$, $\mathcal{X} = \mathrm{L}^2(\Omega)$, *and* $\mathbf{A}u = -(au')'$ *where* $a \in \mathrm{C}^\theta([0,1])$ *with* $a(x) \geq \kappa > 0$ *for all* $x \in \Omega$ *and* $\theta \in (0,1]$. *For* $h \in (0,1]$ *subdivide* Ω *into* k *subintervals of equal length, such that* $1 \geq hk > 1{-}1/k$. *Then, we define* $\mathcal{Q}_h$ *as the continuous and piecewise affine functions on the corresponding intervals. Moreover* $\mathbf{P}_h$ *as the projector defined via* $(\mathbf{P}_h q)'(x) = k\int_{(j-1)/k}^{j/k} q'(y)\,\mathrm{d}y$ *for* $x \in \big(\frac{j-1}{k}, \frac{j}{k}\big)$. *Then,* (3.3) *holds with the exponents* $\alpha_1 = 1$, $\alpha_2 = \theta$, *arbitrary* $\alpha_3 > 0$, *and the constants* $C_0^{\mathbf{P}} = 1$, $C_1^{\mathbf{P}} = 1/\pi$, $C_2^{\mathbf{P}} = \|a\|_{\mathrm{C}^\theta([0,1])}$, *and* $C_3^{\mathbf{P}} = 0$.

The approximation in $\mathcal{Q}$ *provided in Lemma 3.2 is not always optimal. If* $a \in \mathrm{C}^1([0,1])$, *then* $\mathbf{A}q \in \mathcal{X}' = \mathrm{L}^2(\Omega)$ *implies* $q \in \mathrm{H}^2(\Omega)$ *and, hence,* $\|(\mathbf{P}_h{-}\mathbf{I})q\|_{\mathrm{H}^1} \leq Ch\|q\|_{\mathrm{H}^2}$, *while the lemma just gives the bound* $h^{1/2}$.

Using all the above assumptions we will now be able to establish the approximation condition and hence control the space-time discretization error via Theorem 2.3. In addition, we also construct approximate initial conditions $q_h^0 \in \mathcal{S}_h(0)$ which have the same approximation order. Note that the error estimate provides the same order of approximation as is obtained via $\mathbf{P}_h$ in Lemma 3.2.

THEOREM 3.4. *Assume* (3.1), (3.2) *and* (3.3). *Then, there exists* $C_*^{\mathrm{sl}} > 0$ *such that for all* $q^0 \in \mathcal{S}(0)$, $h \in (0,1]$, $q_h^0 \in \mathcal{S}_h(0)$, *and all partitions* Π^τ *we have*

$$\|q_{\tau,h}(t){-}q(t)\|_{\mathcal{Q}} \leq C_*^{\mathrm{sl}}\big(h^{\beta/2}{+}\sqrt{\tau}{+}\|q^0{-}q_h^0\|_{\mathcal{Q}}\big) \text{ for all } t \in [0,T] \tag{3.5}$$

with $\beta \stackrel{\text{def}}{=} \min\{\alpha_1, 2\alpha_2, \alpha_3\}$, *where* $q : [0,T] \to \mathcal{Q}$ *is the solution of* $(\mathcal{Q}, \mathcal{E}, \Psi, q^0)$ *and* $q_{\tau,h} : [0,T] \to \mathcal{Q}_h$ *is defined via* (2.8) *with* $q_{\tau,h}^0 = q_h^0$.

Moreover, there exists a positive constant C_0^{sl} *such that for each* $q^0 \in \mathcal{S}(0)$ *there exists* $q_h^0 \in \mathcal{S}_h(0)$ *such that* $\|q_h^0{-}q^0\|_{\mathcal{Q}} \leq C_0^{\mathrm{sl}} h^{\beta/2}$.

The proof of this result is decomposed in two propositions. The first part, giving the estimate (3.5), follows directly from Theorem 2.2 if we establish the approximation condition (2.10). The second part about the existence of good q_h^0 is contained in Proposition 3.6.

PROPOSITION 3.5. *Assume* (3.1), (3.2), *and* (3.3). *Then, the approximation condition* (2.10) *holds with* $v_h = \mathbf{P}_h w$, $\mathbf{B}_h \stackrel{\text{def}}{=} -\big((\mathbf{A}(\mathbf{P}_h{-}\mathbf{I}){+}(\mathbf{A}(\mathbf{P}_h{-}\mathbf{I}))^*\big)$, *and* $\beta = \min\{\alpha_1, 2\alpha_2, \alpha_3\}$ *where* α_i, $i = 1,2,3$, *are defined in* (3.3).

Proof. We fix $t \in [0,T]$ and take any $q \in \mathcal{S}(t)$, $q_h \in \mathcal{S}_h(t)$, and $w \in \mathcal{Q}$. By Propositions 2.1 and 3.1 and by (3.3b) we have

$$\|q\|_{\mathcal{Q}} \leq C_0^R,\ \|q_h\|_{\mathcal{Q}} \leq C_0^R,\ \|\mathbf{A}q\|_{\mathcal{X}'} \leq C^{\mathcal{X}},\ \|v_h{-}w\|_{\mathcal{X}} \leq C_1^{\mathbf{P}} h^{\alpha_1}\|w\|_{\mathcal{Q}}. \tag{3.6}$$

With the definition (3.1a) of $\mathcal{E}$ and assumptions (3.1c) and (3.2), we get

$$\begin{aligned}&\langle \mathrm{D}_q\mathcal{E}(t,q_h), v_h{-}w\rangle_{\mathcal{Q}} + \Psi(v_h{-}w) = \langle \mathbf{A}q_h{+}\mathrm{D}_q\mathcal{H}(q_h){-}\boldsymbol{\ell}(t), v_h{-}w\rangle_{\mathcal{Q}} + \Psi(v_h{-}w)\\ &\leq \langle \mathbf{A}(q_h{-}q), v_h{-}w\rangle_{\mathcal{Q}} + \big(\|\mathbf{A}q\|_{\mathcal{X}'}{+}\|\mathrm{D}_q\mathcal{H}(q_h)\|_{\mathcal{X}'}{+}\|\boldsymbol{\ell}(t)\|_{\mathcal{X}'}{+}C^{\Psi}\big)\|v_h{-}w\|_{\mathcal{X}}.\end{aligned}$$

Using $C_1^{\mathcal{H}}$ and $C_0^{\boldsymbol{\ell}}$ as defined above, we find

$$\begin{aligned}&\langle \mathrm{D}_q\mathcal{E}(t,q_h), v_h{-}w\rangle_{\mathcal{Q}} + \Psi(v_h{-}w)\\ &\leq \langle \mathbf{A}(q_h{-}q), v_h{-}w\rangle_{\mathcal{Q}} + \big(C^{\mathcal{X}}{+}C_1^{\mathcal{H}}{+}C_0^{\boldsymbol{\ell}}{+}C^{\Psi}\big)C_1^{\mathbf{P}}h^{\alpha_1}\|w\|_{\mathcal{Q}}.\end{aligned} \tag{3.7}$$

Since $\alpha_1 \geq \beta$ the second term in the above right hand side is as required in (2.10b).

Hence, it remains to estimate the first term on the right-hand side of (3.7). Recalling the definition of $\mathbf{B}_h$ and using $v_h = \mathbf{P}_h w$ some elementary rearrangements give

$$\begin{aligned}&\langle \mathbf{A}(q_h{-}q), v_h{-}w\rangle_{\mathcal{Q}} = \langle \mathbf{A}(q_h{-}q), (\mathbf{P}_h{-}\mathbf{I})w\rangle_{\mathcal{Q}}\\ &= \langle (\mathbf{P}_h^*\mathbf{A}{-}\mathbf{A}\mathbf{P}_h)(q_h{-}q), w\rangle_{\mathcal{Q}} + \langle \mathbf{A}(\mathbf{P}_h{-}\mathbf{I})q_h, w\rangle_{\mathcal{Q}} + \langle \mathbf{A}q, (\mathbf{P}_h{-}\mathbf{I})w\rangle_{\mathcal{Q}} + \langle \mathbf{B}_h q, w\rangle_{\mathcal{Q}}.\end{aligned}$$

Using (3.3) and Young's inequality , we obtain

$$\begin{aligned}&\langle \mathbf{A}(q_h{-}q), (\mathbf{P}_h{-}\mathbf{I})w\rangle_{\mathcal{Q}}\\ &\leq \big(C_2^{\mathbf{P}}h^{\alpha_2}\|q_h{-}q\|_{\mathcal{Q}}{+}C_3^{\mathbf{P}}\|\mathbf{A}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')}h^{\alpha_3}\|q_h\|_{\mathcal{Q}}{+}C_1^{\mathbf{P}}\|\mathbf{A}q\|_{\mathcal{X}'}h^{\alpha_1}\big)\|w\|_{\mathcal{Q}} + \langle \mathbf{B}_h q, w\rangle_{\mathcal{Q}}\\ &\leq \Big(\frac{1}{2}\big(C_2^{\mathbf{P}}\big)^2 h^{2\alpha_2}{+}\frac{1}{2}\|q_h{-}q\|_{\mathcal{Q}}^2{+}C_3^{\mathbf{P}}C_0^R\|\mathbf{A}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')}h^{\alpha_3}{+}C_1^{\mathbf{P}}C^{\mathcal{X}}h^{\alpha_1}\Big)\|w\|_{\mathcal{Q}} + \langle \mathbf{B}_h q, w\rangle_{\mathcal{Q}}.\end{aligned}$$

Inserting this into (3.7) we have established (2.10b).

To obtain (2.10a) we simply use $\langle \mathbf{B}_h q, q\rangle_{\mathcal{Q}} = -2\langle \mathbf{A}q, (\mathbf{P}_h{-}\mathbf{I})q\rangle_{\mathcal{Q}}$ and obtain

$$|\langle \mathbf{B}_h q, q\rangle_{\mathcal{Q}}| \leq 2\|\mathbf{A}q\|_{\mathcal{X}'}\|(\mathbf{P}_h{-}\mathbf{I})q\|_{\mathcal{X}} \leq 2C^{\mathcal{X}}C_1^{\mathbf{P}}h^{\alpha_1}\|q\|_{\mathcal{Q}} \leq 2C^{\mathcal{X}}C_1^{\mathbf{P}}C_0^R h^{\alpha_1},$$

which gives (2.10a). This finishes the proof. ∎

The next proposition supplies a useful initial condition q_h^0 for the spatially discretized rate-independent systems $(\mathcal{Q}_h, \mathcal{E}, \Psi)$. For a given $q^0 \in \mathcal{Q}$ and $h \in (0,1]$ we define

$$q_h^0 \stackrel{\text{def}}{=} \mathrm{Argmin}\{\mathcal{E}(0,\widehat{q}_h){+}\Psi(\widehat{q}_h{-}\mathbf{P}_h q^0)\,|\,\widehat{q}_h \in \mathcal{Q}_h\}. \tag{3.8}$$

By the uniform convexity of $\mathcal{E}(0,\cdot)$ the value is uniquely defined. Moreover, the triangle inequality (2.1c) implies

$$\mathcal{E}(0,q_h^0) \leq \mathcal{E}(0,\widehat{q}_h){+}\Psi(\widehat{q}_h{-}\mathbf{P}_h q^0){-}\Psi(q_h^0{-}\mathbf{P}_h q^0) \leq \mathcal{E}(0,\widehat{q}_h){+}\Psi(\widehat{q}_h{-}q_h^0),$$

for all $\widehat{q}_h \in \mathcal{Q}_h$, i.e. $q_h^0 \in \mathcal{S}_h(0)$. We now prove that it is close to $\mathbf{P}_h q^0$ and q^0 if $q^0 \in \mathcal{S}(0)$.

Proposition 3.6. *Assume* (3.1), (3.2), *and* (3.3). *Then, there exists* $C_0^{\mathrm{sl}} > 0$ *such that for all* $q^0 \in \mathcal{S}(0)$ *and all* $h \in (0,1]$ *the value* $q_h^0 \in \mathcal{Q}_h$ *defined via* (3.8) *satisfies*

$$\|q_h^0{-}q^0\|_{\mathcal{Q}} \leq C_0^{\mathrm{sl}}h^{\beta/2}, \tag{3.9}$$

with $\beta=\min\{\alpha_1,2\alpha_2,\alpha_3\}$ *where* α_i, $i=1,2,3$, *are defined in* (3.3).

Proof. Since $q^0\in\mathcal{S}(0)$ we can apply (2.7b) for $\widehat{q}=q_h^0$, we obtain

$$\begin{aligned}\frac{\kappa}{2}\|q_h^0-q^0\|_{\mathcal{Q}}^2&\le\mathcal{E}(0,q_h^0)-\mathcal{E}(0,q^0)+\Psi(q_h^0-q^0)\\&\le\mathcal{E}(0,q_h^0)-\mathcal{E}(0,q^0)+\Psi(q_h^0-\mathbf{P}_hq^0)+\Psi(\mathbf{P}_hq^0-q^0)\\&\le\mathcal{E}(0,\mathbf{P}_hq^0)+\mathcal{E}(0,q^0)+\Psi((\mathbf{P}_h-\mathbf{I})q^0),\end{aligned}\tag{3.10}$$

where we have used the triangle inequality (2.1c) in the second estimate and the fact that q_h^0 is a minimizer in the third. Define

$$\mathcal{I}(q^0,\mathbf{P}_hq^0)\stackrel{\text{def}}{=}\int_0^1\langle\mathrm{D}_q\mathcal{E}(0,q^0+s(\mathbf{P}_h-\mathbf{I})q^0)-\mathrm{D}_q\mathcal{E}(0,q^0),(\mathbf{P}_h-\mathbf{I})q^0\rangle_{\mathcal{Q}}\,\mathrm{d}s.$$

Thus using Taylor's formula, (3.2) and Proposition 3.1, we deduce from (3.10) that

$$\begin{aligned}\frac{\kappa}{2}\|q_h^0-q^0\|_{\mathcal{Q}}^2&\le\mathcal{I}(q^0,\mathbf{P}_hq^0)+\langle\mathrm{D}_q\mathcal{E}(0,q^0),(\mathbf{P}_h-\mathbf{I})q^0\rangle_{\mathcal{Q}}+\Psi((\mathbf{P}_h-\mathbf{I})q^0)\\&\le\mathcal{I}(q^0,\mathbf{P}_hq^0)+\|\mathrm{D}_q\mathcal{E}(0,q^0)\|_{\mathcal{X}'}\|(\mathbf{P}_h-\mathbf{I})q^0\|_{\mathcal{X}}+C^{\Psi}\|(\mathbf{P}_h-\mathbf{I})q^0\|_{\mathcal{X}}\\&\le\mathcal{I}(q^0,\mathbf{P}_hq^0)+2C^{\Psi}C_1^{\mathbf{P}}h^{\alpha_1}\|q^0\|_{\mathcal{Q}}.\end{aligned}\tag{3.11}$$

For $\mathcal{I}(q^0,\mathbf{P}_hq^0))$ we use that, by (3.3a) and (2.9a), we know $\|\mathbf{P}_hq^0\|_{\mathcal{Q}}\le C_0^{\mathbf{P}}C_0^R$. On the ball of radius $(1+C_0^{\mathbf{P}})C_0^R$ the second derivative of $\mathcal{E}$ is bounded by a constant $C_2^{\mathcal{E}}>0$ and we obtain $\mathcal{I}(q^0,\mathbf{P}_hq^0)\le\frac{C_2^{\mathcal{E}}}{2}\|(\mathbf{P}_h-\mathbf{I})q^0\|_{\mathcal{Q}}^2$. Since $q^0\in\mathcal{S}(0)$, Proposition 3.1 yields $\|\mathbf{A}q^0\|_{\mathcal{X}'}\le C^{\mathcal{X}}$. Thus, Lemma 3.2 implies that there exists $C^{\mathcal{I}}>0$ such that $\mathcal{I}(q^0,\mathbf{P}_hq^0)\le C^{\mathcal{I}}h^{\beta}$. Hence we infer from (3.11) the desired result. □

Finally, we notice that the power of h in (3.3b) and (3.3c) depends on the choice of $\mathcal{X}$. Of course, the optimal choice is to make $\mathcal{X}$ as big as allowed by the condition (3.2) for Ψ. We illustrate this in the following example, which gives a first example for convergence rates of space-time discretizations.

EXAMPLE 3.7. *We consider the situation of Example 3.3 with* $\Omega=(0,1)$, $\mathcal{Q}=\mathrm{H}_0^1(\Omega)$, $\mathcal{E}(t,q)=\int_0^1\big(\frac{1}{2}(q'(x))^2+H(q(x))-\ell(t,x)\cdot q(x)\big)\,\mathrm{d}x$ *and* $\Psi(\dot q)=\int_0^1|\dot q|\,\mathrm{d}x$. *We assume that* $H\in\mathrm{C}^3(\mathbb{R};\mathbb{R})$ *is convex and that* $\ell\in\mathrm{C}^1([0,1];\mathrm{L}^\infty(\Omega))$. *Thus, the abstract nonsmooth differential inclusion* (1.1) *takes the explicit form*

$$\begin{aligned}0\in\mathrm{Sign}\big(\dot q(t,x)\big)-q''(t,x)+\mathrm{D}_qH(q(t,x))-\ell(t,x)\quad&\text{for }(t,x)\in[0,T]\times\Omega,\\q(t,0)=q(t,1)=0\quad&\text{for }t\in[0,T].\end{aligned}$$

Here "Sign" denotes the multi-valued signum function with Sign$(0)=[-1,1]$.

As in Example 3.3 the subspaces $\mathcal{Q}_h$ *contain the piecewise affine functions on an equidistant partitions of* $\Omega=(0,1)$ *and* $\mathbf{P}_h:\mathcal{Q}\to\mathcal{Q}_h$ *are the orthogonal Galerkin projectors. Then, taking* $\mathcal{X}=\mathrm{L}^p(\Omega)$ *with* $p\in[1,\infty]$, *we may prove that the power is* $\alpha_1=\widehat{\alpha}(p)\stackrel{\text{def}}{=}\min\big(1,\frac{1}{2}+\frac{1}{p}\big)$ *in* (3.3b). *Since* α_2 *and* α_3 *may be taken as big as we like, our main approximation result* (3.5) *in Theorem 3.4 gives the following error bound*

$$\|q_{\tau,h}(t)-q(t)\|_{\mathrm{H}^1}\le C_*^{\mathrm{sl}}\big(\sqrt{\tau}+h^{\widehat{\alpha}(p)/2}+\|q_{\tau,h}(0)-q(0)\|_{\mathrm{H}^1}\big)\quad\text{for }t\in[0,T].$$

By choosing $p\in[1,2]$ *we obtain the spatial convergence rate* $h^{1/2}$.

4. Application to the isothermal Souza-Auricchio model.

4.1. The isothermal Souza-Auricchio model. Let us start by briefly recalling some modeling issues. The reader is referred to [SMZ98, AuP04, AuP02, ARS07] for additional details and motivation.

We consider a material with a reference configuration $\Omega \subset \mathbb{R}^d$ with $d \in \{2,3\}$. We assume that Ω is an open bounded set with Lipschitz boundary. This body may undergo displacements $u : \Omega \to \mathbb{R}^d$ and phase transformations. The latter will be characterized by a mesoscopic internal variable $z : \Omega \to \mathbb{R}^{d\times d}_{\rm dev}$ where $\mathbb{R}^{d\times d}_{\rm dev}$ is the space of $d\times d$ tensors with vanishing trace. In particular, the tensor z stands as the inelastic part of the deformation due to the martensitic phase transformation.

The set of admissible displacements $\mathcal{F}$ is chosen as a suitable subspace of $\mathrm{H}^1(\Omega;\mathbb{R}^d)$ by prescribing homogeneous Dirichlet data on the measurable subset $\Gamma_{\rm Dir}$ of $\partial\Omega$, i.e.,

$$\mathcal{F} \stackrel{\text{def}}{=} \big\{\, u \in \mathrm{H}^1(\Omega;\mathbb{R}^d) \,\big|\, u = 0 \text{ on } \Gamma_{\rm Dir} \,\big\}.$$

Non-homogeneous Dirichelet conditions could be considered as well by letting $u = \widetilde{u} + u_{\rm Dir}$ with $\widetilde{u} \in \mathcal{F}$. The internal variable z belongs to $\mathcal{Z} \stackrel{\text{def}}{=} \mathrm{H}^1(\Omega;\mathbb{R}^{d\times d}_{\rm dev})$ and we let $\mathcal{Q} \stackrel{\text{def}}{=} \mathcal{F}\times\mathcal{Z}$. We choose $\mathcal{X} \stackrel{\text{def}}{=} \mathcal{X}_{\mathcal{F}} \times \mathcal{X}_{\mathcal{Z}}$ where, given $\zeta \in [0,1/2)$,

$$\mathcal{X}_{\mathcal{F}} \stackrel{\text{def}}{=} \mathrm{L}^2(\Omega;\mathbb{R}^d) \times H^{-\zeta}(\Gamma_{\rm Neu};\mathbb{R}^d), \quad \mathcal{X}_{\mathcal{Z}} \stackrel{\text{def}}{=} \mathrm{L}^2(\Omega;\mathbb{R}^{d\times d}_{\rm dev})$$

where $\Gamma_{\rm Neu} \stackrel{\text{def}}{=} \partial\Omega \setminus \Gamma_{\rm Dir}$. Moreover, we will denote by $\langle\cdot,\cdot\rangle_{\mathcal{X}_{\mathcal{F}}}$ the duality pairing between $\mathcal{X}'_{\mathcal{F}}$ and $\mathcal{X}_{\mathcal{F}}$. In particular, note that the injection $i : \mathcal{Q} \to \mathcal{X}$ given by $iu \stackrel{\text{def}}{=} (u,\gamma u)$, where $\gamma : H^1(\Omega) \to L^2(\Gamma_{\rm Neu})$ in the standard trace operator, is continuous and dense. Hence, one has that

$$\mathcal{X}' = \mathcal{X}'_{\mathcal{F}} \times \mathcal{X}'_{\mathcal{Z}} = \Big(\mathrm{L}^2(\Omega;\mathbb{R}^d) \times H^{\zeta}_0(\Gamma_{\rm Neu};\mathbb{R}^d)\Big) \times \mathrm{L}^2(\Omega;\mathbb{R}^{d\times d}_{\rm dev}) \subset \mathcal{Q}'.$$

We will denote the states by $q \stackrel{\text{def}}{=} (u,z)$. The linearized strain tensor is given by $\mathbf{e}(u) \stackrel{\text{def}}{=} \frac{1}{2}(\nabla u + \nabla u^{\mathsf{T}}) \in \mathbb{R}^{d\times d}_{\rm sym}$ where $\mathbb{R}^{d\times d}_{\rm sym}$ is the space of symmetric $d\times d$ tensors endowed with the scalar product $v{:}w \stackrel{\text{def}}{=} \mathrm{tr}(v^{\mathsf{T}}w)$ and the corresponding norm $|v|^2 \stackrel{\text{def}}{=} v{:}v$ for all $v,w \in \mathbb{R}^{d\times d}_{\rm sym}$. Here $(\cdot)^{\mathsf{T}}$ and $\mathrm{tr}(\cdot)$ denote the transpose and the trace of the tensor , respectively. We assume that $\Gamma_{\rm Dir}$ has positive surface measure so that Korn's inequality holds, i.e. there exists $C^{\rm Korn} > 0$ such that

$$\forall u \in \mathcal{F}: \ \|\mathbf{e}(u)\|^2_{\mathrm{L}^2(\Omega;\mathbb{R}^{d\times d}_{\rm sym})} \geq C^{\rm Korn}\|u\|^2_{\mathrm{H}^1(\Omega;\mathbb{R}^d)}. \tag{4.1}$$

For more details on Korn's inequality and its consequences, we refer to [KoO88] or [DuL76].

The *stored-energy potential* takes the following form

$$\mathcal{E}(t,u,z) \stackrel{\text{def}}{=} \int_\Omega \Big(W(x,\mathbf{e}(u)(x),z(x)) + \frac{\nu}{2}|\nabla z(x)|^2\Big)\,\mathrm{d}x - \langle l(t),u\rangle_{\mathcal{X}_{\mathcal{F}}}. \tag{4.2}$$

Here ν is a positive coefficient that is expected to measure some nonlocal interaction effect for the internal variable z, whereas $W : \Omega \times \mathbb{R}^{d\times d}_{\rm sym} \times \mathbb{R}^{d\times d}_{\rm dev} \to \mathbb{R}$ is the stored-energy density and reads

$$W(x,\mathbf{e}(u)(x),z(x)) \stackrel{\text{def}}{=} \frac{1}{2}\big((\mathbf{e}(u)(x) - z(x)){:}\mathbb{C}(\mathbf{e}(u)(x) - z(x))\big) + \widehat{H}(z(x)).$$

In the latter, $\mathbb{C}$ is the elastic tensor, $\widehat{H} : \mathbb{R}^{d\times d}_{\rm dev} \to \mathbb{R}$ represents the hardening potential. For simplicity, we will omit any dependence on the material point $x \in \Omega$. Moreover $l(t)$ denotes an applied mechanical loading of the form

$$\langle l(t), u\rangle_{\mathcal{X}_{\mathcal{F}}} \stackrel{\rm def}{=} \int_\Omega f_{\rm appl}(t,x)\cdot u(x)\,{\rm d}x + \int_{\Gamma_{\rm Neu}} g_{\rm appl}(t,x)\cdot u(x)\,{\rm d}\Gamma, \tag{4.3}$$

where $f_{\rm appl}$ and $g_{\rm appl}$ are given body forces and a surface tractions on $\Gamma_{\rm Neu}$.

In [SMZ98, AuS01, AuP04], the authors are interested in $\widehat{H} = H_{\rm SoAu}$ with

$$H_{\rm SoAu}(z) \stackrel{\rm def}{=} c_1\sqrt{\delta^2{+}|z|^2} + \frac{c_2}{2}|z|^2 + \frac{((|z|{-}c_3)_+)^4}{\delta(1{+}|z|^2)}, \tag{4.4}$$

where $c_1 > 0$ is an activation threshold for initiation of martensitic phase transformations, $c_2 > 0$ measures the occurrence of hardening with respect to the internal variable z, and $c_3 > 0$ represents the maximum modulus of transformation strain that can be obtained by alignment of martensitic variants. The original model is obtained in the limit $\delta \to 0$ in (4.4) and $\nu \to 0$ in (4.2). More precisely, $\widehat{H} = H_{\rm org}$ is defined as follows

$$H_{\rm org}(z) \stackrel{\rm def}{=} c_1|z| + \frac{c_2}{2}|z|^2 + \chi(z),$$

where $\chi : \mathbb{R}^{d\times d}_{\rm dev} \to [0,\infty]$ is the indicator function of the ball $\big\{z \in \mathbb{R}^{d\times d}_{\rm dev}\,\big|\,|z| \le c_3\big\}$.

To model the hysteretic behavior of shape-memory materials, we define the dissipation potential as follows

$$\psi(v) \stackrel{\rm def}{=} \int_\Omega \rho|v(x)|\,{\rm d}x, \quad \text{where } \rho > 0. \tag{4.5}$$

The material constitutive relation reads as the following *doubly nonlinear differential inclusion*

$$\begin{pmatrix} 0 \\ \partial\psi(\dot z) \end{pmatrix} + \begin{pmatrix} \partial_u\mathcal{E}(t,q) \\ \partial_z\mathcal{E}(t,q) \end{pmatrix} \ni \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{4.6}$$

where $\partial_u\mathcal{E}(t,q) = -{\rm div}\,(\mathbb{C}(\mathbf{e}(u){-}z)) - l(t)$, $\partial_z\mathcal{E}(t,q) = -\mathbb{C}(\mathbf{e}(u){-}z) + \partial_z\widehat{H}(z) - \nu\Delta z$. Hence, the first component provides the elastic equilibrium equations, whereas the second component gives the flow law for the internal variable z.

Defining $q = (u,z)$, $\Psi(\dot q) = \psi(\dot z)$, and $\langle \boldsymbol{\ell}(t), q\rangle_{\mathcal{Q}} = \langle l(t), u\rangle_{\mathcal{X}_{\mathcal{F}}}$, system (4.6) can be rewritten in the abstract form

$$\partial\Psi(\dot q) + \mathbf{A}q + {\rm D}_q\mathcal{H}(q) - \boldsymbol{\ell}(t) \ni 0, \tag{4.7}$$

where $\mathcal{H}(q) \stackrel{\rm def}{=} \int_\Omega H(u(x),z(x))\,{\rm d}x$ with $H(u,z) = \widehat{H}(z) - \frac{c_2}{2}|z|^2$ and

$$\mathbf{A} \stackrel{\rm def}{=} \begin{pmatrix} -{\rm div}(\mathbb{C}\mathbf{e}(\cdot)) & {\rm div}\,(\mathbb{C}(\cdot)) \\ -\mathbb{C}\mathbf{e}(\cdot) & \mathbb{C}(\cdot) - \nu\Delta(\cdot) + c_2\mathbf{I}(\cdot) \end{pmatrix}. \tag{4.8}$$

Here we assume that the elasticity tensor $\mathbb{C}$ is a symmetric positive definite map, i.e.

$$\exists\mu > 0\ \forall e \in \mathbb{R}^{d\times d}_{\rm sym}:\ \ e{:}\mathbb{C}{:}e \ge \mu|e|^2. \tag{4.9}$$

By assuming $f_{\rm appl} \in C^3([0,T]; {\rm L}^2(\Omega;\mathbb{R}^d))$ and $g_{\rm appl} \in C^3([0,T]; H^\zeta_0(\Gamma_{\rm Neu};\mathbb{R}^d))$ in (4.3) we readily check that $\boldsymbol{\ell} \in C^3([0,T];\mathcal{X}')$ (see (3.1d)). Moreover, we may prove that the functional $\mathcal{H}$ built on $H = H_{\rm SoAu}$ satisfies (3.1c). Furthermore, by using (4.9), we infer that (3.1b) holds for $\mathbf{A}$ defined in (4.8). Then it follows that (3.1) is satisfied. This is however not the case for the original model with $H_{\rm org}$, the reader is referred to [AMS08] for some discussion on the limit $(\nu,\delta)\to(0,0)$.

Existence and uniqueness results for a temperature dependent variant of (4.7) were obtained in [MiP07]. Following [AuS01] a function $\widehat{H}(z,\theta) = H_{\rm SoAu}(z,\theta)$ is considered by allowing the constants $c_i(\theta)$, $i = 1, 2, 3$, in (4.4) and $\mathbb{C}(\theta)$ to depend on the temperature θ. Then, the authors assumed that the temperature is given as an applied load, $\theta = \Theta(t, x)$, while here we treat a simpler case where the temperature is constant. The assumption to consider the temperature given as an applied load is acceptable if the changes of the loading are slow and the body is small in at least one direction. Hence, the excessive or missing heat can be balanced through the environment.

4.2. The spatial discretization. Before introducing the spatial discretization, we shall reinforce our assumptions by asking Ω to be a polyhedron. This requirement is quite classical and basically meant to simplify the forthcoming presentation. In particular, our analysis can be generalized to piecewise smooth domains by means of additional technicalities (see, for instance, [BrS94, Cia02]). Moreover, for the sake of definiteness we require that each face of $\partial\Omega$ is contained either in $\Gamma_{\rm Dir}$ or in $\Gamma_{\rm Neu}$.

Our space-discrete analysis will follow from the $\mathrm{H}^{1+\sigma}$ regularity of the associated boundary value problem for linearized elastostatics and the Neumann problem. Namely, we explicitly require that Ω, $\Gamma_{\rm Dir}$, and $\mathbb{C}$ satisfy the following condition:

$$\begin{aligned}&\exists\sigma\in(0,1]\ \exists\widetilde{C}>0\ \forall f\in\mathcal{X}_\mathcal{F}'\ \forall g\in\mathrm{L}^2(\Omega):\\&\|u_f\|_{\mathrm{H}^{1+\sigma}(\Omega;\mathbb{R}^d)}\le\widetilde{C}\|f\|_{\mathcal{X}_\mathcal{F}'}\quad\text{and}\quad\|\zeta_g\|_{\mathrm{H}^{1+\sigma}(\Omega)}\le\widetilde{C}\|g\|_{\mathrm{L}^2(\Omega)},\end{aligned}\tag{4.10}$$

where $u_f\in\mathcal{F}$ and $\zeta_g\in\mathrm{H}^1(\Omega)$ are the unique solution u and ζ, respectively, of

$$\begin{aligned}&\forall v\in\mathcal{F}:\ \int_\Omega \mathbb{C}\mathbf{e}(u){:}\mathbf{e}(v)\,\mathrm{d}x=\langle f,v\rangle_{\mathcal{X}_\mathcal{F}},\\&\forall\eta\in\mathrm{H}^1(\Omega):\ \int_\Omega c_2\zeta\eta+\nu\nabla\zeta{\cdot}\nabla\eta\,\mathrm{d}x=\int_\Omega g\eta\,\mathrm{d}x.\end{aligned}$$

The latter regularity requirement is quite natural and is fulfilled (with $\sigma = 1$) when $\Gamma_{\rm Neu} = \emptyset$ and Ω is either smooth [Cia86, Theorem 2.2-4, p.99] or a convex polyhedron, see [Gri89] in 2D and [DKV88, EbF99] in 3D. Non-convex polyhedrons can also be considered (possibly with $\sigma < 1$) and results for the mixed Neumann-Dirichlet conditions are also available [EbF99]. Additional details on regularity issues and asymptotic developments of solutions near corner points may be found in [Kon67, Dau88, Nic92, Kne06], among others.

Let us start from the following lemma which is crucial to obtain the error estimates for space-time discretizations of the Souza-Auricchio model. The lemma relates to Proposition 3.1 where we now exploit the choice $\mathcal{X} = \mathcal{X}_\mathcal{F}\times\mathcal{X}_\mathcal{Z}$. Another important feature is that the coupling between the elasticity problem and the Neumann problem for the internal variable is of lower order.

Lemma 4.1. *If* (4.10) *holds, then there exists* $C_1^\mathcal{X} > 0$ *such that for* $f\in\mathcal{X}'$ *the unique* $q\in\mathcal{Q}$ *solving* $\mathbf{A}q = f$ *in* $\mathcal{Q}'$ *satisfies*

$$\|q\|_{\mathrm{H}^{1+\sigma}(\Omega;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\rm dev})}\le C_1^\mathcal{X}\|f\|_{\mathcal{X}'},\tag{4.11}$$

where $\sigma \in (0,1]$ *is defined in* (4.10).

Proof. Owing to the coercivity (3.1b) of $\mathbf{A}$ we readily check that there exists $C_1 > 0$ such that

$$\|q\|_{\mathcal{Q}} \leq \|f\|_{\mathcal{Q}'}/\kappa \leq C_1 \|f\|_{\mathcal{X}'}. \tag{4.12}$$

Letting $q = (u, z)$ and $f = (f_1, f_2) \in \mathcal{X}'_{\mathcal{F}} \times \mathcal{X}'_{\mathcal{Z}}$, we have $\mathbf{A}q = f$ if and only if

$$\forall v \in \mathcal{F}: \ \int_\Omega \mathbb{C}\mathbf{e}(v){:}\mathbf{e}(u)\,\mathrm{d}x = \int_\Omega (-\mathrm{div}(\mathbb{C}z)){\cdot}v\,\mathrm{d}x + \langle f_1, v\rangle_{\mathcal{X}_{\mathcal{F}}}, \tag{4.13}$$

$$\forall w \in \mathcal{Z}: \ \int_\Omega (c_2 w{:}z + \nu\nabla w{:}\nabla z)\,\mathrm{d}x = \int_\Omega \big(f_2 + \mathbb{C}(\mathbf{e}(u) - z)\big){:}w\,\mathrm{d}x. \tag{4.14}$$

Using (4.12), the $\mathcal{X}'$-norm of the right-hand side $\big(f_1 - \mathrm{div}(\mathbb{C}z), f_2 + \mathbb{C}(\mathbf{e}(u) - z)\big)$ is bounded by $C_2\|f\|_{\mathcal{X}'}$. Moreover, (4.14) consists of decoupled Neumann problems for the components of z. Thus, employing (4.10) we deduce

$$\|q\|_{\mathrm{H}^{1+\sigma}(\Omega;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\mathrm{dev}})} \leq \widetilde{C}\|(f_1 - \mathrm{div}(\mathbb{C}z), f_2 + \mathbb{C}\mathbf{e}(u))\|_{\mathcal{X}'} \leq \widetilde{C}C_2\|f\|_{\mathcal{X}'},$$

which is the desired result. ◻

We shall define the spatial discretization by letting $\mathcal{F}_h$ and $\mathcal{Z}_h$ be finite-dimensional subspaces of $\mathcal{F}$ and $\mathcal{Z}$, respectively. In particular, assume to be given a regular triangulation $\{\mathcal{T}_k\}$ of Ω (cf. [QuV94]) and choose $\mathcal{F}_h$ and $\mathcal{Z}_h$ to be the subspaces of continuous, piecewise polynomials of fixed degree $m \geq 1$ on $\{\mathcal{T}_k\}$. Finally, let $\mathcal{Q}_h \overset{\mathrm{def}}{=} \mathcal{F}_h \times \mathcal{Z}_h$ and assume to be given linear projectors $\mathbf{\Pi}_h : \mathcal{Q} \to \mathcal{Q}_h$ fulfilling

$$\forall s \in (0,1] \ \exists C^{\mathbf{\Pi}} > 0 \ : \ \|(\mathbf{\Pi}_h - \mathbf{I})q\|_{\mathcal{Q}} \leq C^{\mathbf{\Pi}} h^s \|q\|_{\mathrm{H}^{1+s}(\Omega;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\mathrm{dev}})}. \tag{4.15}$$

The latter can be realized, for instance, by letting $\mathbf{\Pi}_h$ be the L^2 orthogonal projector. The interpolation error control of (4.15) is well-known for $s = 1$ and follows from [HP*05, Lemma 5.6] for $s \in (0,1)$.

The operator $\mathbf{P}_h : \mathcal{Q} \to \mathcal{Q}_h$ is instead defined to be the Galerkin projection via $\mathbf{A}$. Namely, for all $q \in \mathcal{Q}$, we let $\mathbf{P}_h q \overset{\mathrm{def}}{=} \widehat{q}_h$, where $\widehat{q}_h \in \mathcal{Q}_h$ is the unique solution of

$$\langle \mathbf{A}\widehat{q}_h, p_h\rangle_{\mathcal{Q}} = \langle \mathbf{A}q, p_h\rangle_{\mathcal{Q}} \text{ for all } p_h \in \mathcal{Q}_h. \tag{4.16}$$

It remains to prove that $\mathbf{P}_h$ defined above fulfills (3.3); then we are in the position to apply Theorem 3.4 to obtain explicit a priori error bounds for our space-time discretization of the quasistatic evolution problem for the Souza-Auricchio model.

THEOREM 4.2. *Assume that* (4.10) *holds. Then there exist* $C_*^{\mathrm{SoAu}} > 0$ *such that for all* $h \in (0,1]$ *and all partitions* Π^τ *of* $[0,T]$*, we have*

$$\|q_{\tau,h}(t) - q(t)\|_{\mathcal{Q}} \leq C_*^{\mathrm{SoAu}}\big(h^{\sigma/2} + \sqrt{\tau}\big) \textit{ for all } t \in [0,T], \tag{4.17}$$

where $q : [0,T] \to \mathcal{Q}$ *is a solution of* $(\mathcal{Q}, \mathcal{E}, \Psi)$ *and* $q_{\tau,h} : [0,T] \to \mathcal{Q}_h$ *is defined via* (2.8) *and the initial condition* $q_{\tau,h}(0) = \mathrm{Argmin}\,\big\{\mathcal{E}(0, \widehat{q}_h) + \Psi(\widehat{q}_h - \mathbf{P}_h q(0)) \,\big|\, \widehat{q}_h \in \mathcal{Q}_h\big\}$.

Proof. By the definition (4.16) we have $\mathbf{P}_h \circ \mathbf{P}_h = \mathbf{P}_h$ and (3.3d) holds for any $\alpha_3 \geq 0$. Moreover, by using (4.16) we readily check that, for all p, $q \in \mathcal{Q}$,

$$\begin{aligned} \langle(\mathbf{P}_h^*\mathbf{A} - \mathbf{A}\mathbf{P}_h)q, p\rangle_{\mathcal{Q}} &= \langle \mathbf{A}q, \mathbf{P}_h p\rangle_{\mathcal{Q}} - \langle \mathbf{A}\mathbf{P}_h q, p\rangle_{\mathcal{Q}} \\ &\overset{(4.16)}{=} \langle \mathbf{A}q, \mathbf{P}_h p\rangle_{\mathcal{Q}} - \langle \mathbf{A}\mathbf{P}_h q, \mathbf{P}_h p\rangle_{\mathcal{Q}} = \langle \mathbf{A}(q - \mathbf{P}_h q), \mathbf{P}_h p\rangle_{\mathcal{Q}} \overset{(4.16)}{=} 0. \end{aligned}$$

Hence, (3.3c) holds for any $\alpha_2 \geq 0$. Moreover, (3.3a) holds with $C_0^{\mathbf{P}} = \|\mathbf{A}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')}/\kappa$, because

$$\kappa\|\mathbf{P}_h q\|_{\mathcal{Q}}^2 \leq \langle \mathbf{A}\mathbf{P}_h q, \mathbf{P}_h q\rangle_{\mathcal{Q}} \stackrel{(4.16)}{=} \langle \mathbf{A}q, \mathbf{P}_h q\rangle_{\mathcal{Q}} \leq \|\mathbf{A}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')}\|q\|_{\mathcal{Q}}\|\mathbf{P}_h q\|_{\mathcal{Q}}.$$

Finally, let us check for property (3.3b) by means of the classical duality technique by Aubin and Nitsche [Aub67, Nit68]. Fix $q \in \mathcal{Q}$ and, by letting $J^{\mathcal{X}} : \mathcal{X} \to \mathcal{X}'$ be the Riesz mapping, define $\varphi \in \mathcal{Q}$ as the unique solution of $\mathbf{A}\varphi = J^{\mathcal{X}}(\mathbf{P}_h - \mathbf{I})q$. Then, using $\mathbf{A} = \mathbf{A}^*$ for arbitrary $\varphi_h \in \mathcal{Q}_h$ we have

$$\begin{aligned}\|(\mathbf{P}_h-\mathbf{I})q\|_{\mathcal{X}}^2 &= \langle J^{\mathcal{X}}(\mathbf{P}_h-\mathbf{I})q, (\mathbf{P}_h-\mathbf{I})q\rangle_{\mathcal{X}} = \langle \mathbf{A}\varphi, (\mathbf{P}_h-\mathbf{I})q\rangle_{\mathcal{Q}}\\ &= \langle \mathbf{A}(\mathbf{P}_h-\mathbf{I})q, \varphi\rangle_{\mathcal{Q}} \stackrel{(4.16)}{=} \langle \mathbf{A}(\mathbf{P}_h-\mathbf{I})q, \varphi-\varphi_h\rangle_{\mathcal{Q}} \leq C_5^{\mathbf{P}}\|q\|_{\mathcal{Q}}\|\varphi-\varphi_h\|_{\mathcal{Q}}\end{aligned}$$

where $C_5^{\mathbf{P}} \stackrel{\text{def}}{=} \|\mathbf{A}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')}\sup_{h\in(0,1]}\|\mathbf{P}_h-\mathbf{I}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q})}$. Choosing $\varphi_h = \mathbf{\Pi}_h\varphi$ and exploiting (4.15) for $s = \sigma$ with σ from (4.10) we arrive at

$$\|(\mathbf{P}_h-\mathbf{I})q\|_{\mathcal{X}}^2 \leq C_5^{\mathbf{P}}\|q\|_{\mathcal{Q}}\|(\mathbf{\Pi}_h-\mathbf{I})\varphi\|_{\mathcal{Q}} \leq C_5^{\mathbf{P}}\|q\|_{\mathcal{Q}}\, C^{\Pi}h^{\sigma}\|\varphi\|_{\mathrm{H}^{1+\sigma}(\Omega;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\mathrm{dev}})}.$$

Using the definition of φ and the regularity theory provided in Lemma 4.1 we conclude

$$\|(\mathbf{P}_h-\mathbf{I})q\|_{\mathcal{X}}^2 \leq C_5^{\mathbf{P}}\|q\|_{\mathcal{Q}}\, C^{\Pi}h^{\sigma}C_1^{\mathcal{X}}\|(\mathbf{P}_h-\mathbf{I})q\|_{\mathcal{X}},$$

which is the desired approximation result (3.3b) with $\alpha_1 = \sigma$. Hence, applying Theorem 3.4 with $\beta = \alpha_1 = \sigma$, the desired result follows. ☐

REMARK 4.3. *In the special case of a convex reference domain* Ω *for* $\Gamma_{\mathrm{Neu}} = \emptyset$, *we obtain* (4.17) *with* $\sigma = 1$.

Appendix. The aim of this section is to give the proof of (2.9c). We follow the ideas developed in [MiT04] and keep track of all constants to see that they do not depend on h.

Proof. We first recall that there exists $C_0^R > 0$ such that all the solutions satisfy the a priori bound

$$q_{\tau,h}(t) \in \mathcal{B}_{C_0^R} \stackrel{\text{def}}{=} \big\{q \in \mathcal{Q}\,\big|\,\|q\|_{\mathcal{Q}} \leq C_0^R\big\} \text{ for all } \tau \in (0,T],\ h \in [0,1],\ t \in [0,T]$$

(see Theorem 2.2).

Let now the partition $\Pi^{\tau} \stackrel{\text{def}}{=} \{0 = t_0^{\tau} < t_1^{\tau} < \cdots < t_{k_\tau}^{\tau} = T\}$ be given and define Π^{τ_j} by successive bisections, namely

$$\Pi^{\tau_j} \stackrel{\text{def}}{=} \{t_\ell^{\tau} + 2^{-j}r(t_\ell^{\tau} - t_{\ell-1}^{\tau}) \ :\ \ell = 1,\ldots,k_\tau,\ r = 0,1,\ldots,2^j\}.$$

We shall associate to these partitions the corresponding solutions $q_{\tau_j,h}$ of the incremental problems for $(\mathcal{Q}_h, \mathcal{E}, \Psi, q_h(0))$. We want to compare $q_{\tau_j,h}$ and $q_{\tau_{j+1},h}$. To do so, we define $\mathcal{E}^1$ and $\mathcal{E}^2$ as follows: for $t_k^{\tau} \in \Pi^{\tau_{j+1}}$, let $\bar{t}_k^{\tau} \stackrel{\text{def}}{=} \max\{s_n^{\tau} \in \Pi^{\tau_j} \,|\, s_n^{\tau} \leq t_k^{\tau}\}$, $\mathcal{E}^1(t_k^{\tau}, q) \stackrel{\text{def}}{=} \mathcal{E}(\bar{t}_k^{\tau}, q)$ and $\mathcal{E}^2(t_k^{\tau}, q) \stackrel{\text{def}}{=} \mathcal{E}(t_k^{\tau}, q)$ for $t_k^{\tau} \in \Pi^{\tau_{j+1}}$. Notice that $q_{\tau_j,h}$ and $q_{\tau_{j+1},h}$ are the incremental solutions obtained with $\mathcal{E}^1$ and $\mathcal{E}^2$ on the partition $\Pi^{\tau_{j+1}}$.

For the sake of simplicity let us introduce the following notations:

$$\forall t_k^{\tau} \in \Pi^{\tau_{j+1}} :\ q_{\tau,h}^{1,k} \stackrel{\text{def}}{=} q_{\tau_j,h}(t_k^{\tau}) \text{ and } q_{\tau,h}^{2,k} \stackrel{\text{def}}{=} q_{\tau_{j+1},h}(t_k^{\tau}),$$

and $e^k_{\tau,h} \stackrel{\text{def}}{=} q^{1,k}_{\tau,h}-q^{2,k}_{\tau,h}$ and $\eta_k\mu \stackrel{\text{def}}{=} \mu_k-\mu_{k-1}$ where μ stands for t^τ, $q^j_{\tau,h}$ and $e_{\tau,h}$ (and $\gamma_{\tau,h}$, see below). Since $q^j_{\tau,h}$ solves the incremental problems $(\mathrm{IP})^{j,\tau,h}$, we have

$$\forall v_h \in \mathcal{Q}_h : \langle \mathrm{D}_q\mathcal{E}^j(t^\tau_k, q^{j,k}_{\tau,h}), v_h-\eta_k q^j_{\tau,h}\rangle_{\mathcal{Q}} + \Psi(v_h) - \Psi(\eta_k q^j_{\tau,h}) \geq 0. \tag{4.18}$$

Choosing $v_h = \eta_k q^{3-j}_{\tau,h}$ and adding the equations for $j = 1, 2$ gives

$$\langle \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h})-\mathrm{D}_q\mathcal{E}^2(t^\tau_k, q^{2,k}_{\tau,h}), \eta_k q^1_{\tau,h}-\eta_k q^2_{\tau,h}\rangle_{\mathcal{Q}} \leq 0. \tag{4.19}$$

Define

$$\gamma^k_{\tau,h} \stackrel{\text{def}}{=} \langle \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h})-\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h}), q^{1,k}_{\tau,h}-q^{2,k}_{\tau,h}\rangle_{\mathcal{Q}} \geq \kappa\|q^{1,k}_{\tau,h}-q^{2,k}_{\tau,h}\|^2_{\mathcal{Q}} = \kappa\|e^k_{\tau,h}\|^2_{\mathcal{Q}}. \tag{4.20}$$

Let us estimate the increment

$$\begin{aligned}
&\eta_k\gamma_{\tau,h} \stackrel{\text{def}}{=} \gamma^k_{\tau,h} - \gamma^{k-1}_{\tau,h} = \langle \eta_k(\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h})-\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h})), e^{k-1}_{\tau,h}\rangle_{\mathcal{Q}}\\
&- \langle \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h})-\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h}), \eta_k e_{\tau,h}\rangle_{\mathcal{Q}} - 2\langle \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h})-\mathrm{D}_q\mathcal{E}^2(t^\tau_k, q^{2,k}_{\tau,h}), \eta_k e_{\tau,h}\rangle_{\mathcal{Q}}\\
&+ 2\langle \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h})-\mathrm{D}_q\mathcal{E}^2(t^\tau_k, q^{2,k}_{\tau,h}), \eta_k e_{\tau,h}\rangle_{\mathcal{Q}}.
\end{aligned}$$

Let $A_k \in \mathrm{Lin}(\mathcal{Q}, \mathcal{Q}')$ be the symmetric operator defined by

$$A_k \stackrel{\text{def}}{=} \int_0^1 \mathrm{D}^2_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h}+\theta e^k_{\tau,h})\,\mathrm{d}\theta.$$

We get $A_k e^k_{\tau,h} = \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h}) - \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h})$, thus

$$\begin{aligned}
&\langle \eta_k(\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h})-\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h})), e^{k-1}_{\tau,h}\rangle_{\mathcal{Q}} - \langle \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h})-\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h}), \eta_k e_{\tau,h}\rangle_{\mathcal{Q}}\\
&= \langle A_k e^k_{\tau,h}-A_{k-1}e^{k-1}_{\tau,h}, e^{k-1}_{\tau,h}\rangle_{\mathcal{Q}} - \langle A_k e^k_{\tau,h}, \eta_k e_{\tau,h}\rangle_{\mathcal{Q}}\\
&= -\langle A_k\eta_k e_{\tau,h}, \eta_k e_{\tau,h}\rangle_{\mathcal{Q}} + \langle (A_k-A_{k-1})e^{k-1}_{\tau,h}, e^{k-1}_{\tau,h}\rangle_{\mathcal{Q}}.
\end{aligned} \tag{4.21}$$

By convexity of $\mathcal{E}^1(t^\tau_k, \cdot)$, we have

$$\forall y \in \mathcal{Q} : \ \langle A_k y, y\rangle_{\mathcal{Q}} \geq 0,$$

and since $\mathrm{D}^2_q\mathcal{E}$ is Lipschitz continuous on $[0, T] \times \mathcal{B}_R$ for all $R > 0$,

$$\|A_k-A_{k-1}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')} \leq C^{\mathcal{E},R}\big(|t^\tau_k-t^\tau_{k-1}|+\|\eta_k q^1_{\tau,h}\|_{\mathcal{Q}}+\|\eta_k q^2_{\tau,h}\|_{\mathcal{Q}}\big)$$

where $C^{\mathcal{E},R}$ depends only on $\mathcal{E}$ and $R > 0$ such that $R \geq \max_{\tau,h}\big\{\|\eta_k q^j_{\tau,h}\|_{\mathcal{Q}};\ j = 1, 2,\ t^\tau_k \in \Pi^{\tau_{j+1}}\big\}$ and $\mathcal{B}_R$ denotes the ball of radius R. Using (4.19), it follows that

$$\begin{aligned}
\eta_k\gamma_{\tau,h} &\leq C^{\mathcal{E},R}\big(|t^\tau_k-t^\tau_{k-1}|+\|\eta_k q^1_{\tau,h}\|_{\mathcal{Q}}+\|\eta_k q^2_{\tau,h}\|_{\mathcal{Q}}\big)\|e^{k-1}_{\tau,h}\|^2_{\mathcal{Q}}\\
&+ 2\|\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h})-\mathrm{D}_q\mathcal{E}^2(t^\tau_k, q^{2,k}_{\tau,h})\|_{\mathcal{Q}'}\|\eta_k e_{\tau,h}\|_{\mathcal{Q}}.
\end{aligned} \tag{4.22}$$

Since $\mathcal{E}(t, \cdot)$ is κ-uniformly convex, the incremental solutions are Lipschitz continuous, i.e.

$$\forall j = 1, 2 : \ \|\eta_k q^j_{\tau,h}\|_{\mathcal{Q}} \leq C^R_1|t^\tau_k-t^\tau_{k-1}|, \tag{4.23}$$

where $C_1^R > 0$ is independent of h and τ (cf. Theorem 2.2). Carrying (4.23) and (4.20) in (4.22), and observing that $\|\eta_k e_{\tau,h}\|_{\mathcal{Q}} \leq \|\eta_k q^1_{\tau,h}\|_{\mathcal{Q}} + \|\eta_k q^2_{\tau,h}\|_{\mathcal{Q}}$, we obtain

$$\eta_k \gamma_{\tau,h} \leq \frac{C^{\mathcal{E},R}}{\kappa}(1{+}2C_1^R)\gamma^{k-1}_{\tau,h}|t^\tau_k{-}t^\tau_{k-1}| + 4\rho C_1^R|t^\tau_k{-}t^\tau_{k-1}|,$$

where

$$\rho \stackrel{\text{def}}{=} \max_{t^\tau_k\in\Pi^{\tau_{j+1}}} \sup_{q\in\mathcal{B}_{C_0^R}} \|\mathrm{D}_q\mathcal{E}^1(t^\tau_k,q){-}\mathrm{D}_q\mathcal{E}^2(t^\tau_k,q)\|_{\mathcal{Q}'}.$$

Let us denote $C_4 = \max\{C^{\mathcal{E},R}(1{+}2C_1^R)/\kappa, 4C_1^R\}$, we infer

$$\gamma^k_{\tau,h} \leq \gamma^{k-1}_{\tau,h}\big(1{+}C_4(t^\tau_k{-}t^\tau_{k-1})\big) + \rho C_4(t^\tau_k{-}t^\tau_{k-1}).$$

Since $\gamma^0_{\tau,h} = 0$, by induction over k, we find

$$\gamma^k_{\tau,h} \leq C_4\rho\sum_{k=1}^{n}(t^\tau_k{-}t^\tau_{k-1})\prod_{j=k+1}^{n}\big(1{+}C_4(t^\tau_j{-}t^\tau_{j-1})\big) \leq C_4\rho\mathrm{e}^{C_4T}T.$$

Using (4.20), it follows that

$$\|q^{1,k}_{\tau,h}{-}q^{2,k}_{\tau,h}\|^2_{\mathcal{Q}} \leq \frac{C_4\mathrm{e}^{C_4T}T}{\kappa}\rho. \tag{4.24}$$

Owing to the definitions of $\mathcal{E}^1$ and $\mathcal{E}^2$, we infer that there exists a constant $C_5 > 0$ such that

$$\rho \leq C_5 \max_{t^\tau_k\in\Pi^{\tau_{j+1}}}(t^\tau_k{-}t^\tau_{k-1}) \leq C_5 2^{-j}\tau,$$

which implies that

$$\forall t\in[0,T]:\ \|q_{\tau_{j+1},h}(t){-}q_{\tau_j,h}(t)\|_{\mathcal{Q}} \leq C_6 2^{-j/2}\sqrt{\tau}, \quad \text{where } C_6 = \sqrt{\frac{C_4T\mathrm{e}^{C_4T}}{\kappa}C_5}.$$

Note that $(q_{\tau_j,h}(t))_{j\in\mathbb{N}}$ is a Cauchy sequence which limit $q_h : [0,T] \to \mathcal{Q}_h$ is the unique solution for $(\mathcal{Q}_h, \mathcal{E}, \Psi, q_h(0))$. By adding all these estimates, we infer

$$\forall t\in[0,T]:\ \|q_{\tau,h}(t){-}q_h(t)\|_{\mathcal{Q}} \leq \sum_{j=0}^{\infty} C_6 2^{-j/2}\sqrt{\tau} \leq 4C_6\sqrt{\tau}, \tag{4.25}$$

which proves (2.9c). □

Acknowledgments. A.P. was supported by the Deutsche Forschungsgemeinschaft through the project C18 "Analysis and numerics of multidimensional models for elastic phase transformation in a shape-memory alloys" of the Research Center Matheon. U.S. is partially supported by the FP7-IDEAS-ERC-StG grant # 200497 (BioSMA). Moreover, L.P. and U.S. wish to thank for the kind hospitality of WIAS.

References

[ACZ99] J. Alberty, C. Carstensen, and D. Zarrabi. Adaptive numerical analysis in primal elastoplasticity with hardening. *Comput. Methods Appl. Mech. Engrg.*, 171(3-4), 175–204, 1999.

[AlC00] J. Alberty and C. Carstensen. Numerical analysis of time-depending primal elastoplasticity with hardening. *SIAM J. Numer. Anal.*, 37, 1271–1294 (electronic), 2000.

[AMS08] F. Auricchio, A. Mielke, and U. Stefanelli. A rate-independent model for the isothermal quasi-static evolution of shape-memory materials. *Math. Models Methods Appl. Sci.*, 18(1), 125–164, 2008.

[ARS07] F. Auricchio, A. Reali, and U. Stefanelli. A three-dimensional model describing stress-induces solid phase transformation with residual plasticity. *Int. J. Plasticity*, 23(2), 207–226, 2007.

[ARS09] 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.*, 2009, to appear. IMATI Preprint 2PV08/2/0, 2008.

[Aub67] J.-P. Aubin. Behavior of the error of the approximate solutions of boundary value problems for linear elliptic operators by Gelerkin's and finite difference methods. *Ann. Scuola Norm. Sup. Pisa (3)*, 21, 599–637, 1967.

[AuP02] 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 Engng.*, 55, 1255–1284, 2002.

[AuP04] 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. Methods Engng.*, 61, 716–737, 2004.

[AuS01] F. Auricchio and E. Sacco. Thermo-mechanical modelling of a superelastic shape-memory wire under cyclic stretching-bending loadings. *Int. J. Solids Structures*, 38, 6123–6145, 2001.

[AuS04] F. Auricchio and U. Stefanelli. Numerical treatment of a 3D super-elastic constitutive model. *Internat. J. Numer. Methods Engng.*, 61, 142–155, 2004.

[AuS05] F. Auricchio and U. Stefanelli. Well-posedness and approximation for a one-dimensional model for shape memory alloys. *Math. Models Methods Appl. Sci.*, 15, 1301–1327, 2005.

[BrS94] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1994.

[Car99] C. Carstensen. Numerical analysis of the primal problem of elastoplasticity with hardening. *Numer. Math.*, 82(4), 577–597, 1999.

[Cia86] P. G. Ciarlet. *Élasticité tridimensionnelle*, volume 1 of *Recherches en Mathématiques Appliquées [Research in Applied Mathematics]*. Masson, Paris, 1986.

[Cia02] P. G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[CKO06] C. Carstensen, R. Klose, and A. Orlando. Reliable and efficient equilibrated a posteriori finite element error control in elastoplasticity and elastoviscoplasticity with hardening. *Comput. Methods Appl. Mech. Engrg.*, 195(19-22), 2574–2598, 2006.

[COV06] C. Carstensen, A. Orlando, and J. Valdman. A convergent adaptive finite element method for the primal problem of elastoplasticity. *Internat. J. Numer. Methods Engrg.*, 67(13), 1851–1887, 2006.

[Dau88] M. Dauge. *Elliptic boundary value problems on corner domains*, volume 1341 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1988.

[DE*07] J. K. Djoko, F. Ebobisse, A. T. McBride, and B. D. Reddy. A discontinuous Galerkin formulation for classical and gradient plasticity. I. Formulation and analysis. *Comput. Methods Appl. Mech. Engrg.*, 196(37-40), 3881–3897, 2007.

[DKV88] B. E. J. Dahlberg, C. E. Kenig, and G. C. Verchota. Boundary value problems for the systems of elastostatics in Lipschitz domains. *Duke Math. J.*, 57(3), 795–818, 1988.

[DuL76] G. Duvaut and J.-L. Lions. *Inequalities in mechanics and physics*, volume 219 in *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1976.

[EbF99] C. Ebmeyer and J. Frehse. Mixed boundary value problems for nonlinear elliptic equations in multidimensional non-smooth domains. *Math. Nachr.*, 203, 47–74, 1999.

[Gri89] P. Grisvard. Singularités en elasticité. *Arch. Rational Mech. Anal.*, 107(2), 157–180, 1989.

[HaR99] W. Han and B. D. Reddy. *Plasticity (Mathematical Theory and Numerical Analysis)*, volume 9 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 1999.

[HP*05] P. Houston, I. Perugia, A. Schneebeli, and D. Schötzau. Mixed discontinuous Galerkin approximation of the Maxwell operator: the indefinite case. *M2AN Math. Model. Numer. Anal.*, 39(4), 727–753, 2005.

[KMR05] M. Kružík, A. Mielke, and T. Roubíček. Modelling of microstructure and its evolution in shape-memory-alloy single-crystals, in particular in CuAlNi. *Meccanica*, 40, 389–418, 2005.

[Kne06] D. Knees. Global regularity of the elastic fields of a power-law model on Lipschitz domains. *Math. Methods Appl. Sci.*, 29(12), 1363–1391, 2006.

[Kon67] V. A. Kondrat'ev. Boundary value problems for elliptic equations in domains with conical or angular points. *Trudy Moskov. Mat. Obšč.*, 16, 209–292, 1967.

[KoO88] V. Kondrat'ev and O. Oleinik. Boundary-value problems for the system of elasticity theory in unbounded domains. korn's inequailities. *Russian Math. Surveys*, 43(5), 65–119, 1988.

[LiB96] Y. Li and I. Babuška. A convergence analysis of a p-version finite element method for one-dimensional elastoplasticity problem with constitutive laws based on the gauge function method. *SIAM J. Numer. Anal.*, 33(2), 809–842, 1996.

[LiB97] Y. Li and I. Babuška. A convergence analysis of an h-version finite-element method with high-order elements for two-dimensional elastoplasticity problems. *SIAM J. Numer. Anal.*, 34(3), 998–1036, 1997.

[MiP07] A. Mielke and A. Petrov. Thermally driven phase transformation in shape-memory alloys. *Adv. Math. Sci. Appl.*, 17, 667–685, 2007.

[MiR06] A. Mielke and T. Roubíček. Numerical approaches to rate-independent processes and applications in inelasticity. *M2AN Math. Model. Numer. Anal.*, 2006. Submitted. WIAS Preprint 1169.

[MiR07] A. Mielke and R. Rossi. Existence and uniqueness results for a class of rate-independent hysteresis problems. *Math. Models Methods Appl. Sci.*, 17, 81–123, 2007.

[MiT04] A. Mielke and F. Theil. On rate–independent hysteresis models. *NoDEA Nonlinear Differential Equations Appl.*, 11, 151–189, 2004. (Accepted July 2001).

[MPP08] A. Mielke, L. Paoli, and A. Petrov. On the existence and approximation for a 3D model of thermally induced phase transformations in shape-memory alloys. *SIAM J. Math. Anal.*, 2008. Submitted. WIAS preprint 1330.

[Nic92] S. Nicaise. About the Lamé system in a polygonal or a polyhedral domain and a coupled problem between the Lamé system and the plate equation. I. Regularity of the solutions. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 19(3), 327–361, 1992.

[Nit68] J. Nitsche. Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens. *Numer. Math.*, 11, 346–348, 1968.

[OrP04] A. Orlando and D. Perić. Analysis of transfer procedures in elastoplasticity based on the error in the constitutive equations: theory and numerical illustration. *Internat. J. Numer. Methods Engrg.*, 60(9), 1595–1631, 2004.

[QuV94] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*, volume 23 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1994.

[SMZ98] A. Souza, E. Mamiya, and N. Zouain. Three-dimensional model for solids undergoing stress-induced phase transformations. *Europ. J. Mech., A/Solids*, 17, 789–806, 1998.