SIAM J. NUMER. ANAL.
Vol. 48, No. 5, pp. 1625–1646

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 classical strain gradient plasticity and the isothermal Souza–Auricchio model for shape-memory alloys.

Key words. evolutionary variational inequalities, rate-independent processes, space-time discretization, error estimates, strain gradient plasticity, shape-memory alloys

AMS subject classifications. 65M15, 49J40, 74C05, 74N30

DOI. 10.1137/090750238

1. Introduction. The present analysis is concerned with error estimates for space-time discretizations in the context of 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 effects in mechanics, we assume that the system is rate-independent, which amounts to asking that Ψ be 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 nonsmooth 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 nonsmoothness 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., [29, 1, 2, 15, 14, 13, 41, 42, 51].

*Received by the editors February 20, 2009; accepted for publication (in revised form) June 15, 2010; published electronically September 9, 2010.
http://www.siam.org/journals/sinum/48-5/75023.html

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

‡Laboratoire de Mathématiques de l'Université de Saint-Etienne (LaMUSE), 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). This author's work was supported by the Deutsche Forschungsgemeinschaft through the project C18 "Analysis and numerics of multidimensional models for elastic phase transformation in shape-memory alloys" of the Research Center MATHEON.

¶Istituto di Matematica Applicata e Tecnologie Informatiche - CNR, via Ferrata 1, 27100 Pavia, Italy (ulisse.stefanelli@imati.cnr.it). This author's work was partially supported by FP7-IDEAS-ERC-StG grant 200497 (BioSMA).

In particular, we are here specifically interested in a semilinear case, where the potential energy has the 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 nonlinear hardening models in linearized elastoplasticity. In particular, we shall specifically consider the application of the present abstract frame to the case of strain gradient plasticity [23, 24, 26, 27] as well as to the case of the isothermal Souza–Auricchio model for shape-memory alloys (SMA). SMA are metallic alloys showing some surprising thermomechanical 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 *superelastic*, 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 are biomedical devices (vascular stents, archwires, endo-guidewires) and micro-electromechanical systems (MEMS, i.e., actuators, valves, minigrippers, and positioners). The Souza–Auricchio model considered here 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 [54, 9, 6, 8] for the derivation and the mechanics and to [10, 11, 44, 4] for the mathematical analysis.

The paper is organized as follows. After introducing our assumptions more precisely in section 2, we recall a well-posedness result from [47]. 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 $\max\{\, t_k^\tau - t_{k-1}^\tau \ : \ k = 1, \ldots, k^\tau \,\} \le \tau$ and a sequence $(\mathcal{Q}_h)_{h>0}$ of finite-dimensional spaces exhausting $\mathcal{Q}$. Then, the space-time discretized incremental minimization problem

$$q_{\tau,h}^k \stackrel{\text{def}}{=} \mathrm{Argmin}\,\big\{ \mathcal{E}(t_k^\tau, \widehat{q}_h) + \Psi(\widehat{q}_h - q_{\tau,h}^{k-1}) \,\big|\, \widehat{q}_h \in \mathcal{Q}_h \big\}$$

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 a general version, and for brevity we give here a slightly strengthened form:

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

Under a suitable additional assumption we construct a constant $C > 0$ such that

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

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 $\mathrm{D}_q \mathcal{H} \in$

$C^{1,\mathrm{Lip}}(\mathcal{Q},\mathcal{X}')$. 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 elastoplastic problems have been obtained exclusively by assuming higher smoothness-in-time on the solutions. Early results in this direction are in [35, 32] (see also [33]), where $\mathcal{O}(h+\sqrt{\tau})$ convergence for the stress is proved for stress-based formulations under extra regularity in space for σ and z. Strain-based formulations are instead considered in [29], where the optimal-order error in time $\mathcal{O}(\tau)$ is achieved for both u and z by assuming the extra regularity $(u,z)\in \mathrm{W}^{3,1}(0,T,\mathrm{H}^1\times\mathrm{L}^2)$. Subsequently, this extra regularity requirement has been weakened in [28, 31] ($(u,z)\in \mathrm{W}^{2,1}(0,T,\mathrm{H}^1\times\mathrm{L}^2)$ and extra regularity in space for $(\dot{u},\dot{z})$) [2, 13] and ($(\sigma,z)\in \mathrm{H}^1(0,T,\mathrm{L}^2\times\mathrm{L}^2)$ and $u\in \mathrm{W}^{1,\infty}(0,T,\mathrm{H}^2)$). A posteriori estimates are established in [13], indeed paving the way to the possibility of applying adaptive techniques to elastoplastic problems [1, 15, 14, 51]. Error control for strain gradient plasticity is presented in [20] where, nevertheless, the extra regularity $(u,z)\in \mathrm{H}^1(0,T;\mathrm{H}^2\times\mathrm{H}^2)$ is required.

The crucial issue with respect to the mentioned error estimates is that extra regularity in plasticity is generally not to be expected, and one is lead to consider *natural regularity conditions* instead, namely, the regularity ensured by the existence proof. Note that the issue of proving *convergence* of fully discrete schemes under natural regularity conditions has been already tackled in [30] (see also [29, sect. 11.4, p. 253]). The reader is referred to [41, 42] for some additional material in the direction of convergence of finite element methods in elastoplasticity as well as to [53, 50] for results on error control of implicit time-discretizations of nonlinear parabolic problems.

The main point of our analysis is precisely that of proving a priori error estimates for (1.1) under sole *natural regularity conditions*, that is, without assuming extra regularity of the solutions. Note, however, that our overall assumptions will correspond to the occurrence of gradient terms. In particular, classical linearized elastoplasticity cannot be directly accommodated in our setting, as the lack of the plastic strain gradient would prevent us from exhibiting an explicit convergence rate (convergence, however, being ensured) and the reader is referred to [37] for a recent result in this direction. On the other hand, we are in the position to specify our abstract convergence result to the case of strain gradient plasticity and the isothermal Souza–Auricchio model in sections 4–5. Related convergence results for models of phase transformations in shape-memory alloys were obtained in [38, 45, 43]; 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 5, we will show that these conditions are satisfied for the Souza–Auricchio model for phase transformations in the SMA; see [44, 4]. 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

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

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

where $\mathbf{I}$ is the identity in $\mathcal{Q}$. We consider the 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 estimates.

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*

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

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

(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 [47] . 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. [43, Thm. 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 easily 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 now treat 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 $\mathcal{Q}_0 \stackrel{\text{def}}{=} \mathcal{Q}$ 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 [47] 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 $\Pi^\tau = \{0 = t_0^\tau < t_1^\tau < \cdots < t_{k^\tau}^\tau = T\}$ with

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

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

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

By uniform convexity and continuity, 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

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

where $q_{\tau,h}^k$ 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 [47, 46]. In the present case these assumptions hold and we are able to deduce the convergence of the whole sequence $q_{\tau,h}$ to the unique solution of $(\mathcal{Q}_h, \mathcal{E}, \Psi, q_h^0)$. Let us summarize this discussion in the following statement, which is a slight generalization of Theorem 7.1 in [47], 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* $q_h^0 \in \mathcal{S}_h(0)$*, there exists a unique solution* $q_h \in \mathrm{C}^{\mathrm{Lip}}([0,T], \mathcal{Q}_h)$ *of the initial-value problem* $(\mathcal{Q}, \mathcal{E}, \Psi, q_h^0)$. *Moreover, there exist positive constants* C_0^R, C_1^R, *and* $\bar{C}$ *such that, for all* $h \in [0,1]$ *and all partitions* Π^τ, *we have*

$$\|q_{\tau,h}(t)\|_{\mathcal{Q}} \leq C_0^R,\ \|q_h(t)\|_{\mathcal{Q}} \leq C_0^R \quad \textit{for all } t \in [0,T]; \tag{2.9a}$$

$$\|\dot{q}_{\tau,h}(t)\|_{\mathcal{Q}} \leq C_1^R,\ \|\dot{q}_h(t)\|_{\mathcal{Q}} \leq C_1^R \quad \textit{for a.a. } t \in [0,T]; \tag{2.9b}$$

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

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 in obtaining the error estimate of space-time discretizations. Condition (2.9a) follows from Proposition 2.1 by combining parts (b) and (c). Conditions (2.9b) are ensured by [47, Thm. 7.5.b].

Now we address 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 [47]. The *approximation condition* for our error bounds reads as follows:

$$\begin{aligned} &\exists C^{\mathbf{A}} > 0\ \forall h \in (0,1]\ \forall t \in [0,T], \quad q_h \in \mathcal{S}_h(t), \quad w \in \mathcal{Q}, \quad \exists v_h \in \mathcal{Q}_h : \\ &\langle \mathrm{D}_q\mathcal{E}(t,q_h), v_h - w\rangle_{\mathcal{Q}} + \Psi(v_h - w) \leq C^{\mathbf{A}}\, h^\beta \|w\|_{\mathcal{Q}}. \end{aligned} \tag{2.10}$$

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 the stability sets $\mathcal{S}_h$. Moreover, as $\mathcal{S}_h$ are usually much smaller than $\mathcal{Q}_h$ and we have the a priori bound (2.9a), condition (2.10) turns out to be weaker than (1.3).

THEOREM 2.3. *Assume that* $\mathcal{Q}$, $\mathcal{Q}_h$, $\mathcal{E}$, *and* Ψ *satisfy* (2.1), (2.2) *and that* (2.10) *holds. Then, there exists a constant* $C_* > 0$ *such that, for any* $h \in (0,1]$, $q_h^0 \in \mathcal{S}_h(0)$, *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*

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

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

There are two possible strategies for establishing 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 $\|q_h(t) - q(t)\|_{\mathcal{Q}}^2$. This is the approach of the proof given below. Another alternative would be to consider a fixed time-discretization and to estimate $\|q_{\tau,h}^k - q_\tau^k\|_{\mathcal{Q}}^2$ uniformly with respect to τ and $k = 1, \ldots, k_\tau$ (cf. [4]).

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

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

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

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

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

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

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

Employing the triangle inequality (2.1c) we find

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

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

$$\langle \mathrm{D}_q\mathcal{E}(t, q_h(t)) - \mathrm{D}_q\mathcal{E}(t, q(t)), \dot{q}_h(t) - \dot{q}(t)\rangle_{\mathcal{Q}} \leq C^{\mathbf{A}} h^\beta \|\dot{q}(t)\|_{\mathcal{Q}}, \tag{2.15}$$

where we took advantage of the fact that v_h in (2.13) was arbitrary. Now define

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

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

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

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

$$\dot{\gamma} \leq 0 + 2C^{\mathbf{A}} C_1^R h^\beta + C_1\big(\|\dot{q}\|_{\mathcal{Q}} + \|\dot{q}_h\|_{\mathcal{Q}}\big)\|q_h - q\|_{\mathcal{Q}}^2.$$

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

$$\dot{\gamma} \leq \widehat{C}\Big(h^\beta + \frac{\gamma}{\kappa}\Big).$$

In particular, we readily obtain that

$$\gamma(t) \leq \gamma(0) \mathrm{e}^{\widehat{C}t/\kappa} + \kappa\big(\mathrm{e}^{\widehat{C}t/\kappa} - 1\big)h^\beta$$

and we have

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

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 $\gamma(0) \leq C_2 \|q_h^0 - q^0\|_{\mathcal{Q}}^2$. 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

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

Together with (2.12) this completes the proof. □

3. Specification to the semilinear 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 an 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:

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

where

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

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

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

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:

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

3.1. The error estimate. 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 (2.10). For this result, we introduce the notations $C_1^{\mathcal{H}} \overset{\text{def}}{=} \sup_{\|q\|_{\mathcal{Q}} \leq C_0^R} \|\mathrm{D}_q \mathcal{H}(q)\|_{\mathcal{X}'}$ and $C_0^{\boldsymbol{\ell}} \overset{\text{def}}{=} \sup_{t \in [0,T]} \|\boldsymbol{\ell}(t)\|_{\mathcal{X}'}$.

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* $\mathrm{D}_q \mathcal{E}(t, q), \mathbf{A} q \in \mathcal{X}'$, $\|\mathrm{D}_q \mathcal{E}(t, q)\|_{\mathcal{X}'} \leq C^{\Psi}$, *and* $\|\mathbf{A} q\|_{\mathcal{X}'} \leq C^{\mathcal{X}}$.

Proof. By Proposition 2.1 there exists $C_0^R > 0$ such that $\|q\|_{\mathcal{Q}} \leq C_0^R$ 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 $|\langle \eta, v \rangle| \leq C^{\Psi} \|v\|_{\mathcal{X}}$. Thus, we have $\eta \in \mathcal{X}' \subset \mathcal{Q}'$ and $\|\eta\|_{\mathcal{X}'} \leq C^{\Psi}$ 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

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

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

As a corollary, every solution of $(\mathcal{Q}, \mathcal{E}, \Psi, q^0)$ satisfies $\|\mathbf{A} q(t)\|_{\mathcal{X}'} \leq C^{\mathcal{X}}$ 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 $C_1^{\mathbf{P}}$ and a positive exponent α_1 such that for all $h \in (0,1]$, $v \in \mathcal{Q}$, and $v_h \in \mathcal{Q}_h$ we have

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

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

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. Before formulating the main theorem we give the typical situation we have in mind. Note that in Examples 3.2, 3.7, and 3.8 the derivative with respect to x is denoted by $(\cdot)'$.

Example 3.2. *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* $a(x) \ge \kappa > 0$ *for all* $x \in \Omega$ *and* $\theta \in (0,1]$. *For* $k \in \mathbb{N}$ *subdivide* Ω *into* k *subintervals of equal length* $h = 1/k$. *Then, we define* $\mathcal{Q}_h$ *as the set of continuous and piecewise affine functions on the corresponding intervals. Moreover, let* $\mathbf{P}_h$ *be the affine interpolant on the partition, namely,* $(\mathbf{P}_h q)'(x) = k \int_{I_j} q'(y) \,\mathrm{d}y$ *for* $x \in I_j \overset{\mathrm{def}}{=} \big((j-1)/k, j/k\big)$. *Then,* (3.3) *holds with exponent* $\alpha_1 = 1$.

Using all of the above assumptions we are now able to establish the approximation condition and hence control the space-time discretization error via Theorem 2.3.

Theorem 3.3. *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*

$$\|q_{\tau,h}(t) - q(t)\|_{\mathcal{Q}} \le C_*^{\mathrm{sl}} \big(h^{\alpha_1/2} + \sqrt{\tau} + \|q^0 - q_h^0\|_{\mathcal{Q}} \big) \quad \textit{for all } t \in [0,T], \tag{3.4}$$

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* $q_{\tau,h}^0 = q_h^0$.

The proof of this result follows directly from Theorem 2.2 if we establish the approximation condition (2.10). We shall establish this fact in the following proposition.

Proposition 3.4. *Assume* (3.1), (3.2), *and* (3.3). *Then, the approximation condition* (2.10) *holds with* $v_h = \mathbf{P}_h w$ *and* $\beta = \alpha_1$, *where* α_1 *is defined as in* (3.3b).

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 (3.3b) we have

$$\begin{aligned} &\|q\|_{\mathcal{Q}} \le C_0^R, \ \|q_h\|_{\mathcal{Q}} \le C_0^R, \ \|v_h - w\|_{\mathcal{X}} \le C_1^{\mathbf{P}} h^{\alpha_1} \|w\|_{\mathcal{Q}}, \\ &\|\mathbf{A} q_h\|_{\mathcal{X}'} \le C^{\mathcal{X}}, \ \|\mathbf{A} q\|_{\mathcal{X}'} \le C^{\mathcal{X}}. \end{aligned} \tag{3.5}$$

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

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

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

$$\begin{aligned} &\langle \mathrm{D}_q \mathcal{E}(t, q_h), v_h - w \rangle_{\mathcal{Q}} + \Psi(v_h - w) \\ &\le \langle \mathbf{A}(q_h - q), v_h - w \rangle_{\mathcal{Q}} + \big(C^{\mathcal{X}} + C_1^{\mathcal{H}} + C_0^{\boldsymbol{\ell}} + C^\Psi \big) C_1^{\mathbf{P}} h^{\alpha_1} \|w\|_{\mathcal{Q}}. \end{aligned} \tag{3.6}$$

In particular, the second term in the above right-hand side is as required in (2.10).

Hence, it remains to estimate the first term on the right-hand side of (3.6) by letting

$$\begin{aligned}\langle \mathbf{A}(q_h-q), v_h-w\rangle_{\mathcal{Q}} &\le \|\mathbf{A}(q_h-q)\|_{\mathcal{X}'}\|v_h-w\|_{\mathcal{X}}\\ &\le \big(\|\mathbf{A}q_h\|_{\mathcal{X}'}+\|\mathbf{A}q\|_{\mathcal{X}'}\big)\|v_h-w\|_{\mathcal{X}} \overset{(3.5)}{\le} 2C^{\mathcal{X}}C_1^{\mathbf{P}}h^{\alpha_1}\|w\|_{\mathcal{Q}}.\end{aligned}$$

This finishes the proof. ∎

3.2. Control of the initial error. We shall now complement the result of Theorem 3.3 by explicitly exhibiting a convergence rate for the *initial error* term $\|q^0-q_h^0\|_{\mathcal{Q}}$. To this aim, we start by strengthening the requirements on the linear operators $\mathbf{P}_h : \mathcal{Q} \to \mathcal{Q}'$ by asking for two positive constants $C_2^{\mathbf{P}}$, $C_3^{\mathbf{P}}$ and two positive exponents α_2, α_3 such that, in addition to (3.3), for all $h \in (0,1]$, $v \in \mathcal{Q}$, and $v_h \in \mathcal{Q}_h$, one has that

$$\|(\mathbf{P}_h^*\mathbf{A}-\mathbf{A}\mathbf{P}_h)v\|_{\mathcal{Q}'} \le C_2^{\mathbf{P}}h^{\alpha_2}\|v\|_{\mathcal{Q}}, \tag{3.7a}$$

$$\|(\mathbf{P}_h-\mathbf{I})v_h\|_{\mathcal{Q}} \le C_3^{\mathbf{P}}h^{\alpha_3}\|v_h\|_{\mathcal{Q}}. \tag{3.7b}$$

Note that, if $\mathbf{P}_h$ is a projection, then (3.7b) 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 (3.7a) holds with $C_2^{\mathbf{P}} = 0$ and any $\alpha_2 > 0$. Although both of our applications (sections 4–5) rely on Galerkin projections, we shall, however, keep the abstract discussion on the more general setting. This aim for generality is motivated by the observation that the use of Galerkin projectors could result in a weaker convergence result. An example of this circumstance is detailed in Example 3.7 below.

Let us now present a lemma that will be useful in what follows. 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 3.5. *Assume* (3.1b), (3.3)*, and* (3.7)*. 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*

$$\|(\mathbf{P}_h-\mathbf{I})q\|_{\mathcal{Q}} \le C_4^{\mathbf{P}} \max\big\{\big(h^{\alpha_1}\|q\|_{\mathcal{Q}}\|\mathbf{A}q\|_{\mathcal{X}'}\big)^{1/2}, h^{\alpha_2}\|q\|_{\mathcal{Q}}, h^{\alpha_3/2}\|q\|_{\mathcal{Q}}\big\}. \tag{3.8}$$

Proof. To estimate $\eta_h \overset{\text{def}}{=} \|(\mathbf{P}_h-\mathbf{I})q\|_{\mathcal{Q}}$ we employ $\mathbf{A}$ via (3.1b), (3.3), and (3.7). Using the abbreviation $R \overset{\text{def}}{=} \|q\|_{\mathcal{Q}}$ we obtain

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

where we used $y_1y_2 \le \frac{\kappa}{2}y_1^2 + \frac{1}{2\kappa}y_2^2$ in the last passage. Canceling the first term on the right-hand side, we have the desired estimate. ∎

We now present a possible choice for the 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

$$q_h^0 \overset{\text{def}}{=} \mathrm{Argmin}\{\mathcal{E}(0,\widehat{q}_h)+\Psi(\widehat{q}_h-\mathbf{P}_h q^0) \,|\, \widehat{q}_h \in \mathcal{Q}_h\}. \tag{3.9}$$

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

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

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), (3.3), *and* (3.7). *Then, there exists* $C_0^{\rm sl} > 0$ *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* (3.9) *satisfies*

$$\|q_h^0-q^0\|_{\mathcal{Q}} \le C_0^{\rm sl} h^{\beta/2}, \tag{3.10}$$

with $\beta = \min\{\alpha_1, 2\alpha_2, \alpha_3\}$*, where* α_i*,* $i=1,2,3$*, are defined as in* (3.3) *and* (3.7)*.*

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

$$\begin{aligned}\frac{\kappa}{2}\|q_h^0-q^0\|_{\mathcal{Q}}^2 &\le \mathcal{E}(0,q_h^0) - \mathcal{E}(0,q^0) + \Psi(q_h^0-q^0)\\ &\le \mathcal{E}(0,q_h^0) - \mathcal{E}(0,q^0) + \Psi(q_h^0 - \mathbf{P}_h q^0) + \Psi(\mathbf{P}_h q^0 - q^0)\\ &\le \mathcal{E}(0,\mathbf{P}_h q^0) - \mathcal{E}(0,q^0) + \Psi((\mathbf{P}_h-\mathbf{I})q^0),\end{aligned} \tag{3.11}$$

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

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

Thus using Taylor's formula, (3.2), and Proposition 3.1, we deduce from (3.11) that

$$\begin{aligned}\frac{\kappa}{2}\|q_h^0-q^0\|_{\mathcal{Q}}^2 &\le \mathcal{I}(q^0,\mathbf{P}_h q^0) + \langle \mathrm{D}_q\mathcal{E}(0,q^0), (\mathbf{P}_h-\mathbf{I})q^0\rangle_{\mathcal{Q}} + \Psi((\mathbf{P}_h-\mathbf{I})q^0)\\ &\le \mathcal{I}(q^0,\mathbf{P}_h q^0) + \|\mathrm{D}_q\mathcal{E}(0,q^0)\|_{\mathcal{X}'}\|(\mathbf{P}_h-\mathbf{I})q^0\|_{\mathcal{X}} + C^\Psi\|(\mathbf{P}_h-\mathbf{I})q^0\|_{\mathcal{X}}\\ &\le \mathcal{I}(q^0,\mathbf{P}_h q^0) + 2C^\Psi C_1^{\mathbf{P}} h^{\alpha_1}\|q^0\|_{\mathcal{Q}}.\end{aligned} \tag{3.12}$$

For $\mathcal{I}(q^0,\mathbf{P}_h q^0)$ we use that, by (3.3a) and (2.9a), we know $\|\mathbf{P}_h q^0\|_{\mathcal{Q}} \le C_0^{\mathbf{P}} C_0^R$. 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 $\mathcal{I}(q^0,\mathbf{P}_h q^0) \le \frac{C_2^{\mathcal{E}}}{2}\|(\mathbf{P}_h-\mathbf{I})q^0\|_{\mathcal{Q}}^2$. Since $q^0 \in \mathcal{S}(0)$, Proposition 3.1 yields $\|\mathbf{A}q^0\|_{\mathcal{X}'} \le C^{\mathcal{X}}$. Thus, Lemma 3.5 implies that there exists $C^{\mathcal{I}} > 0$ such that $\mathcal{I}(q^0,\mathbf{P}_h q^0) \le C^{\mathcal{I}} h^\beta$. Hence we infer from (3.12) the desired result. ∎

3.3. Discussion via examples. Let us now elaborate on Example 3.2 in order to motivate the generality of (3.7) by showing that the choice of Galerkin projectors for $\mathbf{P}_h$ may not be optimal.

EXAMPLE 3.7. *In the very same setting of Example* 3.2*, by letting* $\mathbf{P}_h$ *be the affine interpolant on the partition one can choose* $\alpha_2 = \theta$*, 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* 3.5 *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,* $\|(\mathbf{P}_h-\mathbf{I})q\|_{\mathrm{H}^1} \le Ch\|q\|_{\mathrm{H}^2}$*, while the lemma just gives the bound* $h^{1/2}$*.*

The use of the Galerkin projection $\mathbf{P}_h^{\rm Gal}$ *results indeed in a weaker convergence* *statement. The Galerkin projection is defined by letting* $\beta_j \stackrel{\text{def}}{=} \int_{I_j} a(y)\,\mathrm{d}y$ *for* $1 \le j \le k$ *as*

$$(\mathbf{P}_h^{\mathrm{Gal}} q)'(x) \stackrel{\text{def}}{=} \frac{1}{\beta_j} \int_{I_j} a(y)\, q'(y)\,\mathrm{d}y - C_h^{\mathrm{Gal}}(q) \quad \text{for } x \in I_j, \quad \text{where}$$

$$C_h^{\mathrm{Gal}}(q) \stackrel{\text{def}}{=} \left(\sum_{i=1}^k \frac{h}{\beta_i} \right)^{-1} \sum_{i=1}^k \frac{h}{\beta_i} \int_{I_i} a(y)\, q'(y)\,\mathrm{d}y$$

is the only (q-dependent) constant letting $(\mathbf{P}_h^{\mathrm{Gal}} q)(0) = (\mathbf{P}_h^{\mathrm{Gal}} q)(1) = 0$. *Indeed, one can check that*

$$\|(\mathbf{P}_h^{\mathrm{Gal}} - \mathbf{I})q\|_{\mathrm{L}^2} \le C_1^{\mathbf{P}^{\mathrm{Gal}}} h^\theta \|q\|_{\mathrm{H}^1}$$

where the constant $C_1^{\mathbf{P}^{\mathrm{Gal}}} > 0$ *scales with the square root of the oscillation of* a. *As we have that* $C_0^{\mathbf{P}^{\mathrm{Gal}}} = 1$, $C_2^{\mathbf{P}^{\mathrm{Gal}}} = C_3^{\mathbf{P}^{\mathrm{Gal}}} = 0$, $\alpha_1 = \theta$, *and* $\alpha_2, \alpha_3 > 0$ *are arbitrary, by using the Galerkin projector* $\mathbf{P}_h^{\mathrm{Gal}}$, *Lemma* 3.5 *provides the rate* $h^{\theta/2}$, *whereas in the case of the affine interpolant* $\mathbf{P}_h$ *we have the stronger* $h^{\min\{1/2,\theta\}}$.

Finally, we notice that the power of h in (3.3b) and (3.7a) depends on the choice of $\mathcal{X}$. Of course, the optimal choice is to make $\mathcal{X}$ as big as allowable by the condition (3.2) for Ψ. We illustrate this in the following, which gives our first example of convergence rates for space-time discretizations.

EXAMPLE 3.8. *We consider the situation of Example* 3.2 *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*

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

Here "Sign" *denotes the multivalued signum function with* $\mathrm{Sign}(0) = [-1,1]$.

As in Example 3.2, *the subspaces* $\mathcal{Q}_h$ *contain the piecewise affine functions on an equidistant partition of* $\Omega = (0,1)$, *and the piecewise affine interpolants* $\mathbf{P}_h : \mathcal{Q} \to \mathcal{Q}_h$ *coincide with 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* (3.4) *in Theorem* 3.3 *gives the 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 strain gradient plasticity.

4.1. Strain gradient plasticity. We shall start by briefly recalling classical strain gradient plasticity. The reader is referred to [23, 24, 26, 27] for additional details and motivation and to [37] for convergence rates under natural regularity for classical linearized elastoplasticity with linear kinematic hardening (no plastic gradient).

We consider a material with a reference configuration $\Omega \subset \mathbb{R}^d$ with $d \in \{2,3\}$, where Ω is an open bounded set with Lipschitz boundary. This body may undergo elastic and plastic deformation. We shall denote by $u : \Omega \to \mathbb{R}^d$ the body displacement. The linearized strain tensor is classically given by $\mathbf{e}(u) \stackrel{\text{def}}{=} \frac{1}{2}(\nabla u + \nabla u^\top) \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, where $\mathbb{R}^{d\times d}_{\rm sym}$ is the space of symmetric $d{\times}d$ tensors endowed with the scalar product $v{:}w \stackrel{\rm def}{=} \operatorname{tr}(v^{\mathsf{T}}w)$ and the corresponding norm $|v|^2 \stackrel{\rm def}{=} v{:}v$ for all $v, w \in \mathbb{R}^{d\times d}_{\rm sym}$. Here $(\cdot)^{\mathsf{T}}$ and $\operatorname{tr}(\cdot)$ denote the transpose and the trace of the tensor, respectively. The strain $\mathbf{e}(u)$ is additively decomposed into $\mathbf{e}(u) = \epsilon^{\rm el} + z$, where $\epsilon^{\rm el} : \Omega \to \mathbb{R}^{d\times d}_{\rm sym}$ is the elastic strain, whereas $z : \Omega \to \mathbb{R}^{d\times d}_{\rm dev}$ stands for the plastic strain. Here, $\mathbb{R}^{d\times d}_{\rm dev}$ is the space of symmetric $d \times d$ tensors with vanishing trace.

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.,

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

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

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

where we assume that $\Gamma_{\rm Neu} \stackrel{\rm def}{=} \partial\Omega \setminus \Gamma_{\rm Dir}$ has a Lipschitz $d{-}2$ boundary in $\partial\Omega$. 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 $i(u, z) \stackrel{\rm def}{=} (u, \gamma u, z)$, where $\gamma : \mathrm{H}^1(\Omega) \to \mathrm{L}^2(\Gamma_{\rm Neu})$ is the standard trace operator, is continuous and dense. Hence, upon identifying

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

we have that $(\mathcal{Q}, \mathcal{X}, \mathcal{Q}')$ forms a classical Gelfand triplet.

We will denote the states by $q \stackrel{\rm def}{=} (u, z)$. 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 the reader to [40] or [21].

The *stored-energy potential* takes the form

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

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 as

$$W(x, \mathbf{e}(u)(x), z(x)) \stackrel{\rm def}{=} \frac{1}{2}\big((\mathbf{e}(u)(x){-}z(x)){:}\mathbb{C}(\mathbf{e}(u)(x){-}z(x))\big) + \frac{c_2}{2}|z(x)|^2.$$

In the latter, $\mathbb{C}$ is the elastic tensor and $c_2 > 0$ is a hardening modulus. 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

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

where f_{appl} and g_{appl} are given body forces and a surface traction on Γ_{Neu}.

The plastic flow rule is enforced by means of the definition of the dissipation potential

$$\psi(v) \stackrel{\mathrm{def}}{=} \int_\Omega \rho |v(x)| \,\mathrm{d}x, \quad \text{where } \rho > 0. \tag{4.4}$$

The material constitutive relation reads as the *differential inclusion*

$$\begin{pmatrix} 0 \\ \partial\psi(\dot z) \end{pmatrix} + \begin{pmatrix} \partial_u \mathcal{E}(t,q) \\ \partial_z \mathcal{E}(t,q) \end{pmatrix} \ni \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{4.5}$$

where $\partial_u \mathcal{E}(t,q) = -\mathrm{div}\,(\mathbb{C}(\mathbf{e}(u){-}z)) - l(t)$, $\partial_z \mathcal{E}(t,q) = -\mathbb{C}(\mathbf{e}(u){-}z) + c_2 z - \nu\Delta z$. Hence, the first component provides the elastic equilibrium equations, whereas the second component gives the plastic flow law.

By letting $q = (u,z)$, $\Psi(\dot q) = \psi(\dot z)$, and $\langle \boldsymbol{\ell}(t), q\rangle_{\mathcal{Q}} = \langle l(t), u\rangle_{\mathcal{X}_{\mathcal{F}}}$, system (4.5) can be rewritten in the abstract form

$$\partial\Psi(\dot q) + \mathbf{A}q - \boldsymbol{\ell}(t) \ni 0, \tag{4.6}$$

where

$$\mathbf{A} \stackrel{\mathrm{def}}{=} \begin{pmatrix} -\mathrm{div}(\mathbb{C}\mathbf{e}(\cdot)) & \mathrm{div}\,(\mathbb{C}(\cdot)) \\ -\mathbb{C}\mathbf{e}(\cdot) & \mathbb{C}(\cdot) - \nu\Delta(\cdot) + c_2 I(\cdot) \end{pmatrix}, \tag{4.7}$$

where I is the identity in $\mathcal{Z}$. Here we assume that the elasticity tensor $\mathbb{C}$ is a symmetric positive definite map, i.e.,

$$\exists \mu > 0 \ \forall e \in \mathbb{R}^{d\times d}_{\mathrm{sym}}: \ e{:}\mathbb{C}{:}e \geq \mu |e|^2. \tag{4.8}$$

By assuming $f_{\mathrm{appl}} \in C^3([0,T]; \mathrm{L}^2(\Omega;\mathbb{R}^d))$ and $g_{\mathrm{appl}} \in C^3([0,T]; \mathrm{H}^\zeta_0(\Gamma_{\mathrm{Neu}};\mathbb{R}^d))$ in (4.3), we readily check that $\boldsymbol{\ell} \in C^3([0,T]; \mathcal{X}')$ (see (3.1d)). By letting $\mathcal{H} = 0$ and using (4.8), we clearly have that (3.1) is satisfied.

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, [12, 17]). Moreover, for the sake of definiteness we require that each face of $\partial\Omega$ is contained either in Γ_{Dir} or in Γ_{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, Γ_{Dir}, and $\mathbb{C}$ satisfy the following condition:

$$\begin{aligned} &\exists \sigma \in (0,1] \ \exists \widetilde{C} > 0 \ \forall f \in \mathcal{X}'_{\mathcal{F}} \ \forall g \in \mathrm{L}^2(\Omega): \\ &\|u_f\|_{\mathrm{H}^{1+\sigma}(\Omega;\mathbb{R}^d)} \leq \widetilde{C}\|f\|_{\mathcal{X}'_{\mathcal{F}}} \ \text{ and } \ \|\zeta_g\|_{\mathrm{H}^{1+\sigma}(\Omega)} \leq \widetilde{C}\|g\|_{\mathrm{L}^2(\Omega)}, \end{aligned} \tag{4.9}$$

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 [16, Thm. 2.2-4, p. 99] or a convex polyhedron; see [25] for the 2D case and [18, 22] for the 3D case. Nonconvex polyhedrons can also be considered (possibly with $\sigma < 1$) and results for the mixed Neumann–Dirichlet conditions are also available [22]. Additional details on regularity issues and asymptotic developments of solutions near corner points may be found in [39, 19, 48, 36], among others.

Let us start from the following lemma which is crucial to obtaining the error estimates for space-time discretizations in strain gradient plasticity. 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* (4.9) *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*

$$\|q\|_{\mathrm{H}^{1+\sigma}(\Omega;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\rm dev})} \leq C_1^\mathcal{X} \|f\|_{\mathcal{X}'}, \tag{4.10}$$

where $\sigma \in (0, 1]$ *is defined in* (4.9).

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

$$\|q\|_\mathcal{Q} \leq \|f\|_{\mathcal{Q}'}/\kappa \leq C_1 \|f\|_{\mathcal{X}'}. \tag{4.11}$$

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

$$\forall v \in \mathcal{F} : \int_\Omega \mathbb{C}\mathbf{e}(v){:}\mathbf{e}(u)\,\mathrm{d}x = \int_\Omega (-\mathrm{div}(\mathbb{C}z)){\cdot}v\,\mathrm{d}x + \langle f_1, v\rangle_{\mathcal{X}_\mathcal{F}}, \tag{4.12}$$

$$\forall w \in \mathcal{Z} : \int_\Omega (c_2 w{:}z + \nu\nabla w{:}\nabla z)\,\mathrm{d}x = \int_\Omega \big(f_2 + \mathbb{C}(\mathbf{e}(u) - z)\big){:}w\,\mathrm{d}x. \tag{4.13}$$

Using (4.11), 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, (4.13) consists of decoupled Neumann problems for the components of z. Thus, employing (4.9) we deduce

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

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 we are given a regular triangulation $\{\mathcal{T}_k\}$ of Ω (cf. [52]) 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 $\mathcal{Q}_h \stackrel{\rm def}{=} \mathcal{F}_h \times \mathcal{Z}_h$ and assume we are given linear projectors $\mathbf{\Pi}_h : \mathcal{Q} \to \mathcal{Q}_h$ fulfilling

$$\forall s \in (0, 1]\ \exists C^\mathbf{\Pi} > 0\ :\ \|(\mathbf{\Pi}_h - \mathbf{I})q\|_\mathcal{Q} \leq C^\mathbf{\Pi} h^s \|q\|_{\mathrm{H}^{1+s}(\Omega;\mathbb{R}^d\times\mathbb{R}^{d\times d}_{\rm dev})}. \tag{4.14}$$

The latter can be realized, for instance, by letting $\mathbf{\Pi}_h$ be the L^2 orthogonal projector. The interpolation error control of (4.14) is well known for $s = 1$ [16] and follows from [34, Lemma 5.6] for $s \in (0, 1)$. Let us explicitly remark that the quasi-uniformity of the mesh is not needed here.

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 $\mathbf{P}_h q \stackrel{\rm def}{=} \widehat{q}_h$, where $\widehat{q}_h \in \mathcal{Q}_h$ is the unique solution of

$$\langle \mathbf{A}\widehat{q}_h, p_h\rangle_\mathcal{Q} = \langle \mathbf{A}q, p_h\rangle_\mathcal{Q} \forall p_h \in \mathcal{Q}_h. \tag{4.15}$$

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

THEOREM 4.2. *Assume that* (4.9) *holds. Then there exists* $C_*^{\rm plast} > 0$ *such that for any* $h \in (0,1]$, $q^0 \in \mathcal{S}(0)$, *and any partition* Π^τ *of* $[0,T]$, *we have*

$$\|q_{\tau,h}(t) - q(t)\|_{\mathcal{Q}} \leq C_*^{\rm plast}\big(h^{\sigma/2} + \sqrt{\tau}\big) \quad \textit{for all } t \in [0,T], \tag{4.16}$$

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

Proof. By the definition (4.15) we have $\mathbf{P}_h \circ \mathbf{P}_h = \mathbf{P}_h$, and (3.7b) holds for any $\alpha_3 \geq 0$. Moreover, by using (4.15) we readily check that, for all p, $q \in \mathcal{Q}$,

$$\begin{aligned}
\langle(\mathbf{P}_h^* \mathbf{A} - \mathbf{A}\mathbf{P}_h)q, p\rangle_{\mathcal{Q}} &= \langle \mathbf{A}q, \mathbf{P}_h p\rangle_{\mathcal{Q}} - \langle \mathbf{A}\mathbf{P}_h q, p\rangle_{\mathcal{Q}} \\
&\overset{(4.15)}{=} \langle \mathbf{A}q, \mathbf{P}_h p\rangle_{\mathcal{Q}} - \langle \mathbf{A}\mathbf{P}_h q, \mathbf{P}_h p\rangle_{\mathcal{Q}} = \langle \mathbf{A}(q - \mathbf{P}_h q), \mathbf{P}_h p\rangle_{\mathcal{Q}} \overset{(4.15)}{=} 0.
\end{aligned}$$

Hence, (3.7a) 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

$$\kappa\|\mathbf{P}_h q\|_{\mathcal{Q}}^2 \leq \langle \mathbf{A}\mathbf{P}_h q, \mathbf{P}_h q\rangle_{\mathcal{Q}} \overset{(4.15)}{=} \langle \mathbf{A}q, \mathbf{P}_h q\rangle_{\mathcal{Q}} \leq \|\mathbf{A}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')}\|q\|_{\mathcal{Q}}\|\mathbf{P}_h q\|_{\mathcal{Q}}.$$

Finally, let us check for property (3.3b) by means of the classical duality technique by Aubin [3] and Nitsche [49]. 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

$$\begin{aligned}
\|(\mathbf{P}_h - \mathbf{I})q\|_{\mathcal{X}}^2 &= \langle J^{\mathcal{X}}(\mathbf{P}_h - \mathbf{I})q, (\mathbf{P}_h - \mathbf{I})q\rangle_{\mathcal{X}} = \langle \mathbf{A}\varphi, (\mathbf{P}_h - \mathbf{I})q\rangle_{\mathcal{Q}} \\
&= \langle \mathbf{A}(\mathbf{P}_h - \mathbf{I})q, \varphi\rangle_{\mathcal{Q}} \overset{(4.15)}{=} \langle \mathbf{A}(\mathbf{P}_h - \mathbf{I})q, \varphi - \varphi_h\rangle_{\mathcal{Q}} \leq C_5^{\mathbf{P}}\|q\|_{\mathcal{Q}}\|\varphi - \varphi_h\|_{\mathcal{Q}},
\end{aligned}$$

where $C_5^{\mathbf{P}} \overset{\mathrm{def}}{=} \|\mathbf{A}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q}')} \sup_{h\in(0,1]} \|\mathbf{P}_h - \mathbf{I}\|_{\mathrm{Lin}(\mathcal{Q},\mathcal{Q})}$. Choosing $\varphi_h = \mathbf{\Pi}_h\varphi$ and exploiting (4.14) for $s = \sigma$ with σ from (4.9) 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 that

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

which is the desired approximation result (3.3b) with $\alpha_1 = \sigma$. Hence, applying Theorem 3.3 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* (4.16) *with* $\sigma = 1$. *Indeed, as already mentioned, in this case we have* H^2 *regularity for the auxiliary problems* (4.9). *Correspondingly, our convergence result in* (4.16) *gives the order* $\mathcal{O}(\sqrt{h} + \sqrt{\tau})$.

5. Application to the isothermal Souza–Auricchio model.

5.1. The isothermal Souza–Auricchio model. The application to strain gradient plasticity can be extended to the *doubly* nonlinear situation of the isothermal SMA. Let us recall the model here by referring the reader to the original papers [54, 6, 5, 7] for additional comments and details.

The evolution of the shape-memory body will be determined by its displacement $u : \Omega \to \mathbb{R}^d$ and suitable martensitic phase transformation. The latter will be characterized by a mesoscopic internal variable $z : \Omega \to \mathbb{R}^{d\times d}_{\mathrm{dev}}$. In particular, the tensor z stands again as the inelastic part of the strain. Still, no plastic evolution occurs and z is completely related to recoverable martensitic phase transformation.

Unless otherwise stated, notation and assumptions are the same as in section 4. In particular, the *stored-energy potential* for the isothermal Souza–Auricchio model is

$$\mathcal{E}(t,u,z) \stackrel{\mathrm{def}}{=} \int_\Omega \Big(W(x,\mathbf{e}(u)(x),z(x)) + \frac{\nu}{2}|\nabla z(x)|^2 \Big)\,\mathrm{d}x - \langle l(t), u\rangle_{\mathcal{X}_\mathcal{F}}, \tag{5.1}$$

where nevertheless the stored-energy density reads

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

Here $\widehat{H} : \mathbb{R}^{d\times d}_{\mathrm{dev}} \to \mathbb{R}$ represents a nonquadratic hardening potential. In [54, 9, 6], the authors are interested in $\widehat{H} = H_{\mathrm{SoAu}}$ with

$$H_{\mathrm{SoAu}}(z) \stackrel{\mathrm{def}}{=} c_1\sqrt{\delta^2 + |z|^2} + \frac{c_2}{2}|z|^2 + \frac{((|z| - c_3)_+)^4}{\delta(1+|z|^2)}, \tag{5.2}$$

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 (5.2) and $\nu \to 0$ in (4.2). More precisely, $\widehat{H} = H_{\mathrm{org}}$ is defined as

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

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

To model the hysteretic behavior of shape-memory materials, we define the dissipation potential as in (4.4). Correspondingly, the material constitutive relation (4.5) is now *doubly nonlinear*. In particular, system (4.5) can be rewritten in the abstract form

$$\partial\Psi(\dot q) + \mathbf{A}q + \mathrm{D}_q\mathcal{H}(q) - \boldsymbol{\ell}(t) \ni 0, \tag{5.3}$$

where $\mathcal{H}(q) \stackrel{\mathrm{def}}{=} \int_\Omega H(u(x),z(x))\,\mathrm{d}x$ with $H(u,z) = \widehat{H}(z) - \frac{c_2}{2}|z|^2$. We may prove that the functional $\mathcal{H}$ built on $H = H_{\mathrm{SoAu}}$ satisfies (3.1c), and thus (3.1a)–(3.1d) are satisfied. This is, however, not the case for the original model with H_{org}; the reader is referred to [4] for some discussion on the limit $(\nu,\delta) \to (0,0)$.

Existence and uniqueness results for a temperature-dependent variant of (4.6) were obtained in [44]. Following [9] a function $\widehat{H}(z,\theta) = H_{\mathrm{SoAu}}(z,\theta)$ is considered by allowing the constants $c_i(\theta)$, $i = 1,2,3$, in (5.2) 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 that the temperature is 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.

As for the domain Ω and the space discretization, we make the very same assumptions of subsection 4.2. In particular, we ask for (4.9), namely, that the auxiliary problems (4.12)–(4.13) have $\mathrm{H}^{1+\sigma}$ regularity. Arguing exactly as in subsection 4.2 we have the following.

THEOREM 5.1. *Assume that* (4.9) *holds. Then there exists* $C_*^{\mathrm{SoAu}} > 0$ *such that for any* $h \in (0,1]$, $q^0 \in \mathcal{S}(0)$, *and any partition* Π^τ *of* $[0,T]$, *we have*

$$\|q_{\tau,h}(t) - q(t)\|_{\mathcal{Q}} \le C_*^{\mathrm{SoAu}}\big(h^{\sigma/2} + \sqrt{\tau}\big) \text{ for all } t \in [0,T], \tag{5.4}$$

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

Once again, the case of a convex reference domain Ω for $\Gamma_{\mathrm{Neu}} = \emptyset$ entails that (5.4) holds with $\sigma = 1$.

Appendix. The aim of this section is to give the proof of (2.9c). We follow the ideas developed in [47] 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

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

(see Theorem 2.2).

Now let the partition $\Pi^\tau \overset{\mathrm{def}}{=} \{0 = t_0^\tau < t_1^\tau < \cdots < t_{k_\tau}^\tau = T\}$ be given, and define Π^{τ_j} by successive bisections, namely,

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

We shall associate with 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 $\bar{t}_k^\tau \overset{\mathrm{def}}{=} \max\{s_n^\tau \in \Pi^{\tau_j} \,|\, s_n^\tau \le t_k^\tau\}$, $\mathcal{E}^1(t_k^\tau,q) \overset{\mathrm{def}}{=} \mathcal{E}(\bar{t}_k^\tau,q)$, and $\mathcal{E}^2(t_k^\tau,q) \overset{\mathrm{def}}{=} \mathcal{E}(t_k^\tau,q)$ 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 notation:

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

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

$$\forall v_h \in \mathcal{Q}_h :\ \langle \mathrm{D}_q\mathcal{E}^j(t_k^\tau, q_{\tau,h}^{j,k}), v_h - \eta_k q_{\tau,h}^j\rangle_{\mathcal{Q}} + \Psi(v_h) - \Psi(\eta_k q_{\tau,h}^j) \ge 0. \tag{A.1}$$

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

$$\langle \mathrm{D}_q\mathcal{E}^1(t_k^\tau, q_{\tau,h}^{1,k}) - \mathrm{D}_q\mathcal{E}^2(t_k^\tau, q_{\tau,h}^{2,k}), \eta_k q_{\tau,h}^1 - \eta_k q_{\tau,h}^2\rangle_{\mathcal{Q}} \le 0. \tag{A.2}$$

Define

(A.3)
$$\gamma^k_{\tau,h} \overset{\text{def}}{=} \langle \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h}) - \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h}), q^{1,k}_{\tau,h} - q^{2,k}_{\tau,h}\rangle_{\mathcal{Q}} \geq \kappa \|q^{1,k}_{\tau,h} - q^{2,k}_{\tau,h}\|^2_{\mathcal{Q}} = \kappa \|e^k_{\tau,h}\|^2_{\mathcal{Q}}.$$

Let us estimate the increment

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

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

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

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

(A.4)
$$\begin{aligned}&\langle \eta_k(\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h}) - \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h})), e^{k-1}_{\tau,h}\rangle_{\mathcal{Q}} - \langle \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{1,k}_{\tau,h}) - \mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h}), \eta_k e_{\tau,h}\rangle_{\mathcal{Q}} \\ &= \langle A_k e^k_{\tau,h} - A_{k-1} e^{k-1}_{\tau,h}, e^{k-1}_{\tau,h}\rangle_{\mathcal{Q}} - \langle A_k e^k_{\tau,h}, \eta_k e_{\tau,h}\rangle_{\mathcal{Q}} \\ &= -\langle A_k \eta_k e_{\tau,h}, \eta_k e_{\tau,h}\rangle_{\mathcal{Q}} + \langle (A_k - A_{k-1}) e^{k-1}_{\tau,h}, e^{k-1}_{\tau,h}\rangle_{\mathcal{Q}}.\end{aligned}$$

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 (A.2), it follows that

$$\text{(A.5)}\qquad \begin{aligned}\eta_k\gamma_{\tau,h} &\leq C^{\mathcal{E},R}\big(|t^\tau_k - t^\tau_{k-1}| + \|\eta_k q^1_{\tau,h}\|_{\mathcal{Q}} + \|\eta_k q^2_{\tau,h}\|_{\mathcal{Q}}\big)\|e^{k-1}_{\tau,h}\|^2_{\mathcal{Q}} \\ &+ 2\|\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q^{2,k}_{\tau,h}) - \mathrm{D}_q\mathcal{E}^2(t^\tau_k, q^{2,k}_{\tau,h})\|_{\mathcal{Q}'}\|\eta_k e_{\tau,h}\|_{\mathcal{Q}}.\end{aligned}$$

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

$$\text{(A.6)}\qquad \forall j = 1,2: \ \|\eta_k q^j_{\tau,h}\|_{\mathcal{Q}} \leq C^R_1 |t^\tau_k - t^\tau_{k-1}|,$$

where $C^R_1 > 0$ is independent of h and τ (cf. Theorem 2.2). Carrying (A.6) and (A.3) in (A.5), 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

$$\eta_k\gamma_{\tau,h} \leq \frac{C^{\mathcal{E},R}}{\kappa}(1 + 2C^R_1)\gamma^{k-1}_{\tau,h}|t^\tau_k - t^\tau_{k-1}| + 4\rho C^R_1 |t^\tau_k - t^\tau_{k-1}|,$$

where

$$\rho \overset{\text{def}}{=} \max_{t^\tau_k \in \Pi^{\tau_{j+1}}} \sup_{q \in \mathcal{B}_{C^R_0}} \|\mathrm{D}_q\mathcal{E}^1(t^\tau_k, q) - \mathrm{D}_q\mathcal{E}^2(t^\tau_k, q)\|_{\mathcal{Q}'}.$$

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

$$\gamma_{\tau,h}^k \le \gamma_{\tau,h}^{k-1}\big(1+C_4(t_k^\tau - t_{k-1}^\tau)\big) + \rho C_4(t_k^\tau - t_{k-1}^\tau).$$

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

$$\gamma_{\tau,h}^k \le C_4\rho \sum_{k=1}^n (t_k^\tau - t_{k-1}^\tau) \prod_{j=k+1}^n \big(1+C_4(t_j^\tau - t_{j-1}^\tau)\big) \le C_4\rho \mathrm{e}^{C_4 T} T.$$

Using (A.3), it follows that

$$\|q_{\tau,h}^{1,k} - q_{\tau,h}^{2,k}\|_{\mathcal{Q}}^2 \le \frac{C_4 \mathrm{e}^{C_4 T} T}{\kappa}\rho. \tag{A.7}$$

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

$$\rho \le C_5 \max_{t_k^\tau \in \Pi^{\tau_{j+1}}} (t_k^\tau - t_{k-1}^\tau) \le C_5 2^{-j}\tau,$$

which implies that

$$\forall t \in [0,T]: \ \|q_{\tau_{j+1},h}(t) - q_{\tau_j,h}(t)\|_{\mathcal{Q}} \le C_6 2^{-j/2}\sqrt{\tau}, \quad \text{where } C_6 = \sqrt{\frac{C_4 T \mathrm{e}^{C_4 T}}{\kappa} C_5}.$$

Note that $(q_{\tau_j,h}(t))_{j\in\mathbb{N}}$ is a Cauchy sequence whose 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

$$\forall t \in [0,T]: \ \|q_{\tau,h}(t) - q_h(t)\|_{\mathcal{Q}} \le \sum_{j=0}^{\infty} C_6 2^{-j/2}\sqrt{\tau} \le 4C_6\sqrt{\tau}, \tag{A.8}$$

which proves (2.9c). ∎

Acknowledgments. L.P. and U.S. wish to thank WIAS for their kind hospitality.

REFERENCES

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

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

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

[4] 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 (2008), pp. 125–164.

[5] 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), pp. 1255–1284.

[6] 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 Engrg., 61 (2004), pp. 716–737.

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

[8] 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), pp. 1631–1637.

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

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

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

[12] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Appl. Math. 15, Springer-Verlag, New York, 2008.

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

[14] 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 (2006), pp. 2574–2598.

[15] 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 (2006), pp. 1851–1887.

[16] P. G. Ciarlet, *Élasticité tridimensionnelle*, Rech. Math. Appl. [Research in Applied Mathematics] 1, Masson, Paris, 1986.

[17] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics in Appl. Math. 40, SIAM, Philadelphia, 2002. Reprint of the 1978 original [North–Holland, Amsterdam].

[18] 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 (1988), pp. 795–818.

[19] M. Dauge, *Elliptic Boundary Value Problems on Corner Domains*, Lecture Notes in Math. 1341, Springer-Verlag, Berlin, 1988.

[20] 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 (2007), pp. 3881–3897.

[21] G. Duvaut and J.-L. Lions, *Inequalities in Mechanics and Physics*, Grundlehren Math. Wiss. 219, Springer-Verlag, Berlin, 1976.

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

[23] N. A. Fleck and J. W. Hutchinson, *Strain gradient plasticity*, Adv. Appl. Mech., 33 (1997), pp. 295–361.

[24] N. A. Fleck and J. W. Hutchinson, *A reformulation of strain gradient plasticity*, J. Mech. Phys. Solids, 49 (2001), pp. 2245–2271.

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

[26] P. Gudmundson, *A unified treatment of strain gradient plasticity*, J. Mech. Phys. Solids, 52 (2004), pp. 1379–1406.

[27] M. E. Gurtin and L. Anand, *A theory of strain-gradient plasticity for isotropic, plastically irrotational materials.* I. *Small deformations*, J. Mech. Phys. Solids, 53 (2005), pp. 1624–1649.

[28] W. Han, S. Jensen, and B. D. Reddy, *Numerical approximations of problems in plasticity: Error analysis and solution algorithms*, Numer. Linear Algebra Appl., 4 (1997), pp. 191–204.

[29] W. Han and B. D. Reddy, *Plasticity (Mathematical Theory and Numerical Analysis)*, Interdiscip. Appl. Math. 9, Springer-Verlag, New York, 1999.

[30] W. Han and B. D. Reddy, *Convergence of approximations to the primal problem of plasticity under conditions of minimal regularity*, Numer. Math., 87 (2000), pp. 283–315.

[31] W. Han, B. D. Reddy, and G. C. Schroeder, *Qualitative and numerical analysis of quasi-static problems in elastoplasticity*, SIAM J. Numer. Anal., 34 (1997), pp. 143–177.

[32] I. Hlaváček, *A finite element solution for plasticity with strain-hardening*, RAIRO Anal. Numer., 14 (1980), pp. 347–368.

[33] I. Hlaváček, J. Haslinger, J. Nečas, and J. Lovíšek, *Solution of Variational Inequalities in Mechanics*, Springer-Verlag, New York, 1988.

[34] 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 (2005), pp. 727–753.
[35] C. Johnson, *On finite element methods for plasticity problems*, Numer. Math., 26 (1976), pp. 79–84.
[36] D. Knees, *Global regularity of the elastic fields of a power-law model on Lipschitz domains*, Math. Methods Appl. Sci., 29 (2006), pp. 1363–1391.
[37] D. Knees, *On global spatial regularity and convergence rates for time dependent elasto-plasticity*, Math. Models Methods Appl. Sci., to appear.
[38] 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 (2005), pp. 389–418.
[39] V. A. Kondrat'ev, *Boundary value problems for elliptic equations in domains with conical or angular points*, Trudy Moskov. Mat. Obšč., 16 (1967), pp. 209–292.
[40] V. Kondrat'ev and O. Oleinik, *Boundary-value problems for the system of elasticity theory in unbounded domains. Korn's inequalities*, Russian Math. Surveys, 43 (1988), pp. 65–119.
[41] 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 (1996), pp. 809–842.
[42] Y. Li and I. Babuška, *A convergence analysis of an h-version finite-element method with high-order elements for two-dimensional elasto-plasticity problems*, SIAM J. Numer. Anal., 34 (1997), pp. 998–1036.
[43] A. Mielke, L. Paoli, and A. Petrov, *On the existence and approximation for a* 3*D model of thermally induced phase transformations in shape-memory alloys*, SIAM J. Math. Anal., 41 (2009), pp. 1388–1414.
[44] A. Mielke and A. Petrov, *Thermally driven phase transformation in shape-memory alloys*, Adv. Math. Sci. Appl., 17 (2007), pp. 667–685.
[45] A. Mielke and T. Roubíček, *Numerical approaches to rate-independent processes and applications in inelasticity*, M2AN Math. Model. Numer. Anal., 43 (2009), pp. 399–428.
[46] A. Mielke and R. Rossi, *Existence and uniqueness results for a class of rate-independent hysteresis problems*, Math. Models Methods Appl. Sci., 17 (2007), pp. 81–123.
[47] A. Mielke and F. Theil, *On rate-independent hysteresis models*, NoDEA Nonlinear Differential Equations Appl., 11 (2004), pp. 151–189.
[48] 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 (1992), pp. 327–361.
[49] J. Nitsche, *Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens*, Numer. Math., 11 (1968), pp. 346–348.
[50] R. H. Nochetto, G. Savaré, and C. Verdi, *A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations*, Comm. Pure Appl. Math., 53 (2000), pp. 525–589.
[51] 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 (2004), pp. 1595–1631.
[52] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Ser. in Comput. Math. 23, Springer-Verlag, Berlin, 1994.
[53] J. Rulla, *Error analysis for implicit approximations to solutions to Cauchy problems*, SIAM J. Numer. Anal., 33 (1996), pp. 68–87.
[54] A. Souza, E. Mamiya, and N. Zouain, *Three-dimensional model for solids undergoing stress-induced phase transformations*, Eur. J. Mech. A Solids, 17 (1998), pp. 789–806.