

**OPTIMAL CONTROL OF QUASISTATIC PLASTICITY WITH
LINEAR KINEMATIC HARDENING
PART I: EXISTENCE AND DISCRETIZATION IN TIME**

GERD WACHSMUTH

ABSTRACT. In this paper we consider the optimal control problem governed by a time-dependent variational inequality arising in quasistatic plasticity with linear kinematic hardening. In this first part of the paper we address certain continuity properties of the forward operator, which implies the existence of an optimal control. Moreover, a discretization in time is derived and we show that every local minimizer of the continuous problem can be approximated by minimizers of modified, time-discrete problems.

1 Introduction

The optimization of elastoplastic systems is of significant importance for industrial deformation processes, e.g., for the control of the springback of deep-drawn metal sheets. In this paper we consider an optimal control problem for the quasistatic problem of small-strain elastoplasticity with linear kinematic hardening. The strong formulation of the forward system reads (cf. [Han and Reddy, 1999, Chapters 2, 3])

$$\left. \begin{aligned}
 \mathbb{C}^{-1}\dot{\boldsymbol{\sigma}} - \boldsymbol{\varepsilon}(\dot{\mathbf{u}}) + \lambda(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D) &= \mathbf{0} && \text{in } (0, T) \times \Omega, \\
 \mathbb{H}^{-1}\dot{\boldsymbol{\chi}} + \lambda(\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D) &= \mathbf{0} && \text{in } (0, T) \times \Omega, \\
 \operatorname{div} \boldsymbol{\sigma} &= -\mathbf{f} && \text{in } (0, T) \times \Omega, \\
 \text{with complementarity } 0 \leq \lambda \perp \phi(\boldsymbol{\Sigma}) \leq 0 &&& \text{in } (0, T) \times \Omega, \\
 \text{and boundary conditions } \mathbf{u} &= \mathbf{0} && \text{on } (0, T) \times \Gamma_D, \\
 \boldsymbol{\sigma} \cdot \mathbf{n} &= \mathbf{g} && \text{on } (0, T) \times \Gamma_N.
 \end{aligned} \right\} \quad (1.1)$$

The system (1.1) is subject to the initial condition $(\boldsymbol{\Sigma}(0), \mathbf{u}(0)) = (\mathbf{0}, \mathbf{0})$. The state variables consist of the stress $\boldsymbol{\sigma}$ and back stress $\boldsymbol{\chi}$, whose values are symmetric matrices. Both matrix functions are combined into the generalized stress $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi})$. The state variables also comprise the vector-valued displacement \mathbf{u} and the scalar-valued plastic multiplier λ associated with the yield condition $\phi(\boldsymbol{\Sigma}) \leq 0$ of von Mises type. The tensors \mathbb{C}^{-1} and \mathbb{H}^{-1} are the inverses of the elasticity tensor and the hardening modulus, respectively, and $\boldsymbol{\sigma}^D$ denotes the deviatoric part of $\boldsymbol{\sigma}$. As usual, $\boldsymbol{\varepsilon}(\mathbf{v}) = (\nabla \mathbf{v} + (\nabla \mathbf{v})^\top)/2$ denotes the (linearized) strain, where $\nabla \mathbf{v}$ is the gradient of \mathbf{v} . By $\operatorname{div} \boldsymbol{\sigma}$ we denote the (row-wise) divergence of the matrix-valued function $\boldsymbol{\sigma}$. The boundary conditions correspond to clamping on Γ_D and the prescription of boundary loads \mathbf{g} on the remainder $\Gamma_N = \partial\Omega \setminus \Gamma_D$, whose outer unit normal vector is denoted by \mathbf{n} .

The boundary loads \mathbf{g} act as control variables. There would be no difficulty in admitting volume forces \mathbf{f} as additional control variables but for practical reasons and simplicity of the presentation, \mathbf{f} is assumed to be zero throughout. The optimal

control, or upper-level problem under consideration reads

$$\left. \begin{array}{l} \text{Minimize } \psi(\mathbf{u}) + \frac{\nu}{2} \|\mathbf{g}\|_{H^1(0,T;L^2(\Gamma_N;\mathbb{R}^d))}^2 \\ \text{such that } (\boldsymbol{\Sigma}, \mathbf{u}, \lambda) \text{ solves the quasistatic plasticity problem (1.1),} \\ \text{and } \mathbf{g} \in U_{\text{ad}}. \end{array} \right\} \quad (\mathbf{P})$$

Here, ψ is a functional favoring certain displacements \mathbf{u} . A typical example would be to track the final deformation after unloading. The set U_{ad} realizes constraints on the control \mathbf{g} . The assumptions on ψ and U_{ad} are given in [Assumption 2.8](#), followed by a number of examples.

We highlight the main contributions of this paper.

- (1) We show that the derivative $(\dot{\boldsymbol{\Sigma}}, \dot{\mathbf{u}})$ of the solution depends continuously on the derivative $\dot{\mathbf{g}}$ of the right-hand side, see [Corollary 2.4](#). This continuity result of the solution operator of (the weak formulation of) [\(1.1\)](#) is superior to the stability results known from the literature.
- (2) Since the solution map of the system [\(1.1\)](#) is nonlinear, its weak continuity is not immediate. We prove this weak continuity in [Theorem 2.7](#). An immediate consequence is the existence of optimal controls of [\(P\)](#), see [Theorem 2.9](#).
- (3) We show the convergence of the solutions of a time-discrete variant of [\(1.1\)](#) w.r.t. the time-step size. In particular, we prove a rate of convergence of the generalized stresses $\boldsymbol{\Sigma}$ in $L^\infty(0, T)$ *without* assuming additional regularity, see [Lemma 3.2](#). Moreover, the convergence of the plastic multiplier λ is shown in [Theorem 3.4](#).
- (4) We study the approximation of local minima of [\(P\)](#) by local minima of its time-discrete variant, see [Section 3.4](#). The results of this section are interesting in their own right and they are also used in order to derive necessary optimality conditions of [\(P\)](#), see [Section 7](#) in Part II of this paper.

Let us sketch the outline of this paper. In the remainder of this section, we introduce the notation and fix the functional analytic framework. Moreover, we state the weak formulation of [\(1.1\)](#) and give some references concerning its analysis. [Section 2](#) is devoted to the continuous optimal control problem [\(P\)](#). We give a brief introduction to evolution variational inequalities (EVIs). Restating [\(1.1\)](#) as an EVI, we are able to prove some new continuity properties of its solution operator. Moreover, we show the weak continuity of the control-to-state map of [\(P\)](#).

In [Section 3](#) we consider a discretization in time of the control problem [\(P\)](#). We show two convergence properties of the solution of the time-discrete forward problems, similar to the strong and weak continuity of the solution map of [\(1.1\)](#). Finally, these results are used in [Section 3.4](#) to prove that *every* local minimizer \mathbf{g} of the control problem [\(P\)](#) can be approximated by local minimizers of a slightly modified time-discrete problem [\(P_g^τ\)](#), see [Theorem 3.10](#).

This paper can be understood as a prerequisite for the analysis contained in Part II. Since the problems under consideration and their analysis are non-standard and non-trivial, the results presented in Part I are believed to be of independent interest.

We give a brief overview over Part II of this work. We regularize the time-discrete forward problem and show the Fréchet differentiability of the associated solution map. This result is non-standard and requires some subtle arguments. The regularized time-discrete optimal control problems are differentiable and consequently, optimality conditions can be derived in a straightforward way. The passage to the

limit in the regularization parameter ε leads to an optimality system of C-stationary type for the time-discrete problem. Finally, we pass to the limit with respect to the discretization in time. This part also requires new convergence arguments. Due to the weak mode of convergence of the adjoint variables, the sign condition for the multipliers is lost in the limit and we finally obtain a system of weak stationarity for the optimal control of (1.1).

1.1. Notation and assumptions. Our notation follows Han and Reddy [1999] and Herzog et al. [2010].

Function spaces. Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with boundary $\Gamma = \partial\Omega$ in dimension $d = 3$. The boundary consists of two disjoint parts Γ_N and Γ_D . We point out that the presented analysis is not restricted to the case $d = 3$, but for reasons of physical interpretation we focus on the three dimensional case. In dimension $d = 2$, the interpretation of (1.1) has to be slightly modified, depending on whether one considers the plane strain or plane stress formulation.

We denote by $\mathbb{S} := \mathbb{R}_{\text{sym}}^{d \times d}$ the space of symmetric d -by- d matrices, endowed with the inner product $\boldsymbol{\sigma} : \boldsymbol{\tau} = \sum_{i,j=1}^d \sigma_{ij} \tau_{ij}$, and we define

$$V = H_D^1(\Omega; \mathbb{R}^d) = \{\mathbf{u} \in H^1(\Omega; \mathbb{R}^d) : \mathbf{u} = \mathbf{0} \text{ on } \Gamma_D\}, \quad S = L^2(\Omega; \mathbb{S})$$

as the spaces for the displacement \mathbf{u} , stress $\boldsymbol{\sigma}$, and back stress $\boldsymbol{\chi}$, respectively. The control \mathbf{g} belongs to the space

$$U = L^2(\Gamma_N; \mathbb{R}^d).$$

The control operator $E : U \rightarrow V'$, $\mathbf{g} \mapsto \ell$, which maps boundary forces (i.e. controls) $\mathbf{g} \in U$ to functionals (i.e. right-hand sides of the weak formulation of (1.1), see Section 1.2) $\ell \in V'$ is given by

$$\langle \mathbf{v}, E\mathbf{g} \rangle_{V, V'} := - \int_{\Gamma_N} \mathbf{v} \cdot \mathbf{g} \, ds \quad \text{for all } \mathbf{v} \in V. \quad (1.2)$$

Hence, $E = -\tau_N^*$, where τ_N is the trace operator from V to $U = L^2(\Gamma_N; \mathbb{R}^d)$. Clearly, $E : U \rightarrow V'$ is compact.

Starting with Section 2.2, we will omit the indices on the duality bracket $\langle \cdot, \cdot \rangle$. We denote by $\langle \cdot, \cdot \rangle$ the dual pairing between V and its dual V' , or the scalar products in S or S^2 , respectively. This will simplify the notation and cause no ambiguities.

For a Banach space X and $p \in [1, \infty]$, we define the Bochner-Lebesgue space

$$L^p(0, T; X) = \{u : [0, T] \rightarrow X, u \text{ Bochner measurable and } p\text{-integrable}\}.$$

In the case $p = \infty$ one has to replace p -integrability by essential boundedness. The norm in $L^p(0, T; X)$ is given by

$$\|u\|_{L^p(0, T; X)} = \left\| \|u(\cdot)\|_X \right\|_{L^p(0, T)}.$$

By $W^{1,p}(0, T; X)$ we denote the Bochner-Sobolev space consisting of functions $u \in L^p(0, T; X)$ which possess a weak derivative $\dot{u} \in L^p(0, T; X)$. Two equivalent norms on $W^{1,p}(0, T; X)$ are given by

$$\left(\|u\|_{L^p(0, T; X)}^p + \|\dot{u}\|_{L^p(0, T; X)}^p \right)^{1/p} \quad \text{and} \quad \left(\|u(0)\|_X^p + \|\dot{u}\|_{L^p(0, T; X)}^p \right)^{1/p}, \quad (1.3)$$

where the extension to the case $p = \infty$ is clear. We use $H^1(0, T; X) = W^{1,2}(0, T; X)$. Moreover, we define

$$H_0^1(0, T; X) = \{u \in H^1(0, T; X) : u(0) = 0\}. \quad (1.4)$$

Details on Bochner-Lebesgue and Bochner-Sobolev spaces can be found in [Yosida \[1965\]](#), [Gajewski et al. \[1974\]](#), [Diestel and Uhl \[1977\]](#), or [Růžička \[2004\]](#).

Yield function and admissible stresses. We restrict our discussion to the von Mises yield function. In the context of linear kinematic hardening, it reads

$$\phi(\boldsymbol{\Sigma}) = (|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D|^2 - \bar{\sigma}_0^2)/2 \quad (1.5)$$

for $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S^2$, where $|\cdot|$ denotes the pointwise Frobenius norm of matrices and

$$\boldsymbol{\sigma}^D = \boldsymbol{\sigma} - \frac{1}{d} (\text{trace } \boldsymbol{\sigma}) \mathbf{I} \quad (1.6)$$

is the deviatoric part of $\boldsymbol{\sigma}$. The yield function gives rise to the set of admissible generalized stresses

$$\mathcal{K} = \{\boldsymbol{\Sigma} \in S^2 : \phi(\boldsymbol{\Sigma}) \leq 0 \text{ a.e. in } \Omega\}.$$

Let us mention that the structure of the yield function ϕ given in (1.5) implies the *shift invariance*

$$\boldsymbol{\Sigma} \in \mathcal{K} \iff \boldsymbol{\Sigma} + (\boldsymbol{\tau}, -\boldsymbol{\tau}) \in \mathcal{K} \text{ for all } \boldsymbol{\tau} \in S. \quad (1.7)$$

This property is exploited quite often in the analysis.

Due to the structure of the yield function ϕ , $\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D$ appears frequently and we abbreviate it and its adjoint by

$$\mathcal{D}\boldsymbol{\Sigma} = \boldsymbol{\sigma}^D + \boldsymbol{\chi}^D \quad \text{and} \quad \mathcal{D}^*\boldsymbol{\sigma} = \begin{pmatrix} \boldsymbol{\sigma}^D \\ \boldsymbol{\sigma}^D \end{pmatrix} \quad (1.8)$$

for matrices $\boldsymbol{\Sigma} \in \mathbb{S}^2$ as well as for functions $\boldsymbol{\Sigma} \in S^2$ and $\boldsymbol{\Sigma} \in L^p(0, T; S^2)$. When considered as an operator in function space, \mathcal{D} maps S^2 and $L^p(0, T; S^2)$ continuously into S and $L^p(0, T; S)$, respectively. For later reference, we also remark that

$$\mathcal{D}^*\mathcal{D}\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma}^D + \boldsymbol{\chi}^D \\ \boldsymbol{\sigma}^D + \boldsymbol{\chi}^D \end{pmatrix} \quad \text{and} \quad (\mathcal{D}^*\mathcal{D})^2 = 2\mathcal{D}^*\mathcal{D}$$

holds. Due to the definition of the operator \mathcal{D} , the constraint $\phi(\boldsymbol{\Sigma}) \leq 0$ can be formulated as $\|\mathcal{D}\boldsymbol{\Sigma}\|_{L^\infty(\Omega; S)} \leq \bar{\sigma}_0$. Hence, we obtain

$$\boldsymbol{\Sigma} \in \mathcal{K} \implies \mathcal{D}\boldsymbol{\Sigma} \in L^\infty(\Omega; S). \quad (1.9)$$

Here and in the sequel we denote linear operators, e.g. $\mathcal{D} : S^2 \rightarrow S$, and the induced Nemytzki operators, e.g. $\mathcal{D} : H^1(0, T; S^2) \rightarrow H^1(0, T; S)$ and $\mathcal{D} : L^2(0, T; S^2) \rightarrow L^2(0, T; S)$, with the same symbol. This will cause no confusion, since the meaning will be clear from the context.

Operators. The linear operators $A : S^2 \rightarrow S^2$ and $B : S^2 \rightarrow V'$ are defined as follows. For $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S^2$ and $\boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in S^2$, let $A\boldsymbol{\Sigma}$ be defined through

$$\langle \boldsymbol{T}, A\boldsymbol{\Sigma} \rangle_{S^2} = \int_{\Omega} \boldsymbol{\tau} : \mathbb{C}^{-1} \boldsymbol{\sigma} \, dx + \int_{\Omega} \boldsymbol{\mu} : \mathbb{H}^{-1} \boldsymbol{\chi} \, dx. \quad (1.10)$$

The term $(1/2) \langle A\boldsymbol{\Sigma}, \boldsymbol{\Sigma} \rangle_{S^2}$ corresponds to the energy associated with the stress state $\boldsymbol{\Sigma}$. Here $\mathbb{C}^{-1}(x)$ and $\mathbb{H}^{-1}(x)$ are linear maps from \mathbb{S} to \mathbb{S} (i.e., they are fourth order tensors) which may depend on the spatial variable x . For $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S^2$ and $\boldsymbol{v} \in V$, let

$$\langle B\boldsymbol{\Sigma}, \boldsymbol{v} \rangle_{V', V} = - \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\boldsymbol{v}) \, dx. \quad (1.11)$$

We recall that $\boldsymbol{\varepsilon}(\boldsymbol{v}) = (\nabla \boldsymbol{v} + (\nabla \boldsymbol{v})^\top)/2$ denotes the (linearized) strain tensor.

Standing assumptions. Throughout the paper, we require

Assumption 1.1.

- (1) The domain $\Omega \subset \mathbb{R}^d$, $d = 3$ is a bounded Lipschitz domain in the sense of [Grisvard, 1985, Chapter 1.2]. The boundary of Ω , denoted by Γ , consists of two disjoint measurable parts Γ_N and Γ_D such that $\Gamma = \Gamma_N \cup \Gamma_D$. While Γ_N is a relatively open subset, Γ_D is a relatively closed subset of Γ . Furthermore Γ_D is assumed to have positive measure. In addition, the set $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, cf. Gröger [1989]. A characterization of regular domains for the case $d \in \{2, 3\}$ can be found in [Haller-Dintelmann et al., 2009, Section 5]. This class of domains covers a wide range of geometries.
- (2) The yield stress $\bar{\sigma}_0$ is assumed to be a positive constant. It equals $\sqrt{2/3} \sigma_0$, where σ_0 is the initial uni-axial yield stress.
- (3) \mathbb{C}^{-1} and \mathbb{H}^{-1} are elements of $L^\infty(\Omega; \mathcal{L}(\mathbb{S}, \mathbb{S}))$, where $\mathcal{L}(\mathbb{S}, \mathbb{S})$ denotes the space of linear operators $\mathbb{S} \rightarrow \mathbb{S}$. Both $\mathbb{C}^{-1}(x)$ and $\mathbb{H}^{-1}(x)$ are assumed to be uniformly coercive. Moreover, we assume that \mathbb{C}^{-1} and \mathbb{H}^{-1} are symmetric, i.e., $\boldsymbol{\tau} : \mathbb{C}^{-1}(x) \boldsymbol{\sigma} = \boldsymbol{\sigma} : \mathbb{C}^{-1}(x) \boldsymbol{\tau}$ and a similar relation for \mathbb{H}^{-1} holds for all $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{S}$.

Assumption 1.1 (1) is not restrictive. It enables us to apply the regularity results in Herzog et al. [2011a] pertaining to systems of nonlinear elasticity. The latter appear in the time-discrete forward problem and its regularizations. Additional regularity leads to a norm gap, which is needed to prove the differentiability of the control-to-state map.

Moreover, **Assumption 1.1 (1)** implies that Korn's inequality holds on Ω , i.e.,

$$\|\mathbf{u}\|_{H^1(\Omega; \mathbb{R}^d)}^2 \leq c_K (\|\mathbf{u}\|_{L^2(\Gamma_D; \mathbb{R}^d)}^2 + \|\boldsymbol{\varepsilon}(\mathbf{u})\|_S^2) \quad (1.12)$$

for all $\mathbf{u} \in H^1(\Omega; \mathbb{R}^d)$, see e.g. [Herzog et al., 2011a, Lemma C.1]. Note that (1.12) entails in particular that $\|\boldsymbol{\varepsilon}(\mathbf{u})\|_S$ is a norm on $H_D^1(\Omega; \mathbb{R}^d)$ equivalent to the standard $H^1(\Omega; \mathbb{R}^d)$ norm. A further consequence is that B^* satisfies the inf-sup condition

$$\|\mathbf{u}\|_V \leq \sqrt{c_K} \|B^* \mathbf{u}\|_{S^2} \quad \text{for all } \mathbf{u} \in V. \quad (1.13)$$

Assumption 1.1 (3) is satisfied, e.g., for isotropic and homogeneous materials, for which

$$\mathbb{C}^{-1} \boldsymbol{\sigma} = \frac{1}{2\mu} \boldsymbol{\sigma} - \frac{\lambda}{2\mu(2\mu + d\lambda)} \text{trace}(\boldsymbol{\sigma}) \mathbf{I}$$

with Lamé constants μ and λ , provided that $\mu > 0$ and $d\lambda + 2\mu > 0$ hold. These constants appear only here and there is no risk of confusion with the plastic multiplier λ . A common example for the hardening modulus is given by $\mathbb{H}^{-1} \boldsymbol{\chi} = \boldsymbol{\chi}/k_1$ with hardening constant $k_1 > 0$, see [Han and Reddy, 1999, Section 3.4].

Clearly, **Assumption 1.1 (3)** shows that $\langle A \boldsymbol{\Sigma}, \boldsymbol{\Sigma} \rangle_{S^2} \geq \underline{\alpha} \|\boldsymbol{\Sigma}\|_{S^2}^2$ for some $\underline{\alpha} > 0$ and all $\boldsymbol{\Sigma} \in S^2$. Hence, the operator A is S^2 -elliptic.

1.2. Weak formulation of (1.1) and known results. Testing the strong formulation of the equilibrium of forces

$$\text{div } \boldsymbol{\sigma}(t) = \mathbf{0} \quad \text{in } \Omega \quad \text{and} \quad \boldsymbol{\sigma}(t) \cdot \mathbf{n} = \mathbf{g}(t) \quad \text{on } \Gamma_N$$

with $\mathbf{v} \in V$ and integrating by parts, we obtain

$$\langle \boldsymbol{\Sigma}(t), B^* \mathbf{v} \rangle_{S^2} = - \int_{\Omega} \boldsymbol{\sigma}(t) : \boldsymbol{\varepsilon}(\mathbf{v}) \, dx = - \int_{\Gamma_N} \mathbf{g}(t) \cdot \mathbf{v} \, ds = \langle \ell(t), \mathbf{v} \rangle_{V', V}$$

for all $\mathbf{v} \in V$, or equivalently $B\Sigma(t) = \ell(t)$ in V' , with $\ell = E\mathbf{g}$, see (1.2).

In order to derive the weak formulation of the first two equations of (1.1), we fix an arbitrary test function $\mathbf{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{K}$. Testing the first and second equation of (1.1) with $\boldsymbol{\tau} - \boldsymbol{\sigma}(t)$ and $\boldsymbol{\mu} - \boldsymbol{\chi}(t)$, respectively, we obtain

$$\langle A\dot{\Sigma}(t) + B^*\dot{\mathbf{u}}(t), \mathbf{T} - \Sigma(t) \rangle_{S^2} + \int_{\Omega} \lambda(t) \mathcal{D}\Sigma(t) : (\mathcal{D}\mathbf{T} - \mathcal{D}\Sigma(t)) \, dx = 0.$$

Due to the complementarity relation in (1.1) and $\Sigma(t), \mathbf{T} \in \mathcal{K}$ we find

$$\int_{\Omega} \lambda(t) \mathcal{D}\Sigma(t) : (\mathcal{D}\mathbf{T} - \mathcal{D}\Sigma(t)) \, dx \leq \int_{\Omega} \lambda(t) (|\mathcal{D}\Sigma(t)| |\mathcal{D}\mathbf{T}| - \tilde{\sigma}_0^2) \, dx \leq 0.$$

Hence, we have

$$\langle A\dot{\Sigma}(t) + B^*\dot{\mathbf{u}}(t), \mathbf{T} - \Sigma(t) \rangle_{S^2} \leq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K}.$$

We have derived the weak formulation of (1.1) in the stress-based (so-called dual) form. It is represented by a time-dependent variational inequality (VI) of mixed type: find generalized stresses $\Sigma \in H^1(0, T; S^2)$ and displacements $\mathbf{u} \in H^1(0, T; V)$ which satisfy $\Sigma(t) \in \mathcal{K}$ and

$$\begin{aligned} \langle A\dot{\Sigma}(t) + B^*\dot{\mathbf{u}}(t), \mathbf{T} - \Sigma(t) \rangle_{S^2} &\geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K}, \\ B\Sigma(t) &= \ell(t) \quad \text{in } V', \end{aligned} \tag{VI}$$

f.a.a. $t \in (0, T)$. Moreover, (VI) is subject to the initial condition $(\Sigma(0), \mathbf{u}(0)) = (\mathbf{0}, \mathbf{0})$. In order to guarantee the existence of a solution, we have to require $\ell(0) = \mathbf{0}$. Note that a weak formulation involving the plastic multiplier λ is given in (1.22).

The remainder of this section is devoted to known results on the analysis of (VI). We start with the results given in [Han and Reddy, 1999, Section 8]. We point out that the authors handle a general situation which includes the case of kinematic hardening as a special case.

Reformulation as a sweeping process in Σ . For $\ell \in V'$, we denote by

$$\mathcal{K}_{\ell} = \{\Sigma \in \mathcal{K} : B\Sigma = \ell\} \tag{1.14}$$

the subset of \mathcal{K} on which the constraint $B\Sigma = \ell$ is fulfilled. Testing the first equation of (VI) with $\mathbf{T} \in \mathcal{K}_{\ell(t)}$ results in: given $\ell \in H_0^1(0, T; V')$, find $\Sigma \in H_0^1(0, T; S^2)$, satisfying $\Sigma(t) \in \mathcal{K}_{\ell(t)}$ and

$$\langle A\dot{\Sigma}(t), \mathbf{T} - \Sigma(t) \rangle_{S^2} \geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K}_{\ell(t)} \text{ and almost all } t \in (0, T). \tag{1.15}$$

This is called the stress problem of plasticity. It is a time-dependent VI, where the associated convex set $\mathcal{K}_{\ell(t)}$ changes in time. Such an equation was introduced in Moreau [1977]. We mention that there are existence and uniqueness results, but no continuity results for the abstract situation considered in Moreau [1977] seem to be available.

Stress problem. [Han and Reddy, 1999, Section 8] deals with the so-called dual formulation (1.15). In [Han and Reddy, 1999, Theorem 8.9] the existence and uniqueness of a solution Σ of (1.15) is shown together with the a-priori bound

$$\|\Sigma\|_{H^1(0, T; S^2)} \leq C \|\ell\|_{H^1(0, T; V')}. \tag{1.16}$$

Let us denote the solution map by \mathcal{G}^{Σ} , i.e. $\Sigma = \mathcal{G}^{\Sigma}(\ell)$. Additionally, [Han and Reddy, 1999, Theorem 8.10] shows the local Hölder continuity of index 1/2 of $\mathcal{G}^{\Sigma} : H_0^1(0, T; V') \rightarrow L^{\infty}(0, T; S^2)$, i.e.,

$$\|\mathcal{G}^{\Sigma}(\ell_1) - \mathcal{G}^{\Sigma}(\ell_2)\|_{L^{\infty}(0, T; S^2)}^2 \leq C (\|\dot{\ell}_1\|_{L^2(0, T; V')} + \|\dot{\ell}_2\|_{L^2(0, T; V')}) \|\ell_1 - \ell_2\|_{L^2(0, T; V')}. \tag{1.17}$$

We mention that this is not a typical estimate. As long as the derivatives of ℓ_i remain bounded in $L^2(0, T; V')$, one can control the L^∞ -norm of $\Sigma_1 - \Sigma_2$ solely through the L^2 -norm of the difference $\ell_1 - \ell_2$. We give a generalization of this estimate in [Theorem 2.3](#) in the context of EVIs.

Finally, [[Han and Reddy, 1999](#), Theorem 8.12] shows that one can introduce a (not necessarily unique) multiplier associated with the equilibrium of forces $B\Sigma = \ell$, which can be interpreted as the displacement \mathbf{u} . As a consequence, (Σ, \mathbf{u}) satisfies [\(VI\)](#). Thus, given $\ell \in H_0^1(0, T; V')$, there exists $(\Sigma, \mathbf{u}) \in H_0^1(0, T; S^2 \times V)$ such that $\Sigma(t) \in \mathcal{K}$ and

$$\langle A\dot{\Sigma}(t) + B^*\dot{\mathbf{u}}(t), \mathbf{T} - \Sigma(t) \rangle_{S^2} \geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K}, \quad (1.18a)$$

$$B\Sigma(t) = \ell(t) \quad \text{in } V' \quad (1.18b)$$

holds for almost all $t \in (0, T)$. This shows the equivalence of [\(1.15\)](#) and [\(VI\)](#). Note that [\(1.18\)](#) is equivalent to [\(VI\)](#). In the sequel we use either reference as appropriate.

Kinematic hardening. In the case of kinematic hardening, the uniqueness of \mathbf{u} is obtained easily. Let us test [\(1.18a\)](#) with $\mathbf{T} = \Sigma(t) + (\boldsymbol{\tau}, -\boldsymbol{\tau}) = (\boldsymbol{\sigma}(t) + \boldsymbol{\tau}, \boldsymbol{\chi}(t) - \boldsymbol{\tau})$. Due to the shift invariance [\(1.7\)](#), we have $\mathbf{T} \in \mathcal{K}$ for all $\boldsymbol{\tau} \in S$. This yields

$$\langle A\dot{\Sigma}(t) + B^*\dot{\mathbf{u}}(t), (\boldsymbol{\tau}, -\boldsymbol{\tau}) \rangle_{S^2} = 0 \quad \text{for all } \boldsymbol{\tau} \in S.$$

Using the definitions of A and B , see [\(1.10\)](#) and [\(1.11\)](#), respectively, we obtain

$$\mathbb{C}^{-1}\dot{\boldsymbol{\sigma}} - \boldsymbol{\varepsilon}(\dot{\mathbf{u}}) - \mathbb{H}^{-1}\dot{\boldsymbol{\chi}} = \mathbf{0} \quad \text{almost everywhere in } (0, T) \times \Omega.$$

Integrating from 0 to t and using the initial condition $(\Sigma(0), \mathbf{u}(0)) = \mathbf{0}$ yields

$$\mathbb{C}^{-1}\boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\mathbf{u}) - \mathbb{H}^{-1}\boldsymbol{\chi} = \mathbf{0} \quad \text{almost everywhere in } (0, T) \times \Omega. \quad (1.19)$$

Together with the inf-sup condition of $B^* = (-\boldsymbol{\varepsilon}, \mathbf{0})$, see [\(1.13\)](#), this proves the uniqueness of \mathbf{u} . Using [\(1.16\)](#) yields the a-priori estimate

$$\|\Sigma\|_{H^1(0, T; S^2)} + \|\mathbf{u}\|_{H^1(0, T; V)} \leq C \|\ell\|_{H^1(0, T; V')}. \quad (1.20)$$

We denote the solution mapping $\ell \mapsto \mathbf{u}$ of [\(1.18\)](#) by $\mathcal{G}^{\mathbf{u}}$. Moreover, the solution operator of [\(VI\)](#) mapping $\ell \rightarrow (\Sigma, \mathbf{u})$ is denoted by $\mathcal{G} = (\mathcal{G}^\Sigma, \mathcal{G}^{\mathbf{u}})$.

Primal problem. In [[Han and Reddy, 1999](#), Section 7] the primal formulation of [\(1.15\)](#) is considered. Both formulations are equivalent, see [[Han and Reddy, 1999](#), Theorem 8.3]. For the primal formulation and under the (additional) assumption of kinematic (or combined kinematic-isotropic) hardening, the (global) Lipschitz continuity of the solution operator from $W^{1,1}(0, T; Z')$ to $L^\infty(0, T; Z)$ was proven, see [[Han and Reddy, 1999](#), pp. 170–171]. Here, Z is the appropriate function space for the analysis of the primal formulation. Due to the equivalence of the primal and the dual problem, this Lipschitz estimate carries over to the dual problem. We obtain

$$\|\mathcal{G}(\ell_1) - \mathcal{G}(\ell_2)\|_{L^\infty(0, T; S^2 \times V)} \leq C \|\ell_1 - \ell_2\|_{W^{1,1}(0, T; V')}. \quad (1.21)$$

This Lipschitz estimate was actually already contained in rather unknown works of Gröger, see [Gröger \[1978a,b, 1979\]](#). In [[Gröger, 1979](#), Section 4] the system [\(VI\)](#) is reformulated as an evolution equation associated with a maximal monotone operator. Then the Lipschitz estimate follows by the classical result [[Brézis, 1973](#), Lemma 3.1].

Remark 1.2. In order to derive optimality conditions, the continuity results mentioned above are not sufficient. In addition, we need the continuity of $\mathcal{G} : H_0^1(0, T; V') \rightarrow H_0^1(0, T; S^2 \times V)$ on two occasions. First, this continuity is needed to prove the approximability of local solutions by time-discrete minimizers in [Lemma 3.8](#). Second, the strong convergence of $(\dot{\Sigma}, \dot{u})$ in $L^2(0, T; S^2 \times V)$ is needed to pass to the limit in the optimality system, see [Section 7](#) in the second part of this paper.

The required continuity of $\mathcal{G} : H_0^1(0, T; V') \rightarrow H_0^1(0, T; S^2 \times V)$ is shown in [Corollary 2.4](#) by building on some results of [Krejčí \[1996, 1998\]](#) about evolution variational inequalities (EVIs). To our knowledge, this is a new result for the analysis of [\(VI\)](#).

Existence, uniqueness and regularity of the plastic multiplier. We have already seen that the generalized plastic strain $\mathbf{P} = -A\Sigma - B^*u$ is unique. Thus by [Herzog et al. \[2011b\]](#) we obtain the existence and uniqueness of the plastic multiplier $\lambda \in L^2(0, T; L^2(\Omega))$ which can be understood as a multiplier associated with the constraint $\Sigma \in \mathcal{K}$ or rather $\phi(\Sigma) \leq 0$. We obtain the system

$$A\dot{\Sigma} + B^*\dot{u} + \lambda \mathcal{D}^* \mathcal{D} \Sigma = \mathbf{0} \quad \text{in } L^2(0, T; S^2), \quad (1.22a)$$

$$B\Sigma = \ell \quad \text{in } L^2(0, T; V'), \quad (1.22b)$$

$$0 \leq \lambda \quad \perp \quad \phi(\Sigma) \leq 0 \quad \text{a.e. in } (0, T) \times \Omega. \quad (1.22c)$$

Endowed with the initial condition $(\Sigma(0), u(0)) = (\mathbf{0}, \mathbf{0})$, this system is equivalent to [\(1.15\)](#) and [\(1.18\)](#).

2 Analysis of the continuous optimal control problem

In this section we study the continuous (i.e. non-discretized) optimal control problem. In the first subsection, we give a brief introduction to evolution variational inequalities (EVIs) and state the continuity of the solution map of an EVI from $H^1(0, T; X)$ into itself, see [Theorem 2.2](#). This enables us to show the continuity of the solution map of [\(VI\)](#), see [Section 2.2](#). [Section 2.3](#) is devoted to show the weak continuity of the control-to-state map of [\(P\)](#). Due to this weak continuity, we are able to conclude the existence of an optimal control in [Section 2.4](#), see [Theorem 2.9](#).

2.1. Introduction to evolution variational inequalities. In this section we give a short introduction to *evolution variational inequalities*. A comprehensive presentation of this topic can be found in [Krejčí \[1998\]](#). We also mention the contributions [Krejčí and Lovicar \[1990\]](#), [Krejčí \[1996\]](#), [Brokate and Krejčí \[1998\]](#). For convenience of the reader who wishes to consult [Krejčí \[1998\]](#) in parallel, we use the notation of [Krejčí \[1998\]](#) in this subsection and turn back to our notation in the next subsection.

Let X be some Hilbert space and $Z \subset X$ be some convex, closed set. Given a function $u : [0, T] \rightarrow X$ and an initial value $x_0 \in Z$, find $x : [0, T] \rightarrow Z \subset X$ such that $x(0) = x_0$ and

$$(\dot{u}(t) - \dot{x}(t), x(t) - \tilde{x})_X \geq 0 \quad \text{for all } \tilde{x} \in Z. \quad (\text{EVI})$$

There are many applications that lead to an EVI, see [[Krejčí, 1998](#), page 2] for references and further comments on this topic. We only mention the strain or stress driven problem of plasticity with linear kinematic hardening. In the strain (stress) driven problem, the strain (stress) is viewed as a known quantity, whereas the stress (strain) has to be determined. We show that the dual formulation [\(1.18\)](#)

is an appropriate starting point for deriving the strain driven problem. Krejčí derived the stress and the strain driven problems in [Krejčí, 1998, (1.29) and (1.31)], respectively.

In the strain driven problem, the strain $\varepsilon(\mathbf{u})$ (or, equivalently, the strain rate $(\varepsilon(\dot{\mathbf{u}}), \mathbf{0}) = -B^*\dot{\mathbf{u}}$) is viewed as a given quantity. The VI (1.18a) reduces to find a function $\Sigma : [0, T] \rightarrow \mathcal{K} \subset S^2$, such that

$$\left\langle \dot{\Sigma}(t) - \begin{pmatrix} \mathbb{C}\varepsilon(\dot{\mathbf{u}}(t)) \\ 0 \end{pmatrix}, \mathbf{T} - \Sigma(t) \right\rangle_A \geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K}, \quad (2.1)$$

where $\langle \cdot, \cdot \rangle_A$ is the scalar product on S^2 induced by A , i.e.,

$$\langle \Sigma, \mathbf{T} \rangle_A := \langle A\Sigma, \mathbf{T} \rangle_{S^2} \quad \text{for } \Sigma, \mathbf{T} \in S^2. \quad (2.2)$$

Using the coercivity of A , which is ensured by Assumption 1.1 (3), we find that $\langle \cdot, \cdot \rangle_A$ is scalar product on S^2 which is equivalent to the standard scalar product. The equation (2.1) is of the form (EVI) in the Hilbert space $X = S^2$ equipped with the scalar product $\langle \cdot, \cdot \rangle_A$ and the feasible set $Z = \mathcal{K}$.

Remark 2.1. We remark that there is a fundamental difference between the dual formulation (1.18) and the strain driven problem (2.1). Whereas in (1.18) both Σ and \mathbf{u} are unknown quantities, $\varepsilon(\mathbf{u})$ has to be *a-priori known* in (2.1). Moreover, the solution Σ of (2.1) has to fulfill (1.18b).

Let us denote the solution map of (2.1) which maps $\mathbf{u} \rightarrow \Sigma$ by S . Therefore, (1.18b) yields $\ell = B\Sigma = BS(\mathbf{u})$. Thus, (1.18) is equivalent to

$$\text{Find } \mathbf{u} \text{ such that } BS(\mathbf{u}) = \ell.$$

This technique is used in [Krejčí, 1996, Chapter III] for hyperbolic equations arising in plasticity, i.e., the author considers plasticity problems where the acceleration term $\rho\ddot{\mathbf{u}}$ is included in the balance of forces (1.18b).

The following theorem summarizes results on (EVI) given in Krejčí [1996, 1998].

Theorem 2.2. Let $u \in W^{1,1}(0, T; X)$ and $x_0 \in Z$ be given. Then there exists a unique solution $S(x_0, u) := x \in W^{1,1}(0, T; X)$ of (EVI). Moreover, $S : Z \times W^{1,1}(0, T; X) \rightarrow L^\infty(0, T; X)$ is globally Lipschitz continuous and $S : Z \times W^{1,p}(0, T; X) \rightarrow W^{1,p}(0, T; X)$ is continuous for all $p < \infty$.

The operator S is called the *stop operator*, whereas $P(x_0, u) = u - S(x_0, u)$ is called the *play operator*.

Let us show a result which generalizes the local Hölder estimate (1.17). In the context of EVIs this is a new result. It shows that the play operator P possesses an additional smoothing property which is not shared by the stop operator S . Another such property is that P maps $C(0, T; X)$ (continuous functions mapping $[0, T] \rightarrow X$) to $CBV(0, T; X)$ (continuous functions of bounded variation mapping $[0, T] \rightarrow X$), see [Krejčí, 1998, Theorem 3.11]. Again, this does not hold for the stop operator S .

Theorem 2.3. Let $p \in [1, \infty]$ be given and denote by q its dual exponent. Let $u_1, u_2 \in W^{1,q}(0, T; X)$ and initial values $x_1^0, x_2^0 \in Z$ be given. Let $\xi_i = P(x_i^0, u_i)$, $i = 1, 2$. Then the estimate

$$\begin{aligned} \|\xi_1 - \xi_2\|_{L^\infty(0, T; X)}^2 &\leq 2 \left(\|\dot{u}_1\|_{L^q(0, T; X)} + \|\dot{u}_2\|_{L^q(0, T; X)} \right) \|u_1 - u_2\|_{L^p(0, T; X)} \\ &\quad + \|\xi_1(0) - \xi_2(0)\|_X^2 \end{aligned}$$

holds, where $\xi_i(0) = u_i(0) - x_i^0$, $i = 1, 2$.

Proof. First we recall [Krejčí, 1998, (3.16ii)], i.e., $(\dot{\xi}_1(t), \dot{x}_1(t))_X = 0$ a.e. in $(0, T)$, where $x_1 = S(x_1^0, u_1)$. Using $x_1 = u_1 - \xi_1$, this implies $\|\dot{\xi}_1(t)\|_X \leq \|\dot{u}_1(t)\|_X$ a.e. in $(0, T)$. Analogously, we obtain $\|\dot{\xi}_2(t)\|_X \leq \|\dot{u}_2(t)\|_X$ a.e. in $(0, T)$.

Using $\tilde{x} = u_2(t) - \xi_2(t) = x_2(t) \in Z$ as a test function in **(EVI)** shows

$$(\dot{\xi}_1(t), u_1(t) - u_2(t) + \xi_2(t) - \xi_1(t))_X \geq 0 \quad \text{a.e. in } (0, T).$$

Similarly, we obtain

$$(\dot{\xi}_2(t), u_2(t) - u_1(t) + \xi_1(t) - \xi_2(t))_X \geq 0 \quad \text{a.e. in } (0, T).$$

Adding these inequalities, we have

$$(\dot{\xi}_1(t) - \dot{\xi}_2(t), \xi_1(t) - \xi_2(t))_X \leq (\dot{\xi}_1(t) - \dot{\xi}_2(t), u_1(t) - u_2(t))_X \quad \text{a.e. in } (0, T).$$

This shows

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\xi_1 - \xi_2\|_X^2 &= (\dot{\xi}_1(t) - \dot{\xi}_2(t), \xi_1(t) - \xi_2(t))_X \\ &\leq (\dot{\xi}_1(t) - \dot{\xi}_2(t), u_1(t) - u_2(t))_X \\ &\leq (\|\dot{\xi}_1(t)\|_X + \|\dot{\xi}_2(t)\|_X) \|u_1(t) - u_2(t)\|_X \\ &\leq (\|\dot{u}_1(t)\|_X + \|\dot{u}_2(t)\|_X) \|u_1(t) - u_2(t)\|_X \quad \text{a.e. in } (0, T). \end{aligned}$$

Integrating from 0 to t yields

$$\begin{aligned} \frac{1}{2} \|\xi_1(t) - \xi_2(t)\|_X^2 - \frac{1}{2} \|\xi_1(0) - \xi_2(0)\|_X^2 \\ \leq \int_0^t (\|\dot{u}_1(s)\|_X + \|\dot{u}_2(s)\|_X) \|u_1(s) - u_2(s)\|_X ds \\ \leq (\|\dot{u}_1\|_{L^q(0,T;X)} + \|\dot{u}_2\|_{L^q(0,T;X)}) \|u_1 - u_2\|_{L^p(0,T;X)}. \end{aligned}$$

Taking the supremum $t \in (0, T)$ yields the claim.

2.2. Quasistatic plasticity as an EVI. In this section we give an equivalent reformulation of (1.18) which fits in the framework of Krejčí [1998]. Once this reformulation is established, the continuity of $\mathcal{G} : H_0^1(0, T; V') \rightarrow H_0^1(0, T; S^2 \times V)$ is a consequence of Theorem 2.2. As mentioned in Remark 1.2, the continuity of \mathcal{G} is used in several places.

In order to reformulate (1.18), we need some preparatory work. Due to the inf-sup condition (1.13) we obtain that for all $\ell \in V'$ there is a $\sigma_\ell \in S$, such that

$$(\sigma_\ell, \mathbf{0}) \in (\ker B)^\perp \quad \text{and} \quad B(\sigma_\ell, \mathbf{0}) = \ell, \quad (2.3)$$

see Brezzi [1974]. Moreover, this mapping $\ell \mapsto \sigma_\ell$ is linear and continuous. For $\ell \in H_0^1(0, T; V')$, we define $\sigma_\ell(t) := \sigma_{\ell(t)}$. Hence, $\sigma_{(\cdot)}$ maps $H_0^1(0, T; V') \rightarrow H_0^1(0, T; S)$ continuously. For convenience, we also introduce the notation

$$\Sigma_\ell = (\sigma_\ell, -\sigma_\ell). \quad (2.4)$$

Let us mention that this is a useful tool to construct test functions \mathbf{T} for (1.15) and (1.18). Indeed, for arbitrary $\Sigma \in \mathcal{K}$ and $\ell \in V'$, let us define the test function $\mathbf{T} = \Sigma + \Sigma_{\ell - B\Sigma}$. The shift invariance (1.7) implies $\mathbf{T} \in \mathcal{K}$ and $B\mathbf{T} = \ell$ is ensured by the definition of Σ_ℓ . Hence we obtain $\mathbf{T} \in \mathcal{K}_\ell$, i.e., it is an admissible test function for (1.15) and (1.18).

Let $\ell \in H_0^1(0, T; V')$ be arbitrary. Using (2.4) we are able to reformulate (1.15) as an EVI. As stated in (1.7), $\Sigma \in \mathcal{K}$ holds if and only if $\Sigma + (\tau, -\tau) \in \mathcal{K}$ holds for all $\tau \in S$. Since $\Sigma_\ell = (\sigma_\ell, -\sigma_\ell)$ by definition, $\Sigma \in \mathcal{K}$ if and only if $\Sigma - \Sigma_\ell \in \mathcal{K}$. This gives rise to the decomposition $\Sigma = \Sigma_0 + \Sigma_\ell$. There are two immediate

consequences: $\Sigma \in \mathcal{K}$ is equivalent to $\Sigma_0 \in \mathcal{K}$ and $B\Sigma = \ell$ is equivalent to $\Sigma_0 \in \ker B$. Therefore, $\Sigma_0 \in \mathcal{K} \cap \ker B$.

We define $\mathcal{K}_B := \mathcal{K} \cap \ker B$. Let $\mathbf{T} \in \mathcal{K}_B$ be arbitrary. Testing (1.18a) or (1.15) with $\mathbf{T} + \Sigma_\ell(t) \in \mathcal{K}$ yields an equivalent reformulation: find $\Sigma_0 \in \mathcal{K}_B$ such that

$$\langle A(\dot{\Sigma}_0 + \dot{\Sigma}_\ell), \mathbf{T} - \Sigma_0 \rangle \geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K}_B. \quad (2.5)$$

Hence, (2.5) is an EVI in the Hilbert space S^2 equipped with $\langle \cdot, \cdot \rangle_A$. Theorem 2.2 yields the continuity of the mapping $\Sigma_\ell \mapsto \Sigma_0$ from $H_0^1(0, T; S^2)$ to $H_0^1(0, T; S^2)$. Since $\ell \mapsto \Sigma_\ell$ is continuous from $H_0^1(0, T; V')$ to $H_0^1(0, T; S^2)$, we obtain the continuity of the mapping $\ell \mapsto \Sigma$ from $H_0^1(0, T; V')$ to $H_0^1(0, T; S^2)$.

The uniqueness and continuous dependence of \mathbf{u} on ℓ follows from (1.19) and the inf-sup condition of B^* , see (1.13). Moreover, Theorem 2.2 also shows the Lipschitz continuity of $\mathcal{G} : \ell \mapsto (\Sigma, \mathbf{u})$, $W^{1,1}(0, T; V') \rightarrow L^\infty(0, T; S^2 \times V)$, see (1.21). Summarizing, we have found

Corollary 2.4. The solution operator \mathcal{G} of (VI) is continuous from $H_0^1(0, T; V')$ to $H_0^1(0, T; S^2 \times V)$ and globally Lipschitz continuous to $L^\infty(0, T; S^2 \times V)$.

Remark 2.5. The technique presented in this subsection is not restricted to kinematic hardening only, the arguments remain valid for all hardening models involving some kinematic part, see also Gröger [1979]. To be precise, the technique is applicable whenever the hardening variable can be split into a kinematic part χ and some other part η , such that the yield function ϕ can be written as $\phi(\Sigma) = \phi(\sigma + \chi, \eta)$, where $\Sigma = (\sigma, \chi, \eta)$.

It is interesting to note that this is exactly the case for which Han and Reddy [1999] were able to prove Lipschitz continuity of the forward operator (in primal formulation) from $W^{1,1}(0, T; Z')$ to $L^\infty(0, T; Z)$, see the discussion in Section 1.2.

2.3. Weak continuity of the forward operator. In this section we show the weak continuity of the control-to-state map $\mathcal{G} \circ E$ of the optimal control problem (P). Here, \mathcal{G} is the solution map of (1.18) and E is the control operator, see (1.2). Since the solution operator $\mathcal{G} \circ E$ is nonlinear, the proof of its weak continuity is non-trivial. The weak continuity of $\mathcal{G} \circ E$ from $H_0^1(0, T; U)$ to $H_0^1(0, T; S^2 \times V)$ is essential for proving the existence of an optimal control in Theorem 2.9.

A main ingredient to prove the weak continuity is the compactness of the control operator E from $U = L^2(\Gamma_N; \mathbb{R}^d)$ to $V' = (H_D^1(\Omega; \mathbb{R}^d))'$, see the proof of Lemma 2.6. Due to this compactness, Aubin's lemma, see, e.g., [Simon, 1986, Equation (6.5)], implies that $H^1(0, T; U)$ embeds compactly into $L^2(0, T; V')$.

Let us mention that the weak continuity of $\mathcal{G} \circ E$ is a non-trivial result: although $E : U \rightarrow V'$ is compact, the operator $E : H^1(0, T; U) \rightarrow H^1(0, T; V')$ (applied in a pointwise sense) is not compact. Hence the weak continuity of $\mathcal{G} \circ E$ does not simply follow from the compactness of E nor the continuity of \mathcal{G} .

First, we need a preliminary result. It can be interpreted as an upper semicontinuity result for the left-hand side of (1.18a).

Lemma 2.6. Let $\{(\Sigma_k, \mathbf{u}_k)\} \subset H^1(0, T; S^2 \times V)$ be a sequence with $\Sigma_k(0) = \mathbf{0}$ and let $(\Sigma, \mathbf{u}) \in H^1(0, T; S^2 \times V)$ be given. Moreover, we assume that there are $\mathbf{g}_k, \mathbf{g} \in H^1(0, T; U)$ such that $B\Sigma_k = E\mathbf{g}_k$ and $B\Sigma = E\mathbf{g}$ hold. If

$$\begin{aligned} (\Sigma_k, \mathbf{u}_k) &\rightharpoonup (\Sigma, \mathbf{u}) \text{ in } H^1(0, T; S^2 \times V) \quad \text{and} \\ \mathbf{g}_k &\rightharpoonup \mathbf{g} \quad \text{in } H^1(0, T; U), \end{aligned}$$

then for all $\mathbf{T} \in L^2(0, T; S^2)$ we have

$$\limsup_{k \rightarrow \infty} \int_0^T \langle A\dot{\Sigma}_k + B^* \dot{\mathbf{u}}_k, \mathbf{T} - \Sigma_k \rangle dt \leq \int_0^T \langle A\dot{\Sigma} + B^* \dot{\mathbf{u}}, \mathbf{T} - \Sigma \rangle dt.$$

Proof. Due to the weak convergence of (Σ_k, \mathbf{u}_k) in $H^1(0, T; S^2 \times V)$ we have

$$\int_0^T \langle A\dot{\Sigma}_k + B^* \dot{\mathbf{u}}_k, \mathbf{T} \rangle dt \rightarrow \int_0^T \langle A\dot{\Sigma} + B^* \dot{\mathbf{u}}, \mathbf{T} \rangle dt.$$

Since U embeds compactly into V' , Aubin's lemma yields that $H^1(0, T; U)$ embeds compactly into $L^2(0, T; V')$ and hence $\mathbf{g}_k \rightharpoonup \mathbf{g}$ in $H^1(0, T; U)$ implies $E\mathbf{g}_k \rightarrow E\mathbf{g}$ in $L^2(0, T; V')$. This shows

$$\begin{aligned} \int_0^T \langle B^* \dot{\mathbf{u}}_k, \Sigma_k \rangle dt &= \int_0^T \langle \dot{\mathbf{u}}_k, B\Sigma_k \rangle dt = \int_0^T \langle \dot{\mathbf{u}}_k, E\mathbf{g}_k \rangle dt \\ &\rightarrow \int_0^T \langle \dot{\mathbf{u}}, E\mathbf{g} \rangle dt = \int_0^T \langle \dot{\mathbf{u}}, B\Sigma \rangle dt = \int_0^T \langle B^* \dot{\mathbf{u}}, \Sigma \rangle dt. \end{aligned}$$

To address the remaining term, we use integration by parts and obtain

$$\int_0^T \langle A\dot{\Sigma}_k, \Sigma_k \rangle dt = \frac{1}{2} (\langle A\Sigma_k(T), \Sigma_k(T) \rangle - \langle A\Sigma_k(0), \Sigma_k(0) \rangle).$$

The functionals $\Sigma \mapsto \langle A\Sigma(t), \Sigma(t) \rangle$ are continuous (w.r.t. the $H^1(0, T; S^2)$ -norm) and convex, hence weakly lower semicontinuous. Due to $\Sigma_k(0) = \mathbf{0}$, this already implies $\langle A\Sigma(0), \Sigma(0) \rangle = 0$. Now the weak lower semicontinuity of $\Sigma \mapsto \langle A\Sigma(T), \Sigma(T) \rangle$ yields

$$\liminf_{k \rightarrow \infty} \langle A\Sigma_k(T), \Sigma_k(T) \rangle \geq \langle A\Sigma(T), \Sigma(T) \rangle.$$

Therefore,

$$\liminf_{k \rightarrow \infty} \int_0^T \langle A\dot{\Sigma}_k, \Sigma_k \rangle dt \geq \frac{1}{2} \langle A\Sigma(T), \Sigma(T) \rangle = \int_0^T \langle A\dot{\Sigma}, \Sigma \rangle dt$$

holds. By combining the results above, we obtain the assertion

$$\limsup_{k \rightarrow \infty} \int_0^T \langle A\dot{\Sigma}_k + B^* \dot{\mathbf{u}}_k, \mathbf{T} - \Sigma_k \rangle dt \leq \int_0^T \langle A\dot{\Sigma} + B^* \dot{\mathbf{u}}, \mathbf{T} - \Sigma \rangle dt.$$

The previous lemma enables us to prove

Theorem 2.7. The operator $\mathcal{G} \circ E$ is weakly continuous from $H_0^1(0, T; U)$ to $H_0^1(0, T; S^2 \times V)$.

Proof. Let $\{\mathbf{g}_k\} \subset H_0^1(0, T; U)$ be a sequence converging weakly towards $\mathbf{g} \in H_0^1(0, T; U)$. The solution of (1.18) with right-hand side $E\mathbf{g}$, $E\mathbf{g}_k$ is denoted by (Σ, \mathbf{u}) , (Σ_k, \mathbf{u}_k) , respectively, i.e., $(\Sigma_k, \mathbf{u}_k) = \mathcal{G}(E\mathbf{g}_k)$ and $(\Sigma, \mathbf{u}) = \mathcal{G}(E\mathbf{g})$. Due to the boundedness of E and \mathcal{G} , see (1.20), there exists a subsequence (for simplicity denoted by the same symbol) $\{(\Sigma_k, \mathbf{u}_k)\}$ which converges weakly towards some $(\tilde{\Sigma}, \tilde{\mathbf{u}})$ in $H^1(0, T; S^2 \times V)$. We shall prove $(\tilde{\Sigma}, \tilde{\mathbf{u}}) = (\Sigma, \mathbf{u})$, therefore the whole sequence converges weakly and this simplification of notation is justified. To this end, we show that $(\tilde{\Sigma}, \tilde{\mathbf{u}})$ is a solution to (1.18) with right-hand side $E\mathbf{g}$.

First we address the admissibility of $\tilde{\Sigma}$ in (1.18): the set $\{\mathbf{T} \in H^1(0, T; S^2) : \mathbf{T}(t) \in \mathcal{K} \text{ f.a.a. } t \in (0, T)\}$ is convex and closed, hence weakly closed. Since all Σ_k

belong to this set, $\tilde{\Sigma}$ belongs to it as well. Therefore $\tilde{\Sigma}(t) \in \mathcal{K}$ holds for almost all $t \in (0, T)$.

Since B is linear and bounded, it is weakly continuous and hence $B\Sigma_k \rightharpoonup B\tilde{\Sigma}$ in $H^1(0, T; V')$. Due to $B\Sigma_k = E\mathbf{g}_k \rightharpoonup E\mathbf{g}$ in $H^1(0, T; V')$, we have $B\tilde{\Sigma} = E\mathbf{g}$ in $H^1(0, T; V')$, this shows (1.18b).

By (1.18a) we have for all $\mathbf{T} \in L^2(0, T; S^2)$ with $\mathbf{T}(t) \in \mathcal{K}$ f.a.a. $t \in (0, T)$

$$\int_0^T \langle A\dot{\Sigma}_k + B^*\dot{\mathbf{u}}_k, \mathbf{T} - \Sigma_k \rangle dt \geq 0.$$

Now we can apply Lemma 2.6. This yields

$$\int_0^T \langle A\dot{\Sigma} + B^*\dot{\mathbf{u}}, \mathbf{T} - \tilde{\Sigma} \rangle dt \geq \limsup_{k \rightarrow \infty} \int_0^T \langle A\dot{\Sigma}_k + B^*\dot{\mathbf{u}}_k, \mathbf{T} - \Sigma_k \rangle dt \geq 0.$$

Since $\mathbf{T} \in L^2(0, T; S^2)$ with $\mathbf{T}(t) \in \mathcal{K}$ f.a.a. $t \in (0, T)$ was arbitrary, we obtain the pointwise variational inequality

$$\langle A\dot{\Sigma} + B^*\dot{\mathbf{u}}, \mathbf{T} - \tilde{\Sigma} \rangle \geq 0 \quad \text{a.e. in } (0, T)$$

for all $\mathbf{T} \in \mathcal{K}$. Hence $(\tilde{\Sigma}, \tilde{\mathbf{u}})$ solves (1.18). Since the solution is unique we obtain $(\tilde{\Sigma}, \tilde{\mathbf{u}}) = (\Sigma, \mathbf{u})$. This implies that the weak limit is indeed independent of the chosen subsequence. Hence, the whole sequence converges weakly, i.e. $(\Sigma_k, \mathbf{u}_k) \rightharpoonup (\Sigma, \mathbf{u})$ in $H^1(0, T; S^2 \times V)$. That is, $\mathcal{G}(E\mathbf{g}_k) = (\Sigma_k, \mathbf{u}_k) \rightharpoonup (\Sigma, \mathbf{u}) = \mathcal{G}(E\mathbf{g})$ in $H^1(0, T; S^2 \times V)$. Since $\{\mathbf{g}_k\}$ and \mathbf{g} were arbitrary, $\mathcal{G} \circ E$ is weakly continuous.

2.4. Existence of optimal controls. In this section we prove the existence of an optimal control. The key tool in the proof of Theorem 2.9 is the weak continuity of the forward operator proven in Theorem 2.7. For convenience, we repeat the definition of the optimal control problem under consideration

$$\left. \begin{array}{l} \text{Minimize } F(\mathbf{u}, \mathbf{g}) = \psi(\mathbf{u}) + \frac{\nu}{2} \|\mathbf{g}\|_{H^1(0, T; U)}^2 \\ \text{such that } (\Sigma, \mathbf{u}) = \mathcal{G}(E\mathbf{g}) \\ \text{and } \mathbf{g} \in U_{\text{ad}}. \end{array} \right\} \quad (\mathbf{P})$$

The admissible set U_{ad} is a convex closed subset of $H_0^1(0, T; U)$. Here, $\|\cdot\|_{H^1(0, T; U)}$ can be any norm such that $H^1(0, T; U)$ is a Hilbert space, see (1.3). Let us fix the assumptions on the objective ψ and on the set of admissible controls U_{ad} .

Assumption 2.8.

- (1) The function $\psi : H^1(0, T; V) \rightarrow \mathbb{R}$ is weakly lower semicontinuous, continuous and bounded from below.
- (2) The cost parameter ν is a positive, real number.
- (3) The admissible set U_{ad} is nonempty, convex and closed in $H_0^1(0, T; U)$.

Let us give some examples which satisfy these conditions. Since the domain of definition of ψ is $H^1(0, T; V)$ we can track the displacements, the strains or even their time derivatives (or point evaluations). As examples for ψ we mention

$$\begin{aligned} \psi^1(\mathbf{u}) &= \frac{1}{2} \|\mathbf{u} - \mathbf{u}_d\|_{L^2(0, T; L^2(\Omega; \mathbb{R}^d))}^2, & \psi^2(\mathbf{u}) &= \frac{1}{2} \|\mathbf{u}(T) - \mathbf{u}_{T,d}\|_{L^2(\Omega; \mathbb{R}^d)}^2, \\ \psi^3(\mathbf{u}) &= \frac{1}{2} \|\varepsilon(\mathbf{u}(T)) - \varepsilon_{T,d}\|_{L^2(\Omega; \mathbb{S})}^2. \end{aligned}$$

Here, $\mathbf{u}_d \in L^2(0, T; L^2(\Omega; \mathbb{R}^d))$, $\mathbf{u}_{T,d} \in L^2(\Omega; \mathbb{R}^d)$ and $\varepsilon_{T,d} \in L^2(\Omega; \mathbb{S})$ are desired displacements and strains, respectively.

Let us give some examples for the admissible set U_{ad} . The sets

$$\begin{aligned} U_{\text{ad}}^1 &= \{\mathbf{g} \in H^1(0, T; U) : \mathbf{g}(0) = \mathbf{0}, \|\mathbf{g}\|_{L^2(\Gamma_N; \mathbb{R}^d)} \leq \rho \text{ a.e. in } (0, T)\} \\ U_{\text{ad}}^2 &= \{\mathbf{g} \in H^1(0, T; U) : \mathbf{g}(0) = \mathbf{g}(T) = \mathbf{0}\} \end{aligned}$$

satisfy [Assumption 2.8](#) for every $\rho \geq 0$.

Of special interest is the combination of ψ^2 and U_{ad}^2 . Since we consider a quasistatic process, the condition $\mathbf{g}(T) = \mathbf{0}$ implies that at time $t = T$ the solid body is unloaded. Due to the plastic behaviour, the remaining (and lasting) deformation $\mathbf{u}(T)$ is typically non-zero. Due to the choice of ψ^2 , the displacement is controlled towards the desired deformation $\mathbf{u}_{T,d}$. Thus, choosing ψ^2 and U_{ad}^2 allows the control of the springback of the solid body. This is of great interest in applications, e.g. deep-drawing of metal sheets.

Let us show the existence of an optimal control. Using standard arguments, see, e.g. [\[Tröltzsch, 2010, Theorem 4.15\]](#), the existence of a global minimizer of [\(P\)](#) is a straightforward consequence of the weak continuity of $\mathcal{G} \circ E$ proven in [Theorem 2.7](#).

Theorem 2.9. There exists a global minimizer of [\(P\)](#).

Proof. In virtue of the control-to-state map $\mathcal{G} \circ E$, we define the reduced objective $f(\mathbf{g}) := F(\mathcal{G}^{\mathbf{u}}(E\mathbf{g}), \mathbf{g})$. Due to [Assumption 2.8](#), f is bounded from below, hence $\inf f$ on U_{ad} exists. We define $j := \inf\{f(\mathbf{g}), \mathbf{g} \in U_{\text{ad}}\} \in \mathbb{R}$.

Let $\{\mathbf{g}_k\} \subset U_{\text{ad}}$ be a minimizing sequence, i.e. $\lim f(\mathbf{g}_k) = j$. Since ψ is bounded from below and $\nu > 0$, the sequence $\{\mathbf{g}_k\}$ is bounded in $H^1(0, T; U)$. Hence, there is a weakly convergent subsequence, denoted by the same symbol. We denote the weak limit by \mathbf{g} . The closeness and convexity of U_{ad} implies the weak closeness. Therefore, $\{\mathbf{g}_k\} \subset U_{\text{ad}}$ implies $\mathbf{g} \in U_{\text{ad}}$.

Since $\mathcal{G}^{\mathbf{u}} \circ E$ is weakly continuous from $H_0^1(0, T; U)$ to $H_0^1(0, T; V)$, see [Theorem 2.7](#), the weak lower semicontinuity of ψ and $\|\cdot\|_{H^1(0, T; U)}^2$ implies that the reduced objective

$$f = F(\mathcal{G}^{\mathbf{u}} \circ E(\cdot), \cdot) = \psi(\mathcal{G}^{\mathbf{u}} \circ E(\cdot)) + \frac{\nu}{2} \|\cdot\|_{H^1(0, T; U)}^2$$

is weakly lower semicontinuous. Therefore,

$$\begin{aligned} j &= \lim f(\mathbf{g}_k) && (\mathbf{g}_k \text{ is a minimizing sequence}) \\ &\geq f(\mathbf{g}) && (\text{weak lower semicontinuity}) \\ &\geq j && (j = \inf f \text{ and } \mathbf{g} \in U_{\text{ad}}). \end{aligned}$$

This shows $F(\mathcal{G}^{\mathbf{u}}(E\mathbf{g}), \mathbf{g}) = f(\mathbf{g}) = j$, hence \mathbf{g} is a global minimizer.

3 Time discretization

In this section we study a discretization in time of the optimal control problem [\(P\)](#). The time-discrete version of the forward system [\(VI\)](#) is introduced in [Section 3.1](#). The strong convergence of the time-discrete solutions $(\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau)$ is shown in [Section 3.2](#). In particular, we prove a new rate of convergence of $(\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau)$ in $L^\infty(0, T; S^2 \times V)$ without assuming additional regularity of $\boldsymbol{\Sigma}$, see [Lemma 3.2](#). Moreover, we show the convergence of $(\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau)$ in $H^1(0, T; S^2 \times V)$ (by using an idea of [Krejčí \[1998\]](#)) and the convergence of λ^τ in $L^2(0, T; L^2(\Omega))$, see [Theorems 3.3](#) and [3.4](#), respectively. These strong convergence results are new for the analysis of the quasistatic problem [\(VI\)](#) and both are essential for passing to the limit in the optimality system in [Section 7](#) in the second part of the paper.

In [Section 3.3](#) we prove the weak convergence of $(\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau)$ in $H^1(0, T; S^2 \times V)$ assuming the weak convergence of the controls \mathbf{g}^τ in $H^1(0, T; U)$. This result is necessary in order to show in [Section 3.4](#) that *every* local minimizer of [\(P\)](#) can be approximated by local minimizers of (slightly modified) time-discrete problems [\(P \$^\tau\$ \)](#), see [Theorem 3.10](#). This approximability of local minimizers is essential for the derivation of *necessary* optimality conditions which are satisfied for *every* local minimizer of [\(P\)](#) in [Section 7](#) of the second part.

Notation. Throughout the paper we partition the time horizon $[0, T]$ into N intervals, each of constant length $\tau = T/N$ for simplicity. We use a superscript τ , i.e. $(\cdot)^\tau$, to indicate variables and operators associated with the discretization in time.

For a time-discrete variable $f^\tau \in X^N$, where X is some Banach space, the components of f^τ are denoted by f_i^τ , $i = 1, \dots, N$. When necessary, we may refer to f_0^τ with a pre-defined value (interpreted as an initial condition), mostly $f_0^\tau = 0$.

If $f^\tau \in X^N$ is the time-discretization of a variable $f \in H_0^1(0, T; X)$, we identify f^τ with its *linear* interpolation, i.e., for $t \in [(i-1)\tau, i\tau]$ we define

$$f^\tau(t) = \frac{t - (i-1)\tau}{\tau} f_i^\tau + \frac{i\tau - t}{\tau} f_{i-1}^\tau, \quad (3.1)$$

with $f_0^\tau = 0$. Therefore, X^N is identified with a subspace of $H_0^1(0, T; X)$. We make use of this identification for the variables $\boldsymbol{\Sigma}$, \mathbf{u} and \mathbf{g} . For later reference, we remark

$$\dot{f}^\tau(t) = \frac{1}{\tau} (f_i^\tau - f_{i-1}^\tau) \quad \text{and} \quad f_i^\tau = f^\tau(t) - (t - i\tau) \dot{f}^\tau(t) \quad (3.2)$$

for almost all $t \in ((i-1)\tau, i\tau)$ and $i = 1, \dots, N$.

On the other hand, if $f^\tau \in X^N$ is the discretization in time of a variable belonging to $L^2(0, T; X)$, we identify f^τ with a piecewise *constant* function in $L^2(0, T; X)$, i.e., for $t \in [(i-1)\tau, i\tau)$ we define

$$f^\tau(t) = f_i^\tau. \quad (3.3)$$

This is used for the plastic multiplier λ .

3.1. Introduction of a discretization in time of the forward problem. In this section we introduce a discretization in time of the forward problem [\(VI\)](#). Replacing the time derivatives by backward differences, we obtain the discretized problem: given $\ell^\tau \in (V')^N$, find $(\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau) \in (S^2 \times V)^N$ such that $\boldsymbol{\Sigma}_i^\tau \in \mathcal{K}$ and

$$\langle A(\boldsymbol{\Sigma}_i^\tau - \boldsymbol{\Sigma}_{i-1}^\tau) + B^*(\mathbf{u}_i^\tau - \mathbf{u}_{i-1}^\tau), \mathbf{T} - \boldsymbol{\Sigma}_i^\tau \rangle \geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K}, \quad (3.4a)$$

$$B\boldsymbol{\Sigma}_i^\tau = \ell_i^\tau \quad \text{in } V', \quad (3.4b)$$

holds for all $i \in \{1, \dots, N\}$, where $(\boldsymbol{\Sigma}_0^\tau, \mathbf{u}_0^\tau) = \mathbf{0}$. We denote the solution operator which maps $\ell^\tau \rightarrow (\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau)$ by \mathcal{G}^τ . Let us mention that, for fixed i , [\(3.4\)](#) can be interpreted as the solution of a *static* plasticity problem. Several properties of this time discretization are proven in [[Han and Reddy, 1999](#), Proof of Theorem 8.12, page 196]. We recall the results which are important for the following analysis.

In order to show an analog to [\(1.19\)](#) for the time-discrete problem, we test [\(3.4a\)](#) with $\mathbf{T} = \boldsymbol{\Sigma}_i^\tau + (\tau, -\tau)$, where $\tau \in S$ is arbitrary. Using $(\boldsymbol{\Sigma}_0^\tau, \mathbf{u}_0^\tau) = \mathbf{0}$ we obtain

$$\mathbb{C}^{-1}\boldsymbol{\sigma}^\tau - \boldsymbol{\varepsilon}(\mathbf{u}^\tau) - \mathbb{H}^{-1}\boldsymbol{\chi}^\tau = \mathbf{0} \quad \text{a.e. in } (0, T) \times \Omega. \quad (3.5)$$

In [[Han and Reddy, 1999](#), Lemma 8.8 and (8.37)], we find the a-priori estimate

$$\|\boldsymbol{\Sigma}^\tau\|_{H^1(0, T; S^2)} \leq C \|\ell^\tau\|_{H^1(0, T; V')}. \quad (3.6)$$

Together with the inf-sup condition [\(1.13\)](#) of B^* and [\(3.5\)](#), we obtain

$$\|\boldsymbol{\Sigma}^\tau\|_{H^1(0, T; S^2)} + \|\mathbf{u}^\tau\|_{H^1(0, T; V)} \leq C \|\ell^\tau\|_{H^1(0, T; V')}. \quad (3.7)$$

Due to this a-priori estimate, [Han and Reddy \[1999\]](#) are able to prove the weak convergence of $(\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau) = \mathcal{G}^\tau(\ell^\tau)$ towards the solution $(\boldsymbol{\Sigma}, \mathbf{u}) = \mathcal{G}(\ell)$ of [\(1.18\)](#) in $H^1(0, T; S^2 \times V)$ if the data ℓ_i^τ in [\(3.4\)](#) is chosen as the point evaluation of ℓ .

From an optimization point of view, this result is too weak for two reasons. First, the convergence is not w.r.t. the strong topology in $H^1(0, T; S^2 \times V)$. Second, the convergence is proven for a restricted choice of ℓ^τ . We therefore strengthen the results of [Han and Reddy \[1999\]](#) in [Theorem 3.3](#).

As for the continuous problem, the time-discrete problem can be formulated as a complementarity system. The results in [[Herzog et al., 2011b](#), Section 2] prove the existence of $\lambda^\tau \in L^2(\Omega)^N$ such that the complementarity system

$$A(\boldsymbol{\Sigma}_i^\tau - \boldsymbol{\Sigma}_{i-1}^\tau) + B^*(\mathbf{u}_i^\tau - \mathbf{u}_{i-1}^\tau) + \tau \lambda_i^\tau \mathcal{D}^* \mathcal{D} \boldsymbol{\Sigma}_i^\tau = \mathbf{0} \quad \text{in } S^2, \quad (3.8a)$$

$$B \boldsymbol{\Sigma}_i^\tau = \ell_i^\tau \quad \text{in } V', \quad (3.8b)$$

$$0 \leq \lambda_i^\tau \quad \perp \quad \phi(\boldsymbol{\Sigma}_i^\tau) \leq 0 \quad \text{a.e. in } \Omega \quad (3.8c)$$

is satisfied. Here, we scaled the multiplier λ^τ appropriately.

Similarly to [\(3.5\)](#) we obtain a representation formula for λ^τ ,

$$-\mathbb{C}^{-1}(\boldsymbol{\sigma}_i^\tau - \boldsymbol{\sigma}_{i-1}^\tau) + \boldsymbol{\varepsilon}(\mathbf{u}_i^\tau - \mathbf{u}_{i-1}^\tau) = -\mathbb{H}^{-1}(\boldsymbol{\chi}_i^\tau - \boldsymbol{\chi}_{i-1}^\tau) = \tau \lambda_i^\tau \mathcal{D} \boldsymbol{\Sigma}_i^\tau. \quad (3.9)$$

Remark 3.1. The same time discretization approach is used in [Krejčí \[1996, 1998\]](#) to prove the existence of a solution of [\(EVI\)](#). Indeed, by applying the time discretization [[Krejčí, 1998](#), (3.6)] to the reduced formulation [\(2.5\)](#), we obtain: find $\boldsymbol{\Sigma}_i^\tau \in \mathcal{K}_{\ell_i^\tau}$, such that

$$\langle A(\boldsymbol{\Sigma}_i^\tau - \boldsymbol{\Sigma}_{i-1}^\tau), \mathbf{T} - \boldsymbol{\Sigma}_i^\tau \rangle \geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K}_{\ell_i^\tau}.$$

As discussed for the continuous problem [\(1.15\)](#), we can introduce a Lagrange multiplier associated with $B \boldsymbol{\Sigma}_i^\tau = \ell_i^\tau$ such that $(\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau)$ satisfy [\(3.4\)](#).

In the following two subsections we prove that the discretization in time converges in a strong and a weak sense, similar to the continuity properties of the forward operator \mathcal{G} proven in [Sections 2.2](#) and [2.3](#). Under the assumption that ℓ^τ converges strongly to ℓ in the space $H^1(0, T; V')$ or that \mathbf{g}^τ converges weakly in the space $H^1(0, T; U)$, the corresponding solutions $(\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau)$ are shown to converge strongly (or weakly) to $(\boldsymbol{\Sigma}, \mathbf{u})$ in $H^1(0, T; S^2 \times V)$, respectively.

Both statements will be needed in [Section 3.4](#) to show the approximability of local solutions of [\(P\)](#) mentioned in the introduction of this section.

3.2. Strong convergence. The strong convergence of the discretization in time of an EVI can be found in [[Krejčí, 1996](#), Proposition I.3.11]. He proves the convergence in $W^{1,1}(0, T; X)$ of the time-discrete solutions. The data of the time-discretized problems are taken as the point evaluations of the data of the continuous problem. We use these ideas to prove the convergence of $(\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau)$ in $H^1(0, T; S^2 \times V)$ whenever the right-hand sides ℓ^τ converge in $H^1(0, T; V')$ as $\tau \searrow 0$. Moreover, we show the convergence of λ^τ in $L^2(0, T; L^2(\Omega))$ as $\tau \searrow 0$. This is a new result for the analysis of [\(VI\)](#) and a crucial ingredient for proving the optimality system in [Section 7](#) in Part II of this paper.

Let us consider a sequence of time steps $\{\tau_k\}$ and a sequence of loads $\ell^{\tau_k} \in (V')^{N_k}$, with $N_k \tau_k = T$, such that the linear interpolants $\ell^{\tau_k} \in H_0^1(0, T; V')$ converge in $H^1(0, T; V')$ towards some $\ell \in H_0^1(0, T; V')$. We denote by $(\boldsymbol{\Sigma}^{\tau_k}, \mathbf{u}^{\tau_k}) \in (S^2 \times V)^N$ the solutions of the time-discrete problem [\(3.4a\)](#), as well as their piecewise linear interpolants. The aim of this section is to prove $(\boldsymbol{\Sigma}^{\tau_k}, \mathbf{u}^{\tau_k}) \rightarrow (\boldsymbol{\Sigma}, \mathbf{u})$ in $H^1(0, T; S^2 \times V)$, where $(\boldsymbol{\Sigma}, \mathbf{u})$ is the solution to the continuous problem [\(1.18\)](#).

For simplicity of notation, we omit the index k and we refer to “ $\ell^{\tau_k} \rightarrow \ell$ if $k \rightarrow \infty$ ” simply as “ $\ell^\tau \rightarrow \ell$ as $\tau \searrow 0$ ”.

The proof of [Theorem 3.3](#) relies on the convergence argument [Lemma A.1](#) which is tailored to the analysis of EVIs. Its prerequisite [\(A.1a\)](#) is the convergence of Σ^τ in $L^\infty(0, T; S^2)$. This is addressed in a preliminary step and required in the proof of [Theorem 3.3](#).

Lemma 3.2. Let $\ell^\tau \rightarrow \ell$ in $W^{1,1}(0, T; V')$ as $\tau \searrow 0$. Then

$$\|\Sigma - \Sigma^\tau\|_{L^\infty(0, T; S^2)}^2 \leq C (\|\dot{\ell} - \dot{\ell}^\tau\|_{L^1(0, T; V')}^2 + \tau). \quad (3.10)$$

In particular, $\Sigma^\tau \rightarrow \Sigma$ in $L^\infty(0, T; S^2)$ as $\tau \searrow 0$.

Proof. We have

$$\frac{1}{2} \frac{d}{dt} \|\Sigma(t) - \Sigma^\tau(t)\|_A^2 = \langle A\dot{\Sigma}(t) - A\dot{\Sigma}^\tau(t), \Sigma(t) - \Sigma^\tau(t) \rangle \quad \text{f.a.a. } t \in (0, T). \quad (3.11)$$

Let us fix some $t \in ((i-1)\tau, i\tau)$, $i \in \{1, \dots, N\}$. Since $\Sigma^\tau(t)$ is a convex combination of Σ_i^τ and Σ_{i-1}^τ , see [\(3.1\)](#), and $\Sigma_{i-1}^\tau, \Sigma_i^\tau \in \mathcal{K}$, we obtain $\Sigma^\tau(t) \in \mathcal{K}$. Hence, $\mathbf{T} = \Sigma^\tau(t) + \Sigma_{\ell-\ell^\tau}(t) \in \mathcal{K}$. Using $B\mathbf{T} = \ell(t)$, [\(1.18a\)](#) implies

$$\langle A\dot{\Sigma}(t), \Sigma^\tau(t) - \Sigma(t) + \Sigma_{\ell-\ell^\tau}(t) \rangle \geq 0. \quad (3.12)$$

Testing [\(3.4a\)](#) with $\mathbf{T} = \Sigma(t) + \Sigma_{\ell^\tau-\ell(t)}$ yields

$$\langle A(\Sigma_i^\tau - \Sigma_{i-1}^\tau), \Sigma(t) - \Sigma_i^\tau + \Sigma_{\ell^\tau-\ell(t)} \rangle \geq 0.$$

Using [\(3.2\)](#) we obtain

$$\langle A\dot{\Sigma}^\tau(t), \Sigma(t) - \Sigma^\tau(t) - (i\tau - t)\dot{\Sigma}^\tau + \Sigma_{\ell^\tau(t)-\ell(t)+(i\tau-t)\dot{\ell}^\tau(t)} \rangle \geq 0.$$

Together with [\(3.12\)](#) this yields

$$\begin{aligned} \langle A\dot{\Sigma}(t) - A\dot{\Sigma}^\tau(t), \Sigma(t) - \Sigma^\tau(t) \rangle &\leq \langle A\dot{\Sigma}(t) - A\dot{\Sigma}^\tau(t), \Sigma_{\ell-\ell^\tau}(t) \rangle \\ &\quad + (i\tau - t) \langle A\dot{\Sigma}^\tau(t), -\dot{\Sigma}^\tau(t) + \Sigma_{\dot{\ell}^\tau}(t) \rangle. \end{aligned}$$

Integrating over t and using [\(3.11\)](#), together with the boundedness of $\dot{\Sigma}^\tau$ and $\Sigma_{\dot{\ell}^\tau}$ in $L^2(0, T; S^2)$, see [\(3.6\)](#), yields

$$\frac{1}{2} \|\Sigma(t) - \Sigma^\tau(t)\|_A^2 \leq \int_0^t \langle A\dot{\Sigma}(s) - A\dot{\Sigma}^\tau(s), \Sigma_{\ell-\ell^\tau}(s) \rangle ds + C\tau.$$

Integrating by parts and using $\Sigma(0) = \Sigma^\tau(0) = \mathbf{0}$ gives

$$\begin{aligned} \frac{1}{2} \|\Sigma(t) - \Sigma^\tau(t)\|_A^2 &\leq \langle A\Sigma(t) - A\Sigma^\tau(t), \Sigma_{\ell-\ell^\tau}(t) \rangle \\ &\quad - \int_0^t \langle A\Sigma(s) - A\Sigma^\tau(s), \Sigma_{\dot{\ell}-\dot{\ell}^\tau}(s) \rangle ds + C\tau \\ &\leq C (\|\Sigma(t) - \Sigma^\tau(t)\|_{S^2} \|\ell(t) - \ell^\tau(t)\|_{V'} \\ &\quad + \|\Sigma - \Sigma^\tau\|_{L^\infty(0, T; S^2)} \|\dot{\ell} - \dot{\ell}^\tau\|_{L^1(0, T; V')} + \tau). \end{aligned}$$

Taking the supremum over $t \in (0, T)$ on both sides shows

$$\begin{aligned} \|\Sigma - \Sigma^\tau\|_{L^\infty(0, T; S^2)}^2 &\leq C (\|\Sigma - \Sigma^\tau\|_{L^\infty(0, T; S^2)} \|\ell - \ell^\tau\|_{L^\infty(0, T; V')} \\ &\quad + \|\Sigma - \Sigma^\tau\|_{L^\infty(0, T; S^2)} \|\dot{\ell} - \dot{\ell}^\tau\|_{L^1(0, T; V')} + \tau). \end{aligned}$$

Finally, Young's inequality and $W^{1,1}(0, T; V') \hookrightarrow L^\infty(0, T; V')$ yields

$$\|\Sigma - \Sigma^\tau\|_{L^\infty(0, T; S^2)}^2 \leq C (\|\dot{\ell} - \dot{\ell}^\tau\|_{L^1(0, T; V')}^2 + \tau).$$

Let us comment on the case when ℓ^τ is the point evaluation of ℓ . The estimate (3.10) shows the order of convergence of $\tau^{1/2}$ w.r.t. the L^∞ -norm provided that $\dot{\ell} \in W^{1,1}(0, T; V')$. This rate of convergence is a new result for the time-discretization of (VI).

In [Han and Reddy, 1999, Theorem 13.1] the estimate

$$\|\Sigma - \Sigma^\tau\|_{L^\infty(0, T; S^2)} \leq c \tau \|\Sigma\|_{W^{2,1}(0, T; S^2)}$$

is shown under the assumption $\Sigma \in W^{2,1}(0, T; S^2)$. However, this assumption is unlikely to hold for the solution of an EVI. There are many examples where the solution of an EVI only belongs to $W^{1,\infty}(0, T; X)$, even if X is one-dimensional and the input is smooth, see e.g. [Krejčí and Lovicar, 1990, Examples 1, 2]. Hence, $\Sigma \in W^{2,1}(0, T; S^2)$ cannot be considered the generic case.

Theorem 3.3. Let us assume $\ell^\tau \rightarrow \ell$ in $H_0^1(0, T; V')$ as $\tau \searrow 0$. Then $(\Sigma^\tau, \mathbf{u}^\tau) \rightarrow (\Sigma, \mathbf{u})$ in $H_0^1(0, T; S^2 \times V)$ as $\tau \searrow 0$.

Proof. We follow the idea of [Krejčí, 1998, Proof of Theorem 3.6]. We will apply Lemma A.1 with the setting

$$\begin{aligned} X &= S_A^2, & u_n &= \Sigma_{\ell^\tau} - 2\Sigma^\tau, & u_0 &= \Sigma_\ell - 2\Sigma, \\ p &= 2, & g_n &= \|\Sigma_{\dot{\ell}^\tau}(\cdot)\|_A, & g_0 &= \|\Sigma_{\dot{\ell}}(\cdot)\|_A. \end{aligned} \quad (3.13)$$

to show the convergence of Σ^τ . Here, S_A^2 denotes the Hilbert space S^2 equipped with the inner product induced by A . To apply Lemma A.1, we have to verify the prerequisites (A.1c) and (A.1d).

By definition of the discrete problem (3.4), we have for a fixed $i \in \{1, 2, \dots, N\}$

$$\langle A(\Sigma_i^\tau - \Sigma_{i-1}^\tau), \mathbf{T} - \Sigma_i^\tau \rangle \geq 0 \quad \text{for all } \mathbf{T} \in \mathcal{K}, \quad B\mathbf{T} = \ell_i^\tau,$$

which is the time-discrete analog to (1.15). We choose $\mathbf{T} = \Sigma_{i-1}^\tau + \Sigma_{\ell_i^\tau - \ell_{i-1}^\tau}$, see (2.4) for the definition of Σ_ℓ . By definition of Σ_ℓ we have $B\mathbf{T} = \ell_i^\tau$ and the shift invariance (1.7) implies $\mathbf{T} \in \mathcal{K}$. This shows

$$\langle A(\Sigma_i^\tau - \Sigma_{i-1}^\tau), \Sigma_{i-1}^\tau + \Sigma_{\ell_i^\tau - \ell_{i-1}^\tau} - \Sigma_i^\tau \rangle \geq 0.$$

Therefore dividing by τ^2 and using the notation of linear interpolants, see (3.1), implies

$$\langle A\dot{\Sigma}^\tau(t), \Sigma_{\dot{\ell}^\tau}(t) - \dot{\Sigma}^\tau(t) \rangle \geq 0 \quad \text{f.a.a. } t \in (0, T).$$

This shows

$$\|\Sigma_{\dot{\ell}^\tau}(t) - 2\dot{\Sigma}^\tau(t)\|_A \leq \|\Sigma_{\dot{\ell}^\tau}(t)\|_A \quad \text{f.a.a. } t \in (0, T), \quad (3.14)$$

where $\|\cdot\|_A$ is the norm on S^2 induced by A , see (2.2).

To derive a similar formula for the continuous solution, we test (1.15) with $\mathbf{T} = \Sigma(t+h) + \Sigma_{\ell(t)-\ell(t+h)}$ for $h > 0$ and obtain

$$\langle A\dot{\Sigma}(t), \Sigma_{\ell(t)-\ell(t+h)} + \Sigma(t+h) - \Sigma(t) \rangle \geq 0.$$

Passing to the limit $h \searrow 0$, using [Krejčí, 1998, Theorem 8.14], yields

$$\langle A\dot{\Sigma}(t), \Sigma_{\dot{\ell}}(t) - \dot{\Sigma}(t) \rangle \geq 0 \quad \text{f.a.a. } t \in (0, T).$$

Analogously, we obtain by $h \nearrow 0$

$$\langle A\dot{\Sigma}(t), \Sigma_{\dot{\ell}}(t) - \dot{\Sigma}(t) \rangle \leq 0 \quad \text{f.a.a. } t \in (0, T).$$

Hence

$$\|\Sigma_{\dot{\ell}}(t) - 2\dot{\Sigma}(t)\|_A = \|\Sigma_{\dot{\ell}}(t)\|_A \quad \text{f.a.a. } t \in (0, T). \quad (3.15)$$

In order to apply [Lemma A.1](#) with the setting [\(3.13\)](#) we check the prerequisites.

[\(A.1a\)](#): [Lemma 3.2](#) ensures $\Sigma^\tau \rightarrow \Sigma$ in $L^\infty(0, T; S_A^2)$ and the assumption $\ell^\tau \rightarrow \ell$ in $H^1(0, T; V')$ implies $\Sigma_{\ell^\tau} \rightarrow \Sigma_\ell$ in $L^\infty(0, T; S^2)$.

[\(A.1b\)](#): By the linearity of $\ell \mapsto \Sigma_\ell$, the assumption $\ell^\tau \rightarrow \ell$ in $H^1(0, T; V')$ implies $\|\Sigma_{\ell^\tau}(\cdot)\|_A \rightarrow \|\Sigma_{\ell}(\cdot)\|_A$ in $L^2(0, T; \mathbb{R})$.

[\(A.1c\)](#): This was shown in [\(3.14\)](#).

[\(A.1d\)](#): This was shown in [\(3.15\)](#).

Therefore, [Lemma A.1](#) yields

$$\|\Sigma_{\ell^\tau} - 2\Sigma^\tau - (\Sigma_\ell - 2\Sigma)\|_{H^1(0, T; S_A^2)} \rightarrow 0.$$

By the assumption $\ell^\tau \rightarrow \ell$ in $H^1(0, T; V')$ we infer

$$\|\Sigma^\tau - \Sigma\|_{H^1(0, T; S_A^2)} \rightarrow 0 \quad \text{as } \tau \searrow 0.$$

Using [\(1.19\)](#), [\(3.5\)](#) and the inf-sup condition of B , we obtain $\mathbf{u}^\tau \rightarrow \mathbf{u}$ in $H^1(0, T; V)$ as $\tau \searrow 0$.

Using the representation formula [\(3.9\)](#) for λ^τ we are able to prove the strong convergence of λ^τ in $L^2(0, T; L^2(\Omega))$ towards λ . We remark that this is a novel result for the analysis of [\(1.22\)](#).

Let us mention that this strong convergence is crucial for the analysis in the second part of the paper. Without this strong convergence, we could neither pass to the limit in the complementarity relation $\mu^\tau \lambda^\tau = 0$ nor in the adjoint equation, see [Section 7](#) in the second part of the paper. Hence, this convergence is essential to prove the necessity of the system of weakly stationary type [\(7.35\)–\(7.40\)](#) in the second part.

Theorem 3.4. Let $\ell^\tau \rightarrow \ell$ in $H_0^1(0, T; V')$ as $\tau \searrow 0$. Then $\lambda^\tau \rightarrow \lambda$ in $L^2(0, T; L^2(\Omega))$ as $\tau \searrow 0$.

Proof. Step (1): We show that $\lambda^\tau \rightarrow \lambda$ in $L^1(0, T; L^1(\Omega))$. Using $\lambda_i^\tau |\mathcal{D}\Sigma_i^\tau|^2 = \tilde{\sigma}_0^2 \lambda_i^\tau$ by [\(3.8c\)](#) and [\(3.9\)](#), we obtain

$$\lambda_i^\tau = \frac{1}{\tilde{\sigma}_0^2} \lambda_i^\tau \mathcal{D}\Sigma_i^\tau : \mathcal{D}\Sigma_i^\tau = \frac{1}{\tau \tilde{\sigma}_0^2} (-\mathbb{H}^{-1}(\boldsymbol{\chi}_i^\tau - \boldsymbol{\chi}_{i-1}^\tau) : \mathcal{D}\Sigma_i^\tau).$$

Using the notion of piecewise linear interpolations $\boldsymbol{\chi}^\tau$, Σ^τ and the piecewise constant interpolation λ^τ , we obtain for $t \in ((i-1)\tau, i\tau)$, see [\(3.2\)](#),

$$\lambda^\tau(t) = \frac{-1}{\tilde{\sigma}_0^2} \left(\mathbb{H}^{-1} \dot{\boldsymbol{\chi}}^\tau(t) : [\mathcal{D}\Sigma^\tau(t) - (t - i\tau) \mathcal{D}\dot{\Sigma}^\tau(t)] \right).$$

Using $\lambda = -\mathbb{H}^{-1} \dot{\boldsymbol{\chi}} : \mathcal{D}\Sigma / \tilde{\sigma}_0^2$, by [\(1.22\)](#), this implies

$$\begin{aligned} \tilde{\sigma}_0^2 |\lambda^\tau(t) - \lambda(t)| &\leq \left| \mathbb{H}^{-1} \dot{\boldsymbol{\chi}}^\tau(t) : \mathcal{D}\Sigma^\tau(t) - \mathbb{H}^{-1} \dot{\boldsymbol{\chi}}(t) : \mathcal{D}\Sigma(t) \right| \\ &\quad + (t - i\tau) \left| \mathbb{H}^{-1} \dot{\boldsymbol{\chi}}^\tau(t) : \mathcal{D}\dot{\Sigma}^\tau(t) \right| \\ &\leq \left| \mathbb{H}^{-1} \dot{\boldsymbol{\chi}}^\tau(t) : \mathcal{D}\Sigma^\tau(t) - \mathbb{H}^{-1} \dot{\boldsymbol{\chi}}(t) : \mathcal{D}\Sigma(t) \right| \\ &\quad + \tau \left| \mathbb{H}^{-1} \dot{\boldsymbol{\chi}}^\tau(t) : \mathcal{D}\dot{\Sigma}^\tau(t) \right|. \end{aligned}$$

Hence,

$$\begin{aligned} \|\lambda^\tau - \lambda\|_{L^1(0,T;L^1(\Omega))} &\leq c \left(\|\dot{\chi}^\tau\| \|\mathcal{D}\Sigma^\tau - \mathcal{D}\Sigma\| + \|\mathcal{D}\Sigma\| \|\dot{\chi}^\tau - \dot{\chi}\| \right. \\ &\quad \left. + \tau \|\dot{\chi}^\tau\| \|\mathcal{D}\dot{\Sigma}^\tau\| \right) \quad \text{a.e. in } \Omega, \end{aligned}$$

where all norms on the right-hand side are those of $L^2(0, T; S)$. Using $\Sigma^\tau \rightarrow \Sigma$ in $H^1(0, T; S^2)$ implies $\lambda^\tau \rightarrow \lambda$ in $L^1(0, T; L^1(\Omega))$.

Step (2): We show that $\|\lambda^\tau\|_{L^2(0,T;L^2(\Omega))} \rightarrow \|\lambda\|_{L^2(0,T;L^2(\Omega))}$. Using that λ^τ is the piecewise constant interpolation of $(\lambda_i^\tau)_{i=1}^N$, we obtain

$$\|\lambda^\tau\|_{L^2(0,T;L^2(\Omega))}^2 = \tau \sum_{i=1}^N \int_{\Omega} (\lambda_i^\tau)^2 dx.$$

Using $\lambda_i^\tau |\mathcal{D}\Sigma_i^\tau| = \tilde{\sigma}_0 \lambda_i^\tau$ by (3.8c), we obtain

$$\begin{aligned} \|\lambda^\tau\|_{L^2(0,T;L^2(\Omega))}^2 &= \frac{\tau}{\tilde{\sigma}_0^2} \sum_{i=1}^N \int_{\Omega} |\lambda_i^\tau \mathcal{D}\Sigma_i^\tau|^2 dx \\ &= \frac{\tau}{\tilde{\sigma}_0^2} \sum_{i=1}^N \int_{\Omega} \left| -\mathbb{H}^{-1} \frac{\chi_i^\tau - \chi_{i-1}^\tau}{\tau} \right|^2 dx \quad (\text{by (3.9)}) \\ &= \frac{1}{\tilde{\sigma}_0^2} \|\mathbb{H}^{-1} \dot{\chi}^\tau\|_{L^2(0,T;S)}^2. \end{aligned}$$

An analogous calculation shows $\|\lambda\|_{L^2(0,T;L^2(\Omega))} = \frac{1}{\tilde{\sigma}_0} \|\mathbb{H}^{-1} \dot{\chi}\|_{L^2(0,T;S)}$. Hence, by $\Sigma^\tau \rightarrow \Sigma$ in $H^1(0, T; S^2)$, see Theorem 3.3, we infer the convergence of norms $\|\lambda^\tau\|_{L^2(0,T;L^2(\Omega))} \rightarrow \|\lambda\|_{L^2(0,T;L^2(\Omega))}$.

Step (3): We show that $\lambda^\tau \rightarrow \lambda$ in $L^2(0, T; L^2(\Omega))$. By Step (2) we know that for every subsequence of τ , there is a subsequence τ_k such that $\lambda^{\tau_k} \rightharpoonup \tilde{\lambda}$ in $L^2(0, T; L^2(\Omega))$ for some $\tilde{\lambda}$. Step (1) implies $\tilde{\lambda} = \lambda$. Hence, the whole sequence converges weakly. In view of the convergence of norms, λ^τ converges strongly to λ in $L^2(0, T; L^2(\Omega))$ as $\tau \searrow 0$.

3.3. Weak convergence. In this section we show the weak convergence of the time-discrete states $(\Sigma^\tau, \mathbf{u}^\tau) = \mathcal{G}^\tau(E\mathbf{g}^\tau)$ in $H_0^1(0, T; S^2 \times V)$, under the assumption that the controls \mathbf{g}^τ converges weakly towards \mathbf{g} in $H_0^1(0, T; U)$. This result is essential for proving the approximability results in Section 3.4. Similarly to the weak continuity of the forward operator proven in Section 2.3, we need the weak convergence of the right-hand sides $E\mathbf{g}^\tau$ with respect to a stronger norm in space, see the discussion in the beginning of Section 2.3.

Theorem 3.5. Let $\mathbf{g}^\tau \in U^n$ such that the linear interpolants converge weakly, i.e., there is $\mathbf{g} \in H_0^1(0, T; U)$ such that $\mathbf{g}^\tau \rightharpoonup \mathbf{g}$ in $H_0^1(0, T; U)$ as $\tau \searrow 0$.

Then the solutions $(\Sigma^\tau, \mathbf{u}^\tau)$ of the time-discrete problem (3.4) with right-hand side $\ell^\tau = E\mathbf{g}^\tau$ converge weakly in $H_0^1(0, T; S^2 \times V)$ towards the solution (Σ, \mathbf{u}) of (1.18) with right-hand side $\ell = E\mathbf{g}$ as $\tau \searrow 0$.

Proof. Due to the a-priori bound (3.7), there exists a weakly convergent subsequence of $(\Sigma^\tau, \mathbf{u}^\tau)$, denoted by the same symbol, i.e. $(\Sigma^\tau, \mathbf{u}^\tau) \rightharpoonup (\tilde{\Sigma}, \tilde{\mathbf{u}})$ in $H^1(0, T; S^2 \times V)$ as $\tau \searrow 0$. As in the proof of Theorem 2.7, we shall show $(\tilde{\Sigma}, \tilde{\mathbf{u}}) = (\Sigma, \mathbf{u})$. Hence the whole sequence converges weakly and this simplification of notation is justified.

Let $\mathbf{T} \in L^2(0, T; S^2)$ with $\mathbf{T} \in \mathcal{K}$ a.e. in $(0, T)$ be arbitrary. Let $i \in \{0, \dots, N-1\}$ and $\kappa \in (0, 1)$ be given. Set $t = (i + \kappa)\tau$. Testing (3.4a) with $\mathbf{T}(t)$ yields

$$\langle A\dot{\Sigma}^\tau(t) + B^*\dot{\mathbf{u}}^\tau(t), \mathbf{T}(t) - \Sigma^\tau((i+1)\tau) \rangle \geq 0.$$

Using $\Sigma^\tau((i+1)\tau) = \Sigma^\tau(t) + (1-\kappa)\tau\dot{\Sigma}^\tau(t)$ by (3.1), and integrating over $\kappa \in (0, 1)$, i.e. $t \in (i\tau, (i+1)\tau)$ yields

$$\begin{aligned} 0 &\leq \int_{i\tau}^{(i+1)\tau} \langle A\dot{\Sigma}^\tau(t) + B^*\dot{\mathbf{u}}^\tau(t), \mathbf{T}(t) - \Sigma^\tau((i+1)\tau) \rangle dt \\ &= \int_{i\tau}^{(i+1)\tau} \langle A\dot{\Sigma}^\tau + B^*\dot{\mathbf{u}}^\tau, \mathbf{T} - \Sigma^\tau \rangle dt - \frac{\tau}{2} \int_{i\tau}^{(i+1)\tau} \langle A\dot{\Sigma}^\tau + B^*\dot{\mathbf{u}}^\tau, \dot{\Sigma}^\tau \rangle dt. \end{aligned}$$

Hence, summing over $i = 0, \dots, N-1$ implies

$$0 \leq \int_0^T \langle A\dot{\Sigma}^\tau + B^*\dot{\mathbf{u}}^\tau, \mathbf{T} - \Sigma^\tau \rangle dt - \frac{\tau}{2} \int_0^T \langle A\dot{\Sigma}^\tau + B^*\dot{\mathbf{u}}^\tau, \dot{\Sigma}^\tau \rangle dt. \quad (3.16)$$

Due to the boundedness of $(\Sigma^\tau, \mathbf{u}^\tau)$ in $H^1(0, T; S^2 \times V)$, the second addend goes to zero as $\tau \searrow 0$. For the first addend we can proceed as in the proof of [Theorem 2.7](#): the application of [Lemma 2.6](#) yields

$$\limsup_{\tau \searrow 0} \int_0^T \langle A\dot{\Sigma}^\tau + B^*\dot{\mathbf{u}}^\tau, \mathbf{T} - \Sigma^\tau \rangle dt \leq \int_0^T \langle A\dot{\Sigma} + B^*\dot{\mathbf{u}}, \mathbf{T} - \tilde{\Sigma} \rangle dt.$$

Together with (3.16) this implies

$$0 \leq \int_0^T \langle A\dot{\Sigma} + B^*\dot{\mathbf{u}}, \mathbf{T} - \tilde{\Sigma} \rangle dt \quad \text{for all } \mathbf{T} \in L^2(0, T; S^2), \mathbf{T} \in \mathcal{K} \text{ a.e. in } (0, T).$$

Therefore, $(\tilde{\Sigma}, \tilde{\mathbf{u}})$ is the solution of (1.18). This implies $(\tilde{\Sigma}, \tilde{\mathbf{u}}) = (\Sigma, \mathbf{u})$.

3.4. The time-discrete optimal control problem. In this section we consider a time discretization of (P). The time-discrete controls belong to the space U^N , which is identified with a subspace of $H_0^1(0, T; U)$ via the piecewise linear interpolation (3.1). The discretization of the admissible set is given by $U_{\text{ad}}^\tau = U_{\text{ad}} \cap U^N$, where $\tau = T/N$. Now, the time-discrete optimal control problem reads

$$\left. \begin{aligned} &\text{Minimize } F(\mathbf{u}^\tau, \mathbf{g}^\tau) = \psi(\mathbf{u}^\tau) + \frac{\nu}{2} \|\mathbf{g}^\tau\|_{H^1(0, T; U)}^2 \\ &\text{such that } (\Sigma^\tau, \mathbf{u}^\tau) = \mathcal{G}^\tau(E\mathbf{g}^\tau) \\ &\text{and } \mathbf{g}^\tau \in U_{\text{ad}}^\tau. \end{aligned} \right\} \quad (\mathbf{P}^\tau)$$

For the analysis of (P $^\tau$) we need the additional assumption that all $\mathbf{g} \in U_{\text{ad}}$ can be approximated by time-discrete controls $\mathbf{g}^\tau \in U_{\text{ad}}^\tau$.

Assumption 3.6. In addition to [Assumption 2.8](#) we suppose that for all $\mathbf{g} \in U_{\text{ad}}$, there exists $\mathbf{g}^\tau \in U_{\text{ad}}^\tau$, such that $\mathbf{g}^\tau \rightarrow \mathbf{g}$ in $H^1(0, T; U)$ as $\tau \searrow 0$.

We remark that this condition is satisfied for both examples of U_{ad} given after [Assumption 2.8](#).

The aim of this section is to answer the question whether optimal controls of (P) can be approximated by optimal controls of (P $^\tau$). Similar results for a regularization of an optimal control problem of static plasticity have been proved in [[Herzog et al., 2010](#), Section 3.2]. These results are very common for the approximation of minimizers of optimal control problems and the technique of proof is applicable to various problems, see also [Casas and Tröltzsch \[2002\]](#), [Barbu \[1981\]](#). To be precise, we prove that

- one global optimum of (\mathbf{P}) can be approximated by global solutions of (\mathbf{P}^τ) , see [Lemma 3.8](#),
- every strict local optimum of (\mathbf{P}) can be approximated by local solutions of (\mathbf{P}^τ) , see [Lemma 3.9](#),
- every local optimum of (\mathbf{P}) can be approximated by local solutions of a perturbed problem (\mathbf{P}_g^τ) , see [Theorem 3.10](#).

These results are very important for the analysis in [Section 7](#) of the second part of the paper. Without these results at hand, one cannot show the necessity of the optimality conditions therein.

Although both of \mathbf{u}^τ and \mathbf{g}^τ are optimization variables in (\mathbf{P}^τ) , for simplicity we refer solely to \mathbf{g}^τ being (locally or globally) optimal. This is justified in view of the continuous solution map \mathcal{G}^τ of [\(3.4\)](#).

First we check that there exists a minimizer of the time-discrete optimal control problem (\mathbf{P}^τ) .

Lemma 3.7. There exists a global minimizer of (\mathbf{P}^τ) .

Proof. In [Section 5.1](#) of the second part we will show in a more general framework that the solution operator \mathcal{G}^τ of [\(3.4\)](#) is Lipschitz continuous from $(V')^N$ to $(S^2 \times V)^N$, see [Proposition 5.8](#). Since the control operator E from U^N into $(V')^N$ is compact, \mathcal{G}^τ is compact from U^N to $(S^2 \times V)^N$. Now we proceed similarly as in the proof of [Theorem 2.9](#).

Now, we are going to prove the approximation properties.

Lemma 3.8. Suppose that [Assumption 3.6](#) is fulfilled. Let $\{\tau\}$ be a sequence tending to 0 and let \mathbf{g}^τ denote a global solution to (\mathbf{P}^τ) .

- (1) Then there exists an accumulation point \mathbf{g} of $\{\mathbf{g}^\tau\}$ in $H^1(0, T; U)$ and
- (2) every weak accumulation point of $\{\mathbf{g}^\tau\}$ in $H^1(0, T; U)$ is in fact a strong accumulation point in $H^1(0, T; U)$ and a global solution of (\mathbf{P}) .

Proof. By [Assumption 2.8](#), we can choose $\mathbf{g}_0 \in U_{\text{ad}}$. By [Assumption 3.6](#), there exists a sequence $\mathbf{g}_0^\tau \in U_{\text{ad}}^\tau$ which converges towards \mathbf{g}_0 in $H_0^1(0, T; U)$. Hence, by [\(3.7\)](#) the corresponding displacements \mathbf{u}_0^τ converge towards \mathbf{u}_0 in $H_0^1(0, T; V)$. Together with the continuity of ψ , this implies the convergence of $F(\mathbf{u}_0^\tau, \mathbf{g}_0^\tau)$. Since \mathbf{g}^τ is a global optimum of (\mathbf{P}^τ) , we have $F(\mathbf{u}^\tau, \mathbf{g}^\tau) \leq F(\mathbf{u}_0^\tau, \mathbf{g}_0^\tau)$. Hence, $\{\mathbf{g}^\tau\}$ is bounded in $H^1(0, T; U)$. This implies the existence of a weakly convergent subsequence in this space. Therefore, assertion (1) follows by assertion (2).

To prove assertion (2), let $\{\mathbf{g}^\tau\}$ converge weakly towards \mathbf{g} in $H^1(0, T; U)$ as $\tau \searrow 0$. We denote by $(\boldsymbol{\Sigma}^\tau, \mathbf{u}^\tau) = \mathcal{G}^\tau(E\mathbf{g}^\tau)$ the (time-discrete) solution to [\(3.4\)](#) and by $(\boldsymbol{\Sigma}, \mathbf{u}) = \mathcal{G}(E\mathbf{g})$ the solution to [\(1.18\)](#). Due to the weak convergence proven in [Theorem 3.5](#), we have $\mathbf{u}^\tau \rightharpoonup \mathbf{u}$ in $H^1(0, T; V)$.

Let $\tilde{\mathbf{g}} \in U_{\text{ad}}$ with corresponding displacement $\tilde{\mathbf{u}}$ be arbitrary. By [Assumption 3.6](#), there is $\tilde{\mathbf{g}}^\tau \in U_{\text{ad}}^\tau$, such that $\tilde{\mathbf{g}}^\tau \rightarrow \tilde{\mathbf{g}}$ in $H^1(0, T; L^2(\Gamma_N; \mathbb{R}^d))$. We denote the corresponding displacements by $\tilde{\mathbf{u}}^\tau$. By [Theorem 3.3](#) we infer $\tilde{\mathbf{u}}^\tau \rightarrow \tilde{\mathbf{u}}$. We have

$$\begin{aligned} F(\mathbf{u}, \mathbf{g}) &\leq \liminf F(\mathbf{u}^\tau, \mathbf{g}^\tau) && \text{by weak lower semicontinuity of } F \\ &\leq \liminf F(\tilde{\mathbf{u}}^\tau, \tilde{\mathbf{g}}^\tau) && \text{by global optimality of } (\mathbf{u}^\tau, \mathbf{g}^\tau) \\ &= F(\tilde{\mathbf{u}}, \tilde{\mathbf{g}}). && \text{by convergence of } (\tilde{\mathbf{u}}^\tau, \tilde{\mathbf{g}}^\tau) \end{aligned}$$

This shows that \mathbf{g} is a global optimal solution. Inserting $\tilde{\mathbf{g}} = \mathbf{g}$ yields the convergence of norms and hence the strong convergence of \mathbf{g}^τ in $H^1(0, T; U)$.

Lemma 3.9. Suppose [Assumption 3.6](#) is fulfilled. Let \mathbf{g} be a strict local optimum of (\mathbf{P}) w.r.t. the topology of $H^1(0, T; U)$. Then, for every sequence $\{\tau\}$ tending to 0, there is a sequence $\{\mathbf{g}^\tau\}$ of local solutions to (\mathbf{P}^τ) , such that $\mathbf{g}^\tau \rightarrow \mathbf{g}$ strongly in $H^1(0, T; U)$ as $\tau \searrow 0$.

Proof. Step (1): Let $\varepsilon > 0$, such that \mathbf{g} is the unique global optimum in the closed ball $B_\varepsilon(\mathbf{g})$ with radius ε centered at \mathbf{g} . We define $\hat{U}_{\text{ad}} = U_{\text{ad}} \cap B_\varepsilon(\mathbf{g})$.

Step (2): We check that [Assumption 3.6](#) is fulfilled for \hat{U}_{ad} . Clearly we have $\mathbf{g} \in \hat{U}_{\text{ad}}$. Let $\tilde{\mathbf{g}} \in \hat{U}_{\text{ad}}$ be arbitrary. We construct a sequence $\{\tilde{\mathbf{g}}^\tau\} \subset \hat{U}_{\text{ad}}$, $\tilde{\mathbf{g}}^\tau \in U^N$, such that $\tilde{\mathbf{g}}^\tau \rightarrow \tilde{\mathbf{g}}$. Since U_{ad} satisfies [Assumption 3.6](#), there are sequences of time-discrete functions $(\mathbf{g}^\tau), (\hat{\mathbf{g}}^\tau) \subset U_{\text{ad}}$, converging towards \mathbf{g} and $\tilde{\mathbf{g}}$, respectively.

If $\|\tilde{\mathbf{g}} - \mathbf{g}\|_{H^1(0, T; U)} < \varepsilon$, the convergence $\hat{\mathbf{g}}^\tau \rightarrow \tilde{\mathbf{g}}$ implies $\hat{\mathbf{g}}^\tau \in B_\varepsilon(\mathbf{g})$, and hence $\hat{\mathbf{g}}^\tau \in \hat{U}_{\text{ad}} \cap U_{\text{ad}}^\tau$ for sufficiently small τ .

Otherwise, if $\|\tilde{\mathbf{g}} - \mathbf{g}\|_{H^1(0, T; U)} = \varepsilon$, we obtain $\lim\|\mathbf{g} - \hat{\mathbf{g}}^\tau\|_{H^1(0, T; U)} = \varepsilon$. Therefore, there exists τ_0 , such that for all $\tau \leq \tau_0$, we have $\|\mathbf{g} - \mathbf{g}^\tau\|_{H^1(0, T; U)} \leq \varepsilon/2$ and $\|\mathbf{g} - \hat{\mathbf{g}}^\tau\|_{H^1(0, T; U)} > \varepsilon/2$. Now, we construct a convex combination $\tilde{\mathbf{g}}^\tau$ of \mathbf{g}^τ and $\hat{\mathbf{g}}^\tau$ which belongs to $B_\varepsilon(\mathbf{g})$ and approximates $\tilde{\mathbf{g}}$. We define the sequence $\tilde{\mathbf{g}}^\tau$ by

$$\tilde{\mathbf{g}}^\tau = (1 - \eta^\tau) \hat{\mathbf{g}}^\tau + \eta^\tau \mathbf{g}^\tau, \quad \text{with } \eta^\tau = \max\left(0, \frac{\|\mathbf{g} - \hat{\mathbf{g}}^\tau\|_{H^1(0, T; U)} - \varepsilon}{\|\mathbf{g} - \hat{\mathbf{g}}^\tau\|_{H^1(0, T; U)} - \varepsilon/2}\right) \in [0, 1].$$

We have $\|\mathbf{g} - \tilde{\mathbf{g}}^\tau\|_{H^1(0, T; U)} \leq \varepsilon$ by construction, and $\eta^\tau \in [0, 1]$ implies $\tilde{\mathbf{g}}^\tau \in U_{\text{ad}}$. Therefore, $\tilde{\mathbf{g}}^\tau \in \hat{U}_{\text{ad}}$. Further, $\lim\|\mathbf{g} - \tilde{\mathbf{g}}^\tau\|_{H^1(0, T; U)} = \varepsilon$ implies $\eta^\tau \rightarrow 0$ and hence $\tilde{\mathbf{g}}^\tau \rightarrow \tilde{\mathbf{g}}$. This shows that \hat{U}_{ad} satisfies [Assumption 3.6](#).

Step (3): We define the auxiliary problem

$$\left. \begin{array}{l} \text{Minimize } F(\mathbf{u}, \mathbf{g}) = \psi(\mathbf{u}) + \frac{\nu}{2} \|\mathbf{g}\|_{H^1(0, T; U)}^2 \\ \text{such that } (\boldsymbol{\Sigma}, \mathbf{u}) = \mathcal{G}(E\mathbf{g}) \\ \text{and } \mathbf{g} \in U_{\text{ad}} \cap B_\varepsilon(\mathbf{g}). \end{array} \right\} \quad (\mathbf{P}_{\mathbf{g}, \varepsilon})$$

Since \mathbf{g} is the unique minimum in $\hat{U}_{\text{ad}} = U_{\text{ad}} \cap B_\varepsilon(\mathbf{g})$, it is the unique global minimizer of $(\mathbf{P}_{\mathbf{g}, \varepsilon})$. Invoking [Lemma 3.8](#) yields the existence of a sequence \mathbf{g}^τ of global solutions of the associated time-discrete problem $(\mathbf{P}_\varepsilon^\tau)$, such that $\mathbf{g}^\tau \rightarrow \mathbf{g}$ in $H^1(0, T; U)$.

Step (4): It remains to check that \mathbf{g}^τ are local solutions of (\mathbf{P}^τ) . The convergence $\mathbf{g}^\tau \rightarrow \mathbf{g}$ in $H^1(0, T; U)$ implies that there is a τ_0 such that $\|\mathbf{g} - \mathbf{g}^\tau\|_{H^1(0, T; U)} \leq \varepsilon/2$ holds for all $\tau \leq \tau_0$. Let $\tau \leq \tau_0$ and $\tilde{\mathbf{g}}^\tau \in B_{\varepsilon/2}(\mathbf{g}^\tau) \cap U_{\text{ad}}^\tau$ be arbitrary. By the triangle inequality we infer $\tilde{\mathbf{g}}^\tau \in B_\varepsilon(\mathbf{g})$. This implies $\tilde{\mathbf{g}}^\tau \in \hat{U}_{\text{ad}} \cap U_{\text{ad}}^\tau$. The global optimality of \mathbf{g}^τ implies $F(\mathcal{G}^{\tau, \mathbf{u}}(\mathbf{g}^\tau), \mathbf{g}^\tau) \leq F(\mathcal{G}^{\tau, \mathbf{u}}(\tilde{\mathbf{g}}^\tau), \tilde{\mathbf{g}}^\tau)$. Hence, \mathbf{g}^τ is a local optimum of (\mathbf{P}^τ) in the neighborhood $B_{\varepsilon/2}(\mathbf{g}^\tau)$.

Finally, we address the approximability of a local minimum, which is not assumed to be strict. Let \mathbf{g} be a local optimum of (\mathbf{P}) w.r.t. the topology of $H^1(0, T; U)$. We define the modified problem, see also [Casas and Tröltzsch \[2002\]](#), [Barbu \[1981\]](#),

$$\left. \begin{array}{l} \text{Minimize } F_{\mathbf{g}}(\mathbf{u}, \tilde{\mathbf{g}}) = \psi(\mathbf{u}) + \frac{\nu}{2} \|\tilde{\mathbf{g}}\|_{H^1(0, T; U)}^2 + \frac{1}{2} \|\tilde{\mathbf{g}} - \mathbf{g}\|_{H^1(0, T; U)}^2 \\ \text{such that } (\boldsymbol{\Sigma}, \mathbf{u}) = \mathcal{G}(E\tilde{\mathbf{g}}) \\ \text{and } \tilde{\mathbf{g}} \in U_{\text{ad}} \end{array} \right\} \quad (\mathbf{P}_{\mathbf{g}})$$

Clearly, \mathbf{g} becomes a *strict* local optimum of $(\mathbf{P}_{\mathbf{g}})$. Analogously to (\mathbf{P}^τ) we define the time-discrete approximation $(\mathbf{P}_{\mathbf{g}}^\tau)$.

Theorem 3.10. Suppose [Assumption 3.6](#) is fulfilled. Let \mathbf{g} be a local optimum of [\(P\)](#) w.r.t. the topology of $H^1(0, T; U)$. Then, for every sequence $\{\tau\}$ tending to 0, there is a sequence (\mathbf{g}^τ) of local solutions to [\(P \$_{\mathbf{g}}^\tau\$ \)](#), such that $\mathbf{g}^\tau \rightarrow \mathbf{g}$ strongly in $H^1(0, T; U)$ as $\tau \searrow 0$.

Proof. Since the additional term $\|\tilde{\mathbf{g}} - \mathbf{g}\|_{H^1(0, T; U)}^2$ is weakly lower semicontinuous, this follows analogously to [Lemma 3.9](#).

Due to this theorem, we are able to derive *necessary* optimality conditions for [\(P\)](#) by passing to the limit with the optimality conditions of [\(P \$_{\mathbf{g}}^\tau\$ \)](#).

A A convergence result in Bochner-Sobolev spaces

For the proof of [Theorem 3.3](#) we need the following result. It is a consequence of [\[Krejčí, 1998, Theorems 8.5, 8.7\]](#) which are cited as [Theorems A.4](#) and [A.5](#) below.

Lemma A.1. Let X be a Hilbert space and $p \in [1, \infty)$ be given. Let $\{u_n\} \subset W^{1,p}(0, T; X)$, $\{g_n\} \subset L^p(0, T; \mathbb{R})$ be given sequences for $n \in \mathbb{N} \cup \{0\}$ such that

$$u_n \rightarrow u_0 \quad \text{in } L^\infty(0, T; X) \quad (\text{A.1a})$$

$$g_n \rightarrow g_0 \quad \text{in } L^p(0, T; \mathbb{R}), \quad (\text{A.1b})$$

$$\|\dot{u}_n(t)\|_X \leq g_n(t) \text{ a.e. in } (0, T), \text{ for all } n \in \mathbb{N}, \quad (\text{A.1c})$$

$$\|\dot{u}_0(t)\|_X = g_0(t) \text{ a.e. in } (0, T). \quad (\text{A.1d})$$

Then $u_n \rightarrow u_0$ in $W^{1,p}(0, T; X)$.

In order to prove [Lemma A.1](#), we show two auxiliary results.

Lemma A.2. Let X be a normed linear space. The sequence $\{x_n\} \subset X$ is assumed to be bounded in X and let $V \subset X'$ be some subset of the dual space X' . If for some $x \in X$

$$f(x_n) \rightarrow f(x) \quad \text{as } n \rightarrow \infty \quad (\text{A.2})$$

holds for all $f \in V$, then [\(A.2\)](#) holds for all $f \in \text{cl} V$, where the closure is taken with respect to the X' -norm.

Proof. Let $f \in \text{cl} V$ be given. Then there exists $\{f_n\} \subset V$ with $f_n \rightarrow f$. Due to the boundedness of $\{x_n\}$, there is a $C > 0$ with $\|x_n\|_X \leq C$ and $\|x\|_X \leq C$. Let $\varepsilon > 0$ be given. Then there is a $m \in \mathbb{N}$ such that $\|f_m - f\|_{X'} \leq \varepsilon/(3C)$. Since [\(A.2\)](#) holds for f_m , there exists $n \in \mathbb{N}$ such that $|f_m(x_n) - f_m(x)| \leq \varepsilon/3$. This implies

$$\begin{aligned} |f(x_n) - f(x)| &\leq |f(x_n) - f_m(x_n)| + |f_m(x_n) - f_m(x)| + |f_m(x) - f(x)| \\ &\leq \|f - f_m\|_{X'} \|x_n\|_X + |f_m(x_n) - f_m(x)| + \|f - f_m\|_{X'} \|x\|_X \leq \varepsilon. \end{aligned}$$

This proves that [\(A.2\)](#) holds for f .

Lemma A.3. Let X be some Hilbert space and let $\{u_n\} \subset W^{1,1}(0, T; X)$ and $u \in W^{1,1}(0, T; X)$ be given. If $u_n \rightarrow u$ in $L^\infty(0, T; X)$ and if $\{u_n\}$ is bounded in $W^{1,1}(0, T; X)$, then

$$\int_0^T \langle \dot{u}_n(t), \varphi(t) \rangle dt \rightarrow \int_0^T \langle \dot{u}(t), \varphi(t) \rangle dt \quad \text{for all } \varphi \in C(0, T; X), \quad (\text{A.3})$$

i.e. the sequence $\{\dot{u}_n\}$ converges in $C(0, T; X)'$ with respect to the weak-*topology.

Proof. Let us prove that (A.3) holds for all $\varphi \in C^1(0, T; X)$. Due to $W^{1,1}(0, T; X) \subset C(0, T; X)$, integrating by parts (see [Krejčí, 1998, (8.24)]) and the convergence of u_n in $L^\infty(0, T; X)$ yields

$$\begin{aligned} \int_0^T \langle \dot{u}_n(t), \varphi(t) \rangle dt &= - \int_0^T \langle u_n(t), \dot{\varphi}(t) \rangle dt + \langle u_n(T), \varphi(T) \rangle - \langle u_n(0), \varphi(0) \rangle \\ &\rightarrow - \int_0^T \langle u(t), \dot{\varphi}(t) \rangle dt + \langle u(T), \varphi(T) \rangle - \langle u(0), \varphi(0) \rangle \\ &= \int_0^T \langle \dot{u}(t), \varphi(t) \rangle dt \end{aligned}$$

for all $\varphi \in C^1(0, T; X)$. Using $\text{cl } C^1(0, T; X) = C(0, T; X)$ with respect to the $\|\cdot\|_{L^\infty(0, T; X)}$ -norm, Lemma A.2 yields the claim.

We remark that Lemma A.3 follows directly from [Krejčí, 1998, Theorem 8.16], with the settings

$$u_n := u := \varphi, \quad \xi_n := v_n, \quad \xi := v,$$

since $\text{Var}_{[0, T]} v_n \leq c$ holds due to the boundedness of v_n in $W^{1,1}(0, T; X)$. The direct proof is presented for convenience of the reader.

Theorem A.4 ([Krejčí, 1998, Theorem 8.7]). Let X be a Hilbert space and $\{v_n\} \subset L^1(0, T; X)$, $\{g_n\} \subset L^1(0, T; \mathbb{R})$ be given sequences for $n \in \mathbb{N} \cup \{0\}$ such that

$$\lim_{n \rightarrow \infty} \int_0^T \langle v_n(t) - v_0(t), \varphi(t) \rangle dt = 0 \quad \text{for all } \varphi \in C(0, T; X), \quad (\text{A.4a})$$

$$g_n \rightarrow g_0 \quad \text{in } L^1(0, T; \mathbb{R}), \quad (\text{A.4b})$$

$$\|v_n(t)\|_X \leq g_n(t) \text{ a.e. in } (0, T), \text{ for all } n \in \mathbb{N}, \quad (\text{A.4c})$$

$$\|v_0(t)\|_X = g_0(t) \text{ a.e. in } (0, T). \quad (\text{A.4d})$$

Then $v_n \rightarrow v_0$ in $L^1(0, T; X)$.

Theorem A.5 ([Krejčí, 1998, Theorem 8.5]). Let B be a Banach space and $p \in [1, \infty)$ be given. Let $\{v_n\} \subset L^p(0, T; B)$, $\{g_n\} \subset L^p(0, T; \mathbb{R})$ be given sequences for $n \in \mathbb{N} \cup \{0\}$ such that

$$v_n(t) \rightarrow v_0(t) \text{ a.e. in } (0, T), \quad (\text{A.5a})$$

$$g_n \rightarrow g_0 \quad \text{in } L^p(0, T; \mathbb{R}), \quad (\text{A.5b})$$

$$\|v_n(t)\|_B \leq g_n(t) \text{ a.e. in } (0, T), \text{ for all } n \in \mathbb{N} \cup \{0\}, \quad (\text{A.5c})$$

Then $v_n \rightarrow v_0$ in $L^p(0, T; B)$.

Corollary A.6. Let X be a Hilbert space and $p \in [1, \infty)$ be given. Let $\{v_n\} \subset L^p(0, T; X)$, $\{g_n\} \subset L^p(0, T; \mathbb{R})$ be given sequences for $n \in \mathbb{N} \cup \{0\}$ such that

$$\lim_{n \rightarrow \infty} \int_0^T \langle v_n(t) - v_0(t), \varphi(t) \rangle dt = 0 \quad \text{for all } \varphi \in C(0, T; X), \quad (\text{A.6a})$$

$$g_n \rightarrow g_0 \quad \text{in } L^p(0, T; \mathbb{R}), \quad (\text{A.6b})$$

$$\|v_n(t)\|_X \leq g_n(t) \text{ a.e. in } (0, T), \text{ for all } n \in \mathbb{N}, \quad (\text{A.6c})$$

$$\|v_0(t)\|_X = g_0(t) \text{ a.e. in } (0, T). \quad (\text{A.6d})$$

Then $v_n \rightarrow v_0$ in $L^p(0, T; X)$.

Proof. Applying [Theorem A.4](#) yields $v_n \rightarrow v$ in $L^1(0, T; X)$. Thus there exists a pointwise convergent subsequence $\{v_{n_k}\}$. Hence, [Theorem A.5](#) yields $v_{n_k} \rightarrow v$ in $L^p(0, T; X)$. Since we can start with an arbitrary subsequence of v_n , we obtain $v_n \rightarrow v$ in $L^p(0, T; X)$.

Proof of [Lemma A.1](#). By [\(A.1b\)](#) and [\(A.1c\)](#) we obtain the boundedness of u_n in $W^{1,1}(0, T; X)$. [Lemma A.3](#) yields [\(A.3\)](#) and hence we can apply [Corollary A.6](#) for $v_n = \dot{u}_n$. This yields $\dot{u}_n \rightarrow \dot{u}_0$ in $L^p(0, T; X)$ and hence $u_n \rightarrow u_0$ in $W^{1,p}(0, T; X)$.

Acknowledgment. The author would like to express his gratitude to Roland Herzog and Christian Meyer for helpful discussions on the topic of this paper and to Dorothee Knees for pointing out reference [Krejčí \[1998\]](#).

This work was supported by a DFG grant within the [Priority Program SPP 1253 \(Optimization with Partial Differential Equations\)](#), which is gratefully acknowledged.

References

- V. Barbu. Necessary conditions for distributed control problems governed by parabolic variational inequalities. *SIAM Journal on Control and Optimization*, 19(1):64–86, 1981. ISSN 0363-0129. doi: [10.1137/0319006](#).
- F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers. *RAIRO Modélisation Mathématique et Analyse Numérique*, 8:129–151, 1974.
- M. Brokate and P. Krejčí. Wellposedness of kinematic hardening models in elastoplasticity. *Mathematical Modelling and Numerical Analysis*, 32:177–209, 1998.
- H. Brézis. *Opérateurs maximaux monotones*, volume 5. North-Holland mathematics studies, Amsterdam, 1973.
- E. Casas and F. Tröltzsch. Error estimates for the finite-element approximation of a semilinear elliptic control problem. *Control and Cybernetics*, 31:695–712, 2002.
- J. Diestel and J. J. Uhl. *Vector Measures*. Mathematical Surveys and Monographs. American Mathematical Society, Providence, 1977.
- H. Gajewski, K. Gröger, and K. Zacharias. *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*. Akademie-Verlag, Berlin, 1974.
- P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston, 1985.
- K. Gröger. Evolution equations in the theory of plasticity. In *Theory of nonlinear operators (Proc. Fifth Internat. Summer School, Central Inst. Math. Mech. Acad. Sci. GDR, Berlin, 1977)*, volume 6 of *Abh. Akad. Wiss. DDR, Abt. Math. Naturwiss. Tech.*, 1978, pages 97–107. Akademie-Verlag, Berlin, 1978a.
- K. Gröger. Zur Theorie des quasi-statischen Verhaltens von elastisch-plastischen Körpern. *Zeitschrift für Angewandte Mathematik und Mechanik. Ingenieurwissenschaftliche Forschungsarbeiten*, 58(2):81–88, 1978b.
- K. Gröger. Initial value problems for elastoplastic and elastoviscoplastic systems. In *Nonlinear analysis, function spaces and applications (Proceedings of Spring School, Horni Bradlo, 1978)*, pages 95–127. Teubner, Leipzig, 1979.
- K. Gröger. A $W^{1,p}$ -estimate for solutions to mixed boundary value problems for second order elliptic differential equations. *Mathematische Annalen*, 283:679–687, 1989. doi: [10.1007/BF01442860](#).
- R. Haller-Dintelmann, C. Meyer, J. Rehberg, and A. Schiela. Hölder continuity and optimal control for nonsmooth elliptic problems. *Applied Mathematics and Optimization*, 60(3):397–428, 2009. doi: [10.1007/s00245-009-9077-x](#).

- W. Han and B. D. Reddy. *Plasticity*. Springer, New York, 1999.
- R. Herzog, C. Meyer, and G. Wachsmuth. C-stationarity for optimal control of static plasticity with linear kinematic hardening. *submitted*, 2010.
- R. Herzog, C. Meyer, and G. Wachsmuth. Integrability of displacement and stresses in linear and nonlinear elasticity with mixed boundary conditions. *Journal of Mathematical Analysis and Applications*, 382(2):802–813, 2011a. doi: [10.1016/j.jmaa.2011.04.074](https://doi.org/10.1016/j.jmaa.2011.04.074).
- R. Herzog, C. Meyer, and G. Wachsmuth. Existence and regularity of the plastic multiplier in static and quasistatic plasticity. *GAMM Reports*, 34(1):39–44, 2011b. doi: [10.1002/gamm.201110006](https://doi.org/10.1002/gamm.201110006).
- P. Krejčí. *Hysteresis, Convexity and Dissipation in Hyperbolic Equations*, volume 8 of *GAKUTO International Series Mathematical Sciences and Applications*. Gakkōtoshō, 1996.
- P. Krejčí. Evolution variational inequalities and multidimensional hysteresis operators. Technical Report 432, Weierstrass Institute for Applied Analysis and Stochastics (WIAS), 1998.
- P. Krejčí and V. Lovicar. Continuity of hysteresis operators in sobolev spaces. *Applications of Mathematics*, 35:60–66, 1990.
- J. J. Moreau. Evolution problem associated with a moving convex set in a hilbert space. *Journal of Differential Equations*, 26:347–374, 1977.
- M. Růžička. *Nichtlineare Funktionalanalysis*. Springer, 2004.
- J. Simon. Compact sets in the space $L^p(0, T; B)$. *Annali di Matematica Pura ed Applicata*, 146:65–96, 1986. doi: [10.1007/BF01762360](https://doi.org/10.1007/BF01762360).
- F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.
- K. Yosida. *Functional Analysis*. Springer, 1965.

CHEMNITZ UNIVERSITY OF TECHNOLOGY, FACULTY OF MATHEMATICS, D–09107 CHEMNITZ, GERMANY

E-mail address: gerd.wachsmuth@mathematik.tu-chemnitz.de

URL: http://www.tu-chemnitz.de/mathematik/part_dgl/people/wachsmuth