Optimal Control of Quasistatic Plasticity with Linear Kinematic Hardening - Part I: Existence and Discretization in Time

July 2011

Preprint-Number SPP1253-118
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 the 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 perturbed, time-discrete problems.

1 Introduction

In this paper we consider an optimal control problem for the quasistatic problem of small-strain elastoplasticity. The forward system in the stress-based (so-called dual) form is represented by a time-dependent variational inequality (VI) of mixed type: find generalized stresses $\Sigma(t) \in S^2$ and displacements $u(t) \in V$ which satisfy $\Sigma(t) \in K$ and

\[(A \dot{\Sigma}(t) + B^* \dot{u}(t), T - \Sigma(t))_{S^2} \geq 0 \quad \text{for all } T \in K,\]

\[B \Sigma(t) = \ell(t) \quad \text{in } V',\]

where $A : S^2 \to S^2$ and $B : S^2 \to V'$ are bounded linear operators. Moreover, (VI) is subject to the initial condition $(\Sigma(0), u(0)) = (0, 0)$. The convex, closed set $K \subset S^2$ of admissible stresses is determined by the von Mises yield condition. The details are made precise below.

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. A first step in this direction is the optimal control of static plasticity which was considered in Herzog et al. [2010, 2011c]. For locally optimal controls, systems of B- and C-stationary type were obtained. These results are of importance here as well, since a discretization in time of the quasistatic problem (VI) leads to a system of static plasticity, see [Han and Reddy, 1999, Section 8.2] and Section 3.

Since the derivation of optimality conditions for optimal control of quasistatic plasticity is more involved, this work is split into two parts. Let us give a brief overview over their contents. In the first part, we review some known results concerning the solution map $G : \ell \mapsto (\Sigma, u)$ of the problem (VI) and derive some new continuity properties. Due to the weak continuity of the solution map $G$, we are able to prove the existence of optimal controls for the optimal control problem (P) defined below. Moreover, we state a time discretized counterpart of (VI) and of the optimal control problem, and show the convergence of these discretizations.

In the second part, we regularize the time-discrete problem and show the Fréchet differentiability of the associated solution map $G^\varepsilon$. 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 \( \varepsilon \) 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 (VI).

We give some references for the analysis of the forward problem (VI). References concerning optimal control of time-dependent variational inequalities can be found in Han and Reddy [1999, Chapters 2, 3]. In Han and Reddy [1999, Chapter 7, 8] continuity properties of the solution map \( G \) of (VI) are derived. They prove that \( G : \ell \mapsto (\Sigma, u) \) is Lipschitz continuous from \( H^1(0,T;V') \) to \( L^\infty(0,T;S^2 \times V) \). 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].

In order to derive optimality conditions, the continuity result mentioned above is not sufficient. In addition, we need the continuity of \( G : H^1(0,T;V') \to H^1(0,T;S^2 \times V) \) on two occasions. First, this continuity is needed to prove the approximability of local solutions by time-discrete minimizers in Theorem 3.8. Second, the strong convergence of \( (\Sigma, \tilde{u}) \) 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 \( G : H^1(0,T;V') \to H^1(0,T;S^2 \times V) \) is shown 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).

Let us sketch the outline of the paper. In the remainder of this section, we introduce the notation and fix the functional analytic framework. In Section 2 we state some properties of the continuous optimal control problem and give a brief introduction to evolution variational inequalities (EVIs). Restating (VI) as an EVI, we are able to show the continuity of its solution operator from \( H^1(0,T;V') \) to \( H^1(0,T;S^2 \times V) \). Afterwards, we prove the weak continuity of the solution operator from \( H^1(0,T;U) \) to \( H^1(0,T;S^2 \times V) \). Here, \( U \) is the control space which is compactly embedded in \( V' \). These continuity properties of the solution map of (VI) are new and needed in Section 2.4 to conclude the existence of an optimal control.

Section 3 is devoted to the discretization in time of the control problem (P). We show two convergence properties of the solution of the time-discrete problems, similar to the continuity properties of (VI). Finally, these results are used in Section 3.4 to prove that every local minimizer \( g \) of the control problem (P) can be approximated by local minimizers of a slightly modified time-discrete problem (P\( g^\varepsilon \)). In the second part this will be an essential ingredient for deducing that the necessary optimality condition holds for every minimizer of (P), see Section 7.

Part I of the paper can be understood as a prerequisite for the analysis of 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.

1.1. Presentation of the optimal control problem. As shown in Herzog et al. [2011b], the time-dependent variational inequality (VI) can be equivalently written
as a complementarity system. We give further comments on this equivalent reformulation in Section 1.3. Since the formulation as a complementarity system is more suited for the analysis of the optimal control problem, a multiplier \( \lambda \) (termed plastic multiplier in mechanical engineering) associated with the constraint \( \phi(\Sigma) \leq 0 \) is introduced. Then, the strong formulation of the quasistatic problem of elastoplasticity with linear kinematic hardening and the von Mises yield condition reads

\[
\begin{aligned}
C^{-1}\dot{\sigma} + \varepsilon(\dot{u}) + \lambda(\sigma^D + \chi^D) &= 0 \quad \text{in} \quad (0,T) \times \Omega, \\
H^{-1}\dot{\chi} + \lambda(\sigma^D + \chi^D) &= 0 \quad \text{in} \quad (0,T) \times \Omega, \\
\text{div} \sigma &= -f \quad \text{in} \quad (0,T) \times \Omega, \\
0 \leq \lambda \perp \phi(\Sigma) &\leq 0 \quad \text{in} \quad (0,T) \times \Omega, \\
\sigma \cdot n &= g \quad \text{on} \quad (0,T) \times \Gamma_N, \quad \text{and boundary conditions} \\
\end{aligned}
\]  

(1.1)

The state variables consist of the stress \( \sigma \) and back stress \( \chi \) which are combined into the generalized stress \( \Sigma = (\sigma, \chi) \). The state variables also comprise the displacement \( u \) and the plastic multiplier \( \lambda \) associated with the yield condition \( \phi(\Sigma) \leq 0 \) of von Mises type. The tensors \( C^{-1} \) and \( H^{-1} \) are the inverses of the elasticity tensor and the hardening modulus, respectively, and \( \sigma^D \) denotes the deviatoric part of \( \sigma \). As usual, \( \varepsilon(v) = (\nabla v + (\nabla v)^\top)/2 \) denotes the (linearized) strain. The boundary conditions correspond to clamping on \( \Gamma_D \) and the prescription of boundary loads \( g \) on the remainder \( \Gamma_N = \Omega \setminus \Gamma_D \).

The boundary loads \( g \) act as control variables. There would be no difficulty in admitting volume forces \( f \) as additional control variables but for practical reasons and simplicity of the presentation, \( f \) is assumed to be zero throughout. The optimal control, or upper-level problem under consideration reads

\[
\begin{aligned}
\text{Minimize} \quad & \psi(u) + \frac{\nu}{2} \|g\|_{H^1(0,T;L^2(\Gamma_N;\mathbb{R}^d))}^2 \\
\text{where} \quad & (\Sigma, u, \lambda) \text{ solves the static plasticity problem (1.1),} \\
\text{and} \quad & g \in U_{ad}. \\
\end{aligned}
\]

(P)

Here, \( \psi \) is a functional favoring certain displacements \( u \). A typical example would be to track the final deformation after unloading. The set \( U_{ad} \) realizes constraints on the control \( g \). The assumptions on \( \psi \) and \( U_{ad} \) are given in Assumption 2.7, followed by a number of examples.

As usual, we refer to the system (1.1) and equivalent reformulations as the forward or lower-level problem and to (P) as the control or upper-level problem.

1.2. 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 \( \Gamma_N \) and \( \Gamma_D \). 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 \( S := \mathbb{R}^{d \times d}_{\text{sym}} \) the space of symmetric \( d \)-by-\( d \) matrices, endowed with the inner product \( A : B = \sum_{i,j=1}^d A_{ij}B_{ij} \), and we define

\[
\begin{aligned}
V &= H_0^1(\Omega; \mathbb{R}^d) = \{ u \in H^1(\Omega; \mathbb{R}^d) : u = 0 \text{ on } \Gamma_D \}, \\
S &= L^2(\Omega; S) \\
\end{aligned}
\]
as the spaces for the displacement $u$, stress $\sigma$, and back stress $\chi$, respectively. The control $g$ belongs to the space
\[ U = L^2(\Gamma_N; \mathbb{R}^d). \]
The embedding $E : U \rightarrow V'$, $g \mapsto \ell$, which maps boundary forces $g \in U$ to functionals $\ell \in V'$ is given by
\[ \langle v, Eg \rangle_{V',V} := -\int_{\Gamma_N} v \cdot g \, ds \quad \text{for all } v \in V. \quad (1.2) \]
Hence, $E = -\tau_N$, where $\tau_N$ 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
\[ L^p(0,T;X) = \{ u : [0,T] \rightarrow X, \text{ } u \text{ is Bochner measurable and } p\text{-integrable} \}. \]
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
\[ \|u\|_{L^p(0,T;X)} = \|\|u(\cdot)\|_X\|_{L^p(0,T)}. \]
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
\[ \left( \|u\|_{L^p(0,T;X)}^p + \|\dot{u}\|_{L^p(0,T;X)}^p \right)^{1/p} \text{ and } \left( \|u(0)\|_X^p + \|\dot{u}\|_{L^p(0,T;X)}^p \right)^{1/p}, \quad (1.3) \]
where the extension to the case $p = \infty$ is clear. 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
\[ \phi(\Sigma) = \left( |\sigma^D + \chi^D|^2 - \sigma^2_0 \right) / 2 \quad (1.4) \]
for $\Sigma = (\sigma, \chi) \in S^2$, where $|\cdot|$ denotes the pointwise Frobenius norm of matrices and
\[ \sigma^D = \sigma - \frac{1}{d} (\text{trace } \sigma) I \quad (1.5) \]
is the deviatoric part of $\sigma$. The yield function gives rise to the set of admissible generalized stresses
\[ \mathcal{K} = \{ \Sigma \in S^2 : \phi(\Sigma) \leq 0 \quad \text{a.e. in } \Omega \}. \]
Let us mention that the structure of the yield function $\phi$ given in (1.4) implies the shift invariance
\[ \Sigma \in \mathcal{K} \quad \Leftrightarrow \quad \Sigma + (\tau, -\tau) \in \mathcal{K} \quad \text{for all } \tau \in S. \quad (1.6) \]
These properties are exploited quite often in the analysis.

Due to the structure of the yield function $\phi$, $\sigma^D + \chi^D$ appears frequently and we abbreviate it and its adjoint by
\[ D\Sigma = \sigma^D + \chi^D \quad \text{and} \quad D^* \sigma = \begin{pmatrix} \sigma^D \\ \sigma_D \end{pmatrix} \quad (1.7) \]
for matrices $\Sigma \in S^2$ as well as for functions $\Sigma \in S^2$ and $\Sigma \in L^p(0,T;S^2)$. When considered as an operator in function space, $\mathcal{D}$ maps $S^2$ and $L^p(0,T;S^2)$ continuously into themselves. For later reference, we also remark that

$$D^* D \Sigma = \left( \sigma^D + \chi^D \right) \text{ and } (D^* D)^2 = 2 D^* D$$

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

$$\Sigma \in \mathcal{K} \Rightarrow D\Sigma \in L^\infty(\Omega;S). \quad (1.8)$$

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

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

$$\langle T, A\Sigma \rangle_{S^2} = \int_\Omega \tau : C^{-1} \sigma \, dx + \int_\Omega \mu : H^{-1} \chi \, dx. \quad (1.9)$$

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

$$\langle B \Sigma, v \rangle_{V',V} = -\int_\Omega \sigma : \epsilon(v) \, dx. \quad (1.10)$$

We recall that $\epsilon(v) = (\nabla v + (\nabla v)^\top)/2$ 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 $\Omega$, denoted by $\Gamma$, consists of two disjoint measurable parts $\Gamma_N$ and $\Gamma_D$ such that $\Gamma = \Gamma_N \cup \Gamma_D$. While $\Gamma_N$ is a relatively open subset, $\Gamma_D$ is a relatively closed subset of $\Gamma$. Furthermore $\Gamma_D$ 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 $\tilde{\sigma}_0$ is assumed to be a positive constant. It equals $\sqrt{2/3} \sigma_0$, where $\sigma_0$ is the initial uni-axial yield stress.

3. $C^{-1}$ and $H^{-1}$ are elements of $L^\infty(\Omega; L(S,S))$, where $L(S,S)$ denotes the space of linear operators $S \to S$. Both $C^{-1}(x)$ and $H^{-1}(x)$ are assumed to be uniformly coercive. Moreover, we assume that $C^{-1}$ and $H^{-1}$ are symmetric, i.e., $\tau : C^{-1}(x) \sigma = \sigma : C^{-1}(x) \tau$ and a similar relation for $H^{-1}$ holds for all $\sigma, \tau \in S$.

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.,
\[ \|u\|_{H^1(\Omega; \mathbb{R}^3)}^2 \leq c_K \left( \|u\|_{L^2(\Omega; \mathbb{R}^3)}^2 + \|\varepsilon(u)\|_S^2 \right) \quad (1.11) \]
for all \( u \in H^1(\Omega; \mathbb{R}^3) \), see e.g. [Herzog et al., 2011a, Lemma C.1]. Note that (1.11) entails in particular that \( \|\varepsilon(u)\|_S \) is a norm on \( H^1_0(\Omega; \mathbb{R}^d) \) equivalent to the standard \( H^1(\Omega; \mathbb{R}^d) \) norm. A further consequence is that \( B^* \) satisfies the inf-sup condition
\[ \|u\|_V \leq \sqrt{c_K} \|B^* u\|_{S^2} \quad \text{for all } u \in V. \quad (1.12) \]

Assumption 1.1 (3) is satisfied, e.g., for isotropic and homogeneous materials, for which
\[ C^{-1} \sigma = \frac{1}{2} \mu \sigma - \frac{\lambda}{2(\mu + d\lambda)} \text{trace}(\sigma) I \]
with Lamé constants \( \mu \) and \( \lambda \), 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 \( \lambda \). A common example for the hardening modulus is given by \( \mathbb{H}^{-1} \chi = \chi/k_1 \) with hardening constant \( k_1 > 0 \), see [Han and Reddy, 1999, Section 3.4].

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

1.3. Preliminaries and known results. As announced above, 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 parabolic VI. For \( \ell \in V' \), we denote by
\[ K_{\ell} = \{ \Sigma \in K : B\Sigma = \ell \} \quad (1.13) \]
the subset of \( K \) on which the constraint \( B\Sigma = \ell \) is fulfilled. Testing the first equation of (VI) with \( T \in K_{\ell(t)} \) results in: given \( \ell \in H^1_0(0, T; V') \) with \( \ell(0) = 0 \), find \( \Sigma \in H^1_0(0, T; V') \), satisfying \( \Sigma(0) = 0 \), \( \Sigma(t) \in K_{\ell(t)} \) and
\[ (A\Sigma(t), T - \Sigma(t))_{S^2} \geq 0 \quad \text{for all } T \in K_{\ell(t)} \text{ and almost all } t \in [0, T]. \quad (1.14) \]
This is called the stress problem of plasticity. It is a time-dependent VI, where the associated convex set \( K_{\ell(t)} \) 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 (1.14). In [Han and Reddy, 1999, Theorem 8.9] the existence and uniqueness of a solution \( \Sigma \) of (1.14) is shown together with the a-priori bound
\[ \|\Sigma\|_{H^1(0, T; S^2)} \leq C \|\ell\|_{H^1(0, T; V')}. \quad (1.15) \]
Let us denote the solution map by \( G^\Sigma \), i.e., \( \Sigma = G^\Sigma(\ell) \). Additionally, [Han and Reddy, 1999, Theorem 8.10] shows the local Hölder continuity of index 1/2 of \( G^\Sigma : H^1_0(0, T; V') \to L^\infty(0, T; S^2) \), i.e.,
\[ \|G^\Sigma(\ell_1) - G^\Sigma(\ell_2)\|^2_{L^\infty(0, T; S^2)} \leq C \left( \|\ell_1\|_{L^2(0, T; V')} + \|\ell_2\|_{L^2(0, T; V')} \right) \|\ell_1 - \ell_2\|_{L^2(0, T; V')} \quad (1.16) \]
We mention that this is not a typical estimate. As long as the derivatives of \( \ell_i \) remain bounded in \( L^2(0, T; V') \), one can control the \( L^\infty \)-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 $u$. As a consequence, $(\Sigma, u)$ satisfies (VI). Thus, given $\ell \in H^1(0,T;V')$ satisfying $\ell(0) = 0$, there exists $(\Sigma, u) \in H^1(0,T;S^2 \times V)$ such that $\Sigma(t) \in \mathcal{K}$, $(\Sigma(0), u(0)) = (0,0)$ and

\[ \langle A\dot{\Sigma}(t) + B^*u(t), T - \Sigma(t) \rangle_{S^2} \geq 0 \quad \text{for all } T \in \mathcal{K}, \]  \[ B\Sigma(t) = \ell(t) \quad \text{in } V', \]  \[ 0 \leq \lambda \perp \phi(\Sigma) \leq 0 \quad \text{a.e. in } [0,T] \times \Omega. \]  

holds for almost all $t \in (0,T)$. This shows the equivalence of (1.14) and (VI). Note that (1.17) is equivalent to (VI). In the sequel we use either reference as appropriate.

**Kinematic hardening.** In the case of kinematic hardening, the uniqueness of $u$ is obtained easily. Let us test (1.17a) with $T = \Sigma(t) + (\tau, -\tau) = (\sigma(t) + \chi(t) - \tau)$. Due to the shift invariance (1.6), we have $T \in \mathcal{K}$ for all $\tau \in S$. This yields

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

Using the definitions of $A$ and $B$, see (1.9) and (1.10), respectively, we obtain

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

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

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

(1.18)

Together with the inf-sup condition of $B^* = (-\varepsilon, 0)$, see (1.11), this proves the uniqueness of $u$. Using (1.15) yields the a-priori estimate

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

(1.19)

We denote the solution mapping $\ell \mapsto u$ of (1.17) by $\mathcal{G}^u$. Moreover, we use $\mathcal{G} = (\mathcal{G}^\Sigma; \mathcal{G}^u)$.

**Primal problem.** In [Han and Reddy, 1999, Section 7] the primal formulation of (1.14) 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

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

(1.20)

This Lipschitz estimate was apparently first proven by Gröger [1978a].

**Existence, uniqueness and regularity of the plastic multiplier.** We already seen that the generalized plastic strain $P = -A\Sigma - B^*u$ 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

\[ A\dot{\Sigma} + B^*u + \lambda D^*D\Sigma = 0 \quad \text{in } L^2(0,T;S^2), \]  \[ B\Sigma = \ell \quad \text{in } L^2(0,T;V'), \]  \[ 0 \leq \lambda \perp \phi(\Sigma) \leq 0 \quad \text{a.e. in } [0,T] \times \Omega. \]  

(1.21)
Endowed with the initial condition \((\Sigma(0), u(0)) = (0, 0)\), this system is equivalent to (1.14) and (1.17).

In the introduction of the second part of the paper we give references showing that a complementarity formulation of the forward problem is better suited for the derivation of first order necessary conditions than the formulation as a time-dependent VI. Thus, the reformulation of (1.17) as (1.21) is essential for the analysis presented in the second part.

Let us briefly comment that (1.1) is the strong formulation of (1.21). Testing the strong formulation of the equilibrium of forces
\[ \text{div} \sigma = 0 \quad \text{in} \quad \Omega \quad \text{and} \quad \sigma \cdot n = g \quad \text{on} \quad \Gamma_N \]
with \(v \in V\) and integrating by parts, we obtain
\[ \langle \Sigma, B^*v \rangle_{S^2} = \langle \ell, v \rangle_{V',V} \quad \text{for all} \quad v \in V, \]
or equivalently \(B\Sigma = \ell\) in \(V'\). Hence, we obtained (1.21b). The reformulation of the remaining equations is straightforward.

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 solution map. Due to this weak continuity, we are able to show the existence of an optimal control in Section 2.4, see Theorem 2.8.

2.1. Introduction to evolution variational inequalities. In this section we give a short introduction to evolution variational inequalities (EVIs). A comprehensive presentation of this topic can be found in Krejčí [1998]. We also mention the contributions Krejčí and Lovíček [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
\[ (\dot{u}(t) - \ddot{x}(t), x(t) - \ddot{x})_X \geq 0 \quad \text{for all} \quad \ddot{x} \in Z. \quad \text{(EVI)} \]
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 (1.17) 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 \(\varepsilon(u)\) (or, equivalently, the strain rate \(\varepsilon(\dot{u}) = -B^*\dot{u}\)) is viewed as a given quantity. The VI (1.17a) reduces to find a function \(\Sigma : [0,T] \rightarrow \mathcal{K} \subset S^2\), such that
\[ \left( \Sigma(t) - \left[ \frac{C \varepsilon(\dot{u}(t))}{0} \right], T - \Sigma(t) \right)_A \geq 0 \quad \text{for all} \quad T \in \mathcal{K}, \quad \text{(2.1)} \]
where \( \langle \cdot, \cdot \rangle_A \) is the scalar product on \( S^2 \) induced by \( A \). This is of the form (EVI) in the space \( X = S^2 \) equipped with \( \langle \cdot, \cdot \rangle_A \) and \( Z = K \).

**Remark 2.1.** We remark that there is a fundamental difference between the dual formulation (1.17) and the strain driven problem (2.1). Whereas in (1.17) both \( \Sigma \) and \( u \) are unknown quantities, \( \varepsilon(u) \) has to be a-priori known in (2.1). Moreover, the solution \( \Sigma \) of (2.1) has to fulfill (1.17b).

Let us denote the solution map of (2.1) which maps \( u \rightarrow \Sigma \) by \( \mathcal{P} \). Therefore, (1.17b) yields \( \ell = B\Sigma = B\mathcal{P}(u) \). Thus, (1.17) is equivalent to

Find \( u \) such that \( B\mathcal{P}(u) = \ell \).

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 \( \rho \ddot{u} \) is included in the balance of forces (1.17b).

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 (1.16). 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 \( C(0,T;X) \) to \( CBV(0,T;X) \), 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 \( x_0^1, x_0^2 \in Z \) be given. Let \( \xi_i = P(x_0^i, u_i), i = 1,2 \). Then the estimate

\[
\|\xi_1 - \xi_2\|_{L^\infty(0,T;X)} \leq 2 \left( \|\hat{u}_1\|_{L^q(0,T;X)} + \|\hat{u}_2\|_{L^q(0,T;X)} \right) \|u_1 - u_2\|_{L^p(0,T;X)} + \|\xi_1(0) - \xi_2(0)\|_X^2
\]

holds, where \( \xi_i(0) = u_i(0) - x_0^i, i = 1,2 \).

**Proof.** First we recall [Krejčí, 1998, (3.16ii)], i.e., \( \left( \dot{x}_1(t), \dot{x}_1(t) \right) \) \( x = 0 \) a.e. in \( (0,T) \), where \( x_1 = S(x_0^1, u) \). 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 \( \ddot{x} = u_2(t) - \xi_2(t) = x_2(t) \in Z \) as a test function in (EVI) shows

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

Similarly, we obtain

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

Adding these inequalities, we have

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

Integrating from 0 to t yields
\[ \frac{1}{2} \| \xi_1(t) - \xi_2(t) \|_X^2 - \frac{1}{2} \| \xi_1(0) - \xi_2(0) \|_X^2 \]
\[ \leq \int_0^t (\| \dot{u}_1(s) \|_X + \| \dot{u}_2(s) \|_X) \| u_1(s) - u_2(s) \|_X \, ds \]
\[ \leq (\| \dot{u}_1 \|_{L^2(0,T,X)} + \| \dot{u}_2 \|_{L^2(0,T,X)}) \| u_1 - u_2 \|_{L^2(0,T,X)}. \]

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 (1.17) which fits in the framework of Krejčí [1998]. This yields the continuity of \( \mathcal{G} : H^1(0, T; V') \to H^1(0, T; S^2 \times V) \). As mentioned in the introduction, this is used in several places.

In order to reformulate (1.17), we need some preparatory work. Due to the inf-sup condition (1.12) we obtain that for all \( \ell \in V' \) there is a \( \sigma_\ell \in S \), such that
\[ (\sigma_\ell, 0) \in (\ker B)^\perp \quad \text{and} \quad B(\sigma_\ell, 0) = \ell, \tag{2.2} \]
see Brezzi [1974]. Moreover, this mapping \( \ell \mapsto \sigma_\ell \) is linear and continuous. For \( \ell \in H^1(0, T; V') \), we define \( \sigma_\ell(t) := \sigma_{\ell(t)} \). Hence \( \sigma(\cdot) \) maps \( H^1(0, T; V') \to H^1(0, T; S) \) continuously. For convenience, we also introduce the notation
\[ \Sigma_\ell = (\sigma_\ell, -\sigma_\ell). \tag{2.3} \]

Let us mention that this is a useful tool to construct test functions \( T \) for (1.14) and (1.17). Indeed, for arbitrary \( \Sigma \in \mathcal{K} \) and \( \ell \in V' \), let us define the test function
\[ T = \Sigma + \Sigma_\ell - B \Sigma. \]  
The shift invariance (1.6) implies \( T \in \mathcal{K} \) and \( BT = \ell \) is ensured by the definition of \( \Sigma_\ell \). Hence we obtain \( T \in \mathcal{K}_\ell \), i.e., it is an admissible test function for (1.14) and (1.17).

Using (2.3) we are able to reformulate (1.14) as an EVI. As stated in (1.6), \( \Sigma \in \mathcal{K} \) 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 \( BS = \ell \) 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 \( T \in \mathcal{K}_B \) be arbitrary. Testing (1.17a) or (1.14) with \( T + \Sigma_\ell(t) \in \mathcal{K} \) yields an equivalent reformulation: find \( \Sigma_0 \in \mathcal{K}_B \) such that
\[ \langle A(\Sigma_0 + \Sigma_\ell), T - \Sigma_0 \rangle \geq 0 \quad \text{for all } T \in \mathcal{K}_B. \tag{2.4} \]

Using the coercivity of \( A \), which is ensured by Assumption 1.1 (3), we find that
\[ \langle \Sigma, T \rangle_A = \langle A \Sigma_0, T \rangle_{S^2} \tag{2.5} \]
is an equivalent scalar product on \( S^2 \). Hence, (2.4) is an EVI in the Hilbert space \( S^2 \) equipped with \( \langle \cdot, \cdot \rangle_A \). Therefore, Theorem 2.2 yields the continuity of the mapping \( \ell \mapsto \Sigma_0 \) from \( H^1(0, T; S^2) \) to \( H^1(0, T; S^2) \). Since \( \ell \mapsto \Sigma_\ell \) is continuous from \( H^1(0, T; V') \to H^1(0, T; S^2) \), we obtain the continuity of the mapping \( \ell \mapsto \Sigma \) from \( H^1(0, T; V') \) to \( H^1(0, T; S^2) \).
The uniqueness and continuous dependence of $u$ on $\ell$ follows from (1.18) and the inf-sup condition of $B^*$, see (1.12). Moreover, Theorem 2.2 also shows the Lipschitz continuity of $G : \ell \mapsto (\Sigma, u), \ W^{1,1}(0,T; V') \to L^\infty(0,T; S^2 \times V)$, see (1.20).

**Remark 2.4.** 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 $\chi$ and some other part $\eta$, such that the yield function $\phi$ 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 1.3.

### 2.3. Weak continuity of the forward operator.

In this section we show the weak continuity of the control-to-state map $G \circ E$ of the optimal control problem (P). Here, $G$ is the solution map of (1.17) and $E$ is the embedding, see (1.2), from the control space $U = L^2(\Gamma; \mathbb{R}^d)$ into $V'$, which is the space of right-hand-sides of (1.17). The weak continuity of $G \circ E$ from $H^1(0,T; U)$ to $H^1(0,T; S^2 \times V)$ is essential for proving the existence of an optimal control in Theorem 2.8.

A main ingredient to prove the weak continuity is the compactness of the embedding $E$ from $U = L^2(\Gamma; \mathbb{R}^d)$ into $V' = (H^1_0(\Omega; \mathbb{R}^d))'$, see the proof of Lemma 2.5. Due to this compactness, Corollary A.7 implies that $H^1(0,T; U)$ embeds compactly into $L^2(0,T; V')$.

Let us mention that the weak continuity of $G \circ E$ is a non-trivial result: although $E : U \to V'$ is compact, the operator $E : H^1(0,T; U) \to H^1(0,T; V')$ (applied in a pointwise sense) is not compact, see also Corollary A.7. Hence the weak continuity of $G \circ E$ does not simply follow from the compactness of $E$ nor the continuity of $G$.

First, we need a preliminary result. It can be interpreted as an upper semicontinuity result for the left-hand side of (1.17a).

**Lemma 2.5.** Let $\{ (\Sigma_k, u_k) \} \subset H^1(0,T; S^2 \times V)$ be a sequence with $\Sigma_k(0) = 0$ and let $(\Sigma, u) \in H^1(0,T; S^2 \times V)$ be given. Moreover, we assume that there are $g_k, g \in H^1(0,T; U)$ such that $B\Sigma_k = Eg_k$ and $B\Sigma = Eg$ hold. If

$$(\Sigma_k, u_k) \to (\Sigma, u) \quad \text{in} \quad H^1(0,T; S^2 \times V) \quad \text{and} \quad g_k \to g \quad \text{in} \quad H^1(0,T; U),$$

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

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

**Proof.** Due to the weak convergence of $(\Sigma_k, u_k)$ in $H^1(0,T; S^2 \times V)$ we have

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

Since $U$ embeds compactly into $V'$, Corollary A.7 yields that $H^1(0,T; U)$ embeds compactly into $L^2(0,T; V')$ and hence $g_k \to g$ in $H^1(0,T; U)$ implies $Eg_k \to Eg$.
in $L^2(0, T; V')$. This shows
\[
\int_0^T \langle B^\ast \dot{u}_k, \Sigma_k \rangle \, dt = \int_0^T \langle \dot{u}_k, B \Sigma_k \rangle \, dt = \int_0^T \langle \dot{u}, B \Sigma \rangle \, dt = \int_0^T \langle B^\ast \dot{u}, \Sigma \rangle \, dt.
\]

To address the remaining term, we use integration by parts and obtain
\[
\int_0^T \langle A \dot{\Sigma}_k, \Sigma_k \rangle \, dt = \frac{1}{2} \langle (A \Sigma_k(T), \Sigma_k(T)) - (A \Sigma_k(0), \Sigma_k(0)) \rangle.
\]
The functionals $\Sigma \mapsto \langle A \Sigma_k(t), \Sigma_k(t) \rangle$ are continuous (w.r.t. the $H^1(0, T; S^2)$-norm) and convex, hence weakly lower semicontinuous. Due to $\Sigma_k(0) = 0$, this already implies $\langle A \Sigma(0), \Sigma(0) \rangle = 0$. Now the weak lower semicontinuity of $\Sigma \mapsto \langle A \Sigma_k(T), \Sigma_k(T) \rangle$ yields
\[
\liminf_{k \to \infty} \langle A \Sigma_k(T), \Sigma_k(T) \rangle \geq \langle A \Sigma(T), \Sigma(T) \rangle.
\]
Therefore,
\[
\liminf_{k \to \infty} \int_0^T \langle A \dot{\Sigma}_k, \Sigma_k \rangle \, dt \geq \frac{1}{2} \langle A \Sigma(T), \Sigma(T) \rangle = \int_0^T \langle A \dot{\Sigma}, \Sigma \rangle \, dt
\]
holds. By combining the results above, we obtain the assertion
\[
\limsup_{k \to \infty} \int_0^T \langle A \dot{\Sigma}_k + B^\ast \dot{u}_k, T - \Sigma_k \rangle \, dt \leq \int_0^T \langle A \dot{\Sigma} + B^\ast \dot{u}, T - \Sigma \rangle \, dt.
\]

The previous lemma enables us to prove

**Theorem 2.6.** The operator $\mathcal{G} \circ E$ is weakly continuous from $H^1(0, T; U)$ to $H^1(0, T; \mathbb{S} \times \mathbb{S})$.

**Proof.** Let $\{g_k\} \subset H^1(0, T; U)$ be a sequence converging weakly towards $g \in H^1(0, T; U)$. The solution of (1.17) with right hand side $Eg$, $Eg_k$ is denoted by $(\Sigma, \dot{u})$, $(\Sigma_k, u_k)$, respectively, i.e., $(\Sigma_k, u_k) = \mathcal{G}(Eg_k)$ and $(\Sigma, \dot{u}) = \mathcal{G}(Eg)$. Due to the boundedness of $E$ and $\mathcal{G}$, see (1.19), there exists a subsequence (for simplicity denoted by the same symbol) $\{((\Sigma_k, u_k))\}$ which converges weakly towards some $(\Sigma, \dot{u})$. We shall prove $(\Sigma, \dot{u}) = (\Sigma, \dot{u})$, therefore the whole sequence converges weakly and this simplification of notation is justified. To this end, we show that $(\Sigma, \dot{u})$ is a solution to (1.17) with right hand side $Eg$.

First we address the admissibility of $\dot{\Sigma}$ in (1.17): the set $\{T \in H^1(0, T; \mathbb{S}) : T(t) \in \mathcal{K} \text{ f.a.a. } t \in (0, T)\}$ is convex and closed, hence weakly closed. Since all $\Sigma_k$ belong to this set, $\dot{\Sigma}$ belongs to it as well. Therefore $\dot{\Sigma}(t) \in \mathcal{K}$ holds for almost all $t \in (0, T)$.

Since $B$ is linear and bounded, it is weakly continuous and hence $B \Sigma_k \rightharpoonup B \dot{\Sigma}$ in $H^1(0, T; V')$. Due to $B \Sigma_k = Eg_k \rightharpoonup Eg$ we have $B \dot{\Sigma} = Eg$ in $H^1(0, T; V')$, this shows (1.17b).

By (1.17a) we have for all $T \in L^2(0, T; \mathbb{S})$ with $T \in \mathcal{K}$ f.a.a. $t \in (0, T)$
\[
\int_0^T \langle A \dot{\Sigma}_k + B^\ast \dot{u}_k, T - \Sigma_k \rangle \, dt \geq 0.
\]
Now we can apply Lemma 2.5. This yields
\[
\int_0^T \langle A \dot{\Sigma} + B^\ast \dot{u}, T - \Sigma \rangle \, dt \geq \limsup_{k \to \infty} \int_0^T \langle A \dot{\Sigma}_k + B^\ast \dot{u}_k, T - \Sigma_k \rangle \, dt \geq 0.
\]
Since \( T \in L^2(0,T;S^2) \) with \( T(t) \in K \) f.a.a. \( t \in (0,T) \) was arbitrary, we obtain the pointwise variational inequality
\[
\langle A \dot{\Sigma} + B^* \dot{\bar{u}}, T - \overline{\Sigma} \rangle \geq 0 \quad \text{a.e. in } (0,T)
\]
for all \( T \in \mathcal{K} \). Hence \((\overline{\Sigma}, \overline{u})\) solves (1.17). Since the solution is unique we obtain \((\overline{\Sigma}, \overline{u}) = (\Sigma, u)\). This implies that the weak limit is indeed independent of the chosen subsequence. Hence, the whole sequence converges weakly, i.e. \((\overline{\Sigma}_k, \overline{u}_k) \rightharpoonup (\Sigma, u)\). That is, \( \mathcal{G}(Eg) = (\Sigma_k, u_k) \rightharpoonup (\Sigma, u) = \mathcal{G}(Eg) \). Since \( \{g_k\} \) and \( g \) 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 2.8 is the weak continuity of the forward operator proven in Theorem 2.6. For convenience, we repeat the definition of the optimal control problem under consideration

\[
\begin{align*}
\text{Minimize} & \quad F(u,g) = \psi(u) + \frac{\nu}{2} \|g\|_{H^1(0,T;U)}^2 \\
\text{such that} & \quad (\Sigma, u) = \mathcal{G}(Eg) \\
\text{and} & \quad g \in U_{ad}.
\end{align*}
\]

The admissible set \( U_{ad} \) is a convex closed subset of \( H^1(0,T; U) \). 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 \( \psi \) and on the set of admissible controls \( U_{ad} \).

**Assumption 2.7.**

1. The function \( \psi: H^1(0,T;V) \to \mathbb{R} \) is weakly lower semicontinuous, continuous and bounded from below.
2. The cost parameter \( \nu \) is a positive, real number.
3. The admissible set \( U_{ad} \) is nonempty, convex and closed in \( H^1(0,T;U) \). Moreover, \( g(0) = 0 \) holds for all \( g \in U_{ad} \).

Let us give some examples which satisfy these conditions. Since the domain of definition of \( \psi \) is \( H^1(0,T;V) \) we can track the displacements, the strains or even their time derivatives (or point evaluations). As examples for \( \psi \) we mention
\[
\begin{align*}
\psi^1(u) &= \frac{1}{2} \|u - u_d\|_{L^2(0,T;L^2(\Omega;\mathbb{R}^d))}^2, \\
\psi^2(u) &= \frac{1}{2} \|u(T) - u_{T,d}\|_{L^2(\Omega;\mathbb{R}^d)}^2, \\
\psi^3(u) &= \frac{1}{2} \|\varepsilon(u(T)) - \varepsilon_{T,d}\|_{L^2(\Omega;\mathbb{S})}^2.
\end{align*}
\]

Here, \( u_d \in L^2(0,T;L^2(\Omega;\mathbb{R}^d)), u_{T,d} \in L^2(\Omega;\mathbb{R}^d) \) 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
\[
\begin{align*}
U_{ad}^1 &= \{g \in H^1(0,T;U) : g(0) = 0, \|g\|_{L^2(\Gamma_N;\mathbb{R}^d)} \leq \rho \text{ a.e. in } (0,T)\} \\
U_{ad}^2 &= \{g \in H^1(0,T;U) : g(0) = g(T) = 0\}
\end{align*}
\]
satisfy Assumption 2.7 for every \( \rho \geq 0 \).

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

The main result of this section is the existence of an optimal control.

**Theorem 2.8.** There exists a global minimizer of (P).

**Proof.** In virtue of the control-to-state map $G \circ E$, we define the reduced objective $f(g) := F(G^u(Eg), g)$. Due to Assumption 2.7, $f$ is bounded from below, hence $\inf f$ on $U_{ad}$ exists. We define $j := \inf \{f(g), g \in U_{ad}\} \in \mathbb{R}$.

Let $\{g_k\} \subset U_{ad}$ be a minimizing sequence, i.e. $\lim f(g_k) = j$. Since $\psi$ is bounded from below and $\nu > 0$, the sequence $\{g_k\}$ 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 $g$. The closeness and convexity of $U_{ad}$ implies the weak closeness. Therefore, $\{g_k\} \subset U_{ad}$ implies $g \in U_{ad}$.

Since the embedding $E : L^2(0, T; R^d) \hookrightarrow V'$ is compact, $G^u \circ E$ is weakly continuous from $H^1(0, T; U)$ to $H^1(0, T; V)$, see Theorem 2.6. Hence, the weak lower semicontinuity of $\psi$ and $\|\cdot\|^2_{H^1(0, T; U)}$ implies that the reduced objective

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

is weakly lower semicontinuous. Therefore,

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

This shows $F(G^u(Eg), g) = f(g) = j$, hence $g$ is a global minimizer.

## 3 Time discretization

In this section we study a discretization in time of (P). After introducing the time-discrete version of (VI) in Section 3.1, we show the strong and weak convergence of this discretization in Sections 3.2 and 3.3, see Theorems 3.3 and 3.5, respectively. These properties are used in Section 3.4 to show that every local minimizer of (P) can be approximated by local minimizers of the (slightly modified) time-discrete problem ($P^\delta_0$), see Corollary 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.

Moreover, the strong convergence of $(\Sigma^T, u^r)$ in $H^1(0, T; S^2 \times V)$ and of $\lambda^r$ in $L^2(0, T; L^2(\Omega))$, see Theorems 3.3 and 3.4, are essential for passing to the limit in the optimality system in Section 7 in the second part of the paper.

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 $\tau$, i.e. $(\cdot)^\tau$, to indicate variables and operators associated with the discretization in time.

**Notation.** For a time-discrete variable $f^\tau \in X^N$, where $X$ is some Banach space, the components of $f^\tau$ are denoted by $f^\tau_i, i = 1, \ldots, N$. When necessary, we may refer to $f^\tau_0$ with a pre-defined value (interpreted as an initial condition), mostly $f^\tau_0 = 0$. 
If \( f^\tau \in X^N \) is the time-discretization of a variable \( f \in H^1(0,T;X) \), we identify \( f^\tau \) with its linear interpolation, i.e., for \( t \in [(i-1)\tau,i\tau] \) we define
\[
f^\tau(t) = \frac{t-(i-1)\tau}{\tau} f_i^\tau + \frac{i\tau-t}{\tau} f_{i-1}^\tau,
\]
with \( f_0^\tau = 0 \). Therefore, \( X^N \) is identified with an \( N \)-dimensional subspace of \( H^1(0,T;X) \). We make use of this identification for the variables \( \Sigma, u \) and \( g \). For later reference, we remark
\[
\dot{f}^\tau(t) = \frac{1}{\tau} (f_i^\tau - f_{i-1}^\tau) \quad \text{and} \quad f_i^\tau = f^\tau(t) - (t - i \tau) \dot{f}^\tau(t)
\]
for almost all \( t \in [(i-1)\tau,i\tau] \) and \( i = 1, \ldots, N \).

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 \( f^\tau \) with a piecewise constant function in \( L^2(0,T;X) \), i.e., for \( t \in [(i-1)\tau,i\tau] \) we define
\[
f^\tau(t) = f_i^\tau.
\]
This is used for the plastic multiplier \( \lambda \).

### 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 \( (\Sigma_i^\tau, u_i^\tau) \in (S^2 \times V)^N \) such that \( \Sigma_i^\tau \in K \) and
\[
(A(\Sigma_i^\tau - \Sigma_{i-1}^\tau) + B^\tau(u_i^\tau - u_{i-1}^\tau), T - \Sigma_i^\tau) \geq 0 \quad \text{for all } T \in K, \tag{3.4a}
\]
\[
B\Sigma_i^\tau = \ell_i^\tau \quad \text{in } V', \tag{3.4b}
\]
holds for all \( i \in \{1,\ldots,N\} \), where \( (\Sigma_0^\tau, u_0^\tau) = 0 \). We denote the solution operator which maps \( \ell^\tau \to (\Sigma^\tau, u^\tau) \) by \( \mathcal{G}^\tau \). 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 (1.18) for the time-discrete problem, we test (3.4a) with \( T = \Sigma_i^\tau + (\tau, -\tau) \), where \( \tau \in S \) is arbitrary. Using \( (\Sigma_0^\tau, u_0^\tau) = 0 \) we obtain
\[
\zeta^{-1} \sigma^\tau - \varepsilon(u^\tau) - H^{-1} \chi^\tau = 0 \quad \text{a.e. in } (0,T) \times \Omega. \tag{3.5}
\]

In [Han and Reddy, 1999, Lemma 8.8 and (8.37)], we find the a-priori estimate
\[
\|\Sigma^\tau\|_{H^1(0,T;S^2)} \leq C \|\ell^\tau\|_{H^1(0,T;V')}.
\]
Together with the inf-sup condition (1.12) of \( B^\tau \) and (3.5), we obtain
\[
\|\Sigma^\tau\|_{H^1(0,T;S^2)} + \|u^\tau\|_{H^1(0,T;V)} \leq C \|\ell^\tau\|_{H^1(0,T;V')}.
\]
Due to this a-priori estimate, Han and Reddy [1999] are able to prove the weak convergence of \( (\Sigma^\tau, u^\tau) = \mathcal{G}^\tau(\ell^\tau) \) towards the solution \( (\Sigma, u) = \mathcal{G}(\ell) \) of (1.17) in \( H^1(0,T;S^2 \times V) \) if the data \( \ell^\tau \) in (3.4) is chosen as the point evaluation of \( \ell \).

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 \( \ell^\tau \). 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
\[
A(\Sigma^\tau_i - \Sigma^\tau_{i-1}) + B^*(u^\tau_i - u^\tau_{i-1}) + \tau \lambda^\tau_i D^* D \Sigma^\tau_i = 0 \quad \text{in} \quad S^2, \tag{3.8a}
\]
\[
B \Sigma^\tau_i = \ell^\tau_i \quad \text{in} \quad V', \tag{3.8b}
\]
\[
0 \leq \lambda^\tau_i \perp \phi(\Sigma^\tau_i) \leq 0 \quad \text{a.e. in} \ \Omega \tag{3.8c}
\]

is satisfied. Here, we scaled the multiplier \( \lambda^\tau \) appropriately.

Similarly to (3.5) we obtain a representation formula for \( \lambda^\tau \),
\[
- C^{-1}(\sigma^\tau_i + \sigma^\tau_{i-1}) + \varepsilon(u^\tau_i - u^\tau_{i-1}) = -H^{-1}(\chi^\tau_i - \chi^\tau_{i-1}) = \tau \lambda^\tau_i D \Sigma^\tau_i. \tag{3.9}
\]

**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 (2.4), we obtain: find \( \Sigma^\tau_i \in K_{E^\tau_i} \), such that
\[
\langle A(\Sigma^\tau_i - \Sigma^\tau_{i-1}), T - \Sigma^\tau_i \rangle \geq 0 \quad \text{for all} \ T \in K_{E^\tau_i}.
\]
As discussed for the continuous problem (1.14), we can introduce a Lagrange multiplier associated with \( B \Sigma^\tau_i = \ell^\tau_i \) such that \( (\Sigma^\tau, u^\tau) \) satisfy (3.4) holds.

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 \( \ell^\tau \) converges to \( \ell \) strongly (or weakly) in the space \( H^1(0, T; U) \), the corresponding solutions \( (\Sigma^\tau, u^\tau) = G^\tau(\ell^\tau) \) are shown to converge strongly (or weakly) to \( (\Sigma, u) \) in \( H^1(0, T; S^2 \times V) \).

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 also 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 shift these ideas to the case of disturbed inputs and prove the convergence of \( (\Sigma^\tau, u^\tau) \) in \( H^1(0, T; S^2 \times V) \).

Let us consider a sequence of time steps \( \{\tau_k\} \) and a sequence of loads \( \ell^{r_k} \in (V')^N \), with \( N_k \tau_k = T \), such that the linear interpolants \( \ell^{r_k} \in H^1(0, T; V') \) converge in \( H^1(0, T; V') \) towards some \( \ell \in H^1(0, T; V') \). We denote by \( (\Sigma^{r_k}, u^{r_k}) \in (S^2 \times V)^N \) the solutions of the time-discrete problem (3.4a), as well as their piecewise linear interpolants. The aim of this section is to prove \( (\Sigma^{r_k}, u^{r_k}) \to (\Sigma, u) \) in \( H^1(0, T; S^2 \times V) \), where \( (\Sigma, u) \) is the solution to the continuous problem (1.17).

For simplicity of notation, we omit the index \( k \) and we refer to \( \ell^{r_k} \to \ell \) if \( k \to \infty \) simply as \( \ell^{r_k} \to \ell \). At some occasions, we have to extract subsequences of \( \{\tau_k\} \). These are denoted by the same symbol. This will cause no confusion, since we show that the limit point \( (\Sigma, u) \) is unique. Hence we obtain the convergence of the whole sequence \( (\Sigma^\tau, u^\tau) \) in Theorem 3.3.

The proof of Theorem 3.3 relies on the convergence argument Corollary A.6 which is tailored to the analysis of EVIs. Its prerequisite (A.6a) is the convergence of \( \Sigma^\tau \) 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 $\ell^\tau \to \ell$ in $W^{1,1}(0,T;V')$. Then
\[
\|\Sigma - \Sigma^\tau\|_{L^\infty(0,T;S^2)}^2 \leq C \left( \|\ell - \hat{\ell}^\tau\|_{L^1(0,T;V')}^2 + \tau \right). \tag{3.10}
\]
In particular, $\Sigma^\tau \to \Sigma$ in $L^\infty(0,T;S^2)$.

Proof. We have
\[
\frac{1}{2} \frac{d}{dt} \|\Sigma(t) - \Sigma^\tau(t)\|^2_A = \langle A\hat{\Sigma}(t) - A\hat{\Sigma}^\tau(t), \Sigma(t) - \Sigma^\tau(t) \rangle \quad \text{f.a.a. } t \in (0,T). \tag{3.11}
\]
Let us fix some $t \in ((i-1)\tau, i\tau)$, $i \in \{1, \ldots, N\}$. Since $\Sigma^\tau(t)$ is a convex combination of $\Sigma^\tau_i$ and $\Sigma^\tau_{i-1}$, see (3.1), and $\Sigma^\tau_i$, $\Sigma^\tau_{i-1} \in K$, we obtain $\Sigma^\tau(t) \in K$. Hence, $T = \Sigma^\tau(t) + \Sigma_{\ell^{-\tau}(t)} \in K$. Using $BT = \ell(t)$, (1.17a) implies
\[
\langle A\hat{\Sigma}(t), \Sigma^\tau(t) - \Sigma(t) + \Sigma_{\ell^{-\tau}(t)} \rangle \geq 0. \tag{3.12}
\]
Testing (3.4a) with $T = \Sigma(t) + \Sigma_{\ell^{-\tau}(t)}$ yields
\[
\langle A(\Sigma^\tau_i - \Sigma^\tau_{i-1}), \Sigma(t) - \Sigma^\tau_i + \Sigma_{\ell^{-\tau}_{i-1}} \rangle \geq 0.
\]
Using (3.2) we obtain
\[
\langle A\hat{\Sigma}^\tau(t), \Sigma(t) - \Sigma^\tau(t) - (i\tau - t) \hat{\Sigma}^\tau + \Sigma_{\ell^{-\tau}(t)} + (i\tau - t) \hat{\ell}^{-\tau}(t) \rangle \geq 0.
\]
Together with (3.12) this yields
\[
\langle A\hat{\Sigma}(t) - A\hat{\Sigma}^\tau(t), \Sigma(t) - \Sigma^\tau(t) \rangle \leq \langle A\hat{\Sigma}(t) - A\hat{\Sigma}^\tau(t), \Sigma_{\ell^{-\tau}(t)} \rangle \nonumber \]
\[
+ (i\tau - t) \langle A\hat{\Sigma}^\tau(t) - \hat{\Sigma}^\tau(t) + \hat{\ell}^{-\tau}(t), \Sigma_{\ell^{-\tau}(t)} \rangle.
\]
Integrating over $t$ and using (3.11), together with the boundedness of $\hat{\Sigma}^\tau$ and $\Sigma_{\ell^{-\tau}}$ in $L^2(0,T;S^2)$, see (3.6), yields
\[
\frac{1}{2} \|\Sigma(t) - \Sigma^\tau(t)\|_A^2 \leq \int_0^t \langle A\hat{\Sigma}(s) - A\hat{\Sigma}^\tau(s), \Sigma_{\ell^{-\tau}(s)} \rangle \, ds + C\tau.
\]
Integrating by parts and using $\Sigma(0) = \Sigma^\tau(0) = 0$ gives
\[
\frac{1}{2} \|\Sigma(t) - \Sigma^\tau(t)\|_A^2 \leq \langle A\hat{\Sigma}(t) - A\hat{\Sigma}^\tau(t), \Sigma_{\ell^{-\tau}(t)} \rangle \nonumber \]
\[
- \int_0^t \langle A\hat{\Sigma}(s) - A\hat{\Sigma}^\tau(s), \Sigma_{\ell^{-\tau}(s)} \rangle \, ds + C\tau \leq C \left( \|\Sigma(t) - \Sigma^\tau(t)\|_{S^2} \|\ell(t) - \ell^\tau(t)\|_{V'} \right. \nonumber \]
\[
+ \|\Sigma - \Sigma^\tau\|_{L^\infty(0,T;S^2)} \|\ell - \hat{\ell}^\tau\|_{L^1(0,T;V')} + \tau. \nonumber
\]
Taking the supremum over $t \in (0,T)$ on both sides shows
\[
\|\Sigma - \Sigma^\tau\|_{L^\infty(0,T;S^2)}^2 \leq C \left( \|\Sigma - \Sigma^\tau\|_{L^\infty(0,T;S^2)} \|\ell - \ell^\tau\|_{L^\infty(0,T;V')} \nonumber \right. \]
\[
+ \|\Sigma - \Sigma^\tau\|_{L^\infty(0,T;S^2)} \|\ell - \hat{\ell}^\tau\|_{L^1(0,T;V')} + \tau. \nonumber
\]
Finally, Young’s inequality and $W^{1,1}(0,T;V') \hookrightarrow L^\infty(0,T;V')$ yields
\[
\|\Sigma - \Sigma^\tau\|_{L^\infty(0,T;S^2)}^2 \leq C \left( \|\ell - \hat{\ell}^\tau\|_{L^1(0,T;V')}^2 + \tau \right). \nonumber
\]

Let us comment on the case when $\ell^\tau$ is the point evaluation of $\ell$. The estimate (3.10) shows the order of convergence of $\tau^{1/2}$ w.r.t. the $L^\infty$-norm provided that $\ell \in W^{1,1}(0,T;V')$. 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 $\ell' \to \ell$ in $H^1(0,T;V')$. Then $(\Sigma',u') \to (\Sigma,u)$ in the space $H^1(0,T;S^2 \times V)$.

**Proof.** We follow the idea of [Krejčí, 1998, Proof of Theorem 3.6]. We will apply Corollary A.6 with the setting

$$X = S^2_0, \quad u_n = \Sigma_{\ell'} - 2 \Sigma_{\ell''}, \quad u_0 = \Sigma_{\ell'} - 2 \Sigma, \quad p = 2, \quad g_n = \|\Sigma_{\ell'}(\cdot)\|_A, \quad g_0 = \|\Sigma_{\ell'}(\cdot)\|_A.$$  \hspace{1cm} (3.13)

To show the convergence of $\Sigma'$. Here, $S^2_0$ denotes the Hilbert space $S^2$ equipped with the inner product induced by $A$. To apply Corollary A.6, we have to verify the prerequisites (A.6c) and (A.6d).

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

$$\langle A(\Sigma_{\ell'}^i - \Sigma_{\ell''}^i), T - \Sigma_{\ell''}^i \rangle \geq 0 \quad \text{for all } T \in K, \quad BT = \ell_i^\prime,$$

which is the time-discrete analog to (1.14). We choose $T = \Sigma_{\ell''}^i + \Sigma_{\ell''}^i - \ell_{i-1}'$, see (2.3) for the definition of $\Sigma_{\ell'}$. By definition of $\Sigma_{\ell'}$ we have $BT = \ell_i^\prime$ and the shift invariance (1.6) implies $T \in K$. This shows

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

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

$$\langle A(\Sigma_{\ell'}(t), \Sigma_{\ell'}(t) - \Sigma(t) \rangle \geq 0 \quad \text{f.a.a. } t \in (0,T).$$

This shows

$$\|\Sigma_{\ell'}(t) - 2 \Sigma'(t)\|_A \leq \|\Sigma_{\ell'}(t)\|_A \quad \text{f.a.a. } t \in (0,T),$$  \hspace{1cm} (3.14)

where $\|\cdot\|_A$ is the norm on $S^2$ induced by $A$, see (2.5).

To derive a similar formula for the continuous solution, we test (1.14) with $T = \Sigma(t + h) + \Sigma(t) - \ell(t + h)$ for $h > 0$ and obtain

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

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

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

Analogously, we obtain by $h \nearrow 0$

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

Hence

$$\|\Sigma_{\ell'}(t) - 2 \Sigma'(t)\|_A \leq \|\Sigma_{\ell'}(t)\|_A$$  \hspace{1cm} (3.15)

In order to apply Corollary A.6 with the setting (3.13) we check the prerequisites.

(A.6a): Lemma 3.2 ensures $\Sigma' \to \Sigma$ in $L^\infty(0,T;S^2_0)$ and the assumption $\ell' \to \ell$ in $H^1(0,T;V')$ implies $\Sigma_{\ell'} \to \Sigma_{\ell}$ in $L^\infty(0,T;S^2)$.

(A.6b): By the linearity of $\ell \mapsto \Sigma_{\ell}$, the assumption $\ell' \to \ell$ in $H^1(0,T;V')$ implies $\|\Sigma_{\ell'}(\cdot)\|_A \to \|\Sigma_{\ell}(\cdot)\|_A$ in $L^2(0,T;R)$.

(A.6c): This was shown in (3.14).

(A.6d): This was shown in (3.15).
Therefore, Corollary A.6 yields 
\[ \| \Sigma_{\ell^*} - 2 \Sigma^\tau - (\Sigma_{\ell} - 2 \Sigma) \|_{H^1(0,T; S_2^\tau)} \rightarrow 0. \] 
By the assumption \( \ell^* \rightarrow \ell \) in \( H^1(0,T; V') \) we infer 
\[ \| \Sigma^\tau - \Sigma \|_{H^1(0,T; S_2^\tau)} \rightarrow 0. \]

Using \( (1.18), (3.5) \) and the inf-sup condition of \( B \), we obtain \( u^\tau \rightarrow u \) in \( H^1(0,T; V) \).

Using the representation formula \( (3.9) \) for \( \lambda^\tau \) we are able to prove the strong convergence of \( \lambda^\tau \) in \( L^2(0,T; L^2(\Omega)) \) towards \( \lambda \). We remark that this is novel result for the analysis of \( (1.21) \).

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 \( (7.36)-(7.41) \) in the second part.

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

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

Using the notion of piecewise linear interpolations \( \chi^\tau, \Sigma^\tau \) and the piecewise constant interpolation \( \lambda^\tau \), we obtain for \( t \in ((i-1)\tau, i\tau) \), see \( (3.2) \),
\[ \lambda^\tau(t) = \frac{-1}{\tilde{\sigma}_0^2} \left( H^{-1}\chi^\tau(t) : \left[ D\Sigma^\tau(t) - (t - i\tau) \cdot D\Sigma^\tau(t) \right] \right). \]

Using \( \lambda = -H^{-1}\chi : D\Sigma/\tilde{\sigma}_0^2 \), by \( (1.21) \), this implies 
\[ |\lambda^\tau(t) - \lambda(t)| \leq \frac{1}{\tilde{\sigma}_0^2} \left| H^{-1}\chi^\tau(t) : D\Sigma^\tau(t) - H^{-1}\chi(t) : D\Sigma(t) \right| 
+ (t - i\tau) \left| H^{-1}\chi^\tau(t) : D\Sigma^\tau(t) \right| 
\leq \frac{1}{\tilde{\sigma}_0^2} \left| H^{-1}\chi^\tau(t) : D\Sigma^\tau(t) - H^{-1}\chi(t) : D\Sigma(t) \right| 
+ (t - i\tau) \left| H^{-1}\chi^\tau(t) : D\Sigma^\tau(t) \right|. \]

Hence, 
\[ \| \lambda^\tau - \lambda \|_{L^1(0,T; L^1(\Omega))} \leq c \left( \| \chi^\tau \| \| D\Sigma^\tau - D\Sigma \| + \| D\Sigma \| \| \chi^\tau - \chi \| 
+ \tau \| \chi^\tau \| \| D\Sigma^\tau \| \right), \]

where all norms on the right hand side are those of \( L^2(0,T; S) \). Using \( \Sigma^\tau \rightarrow \Sigma \) in \( H^1(0,T; S_2^\tau) \) implies \( \lambda^\tau \rightarrow \lambda \) in \( L^1(0,T; L^1(\Omega)) \).
Step (2): We show that \( \| \lambda_T \|_{L^2(0,T;L^2(\Omega))} \to \| \lambda \|_{L^2(0,T;L^2(\Omega))} \). Using that \( \lambda_T \) is the piecewise constant interpolation of \( (\lambda^T_i)_{i=1}^N \), we obtain
\[
\| \lambda^T \|_{L^2(0,T;L^2(\Omega))}^2 &= \int_0^T \int_\Omega (\lambda^T)^2 \, dx \, dt \\
&= \int_\Omega \int_0^T (\lambda^T)^2 \, dt \, dx \\
&= \int_\Omega \sum_{i=1}^N (\lambda^T_i)^2 \, dx \\
&= \sum_{i=1}^N \int_\Omega (\lambda^T_i)^2 \, dx.
\]
Using \( \lambda^T_i \|D\Sigma^T_i\|^2 = \tilde{\sigma}_0^2 \lambda^T_i \) by (3.8c), we obtain
\[
\| \lambda^T \|_{L^2(0,T;L^2(\Omega))}^2 &= \frac{\tau}{\tilde{\sigma}_0^2} \sum_{i=1}^N \int_\Omega |\lambda^T_i D\Sigma^T_i|^2 \, dx \\
&= \frac{\tau}{\tilde{\sigma}_0^2} \sum_{i=1}^N \int_\Omega \left| -\mathbb{H}^{-1} \chi^T_i - \chi^T_{i-1} - \tau \mathbb{H}^{-1} \chi^T \right|^2 \, dx \\
&= \frac{1}{\tilde{\sigma}_0^2} \| \mathbb{H}^{-1} \chi^T \|_{L^2(0,T;S)}^2.
\]
An analogous calculation shows \( \| \lambda \|_{L^2(0,T;L^2(\Omega))} = \frac{1}{\tilde{\sigma}_0} \| \mathbb{H}^{-1} \chi^T \|_{L^2(0,T;S)} \). Hence, by \( \Sigma^T \to \Sigma \) in \( H^1(0,T;S^2) \), see Theorem 3.3, we infer the convergence of norms \( \| \lambda^T \|_{L^2(0,T;L^2(\Omega))} \to \| \lambda \|_{L^2(0,T;L^2(\Omega))} \).

Step (3): We show that \( \lambda^T \to \lambda \) in \( L^2(0,T;L^2(\Omega)) \). By Step (2) we know that for every subsequence of \( \tau \), there is a subsequence \( \tau_k \) such that \( \lambda^T_k \to \lambda \) in \( L^2(0,T;L^2(\Omega)) \) for some \( \lambda \). Step (1) implies \( \lambda \to \lambda \). Hence, the whole sequence converges weakly. In view of the convergence of norms, \( \lambda^T \) converges strongly to \( \lambda \) in \( L^2(0,T;L^2(\Omega)) \).

3.3. Weak convergence. In this section we show the weak convergence of the time-discrete states \( (\Sigma^T, u^T) = G^T(Eg^T) \) in \( H^1(0,T;S^2 \times V) \), under the assumption that the controls \( g^T \) converges weakly towards \( g \) in \( H^1(0,T;U) \). Similarly to the weak continuity of the forward operator proven in Section 2.3, we need the weak convergence of the right hand sides \( Eg^T \) with respect to a stronger norm in space, see the discussion in the beginning of Section 2.3.

Theorem 3.5. Let \( g^T \in U^n \) such that the linear interpolants converge weakly, i.e., there is \( g \in H^1(0,T;U) \) such that \( g^T \to g \) in \( H^1(0,T;U) \).

Then the solutions \( (\Sigma^T, u^T) \) of the time-discrete problem (3.4) with right hand side \( \ell^T = Eg^T \) converge weakly in \( H^1(0,T;S^2 \times V) \) towards the solution \( (\Sigma, u) \) of (1.17) with right hand side \( \ell = Eg \).

Proof. Due to the a-priori bound (3.7), there exists a weakly convergent subsequence of \( (\Sigma^T, u^T) \), denoted by the same symbol, i.e. \( (\Sigma^T, u^T) \to (\Sigma, \tilde{u}) \). As in the proof of Theorem 2.6, we shall show \( \tilde{u} = \tilde{u} \). Hence the whole sequence converges weakly and this simplification of notation is justified.
Let \( T \in L^2(0,T; S^2) \) with \( T \in K \) a.e. in \((0,T)\) be arbitrary. Let \( i \in \{1, \ldots, N\} \) and \( \kappa \in (0,1) \) be given. Set \( t = (i + \kappa) \tau \). Testing (3.4a) with \( T(t) \) yields

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

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

\[
0 \leq \int_{(i+1)\tau}^{(i+1)\tau} \langle A\dot{\Sigma}^\tau(t) + B^*\dot{u}^\tau(t), \ T(t) - \Sigma^\tau((i + 1) \tau) \rangle \, dt
\]

\[
= \int_{i\tau}^{(i+1)\tau} \langle A\dot{\Sigma}^\tau + B^*\dot{u}^\tau, \ T - \Sigma^\tau \rangle \, dt - \frac{\tau}{2} \int_{i\tau}^{(i+1)\tau} \langle A\dot{\Sigma}^\tau + B^*\dot{u}^\tau, \ \dot{\Sigma}^\tau \rangle \, dt.
\]

Hence, summing over \( i = 0, \ldots, N-1 \) yields

\[
0 \leq \int_0^T \langle A\dot{\Sigma}^\tau + B^*\dot{u}^\tau, \ T - \Sigma^\tau \rangle \, dt - \frac{\tau}{2} \int_0^T \langle A\dot{\Sigma}^\tau + B^*\dot{u}^\tau, \ \dot{\Sigma}^\tau \rangle \, dt.
\]

Due to the boundedness of \((\Sigma^\tau, u^\tau)\) in \(H^1(0,T; S^2 \times V)\), the second addend goes to zero as \( \tau \downarrow 0 \). For the first addend we can proceed as in the proof of Theorem 2.6: the application of Lemma 2.5 yields

\[
\limsup_{\tau \downarrow 0} \int_0^T \langle A\dot{\Sigma}^\tau + B^*\dot{u}^\tau, \ T - \Sigma^\tau \rangle \, dt \leq \int_0^T \langle \dot{\Sigma}^\tau + B^*\dot{u}^\tau, \ T - \Sigma^\tau \rangle \, dt.
\]

Together with (3.16) this implies

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

Therefore, \((\tilde{\Sigma}, \tilde{u}) = (\Sigma, u)\).

### 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 \(L^2(\Gamma_N; \mathbb{R}^d)^N\), which is identified with a subspace of \(H^1(0,T; U)\) via the piecewise linear interpolation (3.1). The discretization of the admissible set is given by \(U^\tau_{ad} = U_{ad} \cap L^2(\Gamma_N; \mathbb{R}^d)^N\), where \(\tau = T/N\). Now, the time-discrete optimal control problem reads

\[
\begin{align*}
\text{Minimize} \quad & F(u^\tau, g^\tau) = \psi(u^\tau) + \frac{\nu}{2} \|g^\tau\|^2_{H^1(0,T; U)} \\
\text{such that} \quad & (\Sigma^\tau, u^\tau) = G^\tau(Eg^\tau) \quad \text{(P*)}
\end{align*}
\]

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

**Assumption 3.6.** In addition to Assumption 2.7 we suppose that for all \(g \in U_{ad}\), there exists \(g^\tau \in U^\tau_{ad}\), such that \(g^\tau \to g\) in \(H^1(0,T; U)\) as \(\tau \downarrow 0\).

We remark that this condition is satisfied for both examples of \(U_ad\) given after Assumption 2.7.

The aim of this section is to answer the question whether optimal controls of (P) can be approximated by optimal controls of (P*). 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 \((P)\) can be approximated by global solutions of \((P^\tau)\), see Theorem 3.8.
• every strict local optimum of \((P)\) can be approximated by local solutions of \((P^\tau)\), see Theorem 3.9.
• every local optimum of \((P)\) can be approximated by local solutions of a perturbed problem \((P^\tau_0)\), see Corollary 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 \(u^\tau\) and \(g^\tau\) are optimization variables in \((P^\tau)\), for simplicity we refer solely to \(g^\tau\) being (locally or globally) optimal. This is justified in view of the continuous solution map \(G^\tau\) of (3.4).

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

**Lemma 3.7.** There exists a global minimizer of \((P^\tau)\).

**Proof.** In Section 5.1 of the second part we will show in a more general framework that the solution operator \(G^\tau\) of (3.4) is Lipschitz continuous from \((V^1)^N\) to \((S^2 \times V)^N\), see Proposition 5.8. Since the embedding \(E\) from \(U^N\) into \((V^1)^N\) is compact, \(G^\tau\) is compact from \(U^N\) to \((S^2 \times V)^N\). Now we proceed similarly as in the proof of Theorem 2.8.

Now, we are going to prove the approximation properties.

**Theorem 3.8.** Suppose that Assumption 3.6 is fulfilled. Let \(\{\tau\}\) be a sequence tending to 0 and let \(g^\tau\) denote a global solution to \((P^\tau)\).

1. Then there exists an accumulation point \(g\) of \(\{g^\tau\}\) and
2. every weak accumulation point of \(\{g^\tau\}\) is in fact a strong accumulation point and a global solution of \((P)\).

**Proof.** By Assumption 2.7, we can choose \(g_0 \in U_{ad}\). By Assumption 3.6, there exists a sequence \(g^0_\tau \in U^\tau_{ad}\) which converges towards \(g_0\) in \(H^1(0,T;U)\). Hence, by (3.7) the corresponding displacements \(u^0_\tau\) converge towards \(u_0\) in \(H^1(0,T;V)\). Together with the continuity of \(\psi\), this implies the convergence of \(F(u^0_\tau, g_0)\). Since \(g^\tau\) is a global optimum of \((P^\tau)\), we have \(F(u^\tau, g^\tau) \leq F(u^0_\tau, g^\tau)\). Hence, \(\{g^\tau\}\) is bounded in \(H^1(0,T;U)\). This implies the existence of a weakly convergent subsequence. Therefore assertion (1) follows by assertion (2).

To prove assertion (2), let \(\{g^\tau\}\) converge weakly towards \(g\). We denote by \((\Sigma^\tau, u^\tau) = G^\tau(g^\tau)\) the (time-discrete) solution to (3.4) and by \((\Sigma, u)\) the solution to (1.17). Due to the weak convergence proven in Theorem 3.5, we have \(u^\tau \rightharpoonup u\) in \(H^1(0,T;V)\).

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

\[
F(u, g) \leq \liminf F(u^\tau, g^\tau) \quad \text{by weak lower semicontinuity of } F
\]

\[
\leq \liminf F(\bar{u}^\tau, \bar{g}^\tau) \quad \text{by global optimality of } (u^\tau, g^\tau)
\]

\[
= F(u, \bar{g}) \quad \text{by convergence of } (u^\tau, g^\tau)
\]

This shows that \(g\) is a global optimal solution. Inserting \(\bar{g} = g\) yields the convergence of norms and hence the strong convergence of \(g^\tau\) in \(H^1(0,T;U)\).
Theorem 3.9. Suppose Assumption 3.6 is fulfilled. Let $g$ be a strict local optimum of $(P)$ w.r.t. the topology of $H^1(0,T;U)$. Then, for every sequence $\{g^\tau\}$ tending to 0, there is a sequence $\{g^\tau\}$ of local solutions to $(P^\tau)$, such that $g^\tau \to g$ strongly in $H^1(0,T;U)$.

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

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

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

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

$$\tilde{g}^\tau = (1 - \eta^\tau) g^\tau + \eta^\tau \tilde{g}^\tau,$$

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

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

Step (3): We define the auxiliary problem

Minimize $F(u,g) = \psi(u) + \frac{\nu}{2} \|g\|^2_{H^1(0,T;U)}$

such that $(\Sigma, u) = \mathcal{G}(Eg)$

and $g \in U_{ad} \cap B_\varepsilon(g).$

Since $g$ is the unique minimum in $\hat{U}_{ad} \cap B_\varepsilon(g)$, it is the unique global minimizer of $(P_{g,\varepsilon})$. Invoking Theorem 3.8 yields the existence of a sequence $g^\tau$ of global solutions of the associated time-discrete problem $(P^\tau_{\varepsilon})$, such that $g^\tau \to g$ in $H^1(0,T;U)$.

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

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

Minimize $F_g(u,\bar{g}) = \psi(u) + \frac{\nu}{2} \|ar{g}\|^2_{H^1(0,T;U)} + \frac{1}{2} \|\bar{g} - g\|^2_{H^1(0,T;U)}$

such that $(\Sigma, u) = \mathcal{G}(E\bar{g})$

and $\bar{g} \in U_{ad}$

Clearly, $g$ becomes a strict local optimum of $(P_g)$. Analogously to $(P^\tau)$ we define the time-discrete approximation $(P^\tau_{g}$).
Corollary 3.10. Suppose Assumption 3.6 is fulfilled. Let $g$ 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 $(g^n)$ of local solutions to $(P_g^n)$, such that $g^n \to g$ strongly in $H^1(0,T;U)$.

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

Due to this corollary, we are able to derive necessary optimality conditions for (P) by passing to the limit with the optimality conditions of $(P_g^n)$.

A Tools from functional analysis about Bochner spaces

This section is devoted to some convergence properties in Bochner spaces. We cite two theorems given in Krejčí [1998] and state Corollary A.6 which is be used to infer the convergence in $H^1(0,T;X)$. These results are tailored to the analysis of EVIs.

Lemma A.1. 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 $x \in X$

\[ f(x_n) \to f(x) \quad \text{as} \quad n \to \infty \quad (A.1) \]

holds for all $f \in V$, then (A.1) holds for all $f \in \cl V$, where the closure is taken with respect to the $X'$-norm.

Proof. Let $f \in \cl V$ be given. Then there exists $\{f_n\} \subset V$ with $f_n \to f$. 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 an $n \in \mathbb{N}$ such that $\|f_m - f\|_{X'} \leq \varepsilon/(3C)$. Since (A.1) holds for $f_m$, there exists $n \in \mathbb{N}$ such that $\|f_m(x_n) - f_m(x)\| \leq \varepsilon/3$. This implies

\[ |f(x_n) - f(x)| \leq |f(x_n) - f_m(x_n)| + |f_m(x_n) - f_m(x)| + |f_m(x) - f(x)| \]

\[ \leq \|f - f_m\|_{X'} \|x_n\|_X + |f_m(x_n) - f_m(x)| + \|f - f_m\|_{X'} \|x\|_X \leq \varepsilon. \]

This proves that (A.1) holds for $f$.

Lemma A.2. 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 $u_n \to u$ in $L^\infty(0,T;X)$ and if $\{u_n\}$ is bounded in $W^{1,1}(0,T;X)$, then

\[ \int_0^T \langle \dot{u}_n(t), \varphi(t) \rangle \, dt \to \int_0^T \langle \dot{u}(t), \varphi(t) \rangle \, dt \quad \text{for all} \quad \varphi \in C(0,T;X), \quad (A.2) \]

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.2) 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
Let \( u_n \) be given sequences for \( n \in \mathbb{N} \cup \{0\} \) such that

\[
\lim_{n \to \infty} \int_0^T (\dot{u}_n(t), \varphi(t)) \, dt = 0 \quad \text{for all } \varphi \in C(0, T; X),
\]

\[
g_n \to g \quad \text{in } L^1(0, T; \mathbb{R}),
\]

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

\[
\|v_0(t)\|_X = g_0(t) \quad \text{a.e. in } [0, T].
\]

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

\[\text{Theorem A.3 ([Krejčí, 1998, Theorem 8.7])}\]

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

\[
v_n(t) \to v_0(t) \quad \text{a.e. in } [0, T],
\]

\[
g_n \to g \quad \text{in } L^p(0, T; \mathbb{R}),
\]

\[
\|v_n(t)\|_B \leq g_n(t) \quad \text{a.e. in } [0, T], \text{ for all } n \in \mathbb{N},
\]

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

\[\text{Theorem A.4 ([Krejčí, 1998, Theorem 8.5])}\]

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

\[
\lim_{n \to \infty} \int_0^T (v_n(t) - v_0(t), \varphi(t)) \, dt = 0 \quad \text{for all } \varphi \in C(0, T; X),
\]

\[
g_n \to g \quad \text{in } L^p(0, T; \mathbb{R}),
\]

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

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

\[\text{Corollary A.5}\]

Then \( v_n \to v \) in \( L^1(0, T; X) \). Thus there exists a pointwise convergent subsequence \( \{v_{n_k}\} \). Hence, \textbf{Theorem A.4} yields \( v_{n_k} \to v \) in
Corollary A.6. 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

\begin{align*}
    u_n &\to u_0 \quad \text{in } L^\infty(0,T; X) \quad \text{(A.6a)} \\
    g_n &\to g_0 \quad \text{in } L^p(0,T; \mathbb{R}), \quad \text{(A.6b)} \\
    \|\dot{u}_n(t)\|_X &\leq g_n(t) \quad \text{a.e. in } [0,T], \quad \text{for all } n \in \mathbb{N}, \quad \text{(A.6c)} \\
    \|\dot{u}_0(t)\|_X &= g_0(t) \quad \text{a.e. in } [0,T]. \quad \text{(A.6d)}
\end{align*}

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

Proof. By (A.6b) and (A.6c) we obtain the boundedness of $u_n$ in $W^{1,1}(0,T; X)$. Lemma A.2 yields (A.2) and hence we can apply Corollary A.5 for $v_n = \dot{u}_n$. This yields $\dot{u}_n \to \dot{u}_0$ in $L^p(0,T; X)$ and hence $u_n \to u_0$ in $W^{1,p}(0,T; X)$.

In [Simon, 1986, (6.5)] we find the compact embedding of $W^{1,1}(0,T; Y)$ into $L^p(0,T; X)$ if $Y$ is compactly embedded into $X$. This is also known as Aubin’s Lemma.

Corollary A.7. Let $Y, X$ be Banach spaces, such that $Y$ is compactly embedded into $X$. Then, $W^{1,1}(0,T; Y)$ is compactly embedded into $L^p(0,T; X)$ for all $p \in [1, \infty)$.

Moreover, for all $p \in (1, \infty]$, $W^{1,p}(0,T; Y)$ is compactly embedded into $C(0,T; X)$.

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


