Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Optimal control of static plasticity with linear kinematic hardening

Roland Griesse[1], Christian Meyer[2]

submitted: November 5, 2008

[1] Chemnitz University of Technology
Faculty of Mathematics
09107 Chemnitz
Germany
E-Mail: roland.griesse@mathematik.tu-chemnitz.de

[2] Weierstrass Institute
for Applied Analysis
and Stochastics
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: meyer@wias-berlin.de

No. 1370
Berlin 2008

2000 *Mathematics Subject Classification.* 49J40, 49M30, 35J85, 35Q72.

Key words and phrases. Optimal control of variational inequalities, static plasticity, Yosida approximation.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract. An optimal control problem for the static problem of infinitesimal elastoplasticity with linear kinematic hardening is considered. The variational inequality arising on the lower-level is regularized using a Yosida-type approach, and an optimal control problem for the so-called viscoplastic model is obtained. Existence of a global optimizer is proved for both the regularized and original problems, and strong convergence of the solutions is established.

1 Introduction

This paper is concerned with an optimal control problem for the static model of infinitesimal elastoplasticity with linear kinematic hardening. Static, or incremental, plast icity models arise through discretization in time of quasi-static plasticity problems. For a detailed physical motivation, we refer to Simo and Hughes [1998] and Han and Reddy [1999]. The *forward* problem is characterized by the unique solution of

$$\left.\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\int_\Omega \boldsymbol{\sigma} : \mathbb{C}^{-1} : \boldsymbol{\sigma}\, dx + \frac{1}{2}\int_\Omega \boldsymbol{\chi} : \mathbb{H}^{-1} : \boldsymbol{\chi}\, dx \\ \text{s.t.} \quad & \begin{cases} -\operatorname{div} \boldsymbol{\sigma} = \boldsymbol{f} & \text{in } \Omega \\ \boldsymbol{\sigma} \cdot \boldsymbol{n} = \boldsymbol{g} & \text{on } \Gamma_N. \end{cases} \\ \text{and} \quad & (\boldsymbol{\sigma}(x), \boldsymbol{\chi}(x)) \in K \quad \text{a.e. in } \Omega. \end{aligned}\right\} \tag{1.1}$$

Hence, on the one hand, the optimal control of (1.1) leads to a bi-level optimization problem. On the other hand, (1.1) can be replaced by its necessary and sufficient optimality conditions, and thus we obtain an optimal control problem for a variational inequality.

We work under the assumption of infinitesimal strains. Hence, Ω is the domain occupied by the body in both the undeformed and deformed states. The volume and boundary loads $\boldsymbol{f}$ and $\boldsymbol{g}$ serve as control variables, and $\boldsymbol{\sigma}(x) \in \mathbb{R}^{d\times d}_{\text{sym}}$ denotes the stress tensor of the body resulting from these loads. The fourth order tensors $\mathbb{C}$ and $\mathbb{H}$ are the elasticity tensor and the hardening modulus, respectively. Conditions motivated by physical considerations ensure that (1.1) is uniquely solvable.

The closed and convex set K imposes bounds on the generalized stress $(\boldsymbol{\sigma}, \boldsymbol{\chi})$ the body can take. The variable $\boldsymbol{\chi}$ denotes an internal force which arises during hardening. It is termed the back stress in the case of kinematic hardening, and it causes a translation of the initial yield surface, compare (2.2). Plastic deformation is characterized by material points satisfying $(\boldsymbol{\sigma}, \boldsymbol{\chi}) \in \partial K$. When deleting the variable $\boldsymbol{\chi}$, one obtains the static problem of perfect plasticity (Hencky model), which provides substantially less regular solutions, see Temam [1983]. If, in addition, the condition $(\boldsymbol{\sigma}, \boldsymbol{\chi}) \in K$ is neglected, (1.1) reduces to the problem of static linear elasticity.

We consider the following optimal control problem:

$$\left.\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|\boldsymbol{u} - \boldsymbol{u}_d\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_1}{2}\|\boldsymbol{f}\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_2}{2}\|\boldsymbol{g}\|^2_{L^2(\Gamma_N;\mathbb{R}^d)} \\ \text{s.t.} \quad & \text{the static plasticity problem (1.1).} \end{aligned}\right\} \tag{1.2}$$

The lower-level problem (1.1) can be equivalently replaced by the variational formulation of its necessary and sufficient optimality conditions

$$-\operatorname{div} \boldsymbol{\sigma} = \boldsymbol{f} \quad \text{in } \Omega, \qquad \boldsymbol{\sigma} \cdot \boldsymbol{n} = \boldsymbol{g} \quad \text{on } \Gamma_N, \tag{1.3a}$$

$$\begin{aligned} &\boldsymbol{\varepsilon}(\boldsymbol{u}(x)) : \boldsymbol{\sigma}(x) - \boldsymbol{\sigma}(x) : \mathbb{C}^{-1} : \boldsymbol{\sigma}(x) - \boldsymbol{\chi}(x) : \mathbb{H}^{-1} : \boldsymbol{\chi}(x) \\ &\quad = \max_{(\boldsymbol{\tau},\boldsymbol{\mu})\in K} \{\boldsymbol{\varepsilon}(\boldsymbol{u}(x)) : \boldsymbol{\tau} - \boldsymbol{\sigma}(x) : \mathbb{C}^{-1} : \boldsymbol{\tau} - \boldsymbol{\chi}(x) : \mathbb{H}^{-1} : \boldsymbol{\mu}\} \quad \text{a.e. in } \Omega. \end{aligned} \tag{1.3b}$$

where $\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}\big(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\top\big)$ denotes the strain tensor. Note that (1.3) is equivalent to a mixed variational inequality of the first kind, see (2.11) below. The variable $\boldsymbol{u}$ is the Lagrange multiplier associated to the equality constraints in (1.1) and it can be physically interpreted as the displacement field by means of duality techniques, see Appendix A. We emphasize that the occurence of the Lagrange multiplier for the lower-level problem in the upper-level objective is a particular feature of the problem at hand.

In the present paper, we prove the existence of a global optimizer of (1.2), and also of a family of regularized problems. The regularization consists in replacing the constraint $(\boldsymbol{\sigma}, \boldsymbol{\chi}) \in K$ by a penalty term based on the Yosida approximation of the indicator function. Remarkably, the Yosida approximation leads to a lower-level problem which allows a physical interpretation in its own right, the so-called viscoplastic approximation of (1.1). We also prove that *every* strict local optimum of the original problem (1.2) is the strong limit of local solutions of the viscoplastic optimal control problems, as the regularization parameter tends to ∞. The paper can thus be viewed as a preparatory step for the derivation of first-order necessary optimality conditions for (1.2) in the spirit of Ito and Kunisch [2000] and Hintermüller [2008], which will be the subject of a subsequent paper.

Let us put our work into perspective. As was noted above, the weak formulation of (1.3) is equivalent to a mixed variational inequality of the first kind. The bi-level optimization problem (1.2) thus represents an optimal control problem governed by an elliptic variational inequality. This class of problems has been adressed by many authors under different aspects. We only mention Mignot [1976], Barbu [1984], Mignot and Puel [1984], Bonnans and Tiba [1991], Bonnans and Casas [1995], Bergounioux [1998], Bergounioux and Zidani [1999], Ito and Kunisch [2000], Hintermüller [2008], and the references therein. In these contributions, various techniques were used to establish first-order conditions for optimal control of elliptic variational inequalities of both, first and second kind. To the best of our knowledge, the optimal control of mixed variational problems has not been addressed, let alone problems in the context of elastoplasticity. As another distinguishing feature, we note that the Lagrange multiplier associated to the equality constraint in (1.1) appears in the objective of the upper-level problem (1.2). Thus, the discussion of (1.2) offers a genuine contribution to the theory of optimal control for variational inequalities.

For the analysis, we follow the classical approach of Barbu [1984], who employs a two-fold regularization to the lower-level problem, consisting of a Yosida approximation of the indicator function of the admissible set and the subsequent convolution with a smoothing kernel. As was already noted, the Yosida approximation leads to a lower-level problem which allows a physical interpretation in its own right, the so-called viscoplastic approximation of (1.1). This provides another motivation to analyze optimal controls for the viscoplastic model.

The paper is organized as follows. Notations, assumptions, and the weak formulation associated to (1.3) are collected in Section 2. Section 3 starts with a discussion of the lower-level problem (1.1) by Lagrange techniques. We emphasize here that the same results could also be obtained using Fenchel duality, cf. for instance [Temam, 1983, Chapter III]. The existence of solutions to (1.2) then follows from standard arguments. Section 4 is devoted to the analysis of the viscoplastic approximation of the lower-level and bi-level problems (1.1) and (1.2), respectively. Using these results, the strong convergence of solutions as the regularization parameter tends to ∞ is established in Section 5.

We remark that the lower-level problem (1.1) is called the dual, or stress-based formulation. It is well known that an equivalent primal formulation exists, which justifies the existence of the Lagrange multiplier $\boldsymbol{u}$ and its interpretation as the displacement field. The primal form is traditionally derived by means of Fenchel duality, cf. for instance [Temam, 1983, Chapter III]. By employing Lagrange techniques in Sections 3 and 4, we offer an alternative approach. In Appendix A, we also give an alternative form of the primal problem, using Lagrangian duality. For convenience of the reader, some results on orthogonal projections in Hilbert spaces that are used in Section 4 are collected in Appendix B.

2 Notation and Preliminary Results

Our notation follows Han and Reddy [1999]. We begin by recalling some elements of tensor calculus. By $\mathbb{R}^{d\times d}$ we denote the space of real $d \times d$ matrices, and $\mathbb{R}^{d\times d}_{\text{sym}}$ is the subspace of symmetric matrices. Throughout, all tensors will be considered with respect to the standard Cartesian basis. Therefore, second-order tensors can be identified with elements of $\mathbb{R}^{d\times d}$. They will be denoted by bold-face upper-case letters, or by bold-face lower-case Greek letters. The standard scalar product of two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ is denoted by $\boldsymbol{a} \cdot \boldsymbol{b}$. Moreover, the scalar product of two matrices $\boldsymbol{A} = (A_{ij})$ and $\boldsymbol{B} = (B_{ij})$ in $\mathbb{R}^{d\times d}$ is defined by

$$\boldsymbol{A} : \boldsymbol{B} = A_{ij}\, B_{ij},$$

where Einstein's summation convention is used. This scalar product gives rise to the Frobenius norm on $\mathbb{R}^{d\times d}$, denoted by $|\boldsymbol{A}| = (\boldsymbol{A} : \boldsymbol{A})^{1/2}$.

Every tensor $\boldsymbol{A}$ can be uniquely decomposed into its spherical and deviatoric parts $\boldsymbol{A} = \boldsymbol{A}^S + \boldsymbol{A}^D$, where

$$\boldsymbol{A}^S = \frac{1}{d}\,(\text{tr}\boldsymbol{A})\,\boldsymbol{I} = \frac{1}{d} A_{kk}\,(\delta_{ij}), \qquad \boldsymbol{A}^D = \boldsymbol{A} - \frac{1}{d}\,(\text{tr}\boldsymbol{A})\,\boldsymbol{I} = (A_{ij}) - \frac{1}{d} A_{kk}\,(\delta_{ij}).$$

Here δ_{ij} is the Kronecker delta, $\boldsymbol{I} = (\delta_{ij})$ is the unit tensor, and $\text{tr}(\boldsymbol{A}) = A_{kk}$ is the trace of $\boldsymbol{A}$.

A real tensor of fourth order is identified with an element of $\mathbb{R}^{d\times d\times d\times d}$ and it is denoted by $\mathbb{A} = (A_{ijkl})$. We define the products $\mathbb{A} : \mathbb{B} = (A_{ijkl}\, B_{klmn})$ and $\mathbb{A} : \boldsymbol{B} = (A_{ijkl}\, B_{kl})$.

Definition 2.1. *We say that a fourth-order tensor* $\mathbb{A}$ *is*

(a) symmetric *if it has the following symmetry properties:*

$$A_{ijkl} = A_{jikl} = A_{ijlk} = A_{jilk},$$

which imply that $\mathbb{A} : \boldsymbol{B}$ *is a symmetric second-order tensor whenever* $\boldsymbol{B}$ *is;*

(b) coercive *if*

$$\boldsymbol{B} : \mathbb{A} : \boldsymbol{B} \geq \underline{c} \, |\boldsymbol{B}|^2 \quad \text{for all } \boldsymbol{B} \in \mathbb{R}^{d \times d} \tag{2.1}$$

holds for some constant $\underline{c} > 0$.

Remark 2.2. (a) *If* $\mathbb{A}$ *is symmetric, then it is sufficient that* (2.1) *holds for all symmetric matrices* $\boldsymbol{B}$ *in order for* $\mathbb{A}$ *to be coercive.*

(b) *If* $\mathbb{A}$ *is coercive, then it is invertible in the following sense: there exists a unique fourth-order tensor* $\mathbb{A}^{-1}$ *such that*

$$\mathbb{A}^{-1} : \mathbb{A} : \boldsymbol{B} = \mathbb{A} : \mathbb{A}^{-1} : \boldsymbol{B} = \boldsymbol{B} \quad \text{for all } \boldsymbol{B} \in \mathbb{R}^{d \times d}.$$

$\mathbb{A}^{-1}$ *is coercive as well. If, in addition,* $\mathbb{A}$ *is symmetric, then* $\mathbb{A}^{-1}$ *is symmetric, too.*

Now we turn to the functional analytic setting.

Assumption 2.3. *Let* $\Omega \subset \mathbb{R}^d$ *be a bounded domain with Lipschitz boundary* Γ *in dimension* $d \in \{2, 3\}$. *The boundary consists of two disjoint parts* Γ_N *and* Γ_D, *where* Γ_D *is a relatively closed set in* Γ *of positive measure, and no connected component of* Γ_D *consists of isolated points.*

Assumption 2.4. (a) *The components of the elasticity tensor* $\mathbb{C}$ *in* (1.1) *are assumed to satisfy* $C_{ijkl} \in L^\infty(\Omega)$. *For almost all* $x \in \Omega$, *we assume that* $\mathbb{C}(x)$ *is coercive according to Definition 2.1, with a constant* $\underline{c} > 0$ *independent of* x. *(By Remark 2.2, the so-called compliance tensor* $\mathbb{C}^{-1}$ *exists almost everywhere with components in* $L^\infty(\Omega)$, *and it satisfies a coercivity condition with a constant* $\underline{c}' > 0$ *independent of* x.)
The same is assumed for the hardening modulus $\mathbb{H}$.

(b) *In addition, we assume that* $\mathbb{C}(x)$ *is symmetric in the sense of Definition 2.1 (a). (This implies that* $\mathbb{C}^{-1}(x)$ *is symmetric as well.)*

(c) *Without loss of generality, we infer from the objective in* (1.1) *that* $\mathbb{C}^{-1}$ *satisfies* $(C^{-1})_{ijkl} = (C^{-1})_{klij}$, *i.e.,* $\boldsymbol{\sigma} : \mathbb{C}^{-1} : \boldsymbol{\tau} = \boldsymbol{\tau} : \mathbb{C}^{-1} : \boldsymbol{\sigma}$ *holds for all* $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{R}^{d \times d}$. *The same is true for* $\mathbb{H}^{-1}$.

In homogeneous isotropic materials, $\mathbb{C}$ is given by

$$C_{ijkl} = \lambda \, \delta_{ij} \, \delta_{kl} + \mu \, (\delta_{ik} \, \delta_{jl} + \delta_{il} \, \delta_{jk})$$

where μ and λ are the Lamé constants. When $\mu > 0$ and $d \, \lambda + 2 \, \mu > 0$ hold, then $\mathbb{C}$ satisfies Assumption 2.4. A common example for the hardening modulus is given by $\mathbb{H} = \operatorname{diag}(k_1)$ with hardening constant $k_1 > 0$, see [Han and Reddy, 1999, Section 3.4].

Assumption 2.5 (Set of admissible generalized stresses)**.**

(a) *The set* $K \subset \mathbb{R}^{d \times d}_{\text{sym}} \times \mathbb{R}^{d \times d}$ *is assumed to be nonempty, closed and convex with* $(\boldsymbol{0}, \boldsymbol{0}) \in K$.

(b) *For all* $\boldsymbol{A} \in \mathbb{R}^{d \times d}_{\text{sym}}$, *we assume that*

$$(\boldsymbol{\sigma}, \boldsymbol{\chi}) \in K \quad \Rightarrow \quad (\boldsymbol{\sigma} + \boldsymbol{A}, \boldsymbol{\chi} - \boldsymbol{A}) \in K.$$

Remark 2.6 (Safe load condition)**.** *Kinematic hardening is characterized by a translation of the initial yield surface during plastic loading, see [Han and Reddy, 1999, p. 69]. In other words, whether or not a generalized stress state* $(\boldsymbol{\sigma}, \boldsymbol{\chi})$ *belongs to the admissible set* K *depends only on* $\boldsymbol{\sigma}+\boldsymbol{\chi}$*. Assumption 2.5 (b) is thus natural for problems with kinematic hardening. Note that is equivalent to* $K + (\boldsymbol{A}, -\boldsymbol{A}) = K$ *for all* $\boldsymbol{A} \in \mathbb{R}^{d\times d}_{\text{sym}}$*.*

Moreover, Assumption 2.5 (b) can be interpreted as a particular form of the safe load condition. In fact, Assumption 2.5 implies that an admissible generalized stress exists for arbitrary loads, see Proposition 3.1.

Example 2.7. *Assumption 2.5 is satisfied, for instance, by the von Mises yield condition in case of linear kinematic hardening, i.e.,*

$$K = \{(\boldsymbol{\sigma}, \boldsymbol{\chi}) \in \mathbb{R}^{d\times d} : |\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D| \leq \sqrt{2/3}\,\sigma_0\}, \tag{2.2}$$

where σ_0 *is the initial uni-axial yield stress, compare [Han and Reddy, 1999, p.69, p.182].*

Definition 2.8. *We define*

$$\begin{aligned} V &= H^1_D(\Omega; \mathbb{R}^d) = \{\boldsymbol{u} \in H^1(\Omega; \mathbb{R}^d) : \boldsymbol{u} = 0 \text{ on } \Gamma_D\}, \\ S &= L^2(\Omega; \mathbb{R}^{d\times d}_{\text{sym}}), \qquad M = L^2(\Omega; \mathbb{R}^{d\times d}). \end{aligned}$$

as spaces for the displacement $\boldsymbol{u}$*, stress* $\boldsymbol{\sigma}$*, and back stress* $\boldsymbol{\chi}$*, respectively.*

Now we are in the position to define the following bilinear forms associated to the static plasticity problem.

Definition 2.9. *For* $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S \times M$ *and* $\boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in S \times M$*, define*

$$a(\boldsymbol{\Sigma}, \boldsymbol{T}) = \int_\Omega \boldsymbol{\sigma} : \mathbb{C}^{-1} : \boldsymbol{\tau}\, dx + \int_\Omega \boldsymbol{\chi} : \mathbb{H}^{-1} : \boldsymbol{\mu}\, dx. \tag{2.3}$$

For $\boldsymbol{\sigma} \in S$ *and* $\boldsymbol{v} \in V$*, let*

$$b(\boldsymbol{\sigma}, \boldsymbol{v}) = -\int_\Omega \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\boldsymbol{v})\, dx. \tag{2.4}$$

We recall that $\boldsymbol{\varepsilon}(\boldsymbol{v}) = \frac{1}{2}\big(\nabla \boldsymbol{v} + (\nabla \boldsymbol{v})^\top\big)$ *denotes the strain tensor. We also define an operator* $B : S \to V'$ *associated to* b*, by*

$$\langle B\boldsymbol{\sigma}, \boldsymbol{v}\rangle = b(\boldsymbol{v}, \boldsymbol{\sigma}).$$

Here and in the following, $\langle \cdot\,, \cdot\rangle$ *denotes the dual pairing between* V' *and* V*, and* V' *is the dual space of* V *w.r.t. the topology of* $L^2(\Omega; \mathbb{R}^d)$*.*

Note that the objective in (1.1) can be expressed as $\frac{1}{2}a(\boldsymbol{\Sigma}, \boldsymbol{\Sigma})$. As a consequence of Assumption 2.4, a is coercive on $S \times M$, i.e., there exists $\underline{\alpha} > 0$ such that

$$a(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}) \geq \underline{\alpha}\, \|\boldsymbol{\Sigma}\|^2_{S\times M} \tag{2.5}$$

holds for all $\boldsymbol{\Sigma} \in S \times M$. Moreover, a is bounded on $S \times M$, i.e., there exists $\overline{\alpha} > 0$ such that

$$|a(\boldsymbol{\Sigma}, \boldsymbol{T})| \leq \overline{\alpha}\, \|\boldsymbol{\Sigma}\|_{S\times M} \|\boldsymbol{T}\|_{S\times M} \tag{2.6}$$

holds for all $\boldsymbol{\Sigma}, \boldsymbol{T} \in S \times M$. The bilinear form b is bounded on $S \times V$,

$$b(\boldsymbol{\sigma}, \boldsymbol{v}) \leq \overline{\beta}\, \|\boldsymbol{\sigma}\|_S \|\boldsymbol{v}\|_V \tag{2.7}$$

and it satisfies the condition of Babuška-Brezzi, i.e., there exists $\underline{\beta} > 0$ such that

$$\sup_{\boldsymbol{\sigma} \in S \setminus \{0\}} \frac{b(\boldsymbol{\sigma}, \boldsymbol{v})}{\|\boldsymbol{\sigma}\|_S} \geq \underline{\beta} \, \|\boldsymbol{v}\|_V \quad \text{for all } \boldsymbol{v} \in V. \tag{2.8}$$

This follows from Korn's inequality, see e.g., [Temam, 1983, Proposition 1.1 and Remark 1.1],

$$\|\boldsymbol{u}\|^2_{H^1(\Omega;\mathbb{R}^d)} \leq c_K \left(\|\boldsymbol{u}\|^2_{L^2(\Gamma_D;\mathbb{R}^d)} + \|\boldsymbol{\varepsilon}(\boldsymbol{u})\|^2_{L^2(\Omega;\mathbb{R}^{d\times d})} \right) \tag{2.9}$$

for all $\boldsymbol{u} \in H^1(\Omega;\mathbb{R}^d)$. Note that (2.9) also implies that $\|\boldsymbol{\varepsilon}(\boldsymbol{u})\|_{L^2(\Omega;\mathbb{R}^{d\times d})}$ is a norm on $H^1_D(\Omega;\mathbb{R}^d)$ equivalent to the natural norm.

As a consequence of (2.8), the following lemma holds, see Brezzi [1974]:

Lemma 2.10. *For any given $\ell \in V'$, the equation*

$$b(\boldsymbol{\sigma}, \boldsymbol{v}) = \langle \ell, \boldsymbol{v} \rangle \quad \textit{for all } \boldsymbol{v} \in V$$

has a unique solution $\boldsymbol{\sigma} \in (\ker B)^\perp$ and the estimate

$$\|\boldsymbol{\sigma}\|_S \leq c_B \, \|\ell\|_{V'}. \tag{2.10}$$

holds with a constant c_B independent of ℓ.

Definition 2.11 (Weak solution)**.** *Let $\boldsymbol{f} \in L^2(\Omega;\mathbb{R}^d)$ and $\boldsymbol{g} \in L^2(\Gamma_N;\mathbb{R}^d)$ be given. A triple $(\boldsymbol{\Sigma}, \boldsymbol{u}) = (\boldsymbol{\sigma}, \boldsymbol{\chi}, \boldsymbol{u}) \in S \times M \times V$ is called a weak solution of (1.3) if $\boldsymbol{\Sigma} \in \mathcal{K}$ and*

$$a(\boldsymbol{\Sigma}, \boldsymbol{T} - \boldsymbol{\Sigma}) + b(\boldsymbol{\tau} - \boldsymbol{\sigma}, \boldsymbol{u}) \geq 0 \quad \textit{for all } \boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{K} \tag{2.11a}$$

$$b(\boldsymbol{\sigma}, \boldsymbol{v}) = \langle \ell, \boldsymbol{v} \rangle \quad \textit{for all } \boldsymbol{v} \in V, \tag{2.11b}$$

where

$$\mathcal{K} = \{\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S \times M : (\boldsymbol{\sigma}(x), \boldsymbol{\chi}(x)) \in K \ \textit{a.e. in } \Omega\} \tag{2.12}$$

is the set of admissible generalized stresses and

$$\langle \ell, v \rangle = -\int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \, dx - \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v} \, ds. \tag{2.13}$$

The reader will verify that the above variational formulation is obtained from (1.3) by formal integration by parts. Similarly, we reformulate the stress problem (1.1) as

$$\left.\begin{aligned} \text{Minimize} \quad & \frac{1}{2} a(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}) \\ \text{s.t.} \quad & b(\boldsymbol{\sigma}, \boldsymbol{v}) = \langle \ell, \boldsymbol{v} \rangle \quad \text{for all } \boldsymbol{v} \in V \\ \text{and} \quad & \boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in \mathcal{K}. \end{aligned}\right\} \tag{$\mathbf{L}$}$$

Consequently, we consider from now on the following bi-level optimization problem, which is the weak form of (1.2):

$$\left.\begin{aligned} \text{Minimize} \quad & F(\boldsymbol{u}, \boldsymbol{f}, \boldsymbol{g}) := \frac{1}{2} \|\boldsymbol{u} - \boldsymbol{u}_d\|^2_{L^2(\Omega;\mathbb{R}^d)} \\ & \qquad + \frac{\nu_1}{2} \|\boldsymbol{f}\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_2}{2} \|\boldsymbol{g}\|^2_{L^2(\Gamma_N;\mathbb{R}^d)} \\ \text{s.t.} \quad & \text{the plasticity problem } (\mathbf{L}) \text{ with } \ell \text{ as in (2.13).} \end{aligned}\right\} \tag{$\mathbf{P}$}$$

3 Existence of Solutions

The theory of (**P**) is clearly based on the existence and uniqueness results for (**L**). Here, we will take the convex optimization point of view and derive necessary and sufficient optimality conditions for (**L**) by means of a Lagrange multiplier approach. As pointed out in the introduction, it is to be noted that these results are not genuine and can also be obtained by means of Fenchel duality (cf. [Temam, 1983, Chapter III]) or standard arguments for variational inequalities (see for instance Kinderlehrer and Stampacchia [1980]). Nevertheless, it is interesting to see how standard techniques in convex optimization yield that the displacement field can be viewed as a Lagrange multiplier associated to the equality constraints in (**L**), cf. Proposition 3.2.

3.1. Analysis of the Lower-Level Problem. In this section we discuss the existence and uniqueness of solutions for the lower-level problem (**L**).

Proposition 3.1 (Existence and uniqueness). *For every $\ell \in V'$, problem* (**L**) *possesses a unique solution* $(\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S \times M$.

Proof. The proof uses standard arguments. The objective in (**L**) is uniformly convex due to the coercivity of a, and radially unbounded. The admissible set

$$\mathcal{K}_\ell = \{\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in \mathcal{K} : b(\boldsymbol{\sigma}, \boldsymbol{v}) = \langle \ell, \boldsymbol{v} \rangle \quad \text{for all } \boldsymbol{v} \in V\} \tag{3.1}$$

is closed and convex (hence weakly closed) due to Assumption 2.5. From Lemma 2.10, we obtain $\widetilde{\boldsymbol{\sigma}} \in (\ker B)^\perp \subset S$ such that the equality constraint is satisfied. With $\widetilde{\boldsymbol{\Sigma}} = (\widetilde{\boldsymbol{\sigma}}, -\widetilde{\boldsymbol{\sigma}})$, the conditions on K in Assumption 2.5 imply $\widetilde{\boldsymbol{\Sigma}} \in \mathcal{K}$, and thus $\mathcal{K}$ is nonempty. The weak lower semicontinuity of the objective therefore yields the existence of a solution, which is unique due to the uniform convexity. □

Next we address the first-order necessary and sufficient optimality conditions for problem (**L**).

Proposition 3.2 (Optimality conditions, existence of the displacement field). *For given $\ell \in V'$ and $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in \mathcal{K}$, the following are equivalent:*

(*i*) $\boldsymbol{\Sigma}$ *is the unique solution of* (**L**),
(*ii*) *there exists a Lagrange multiplier $\boldsymbol{u} \in V$ such that* (2.11) *holds,*
(*iii*) $\boldsymbol{\Sigma} \in \mathcal{K}_\ell$ *and the variational inequality*

$$a(\boldsymbol{\Sigma}, \boldsymbol{T} - \boldsymbol{\Sigma}) \geq 0 \quad \textit{for all } \boldsymbol{T} \in \mathcal{K}_\ell \tag{3.2}$$

holds, with $\mathcal{K}_\ell$ as defined in (3.1).

Proof. (*i*) ⇔ (*ii*): We apply the generalized Karush-Kuhn-Tucker theory. To this end, we verify the constraint qualification according to Zowe and Kurcyusz. For problem (**L**), this amounts to verifying the surjectivity of B, which follows from Lemma 2.10. We associate to (**L**) the Lagrangian

$$\mathcal{L}(\boldsymbol{\sigma}, \boldsymbol{\chi}, \boldsymbol{u}) = \frac{1}{2} a((\boldsymbol{\sigma}, \boldsymbol{\chi}), (\boldsymbol{\sigma}, \boldsymbol{\chi})) + b(\boldsymbol{\sigma}, \boldsymbol{u}) - \langle \ell, \boldsymbol{u} \rangle.$$

Theorem 4.1 in Zowe and Kurcyusz [1979] implies the existence of a Lagrange multiplier $\boldsymbol{u} \in V$, such that the optimality system

$$\begin{aligned}&\mathcal{L}_{(\boldsymbol{\sigma},\boldsymbol{\chi})}(\boldsymbol{\sigma},\boldsymbol{\chi},\boldsymbol{u})(\boldsymbol{\tau}-\boldsymbol{\sigma},\boldsymbol{\mu}-\boldsymbol{\chi}) \geq 0 \quad \text{for all } (\boldsymbol{\tau},\boldsymbol{\mu}) \in \mathcal{K}\\ &\mathcal{L}_{\boldsymbol{u}}(\boldsymbol{\sigma},\boldsymbol{\chi},\boldsymbol{u}) = 0,\end{aligned}$$

is satisfied, which is the same as (2.11). The sufficiency of (2.11) for optimality of $(\boldsymbol{\sigma},\boldsymbol{\chi})$ is standard for convex problems.

$(i) \Leftrightarrow (iii)$: Since $\mathcal{K}_\ell$ is convex, (3.2) are necessary and sufficient for optimality by standard arguments. □

The above theorem does not imply the uniqueness of the Lagrange multiplier $\boldsymbol{u}$. The uniqueness follows, however, from the following lemma.

Lemma 3.3 (Lipschitz stability)**.** *For any given $\ell_1, \ell_2 \in V'$, the associated solutions $(\boldsymbol{\sigma}_1,\boldsymbol{\chi}_1,\boldsymbol{u}_1)$ and $(\boldsymbol{\sigma}_2,\boldsymbol{\chi}_2,\boldsymbol{u}_2)$ satisfy*

$$\|\boldsymbol{\sigma}_1-\boldsymbol{\sigma}_2\|_S + \|\boldsymbol{\chi}_1-\boldsymbol{\chi}_2\|_M + \|\boldsymbol{u}_1-\boldsymbol{u}_2\|_V \leq L\,\|\ell_1-\ell_2\|_{V'}, \tag{3.3}$$

where L is independent of ℓ_1, ℓ_2.

Proof. Step 1: Estimate for $(\boldsymbol{\sigma},\boldsymbol{\chi})$
The estimate for $(\boldsymbol{\sigma},\boldsymbol{\chi})$ is obtained by choosing appropriate test function $\boldsymbol{T} = (\boldsymbol{\tau},\boldsymbol{\mu})$ in the variational inequality (3.2). Let $\widetilde{\boldsymbol{\sigma}}$ be the unique solution in $(\ker B)^\perp \subset S$ of $b(\widetilde{\boldsymbol{\sigma}},\boldsymbol{v}) = \langle \ell_1-\ell_2,\boldsymbol{v}\rangle$ for all $\boldsymbol{v} \in V$ (see Lemma 2.10), and set $\widetilde{\boldsymbol{\chi}} = -\widetilde{\boldsymbol{\sigma}}$. Then $\boldsymbol{T}_1 = (\boldsymbol{\tau}_1,\boldsymbol{\mu}_1) := (\boldsymbol{\sigma}_2,\boldsymbol{\chi}_2) + (\widetilde{\boldsymbol{\sigma}},\widetilde{\boldsymbol{\chi}}) \in \mathcal{K}_{\ell_1}$ is an admissible test function for (3.2), evaluated at $\boldsymbol{\Sigma}_1 = (\boldsymbol{\sigma}_1,\boldsymbol{\chi}_1)$ since

$$b(\boldsymbol{\tau}_1,\boldsymbol{v}) = b(\boldsymbol{\sigma}_2,\boldsymbol{v}) + b(\widetilde{\boldsymbol{\sigma}},\boldsymbol{v}) = \langle \ell_2,\boldsymbol{v}\rangle + \langle \ell_1-\ell_2,\boldsymbol{v}\rangle = \langle \ell_1,\boldsymbol{v}\rangle$$

holds for all $\boldsymbol{v} \in V$, and

$$(\boldsymbol{\sigma}_2,\boldsymbol{\chi}_2) \in \mathcal{K} \quad\Rightarrow\quad \boldsymbol{T}_1 = (\boldsymbol{\sigma}_2+\widetilde{\boldsymbol{\sigma}},\boldsymbol{\chi}_2+\widetilde{\boldsymbol{\chi}}) = (\boldsymbol{\sigma}_2+\widetilde{\boldsymbol{\sigma}},\boldsymbol{\chi}_2-\widetilde{\boldsymbol{\sigma}}) \in \mathcal{K}$$

by Assumption 2.5. Consequently we obtain from (3.2)

$$a(\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_1) \geq -a(\boldsymbol{\Sigma}_1,(\widetilde{\boldsymbol{\sigma}},\widetilde{\boldsymbol{\chi}})).$$

Similarly, one shows that $\boldsymbol{T}_2 = (\boldsymbol{\tau}_2,\boldsymbol{\mu}_2) := (\boldsymbol{\sigma}_1,\boldsymbol{\chi}_1) - (\widetilde{\boldsymbol{\sigma}},\widetilde{\boldsymbol{\chi}})$ lies in $\mathcal{K}_{\ell_2}$, i.e., it is an admissible test function for (3.2) evaluated at $(\boldsymbol{\sigma}_2,\boldsymbol{\chi}_2)$. This yields

$$a(\boldsymbol{\Sigma}_2,\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2) \geq a(\boldsymbol{\Sigma}_2,(\widetilde{\boldsymbol{\sigma}},\widetilde{\boldsymbol{\chi}})).$$

Adding these inequalities gives

$$a(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2,\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2) \leq a(\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2,(\widetilde{\boldsymbol{\sigma}},\widetilde{\boldsymbol{\chi}}))$$

The left hand side can be estimated by (2.5). By construction and the a priori estimate (2.10), $\|\widetilde{\boldsymbol{\sigma}}\|_S \leq c_B\,\|\ell_1-\ell_2\|_{V'}$ holds. Hence we can estimate

$$\underline{\alpha}\,\|\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2\|^2_{S\times M} \leq c_B\,\overline{\alpha}\,\|\ell_1-\ell_2\|_{V'}\|\boldsymbol{\Sigma}_1-\boldsymbol{\Sigma}_2\|_{S\times M}.$$

Step 2: Estimate for $\boldsymbol{u}$
The estimates for the displacement are obtained by using

$$\begin{aligned}\boldsymbol{T}_1 &= (\boldsymbol{\tau}_1,\boldsymbol{\mu}_1) := (\boldsymbol{\sigma}_2+\boldsymbol{\varepsilon}(\boldsymbol{u}_1-\boldsymbol{u}_2),\boldsymbol{\chi}_2-\boldsymbol{\varepsilon}(\boldsymbol{u}_1-\boldsymbol{u}_2)),\\ \boldsymbol{T}_2 &= (\boldsymbol{\tau}_2,\boldsymbol{\mu}_2) := (\boldsymbol{\sigma}_1+\boldsymbol{\varepsilon}(\boldsymbol{u}_2-\boldsymbol{u}_1),\boldsymbol{\chi}_1-\boldsymbol{\varepsilon}(\boldsymbol{u}_2-\boldsymbol{u}_1))\end{aligned}$$

as test functions in (2.11a). Note that $\boldsymbol{\tau}_1$ is symmetric, and $(\boldsymbol{\tau}_1, \boldsymbol{\mu}_1) \in K$ holds a.e. in Ω in view of $(\boldsymbol{\sigma}_2, \boldsymbol{\chi}_2) \in \mathcal{K}$ and Assumption 2.5, applied with $\boldsymbol{A} = -\boldsymbol{\varepsilon}(\boldsymbol{u}_1 - \boldsymbol{u}_2)$. This shows that $\boldsymbol{T}_1 \in \mathcal{K}$ holds, and a similar argument applies when verifying $\boldsymbol{T}_2 \in \mathcal{K}$. We obtain from (2.11a) the estimates

$$a(\boldsymbol{\Sigma}_1, \boldsymbol{T}_1 - \boldsymbol{\Sigma}_1) + b(\boldsymbol{\tau}_1 - \boldsymbol{\sigma}_1, \boldsymbol{u}_1) \geq 0$$
$$a(\boldsymbol{\Sigma}_2, \boldsymbol{T}_2 - \boldsymbol{\Sigma}_2) + b(\boldsymbol{\tau}_2 - \boldsymbol{\sigma}_2, \boldsymbol{u}_2) \geq 0$$

or equivalently

$$\begin{aligned}
&a(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2 - \boldsymbol{\Sigma}_1) - \int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{u}_1 - \boldsymbol{u}_2) : \boldsymbol{\varepsilon}(\boldsymbol{u}_1)\, dx \\
&\qquad - b(\mathbb{C}^{-1} : \boldsymbol{\sigma}_1, \boldsymbol{u}_1 - \boldsymbol{u}_2) + b(\mathbb{H}^{-1} : \boldsymbol{\chi}_1, \boldsymbol{u}_1 - \boldsymbol{u}_2) + b(\boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1, \boldsymbol{u}_1) \geq 0, \\
&a(\boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) - \int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{u}_2 - \boldsymbol{u}_1) : \boldsymbol{\varepsilon}(\boldsymbol{u}_2)\, dx \\
&\qquad - b(\mathbb{C}^{-1} : \boldsymbol{\sigma}_2, \boldsymbol{u}_2 - \boldsymbol{u}_1) + b(\mathbb{H}^{-1} : \boldsymbol{\chi}_2, \boldsymbol{u}_2 - \boldsymbol{u}_1) + b(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \boldsymbol{u}_2) \geq 0.
\end{aligned}$$

Adding both inequalities yields

$$\begin{aligned}
a(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) &+ \int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{u}_1 - \boldsymbol{u}_2) : \boldsymbol{\varepsilon}(\boldsymbol{u}_1 - \boldsymbol{u}_2)\, dx \\
&\leq b(\mathbb{H}^{-1} : (\boldsymbol{\chi}_1 - \boldsymbol{\chi}_2), \boldsymbol{u}_1 - \boldsymbol{u}_2) \\
&\quad - b(\mathbb{C}^{-1} : (\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2), \boldsymbol{u}_1 - \boldsymbol{u}_2) - b(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \boldsymbol{u}_1 - \boldsymbol{u}_2).
\end{aligned} \tag{3.4}$$

The left hand side can be estimated by (2.5) and (2.9). In view of (2.7) and Assumption 2.4, the right hand side is bounded by a multiple of $\|\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2\| \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|$. We thus obtain

$$\underline{a} \, \|\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2\|^2 + c_K^{-1} \, \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|^2_{H^1(\Omega;\mathbb{R}^d)} \leq C \, \|\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2\| \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|.$$

Young's inequality and the estimate for $\|\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2\|$ from Step 1 yield the desired estimate (3.3). ☐

Remark 3.4. *We point out that the Lagrange mutiplier $\boldsymbol{u}$ can be physically interpreted as displacement field. It solves an optimization problem which is known as primal problem of static elastoplasticity, see for instance [Han and Reddy, 1999, Section 7]. As shown in Appendix A this optimization problem is equivalent to the dual problem of* (**L**).

3.2. Discussion of the Bi-Level Problem (P). Based on the results of Subsection 3.1, we derive the existence of a global optimizer of problem (**P**). As a consequence of Proposition 3.2, we can replace the lower-level problem (**L**) by its necessary and sufficient optimality conditions (2.11). We introduce the space of admissible controls

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

As in (2.13), we associate to given $(\boldsymbol{f}, \boldsymbol{g}) \in U$ a functional $\ell = R(\boldsymbol{f}, \boldsymbol{g})$ through

$$\langle R(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{v} \rangle := -\int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\, dx - \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v}\, ds, \quad \boldsymbol{v} \in V.$$

Lemma 3.5. *The operator $R : U \to V'$ is linear and compact.*

Proof. The embedding $V \hookrightarrow L^2(\Omega; \mathbb{R}^d)$ and the trace operator $V \to L^2(\Gamma_N)$ are compact. The operator R is the (negative) adjoint, and thus it is compact as well. ☐

The results of the previous section give rise the definition of a solution operator for the lower-level problem ($\mathbf{L}$),

$$S : U \ni (\boldsymbol{f}, \boldsymbol{g}) \mapsto (\boldsymbol{\sigma}, \boldsymbol{\chi}, \boldsymbol{u}) \in Y := S \times M \times V.$$

The individual components of S will be denoted by $S^{\boldsymbol{\sigma}}$, $S^{\boldsymbol{\chi}}$ and $S^{\boldsymbol{u}}$. Note that S is nonlinear due to the presence of the constraint $(\boldsymbol{\sigma}, \boldsymbol{\chi}) \in \mathcal{K}$. By Lemma 3.3 and 3.5, S is Lipschitz continuous and compact.

Based on the properties of S, we obtain a global minimizer of ($\mathbf{P}$) as in [Hintermüller, 2001, Theorem 2.2]. Due to the nonlinearity of S, the minimizer can not be expected to be unique.

Proposition 3.6. *Problem* ($\mathbf{P}$) *possesses a global optimal solution* $(\boldsymbol{f}^*, \boldsymbol{g}^*) \in U$.

Proof. Let $j := \inf F(S^{\boldsymbol{u}}(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{f}, \boldsymbol{g})$, where the infimum extends over the space U, and let $\{(\boldsymbol{f}_n, \boldsymbol{g}_n)\}$ be a minimizing sequence. Then $\{(\boldsymbol{f}_n, \boldsymbol{g}_n)\}$ is bounded in U, and hence it possesses a weakly convergent subsequence $(\boldsymbol{f}_{n'}, \boldsymbol{g}_{n'}) \rightharpoonup (\boldsymbol{f}^*, \boldsymbol{g}^*)$ in U. The compactness of S implies that the corresponding solutions $(\boldsymbol{\sigma}_{n'}, \boldsymbol{\chi}_{n'}, \boldsymbol{u}_{n'})$ of the lower-level problem ($\mathbf{L}$) converge to $(\boldsymbol{\sigma}^*, \boldsymbol{\chi}^*, \boldsymbol{u}^*)$ in $S \times M \times V$. The weak lower semicontinuity of the objective implies that $(\boldsymbol{f}^*, \boldsymbol{g}^*)$ is a global optimum of ($\mathbf{P}$). □

4 Viscoplastic Approximation

Before we turn to the viscoplastic approximation of ($\mathbf{L}$), let us state some known results on orthogonal projections in Hilbert spaces that will be useful in the following. The associated proofs are given in Appendix B.

Lemma 4.1 (Differentiability and Shift-Invariance)**.** *Let H be a Hilbert space, $C \subset H$ be a nonempty closed convex set, and denote by $P_C(x)$ the orthogonal projection of x onto C.*

(a) *The function $F(x) = \frac{1}{2}\|x - P_C\, x\|^2$ is convex and Fréchet differentiable with derivative $F'(x) = x - P_C\, x$.*
(b) *The derivative F' is a monotone operator, i.e., $(F'(x) - F'(y), x - y) \geq 0$ holds for all $x, y \in H$.*
(c) *$(F'(x), x - y) \geq 0$ for all $x \in H$ and all $y \in C$.*
(d) *If $a + C = C$ for some element $a \in H$ holds, then $P_C\, x = P_C(x + a) - a$ and $F'(x) = F'(x + a)$ for all $x \in H$.*

Next, we introduce the viscoplastic regularization of ($\mathbf{L}$):

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} a(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}) + \frac{\gamma}{2} \|\boldsymbol{\Sigma} - P_{\mathcal{K}}(\boldsymbol{\Sigma})\|^2_{S \times M} \\ \text{s.t.} \quad & b(\boldsymbol{\sigma}, \boldsymbol{v}) = \langle \ell, \boldsymbol{v} \rangle \quad \text{for all } \boldsymbol{v} \in V, \end{aligned} \tag{$\mathbf{L}_\gamma$}$$

where $\gamma > 0$ is a given real number and $P_{\mathcal{K}}$ denotes the orthogonal projection on $\mathcal{K}$. As pointed out in the introduction, ($\mathbf{L}_\gamma$) has a physical motivation in its own right, see for instance [Simo and Hughes, 1998, Section 2.7].

Remark 4.2. *The viscoplastic problem* $(\mathbf{L}_\gamma)$ *represents a penalized version of* $(\mathbf{L})$ *in the sense that the inequality constraints in* $(\mathbf{L})$ *are replaced by a quadratic penalty term in the objective functional. This is also known as Yosida regularization of the indicator function associated to* $\mathcal{K}$. *This type of regularization is particularly well-suited for the optimal control of variational inequalities, as demonstrated for instance in the classical book Barbu [1984], or more recently in Ito and Kunisch [2000], Hintermüller [2008], where Barbu's approach is modified by means of a feasibility shift. However, due to the non-smoothness of the projection, an additional regularization will be necessary to derive first-order optimality conditions, see e.g. Barbu [1984] or Mignot and Puel [1984]. As optimality conditions will be topic of a subsequent paper, two-fold smoothing is not considered in this work.*

For convenience, we define

$$J_\gamma(\mathbf{\Sigma}) = \frac{\gamma}{2}\|\mathbf{\Sigma} - P_{\mathcal{K}}(\mathbf{\Sigma})\|^2.$$

Since J_γ is convex by Lemma 4.1, we find the following analog to Proposition 3.1:

Proposition 4.3. *For every* $\ell \in V'$ *and every* $\gamma > 0$, *there exists a unique solution* $\mathbf{\Sigma}_\gamma = (\boldsymbol{\sigma}_\gamma, \boldsymbol{\chi}_\gamma) \in S \times M$ *of problem* $(\mathbf{L}_\gamma)$.

On the basis of Lemma 2.10 and [Zowe and Kurcyusz, 1979, Theorem 4.1], one obtains necessary and sufficient optimality conditions for $(\mathbf{L}_\gamma)$, similarly to Theorem 3.2. To this end, let us define

$$\mathcal{C}_\ell^{\mathrm{eq}} = \{\mathbf{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S \times M \,:\, b(\boldsymbol{\sigma}, \boldsymbol{v}) = \langle \ell, \boldsymbol{v}\rangle \quad \text{for all } \boldsymbol{v} \in V\}.$$

Proposition 4.4. *Let* $\ell \in V'$ *and* $\gamma > 0$ *be given. For* $\mathbf{\Sigma}_\gamma \in S \times M$, *the following are equivalent:*

(*i*) $\mathbf{\Sigma}_\gamma$ *is the unique solution of* $(\mathbf{L}_\gamma)$,

(*ii*) *there exists a Lagrange multiplier* $\boldsymbol{u}_\gamma \in V$ *such that the following optimality system is fulfilled:*

$$a(\mathbf{\Sigma}_\gamma, \boldsymbol{T}) + b(\boldsymbol{\tau}, \boldsymbol{u}_\gamma) + (J_\gamma'(\mathbf{\Sigma}_\gamma), \boldsymbol{T}) = 0 \quad \textit{for all } \boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in S \times M \tag{4.1a}$$

$$b(\boldsymbol{\sigma}_\gamma, \boldsymbol{v}) = \langle \ell, \boldsymbol{v}\rangle \quad \textit{for all } \boldsymbol{v} \in V, \tag{4.1b}$$

where $J_\gamma'(\mathbf{\Sigma}) = \gamma\,(\mathbf{\Sigma} - P_{\mathcal{K}}(\mathbf{\Sigma})) \in S \times M$ *is the derivative of* J_γ,

(*iii*) $\mathbf{\Sigma}_\gamma \in \mathcal{C}_\ell^{\mathrm{eq}}$ *satisfies*

$$a(\mathbf{\Sigma}_\gamma, \boldsymbol{T} - \mathbf{\Sigma}_\gamma) + (J_\gamma'(\mathbf{\Sigma}_\gamma), \boldsymbol{T} - \mathbf{\Sigma}_\gamma) \geq 0 \quad \textit{for all } \boldsymbol{T} \in \mathcal{C}_\ell^{\mathrm{eq}}. \tag{4.2}$$

As in case of $(\mathbf{L})$, the uniqueness of $\boldsymbol{u}_\gamma$ follows from the following Lipschitz property of $\ell \mapsto (\mathbf{\Sigma}_\gamma, \boldsymbol{u}_\gamma)$:

Lemma 4.5. *Let* $\ell_1, \ell_2 \in V'$ *and* $\gamma > 0$ *be given. Let* $(\boldsymbol{\sigma}_{\gamma,1}, \boldsymbol{\chi}_{\gamma,1}, \boldsymbol{u}_{\gamma,1})$ *and* $(\boldsymbol{\sigma}_{\gamma,2}, \boldsymbol{\chi}_{\gamma,2}, \boldsymbol{u}_{\gamma,2})$ *denote the solutions of* $(\mathbf{L}_\gamma)$ *associated to* ℓ_1 *and* ℓ_2, *respectively. Then*

$$\|\boldsymbol{\sigma}_{\gamma,1} - \boldsymbol{\sigma}_{\gamma,2}\|_S + \|\boldsymbol{\chi}_{\gamma,1} - \boldsymbol{\chi}_{\gamma,2}\|_M + \|\boldsymbol{u}_{\gamma,1} - \boldsymbol{u}_{\gamma,2}\|_V \leq L\,\|\ell_1 - \ell_2\|_{V'}$$

holds with the same constant L *as in Lemma 3.3. In particular, this yields the uniqueness of the displacement fields.*

Proof. The proof proceeds similarly to the proof of Lemma 3.3 so we can focus here on the arguments which differ.

Step 1: Estimate for $(\boldsymbol{\sigma}, \boldsymbol{\chi})$
Let again $\widetilde{\boldsymbol{\sigma}}$ be the unique solution in $(\ker B)^\perp \subset S$ of $b(\widetilde{\boldsymbol{\sigma}}, \boldsymbol{v}) = \langle \ell_1 - \ell_2, \boldsymbol{v}\rangle$ for all $\boldsymbol{v} \in V$, and set $\widetilde{\boldsymbol{\Sigma}} = (\widetilde{\boldsymbol{\sigma}}, -\widetilde{\boldsymbol{\sigma}})$. We set $\boldsymbol{\Sigma}_{\gamma,i} = (\boldsymbol{\sigma}_{\gamma,i}, \boldsymbol{\chi}_{\gamma,i})$ for $i = 1, 2$ and use

$$\boldsymbol{T}_1 = \boldsymbol{\Sigma}_{\gamma,2} + \widetilde{\boldsymbol{\Sigma}}, \qquad \boldsymbol{T}_2 = \boldsymbol{\Sigma}_{\gamma,1} - \widetilde{\boldsymbol{\Sigma}}$$

as test functions in (4.2), which yields

$$\begin{aligned} &a(\boldsymbol{\Sigma}_{\gamma,1} - \boldsymbol{\Sigma}_{\gamma,2}, \boldsymbol{\Sigma}_{\gamma,1} - \boldsymbol{\Sigma}_{\gamma,2}) \\ &\qquad \le (J'_\gamma(\boldsymbol{\Sigma}_{\gamma,1}) - J'_\gamma(\boldsymbol{\Sigma}_{\gamma,2}), \boldsymbol{\Sigma}_{\gamma,2} - \boldsymbol{\Sigma}_{\gamma,1} + \widetilde{\boldsymbol{\Sigma}}) + a(\boldsymbol{\Sigma}_{\gamma,1} - \boldsymbol{\Sigma}_{\gamma,2}, \widetilde{\boldsymbol{\Sigma}}). \end{aligned}$$

The first term on the right hand side was not present in Lemma 3.3. Assumption 2.5 implies that $\mathcal{K} + \widetilde{\boldsymbol{\Sigma}} = \mathcal{K}$ holds. Using the shift invariance of J'_γ from part (d) of Lemma 4.1, we infer that $J'_\gamma(\boldsymbol{\Sigma}_{\gamma,2}) = J'_\gamma(\boldsymbol{\Sigma}_{\gamma,2} + \widetilde{\boldsymbol{\Sigma}})$ holds. Thus we have

$$\begin{aligned} &(J'_\gamma(\boldsymbol{\Sigma}_{\gamma,1}) - J'_\gamma(\boldsymbol{\Sigma}_{\gamma,2}), \boldsymbol{\Sigma}_{\gamma,2} - \boldsymbol{\Sigma}_{\gamma,1} + \widetilde{\boldsymbol{\Sigma}}) \\ &\qquad = -(J'_\gamma(\boldsymbol{\Sigma}_{\gamma,1}) - J'_\gamma(\boldsymbol{\Sigma}_{\gamma,2} + \widetilde{\boldsymbol{\Sigma}}), \boldsymbol{\Sigma}_{\gamma,1} - (\boldsymbol{\Sigma}_{\gamma,2} + \widetilde{\boldsymbol{\Sigma}})) \le 0, \end{aligned}$$

and the inequality follows from the monotonicity of the derivative, see part (b) of Lemma 4.1. From here we can continue as in the proof of Lemma 3.3 until the end of step 1.

Step 2: Estimate for $\boldsymbol{u}$
In order to derive the esimates for the displacements, we set

$$\widetilde{\boldsymbol{T}} := (\boldsymbol{\varepsilon}(\boldsymbol{u}_{\gamma,1} - \boldsymbol{u}_{\gamma,2}),\ -\boldsymbol{\varepsilon}(\boldsymbol{u}_{\gamma,1} - \boldsymbol{u}_{\gamma,2}))$$

and use

$$\boldsymbol{T}_1 = (\boldsymbol{\tau}_1, \boldsymbol{\mu}_1) := \boldsymbol{\Sigma}_{\gamma,2} - \boldsymbol{\Sigma}_{\gamma,1} + \widetilde{\boldsymbol{T}}, \qquad \boldsymbol{T}_2 = (\boldsymbol{\tau}_2, \boldsymbol{\mu}_2) := \boldsymbol{\Sigma}_{\gamma,1} - \boldsymbol{\Sigma}_{\gamma,2} - \widetilde{\boldsymbol{T}}$$

as test functions in (4.1a). Adding both equations yields

$$\begin{aligned} a(\boldsymbol{\Sigma}_{\gamma,1} - \boldsymbol{\Sigma}_{\gamma,2}, \boldsymbol{\Sigma}_{\gamma,1} - \boldsymbol{\Sigma}_{\gamma,2}) &+ \int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{u}_{\gamma,1} - \boldsymbol{u}_{\gamma,2}) : \boldsymbol{\varepsilon}(\boldsymbol{u}_{\gamma,1} - \boldsymbol{u}_{\gamma,2})\, dx \\ &+ (J'_\gamma(\boldsymbol{\Sigma}_{\gamma,1}) - J'_\gamma(\boldsymbol{\Sigma}_{\gamma,2}), \boldsymbol{\Sigma}_{\gamma,1} - \boldsymbol{\Sigma}_{\gamma,2} - \widetilde{\boldsymbol{T}}) \\ = b(\mathbb{H}^{-1} : (\boldsymbol{\chi}_{\gamma,1} - \boldsymbol{\chi}_{\gamma,2}), \boldsymbol{u}_{\gamma,1} - \boldsymbol{u}_{\gamma,2}) &- b(\mathbb{C}^{-1} : (\boldsymbol{\sigma}_{\gamma,1} - \boldsymbol{\sigma}_{\gamma,2}), \boldsymbol{u}_{\gamma,1} - \boldsymbol{u}_{\gamma,2}) \\ &- b(\boldsymbol{\sigma}_{\gamma,1} - \boldsymbol{\sigma}_{\gamma,2}, \boldsymbol{u}_{\gamma,1} - \boldsymbol{u}_{\gamma,2}). \end{aligned} \tag{4.3}$$

Except for the term involving J'_γ, this is the same as (3.4) in the proof of Lemma 3.3. Similarly as in the discussion in Step 1 above, we infer that $J'_\gamma(\boldsymbol{\Sigma}_{\gamma,2}) = J'_\gamma(\boldsymbol{\Sigma}_{\gamma,2} + \widetilde{\boldsymbol{T}})$ and hence

$$\begin{aligned} &(J'_\gamma(\boldsymbol{\Sigma}_{\gamma,1}) - J'_\gamma(\boldsymbol{\Sigma}_{\gamma,2}), \boldsymbol{\Sigma}_{\gamma,1} - \boldsymbol{\Sigma}_{\gamma,2} - \widetilde{\boldsymbol{T}}) \\ &\qquad = (J'_\gamma(\boldsymbol{\Sigma}_{\gamma,1}) - J'_\gamma(\boldsymbol{\Sigma}_{\gamma,2} + \widetilde{\boldsymbol{T}}), \boldsymbol{\Sigma}_{\gamma,1} - (\boldsymbol{\Sigma}_{\gamma,2} + \widetilde{\boldsymbol{T}})) \ge 0 \end{aligned}$$

holds. Now (4.3) has exactly the same structure as (3.4), and we get the desired estimate with the same Lipschitz constant L. $\square$

It is easy to see that $(\boldsymbol{\Sigma}, \boldsymbol{u}) = 0$ solves (4.1) for $\ell = 0$. Hence Lemma 4.5 yields the following a priori estimate:

Corollary 4.6. *For every $\ell \in V'$, one has*

$$\|\boldsymbol{\sigma}_\gamma\|_S + \|\boldsymbol{\chi}_\gamma\|_M + \|\boldsymbol{u}_\gamma\|_V \leq L\,\|\ell\|_{V'}.$$

Remark 4.7. *We point out that the monotonicity and shift-invariance of J'_γ are essential for the analysis above. The assertion of Lemma 4.5 would also follow from a boundedness property of J'_γ, which was used in the proof of Theorem 8.12 in Han and Reddy [1999]. However, the verification of this property remains in doubt.*

By combining the analysis of [Han and Reddy, 1999, Section 8] and Hintermüller [2001], we now prove the strong convergence of $(\boldsymbol{\Sigma}_\gamma, \boldsymbol{u}_\gamma)$ to the solution of $(\mathbf{L})$, denoted as above by $(\boldsymbol{\Sigma}, \boldsymbol{u})$. A similar result is proved in [Temam, 1983, Theorem III.1.1] for the Hencky model.

Theorem 4.8. *Let $\ell \in V'$ be fixed, but arbitrary. Then, the solution and the Lagrange multiplier of $(\mathbf{L}_\gamma)$ converge strongly to the solution and the Lagrange multiplier of $(\mathbf{L})$ as γ tends to ∞, i.e.,*

$$(\boldsymbol{\sigma}_\gamma, \boldsymbol{\chi}_\gamma, \boldsymbol{u}_\gamma) \to (\boldsymbol{\sigma}, \boldsymbol{\chi}, \boldsymbol{u}) \quad \text{in } S \times M \times V, \text{ as } \gamma \to \infty.$$

Proof. The following analysis relies on a combination of arguments introduced in Han and Reddy [1999] and Hintermüller [2001]. We start with a given sequence of penalty parameters $\{\gamma_k\}$ tending to ∞ as $k \to \infty$. The associated solution and Lagrange multiplier of $(\mathbf{L}_{\gamma_k})$ is denoted by $(\boldsymbol{\Sigma}_k, \boldsymbol{u}_k) = (\boldsymbol{\sigma}_k, \boldsymbol{\chi}_k, \boldsymbol{u}_k)$. As before, we split the proof into two steps. First we prove the convergence of $\{\boldsymbol{\Sigma}_k\}$ by employing (4.2). Secondly, the strong convergence of $\{\boldsymbol{u}_k\}$ is derived by similar arguments as in the proof of Lemma 4.5.

Step 1: Convergence of $\{\boldsymbol{\Sigma}_k\}$
By Corollary 4.6, the sequence $\{\boldsymbol{\Sigma}_k\}$ is bounded in $S \times M$. Hence, there is a weakly converging subsequence, for simplicity it is also denoted by $\{\boldsymbol{\Sigma}_k\}$. The weak limit is denoted by $\overline{\boldsymbol{\Sigma}}$ and we show $\overline{\boldsymbol{\Sigma}} \in \mathcal{K}_\ell$. As in the proof of Lemma 3.3, Lemma 2.10 gives the existence of a unique $\widetilde{\boldsymbol{\sigma}} \in (\ker B)^\perp$ such that $\widetilde{\boldsymbol{\Sigma}} = (\widetilde{\boldsymbol{\sigma}}, -\widetilde{\boldsymbol{\sigma}}) \in \mathcal{C}_\ell^{\text{eq}}$. Moreover, $\widetilde{\boldsymbol{\Sigma}} \in \mathcal{K}$ holds thanks to Assumption 2.5 and therefore, one has $J_{\gamma_k}(\widetilde{\boldsymbol{\Sigma}}) = 0$ for all $k \in \mathbb{N}$. The convexity of J_γ thus implies

$$J_{\gamma_k}(\boldsymbol{\Sigma}_k) \leq \big(J'_{\gamma_k}(\boldsymbol{\Sigma}_k), \boldsymbol{\Sigma}_k - \widetilde{\boldsymbol{\Sigma}}\big) \leq a(\boldsymbol{\Sigma}_k, \widetilde{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_k).$$

We used (4.2) for the last estimate, which is applicable here, since $\widetilde{\boldsymbol{\Sigma}} \in \mathcal{C}_\ell^{\text{eq}}$ by construction. Hence, (2.6), (2.10) and Corollary 4.6 allow us to conclude

$$J_{\gamma_k}(\boldsymbol{\Sigma}_k) \leq (\overline{\alpha}\, L^2 + \sqrt{2}\, c_B\, L)\, \|\ell\|_{V'}^2 =: C,$$

which implies the boundedness of $J_{\gamma_k}(\boldsymbol{\Sigma}_k)$. By definition of J_γ, we therefore obtain

$$\begin{aligned} 0 \leq \|\overline{\boldsymbol{\Sigma}} - P_{\mathcal{K}}(\overline{\boldsymbol{\Sigma}})\|_{S\times M}^2 &\leq \liminf_{k\to\infty} \|\boldsymbol{\Sigma}_k - P_{\mathcal{K}}(\boldsymbol{\Sigma}_k)\|_{S\times M}^2 \\ &\leq \limsup_{k\to\infty} \|\boldsymbol{\Sigma}_k - P_{\mathcal{K}}(\boldsymbol{\Sigma}_k)\|_{S\times M}^2 \leq \frac{2C}{\gamma_k} \to 0, \quad \text{as } k \to \infty. \end{aligned}$$

Here, we used the weak lower semicontinuity of $\| \cdot - P_{\mathcal{K}}(\cdot)\|$ which follows from Lemma 4.1. Hence, $\overline{\boldsymbol{\Sigma}} = P_{\mathcal{K}}(\overline{\boldsymbol{\Sigma}})$ holds, which implies $\overline{\boldsymbol{\Sigma}} \in \mathcal{K}$. Since $\overline{\boldsymbol{\Sigma}} \in \mathcal{C}_\ell^{\mathrm{eq}}$ due to the weak convergence $\boldsymbol{\sigma}_k \rightharpoonup \overline{\boldsymbol{\sigma}}$ in S, we obtain $\overline{\boldsymbol{\Sigma}} \in \mathcal{K}_\ell$, i.e., $\overline{\boldsymbol{\Sigma}}$ is feasible for $(\mathbf{L})$.

The optimality of $\boldsymbol{\Sigma}_k$ gives

$$\frac{1}{2}\,a(\boldsymbol{\Sigma}_k, \boldsymbol{\Sigma}_k) + J_{\gamma_k}(\boldsymbol{\Sigma}_k) \le \frac{1}{2}\,a(\boldsymbol{T}, \boldsymbol{T}) + J_{\gamma_k}(\boldsymbol{T}) \quad \text{for all } \boldsymbol{T} \in \mathcal{C}_\ell^{\mathrm{eq}}.$$

The above inequality holds in particular for all $\boldsymbol{T} \in \mathcal{C}_\ell^{\mathrm{eq}} \cap \mathcal{K} = \mathcal{K}_\ell$, and consequently

$$\frac{1}{2}\,a(\boldsymbol{\Sigma}_k, \boldsymbol{\Sigma}_k) \le \frac{1}{2}\,a(\boldsymbol{T}, \boldsymbol{T}) \quad \text{for all } \boldsymbol{T} \in \mathcal{K}_\ell,$$

where we used the non-negativity of $J_{\gamma_k}(\boldsymbol{\Sigma}_k)$ and $J_{\gamma_k}(\boldsymbol{T}) = 0$ for $\boldsymbol{T} \in \mathcal{K}$ for all $k \in \mathbb{N}$. Since $\overline{\boldsymbol{\Sigma}} \in \mathcal{K}_\ell$ as seen above, the weak lower semicontinuity of $a(\,\cdot\,,\,\cdot\,)$ thus implies

$$\begin{aligned} \frac{1}{2}\,a(\overline{\boldsymbol{\Sigma}}, \overline{\boldsymbol{\Sigma}}) &\le \liminf_{k\to\infty} \frac{1}{2}\,a(\boldsymbol{\Sigma}_k, \boldsymbol{\Sigma}_k) \\ &\le \limsup_{k\to\infty} \frac{1}{2}\,a(\boldsymbol{\Sigma}_k, \boldsymbol{\Sigma}_k) \le \frac{1}{2}\,a(\boldsymbol{T}, \boldsymbol{T}) \text{ for all } \boldsymbol{T} \in \mathcal{K}_\ell. \end{aligned} \tag{4.4}$$

Therefore, $\overline{\boldsymbol{\Sigma}}$ is the unique solution of $(\mathbf{L})$, so as before, we simply denote it by $\boldsymbol{\Sigma}$ for the rest of the proof. By standard arguments, the uniqueness of $\boldsymbol{\Sigma}$ guarantees the weak convergence of the whole sequence.

By inserting $\boldsymbol{T} = \boldsymbol{\Sigma}$ in (4.4), the convergence $a(\boldsymbol{\Sigma}_k, \boldsymbol{\Sigma}_k) \to a(\boldsymbol{\Sigma}, \boldsymbol{\Sigma})$ follows. Since $a(\boldsymbol{\Sigma}, \boldsymbol{\Sigma})$ is an equivalent norm on S by (2.5) and (2.6), this implies convergence of the norm, i.e., $\|\boldsymbol{\Sigma}_k\|_{S\times M} \to \|\boldsymbol{\Sigma}\|_{S\times M}$. Together with the weak convergence, strong convergence is obtained.

Step 2: Convergence of $\{\boldsymbol{u}_k\}$
If we insert $\boldsymbol{T}_1 - \boldsymbol{\Sigma}_k$ with $\boldsymbol{T}_1 \in \mathcal{K}$ as test function in (4.1a), then part (c) of Lemma 4.1 implies

$$a(\boldsymbol{\Sigma}_k, \boldsymbol{T}_1 - \boldsymbol{\Sigma}_k) + b(\boldsymbol{\tau}_1 - \boldsymbol{\sigma}_k, \boldsymbol{u}_k) \ge 0 \quad \text{for all } \boldsymbol{T}_1 = (\boldsymbol{\tau}_1, \boldsymbol{\mu}_1) \in \mathcal{K}. \tag{4.5}$$

Moreover, as $\boldsymbol{\Sigma}$ is the unique solution of $(\mathbf{L})$, it fulfills (2.11a), i.e.,

$$a(\boldsymbol{\Sigma}, \boldsymbol{T}_2 - \boldsymbol{\Sigma}) + b(\boldsymbol{\tau}_2 - \boldsymbol{\sigma}, \boldsymbol{u}) \ge 0 \quad \text{for all } \boldsymbol{T}_2 = (\boldsymbol{\tau}_2, \boldsymbol{\mu}_2) \in \mathcal{K}. \tag{4.6}$$

Here $\boldsymbol{u}$ is the associated Lagrange multiplier, which is unique by Lemma 3.3. Next, we proceed similarly to step 2 in the proof of Lemma 3.3. Since $\boldsymbol{\Sigma} \in \mathcal{K}$, Assumption 2.5 allows us to insert $\boldsymbol{T}_1 = (\boldsymbol{\sigma} + \boldsymbol{\varepsilon}(\boldsymbol{u}_k - \boldsymbol{u}), \boldsymbol{\chi} - \boldsymbol{\varepsilon}(\boldsymbol{u}_k - \boldsymbol{u}))$ into (4.5). Unfortunately, $\boldsymbol{\Sigma}_k$ is not feasible for (4.6), but we can use $P_{\mathcal{K}}(\boldsymbol{\Sigma}_k)$ instead and insert

$$\boldsymbol{T}_2 = (P_{\mathcal{K}}^{\boldsymbol{\sigma}}(\boldsymbol{\sigma}_k) + \boldsymbol{\varepsilon}(\boldsymbol{u} - \boldsymbol{u}_k), P_{\mathcal{K}}^{\boldsymbol{\chi}}(\boldsymbol{\chi}_k) - \boldsymbol{\varepsilon}(\boldsymbol{u} - \boldsymbol{u}_k)).$$

Here, $P_{\mathcal{K}}^{\boldsymbol{\sigma}}$ and $P_{\mathcal{K}}^{\boldsymbol{\chi}}$ refer to the components of $P_{\mathcal{K}}$. Adding the arising inequalities give

$$\begin{aligned} &a(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_k, \boldsymbol{\Sigma} - \boldsymbol{\Sigma}_k) + \int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{u} - \boldsymbol{u}_k) : \boldsymbol{\varepsilon}(\boldsymbol{u} - \boldsymbol{u}_k)\,dx \\ &\quad \le b(\mathbb{H}^{-1} : (\boldsymbol{\chi} - \boldsymbol{\chi}_k), \boldsymbol{u} - \boldsymbol{u}_k) - b(\mathbb{C}^{-1} : (\boldsymbol{\sigma} - \boldsymbol{\sigma}_k), \boldsymbol{u} - \boldsymbol{u}_k) - b(\boldsymbol{\sigma} - \boldsymbol{\sigma}_k, \boldsymbol{u} - \boldsymbol{u}_k) \\ &\qquad + a(\boldsymbol{\Sigma}, P_{\mathcal{K}}(\boldsymbol{\Sigma}_k) - \boldsymbol{\Sigma}_k) + b(P_{\mathcal{K}}^{\boldsymbol{\sigma}}(\boldsymbol{\sigma}_k) - \boldsymbol{\sigma}_k, \boldsymbol{u}). \end{aligned}$$

The terms on the left hand side as well as the first three addends on the right hand side can be estimated as in the proof of Lemma 3.3. Using (2.6) and (2.7) for the remaing terms, we obtain

$$\|\boldsymbol{u} - \boldsymbol{u}_k\|_V^2 \le c\Big(\|\boldsymbol{\Sigma}\|_{S\times M}\|P_{\mathcal{K}}(\boldsymbol{\Sigma}_k) - \boldsymbol{\Sigma}_k\|_{S\times M} + \|\boldsymbol{u}\|_V\|\boldsymbol{\sigma}_k - P_{\mathcal{K}}^{\boldsymbol{\sigma}}(\boldsymbol{\sigma}_k)\|_S \\ + \|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_k\|_{S\times M}\|\boldsymbol{u} - \boldsymbol{u}_k\|_V\Big).$$

Now, the continuity of $P_{\mathcal{K}}$ implies $P_{\mathcal{K}}(\boldsymbol{\Sigma}_k) \to P_{\mathcal{K}}(\boldsymbol{\Sigma}) = \boldsymbol{\Sigma}$ such that the convergence of $\{\boldsymbol{\Sigma}_k\}$ and an application of Young's inequality yield the desired convergence of $\{\boldsymbol{u}_k\}$.

The above argument is valid for arbitrary sequences $\gamma_k \to \infty$. The limit $(\boldsymbol{\sigma}, \boldsymbol{\chi}, \boldsymbol{u})$ is the unique solution and Lagrange multiplier of $(\mathbf{L})$. Therefore, $(\boldsymbol{\sigma}_\gamma, \boldsymbol{\chi}_\gamma, \boldsymbol{u}_\gamma) \to (\boldsymbol{\sigma}, \boldsymbol{\chi}, \boldsymbol{u})$ holds as claimed. $\square$

Remark 4.9. *We point out that the arguments for the convergence of $\boldsymbol{\sigma}$ are similar to those in the proof of Theorem III.1.1 in Temam [1983] for the Hencky model. However, due to the low regularity of the displacement field $\boldsymbol{u}$, corresponding convergence result for $\boldsymbol{u}$ cannot be expected in that case.*

Next, we turn to the bi-level problem associated to $(\mathbf{L}_\gamma)$ which is given by

$$\left.\begin{aligned} \text{Minimize}\quad & F(\boldsymbol{u}, \boldsymbol{f}, \boldsymbol{g}) := \frac{1}{2}\,\|\boldsymbol{u} - \boldsymbol{u}_d\|^2_{L^2(\Omega,\mathbb{R}^d)} \\ & \qquad + \frac{\nu_1}{2}\|\boldsymbol{f}\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_2}{2}\|\boldsymbol{g}\|^2_{L^2(\Gamma_N;\mathbb{R}^d)} \\ \text{s.t.}\quad & (\mathbf{L}_\gamma) \text{ with } \ell = R(\boldsymbol{f}, \boldsymbol{g}). \end{aligned}\right\} \qquad (\mathbf{P}_\gamma)$$

Based on the above results, it is straightforward to adapt the proof of Proposition 3.6 to problem $(\mathbf{P}_\gamma)$ to obtain:

Proposition 4.10. *For each $\gamma > 0$, problem $(\mathbf{P}_\gamma)$ has a global optimal solution $(\boldsymbol{f}^*_\gamma, \boldsymbol{g}^*_\gamma) \in U$.*

Remark 4.11. *As a consequence of Proposition 4.4, we can replace the lower-level problem $(\mathbf{L}_\gamma)$ by its necessary and sufficient optimality conditions* (4.1). *We point out that $(\mathbf{P}_\gamma)$ then becomes an optimal control problem for a partial differential equation in mixed variational form* (4.1).

5 Convergence for the Upper-Level Solutions

The results of the previous section give rise the definition of a solution operator for the viscoplastic lower-level problem $(\mathbf{L}_\gamma)$,

$$S_\gamma : U \ni (\boldsymbol{f}, \boldsymbol{g}) \mapsto (\boldsymbol{\sigma}_\gamma, \boldsymbol{\chi}_\gamma, \boldsymbol{u}_\gamma) \in Y := S \times M \times V.$$

Note that S_γ is nonlinear due to the term involving J'_γ. By Lemma 4.5 and 3.5, S_γ is Lipschitz continuous and compact, both uniformly with respect to γ.

Theorem 5.1. *Let $\{\gamma_k\}$ be a sequence tending to ∞ and let $(\boldsymbol{f}^*_k, \boldsymbol{g}^*_k)$ denote a global solution to $(\mathbf{P}_\gamma)$ with $\gamma = \gamma_k$.*

(a) *There exists an accumulation point $(\boldsymbol{f}^*, \boldsymbol{g}^*)$ in the strong topology of U.*

(b) *Every weak accumulation point of* $\{(\boldsymbol{f}_k^*, \boldsymbol{g}_k^*)\}$ *is a global optimal solution of* $(\mathbf{P})$.

Proof. We denote by S_k the solution operator associated to γ_k. The optimality of $(\boldsymbol{f}_k^*, \boldsymbol{g}_k^*)$, the feasibility of $(\mathbf{0}, \mathbf{0})$ and $S_k(\mathbf{0}, \mathbf{0}) = (\mathbf{0}, \mathbf{0}, \mathbf{0})$ for every k (see Corollary 4.6) imply

$$\begin{aligned}\frac{1}{2}\|S_k^{\boldsymbol{u}}(\boldsymbol{f}_k^*, \boldsymbol{g}_k^*) - \boldsymbol{u}_d\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_1}{2}\|\boldsymbol{f}_k^*\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_2}{2}\|\boldsymbol{g}_k^*\|^2_{L^2(\Gamma_N;\mathbb{R}^d)}& \\ \leq \frac{1}{2}\|S_k^{\boldsymbol{u}}(\boldsymbol{f}, \boldsymbol{g}) - \boldsymbol{u}_d\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_1}{2}\|\boldsymbol{f}\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_2}{2}\|\boldsymbol{g}\|^2_{L^2(\Gamma_N;\mathbb{R}^d)}& \quad (5.1)\end{aligned}$$

for all $(\boldsymbol{f}, \boldsymbol{g}) \in U$. Inserting $(\boldsymbol{f}, \boldsymbol{g}) = (\mathbf{0}, \mathbf{0})$ and observing that $S_k(\mathbf{0}, \mathbf{0}) = \mathbf{0}$, we obtain the boundedness of the sequence $\{(\boldsymbol{f}_k^*, \boldsymbol{g}_k^*)\}$ in U. And thus there exists a weakly convergent subsequence, which we denote by $\{(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*)\}$. The weak limit in U is called $(\boldsymbol{f}^*, \boldsymbol{g}^*)$. Using Lemma 4.5, we infer

$$\|S_{k'}(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*) - S_{k'}(\boldsymbol{f}^*, \boldsymbol{g}^*)\|_Y \leq L\,\|R(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*) - R(\boldsymbol{f}^*, \boldsymbol{g}^*)\|_{V'},$$

which converges to zero as $k' \to \infty$ due to the compactness of R (Lemma 3.5). Moreover, we have $S_{k'}(\boldsymbol{f}^*, \boldsymbol{g}^*) \to S(\boldsymbol{f}^*, \boldsymbol{g}^*)$ for $k' \to \infty$ due to Theorem 4.8. Thus we conclude

$$\begin{aligned}&\|S_{k'}(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*) - S(\boldsymbol{f}^*, \boldsymbol{g}^*)\|_Y \\ &\qquad \leq \|S_{k'}(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*) - S_{k'}(\boldsymbol{f}^*, \boldsymbol{g}^*)\|_Y + \|S_{k'}(\boldsymbol{f}^*, \boldsymbol{g}^*) - S(\boldsymbol{f}^*, \boldsymbol{g}^*)\|_Y \to 0 \quad \text{as } k' \to \infty.\end{aligned}$$

Together with the weak lower semicontinuity of norms, this implies

$$\begin{aligned}F(S^{\boldsymbol{u}}(\boldsymbol{f}^*, \boldsymbol{g}^*), \boldsymbol{f}^*, \boldsymbol{g}^*) &\leq \liminf_{k'\to\infty} F(S_{k'}^{\boldsymbol{u}}(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*), \boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*) \\ &\leq \limsup_{k'\to\infty} F(S_{k'}^{\boldsymbol{u}}(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*), \boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*) \\ &\leq \limsup_{k'\to\infty} F(S_{k'}^{\boldsymbol{u}}(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{f}, \boldsymbol{g}) && \text{by (5.1)} \\ &= F(S^{\boldsymbol{u}}(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{f}, \boldsymbol{g}) && \text{by Theorem 4.8}\end{aligned}$$

for all $(\boldsymbol{f}, \boldsymbol{g}) \in U$. Therefore, $(\boldsymbol{f}^*, \boldsymbol{g}^*)$ is a global optimal solution of $(\mathbf{P})$.

Inserting $(\boldsymbol{f}, \boldsymbol{g}) = (\boldsymbol{f}^*, \boldsymbol{g}^*)$ in the inequality above, we infer the convergence

$$F(S_{k'}^{\boldsymbol{u}}(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*), \boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*) \to F(S^{\boldsymbol{u}}(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{f}, \boldsymbol{g}).$$

Together with the strong convergence $S_{k'}(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*) \to S(\boldsymbol{f}^*, \boldsymbol{g}^*)$, this yields convergence of norms $\|(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*)\| \to \|(\boldsymbol{f}^*, \boldsymbol{g}^*)\|$. Using the weak convergence shown above, we obtain the strong convergence

$$(\boldsymbol{f}_{k'}^*, \boldsymbol{g}_{k'}^*) \to (\boldsymbol{f}^*, \boldsymbol{g}^*) \quad \text{in } U \qquad \text{as } k' \to \infty.$$

This proves assertion (a). Since the above arguments leading to the optimality of the weak limit hold for every weakly convergent subsequence of $(\boldsymbol{f}_k^*, \boldsymbol{g}_k^*)$, assertion (b) is also proved. $\square$

Remark 5.2. *The proof of Theorem 5.1 shows that every weak accumulation point of* $\{(f_k^*, g_k^*)\}$ *is automatically a strong accumulation point.*

The necessary modifications of the above arguments are obvious in case of additional control constraints in $(\mathbf{P})$ of the form

$$(\boldsymbol{f}, \boldsymbol{g}) \in \mathcal{U}_{ad}$$

with a closed and convex subset $\mathcal{U}_{ad} \subset U = L^2(\Omega;\mathbb{R}^d)\times L^2(\Gamma_N;\mathbb{R}^d)$, as is given for example by constraints of the form

$$\mathcal{U}_{ad} = \{(\boldsymbol{f},\boldsymbol{g}) \in U \,:\, |\boldsymbol{f}(x)|_{\mathbb{R}^d} \le \rho_1 \text{ a.e. in } \Omega \text{ and } |\boldsymbol{g}(x)|_{\mathbb{R}^d} \le \rho_2 \text{ a.e. on } \Gamma_N\}. \tag{5.2}$$

Since U is reflexive and $\mathcal{U}_{ad}$ is weakly closed, the weak limit $(\boldsymbol{f}^*,\boldsymbol{g}^*)$ in the proof of Theorem 5.1 clearly satisfies the additional control constraints. The rest of the theory above is not affected by the control constraints, and hence we obtain following result:

Corollary 5.3. *Suppose that* $(\mathbf{P})$ *contains additional control constraints, i.e.,*

$$\left.\begin{aligned} \text{Minimize}\quad & F(\boldsymbol{u},\boldsymbol{f},\boldsymbol{g}) := \frac{1}{2}\,\|\boldsymbol{u}-\boldsymbol{u}_d\|^2_{L^2(\Omega,\mathbb{R}^d)} \\ &\qquad + \frac{\nu_1}{2}\|\boldsymbol{f}\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_2}{2}\|\boldsymbol{g}\|^2_{L^2(\Gamma_N;\mathbb{R}^d)} \\ \text{s.t.}\quad & \text{the plasticity problem } (\mathbf{L}) \text{ with } \ell \text{ as in (2.13)} \\ \text{and}\quad & (\boldsymbol{f},\boldsymbol{g}) \in \mathcal{U}_{ad} \end{aligned}\right\} \qquad (\mathbf{P})$$

with a closed and convex subset $\mathcal{U}_{ad} \subset U$*, and let the regularized problems be defined analogously to* $(\mathbf{P}_\gamma)$*. Then the assertion of Theorem 5.1 remains true, i.e., if* $\{\gamma_k\}$ *is a sequence of numbers tending to* ∞ *and* $(\boldsymbol{f}_k,\boldsymbol{g}_k)$ *are global solutions to* $(\mathbf{P}_\gamma)$ *with* $\gamma=\gamma_k$*, then there exists a weak accumulation point* $(\boldsymbol{f}^*,\boldsymbol{g}^*)$*, which is a strong accumulation point and in addition a solution of* $(\mathbf{P})$*.*

The following theorem answers the question which optima of $(\mathbf{P})$ can be approximated by a sequence of solutions of viscoplastic problems. The underlying analysis is standard and follows a classical argument which was, for instance, given in Casas and Tröltzsch [2002].

Theorem 5.4. *Suppose that* $(\boldsymbol{f}^*,\boldsymbol{g}^*)$ *is a strict local optimum of* $(\mathbf{P})$ *in the topology of* U*. Let* γ_k *be an arbitrary sequence tending to* ∞*. Then there exists a sequence* $(\boldsymbol{f}_k^*,\boldsymbol{g}_k^*)$ *of local optimal solutions of* $(\mathbf{P}_{\gamma_k})$ *such that* $(\boldsymbol{f}_k^*,\boldsymbol{g}_k^*)\to(\boldsymbol{f}^*,\boldsymbol{g}^*)$ *strongly in* U*.*

Proof. Let $\varepsilon>0$ be the radius of the neighborhood of strict local optimality of $(\boldsymbol{f}^*,\boldsymbol{g}^*)$. We start by defining the following auxiliary problem:

$$\left.\begin{aligned} \text{Minimize}\quad & F(\boldsymbol{u},\boldsymbol{f},\boldsymbol{g}) := \frac{1}{2}\,\|\boldsymbol{u}-\boldsymbol{u}_d\|^2_{L^2(\Omega,\mathbb{R}^d)} \\ &\qquad + \frac{\nu_1}{2}\|\boldsymbol{f}\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_2}{2}\|\boldsymbol{g}\|^2_{L^2(\Gamma_N;\mathbb{R}^d)} \\ \text{s.t.}\quad & \text{the plasticity problem } (\mathbf{L}) \text{ with } \ell \text{ as in (2.13)} \\ \text{and}\quad & (\boldsymbol{f},\boldsymbol{g}) \in B_\delta(\boldsymbol{f}^*,\boldsymbol{g}^*), \end{aligned}\right\} \qquad (\mathbf{P}^\delta)$$

where δ satisfies $0<\delta<\varepsilon$ and $B_\delta(\boldsymbol{f}^*,\boldsymbol{g}^*)\subset U$ is the closed ball of radius δ centered at $(\boldsymbol{f}^*,\boldsymbol{g}^*)$ in the topology of U. Thus the assumption on $(\boldsymbol{f}^*,\boldsymbol{g}^*)$ implies

$$\begin{aligned} F(S^{\boldsymbol{u}}(\boldsymbol{f},\boldsymbol{g}),\boldsymbol{f},\boldsymbol{g}) >& F(S^{\boldsymbol{u}}(\boldsymbol{f}^*,\boldsymbol{g}^*),\boldsymbol{f}^*,\boldsymbol{g}^*) \\ & \text{for all } (\boldsymbol{f},\boldsymbol{g}) \in B_\delta(\boldsymbol{f}^*,\boldsymbol{g}^*)\setminus\{(\boldsymbol{f}^*,\boldsymbol{g}^*)\}, \end{aligned} \tag{5.3}$$

to the effect that $(\boldsymbol{f}^*,\boldsymbol{g}^*)$ is the *unique* global optimum of $(\mathbf{P}^\delta)$. Since $B_\delta(\boldsymbol{f}^*,\boldsymbol{g}^*)$ is closed and convex, Corollary 5.3 yields the existence of a sequence $(\boldsymbol{f}_k^*,\boldsymbol{g}_k^*)$ of solutions to the associated regularized problems $(\mathbf{P}^\delta_{\gamma_k})$, that converges strongly in U to $(\boldsymbol{f}^*,\boldsymbol{g}^*)$. It remains to show that $(\boldsymbol{f}_k^*,\boldsymbol{g}_k^*)$ is a local optimum of $(\mathbf{P}^\delta_{\gamma_k})$ for all sufficiently large k. To this end, take an arbitrary $(\boldsymbol{f}, \boldsymbol{g}) \in U$ with $\|(\boldsymbol{f}, \boldsymbol{g}) - (\boldsymbol{f}_k^*, \boldsymbol{g}_k^*)\|_U < \delta/2$. Then, Corollary 5.3 yields that, for sufficiently large k,

$$\|(\boldsymbol{f}, \boldsymbol{g}) - (\boldsymbol{f}^*, \boldsymbol{g}^*)\|_U \leq \|(\boldsymbol{f}, \boldsymbol{g}) - (\boldsymbol{f}_k^*, \boldsymbol{g}_k^*)\|_U + \|(\boldsymbol{f}_k^*, \boldsymbol{g}_k^*) - (\boldsymbol{f}^*, \boldsymbol{g}^*)\|_U < \delta.$$

This in turn implies that $(\boldsymbol{f}, \boldsymbol{g}) \in B_\delta(\boldsymbol{f}^*, \boldsymbol{g}^*)$, i.e. $(\boldsymbol{f}, \boldsymbol{g})$ is feasible for $(\mathbf{P}^\delta_{\gamma_k})$. Since $(\boldsymbol{f}, \boldsymbol{g})$ was chosen arbitrary, the (global) optimality of $(\boldsymbol{f}_k^*, \boldsymbol{g}_k^*)$ for $(\mathbf{P}^\delta_{\gamma_k})$ ensures

$$F(S^{\boldsymbol{u}}(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{f}, \boldsymbol{g}) \geq F(S^{\boldsymbol{u}}(\boldsymbol{f}_k^*, \boldsymbol{g}_k^*), \boldsymbol{f}_k^*, \boldsymbol{g}_k^*) \quad \forall\, (\boldsymbol{f}, \boldsymbol{g}) \text{ with } \|(\boldsymbol{f}, \boldsymbol{g}) - (\boldsymbol{f}_k^*, \boldsymbol{g}_k^*)\|_U < \frac{\delta}{2},$$

which amounts to local optimality of $(\boldsymbol{f}_k^*, \boldsymbol{g}_k^*)$. □

Remark 5.5. *Similarly to Corollary 5.3, it is straightforward to incoporate additional control constraints into Theorem 5.4 as for instance constraints of the form* (5.2).

A Lagrange Duality

The Lagrange multiplier $\boldsymbol{u}$ associated to the equality constraint in $(\mathbf{L})$ can be viewed as the displacement field induced by the load functional ℓ. This is due to the fact that $\boldsymbol{u}$ solves the so-called primal optimization problem, which is obtained traditionally by means of Fenchel duality, cf. [Temam, 1983, Theorem III.1.3]. Here, we consider another notion of primal problem that involves the so-called *plastic strain* $\boldsymbol{p}$ and an *internal hardening variable* $\boldsymbol{\xi}$ as optimization variables, in addition to $\boldsymbol{u}$. This definition of the primal problem coincides with the one used in [Han and Reddy, 1999, Section 7] or Carstensen [1999]. By means of Lagrange duality, we will see that the primal problem can be identified with the dual problem associated to $(\mathbf{L})$. We point out that the same result could also be obtained using Fenchel duality, similarly to [Temam, 1983, Theorem III.1.3].

We start by introducing the Lagrange function $\mathcal{L} : S \times M \times V \to \mathbb{R}$ associated to $(\mathbf{L})$

$$\mathcal{L}(\boldsymbol{\Sigma}, \boldsymbol{u}) = \frac{1}{2} a(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}) + b(\boldsymbol{\sigma}, \boldsymbol{u}) - \langle \ell, \boldsymbol{u} \rangle. \tag{A.1}$$

Since $(\mathbf{L})$ is a strictly convex problem, the solution, together with the Lagrange multiplier, is a saddle point of the Lagrange function. That is, the solution of $(\mathbf{L})$, denoted as before by $\boldsymbol{\Sigma}$, and $\boldsymbol{u}$ satisfy

$$\mathcal{L}(\boldsymbol{\Sigma}, \boldsymbol{v}) \leq \mathcal{L}(\boldsymbol{\Sigma}, \boldsymbol{u}) \leq \mathcal{L}(\boldsymbol{T}, \boldsymbol{u}) \quad \text{for all } \boldsymbol{v} \in V \text{ and } \boldsymbol{T} \in \mathcal{K}. \tag{A.2}$$

By standard arguments, $(\mathbf{L})$ is equivalent to

$$\inf_{\boldsymbol{T} \in \mathcal{K}} \sup_{\boldsymbol{v} \in V} \mathcal{L}(\boldsymbol{T}, \boldsymbol{v}). \tag{\textbf{LD}}$$

The dual problem associated to $(\mathbf{LD})$ arises by interchanging inf and sup:

$$\sup_{\boldsymbol{v} \in V} \inf_{\boldsymbol{T} \in \mathcal{K}} \mathcal{L}(\boldsymbol{T}, \boldsymbol{v}). \tag{\textbf{LP}}$$

Due to (A.2), there is no duality gap, which implies that $(\boldsymbol{\Sigma}, \boldsymbol{u})$ is the unique solution of the dual problem $(\mathbf{LP})$.

Definition A.1. *For given* $(\boldsymbol{u},\boldsymbol{P}) = (\boldsymbol{u},\boldsymbol{p},\boldsymbol{\xi}) \in V \times S \times M$ *and* $(\boldsymbol{v},\boldsymbol{Q}) = (\boldsymbol{v},\boldsymbol{q},\boldsymbol{\eta}) \in V \times S \times M$*, we define*

$$\mathfrak{a}\big((\boldsymbol{u},\boldsymbol{P}),(\boldsymbol{v},\boldsymbol{Q})\big) := \int_\Omega \big[(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}) : \mathbb{C} : (\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{q}) + \boldsymbol{\xi} : \mathbb{H} : \boldsymbol{\eta}\big]\, dx.$$

Moreover, let $I^*_{\mathcal{K}}$ *denote the support functional of* $\mathcal{K}$*, i.e.,*

$$I^*_{\mathcal{K}}(\boldsymbol{P}) = \sup_{\boldsymbol{T}\in\mathcal{K}} \int_\Omega \boldsymbol{P} : \boldsymbol{T}\, dx.$$

Here and in the following, the expression $\boldsymbol{P} : \boldsymbol{T}$ *with* $\boldsymbol{P} = (\boldsymbol{p},\boldsymbol{\xi})$ *and* $\boldsymbol{T} = (\boldsymbol{\tau},\boldsymbol{\mu})$ *with given tensors* $\boldsymbol{p},\boldsymbol{\tau} \in \mathbb{R}^{d\times d}_{\text{sym}}$ *and* $\boldsymbol{\xi},\boldsymbol{\mu} \in \mathbb{R}^{d\times d}$ *refers to* $\boldsymbol{P} : \boldsymbol{T} := \boldsymbol{p} : \boldsymbol{\tau} + \boldsymbol{\xi} : \boldsymbol{\mu}$.

Next, we introduce the following optimization problem

$$\inf_{(\boldsymbol{v},\boldsymbol{Q})\in V\times S\times M} \frac{1}{2}\,\mathfrak{a}\big((\boldsymbol{v},\boldsymbol{Q}),(\boldsymbol{v},\boldsymbol{Q})\big) + \langle \ell\,,\, \boldsymbol{v}\rangle + I^*_{\mathcal{K}}(\boldsymbol{Q}). \tag{$\mathbf{LP'}$}$$

The following lemma shows in which sense ($\mathbf{LP}$) and ($\mathbf{LP'}$) coincide.

Lemma A.2. *Problem* ($\mathbf{LP'}$) *admits a unique solution, denoted by* $(\boldsymbol{u},\boldsymbol{P}) = (\boldsymbol{u},\boldsymbol{p},\boldsymbol{\xi})$*, which is related to the unique solution* $(\boldsymbol{u},\boldsymbol{\sigma},\boldsymbol{\chi})$ *of* ($\mathbf{LP}$) *by*

$$\boldsymbol{p} = \boldsymbol{\varepsilon}(\boldsymbol{u}) - \mathbb{C}^{-1} : \boldsymbol{\sigma} \quad \text{and} \quad \boldsymbol{\xi} = -\mathbb{H}^{-1} : \boldsymbol{\chi}. \tag{A.3}$$

Proof. We start with the inf-problem in ($\mathbf{LP}$), i.e.,

$$\inf_{\boldsymbol{T}\in\mathcal{K}} \mathcal{L}(\boldsymbol{T},\boldsymbol{v})$$

for given $\boldsymbol{v} \in V$. In view of the definition of $\mathcal{L}$, the necessary and sufficient optimality conditions for this problem are given by

$$\boldsymbol{T} \in \mathcal{K} \quad \text{and} \quad a(\boldsymbol{T},\boldsymbol{\Upsilon} - \boldsymbol{T}) + b(\boldsymbol{\varsigma} - \boldsymbol{\tau},\boldsymbol{v}) \geq 0 \quad \forall\, \boldsymbol{\Upsilon} = (\boldsymbol{\varsigma},\boldsymbol{\psi}) \in \mathcal{K}.$$

Hence, ($\mathbf{LP}$) is equivalent to

$$(\mathbf{LP}) \quad \iff \quad \begin{cases} \sup\limits_{\boldsymbol{v},\boldsymbol{T}} & \mathcal{L}(\boldsymbol{T},\boldsymbol{v}) \\ \text{s.t.} & \boldsymbol{v} \in V,\ \boldsymbol{T} \in \mathcal{K} \\ \text{and} & a(\boldsymbol{T},\boldsymbol{\Upsilon} - \boldsymbol{T}) + b(\boldsymbol{\varsigma} - \boldsymbol{\tau},\boldsymbol{v}) \geq 0 \quad \forall\, \boldsymbol{\Upsilon} \in \mathcal{K}. \end{cases} \tag{A.4}$$

Next, we turn to ($\mathbf{LP'}$). Since the objective of ($\mathbf{LP'}$) is strictly convex, there is a unique solution $(\boldsymbol{u},\boldsymbol{P}) = (\boldsymbol{u},\boldsymbol{p},\boldsymbol{\xi})$ of ($\mathbf{LP'}$). The necessary and sufficient conditions are given by

$$\begin{aligned} \mathbf{0} &\in \partial\Big(\frac{1}{2}\,\mathfrak{a}\big((\boldsymbol{u},\boldsymbol{P}),(\boldsymbol{u},\boldsymbol{P})\big) + \langle \ell\,,\, \boldsymbol{u}\rangle + I^*_{\mathcal{K}}(\boldsymbol{P})\Big) \\ &\qquad = \partial\Big(\frac{1}{2}\,\mathfrak{a}\big((\boldsymbol{u},\boldsymbol{P}),(\boldsymbol{u},\boldsymbol{P})\big) + \langle \ell\,,\, \boldsymbol{u}\rangle\Big) + \partial I^*_{\mathcal{K}}(\boldsymbol{P}). \end{aligned} \tag{A.5}$$

Here we used the sum rule of subdifferential calculus, which holds since $\mathbf{0} \in \operatorname{dom}(1/2\,\mathfrak{a}(\cdot\,,\,\cdot) + \langle \ell\,,\,\cdot\rangle) \cap \operatorname{dom} I^*_{\mathcal{K}}$ and due to the continuity of $\mathfrak{a}$. Clearly, the first addend is a singleton set consisting of the Fréchet-derivative of $\mathfrak{a}(\cdot\,,\,\cdot) + \langle \ell\,,\,\cdot\rangle$ and hence, (A.5) is equivalent to

$$\begin{aligned} I^*_{\mathcal{K}}(\boldsymbol{Q}) \geq I^*_{\mathcal{K}}(\boldsymbol{P}) - \mathfrak{a}\big((\boldsymbol{u},\boldsymbol{P}),(\boldsymbol{v},\boldsymbol{Q}) - (\boldsymbol{u},\boldsymbol{P})\big) - \langle \ell, \boldsymbol{v} - \boldsymbol{u}\rangle& \\ \text{for all } (\boldsymbol{v},\boldsymbol{Q}) = (\boldsymbol{v},\boldsymbol{q},\boldsymbol{\eta}) \in V \times S \times M.& \end{aligned} \tag{A.6}$$

Next, we introduce a mapping $\Sigma : V \times S \times M \to S \times M$ by

$$\Sigma(\boldsymbol{u}, \boldsymbol{P}) = (\sigma(\boldsymbol{u}, \boldsymbol{P}), \chi(\boldsymbol{u}, \boldsymbol{P})) := (\mathbb{C} : (\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{p}), -\mathbb{H} : \boldsymbol{\xi}). \tag{A.7}$$

With this setting, (A.6) implies

$$\int_\Omega \Sigma(\boldsymbol{u}, \boldsymbol{P}) : (\boldsymbol{Q} - \boldsymbol{P})\, dx + \sup_{\boldsymbol{\Upsilon} \in \mathcal{K}} \int_\Omega \boldsymbol{P} : \boldsymbol{\Upsilon}\, dx - \sup_{\boldsymbol{\Upsilon} \in \mathcal{K}} \int_\Omega \boldsymbol{Q} : \boldsymbol{\Upsilon}\, dx \le 0 \qquad \forall \boldsymbol{Q} \in S \times M. \tag{A.8}$$

If we choose $\boldsymbol{Q} = \boldsymbol{0}$ and $\boldsymbol{Q} = 2\,\boldsymbol{P}$ in (A.8), we obtain

$$\sup_{\boldsymbol{\Upsilon} \in \mathcal{K}} \int_\Omega \boldsymbol{P} : \boldsymbol{\Upsilon}\, dx - \int_\Omega \boldsymbol{P} : \Sigma(\boldsymbol{u}, \boldsymbol{P}) dx = 0. \tag{A.9}$$

The closed and convex set $\mathcal{K}$ is equal to the intersection of all closed half-spaces containing it, i.e.,

$$\mathcal{K} = \bigcap_{\boldsymbol{Q} \in S \times M} \left\{ \boldsymbol{\Sigma} \in S \times M \ : \ \int_\Omega \boldsymbol{Q} : \boldsymbol{\Sigma}\, dx \le \sup_{\boldsymbol{\Upsilon} \in \mathcal{K}} \int_\Omega \boldsymbol{Q} : \boldsymbol{\Upsilon}\, dx \right\} \tag{A.10}$$

Now, (A.8) and (A.9) imply for all $\boldsymbol{Q} \in S \times M$

$$\int_\Omega \boldsymbol{Q} : \Sigma(\boldsymbol{u}, \boldsymbol{P})\, dx \le \sup_{\boldsymbol{\Upsilon} \in \mathcal{K}} \int_\Omega \boldsymbol{Q} : \boldsymbol{\Upsilon}\, dx,$$

so that (A.10) gives $\Sigma(\boldsymbol{u}, \boldsymbol{P}) \in \mathcal{K}$. Moreover, (A.9) implies $\int_\Omega \boldsymbol{P} : \Sigma(\boldsymbol{u}, \boldsymbol{P})\, dx \ge \int_\Omega \boldsymbol{P} : \boldsymbol{\Upsilon}\, dx$ for all $\boldsymbol{\Upsilon} \in \mathcal{K}$ and, in view of (A.7), this is equivalent to

$$a(\Sigma(\boldsymbol{u}, \boldsymbol{P}), \boldsymbol{\Upsilon} - \Sigma(\boldsymbol{u}, \boldsymbol{P})) + b(\boldsymbol{\varsigma} - \sigma(\boldsymbol{u}, \boldsymbol{P}), \boldsymbol{u}) \ge 0 \quad \forall \boldsymbol{\Upsilon} = (\boldsymbol{\varsigma}, \boldsymbol{\psi}) \in \mathcal{K}$$

thanks to the definition of a in (2.9). Hence, it suffices to search the optimum of $(\mathbf{LP}')$ on the set given by

$$\begin{aligned} \mathcal{P}_{ad} = \{ \boldsymbol{v} \in V, \boldsymbol{Q} \in S \times M \ : \Sigma(\boldsymbol{v}, \boldsymbol{Q}) \in \mathcal{K} \\ \text{and } a(\Sigma(\boldsymbol{v}, \boldsymbol{Q}), \boldsymbol{\Upsilon} - \Sigma(\boldsymbol{v}, \boldsymbol{Q})) + b(\boldsymbol{\varsigma} - \sigma(\boldsymbol{v}, \boldsymbol{Q}), \boldsymbol{v}) \ge 0 \quad \forall \boldsymbol{\Upsilon} \in \mathcal{K} \}. \end{aligned}$$

The definitions of a, b, and Σ imply for the objective of $(\mathbf{LP}')$

$$\begin{aligned} &\frac{1}{2}\, \mathfrak{a}\big((\boldsymbol{v}, \boldsymbol{Q}), (\boldsymbol{v}, \boldsymbol{Q})\big) + \langle \ell\,,\, \boldsymbol{v} \rangle + I^*_{\mathcal{K}}(\boldsymbol{Q}) \\ &\quad = \frac{1}{2} \int_\Omega \big[(\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{q}) : \mathbb{C} : (\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{q}) + \boldsymbol{\eta} : \mathbb{H} : \boldsymbol{\eta}\big] dx + \langle \ell\,,\, \boldsymbol{v} \rangle + \sup_{\boldsymbol{\Upsilon} \in \mathcal{K}} \int_\Omega \boldsymbol{Q} : \boldsymbol{\Upsilon}\, dx \\ &\quad = \frac{1}{2}\, a\big(\Sigma(\boldsymbol{v}, \boldsymbol{Q}), \Sigma(\boldsymbol{v}, \boldsymbol{Q})\big) + \langle \ell\,,\, \boldsymbol{v} \rangle - \inf_{\boldsymbol{\Upsilon} \in \mathcal{K}} \big[a\big(\Sigma(\boldsymbol{v}, \boldsymbol{Q}), \boldsymbol{\Upsilon}\big) + b(\boldsymbol{\varsigma}, \boldsymbol{v})\big]. \end{aligned}$$

Hence, one obtains for all $(\boldsymbol{v}, \boldsymbol{Q}) \in \mathcal{P}_{ad}$

$$\begin{aligned} &\frac{1}{2}\, \mathfrak{a}\big((\boldsymbol{v}, \boldsymbol{Q}), (\boldsymbol{v}, \boldsymbol{Q})\big) + \langle \ell\,,\, \boldsymbol{v} \rangle + I^*_{\mathcal{K}}(\boldsymbol{Q}) \\ &\quad = -\frac{1}{2}\, a\big(\Sigma(\boldsymbol{v}, \boldsymbol{Q}), \Sigma(\boldsymbol{v}, \boldsymbol{Q})\big) - b(\sigma(\boldsymbol{v}, \boldsymbol{Q}), \boldsymbol{v}) + \langle \ell\,,\, \boldsymbol{v} \rangle = -\mathcal{L}(\Sigma(\boldsymbol{v}, \boldsymbol{Q}), \boldsymbol{v}), \end{aligned}$$

where $\mathcal{L}$ is the Lagrangian defined in (A.1). Thus we have shown that $(\boldsymbol{u}, \boldsymbol{P})$ solves

$$\begin{cases} \sup\limits_{\boldsymbol{v},\boldsymbol{Q}} & \mathcal{L}(\Sigma(\boldsymbol{v},\boldsymbol{Q}),\boldsymbol{v}) \\ \text{s.t.} & \boldsymbol{v} \in V,\ \Sigma(\boldsymbol{v},\boldsymbol{Q}) \in \mathcal{K} \\ \text{and} & a(\Sigma(\boldsymbol{v},\boldsymbol{Q}), \boldsymbol{\Upsilon} - \Sigma(\boldsymbol{v},\boldsymbol{Q})) + b(\boldsymbol{\varsigma} - \sigma(\boldsymbol{v},\boldsymbol{Q}),\boldsymbol{v}) \geq 0 \quad \forall\, \boldsymbol{\Upsilon} \in \mathcal{K}. \end{cases}$$

Since $\Sigma(\boldsymbol{v}, \cdot\,) : S \times M \to S \times M$ is surjective for every $\boldsymbol{v} \in V$ in view of Assumption 2.4 for $\mathbb{C}$ and $\mathbb{H}$, $(\boldsymbol{u}, \Sigma(\boldsymbol{u},\boldsymbol{P}))$ solves $(\mathbf{LP})$ according to (A.4), and hence coincides with its unique solution $(\boldsymbol{u}, \boldsymbol{\Sigma})$. The definition of Σ finally gives (A.3). □

As already mentioned above, $(\mathbf{LP}')$ is a strictly convex problem such that the variational inequality of the second kind (A.6) admits the optimum of $(\mathbf{LP}')$ as its unique solution. For the sake of completeness, we convert (A.6) into a form which is found elsewhere in the literature (cf. for instance [Han and Reddy, 1999, Section 7] for the quasi-static counterpart of (A.6)). According to [Temam, 1983, Prop. I.2.5], there holds

$$\sup_{\boldsymbol{T} \in \mathcal{K}} \int_\Omega \boldsymbol{T}(x) : \boldsymbol{P}(x)\, dx = \int_\Omega \sup_{\boldsymbol{T} \in K} (\boldsymbol{T} : \boldsymbol{P}(x))\, dx.$$

Note that this implies in particular that $\sup_{\boldsymbol{T} \in K} (\boldsymbol{T} : \boldsymbol{P}(\,\cdot\,))$ is measureable if $I^*_{\mathcal{K}}(\boldsymbol{P}) < \infty$. With this result at hand, we may rewrite (A.6) by

$$\begin{aligned} \mathfrak{a}\big((\boldsymbol{u},\boldsymbol{P}),(\boldsymbol{v},\boldsymbol{Q}) - (\boldsymbol{u},\boldsymbol{P})\big) + j(\boldsymbol{Q}) - j(\boldsymbol{P}) &\geq \langle -\ell, \boldsymbol{v} - \boldsymbol{u} \rangle \\ \text{for all } (\boldsymbol{v},\boldsymbol{Q}) &= (\boldsymbol{v},\boldsymbol{q},\boldsymbol{\eta}) \in V \times S \times M, \end{aligned} \tag{A.11}$$

where $j : S \times M \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$j(\boldsymbol{P}) := \int_\Omega \sup_{\boldsymbol{T} \in K} (\boldsymbol{T} : \boldsymbol{P}(x))\, dx.$$

Remark A.3. *It is to be noted that* (A.11) *represents a variational inequality of the second kind. There are several contributions concerning the theory of optimal control problems governed by variational inequalities of the second kind. We only refer to Barbu [1984], Bonnans and Tiba [1991], Bonnans and Casas [1995], and Bergounioux [1998]. Nevertheless, since the stress field is a physically important quantity in various applications, we focus on the dual problem of infinitesimal elastoplasticity in form of* $(\mathbf{L})$ *and* (2.11)*, respectively, which explicitly contains the stress field instead of the plastic strain.*

B Proof of Lemma 4.1

The projection $P_C x$ of an element $x \in H$ is uniquely characterized by $P_C\, x \in C$ and the variational inequality

$$(x - P_C\, x, y - P_C\, x) \leq 0 \quad \text{for all } y \in C. \tag{B.1}$$

As a consequence, the projection is non-expansive:

$$\|P_C\, x - P_C\, y\| \leq \|x - y\| \quad \text{for all } x, y \in H.$$

Now we address the differentiability of $F(x) = \frac{1}{2}\|x - P_C\, x\|^2$. We observe

$$\begin{aligned}
&\|x + \delta x - P_C(x + \delta x)\|^2 \\
&\qquad = \|(x - P_C x) + \delta x + P_C x - P_C(x + \delta x)\|^2 \\
&\qquad = \|x - P_C x\|^2 + 2\,(x - P_C x, \delta x) + 2\,(x - P_C x, P_C x - P_C(x + \delta x)) \\
&\qquad\quad + \|P_C x - P_C(x + \delta x)\|^2.
\end{aligned}$$

It follows that

$$\begin{aligned}
&F(x + \delta x) - F(x) - (x - P_C x, \delta x) \\
&\qquad = (x - P_C x, P_C x - P_C(x + \delta x)) + \frac{1}{2}\|P_C x - P_C(x + \delta x)\|^2 \ge 0
\end{aligned}$$

by (B.1) since $P_C(x + \delta x) \in C$. On the other hand, we can estimate

$$\begin{aligned}
&F(x + \delta x) - F(x) - (x - P_C\, x, \delta x) \\
&\qquad = (x - P_C(x + \delta x), P_C\, x - P_C(x + \delta x)) - \frac{1}{2}\|P_C\, x - P_C(x + \delta x)\|^2 \\
&\qquad \le (x + \delta x - P_C(x + \delta x), P_C\, x - P_C(x + \delta x)) - (\delta x, P_C\, x - P_C(x + \delta x)) \\
&\qquad \le 0 + \|\delta x\| \|P_C\, x - P_C(x + \delta x)\| \\
&\qquad \le \|\delta x\|^2
\end{aligned}$$

where we used (B.1), the Cauchy-Schwarz inequality and the non-expansiveness of the projection. We conclude that

$$|F(x + \delta x) - F(x) - (x - P_C\, x, \delta x)| \le \|\delta x\|^2$$

holds, which confirms Fréchet differentiability of F with derivative $F'(x) = x - P_C\, x$. The monotonicity of F' follows from the estimate

$$\begin{aligned}
(F'(x) - F'(y), x - y) &= \|x - y\|^2 - (P_C\, x - P_C\, y, x - y) \\
&\ge \|x - y\|^2 - \|P_C\, x - P_C\, y\| \|x - y\| \ge 0,
\end{aligned}$$

where the last inequality is due to the non-expansiveness of the projection. As a consequence, F is a convex function, which completes the proof of parts (a) and (b). To prove part (c), let $x \in H$ and $y \in C$. Then $F'(y) = 0$ and the monotonicity of F' imply

$$(F'(x), x - y) = (F'(x) - F'(y), x - y) \ge 0.$$

For part (d), let $x \in H$ be arbitrary and suppose that $a + C = C$ holds. Then we have

$$\begin{aligned}
&(x + a - P_C(x + a), z - P_C(x + a)) \le 0 \quad \text{for all } z \in C \qquad \text{by (B.1)} \\
\Rightarrow\quad &(x - (P_C(x + a) - a), y - (P_C(x + a) - a)) \le 0
\end{aligned}$$

for all $y \in H$ such that $y + a = z$ with some $z \in C$, i.e., for all $y \in C$. This inequality together with the fact $P_C(x+a)-a \in C-a = C$ confirms that $P_C\, x = P_C(x+a)-a$ holds. As a consequence, we obtain from part (a) that $F'(x) = x - P_C\, x = x-(P_C(x+a)-a) = F'(x + a)$ as claimed.

References

V. Barbu. *Optimal Control of Variational Inequalities*, volume 100 of *Research Notes in Mathematics*. Pitman, Boston, 1984.

M. Bergounioux. Optimal control problems governed by abstract elliptic variational inequalities with state constraints. *SIAM Journal on Control and Optimization*, 36(1): 273–289, 1998.

M. Bergounioux and H. Zidani. Pontryagin maximum principle for optimal control of variational inequalities. *SIAM Journal on Control and Optimization*, 37(4):1273–1290, 1999.

F. Bonnans and E. Casas. An extension of Pontryagin's principle for state-constrained optimal control of semilinear elliptic equations and variational inequalities. *SIAM Journal on Control and Optimization*, 33(1):274–298, 1995.

F. Bonnans and D. Tiba. Pontryagin's principle in the control of semilinear elliptic variational inequalities. *Applied Mathematics and Optimization*, 23:299–312, 1991.

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.

C. Carstensen. Numerical analysis of the primal problem of elastoplasticity with hardening. *Numerische Mathematik*, 82(4):577–597, 1999.

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.

W. Han and B. D. Reddy. *Plasticity*. Springer, New York, 1999.

M. Hintermüller. Inverse coefficient problems for variational inequalities: Optimality conditions and numerical realization. *ESAIM Mathematical Modelling and Numerical Analysis*, 35(1):129–152, 2001.

M. Hintermüller. An active-set equality constrained Newton solver with feasibility restoration for inverse coefficient problems in elliptic variational inequalities. *Inverse Problems*, 24(3):034017, 23, 2008.

K. Ito and K. Kunisch. Optimal control of elliptic variational inequalities. *Applied Mathematics and Optimization*, 41:343–364, 2000.

D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York, 1980.

F. Mignot. Contrôle dans les inéquations variationelles èlliptiques. *Journal of Functional Analysis*, 22(2):130–185, 1976.

F. Mignot and J.-P. Puel. Optimal control in some variational inequalities. *SIAM Journal on Control and Optimization*, 22(3):466–476, 1984.

J. C. Simo and T. J. R. Hughes. *Computational Inelasticity*. Springer, New York, 1998.

R. Temam. *Mathematical Problems in Plasticity*. Gauthier-Villars, Montrouge, 1983.

J. Zowe and S. Kurcyusz. Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics and Optimization*, 5(1):49–62, 1979.