Deutsche Forschungsgemeinschaft

Priority Program 1253

Optimization with Partial Differential Equations

R. Herzog, C. Meyer and G. Wachsmuth

B- and Strong Stationarity for Optimal Control of Static Plasticity with Hardening

January 2011

Preprint-Number SPP1253-114

---

http://www.am.uni-erlangen.de/home/spp1253

B- AND STRONG STATIONARITY FOR OPTIMAL CONTROL OF STATIC PLASTICITY WITH HARDENING

ROLAND HERZOG, CHRISTIAN MEYER, AND GERD WACHSMUTH

ABSTRACT. Optimal control problems for the variational inequality of static elastoplasticity with linear kinematic hardening are considered. The control-to-state map is shown to be weakly directionally differentiable, and local optimal controls are proved to verify an optimality system of B-stationary type. For a modified problem, local minimizers are shown to even satisfy an optimality system of strongly stationary type.

1 Introduction

In this paper we continue the investigation of first-order necessary optimality conditions for optimal control problems in static elastoplasticity. The forward system in the stress-based (so-called dual) form is represented by a variational inequality (VI) of mixed type: find generalized stresses $\boldsymbol{\Sigma} \in S^2$ and displacements $\boldsymbol{u} \in V$ which satisfy $\boldsymbol{\Sigma} \in \mathcal{K}$ and

$$\left.\begin{aligned} \langle A\boldsymbol{\Sigma}, \boldsymbol{T} - \boldsymbol{\Sigma}\rangle + \langle B^\star \boldsymbol{u}, \boldsymbol{T} - \boldsymbol{\Sigma}\rangle \geq 0 \quad &\text{for all } \boldsymbol{T} \in \mathcal{K} \\ B\boldsymbol{\Sigma} = \ell \quad &\text{in } V'. \end{aligned}\right\} \qquad \textbf{(VI)}$$

where A and B are linear operators. The closed, convex set $\mathcal{K} \subset S^2$ of admissible stresses is determined by the von Mises yield condition.

The optimization of elastoplastic systems is of significant importance for industrial deformation processes. We emphasize that, in spite of its limited physical importance itself, the static problem **(VI)** appears as a time step of its quasi-static variant, which will be investigated elsewhere. We will consider primarily the following prototypical optimal control problem

$$\left.\begin{aligned} \text{Minimize} \quad & J(\boldsymbol{u}, \boldsymbol{g}) := \frac{1}{2}\|\boldsymbol{u} - \boldsymbol{u}_d\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu}{2}\|\boldsymbol{g}\|^2_{L^2(\Gamma_N;\mathbb{R}^d)} \\ \text{s.t.} \quad & \text{the plasticity problem } \textbf{(VI)} \text{ with } \ell \in V' \text{ defined by} \\ & \langle \ell, \boldsymbol{v}\rangle = -\int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v} \, \mathrm{d}s, \quad \boldsymbol{v} \in V \\ \text{and} \quad & \boldsymbol{g} \in U_{\text{ad}} \end{aligned}\right\} \qquad \textbf{(P)}$$

in which the boundary loads $\boldsymbol{g}$ appear as control variables. The details are made precise below.

The optimal control of **(VI)** leads to an infinite dimensional MPEC (mathematical program with equilibrium constraints). The derivation of necessary optimality conditions is challenging due to the lack of Fréchet differentiability of the associated control-to-state map $\ell \mapsto (\boldsymbol{\Sigma}, \boldsymbol{u})$. The same is true for the re-formulation of **(VI)** as a complementarity system involving the so-called plastic multiplier. It is well known that for the resulting MPCC (mathematical program with complementarity constraints) classical constraint qualifications fail to hold.

Date: January 14, 2011.

To overcome these difficulties, several competing stationarity concepts for MPCCs have been developed, see for instance Scheel and Scholtes [2000] for an overview in the finite dimensional case. It was shown recently in Herzog et al. [2010a] that local optima of (**P**) satisfy first-order optimality conditions of C-(Clarke)-stationary type. This was achieved by approximating them by sequences of solutions to regularized problems. In these regularized problems, (**VI**) is replaced by a smooth equation. We exemplarily refer to Barbu [1984], Hintermüller [2001] for related results for optimal control of the obstacle problem.

In the present paper, we pursue a different approach. We prove that local optima of (**P**) satisfy first-order optimality conditions of B-(Bouligand)-stationary type. This optimality concept is based solely on primal variables, and the main step in the proof is to establish the weak directional differentiability of the control-to-state map. In order to achieve this, we make use of a recent regularity result Herzog et al. [2010b] for nonlinear elasticity systems, which applies to (**VI**) as shown in Section 2.2.

Moreover, we show that for a modified problem, local optima are even strongly stationary. In order to obtain this result, we suppose that the modified problem has so-called "ample" controls, i.e., distributed control functions which act on both right hand sides of (**VI**). In addition, we dispose of control constraints in the modified problem. These modifications are in accordance with previous strong stationarity results for the optimal control of the obstacle problem, see Mignot and Puel [1984]. However, we present a different and more elementary technique of proof.

Let us put our work into perspective. In contrast to the multitude of papers concerning regularization and C-stationarity conditions for infinite dimensional MPECs, there are fewer contributions which address the question of stricter optimality conditions for optimal control problems governed by variational inequalities. We refer to the classical paper of Mignot and Puel [1984], where the obstacle problem is discussed. In Hintermüller and Kopacka [2008] conditions are derived which guarantee the convergence of stationary points of regularized problems to strongly stationary points in case of optimal control of the obstacle problem. It is to be noted that these conditions depend on the regularized sequence itself and cannot be guaranteed a priori. Recently, Jarušek et al. [2010] confirmed the strong stationarity result of Mignot and Puel [1984] for the obstacle problem by using a completely different technique based on results of Jarušek and Outrata [2007]. To the authors' knowledge, B- and strong stationarity results for optimal control problems governed by variational inequalities other than of obstacle type (as for instance (**VI**)) have not been discussed so far.

The paper is organized as follows. In the remainder of this section we introduce the notation and state our generic assumptions. In Section 2, we review some results about the forward problem (**VI**). Moreover, we establish a regularity result concerning the solution of the static elastoplasticity system (**VI**) which is essential for the B-stationarity analysis in Section 3. Section 4 is devoted to the investigation of strong stationarity for the modified problem.

Notation and Assumptions. Our notation for the forward problem follows Han and Reddy [1999] and Herzog and Meyer [2009]. We restrict the discussion to the case of linear kinematic hardening.

Function Spaces. Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with boundary Γ in dimension $d \in \{2, 3\}$. We point out that the presented analysis is not restricted to the case $d \le 3$, but for reasons of physical interpretation we focus on the two and three dimensional case. The boundary consists of two disjoint parts Γ_N and Γ_D, on which boundary loads and zero displacement conditions are imposed, respectively. We denote by $\mathbb{S} := \mathbb{R}^{d\times d}_{\text{sym}}$ the space of symmetric d-by-d matrices, endowed with the inner product $\boldsymbol{A} : \boldsymbol{B} = \sum_{i,j=1}^d A_{ij} B_{ij}$, and we define

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

as the spaces for the displacement $\boldsymbol{u}$, stress $\boldsymbol{\sigma}$, and back stress $\boldsymbol{\chi}$, respectively. We refer to $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}, \boldsymbol{\chi}) \in S^2$ as the generalized stress. The boundary control $\boldsymbol{g}$ belongs to the space $L^2(\Gamma_N;\mathbb{R}^d)$.

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

$$\phi(\boldsymbol{\Sigma}) = \big(|\boldsymbol{\sigma}^D + \boldsymbol{\chi}^D|^2 - \tilde{\sigma}_0^2\big)/2 \tag{1.1}$$

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

$$\boldsymbol{\sigma}^D = \boldsymbol{\sigma} - (1/d)\,(\operatorname{trace}\boldsymbol{\sigma})\,\boldsymbol{I}$$

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

$$\mathcal{K} = \{\boldsymbol{\Sigma} \in S^2 : \phi(\boldsymbol{\Sigma}) \le 0 \quad \text{a.e. in } \Omega\}. \tag{1.2}$$

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

$$\mathcal{D}\boldsymbol{\Sigma} = \boldsymbol{\sigma}^D + \boldsymbol{\chi}^D \quad \text{and} \quad \mathcal{D}^\star\boldsymbol{\sigma} = \begin{pmatrix} \boldsymbol{\sigma}^D \\ \boldsymbol{\sigma}^D \end{pmatrix}$$

for matrices $\boldsymbol{\Sigma} \in \mathbb{S}^2$ as well as for functions $\boldsymbol{\Sigma} \in S^2$. When considered as an operator in function space, $\mathcal{D}$ maps $S^2 \to S$. For later reference, we also remark that

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

holds.

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

$$\langle \boldsymbol{T},\, A\boldsymbol{\Sigma}\rangle = \int_\Omega \boldsymbol{\tau} : \mathbb{C}^{-1}\boldsymbol{\sigma}\,\mathrm{d}x + \int_\Omega \boldsymbol{\mu} : \mathbb{H}^{-1}\boldsymbol{\chi}\,\mathrm{d}x. \tag{1.3}$$

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

$$\langle B\boldsymbol{\Sigma},\, \boldsymbol{v}\rangle = -\int_\Omega \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\boldsymbol{v})\,\mathrm{d}x. \tag{1.4}$$

We recall that $\boldsymbol{\varepsilon}(\boldsymbol{v}) = \frac{1}{2}\big(\nabla\boldsymbol{v} + (\nabla\boldsymbol{v})^\top\big)$ denotes the (linearized) strain tensor.

Here and throughout, $\langle\cdot,\,\cdot\rangle$ denotes the dual pairing between V and its dual V', or the scalar products in S or S^2, respectively. Moreover, $(\cdot,\cdot)_E$ refers to the scalar product of $L^2(E)$ where $E \subset \Omega$ or $E = \Gamma_N$.

Assumption 1.1. (1) *The domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$ is a bounded Lipschitz domain in the sense of [Grisvard, 1985, Chapter 1.2]. The boundary of Ω, denoted by Γ, consists of two disjoint measurable parts Γ_N and Γ_D such that $\Gamma = \Gamma_N \cup \Gamma_D$. While Γ_N is a relatively open subset, Γ_D is a relatively closed subset of Γ. Furthermore Γ_D is assumed to have positive measure.*

(2) *The yield stress $\tilde{\sigma}_0$ is assumed to be a positive constant. It equals $\sqrt{2/3}\,\sigma_0$, where σ_0 is the initial uni-axial yield stress.*

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

(4) *The desired displacement $\boldsymbol{u}_d$ is an element of $L^2(\Omega;\mathbb{R}^d)$. Moreover, ν is a positive constant and $U_{\mathrm{ad}} \subset L^2(\Gamma_N;\mathbb{R}^d)$ is a nonempty, closed, and convex set.*

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

$$\|\boldsymbol{u}\|^2_{H^1(\Omega;\mathbb{R}^d)} \le c_K \left(\|\boldsymbol{u}\|^2_{L^2(\Gamma_D;\mathbb{R}^d)} + \|\boldsymbol{\varepsilon}(\boldsymbol{u})\|^2_S\right) \tag{1.5}$$

for all $\boldsymbol{u} \in H^1(\Omega;\mathbb{R}^d)$. Note that (1.5) entails in particular that $\|\boldsymbol{\varepsilon}(\boldsymbol{u})\|_S$ is a norm on $H^1_D(\Omega;\mathbb{R}^d)$ equivalent to the $H^1(\Omega;\mathbb{R}^d)$ norm.

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

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

with Lamé constants μ and λ, provided that $\mu > 0$ and $d\,\lambda + 2\,\mu > 0$ hold. These constants appear only here and there is no risk of confusion with the plastic multiplier λ or the Lagrange multiplier μ, which are introduced in Section 2.1 and Section 4, respectively. A common example for the hardening modulus is given by $\mathbb{H}^{-1}\boldsymbol{\chi} = \boldsymbol{\chi}/k_1$ with hardening constant $k_1 > 0$, see [Han and Reddy, 1999, Section 3.4]. Assumption (3) implies that $\langle A\boldsymbol{\Sigma},\, \boldsymbol{\Sigma}\rangle \ge \underline{\alpha}\,\|\boldsymbol{\Sigma}\|^2_{S^2}$ for some $\underline{\alpha} > 0$, i.e. A is a coercive operator.

As an example for U_{ad}, we mention the set of boundary stresses with modulus bounded by $\rho > 0$, i.e.,

$$U_{\mathrm{ad}} = \{\boldsymbol{g} \in L^2(\Gamma_N;\mathbb{R}^d) \,:\, |\boldsymbol{g}(x)|_{\mathbb{R}^d} \le \rho \text{ a.e. on } \Gamma_N\}.$$

2 Results Concerning the Forward Problem

2.1. Known Results. In this section we collect some results concerning **(VI)**. Given $\ell \in V'$, the existence and uniqueness of a solution to **(VI)** is well known, see for instance [Han and Reddy, 1999, Section 8.1]. As a consequence, we may introduce the control-to-state map

$$G : V' \to S^2 \times V, \quad \text{mapping } \ell \mapsto (G^{\boldsymbol{\Sigma}}, G^{\boldsymbol{u}})(\ell) = (\boldsymbol{\Sigma}, \boldsymbol{u}).$$

The following result can be found in Herzog and Meyer [2009].

Theorem 2.1. *The solution operator $G : V' \to S^2 \times V$ is Lipschitz continuous, i.e.*

$$\|G(\ell_1) - G(\ell_2)\|_{S^2 \times V} = \|\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2\|_{S^2} + \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_V \le L\,\|\ell_1 - \ell_2\|_{V'}$$

holds with a Lipschitz constant $L > 0$.

In our subsequent analysis we will frequently make use of an equivalent formulation of (VI) which involves a Lagrange multiplier for the yield condition, termed the plastic multiplier. We refer to Herzog et al. [2010a] and Herzog et al. [2010c] for the following result.

Theorem 2.2. *Let $\ell \in V'$ be given. The pair $(\boldsymbol{\Sigma}, \boldsymbol{u}) \in S^2 \times V$ is the unique solution of (VI) if and only if there exists a plastic multiplier $\lambda \in L^2(\Omega)$ such that*

$$A\boldsymbol{\Sigma} + B^\star \boldsymbol{u} + \lambda\, \mathcal{D}^\star \mathcal{D} \boldsymbol{\Sigma} = \mathbf{0} \quad \text{in } S^2, \tag{2.1a}$$

$$B\boldsymbol{\Sigma} = \ell \quad \text{in } V', \tag{2.1b}$$

$$0 \le \lambda(x) \;\perp\; \phi(\boldsymbol{\Sigma}(x)) \le 0 \quad \text{a.e. in } \Omega \tag{2.1c}$$

holds. Moreover, λ is unique.

Note that $\lambda(x) \perp \phi(\boldsymbol{\Sigma}(x))$ is a shorthand notation for $\lambda(x)\, \phi(\boldsymbol{\Sigma}(x)) = 0$.

The complementarity condition (2.1c) gives rise to the following definition of subsets of Ω:

$$\mathcal{A}(\ell) := \{x \in \Omega : \phi(\boldsymbol{\Sigma}(x)) = 0\}, \qquad \text{(active set)} \tag{2.2a}$$

$$\mathcal{A}_s(\ell) := \{x \in \Omega : \lambda(x) > 0\}, \qquad \text{(strongly active set)} \tag{2.2b}$$

$$\mathcal{B}(\ell) := \{x \in \Omega : \phi(\boldsymbol{\Sigma}(x)) = \lambda(x) = 0\}, \qquad \text{(biactive set)} \tag{2.2c}$$

$$\mathcal{I}(\ell) := \{x \in \Omega : \phi(\boldsymbol{\Sigma}(x)) < 0\}, \qquad \text{(inactive set)} \tag{2.2d}$$

where $\boldsymbol{\Sigma}$ and λ are given by the solution of (2.1). The notation for the sets in (2.2) is driven by the point of view that (VI) and equivalently (2.1) are the necessary and sufficient optimality conditions for the lower-level problem

$$\text{Minimize} \quad \tfrac{1}{2}\langle A\boldsymbol{\Sigma},\, \boldsymbol{\Sigma}\rangle \quad \text{s.t.} \quad B\boldsymbol{\Sigma} = \ell \quad \text{and} \quad \phi(\boldsymbol{\Sigma}) \le 0, \tag{2.3}$$

in which $\phi(\boldsymbol{\Sigma}) \le 0$ appears as a constraint with Lagrange multiplier λ. We remark that $\mathcal{A}_s(\ell)$, $\mathcal{B}(\ell)$, and $\mathcal{I}(\ell)$ are pairwise disjoint sets. Furthermore $\mathcal{A}(\ell) = \mathcal{A}_s(\ell) \cup \mathcal{B}(\ell)$ and $\Omega = \mathcal{A}(\ell) \cup \mathcal{I}(\ell)$.

Remark 2.3. *The component-wise evaluation of (2.1a) yields*

$$\mathbb{C}^{-1}\boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda\, \mathcal{D}\boldsymbol{\Sigma} = \mathbf{0}, \tag{2.4a}$$

$$\mathbb{H}^{-1}\boldsymbol{\chi} + \lambda\, \mathcal{D}\boldsymbol{\Sigma} = \mathbf{0}. \tag{2.4b}$$

Combining both equations, we find

$$\mathbb{C}^{-1}\boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\boldsymbol{u}) - \mathbb{H}^{-1}\boldsymbol{\chi} = \mathbf{0}. \tag{2.5}$$

2.2. Higher integrability of solutions. In this section we show that under a mildly strengthened assumption on the domain Ω, the unique solution $(\boldsymbol{\Sigma}, \boldsymbol{u})$ of (VI) belongs to $L^p(\Omega; \mathbb{S}^2) \times W^{1,p}_D(\Omega; \mathbb{R}^d)$ for some $p > 2$, provided that $\ell \in W^{-1,p}_D(\Omega; \mathbb{R}^d) := W^{1,p'}_D(\Omega; \mathbb{R}^d)'$, where $p' = p/(p-1)$ is the conjugate exponent to p.

Assumption 2.4. *In addition to Assumption 1.1 (1), suppose that the set $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, cf. [Gröger, 1989, Definition 2].*

A characterization of regular domains for the case $d \in \{2, 3\}$ can be found in [Haller-Dintelmann et al., 2009, Section 5]. This class covers a wide range of geometries, but it excludes, for instance, domains in which the Dirichlet part Γ_D of the boundary has isolated points.

In order to obtain the higher integrability result, we state a reformulation of (VI) as a quasilinear equation in the displacement $\boldsymbol{u}$ with a nonsmooth quasilinearity and apply a recent regularity result from Herzog et al. [2010b].

We begin by deriving the nonsmooth reformulation. To this end, we test (VI) with $\boldsymbol{T} = (\boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{K}$ with $\boldsymbol{\tau} = \boldsymbol{\sigma}$. This yields

$$\int_\Omega (\mathbb{H}^{-1}\boldsymbol{\chi}) : (\boldsymbol{\mu} - \boldsymbol{\chi}) \, \mathrm{d}x \geq 0 \quad \text{for all } \boldsymbol{\mu} \in S \text{ such that } \boldsymbol{\sigma} + \boldsymbol{\mu} \in \bar{\mathcal{K}}, \tag{2.6}$$

where $\bar{\mathcal{K}} = \{\boldsymbol{\tau} \in S : (\boldsymbol{\tau}, \boldsymbol{0}) \in \mathcal{K}\}$. The structure of ϕ implies, that $\boldsymbol{\sigma} + \boldsymbol{\mu} \in \bar{\mathcal{K}}$ is equivalent to $(\boldsymbol{\sigma}, \boldsymbol{\mu}) \in \mathcal{K}$, see (1.1). The condition $\boldsymbol{\sigma} + \boldsymbol{\mu} \in \bar{\mathcal{K}}$ is equivalent to $\boldsymbol{\mu} \in \bar{\mathcal{K}} - \boldsymbol{\sigma}$ and hence (2.6) implies

$$\boldsymbol{\chi} = \operatorname{Proj}^{\mathbb{H}^{-1}}_{\bar{\mathcal{K}}-\boldsymbol{\sigma}}(\boldsymbol{0}) = \operatorname{Proj}^{\mathbb{H}^{-1}}_{\bar{\mathcal{K}}} \boldsymbol{\sigma} - \boldsymbol{\sigma}. \tag{2.7}$$

Here and below, $\operatorname{Proj}^{\mathbb{H}^{-1}}_{\bar{\mathcal{K}}}$ denotes the pointwise projection on the set $\bar{\mathcal{K}}$ with respect to the norm induced by $\mathbb{H}^{-1}$. Plugging (2.7) into (2.5) yields

$$\mathbb{C}^{-1}\boldsymbol{\sigma} + \mathbb{H}^{-1}(\boldsymbol{\sigma} - \operatorname{Proj}^{\mathbb{H}^{-1}}_{\bar{\mathcal{K}}} \boldsymbol{\sigma}) = \boldsymbol{\varepsilon}(\boldsymbol{u}) \quad \text{a.e. in } \Omega. \tag{2.8}$$

We define $M : \Omega \times \mathbb{S} \to \mathbb{S}$ such that the left hand becomes $M(\cdot, \boldsymbol{\sigma})$, i.e.

$$M : (x, \boldsymbol{\sigma}) \mapsto \mathbb{C}^{-1}(x)\, \boldsymbol{\sigma} + \mathbb{H}^{-1}(x)(\boldsymbol{\sigma} - \operatorname{Proj}^{\mathbb{H}^{-1}(x)}_{\bar{K}} \boldsymbol{\sigma}),$$

where $\bar{K} = \{\boldsymbol{T} \in \mathbb{S} : \phi(\boldsymbol{T}) \leq 0\}$ is the pointwise condition in $\bar{\mathcal{K}} = \{\boldsymbol{T} \in S : \boldsymbol{T}(x) \in \bar{K} \text{ a.e. in } \Omega\}$. Note that the term $\mathbb{H}^{-1}(x)(\boldsymbol{\sigma} - \operatorname{Proj}^{\mathbb{H}^{-1}(x)}_{\bar{K}} \boldsymbol{\sigma})$ is monotone and Lipschitz continuous w.r.t. $\boldsymbol{\sigma}$, see [Herzog and Meyer, 2009, Lemma 4.1(b)]. Moreover, $\mathbb{C}^{-1}(x)$ is uniformly coercive and bounded. Altogether, we infer that $M(x, \cdot)$ is strongly monotone and Lipschitz continuous, i.e.

$$\begin{aligned} \|M(x,\boldsymbol{\sigma}) - M(x,\boldsymbol{\tau})\|_{\mathbb{S}} &\leq \overline{m}\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbb{S}}, \\ \langle M(x,\boldsymbol{\sigma}) - M(x,\boldsymbol{\tau}),\, \boldsymbol{\sigma} - \boldsymbol{\tau}\rangle_{\mathbb{S}} &\geq \underline{m}\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbb{S}}^2 \end{aligned}$$

holds for all $(\boldsymbol{\sigma}, \boldsymbol{\tau}) \in \mathbb{S}^2$ with constants $\overline{m}, \underline{m}$ uniformly in Ω. Due to the Browder-Minty theorem, the pointwise inverse $M^{-1}(x, \cdot)$ w.r.t. the second variable $\boldsymbol{\sigma}$ exists and it is Lipschitz continuous with Lipschitz constant $1/\underline{m}$. The strong monotonicity of $M^{-1}(x, \cdot)$ is confirmed by

$$\begin{aligned} &\langle M^{-1}(x,\boldsymbol{\sigma}) - M^{-1}(x,\boldsymbol{\tau}),\, \boldsymbol{\sigma} - \boldsymbol{\tau}\rangle_{\mathbb{S}} \\ &\quad = \langle M^{-1}(x,\boldsymbol{\sigma}) - M^{-1}(x,\boldsymbol{\tau}),\, M(x, M^{-1}(x,\boldsymbol{\sigma})) - M(x, M^{-1}(x,\boldsymbol{\tau}))\rangle_{\mathbb{S}} \\ &\quad \geq \underline{m}\, \|M^{-1}(x,\boldsymbol{\sigma}) - M^{-1}(x,\boldsymbol{\tau})\|_{\mathbb{S}}^2 \geq \frac{\underline{m}}{\overline{m}^2}\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbb{S}}^2. \end{aligned}$$

To summarize, $M^{-1}(x, \cdot)$ is strongly monotone and Lipschitz continuous, i.e.

$$\|M^{-1}(x,\boldsymbol{\sigma}) - M^{-1}(x,\boldsymbol{\tau})\|_{\mathbb{S}} \leq \frac{1}{\underline{m}}\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbb{S}}, \tag{2.9a}$$

$$\langle M^{-1}(x,\boldsymbol{\sigma}) - M^{-1}(x,\boldsymbol{\tau}),\, \boldsymbol{\sigma} - \boldsymbol{\tau}\rangle_{\mathbb{S}} \geq \frac{\underline{m}}{\overline{m}^2}\, \|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_{\mathbb{S}}^2, \tag{2.9b}$$

with constants independent of $x \in \Omega$.

With M^{-1} at hand, we see that (2.8) is equivalent to $\boldsymbol{\sigma} = M^{-1}(-B^\star \boldsymbol{u})$ and hence $\boldsymbol{u}$ is a solution of

$$B M^{-1}(-B^\star \boldsymbol{u}) = \ell.$$

Let us check that M^{-1} satisfies [Herzog et al., 2010b, Assumption 1.5(2)]:

- $M^{-1}(\cdot, \boldsymbol{0}) = \boldsymbol{0} \in L^\infty(\Omega; \mathbb{S})$.
- $M^{-1}(\cdot, \boldsymbol{\sigma})$ is measurable due to the measurability of $\mathbb{C}$ and $\mathbb{H}$.
- The strong monotonicity and the Lipschitz continuity of M^{-1} was proven in (2.9).

Hence, by [Herzog et al., 2010b, Theorem 1.1], there exists $\overline{p} > 2$ such that $\ell \mapsto \boldsymbol{u}$ is Lipschitz from $W_D^{-1,p}(\Omega;\mathbb{R}^d) \to W_D^{1,p}(\Omega;\mathbb{R}^d)$ for all $p \in [2,\overline{p}]$. Due to $\boldsymbol{\sigma} = M^{-1}(-B^\star \boldsymbol{u})$ and since M^{-1} maps $L^p(\Omega;\mathbb{S})$ Lipschitz continuously into itself by (2.9a), we conclude $\boldsymbol{\sigma} \in L^p(\Omega;\mathbb{S})$. Due to (2.7), $\boldsymbol{\chi} \in L^p(\Omega;\mathbb{S})$ holds, and both $\boldsymbol{\sigma}$ and $\boldsymbol{\chi}$ depend Lipschitz continuously on $\boldsymbol{u}$ and thus on ℓ. We have proved the following

Theorem 2.5. *Suppose that Assumption 2.4 holds. Then there exists $\overline{p} > 2$ such that for all $p \in [2,\overline{p}]$ for any $\ell \in W_D^{-1,p}(\Omega;\mathbb{R}^d)$, the unique solution $(\boldsymbol{u},\boldsymbol{\Sigma})$ of* **(VI)** *belongs to $W_D^{1,p}(\Omega;\mathbb{R}^d) \times L^p(\Omega;\mathbb{S}^2)$. There exists $L > 0$ such that*

$$\|\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2\|_{L^p(\Omega;\mathbb{S}^2)} + \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_{W_D^{1,p}(\Omega;\mathbb{R}^d)} \le L\,\|\ell_1 - \ell_2\|_{W_D^{-1,p}(\Omega;\mathbb{R}^d)}.$$

3 Bouligand Stationarity

In virtue of the control-to-state map G with components $(G^{\boldsymbol{\Sigma}}, G^{\boldsymbol{u}})$, we may reduce problem **(P)** to the control variable $\boldsymbol{g}$

$$\begin{aligned} \text{Minimize} \quad & j(\boldsymbol{g}) := \frac{1}{2}\,\|G^{\boldsymbol{u}}(-\tau_N^\star \boldsymbol{g}) - \boldsymbol{u}_d\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu}{2}\|\boldsymbol{g}\|^2_{L^2(\Gamma_N;\mathbb{R}^d)} \\ \text{s.t.} \quad & \boldsymbol{g} \in U_{\text{ad}}. \end{aligned}$$

The term $-\tau_N^\star \boldsymbol{g}$ denotes the load ℓ induced by the boundary stresses $\boldsymbol{g}$, i.e.,

$$\tau_N^\star : L^2(\Gamma_N;\mathbb{R}^d) \to V', \quad \langle \tau_N^\star \boldsymbol{g},\, \boldsymbol{v}\rangle := \int_{\Gamma_N} \boldsymbol{g}\cdot\boldsymbol{v}\,\mathrm{d}s, \quad \boldsymbol{v} \in V,$$

compare **(P)**. Note that the bounded linear operator $\tau_N^\star$ is the adjoint of the boundary trace map $\tau_N : V \to L^2(\Gamma_N;\mathbb{R}^d)$ onto the Neumann part Γ_N of the boundary.

Remark 3.1. (1) *It can be easily shown that* **(P)** *possesses at least one global optimal solution, see [Herzog and Meyer, 2009, Proposition 3.6]. Notice however that one cannot expect the solution to be unique due to the nonlinearity of G.*
(2) *To keep the presentation concise, we restrict the discussion to the control of boundary stresses only. There would be no difficulty in including additional volume forces as control variables as in Herzog and Meyer [2009].*

The aim of this section is to prove that j is directionally differentiable so that local minimizers $\overline{\boldsymbol{g}}$ necessarily satisfy

$$\delta j(\overline{\boldsymbol{g}}; \boldsymbol{g} - \overline{\boldsymbol{g}}) \ge 0 \quad \text{for all } \boldsymbol{g} \in U_{\text{ad}}. \tag{3.1}$$

Like the Bouligand, or B-stationarity concept for MPECs, this optimality condition involves solely primal variables, cf. [Mignot and Puel, 1984, Lemma 3.1] in case of the obstacle problem. Indeed, we will show later on that (3.1) is equivalent to B-stationarity for MPECs, see for instance Scheel and Scholtes [2000].

In order to establish (3.1), the main step is to prove the weak directional differentiability of G. This is achieved in the following subsection. The variational inequality (3.1) then follows easily by a chain rule argument which exploits the quadratic structure of the objective, see Section 3.2. In Section 3.3 we confirm that (3.1) is indeed equivalent to the concept of B-stationarity.

Throughout Section 3, we suppose that Assumptions 1.1 and 2.4 hold.

3.1. Weak Directional Differentiability of the Control-to-State Map. For the remainder of this subsection, we fix $p \in (2, \overline{p}]$ as in Theorem 2.5 and suppose that the load ℓ belongs to $W_D^{-1,p}(\Omega; \mathbb{R}^d)$. This will cause no difficulty later on since in fact, $\tau_N^\star$ maps into $W_D^{-1,p}(\Omega; \mathbb{R}^d)$ for some $p > 2$ by Sobolev's embedding theorem. Under this assumption, we will show that G is directionally differentiable in a weak sense in all directions $\delta\ell \in V'$, i.e.,

$$\frac{G(\ell + t\,\delta\ell) - G(\ell)}{t} \rightharpoonup \delta_w G(\ell; \delta\ell) \quad \text{in } S^2 \times V \quad \text{as } t \searrow 0.$$

The weak limit $\delta_w G(\ell; \delta\ell) = (\boldsymbol{\Sigma}', \boldsymbol{u}')$ is given by the unique solution $(\boldsymbol{\Sigma}', \boldsymbol{u}') \in \mathcal{S}_\ell \times V$ of the following variational inequality:

$$\begin{aligned} \langle A\boldsymbol{\Sigma}', \boldsymbol{T} - \boldsymbol{\Sigma}' \rangle + \langle B^\star \boldsymbol{u}', \boldsymbol{T} - \boldsymbol{\Sigma}' \rangle & \\ + \big(\lambda, \mathcal{D}\boldsymbol{\Sigma}' : \mathcal{D}(\boldsymbol{T} - \boldsymbol{\Sigma}')\big)_\Omega \geq 0 \quad & \text{for all } \boldsymbol{T} \in \mathcal{S}_\ell, \qquad (3.2a) \\ B\boldsymbol{\Sigma}' = \delta\ell, & \qquad (3.2b) \end{aligned}$$

where the convex cone $\mathcal{S}_\ell$ is defined by

$$\begin{aligned} \mathcal{S}_\ell := \{\boldsymbol{T} \in S^2 \colon \sqrt{\lambda}\,\mathcal{D}\boldsymbol{T} \in S, \quad & \mathcal{D}\boldsymbol{\Sigma}(x) : \mathcal{D}\boldsymbol{T}(x) \leq 0 \text{ a.e. in } \mathcal{B}(\ell), \\ & \mathcal{D}\boldsymbol{\Sigma}(x) : \mathcal{D}\boldsymbol{T}(x) = 0 \text{ a.e. in } \mathcal{A}_s(\ell)\}. \end{aligned} \qquad (3.3)$$

Here, $(\boldsymbol{\Sigma}, \boldsymbol{u}, \lambda) \in S^2 \times V \times L^2(\Omega)$ is the unique solution of (2.1), i.e. $(\boldsymbol{\Sigma}, \boldsymbol{u}) = G(\ell)$ and $\lambda \in L^2(\Omega)$ is the associated plastic multiplier. The structure of $\mathcal{S}_\ell$ is typical for directional derivatives of solutions to optimization problems such as (2.3). Concerning the linearization of inequality constraints, one needs to distinguish three cases. Inactive constraints impose no restrictions on the derivative, while (strongly active) active constraints have to remain (active) feasible to first order, see Jittorntrum [1984]. The additional condition $\sqrt{\lambda}\,\mathcal{D}\boldsymbol{T} \in S$ serves to make the third term in (3.2a) a priori well defined. It will be shown below that this condition is indeed satisfied for the weak directional derivative of G and therefore it does not introduce an artificial restriction. Note that in case $\lambda \notin L^\infty(\Omega)$, the set $\mathcal{S}_\ell$ is not closed in S^2. Nevertheless, as shown in Theorem 3.2, there exists a unique solution of (3.2).

We also refer to Mignot and Puel [1984] for the control of the obstacle problem where the structure of (3.2) and (3.3) is simpler due to the linearity of the inequality constraint. Finally, we point out that an equivalent characterization of the weak derivative involving a derivative of the plastic multiplier is given in (3.28).

The main result of this subsection is the following.

Theorem 3.2. *For every $\ell \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$ and $\delta\ell \in V'$, the control-to-state map $G : V' \to S^2 \times V$ is weakly directionally differentiable at ℓ in direction $\delta\ell$ and the weak directional derivative is given by the unique solution of (3.2).*

Before we are in the position to prove Theorem 3.2, we need several auxiliary results. Let us consider a fixed but arbitrary sequence of positive real numbers $\{t_n\}$ tending to zero as $n \to \infty$. We introduce a perturbed problem associated with t_n by

$$\langle A\boldsymbol{\Sigma}_n, \boldsymbol{T} - \boldsymbol{\Sigma}_n \rangle + \langle B^* \boldsymbol{u}_n, \boldsymbol{T} - \boldsymbol{\Sigma}_n \rangle \geq 0 \quad \text{for all } \boldsymbol{T} \in \mathcal{K}, \qquad (3.4a)$$

$$B\boldsymbol{\Sigma}_n = \ell + t_n \delta\ell. \qquad (3.4b)$$

with $\boldsymbol{\Sigma}_n \in \mathcal{K}$. Clearly, (3.4) admits a unique solution and, in view of Theorem 2.1, we have

$$\left\| \frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n} \right\|_{S^2} + \left\| \frac{\boldsymbol{u}_n - \boldsymbol{u}}{t_n} \right\|_V \leq L\, \|\delta\ell\|_{V'} < \infty. \qquad (3.5)$$

Therefore, the sequence $\big\{\big((\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma})/t_n, (\boldsymbol{u}_n - \boldsymbol{u})/t_n\big)\big\}$ is bounded in $S^2 \times V$ and there exists a weakly convergent subsequence, which is denoted by the same symbol to simplify the notation. At the end of the proof of Theorem 3.2, we shall see that *every* weakly convergent subsequence of difference quotients has the same limit so that a well known argument yields the weak convergence of the whole sequence. This justifies the simplification of notation. The weak limit is denoted by $(\widetilde{\boldsymbol{\Sigma}}, \widetilde{\boldsymbol{u}})$, i.e.

$$\Big(\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}, \frac{\boldsymbol{u}_n - \boldsymbol{u}}{t_n}\Big) \rightharpoonup (\widetilde{\boldsymbol{\Sigma}}, \widetilde{\boldsymbol{u}}) \quad \text{in } S^2 \times V \text{ for } n \to \infty. \tag{3.6}$$

Our goal is to show that $(\widetilde{\boldsymbol{\Sigma}}, \widetilde{\boldsymbol{u}})$ satisfies (3.2). To this end, we proceed as follows.

(1) We verify that $\widetilde{\boldsymbol{\Sigma}}$ satisfies the sign conditions in (3.3) (Proposition 3.3).
(2) We introduce the plastic multipliers for the perturbed problems, see (3.12).
(3) We show the weak directional differentiability of the plastic multiplier (Proposition 3.4).
(4) We establish $\widetilde{\boldsymbol{\Sigma}} \in \mathcal{S}_\ell$ as well as a complementarity relation for $\widetilde{\boldsymbol{\Sigma}}$ and the weak directional derivative of the plastic multiplier (Proposition 3.8).
(5) In the proof of Theorem 3.2, this complementarity relation is used to show that $(\widetilde{\boldsymbol{\Sigma}}, \widetilde{\boldsymbol{u}})$ satisfies (3.2). Moreover, we prove the uniqueness of the solution of (3.2).

Proposition 3.3. *The weak limit $(\widetilde{\boldsymbol{\Sigma}}, \widetilde{\boldsymbol{u}})$ satisfies $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}} \leq 0$ a.e. in $\mathcal{B}(\ell)$ and $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}} = 0$ a.e. in $\mathcal{A}_s(\ell)$.*

Proof. Due to $\boldsymbol{\Sigma}, \boldsymbol{\Sigma}_n \in \mathcal{K}$ for every $n \in \mathbb{N}$, one finds

$$\mathcal{D}\boldsymbol{\Sigma}(x) : (\mathcal{D}\boldsymbol{\Sigma}_n(x) - \mathcal{D}\boldsymbol{\Sigma}(x)) \leq |\mathcal{D}\boldsymbol{\Sigma}(x)|\,|\mathcal{D}\boldsymbol{\Sigma}_n(x)| - \tilde{\sigma}_0^2 \leq 0$$

on the active set $\mathcal{A}(\ell)$ and thus

$$\mathcal{D}\boldsymbol{\Sigma}(x) : \frac{\mathcal{D}\boldsymbol{\Sigma}_n(x) - \mathcal{D}\boldsymbol{\Sigma}(x)}{t_n} \leq 0 \quad \text{a.e. in } \mathcal{A}(\ell) \tag{3.7}$$

for all $n \in \mathbb{N}$. We note that the set $\{\boldsymbol{T} \in S^2 : \mathcal{D}\boldsymbol{\Sigma}(x) : \mathcal{D}\boldsymbol{T}(x) \leq 0 \text{ a.e. in } \mathcal{A}(\ell)\}$ is convex and closed, thus weakly closed. This implies

$$\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}} \leq 0 \quad \text{a.e. in } \mathcal{A}(\ell). \tag{3.8}$$

Since $\lambda(x) = 0$ a.e. in $\mathcal{I}(\ell)$ and $\lambda(x) \geq 0$ a.e. in $\mathcal{A}(\ell)$, (3.7) yields

$$\Big(\lambda, \mathcal{D}\boldsymbol{\Sigma} : \frac{\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\boldsymbol{\Sigma}}{t_n}\Big)_\Omega \leq 0 \quad \text{for all } n \in \mathbb{N}. \tag{3.9}$$

Now we test (3.4a) with $\boldsymbol{T} = \boldsymbol{\Sigma}$ which is clearly feasible since $\boldsymbol{\Sigma} \in \mathcal{K}$. Then (3.9) together with (2.1a) implies

$$\begin{aligned}
0 &\leq -\Big(\lambda, \mathcal{D}\boldsymbol{\Sigma} : \frac{\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\boldsymbol{\Sigma}}{t_n}\Big)_\Omega \\
&\leq t_n\Big[-\Big\langle A\Big(\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}\Big), \Big(\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}\Big)\Big\rangle - \Big\langle B^\star\Big(\frac{\boldsymbol{u}_n - \boldsymbol{u}}{t_n}\Big), \Big(\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}\Big)\Big\rangle\Big] \\
&\leq c\,t_n \Big\|\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}\Big\|_{S^2}\Big\|\frac{\boldsymbol{u}_n - \boldsymbol{u}}{t_n}\Big\|_V \leq c\,t_n\,\|\delta\ell\|_{V'}^2 \to 0 \quad \text{as } n \to \infty
\end{aligned} \tag{3.10}$$

because of (3.5). Furthermore, (3.6) implies the weak convergence of $(\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\boldsymbol{\Sigma})/t_n$ in S to $\mathcal{D}\widetilde{\boldsymbol{\Sigma}}$. Hence, we obtain

$$\Big(\lambda, \mathcal{D}\boldsymbol{\Sigma} : \frac{\mathcal{D}\boldsymbol{\Sigma}_n - \mathcal{D}\boldsymbol{\Sigma}}{t_n}\Big)_\Omega \to (\lambda, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}})_\Omega$$

and therefore, (3.10) gives

$$(\lambda, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}})_\Omega = 0. \tag{3.11}$$

Hence, since $\lambda > 0$ on $\mathcal{A}_s(\ell)$ holds, (3.8) implies $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}} = 0$ a.e. in $\mathcal{A}_s(\ell)$. □

In order to prove the existence of the weak directional derivative of the plastic multiplier, we reformulate the perturbed problem (3.4) by introducing the plastic multiplier λ_n, see Theorem 2.2:

$$A\boldsymbol{\Sigma}_n + B^\star \boldsymbol{u}_n + \lambda_n \, \mathcal{D}^\star \mathcal{D}\boldsymbol{\Sigma}_n = \mathbf{0} \tag{3.12a}$$

$$B\boldsymbol{\Sigma}_n = \ell + t_n \delta\ell \tag{3.12b}$$

$$0 \le \lambda_n(x) \ \perp \ \phi(\boldsymbol{\Sigma}_n(x)) \le 0 \quad \text{a.e. in } \Omega. \tag{3.12c}$$

Arguing as in Remark 2.3 yields

$$\begin{aligned} \lambda \, \mathcal{D}\boldsymbol{\Sigma} &= -\mathbb{H}^{-1}\boldsymbol{\chi}, \\ \lambda_n \, \mathcal{D}\boldsymbol{\Sigma}_n &= -\mathbb{H}^{-1}\boldsymbol{\chi}_n. \end{aligned} \tag{3.13}$$

These relations define the starting point for the proof of convergence of the plastic multipliers in the following proposition.

Proposition 3.4. *Given $\ell \in W_D^{-1,p}(\Omega;\mathbb{R}^d)$ and $\delta\ell \in V'$, we have convergence of the plastic multipliers $\lambda_n \to \lambda$ in $L^2(\Omega)$. Their difference quotient $(\lambda_n - \lambda)/t_n \rightharpoonup \widetilde{\lambda}$ converges weakly in $L^1(\Omega)$. Moreover, $\widetilde{\lambda} \in L^2(\Omega)$ and $\widetilde{\lambda}$ is uniquely determined by the weak limit $\widetilde{\boldsymbol{\Sigma}}$ of the stresses.*

Proof. We address $\lambda_n \to \lambda$ first. Since $\lambda = 0$ on $\mathcal{I}(\ell)$ and $\lambda_n = 0$ on $\mathcal{I}(\ell + t_n \, \delta\ell)$, we get by (3.13) the following characterization of λ and λ_n,

$$\begin{aligned} \tilde\sigma_0^2 \, \lambda &= \lambda \, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{\Sigma} = -\mathbb{H}^{-1}\boldsymbol{\chi} : \mathcal{D}\boldsymbol{\Sigma} && \text{a.e. in } \Omega, \\ \tilde\sigma_0^2 \, \lambda_n &= \lambda_n \, \mathcal{D}\boldsymbol{\Sigma}_n : \mathcal{D}\boldsymbol{\Sigma}_n = -\mathbb{H}^{-1}\boldsymbol{\chi}_n : \mathcal{D}\boldsymbol{\Sigma}_n && \text{a.e. in } \Omega. \end{aligned}$$

Taking the difference yields

$$\begin{aligned} \lambda_n - \lambda &= \frac{1}{\tilde\sigma_0^2}\big(-\mathbb{H}^{-1}\boldsymbol{\chi}_n : \mathcal{D}\boldsymbol{\Sigma}_n + \mathbb{H}^{-1}\boldsymbol{\chi} : \mathcal{D}\boldsymbol{\Sigma}\big) \\ &= \frac{1}{\tilde\sigma_0^2}\big(\mathbb{H}^{-1}\boldsymbol{\chi} : (\mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\boldsymbol{\Sigma}_n) + \mathbb{H}^{-1}(\boldsymbol{\chi} - \boldsymbol{\chi}_n) : \mathcal{D}\boldsymbol{\Sigma}_n\big). \end{aligned} \tag{3.14}$$

Clearly, the right hand side is bounded in $L^2(\Omega)$ since $\mathcal{D}\boldsymbol{\Sigma}, \mathcal{D}\boldsymbol{\Sigma}_n \in L^\infty(\Omega;\mathbb{S})$. Thus there exists a subsequence λ_{n_k} of λ_n which converges weakly in $L^2(\Omega)$ to some λ°. We will show $\lambda^\circ = \lambda$, therefore the weak limit is independent of the chosen subsequence, and hence the whole sequence $\{\lambda_n\}$ converges weakly in $L^2(\Omega)$, i.e. $\lambda_n \rightharpoonup \lambda$. Using again (3.13) and the convergence of $\boldsymbol{\chi}_n \to \boldsymbol{\chi}$ in S gives

$$\lambda_n \, \mathcal{D}\boldsymbol{\Sigma}_n = -\mathbb{H}^{-1}\boldsymbol{\chi}_n \to -\mathbb{H}^{-1}\boldsymbol{\chi} = \lambda \, \mathcal{D}\boldsymbol{\Sigma} \quad \text{in } S. \tag{3.15}$$

On the other hand, the weak convergence $\lambda_{n_k} \rightharpoonup \lambda^\circ$ in $L^2(\Omega)$ implies

$$\lambda_{n_k} \, \mathcal{D}\boldsymbol{\Sigma}_{n_k} \rightharpoonup \lambda^\circ \, \mathcal{D}\boldsymbol{\Sigma} \quad \text{in } L^1(\Omega;\mathbb{S}).$$

Therefore $\lambda = \lambda^\circ$ holds on $\mathcal{A}(\ell)$. Due to the complementarity condition (2.1c), $\lambda = 0$ holds on $\mathcal{I}(\ell)$ and this implies $\|\lambda\|_{L^2(\Omega)} \le \|\lambda^\circ\|_{L^2(\Omega)}$.

The complementarity relations (2.1c) and (3.12c) together with (3.15) imply

$$\|\lambda\|_{L^2(\Omega)} = \frac{1}{\tilde\sigma_0}\|\lambda \, \mathcal{D}\boldsymbol{\Sigma}\|_S = \lim_{n\to\infty} \frac{1}{\tilde\sigma_0}\|\lambda_n \, \mathcal{D}\boldsymbol{\Sigma}_n\|_S = \lim_{n\to\infty} \|\lambda_n\|_{L^2(\Omega)}.$$

Since the norm is weakly lower semicontinuous this implies

$$\|\lambda\|_{L^2(\Omega)} = \lim_{n\to\infty} \|\lambda_n\|_{L^2(\Omega)} = \liminf_{k\to\infty} \|\lambda_{n_k}\|_{L^2(\Omega)} \ge \|\lambda^\circ\|_{L^2(\Omega)}.$$

We conclude $\|\lambda^\circ\|_{L^2(\Omega)} = \|\lambda\|_{L^2(\Omega)}$. Due to $\lambda^\circ = \lambda$ on $\mathcal{A}(\ell)$ and $\lambda = 0$ on $\mathcal{I}(\ell)$, $\lambda^\circ = \lambda$ is satisfied. This shows the independence of the weak limit from the chosen subsequence, thus the whole sequence $\{\lambda_n\}$ converges weakly to λ. Additionally, in view of the convergence of norms, $\lambda_n \to \lambda$ strongly in $L^2(\Omega)$.

We proceed by showing the weak convergence of the difference quotients for λ. Again, we obtain this convergence for a subsequence first, and derive a representation formula for the weak limit $\widetilde{\lambda}$, see (3.19). Given $\widetilde{\boldsymbol{\Sigma}}$, the representation formula implies the uniqueness of the corresponding $\widetilde{\lambda}$. For this purpose, fix $\varepsilon = p - 2 > 0$ (see Theorem 2.5) and define $\hat{\varepsilon} = \varepsilon/(4+\varepsilon) > 0$. Due to (3.14) and Assumption 1.1 (3), we get

$$\begin{aligned}\|\lambda_n - \lambda\|_{L^{1+\hat{\varepsilon}}(\Omega)} &\le c\Big(\|\boldsymbol{\chi} : (\mathcal{D}\boldsymbol{\Sigma} - \mathcal{D}\boldsymbol{\Sigma}_n)\|_{L^{1+\hat{\varepsilon}}(\Omega)} + \|(\boldsymbol{\chi} - \boldsymbol{\chi}_n) : \mathcal{D}\boldsymbol{\Sigma}_n\|_{L^{1+\hat{\varepsilon}}(\Omega)}\Big)\\ &\le c\Big(\|\boldsymbol{\chi}\|_{L^{2+\varepsilon}(\Omega;\mathbb{S})} + \|\mathcal{D}\boldsymbol{\Sigma}_n\|_{L^\infty(\Omega;\mathbb{S})}\Big)\|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_n\|_{S^2}\\ &\le c\, t_n.\end{aligned}$$

Here we applied Hölder's inequality with $2/(4+\varepsilon) + (2+\varepsilon)/(4+\varepsilon) = 1$ and we used the boundedness of $\mathcal{D}\boldsymbol{\Sigma}_n$ in $L^\infty(\Omega;\mathbb{S})$ and the regularity $\boldsymbol{\chi} \in L^{2+\varepsilon}(\Omega;\mathbb{S})$ (see Theorem 2.5). Moreover, we used the Lipschitz estimate $\|\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}\|_{S^2} \le L\, t_n \|\delta\ell\|_{V'}$ (see Theorem 2.1). The reflexivity of $L^{1+\hat{\varepsilon}}(\Omega)$ implies the existence of a subsequence t_{n_k} of t_n and some $\widetilde{\lambda} \in L^{1+\hat{\varepsilon}}(\Omega)$ such that $(\lambda_{n_k} - \lambda)/t_{n_k} \rightharpoonup \widetilde{\lambda}$ in $L^{1+\hat{\varepsilon}}(\Omega)$.

Now we derive a representation formula for $\widetilde{\lambda}$. Using the above convergence results, we obtain

$$\begin{aligned}\frac{\lambda_{n_k}\mathcal{D}\boldsymbol{\Sigma}_{n_k} - \lambda\,\mathcal{D}\boldsymbol{\Sigma}}{t_{n_k}} &= \frac{\lambda_{n_k} - \lambda}{t_{n_k}}\mathcal{D}\boldsymbol{\Sigma} + \lambda_{n_k}\frac{\mathcal{D}\boldsymbol{\Sigma}_{n_k} - \mathcal{D}\boldsymbol{\Sigma}}{t_{n_k}}\\ &\rightharpoonup \widetilde{\lambda}\,\mathcal{D}\boldsymbol{\Sigma} + \lambda\,\mathcal{D}\widetilde{\boldsymbol{\Sigma}} \quad \text{in } L^1(\Omega;\mathbb{S}).\end{aligned} \tag{3.16}$$

Due to (3.13) we find

$$\frac{\lambda_n\,\mathcal{D}\boldsymbol{\Sigma}_n - \lambda\,\mathcal{D}\boldsymbol{\Sigma}}{t_n} = \frac{\mathbb{H}^{-1}\boldsymbol{\chi} - \mathbb{H}^{-1}\boldsymbol{\chi}_n}{t_n} \rightharpoonup -\mathbb{H}^{-1}\widetilde{\boldsymbol{\chi}} \quad \text{in } S.$$

Consequently,

$$-\,\mathbb{H}^{-1}\widetilde{\boldsymbol{\chi}} = \widetilde{\lambda}\,\mathcal{D}\boldsymbol{\Sigma} + \lambda\,\mathcal{D}\widetilde{\boldsymbol{\Sigma}}. \tag{3.17}$$

Note that (3.17) implies the uniqueness of $\widetilde{\lambda}\,\mathcal{D}\boldsymbol{\Sigma}$ for any given $\widetilde{\boldsymbol{\Sigma}}$. To verify the uniqueness of $\widetilde{\lambda}$, we show next that $\widetilde{\lambda} = 0$ on $\mathcal{I}(\ell)$. For convenience, let us abbreviate

$$\mathcal{A}_{s,k} := \mathcal{A}_s(\ell + t_{n_k}\delta\ell).$$

We consider the sets $\mathcal{I}(\ell) \cap \mathcal{A}_{s,k}$, on which $|\mathcal{D}\boldsymbol{\Sigma}| < \tilde{\sigma}_0$ and $|\mathcal{D}\boldsymbol{\Sigma}_{n_k}| = \tilde{\sigma}_0$ hold. From $\boldsymbol{\Sigma}_{n_k} \to \boldsymbol{\Sigma}$ in S^2 we infer $|\mathcal{D}\boldsymbol{\Sigma}_{n_k}| \to |\mathcal{D}\boldsymbol{\Sigma}|$ in $L^2(\Omega)$. Lemma A.2 with $M = \mathcal{I}(\ell)$, $f = \tilde{\sigma}_0 - |\mathcal{D}\boldsymbol{\Sigma}|$, $f_k = \tilde{\sigma}_0 - |\mathcal{D}\boldsymbol{\Sigma}_{n_k}|$ yields $|\mathcal{I}(\ell) \cap \mathcal{A}_{s,k}| \to 0$ as $k \to \infty$. We may now invoke Lemma A.3 with $f_{n_k} := (\lambda_{n_k} - \lambda)/t_{n_k} \ge 0$ and $f := \widetilde{\lambda}$, both restricted to $\mathcal{I}(\ell)$, to conclude that

$$\widetilde{\lambda}(x) = 0 \quad \text{a.e. in } \mathcal{I}(\ell). \tag{3.18}$$

Testing (3.17) with $\mathcal{D}\boldsymbol{\Sigma} \in L^\infty(\Omega;\mathbb{S})$ and using (3.11) and (3.18) implies

$$\widetilde{\lambda} = -\frac{\mathcal{D}\boldsymbol{\Sigma} : \mathbb{H}^{-1}\widetilde{\boldsymbol{\chi}}}{\tilde{\sigma}_0^2} \quad \text{a.e. in } \Omega. \tag{3.19}$$

To summarize, we have shown the following: Given a sequence $t_n \searrow 0$ and a weak limit $\widetilde{\boldsymbol{\Sigma}}$ according to (3.6), there exists a subsequence t_{n_k} such that the difference quotient $(\lambda_{n_k} - \lambda)/t_{n_k}$ converges weakly to some $\widetilde{\lambda}$ in $L^1(\Omega)$. Moreover, the weak limit $\widetilde{\lambda}$ is independent of the subsequence (it is uniquely determined by $\widetilde{\boldsymbol{\Sigma}}$) and it belongs to $L^2(\Omega)$ by (3.19). Hence the whole sequence of difference quotients converges weakly. □

We are now in the position to state an equation relating the terms for the directional derivative problem.

Lemma 3.5. *The weak limit* $(\widetilde{\boldsymbol{\Sigma}}, \widetilde{\boldsymbol{u}}, \widetilde{\lambda})$ *satisfies*

$$A\widetilde{\boldsymbol{\Sigma}} + B^\star\widetilde{\boldsymbol{u}} + \lambda\,\mathcal{D}^\star\mathcal{D}\widetilde{\boldsymbol{\Sigma}} + \widetilde{\lambda}\,\mathcal{D}^\star\mathcal{D}\boldsymbol{\Sigma} = \boldsymbol{0} \quad \text{in } S^2. \tag{3.20}$$

Proof. Let us start with an arbitrary $\boldsymbol{T} \in L^\infty(\Omega;\mathbb{S}^2)$. In view of (2.1a) and (3.12a) we have

$$\Big\langle A\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}, \boldsymbol{T}\Big\rangle + \Big\langle B^\star\frac{\boldsymbol{u}_n - \boldsymbol{u}}{t_n}, \boldsymbol{T}\Big\rangle + \Big(\frac{\lambda_n\,\mathcal{D}\boldsymbol{\Sigma}_n - \lambda\,\mathcal{D}\boldsymbol{\Sigma}}{t_n}, \mathcal{D}\boldsymbol{T}\Big)_\Omega = 0.$$

Due to (3.16) we can pass to the limit to obtain

$$\langle A\widetilde{\boldsymbol{\Sigma}}, \boldsymbol{T}\rangle + \langle B^\star\widetilde{\boldsymbol{u}}, \boldsymbol{T}\rangle + \big(\lambda, \mathcal{D}\widetilde{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{T}\big)_\Omega + \big(\widetilde{\lambda}, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{T}\big)_\Omega = 0 \quad \text{for all } \boldsymbol{T} \in L^\infty(\Omega;\mathbb{S}^2). \tag{3.21}$$

By (3.17), we know that $\lambda\,\mathcal{D}\widetilde{\boldsymbol{\Sigma}}$ belongs to S^2 and thus the left hand side of (3.21) is a continuous mapping from S^2 to $\mathbb{R}$ with respect to $\boldsymbol{T}$. Therefore the density of $L^\infty(\Omega;\mathbb{S}^2)$ in S^2 implies that (3.21) holds for all $\boldsymbol{T} \in S^2$. □

In order to pass from (3.20) to the variational inequality (3.2a), we need to verify a certain complementarity condition, see Proposition 3.8 below. To this end, we recall a particular case of [Herzog et al., 2010a, Proposition 3.15]:

Lemma 3.6. *Let* $\varphi \in C_0^\infty(\Omega)$ *with* $\varphi \ge 0$ *be arbitrary and suppose that*

$$\boldsymbol{\Phi}_k \rightharpoonup \boldsymbol{\Phi} \quad \text{in } S^2, \qquad \boldsymbol{z}_k \rightharpoonup \boldsymbol{z} \quad \text{in } V, \qquad B\boldsymbol{\Phi}_k \equiv h \quad \text{in } V'.$$

Assume further that

$$\langle A\boldsymbol{\Phi}_k, \varphi\,\boldsymbol{\Phi}_k\rangle + \langle B^\star\boldsymbol{z}_k, \varphi\,\boldsymbol{\Phi}_k\rangle \le 0 \quad \text{for all } k \in \mathbb{N}.$$

Then $\langle A\boldsymbol{\Phi}, \varphi\,\boldsymbol{\Phi}\rangle + \langle B^\star\boldsymbol{z}, \varphi\,\boldsymbol{\Phi}\rangle \le 0$.

Lemma 3.7. *For all* $\boldsymbol{T} \in \mathcal{S}_\ell$, *there holds*

$$(\widetilde{\lambda}\,\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{T})(x) \begin{cases} \le 0 & \text{a.e. in } \mathcal{B}(\ell) \\ = 0 & \text{a.e. in } \Omega \setminus \mathcal{B}(\ell). \end{cases} \tag{3.22}$$

Proof. Recall that $\widetilde{\lambda} = 0$ a.e. in $\mathcal{I}(\ell)$. Moreover we have $\widetilde{\lambda} \ge 0$ a.e. in $\Omega \setminus \mathcal{A}_s(\ell) = \{x \in \Omega : \lambda(x) = 0\}$, which follows from

$$0 \le \frac{\lambda_n}{t_n} = \frac{\lambda_n - \lambda}{t_n} \rightharpoonup \widetilde{\lambda} \quad \text{in } L^1(\Omega \setminus \mathcal{A}_s(\ell)) \tag{3.23}$$

and the weak closedness in L^1 of the set of nonnegative functions. Now the assertion follows from $\boldsymbol{T} \in \mathcal{S}_\ell$, i.e., $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{T} = 0$ on $\mathcal{A}_s(\ell)$ and $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{T} \le 0$ on $\mathcal{B}(\ell)$. □

Proposition 3.8. *The weak limit* $\widetilde{\boldsymbol{\Sigma}}$ *belongs to* $\mathcal{S}_\ell$. *Moreover, the relation* $\widetilde{\lambda}\,\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}} = 0$ *holds a.e. in* Ω.

Proof. We have already verified in Proposition 3.3 that $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}} \le 0$ in $\mathcal{B}(\ell)$ and $\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}} = 0$ in $\mathcal{A}_s(\ell)$. To prove $\widetilde{\boldsymbol{\Sigma}} \in \mathcal{S}_\ell$, it remains to show $\sqrt{\lambda}\,\mathcal{D}\widetilde{\boldsymbol{\Sigma}} \in S$. This follows by testing (3.17) with $\mathcal{D}\widetilde{\boldsymbol{\Sigma}}$ and using $\mathcal{D}\boldsymbol{\Sigma} \in L^\infty(\Omega;\mathbb{S})$.

To show the slackness condition, let $\varphi \in C_0^\infty(\Omega)$ with $0 \le \varphi \le 1$ be given. Since $\boldsymbol{\Sigma}, \boldsymbol{\Sigma}_n \in \mathcal{K}$ for all $n \in \mathbb{N}$, we may test **(VI)** with $\boldsymbol{T} = \varphi\,\boldsymbol{\Sigma}_n + (1-\varphi)\,\boldsymbol{\Sigma} \in \mathcal{K}$ and (3.4a) with $\boldsymbol{T} = \varphi\,\boldsymbol{\Sigma} + (1-\varphi)\,\boldsymbol{\Sigma}_n \in \mathcal{K}$. Adding the arising inequalities implies

$$\Big\langle A\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}, \varphi\,\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}\Big\rangle + \Big\langle B^\star\frac{\boldsymbol{u}_n - \boldsymbol{u}}{t_n}, \varphi\,\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}\Big\rangle \le 0.$$

Now we apply Lemma 3.6 with

$$\boldsymbol{\Phi}_n := (\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma})/t_n \rightharpoonup \widetilde{\boldsymbol{\Sigma}} \text{ in } S^2, \qquad \boldsymbol{z}_n := (\boldsymbol{u}_n - \boldsymbol{u})/t_n \rightharpoonup \widetilde{\boldsymbol{u}} \text{ in } V.$$

This yields

$$\langle A\widetilde{\boldsymbol{\Sigma}}, \varphi\,\widetilde{\boldsymbol{\Sigma}}\rangle + \langle B^\star\widetilde{\boldsymbol{u}}, \varphi\,\widetilde{\boldsymbol{\Sigma}}\rangle \le 0 \quad \text{for all } \varphi \in C_0^\infty(\Omega) \text{ with values in } [0,1].$$

A simple scaling argument shows the same inequality for all nonnegative $\varphi \in C_0^\infty(\Omega)$, and thus

$$\widetilde{\boldsymbol{\Sigma}}(x) : (A\widetilde{\boldsymbol{\Sigma}})(x) + \widetilde{\boldsymbol{\Sigma}}(x) : (B^\star\widetilde{\boldsymbol{u}})(x) \le 0 \quad \text{a.e. in } \Omega.$$

Since $\lambda = 0$ on $\mathcal{B}(\ell)$, (3.20) implies

$$\widetilde{\boldsymbol{\Sigma}}(x) : (A\widetilde{\boldsymbol{\Sigma}})(x) + \widetilde{\boldsymbol{\Sigma}}(x) : (B^\star\widetilde{\boldsymbol{u}})(x) + \widetilde{\lambda}(x)\,(\mathcal{D}\boldsymbol{\Sigma})(x) : (\mathcal{D}\widetilde{\boldsymbol{\Sigma}})(x) = 0 \quad \text{a.e. on } \mathcal{B}(\ell).$$

Therefore $\widetilde{\lambda}\,\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}} \ge 0$ on the biactive set $\mathcal{B}(\ell)$ and due to $\widetilde{\boldsymbol{\Sigma}} \in \mathcal{S}_\ell$ and (3.22) we have

$$\widetilde{\lambda}\,\mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}} = 0 \qquad \text{a.e. in } \Omega. \tag{3.24}$$

□

Finally we are in the position to prove Theorem 3.2.

Proof of Theorem 3.2. Let $\boldsymbol{T} \in \mathcal{S}_\ell$ be arbitrary. We test (3.20) with $\boldsymbol{T} - \widetilde{\boldsymbol{\Sigma}}$ which leads to

$$\langle A\widetilde{\boldsymbol{\Sigma}}, \boldsymbol{T} - \widetilde{\boldsymbol{\Sigma}}\rangle + \langle B^\star\widetilde{\boldsymbol{u}}, \boldsymbol{T} - \widetilde{\boldsymbol{\Sigma}}\rangle + \big(\lambda, \mathcal{D}\widetilde{\boldsymbol{\Sigma}} : \mathcal{D}(\boldsymbol{T} - \widetilde{\boldsymbol{\Sigma}})\big)_\Omega \\ = \big(\widetilde{\lambda}, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}}\big)_\Omega - \big(\widetilde{\lambda}, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{T}\big)_\Omega.$$

The first addend on the right hand side vanishes due to Proposition 3.8. In view of Lemma 3.7, we conclude

$$\langle A\widetilde{\boldsymbol{\Sigma}}, \boldsymbol{T} - \widetilde{\boldsymbol{\Sigma}}\rangle + \langle B^\star\widetilde{\boldsymbol{u}}, \boldsymbol{T} - \widetilde{\boldsymbol{\Sigma}}\rangle + \big(\lambda, \mathcal{D}\widetilde{\boldsymbol{\Sigma}} : \mathcal{D}(\boldsymbol{T} - \widetilde{\boldsymbol{\Sigma}})\big)_\Omega \ge 0,$$

which is the claimed variational inequality for the derivative (3.2a).

The equation in (**VI**) and (3.4b) imply

$$B\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n} = \delta\ell \quad \text{in } V'$$

and the weak convergence immediately gives $B\widetilde{\boldsymbol{\Sigma}} = \delta\ell$.

We have shown that the weak limit of the difference quotients $(\widetilde{\boldsymbol{\Sigma}}, \widetilde{\boldsymbol{u}})$ satisfies (3.2).

It remains to verify that (3.2) does not admit other solutions. Suppose on the contrary that $(\boldsymbol{\Sigma}', \boldsymbol{u}')$ and $(\boldsymbol{\Sigma}'', \boldsymbol{u}'')$ are two solutions, then a simple testing argument using $\lambda \ge 0$ and the coercivity of A, thanks to Assumption 1.1 (3), shows that $\boldsymbol{\Sigma}'$ and $\boldsymbol{\Sigma}''$ must coincide. To verify the uniqueness of the displacement field, we define $\boldsymbol{\tau} = \boldsymbol{\varepsilon}(\boldsymbol{u}' - \boldsymbol{u}'')$ and $\boldsymbol{T}' = (\boldsymbol{\tau}, -\boldsymbol{\tau}) + \boldsymbol{\Sigma}'$ and $\boldsymbol{T}'' = \boldsymbol{\Sigma}''$. These are feasible as test functions in (3.2a) due to the structure of $\mathcal{K}$. This implies

$$0 \le \langle B^\star(\boldsymbol{u}' - \boldsymbol{u}''), \boldsymbol{T}' - \boldsymbol{T}''\rangle = -\int_\Omega \boldsymbol{\varepsilon}(\boldsymbol{u}' - \boldsymbol{u}'') : \boldsymbol{\varepsilon}(\boldsymbol{u}' - \boldsymbol{u}'')\,\mathrm{d}x \le 0.$$

From here, Korn's inequality (1.5) shows $\boldsymbol{u}' = \boldsymbol{u}''$.

Thus the weak limit is unique and a well known argument implies the convergence of the whole sequence, i.e.

$$\frac{G(\ell + t_n\delta\ell) - G(\ell)}{t_n} = \Big(\frac{\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}}{t_n}, \frac{\boldsymbol{u}_n - \boldsymbol{u}}{t_n}\Big) \rightharpoonup (\boldsymbol{\Sigma}', \boldsymbol{u}') = \delta_w G(\ell; \delta\ell),$$

which is the assertion of Theorem 3.2. □

3.2. B-Stationarity Conditions. We start the discussion of (**P**) with a differentiability result for general quadratic functionals:

Lemma 3.9. *Let W, H be Hilbert spaces and $\mathcal{G} : W \to H$ weakly directionally differentiable at $w \in W$, i.e.*

$$\frac{\mathcal{G}(w + t\,\delta w) - \mathcal{G}(w)}{t} \rightharpoonup \delta_w \mathcal{G}(w; \delta w) \quad \textit{in } H \quad \textit{as } t \searrow 0 \tag{3.25}$$

for every $\delta w \in W$. Then the functional $j : W \to \mathbb{R}$, defined by

$$j(w) = \frac{1}{2}\|\mathcal{G}(w) - z\|_H^2 + \frac{\nu}{2}\|w\|_W^2 \quad \textit{with } z \in H \textit{ and } \nu \in \mathbb{R} \textit{ given}$$

is directionally differentiable at w, and its directional derivative in the direction $\delta w \in W$ is given by

$$\delta j(w; \delta w) = \big\langle \mathcal{G}(w) - z,\, \delta_w \mathcal{G}(w; \delta w)\big\rangle_H + \nu\, \langle w,\, \delta w\rangle_W.$$

Proof. The proof is almost trivial. Nevertheless, for convenience of the reader, we present the arguments in the following. We have to verify that

$$\Big|\frac{j(w + t\,\delta w) - j(w)}{t} - \delta j(w; \delta w)\Big| \to 0 \text{ in } \mathbb{R} \quad \text{as } t \searrow 0.$$

First of all, note that (3.25) yields

$$\|\mathcal{G}(w + t\,\delta w) - \mathcal{G}(w)\|_H = t\,\Big\|\frac{\mathcal{G}(w + t\,\delta w) - \mathcal{G}(w)}{t}\Big\|_H \to 0 \text{ as } t \searrow 0, \tag{3.26}$$

since weak convergence implies boundedness. By a straightforward computation, one obtains

$$\begin{aligned}\frac{j(w + t\,\delta w) - j(w)}{t} \\ &= \frac{1}{2}\,\big\|\mathcal{G}(w + t\,\delta w) - \mathcal{G}(w)\big\|_H\,\Big\|\frac{\mathcal{G}(w + t\,\delta w) - \mathcal{G}(w)}{t}\Big\|_H \\ &\quad + \Big\langle \frac{\mathcal{G}(w + t\,\delta w) - \mathcal{G}(w)}{t},\, \mathcal{G}(w) - z\Big\rangle_H + \frac{\nu}{2}\,t\,\|\delta w\|_W^2 + \nu\,\langle w,\, \delta w\rangle_W\end{aligned}$$

and hence

$$\begin{aligned}\Big|\frac{j(w + t\,\delta w) - j(w)}{t} - \delta j(w; \delta w)\Big| \\ &\le \frac{1}{2}\,\big\|\mathcal{G}(w + t\,\delta w) - \mathcal{G}(w)\big\|_H\Big\|\frac{\mathcal{G}(w + t\,\delta w) - \mathcal{G}(w)}{t}\Big\|_H \\ &\quad + \Big|\Big\langle \frac{\mathcal{G}(w + t\,\delta w) - \mathcal{G}(w)}{t} - \delta_w \mathcal{G}(w; \delta w),\, \mathcal{G}(w) - z\Big\rangle_H\Big| + \frac{\nu}{2}\,t\,\|\delta w\|_W^2 \to 0\end{aligned}$$

as $t \searrow 0$ because of (3.25) and (3.26). ∎

In order to apply the general setting of Lemma 3.9 to our optimal control problem (**P**), we set

$$\begin{aligned} W &= L^2(\Gamma_N; \mathbb{R}^d), \quad w = \boldsymbol{g}, \quad z = \boldsymbol{u}_d, \\ H &= L^2(\Omega; \mathbb{R}^d), \qquad \mathcal{G} = G^{(\boldsymbol{u})} \circ (-\tau_N^\star). \end{aligned}$$

The weak directional differentiability of $\mathcal{G}$ follows from Theorem 3.2, taking into account that $\tau_N^\star$ maps W into $W_D^{-1,p}(\Omega; \mathbb{R}^d)$ by Sobolev's embedding theorem.

We conclude that local minimizers $\overline{\boldsymbol{g}}$ of (**P**) necessarily satisfy (3.1). The following theorem states this in more explicit terms.

Theorem 3.10. *Let $\overline{\boldsymbol{g}} \in U_{\text{ad}}$ be a local optimal solution of* (**P**) *with associated state $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}) = G(\ell)$ for $\ell := -\tau_N^\star \overline{\boldsymbol{g}}$. Then the following variational inequality is satisfied:*

$$\int_\Omega (\bar{\boldsymbol{u}} - \boldsymbol{u}_d) \cdot \boldsymbol{u}' \, \mathrm{d}x + \nu \int_{\Gamma_N} \overline{\boldsymbol{g}} \cdot (\boldsymbol{g} - \overline{\boldsymbol{g}}) \, \mathrm{d}s \geq 0 \quad \text{for all } \boldsymbol{g} \in U_{\text{ad}}, \tag{3.27}$$

where $(\boldsymbol{\Sigma}', \boldsymbol{u}')$ solves the derivative problem (3.2) with $\delta\ell := -\tau_N^\star(\boldsymbol{g} - \overline{\boldsymbol{g}})$ as right hand side.

Remark 3.11. *We point out that an objective acting on the strain such as for instance*

$$J(\boldsymbol{u}, \boldsymbol{g}) := \frac{1}{2} \|\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{z}\|_S^2 + \frac{\nu}{2} \|\boldsymbol{g}\|^2_{L^2(\Gamma_N;\mathbb{R}^d)}$$

with given $\boldsymbol{z} \in S$ would also fit into the analysis since $G \circ (-\tau_N^\star)$ is weakly directionally differentiably from $L^2(\Gamma_N; \mathbb{R}^d)$ to $S^2 \times V$ and $\boldsymbol{\varepsilon} : V \to S$ is linear and bounded, thus weakly continuous.

3.3. Discussion of B-Stationarity. In the remaining part of this section, we briefly reformulate the B-stationarity result of Theorem 3.10 in order to allow a comparison with B-stationarity conditions known for finite dimensional MPECs. We start with an equivalent formulation of the variational inequality (3.2) for the derivative, which involves the derivative λ' of the plastic multiplier.

Proposition 3.12. *Let $\ell \in W_D^{-1,p}(\Omega; \mathbb{R}^d)$ and $\delta\ell \in V'$ be given. Let $(\boldsymbol{\Sigma}, \boldsymbol{u}, \lambda)$ be the state and plastic multiplier associated with ℓ. A pair $(\boldsymbol{\Sigma}', \boldsymbol{u}') \in S^2 \times V$ is the unique solution of (3.2) if and only if there exists a multiplier $\lambda' \in L^2(\Omega)$ such that*

$$A\boldsymbol{\Sigma}' + B^\star \boldsymbol{u}' + \lambda \, \mathcal{D}^\star \mathcal{D} \boldsymbol{\Sigma}' + \lambda' \, \mathcal{D}^\star \mathcal{D} \boldsymbol{\Sigma} = \boldsymbol{0} \quad \text{in } S^2, \tag{3.28a}$$

$$B\boldsymbol{\Sigma}' = \delta\ell \quad \text{in } V', \tag{3.28b}$$

$$\mathbb{R} \ni \lambda'(x) \;\perp\; \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{\Sigma}'(x) = 0 \quad \text{a.e. in } \mathcal{A}_s(\ell), \tag{3.28c}$$

$$0 \leq \lambda'(x) \;\perp\; \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{\Sigma}'(x) \leq 0 \quad \text{a.e. in } \mathcal{B}(\ell), \tag{3.28d}$$

$$0 = \lambda'(x) \;\perp\; \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{\Sigma}'(x) \in \mathbb{R} \quad \text{a.e. in } \mathcal{I}(\ell). \tag{3.28e}$$

Moreover, λ' is unique.

Remark 3.13. *By setting $F = (F_1, F_2) : \mathbb{S} \times \mathbb{R} \to \mathbb{R}^2$, $F_1(\boldsymbol{\Sigma}, \lambda) := -\phi(\boldsymbol{\Sigma})$, $F_2(\boldsymbol{\Sigma}, \lambda) := \lambda$, we find that the complementarity relations (3.28c)–(3.28e) are equivalent to*

$$\min \left\{ F_i'\big(\boldsymbol{\Sigma}(x), \lambda(x)\big)\big(\boldsymbol{\Sigma}'(x), \lambda'(x)\big) : i \in \{1,2\} \text{ with } F_i\big(\boldsymbol{\Sigma}(x), \lambda(x)\big) = 0 \right\} = 0$$

a.e. in Ω, which corresponds to the notation of [Scheel and Scholtes, 2000, Section 2.1].

Proof of Proposition 3.12. Suppose that $(\boldsymbol{\Sigma}', \boldsymbol{u}') \in \mathcal{S}_\ell \times V$ is the unique solution of (3.2), then (3.28b) follows immediately. As seen in Section 3.1, $(\boldsymbol{\Sigma}', \boldsymbol{u}')$ equals the weak limit $(\widetilde{\boldsymbol{\Sigma}}, \widetilde{\boldsymbol{u}})$ of the difference quotient in (3.6). Proposition 3.4 and Lemma 3.5 imply that there exists a unique $\lambda' = \widetilde{\lambda} \in L^2(\Omega)$ such that (3.28a) holds true. Moreover, by (3.18) we have $\lambda' = 0$ on $\mathcal{I}(\ell)$, which is (3.28e). Equation (3.28c) follows from $\boldsymbol{\Sigma}' \in \mathcal{S}_\ell$. The relations on $\mathcal{B}(\ell)$ in (3.28d) follow from $\lambda' \geq 0$ on $\mathcal{B}(\ell)$ by (3.23), $\boldsymbol{\Sigma}' \in \mathcal{S}_\ell$, and from $\lambda' \, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\widetilde{\boldsymbol{\Sigma}} = 0$ by Proposition 3.8. Thus $(\boldsymbol{\Sigma}', \boldsymbol{u}')$, together with λ', indeed solves (3.28).

If on the other hand $(\boldsymbol{\Sigma}', \boldsymbol{u}')$ is a solution of (3.28), then the same arguments as in the proof of Lemma 3.7 yield $\lambda' \, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{T} \leq 0$ a.e. in Ω for all $\boldsymbol{T} \in \mathcal{S}_\ell$.

Furthermore, the complementarity relations in (3.28c)–(3.28e) immediately imply $\lambda' \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{\Sigma}' = 0$, hence the variational inequality (3.2a) follows from (3.28a) tested with $\boldsymbol{T} - \boldsymbol{\Sigma}'$. Finally, $\boldsymbol{\Sigma}' \in \mathcal{S}_\ell$ is readily obtained from (3.28c) and (3.28d), and by using that (3.28a), tested with $\boldsymbol{\Sigma}'$, implies $\sqrt{\lambda}\,\mathcal{D}\boldsymbol{\Sigma}' \in S$. □

Thus, in view of Theorem 3.10, we have found the following

Corollary 3.14. *Let $\overline{\boldsymbol{g}} \in U_{\text{ad}}$ be a local optimal solution of* (**P**) *with associated state $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}) = G(\ell)$ for $\ell := -\tau_N^\star \overline{\boldsymbol{g}}$. Then the following variational inequality is satisfied:*

$$\int_\Omega (\bar{\boldsymbol{u}} - \boldsymbol{u}_d) \cdot \boldsymbol{u}' \,\mathrm{d}x + \nu \int_{\Gamma_N} \overline{\boldsymbol{g}} \cdot (\boldsymbol{g} - \overline{\boldsymbol{g}}) \,\mathrm{d}s \geq 0 \quad \text{for all } \boldsymbol{g} \in U_{\text{ad}}, \tag{3.29}$$

where $(\boldsymbol{\Sigma}', \boldsymbol{u}', \lambda')$ solves the derivative problem (3.28) *with $\delta\ell = -\tau_N^\star(\boldsymbol{g} - \overline{\boldsymbol{g}})$ as right hand side.*

Remark 3.15. *This notion of B-stationarity is the infinite dimensional version of the B-stationary concept given in [Scheel and Scholtes, 2000, Section 2.1]. We observe that, as in the finite dimensional setting, B-stationarity is a purely primal concept which does not involve any dual variables. Already in the case of a finite dimensional MPEC, this observation implies that the verification of B-stationarity conditions involves the evaluation of a possibly large number of inequality systems. In the infinite dimensional setting, however, one has to deal even with infinitely many inequalities. Thus, B-stationarity is in general not useful for numerical computations.*

4 Strong Stationarity

The main part of this section is devoted to the derivation of strong stationarity conditions for a modified problem in Section 4.1. We obtain a system consistent with the notion of strong stationarity for finite dimensional MPECs as in Scheel and Scholtes [2000]. Subsequently, we derive in Section 4.2 an equivalent formulation which coincides with the optimality conditions given in Mignot and Puel [1984] for the obstacle problem, but which were not termed 'strong stationarity conditions' at the time. Compared to Mignot and Puel [1984], we present a different and more elementary technique of proof which avoids the concept of conical derivatives. In Section 4.3, we state strong stationarity conditions for the original problem (**P**) (without proving their necessity) and show that they imply the B-stationarity conditions.

We suppose Assumption 1.1 to hold throughout this section.

4.1. Strong Stationarity for a Modified Problem. We consider the following modification of the optimal control problem (**P**):

$$\left.\begin{aligned} \text{Minimize}\quad & \widetilde{J}(\boldsymbol{u}, \ell, \mathscr{L}) := \frac{1}{2}\|\boldsymbol{u} - \boldsymbol{u}_d\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_\ell}{2}\|R\,\ell\|^2_V + \frac{\nu_{\mathscr{L}}}{2}\|\mathscr{L}\|^2_{S^2} \\ \text{s.t.}\quad & \langle A\boldsymbol{\Sigma}, \boldsymbol{T} - \boldsymbol{\Sigma}\rangle + \langle B^\star \boldsymbol{u}, \boldsymbol{T} - \boldsymbol{\Sigma}\rangle \geq \langle \mathscr{L}, \boldsymbol{T} - \boldsymbol{\Sigma}\rangle \quad \forall \boldsymbol{T} \in \mathcal{K} \\ & B\boldsymbol{\Sigma} = \ell \end{aligned}\right\} \quad (\widetilde{\mathbf{P}})$$

where $R : V' \to V$ is the Riesz isomorphism defined through

$$\boldsymbol{z} = R\,\ell \quad \Leftrightarrow \quad \int_\Omega (\nabla \boldsymbol{z} : \nabla \boldsymbol{v} + \boldsymbol{z} \cdot \boldsymbol{v}) \,\mathrm{d}x = \langle \ell, \boldsymbol{v}\rangle \quad \forall \boldsymbol{v} \in V.$$

Due to the isometry property of R, $\|\ell\|_{V'} = \|R\,\ell\|_V$ holds for all $\ell \in V'$.

Problem $(\widetilde{\mathbf{P}})$ differs from $(\mathbf{P})$ in the following ways. First of all, the term $\ell \in V'$ is used directly as a control variable, whereas in $(\mathbf{P})$, only loads $\ell = -\tau_N^\star \boldsymbol{g}$ induced by boundary control functions $\boldsymbol{g}$ where used. Secondly, $\mathscr{L} \in S^2$ appears as an additional control variable on the right hand side of the variational inequality. Finally, no control constraints are present. These modifications are in accordance with previous strong stationarity results for control of the obstacle problem, see Mignot and Puel [1984], where it was also required that the set of feasible controls maps onto the range space of the variational inequality.

In the elastic case, i.e. $\mathcal{K} = S^2$ and $\boldsymbol{\chi} \equiv \mathbf{0}$, the control $\mathscr{L}$ can be interpreted as a prestress applied to the work piece Ω. In the context of elastoplasticity however, problem $(\widetilde{\mathbf{P}})$ is of rather academic nature. Nevertheless, it is worth to be investigated since the particular structure of $(\widetilde{\mathbf{P}})$ allows to establish first-order optimality conditions in strongly stationary form. In contrast to the results in Section 3, strongly stationary optimality systems involve adjoint states and Lagrange multipliers. Our technique for the derivation of strong stationarity conditions differs completely from the one used in Mignot and Puel [1984] for optimal control of the obstacle problem. Instead of transforming the B-stationarity conditions as done in Mignot and Puel [1984], we introduce two auxiliary problems, which turn out to be "standard" control problems that allow the application of the generalized Karush-Kuhn-Tucker (KKT) theory. As a local optimal control of $(\widetilde{\mathbf{P}})$ is also locally optimal for these auxiliary problems, the associated KKT systems apply and strong stationarity is a consequence of these KKT systems.

Let us consider a fixed local optimum $(\bar\ell, \bar{\mathscr{L}}) \in V' \times S^2$. We do not address the existence of a global solution to $(\widetilde{\mathbf{P}})$. This is a delicate question since the solution operator $(\ell, \mathscr{L}) \mapsto (\boldsymbol{\Sigma}, \boldsymbol{u})$ associated with the variational inequality in $(\widetilde{\mathbf{P}})$ is not completely continuous from $S^2 \times V'$ to $S^2 \times V$. Thus, the discussion of global existence for $(\widetilde{\mathbf{P}})$ would go beyond the scope of this paper.

The results for (2.1) readily transfer to the variational inequality in $(\widetilde{\mathbf{P}})$ since the underlying analysis is not affected by the additional inhomogeneity $\mathscr{L} \in S^2$. Therefore, we find the following result analogous to Theorem 2.2:

Lemma 4.1. *For every $(\ell, \mathscr{L}) \in V' \times S^2$, there is a unique solution $(\boldsymbol{\Sigma}, \boldsymbol{u}) \in S^2 \times V$ of the variational inequality in $(\widetilde{\mathbf{P}})$. The pair $(\boldsymbol{\Sigma}, \boldsymbol{u}) \in S^2 \times V$ is the unique solution if and only if there exists a plastic multiplier $\lambda \in L^2(\Omega)$ such that*

$$A\boldsymbol{\Sigma} + B^\star \boldsymbol{u} + \lambda\, \mathcal{D}^\star \mathcal{D} \boldsymbol{\Sigma} = \mathscr{L} \quad \text{in } S^2, \tag{4.1a}$$

$$B\boldsymbol{\Sigma} = \ell \quad \text{in } V', \tag{4.1b}$$

$$0 \le \lambda(x) \;\perp\; \phi(\boldsymbol{\Sigma}(x)) \le 0 \quad \text{a.e. in } \Omega \tag{4.1c}$$

holds. Moreover, λ is unique.

In view of this lemma, $(\widetilde{\mathbf{P}})$ is equivalent to

$$\begin{aligned} \text{Minimize} \quad & \widetilde{J}(\boldsymbol{u}, \ell, \mathscr{L}) := \frac{1}{2} \|\boldsymbol{u} - \boldsymbol{u}_d\|^2_{L^2(\Omega;\mathbb{R}^d)} + \frac{\nu_\ell}{2} \|R\,\ell\|^2_V + \frac{\nu_{\mathscr{L}}}{2} \|\mathscr{L}\|^2_{S^2} \\ \text{s.t.} \quad & (4.1). \end{aligned}$$

Next, we introduce two auxiliary problems. To this end, let us denote by $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}, \bar\lambda)$ the solution of (4.1) for controls $(\bar\ell, \bar{\mathscr{L}})$. Moreover, similarly to (2.2), we define up to sets of measure zero

$$\bar{\mathcal{A}} := \{x \in \Omega : \phi(\bar{\boldsymbol{\Sigma}}(x)) = 0\} \qquad \text{(active set)} \tag{4.2a}$$

$$\bar{\mathcal{A}}_s := \{x \in \Omega : \bar{\lambda}(x) > 0\} \qquad \text{(strongly active set)} \tag{4.2b}$$

$$\bar{\mathcal{B}} := \{x \in \Omega : \phi(\bar{\boldsymbol{\Sigma}}(x)) = \bar{\lambda}(x) = 0\} = \bar{\mathcal{A}} \setminus \bar{\mathcal{A}}_s \qquad \text{(biactive set)} \tag{4.2c}$$

$$\bar{\mathcal{I}} := \{x \in \Omega : \phi(\bar{\boldsymbol{\Sigma}}(x)) < 0\} = \Omega \setminus \bar{\mathcal{A}} \qquad \text{(inactive set).} \tag{4.2d}$$

The two auxiliary problems ($\widetilde{\mathbf{P}}_i$) are derived from ($\widetilde{\mathbf{P}}$) in the following way. In each of them, the inequality constraint $\phi(\boldsymbol{\Sigma}) \le 0$ is tightened and replaced with an equality constraint on some set $\bar{\mathcal{A}}_i \subset \Omega$, while the inequality constraint $\lambda \ge 0$ is replaced with an equality constraint on the complement $\Omega \setminus \bar{\mathcal{A}}_i$. The two problems ($\widetilde{\mathbf{P}}_1$) and ($\widetilde{\mathbf{P}}_2$) differ by the choice of these sets, viz. $\bar{\mathcal{A}}_1 := \bar{\mathcal{A}}_s$ and $\bar{\mathcal{A}}_2 := \bar{\mathcal{A}}$.

These choices lead to the definitions of the following convex feasible sets:

$$Z_1 := \{\boldsymbol{\Sigma} \in S^2 : \phi(\boldsymbol{\Sigma}(x)) \le 0 \text{ a.e. in } \bar{\mathcal{B}} \cup \bar{\mathcal{I}}\},$$
$$M_1 := \{\lambda \in L^2(\Omega) : \lambda(x) \ge 0 \text{ a.e. in } \bar{\mathcal{A}}_s, \quad \lambda(x) = 0 \text{ a.e. in } \bar{\mathcal{B}} \cup \bar{\mathcal{I}}\},$$

and

$$Z_2 := \{\boldsymbol{\Sigma} \in S^2 : \phi(\boldsymbol{\Sigma}(x)) \le 0 \text{ a.e. in } \mathcal{I}\},$$
$$M_2 := \{\lambda \in L^2(\Omega) : \lambda(x) \ge 0 \text{ a.e. in } \mathcal{A}_s \cup \mathcal{B}, \quad \lambda(x) = 0 \text{ a.e. in } \mathcal{I}\}.$$

The auxiliary problems can now be stated as

$$\left.\begin{aligned} &\text{Minimize} \quad \widetilde{J}(\boldsymbol{u}, \ell, \mathscr{L}) \\ &\qquad \text{s.t.} \quad A\boldsymbol{\Sigma} + B^\star \boldsymbol{u} + \lambda\, \mathcal{D}^\star\mathcal{D}\boldsymbol{\Sigma} = \mathscr{L} \quad \text{in } S^2, \\ &\qquad\qquad B\boldsymbol{\Sigma} = \ell \quad \text{in } V', \\ &\qquad\qquad \phi(\boldsymbol{\Sigma}(x)) = 0 \quad \text{a.e. in } \bar{\mathcal{A}}_i, \\ &\qquad \text{and} \quad \boldsymbol{\Sigma} \in Z_i, \quad \lambda \in M_i, \end{aligned}\right\} \tag{$\widetilde{\mathbf{P}}_i$}$$

with $i = 1, 2$. Note that we treat the non-convex constraint $\phi(\boldsymbol{\Sigma}(x)) = 0$ in an explicit way, while the remaining convex constraints enter implicitly through the definition of the sets Z_i and M_i.

Since every feasible point of ($\widetilde{\mathbf{P}}_1$) and ($\widetilde{\mathbf{P}}_2$) is feasible for ($\widetilde{\mathbf{P}}$) as well, we have the following result.

Lemma 4.2. *Let $(\bar{\ell}, \bar{\mathscr{L}}) \in V' \times S^2$ be a local optimal solution to ($\widetilde{\mathbf{P}}$) with associated state $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}, \bar{\lambda})$. Then $(\bar{\ell}, \bar{\mathscr{L}})$ is also locally optimal for both auxiliary problems ($\widetilde{\mathbf{P}}_1$) and ($\widetilde{\mathbf{P}}_2$).*

In order to apply the KKT theory in Banach spaces to ($\widetilde{\mathbf{P}}_1$) and ($\widetilde{\mathbf{P}}_2$), we verify the constraint qualification of Zowe and Kurcyusz [1979] which is frequently also termed regular point condition. To this end, let us introduce the space S^2_∞ by

$$S^2_\infty := \{\boldsymbol{T} \in S^2 : \mathcal{D}\boldsymbol{T} \in L^\infty(\Omega; \mathbb{S})\}.$$

Endowed with the norm $\|\boldsymbol{T}\|_{S^2} + \|\mathcal{D}\boldsymbol{T}\|_{L^\infty(\Omega;\mathbb{S})}$, S^2_∞ becomes a Banach space. Note that every $\boldsymbol{\Sigma}$ satisfying the constraints in ($\widetilde{\mathbf{P}}$), ($\widetilde{\mathbf{P}}_1$), or ($\widetilde{\mathbf{P}}_2$), respectively, is an element of S^2_∞. This is due to the structure of ϕ, see (1.1).

Let us abbreviate $\boldsymbol{x} := (\boldsymbol{\Sigma}, \boldsymbol{u}, \lambda, \ell, \mathscr{L})$ and define $e_i : S^2_\infty \times V \times L^2(\Omega) \times V' \times S^2 \to S^2 \times V' \times L^\infty(\bar{\mathcal{A}}_i)$, $i = 1, 2$, by

$$e_i(\boldsymbol{x}) := \begin{pmatrix} A\boldsymbol{\Sigma} + B^\star \boldsymbol{u} + \lambda\, \mathcal{D}^\star\mathcal{D}\boldsymbol{\Sigma} - \mathscr{L} \\ B\boldsymbol{\Sigma} - \ell \\ \phi(\boldsymbol{\Sigma})|_{\bar{\mathcal{A}}_i} \end{pmatrix}$$

where $\phi(\mathbf{\Sigma})|_{\bar{\mathcal{A}}_i}$ denotes the restriction of $\phi(\mathbf{\Sigma})$ to $\bar{\mathcal{A}}_i$. Note that the equality constraints in ($\widetilde{\mathbf{P}}_1$) and ($\widetilde{\mathbf{P}}_2$) are equivalent to $e_i(\boldsymbol{x}) = \mathbf{0}$. It is easy to see that e_i is of class C^1 since the nonlinear terms are differentiable with respect to $\mathbf{\Sigma} \in S^2_\infty$ and $\lambda \in L^2(\Omega)$. Finally, we define the cones

$$\begin{aligned} C_i(\overline{\boldsymbol{x}}) := \{t\,(\boldsymbol{x} - \overline{\boldsymbol{x}}) \colon t \geq 0,\ \boldsymbol{x} = (\mathbf{\Sigma}, \boldsymbol{u}, \lambda, \ell, \mathscr{L}) \in S^2_\infty \times V \times L^2(\Omega) \times V' \times S^2, \\ \mathbf{\Sigma} \in Z_i,\ \lambda \in M_i\} \end{aligned}$$

for $i = 1, 2$.

Lemma 4.3. *Let $\overline{\boldsymbol{x}} \in S^2_\infty \times V \times L^2(\Omega) \times V' \times S^2$ be feasible for ($\widetilde{\mathbf{P}}$). Then there holds*

$$e_i'(\overline{\boldsymbol{x}})\, C_i(\overline{\boldsymbol{x}}) = S^2 \times V' \times L^\infty(\bar{\mathcal{A}}_i)$$

for $i = 1, 2$. Consequently, the regular point condition is fulfilled for ($\widetilde{\mathbf{P}}_1$) and ($\widetilde{\mathbf{P}}_2$).

Proof. The derivative of e_i at $\overline{\boldsymbol{x}}$ in the direction $\delta\boldsymbol{x} = (\delta\mathbf{\Sigma}, \delta\boldsymbol{u}, \delta\lambda, \delta\ell, \delta\mathscr{L})$ is given by

$$e_i'(\overline{\boldsymbol{x}})\, \delta\boldsymbol{x} = \begin{pmatrix} A\delta\mathbf{\Sigma} + B^\star\delta\boldsymbol{u} + \bar{\lambda}\,\mathcal{D}^\star\mathcal{D}\,\delta\mathbf{\Sigma} + \delta\lambda\,\mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}} - \delta\mathscr{L} \\ B\delta\mathbf{\Sigma} - \delta\ell \\ (\mathcal{D}\bar{\mathbf{\Sigma}} : \mathcal{D}\,\delta\mathbf{\Sigma})|_{\bar{\mathcal{A}}_i} \end{pmatrix},$$

with $i = 1, 2$. Now let $(\mathscr{L}, \ell, f) \in S^2 \times V' \times L^\infty(\bar{\mathcal{A}}_i)$ be arbitrary. By construction of $C_i(\overline{\boldsymbol{x}})$, we see that

$$\delta\boldsymbol{x}_i = \begin{pmatrix} \delta\mathbf{\Sigma}_i \\ \delta\boldsymbol{u}_i \\ \delta\lambda_i \\ \delta\ell_i \\ \delta\mathscr{L}_i \end{pmatrix} := \begin{pmatrix} \chi_{\bar{\mathcal{A}}_i}\,\bar{\mathbf{\Sigma}}/\tilde{\sigma}_0^2\, f \\ \mathbf{0} \\ 0 \\ -\ell + B(\chi_{\bar{\mathcal{A}}_i}\,\bar{\mathbf{\Sigma}}/\tilde{\sigma}_0^2\, f) \\ -\mathscr{L} + (A + \bar{\lambda}\,\mathcal{D}^\star\mathcal{D})\chi_{\bar{\mathcal{A}}_i}\,\bar{\mathbf{\Sigma}}/\tilde{\sigma}_0^2\, f \end{pmatrix} \in C_i(\overline{\boldsymbol{x}}) \qquad (4.3)$$

belongs to $C_i(\overline{\boldsymbol{x}})$ for $i = 1, 2$. Here $\chi_{\bar{\mathcal{A}}_i}$ denotes the characteristic function on $\bar{\mathcal{A}}_i$. Note that $C_i(\overline{\boldsymbol{x}})$ does not contain any restriction on the $\mathbf{\Sigma}$ component on the set $\bar{\mathcal{A}}_i$. In view of $|\mathcal{D}\bar{\mathbf{\Sigma}}|^2 = \tilde{\sigma}_0^2$ on $\bar{\mathcal{A}}_i$ for $i = 1, 2$, we have $(\mathcal{D}\bar{\mathbf{\Sigma}} : \mathcal{D}\,\delta\mathbf{\Sigma}_i)|_{\bar{\mathcal{A}}_i} = f$ and hence

$$e_i'(\overline{\boldsymbol{x}})\,\delta\boldsymbol{x}_i = (\mathscr{L}, \ell, f),$$

which establishes the case. ☐

Remark 4.4. *We point out that the presence of "ample" controls which cover the entire range space $S^2 \times V'$ of the variational inequality, is essential for the verification of the regular point condition. To our best knowledge, it is an open question how to verify a suitable constraint qualification if additional restrictions on the control are present, as for instance in the case of* (**P**)*, where $\mathscr{L} = \mathbf{0}$ and ℓ is induced by a boundary control. This is the main reason why the following analysis does not apply to* (**P**).

Similarly to our result, the technique of Mignot and Puel [1984] for the derivation of strong stationarity conditions for optimal control of the obstacle problem also requires "ample" controls, i.e. distributed controls that are not restricted by additional control constraints, see [Mignot and Puel, 1984, Section 4]. It is straightforward to see that the upcoming analysis can be adapted to optimal control of the obstacle problem and delivers the same result as in Mignot and Puel [1984] but by a different technique, provided that the obstacle problem under consideration is H^2-regular.

The aim of this section is to derive an optimality system for ($\widetilde{\mathbf{P}}$) which involves an adjoint state and Lagrange multipliers associated with the inequality constraints in (4.1). For convenience, we summarize our notation for primal and dual quantities in

Table 4.1. Note that — as is usual in the study of MPCCs — there is no Lagrange multiplier associated with the complementarity constraint $\lambda\,\phi(\boldsymbol{\Sigma}) = 0$.

	state variable	adjoint variable
generalized stresses	$\boldsymbol{\Sigma}$	$\boldsymbol{\Upsilon}$
displacement field	$\boldsymbol{u}$	$\boldsymbol{w}$
	constraint	associated multiplier
plastic multiplier	$\lambda \ge 0$	μ
yield condition	$\phi(\boldsymbol{\Sigma}) \le 0$	θ

Table 4.1. Summary of primal and dual variable names

We are now in the position to prove the following first-order necessary optimality system of strongly stationary type for $(\widetilde{\mathbf{P}})$:

Theorem 4.5. *Let $(\bar{\mathscr{L}}, \bar{\ell}) \in S^2 \times V'$ be a locally optimal solution to $(\widetilde{\mathbf{P}})$ with associated optimal state $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}, \bar{\lambda}) \in S^2_\infty \times V \times L^2(\Omega)$. Then there exists an adjoint state $(\boldsymbol{\Upsilon}, \boldsymbol{w}) \in S^2 \times V$ and Lagrange multipliers $\mu \in L^2(\Omega)$ and $\theta \in L^2(\Omega)$ such that the following optimality system is fulfilled:*

$$A\bar{\boldsymbol{\Sigma}} + \bar{\lambda}\,\mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}} + B^\star\bar{\boldsymbol{u}} = \bar{\mathscr{L}} \tag{4.4a}$$

$$B\bar{\boldsymbol{\Sigma}} = \bar{\ell} \tag{4.4b}$$

$$0 \le \bar{\lambda} \perp \phi(\bar{\boldsymbol{\Sigma}}) \le 0 \quad \textit{a.e. in } \Omega \tag{4.4c}$$

$$A\boldsymbol{\Upsilon} + B^\star\boldsymbol{w} + \bar{\lambda}\,\mathcal{D}^\star\mathcal{D}\boldsymbol{\Upsilon} + \theta\,\mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}} = \mathbf{0} \tag{4.5a}$$

$$B\boldsymbol{\Upsilon} = -(\bar{\boldsymbol{u}} - \boldsymbol{u}_d) \tag{4.5b}$$

$$\nu_{\mathscr{L}}\bar{\mathscr{L}} - \boldsymbol{\Upsilon} = \mathbf{0} \tag{4.6a}$$

$$\nu_\ell\,\langle\bar{\ell},\,\boldsymbol{v}\rangle - \int_\Omega (\nabla\boldsymbol{w} : \nabla\boldsymbol{v} + \boldsymbol{w}\cdot\boldsymbol{v})\,\mathrm{d}x = 0 \quad \textit{for all } \boldsymbol{v} \in V \tag{4.6b}$$

$$\mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon} - \mu = 0 \tag{4.7a}$$

$$\mu\,\bar{\lambda} = 0 \quad \textit{a.e. in } \Omega \tag{4.7b}$$

$$\theta\,\phi(\bar{\boldsymbol{\Sigma}}) = 0 \quad \textit{a.e. in } \Omega \tag{4.7c}$$

$$\theta \ge 0, \quad \mu \ge 0 \quad \textit{a.e. in } \bar{\mathcal{B}} \tag{4.7d}$$

Moreover, the adjoint states $\boldsymbol{\Upsilon}$ and $\boldsymbol{w}$ and Lagrange multipliers μ and θ are unique.

Remark 4.6. *In contrast to the C-stationarity conditions discussed in Herzog et al. [2010a], the strong stationarity conditions provide a sign on the biactive set not only for the product of the multipliers, but even for each multiplier individually, cf. (4.7d). This is the essential difference between C- and strong stationarity, see also Scheel and Scholtes [2000].*

Proof of Theorem 4.5. We start by defining the Lagrangian $\mathcal{L}_i : S^2_\infty \times V \times L^2(\Omega) \times V' \times S^2 \times S^2 \times V \times L^\infty(\bar{\mathcal{A}}_i)' \to \mathbb{R}$ associated with $(\widetilde{\mathbf{P}}_i)$, $i = 1, 2$, by

$$\begin{aligned}\mathcal{L}_i(\boldsymbol{\Sigma}, \boldsymbol{u}, \lambda, \ell, \mathscr{L}, \boldsymbol{\Upsilon}, \boldsymbol{w}, \theta) := {} & \widetilde{J}(\boldsymbol{u}, \ell, \mathscr{L}) \\ & + \langle A\boldsymbol{\Sigma},\,\boldsymbol{\Upsilon}\rangle + \langle B^\star\boldsymbol{u},\,\boldsymbol{\Upsilon}\rangle + (\lambda, \mathcal{D}\boldsymbol{\Sigma} : \mathcal{D}\boldsymbol{\Upsilon})_\Omega - \langle\mathscr{L},\,\boldsymbol{\Upsilon}\rangle \\ & + \langle B\boldsymbol{\Sigma},\,\boldsymbol{w}\rangle - \langle\ell,\,\boldsymbol{w}\rangle + \langle\phi(\boldsymbol{\Sigma}),\,\theta\rangle_{L^\infty(\bar{\mathcal{A}}_i), L^\infty(\bar{\mathcal{A}}_i)'}.\end{aligned}$$

As was mentioned before, the convex constraints $\boldsymbol{\Sigma} \in Z_i$ and $\lambda \in M_i$ are treated as abstract constraints and hence they are not included in the definition of $\mathcal{L}_i$. According to Lemma 4.3, the constraint qualification of Zowe and Kurcyusz holds for $(\widetilde{\mathbf{P}}_i)$, $i = 1, 2$.

Let us abbreviate $\overline{\boldsymbol{x}} := (\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}, \bar{\lambda}, \bar{\ell}, \bar{\mathscr{L}})$. Since $(\bar{\mathscr{L}}, \bar{\ell})$ is locally optimal for $(\widetilde{\mathbf{P}}_i)$, $i = 1, 2$, there exist Lagrange multipliers $(\boldsymbol{\Upsilon}_i, \boldsymbol{w}_i, \theta_i) \in S^2 \times V \times L^\infty(\bar{\mathcal{A}}_i)'$ such that the following optimality systems are satisfied:

$$\partial_{\mathscr{L}} \mathcal{L}_i(\overline{\boldsymbol{x}}, \boldsymbol{\Upsilon}_i, \boldsymbol{w}_i, \theta_i) = 0 \tag{4.8a}$$

$$\partial_{\ell} \mathcal{L}_i(\overline{\boldsymbol{x}}, \boldsymbol{\Upsilon}_i, \boldsymbol{w}_i, \theta_i) = 0 \tag{4.8b}$$

$$\partial_{\boldsymbol{u}} \mathcal{L}_i(\overline{\boldsymbol{x}}, \boldsymbol{\Upsilon}_i, \boldsymbol{w}_i, \theta_i) = 0 \tag{4.8c}$$

$$\partial_{\boldsymbol{\Sigma}} \mathcal{L}_i(\overline{\boldsymbol{x}}, \boldsymbol{\Upsilon}_i, \boldsymbol{w}_i, \theta_i)(\boldsymbol{T} - \bar{\boldsymbol{\Sigma}}) \geq 0 \quad \text{for all } \boldsymbol{T} \in S^2_\infty \cap Z_i \tag{4.8d}$$

$$\partial_{\lambda} \mathcal{L}_i(\overline{\boldsymbol{x}}, \boldsymbol{\Upsilon}_i, \boldsymbol{w}_i, \theta_i)(\xi - \bar{\lambda}) \geq 0 \quad \text{for all } \xi \in M_i \tag{4.8e}$$

with $i = 1, 2$. It remains to verify that (4.8) implies (4.5)–(4.7) and that the dual variables are unique as claimed. This is done in the following four steps.

Step (1): We begin by verifying (4.6) and (4.5b).

The evaluation (4.8a) and (4.8b) gives

$$\nu_{\mathscr{L}} \bar{\mathscr{L}} - \boldsymbol{\Upsilon}_i = 0 \quad \text{and} \quad \nu_\ell \langle \bar{\ell}, \boldsymbol{v} \rangle - \int_\Omega (\nabla \boldsymbol{w}_i : \nabla \boldsymbol{v} + \boldsymbol{w}_i \cdot \boldsymbol{v}) \, \mathrm{d}x = 0 \quad \forall \boldsymbol{v} \in V$$

and $i = 1, 2$. Since $(\bar{\ell}, \mathscr{L})$ is fixed, this yields $\boldsymbol{\Upsilon}_1 = \boldsymbol{\Upsilon}_2 =: \boldsymbol{\Upsilon}$ and $\boldsymbol{w}_1 = \boldsymbol{w}_2 =: \boldsymbol{w}$, which proves (4.6a) and (4.6b) and the uniqueness of $\boldsymbol{\Upsilon}$ and $\boldsymbol{w}$. Furthermore, (4.5b) immediately follows from (4.8c).

Step (2): Next we confirm (4.7a), (4.7b) and the second part of (4.7d).

We set $\mu := \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon}$ to fulfill (4.7a). Note that $\mu \in L^2(\Omega)$ holds since $\bar{\boldsymbol{\Sigma}} \in S^2_\infty$. Now let us use (4.8e) with $i = 2$, which is equivalent to

$$\int_\Omega \mu \, (\xi - \bar{\lambda}) \, \mathrm{d}x = \int_\Omega (\mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon})(\xi - \bar{\lambda}) \, \mathrm{d}x \geq 0 \quad \forall \xi \in M_2. \tag{4.9}$$

By construction of M_2, we have $0 \in M_2$ and $2\bar{\lambda} \in M_2$. Inserting these as test functions into (4.9) yields $(\mu, \bar{\lambda})_\Omega = 0$ so that (4.9) results in

$$\int_\Omega \mu \, \xi \, \mathrm{d}x \geq 0 \quad \forall \xi \in M_2. \tag{4.10}$$

To evaluate this inequality pointwise, let E be an arbitrary measurable subset of $\bar{\mathcal{A}}$ and choose $\xi = \chi_E$ as test function in (4.10), where χ_E denotes the characteristic function of E. This is clearly feasible since $\chi_E(x) = 0$ a.e. in $\mathcal{I}$. Then we obtain $\int_E \mu \, \mathrm{d}x \geq 0$ for all $E \subset \mathcal{A}$, giving in turn $\mu(x) \geq 0$ a.e. in $\bar{\mathcal{A}}$ and thus $\mu(x) \geq 0$ a.e. in $\bar{\mathcal{B}}$ as claimed in (4.7d). Moreover, since $\bar{\lambda} = 0$ in $\bar{\mathcal{I}}$, $\mu \geq 0$ in $\bar{\mathcal{A}}$, and $\bar{\lambda} \geq 0$, the equation $(\bar{\lambda}, \mu)_\Omega = 0$ implies $\mu(x) \, \bar{\lambda}(x) = 0$ a.e. in Ω, which is (4.7b).

Step (3): We proceed to prove (4.5a) and (4.7c).

Next let us consider (4.8d) with $i = 2$ which reads (in view of $\bar{\mathcal{A}}_2 = \bar{\mathcal{A}}$)

$$\begin{aligned} &\langle A\boldsymbol{\Upsilon}, \boldsymbol{T} - \bar{\boldsymbol{\Sigma}} \rangle + \langle B^\star \boldsymbol{w}, \boldsymbol{T} - \bar{\boldsymbol{\Sigma}} \rangle + \big(\bar{\lambda}, \mathcal{D}\boldsymbol{\Upsilon} : (\mathcal{D}\boldsymbol{T} - \mathcal{D}\bar{\boldsymbol{\Sigma}})\big)_\Omega \\ &\qquad + \langle \phi'(\bar{\boldsymbol{\Sigma}})(\boldsymbol{T} - \bar{\boldsymbol{\Sigma}}), \, \theta_2 \rangle_{L^\infty(\bar{\mathcal{A}}), L^\infty(\bar{\mathcal{A}})'} \geq 0 \quad \forall \boldsymbol{T} \in S^2_\infty \cap Z_2. \end{aligned} \tag{4.11}$$

Now let $\varphi \in L^\infty(\bar{\mathcal{A}})$ be arbitrary and choose

$$\boldsymbol{T}(x) = \begin{cases} \varphi(x) \, \mathcal{D}^\star \mathcal{D} \bar{\boldsymbol{\Sigma}}(x) + \bar{\boldsymbol{\Sigma}}(x), & x \in \bar{\mathcal{A}} \\ \bar{\boldsymbol{\Sigma}}(x), & x \in \bar{\mathcal{I}} \end{cases}$$

as test function in (4.11) which belongs to Z_2 due to the feasibility of $\bar{\boldsymbol{\Sigma}}$. Since $\varphi \in L^\infty(\bar{\mathcal{A}})$ is arbitrary, one obtains

$$\int_{\bar{\mathcal{A}}} \varphi\, A\boldsymbol{\Upsilon} : \mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}}\, \mathrm{d}x + \int_{\bar{\mathcal{A}}} \varphi\, (B^\star \boldsymbol{w}) : \mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}}\, \mathrm{d}x + 2\int_{\bar{\mathcal{A}}} \varphi\, \bar{\lambda}\, \mathcal{D}\boldsymbol{\Upsilon} : \mathcal{D}\bar{\boldsymbol{\Sigma}}\, \mathrm{d}x \\ + 2\langle \varphi\, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\bar{\boldsymbol{\Sigma}},\, \theta_2\rangle_{L^\infty(\bar{\mathcal{A}}), L^\infty(\bar{\mathcal{A}})'} = 0 \quad \forall\, \varphi \in L^\infty(\bar{\mathcal{A}}) \tag{4.12}$$

where we used $\mathcal{D}\mathcal{D}^\star\mathcal{D} = 2\mathcal{D}$. Thanks to (4.7a), (4.7b), and $|\mathcal{D}\bar{\boldsymbol{\Sigma}}|^2 = \tilde{\sigma}_0^2$ on $\bar{\mathcal{A}}$, (4.12) results in

$$\langle \varphi,\, \theta_2\rangle = -\frac{1}{2\,\tilde{\sigma}_0^2} \int_{\bar{\mathcal{A}}} \varphi\, (A\boldsymbol{\Upsilon} + B^\star \boldsymbol{w}) : \mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}}\, \mathrm{d}x \quad \forall\, \varphi \in L^\infty(\bar{\mathcal{A}}). \tag{4.13}$$

Due to $\bar{\boldsymbol{\Sigma}} \in S^2_\infty$, $\boldsymbol{\Upsilon} \in S^2$, and $\boldsymbol{w} \in V$, we have $(A\boldsymbol{\Upsilon} + B^\star\boldsymbol{w}) : \mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}} \in L^2(\bar{\mathcal{A}})$. This implies

$$|\langle \varphi,\, \theta_2\rangle| \le \frac{1}{2\,\tilde{\sigma}_0} \|\varphi\|_{L^2(\Omega)}\, \|(A\boldsymbol{\Upsilon} + B^\star\boldsymbol{w}) : \mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}}\|_{L^2(\Omega)} \quad \forall\, \varphi \in L^\infty(\bar{\mathcal{A}}).$$

Since $L^\infty(\bar{\mathcal{A}})$ is dense in $L^2(\bar{\mathcal{A}})$ this implies $\theta_2 \in L^2(\bar{\mathcal{A}})$. We define

$$\theta(x) := \begin{cases} \theta_2(x), & x \in \bar{\mathcal{A}} \\ 0, & x \notin \bar{\mathcal{A}}. \end{cases}$$

Then, due to $\theta \in L^2(\Omega)$ and due to the density of S^2_∞ in S^2, (4.11) implies

$$\langle A\boldsymbol{\Upsilon},\, \boldsymbol{T} - \bar{\boldsymbol{\Sigma}}\rangle + \langle B^\star\boldsymbol{w},\, \boldsymbol{T} - \bar{\boldsymbol{\Sigma}}\rangle + \big(\bar{\lambda}, \mathcal{D}\boldsymbol{\Upsilon} : (\mathcal{D}\boldsymbol{T} - \mathcal{D}\bar{\boldsymbol{\Sigma}})\big)_\Omega \\ + \big(\theta, \mathcal{D}\bar{\boldsymbol{\Sigma}} : (\mathcal{D}\boldsymbol{T} - \mathcal{D}\bar{\boldsymbol{\Sigma}})\big)_\Omega \ge 0 \quad \forall\, \boldsymbol{T} \in Z_2. \tag{4.14}$$

Since Z_2 does not involve a condition on the set $\bar{\mathcal{A}}$, the above inequality readily yields

$$A\boldsymbol{\Upsilon} + B^\star\boldsymbol{w} + \bar{\lambda}\, \mathcal{D}^\star\mathcal{D}\boldsymbol{\Upsilon} + \theta\, \mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}} = \boldsymbol{0} \quad \text{a.e. in } \bar{\mathcal{A}}. \tag{4.15}$$

To derive a pointwise version of (4.14) on the inactive set $\bar{\mathcal{I}}$, we observe that almost every $x_0 \in \bar{\mathcal{I}}$ is a common Lebesgue point of $A\boldsymbol{\Upsilon} + B^\star\boldsymbol{w} + \bar{\lambda}\, \mathcal{D}^\star\mathcal{D}\boldsymbol{\Upsilon} + \theta\, \mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}}$ and $(A\boldsymbol{\Upsilon} + B^\star\boldsymbol{w} + \bar{\lambda}\, \mathcal{D}^\star\mathcal{D}\boldsymbol{\Upsilon} + \theta\, \mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}}) : \bar{\boldsymbol{\Sigma}}$. Fix any such x_0 and $\boldsymbol{T} \in \mathbb{S}^2$ with $\phi(\boldsymbol{T}) \le 0$, and define

$$\boldsymbol{T}_r(x) = \begin{cases} \boldsymbol{T}, & x \in B_r(x_0) \\ \bar{\boldsymbol{\Sigma}}(x), & \text{otherwise}, \end{cases}$$

where $r > 0$ is sufficient small such that $B_r(x_0) \subset \Omega$. Note that $\boldsymbol{T}_r \in Z_2$ so that it is feasible for (4.14). Inserting this test function into (4.14) and taking the limit $r \searrow 0$ gives the pointwise form of (4.14) on $\bar{\mathcal{I}}$, i.e.

$$(A\boldsymbol{\Upsilon} + B^\star\boldsymbol{w})(x) : (\boldsymbol{T} - \bar{\boldsymbol{\Sigma}}(x)) \\ + (\bar{\lambda}\, \mathcal{D}\boldsymbol{\Upsilon} + \theta\, \mathcal{D}\bar{\boldsymbol{\Sigma}})(x) : (\mathcal{D}\boldsymbol{T} - \mathcal{D}\bar{\boldsymbol{\Sigma}}(x)) \ge 0 \quad \text{a.e. in } \bar{\mathcal{I}} \tag{4.16}$$

for all $\boldsymbol{T} \in \mathbb{S}^2$ satisfying $\phi(\boldsymbol{T}) \le 0$. For almost all $x \in \bar{\mathcal{I}}$, we have $\phi(\bar{\boldsymbol{\Sigma}}(x)) < 0$. Therefore, for $\rho > 0$ sufficiently small (depending on x), there holds

$$\phi(\bar{\boldsymbol{\Sigma}}(x) + \boldsymbol{T}) \le 0 \quad \text{for all } \boldsymbol{T} \in \mathbb{S}^2 \text{ with } |\boldsymbol{T}| \le \rho.$$

Thus, by (4.16), one deduces

$$(A\boldsymbol{\Upsilon} + B^\star\boldsymbol{w})(x) : \boldsymbol{T} + (\bar{\lambda}\, \mathcal{D}\boldsymbol{\Upsilon} + \theta\, \mathcal{D}\bar{\boldsymbol{\Sigma}})(x) : \mathcal{D}\boldsymbol{T} \ge 0 \quad \text{for all } \boldsymbol{T} \in \mathbb{S}^2 \text{ with } |\boldsymbol{T}| \le \rho.$$

for almost all $x \in \bar{\mathcal{I}}$. Together with (4.15) this implies

$$A\boldsymbol{\Upsilon} + B^\star\boldsymbol{w} + \bar{\lambda}\, \mathcal{D}^\star\mathcal{D}\boldsymbol{\Upsilon} + \theta\, \mathcal{D}^\star\mathcal{D}\bar{\boldsymbol{\Sigma}} = \boldsymbol{0} \quad \text{a.e. in } \Omega,$$

i.e. (4.5a).

Since we extended θ on $\bar{\mathcal{I}}$ by zero and $\phi(\bar{\boldsymbol{\Sigma}}(x)) = 0$ holds a.e. in $\bar{\mathcal{A}}$, we obtain $\theta(x)\, \phi(\bar{\boldsymbol{\Sigma}}(x)) = 0$ a.e. in Ω, which coincides with (4.7c). We observe that θ is uniquely defined since it is determined by (4.5a) on $\bar{\mathcal{A}}$ and necessarily $\theta = 0$ on $\bar{\mathcal{I}}$ by (4.7c).

Step (4): It remains to prove the sign condition for θ in (4.7d).

To this end, consider (4.8d) with $i = 1$, i.e. (using $\bar{\mathcal{A}}_1 = \bar{\mathcal{A}}_s$)

$$\begin{aligned}\langle A\boldsymbol{\Upsilon}, \boldsymbol{T} - \bar{\boldsymbol{\Sigma}}\rangle + \langle B^\star \boldsymbol{w}, \boldsymbol{T} - \bar{\boldsymbol{\Sigma}}\rangle + \big(\bar{\lambda}, \mathcal{D}\boldsymbol{\Upsilon} : (\mathcal{D}\boldsymbol{T} - \mathcal{D}\bar{\boldsymbol{\Sigma}})\big)_\Omega \\ + \langle \phi'(\bar{\boldsymbol{\Sigma}})(\boldsymbol{T} - \bar{\boldsymbol{\Sigma}}), \theta_1\rangle_{L^\infty(\bar{\mathcal{A}}_s), L^\infty(\bar{\mathcal{A}}_s)'} \geq 0 \quad \forall \boldsymbol{T} \in S^2_\infty \cap Z_1.\end{aligned} \tag{4.17}$$

Now let $\varphi \in L^\infty(\bar{\mathcal{B}})$ be arbitrary with $\varphi(x) \in [0,1]$ a.e. in $\bar{\mathcal{B}}$ and test (4.17) with

$$\boldsymbol{T}(x) = \begin{cases} \bar{\boldsymbol{\Sigma}}(x), & x \in \bar{\mathcal{I}} \cup \bar{\mathcal{A}}_s \\ (1 - \varphi(x))\bar{\boldsymbol{\Sigma}}(x), & x \in \bar{\mathcal{B}}. \end{cases}$$

Note that this test function is feasible since $\boldsymbol{T}$ is a convex combination on $\bar{\mathcal{B}}$ of the two functions $\boldsymbol{0}$ and $\bar{\boldsymbol{\Sigma}}$ which belong to $S^2_\infty \cap Z_1$. In this way, we obtain

$$0 \overset{(4.17)}{\leq} -\int_{\bar{\mathcal{B}}} \varphi \left(A\boldsymbol{\Upsilon} : \bar{\boldsymbol{\Sigma}} + (B^\star \boldsymbol{w}) : \bar{\boldsymbol{\Sigma}} + \bar{\lambda}\, \mathcal{D}\boldsymbol{\Upsilon} : \mathcal{D}\bar{\boldsymbol{\Sigma}}\right) \mathrm{d}x \overset{(4.5a)}{=} \int_{\bar{\mathcal{B}}} \varphi\, \theta\, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\bar{\boldsymbol{\Sigma}}\, \mathrm{d}x.$$

Due to $|\mathcal{D}\bar{\boldsymbol{\Sigma}}| = \tilde{\sigma}_0$ on $\bar{\mathcal{B}}$, we arrive at

$$\int_{\bar{\mathcal{B}}} \varphi\, \theta\, \mathrm{d}x \geq 0 \quad \forall\, \varphi \in L^\infty(\bar{\mathcal{B}}) \text{ with } \varphi \in [0,1] \text{ a.e. in } \bar{\mathcal{B}}, \tag{4.18}$$

which proves the nonnegativity of θ on $\bar{\mathcal{B}}$. □

4.2. Discussion of Strong Stationarity. In the following we reformulate the strong stationarity conditions (4.4)–(4.7) in order to allow a comparison to the optimality conditions given in Mignot and Puel [1984] for optimal control of the obstacle problem.

Proposition 4.7. *The strong stationarity system (4.4)–(4.7) is equivalent to the following set of conditions*

$$A\bar{\boldsymbol{\Sigma}} + \bar{\lambda}\, \mathcal{D}^\star \mathcal{D}\bar{\boldsymbol{\Sigma}} + B^\star \bar{\boldsymbol{u}} = \bar{\mathscr{L}} \tag{4.19a}$$
$$B\bar{\boldsymbol{\Sigma}} = \bar{\ell} \tag{4.19b}$$
$$0 \leq \bar{\lambda} \ \perp\ \phi(\boldsymbol{\Sigma}) \leq 0 \quad \textit{a.e. in } \Omega \tag{4.19c}$$

$$\langle A\boldsymbol{\Upsilon}, \boldsymbol{T}\rangle + \langle B^\star \boldsymbol{w}, \boldsymbol{T}\rangle + (\bar{\lambda}, \mathcal{D}\boldsymbol{\Upsilon} : \mathcal{D}\boldsymbol{T})_\Omega \geq 0 \quad \textit{for all } \boldsymbol{T} \in \bar{\mathcal{S}} \tag{4.20a}$$
$$B\boldsymbol{\Upsilon} = -(\bar{\boldsymbol{u}} - \boldsymbol{u}_d) \tag{4.20b}$$
$$-\boldsymbol{\Upsilon} \in \bar{\mathcal{S}} \tag{4.20c}$$

$$\nu_{\mathscr{L}} \bar{\mathscr{L}} - \boldsymbol{\Upsilon} = \boldsymbol{0} \tag{4.21a}$$
$$\nu_\ell \,\langle \bar{\ell}, \boldsymbol{v}\rangle - \int_\Omega (\nabla \boldsymbol{w} : \nabla \boldsymbol{v} + \boldsymbol{w} \cdot \boldsymbol{v})\, \mathrm{d}x = 0 \quad \textit{for all } \boldsymbol{v} \in V, \tag{4.21b}$$

where $\bar{\mathcal{S}}$ is defined similarly to (3.3) by

$$\begin{aligned}\bar{\mathcal{S}} := \{\boldsymbol{T} \in S^2 \colon \sqrt{\bar{\lambda}}\, \mathcal{D}\boldsymbol{T} \in S, \quad & \mathcal{D}\bar{\boldsymbol{\Sigma}}(x) : \mathcal{D}\boldsymbol{T}(x) \leq 0 \textit{ a.e. in } \bar{\mathcal{B}}, \\ & \mathcal{D}\bar{\boldsymbol{\Sigma}}(x) : \mathcal{D}\boldsymbol{T}(x) = 0 \textit{ a.e. in } \bar{\mathcal{A}}_s\}\end{aligned}$$

with $\bar{\mathcal{B}}$ and $\bar{\mathcal{A}}_s$ as in (4.2).

Remark 4.8. *The above notion of strong stationarity is equivalent to the one introduced by Mignot and Puel in case of the obstacle problem, cf. [Mignot and Puel, 1984, Theorem 2.2]. We point out that the adjoint system (4.20a)–(4.20c) cannot be written in form of a variational inequality.*

Proof of Proposition 4.7. We only have to show that (4.20) is equivalent to (4.5) and (4.7).

Let us start by assuming that (4.5) and (4.7) are fulfilled. By (4.7a), (4.7b), and (4.7d) we know that $\mathcal{D}\bar{\mathbf{\Sigma}} : \mathcal{D}\mathbf{\Upsilon} \geq 0$ a.e. in $\bar{\mathcal{B}}$ and $\bar{\lambda}\,(\mathcal{D}\bar{\mathbf{\Sigma}} : \mathcal{D}\mathbf{\Upsilon}) = 0$ a.e. in Ω. The latter equality immediately implies $\mathcal{D}\bar{\mathbf{\Sigma}} : \mathcal{D}\mathbf{\Upsilon} = 0$ a.e. in $\bar{\mathcal{A}}_s$. Moreover, if we test (4.5a) with $\mathbf{\Upsilon}$ and use $\theta \in L^2(\Omega)$ and $\mathcal{D}\bar{\mathbf{\Sigma}} \in L^\infty(\Omega, \mathbb{S})$, then $\sqrt{\bar{\lambda}}\,\mathcal{D}\mathbf{\Upsilon} \in S$ is obtained, giving in turn $-\mathbf{\Upsilon} \in \bar{\mathcal{S}}$. To verify (4.20a), let $\boldsymbol{T} \in \bar{\mathcal{S}}$ be arbitrary. Note that (4.7c) yields $\theta = 0$ a.e. in $\bar{\mathcal{I}}$ and (4.7d) implies $\theta \geq 0$ a.e. in $\bar{\mathcal{B}}$. Using the sign conditions on $\mathcal{D}\bar{\mathbf{\Sigma}} : \mathcal{D}\boldsymbol{T}$ implied by $\boldsymbol{T} \in \bar{\mathcal{S}}$, we get

$$(\theta, \mathcal{D}\bar{\mathbf{\Sigma}} : \mathcal{D}\boldsymbol{T})_\Omega \leq 0.$$

Inserting this into (4.5a) results in (4.20a) so that (4.20) is indeed verified.

The opposite direction is shown in three steps. To this end assume that $(\mathbf{\Upsilon}, \boldsymbol{w}) \in S^2 \times V$ satisfies (4.20).

Step (1): We confirm the conditions on μ in (4.7a), (4.7b), and (4.7d).

First let us define μ according to (4.7a) by $\mu := \mathcal{D}\bar{\mathbf{\Sigma}}:\mathcal{D}\mathbf{\Upsilon}$. Due to $\mathcal{D}\bar{\mathbf{\Sigma}} \in L^\infty(\Omega, \mathbb{S})$ in virtue of $\phi(\bar{\mathbf{\Sigma}}) \leq 0$, we have $\mu \in L^2(\Omega)$. Moreover, because of $-\mathbf{\Upsilon} \in \bar{\mathcal{S}}$ by (4.20c), one obtains $\mu = 0$ a.e. in $\bar{\mathcal{A}}_s = \{x \in \Omega : \bar{\lambda}(x) \neq 0\}$ so that $\mu\,\bar{\lambda} = 0$, i.e. (4.7b) is satisfied. The sign condition on μ on the biactive set $\bar{\mathcal{B}}$ finally follows directly from $-\mathbf{\Upsilon} \in \bar{\mathcal{S}}$.

Step (2): We verify the first adjoint equation (4.5a).

Let us start by defining

$$\boldsymbol{M} := A\mathbf{\Upsilon} + B^\star \boldsymbol{w} + \bar{\lambda}\,\mathcal{D}^\star\mathcal{D}\mathbf{\Upsilon}$$

so that

$$\int_\Omega \boldsymbol{M} : \boldsymbol{T} \,\mathrm{d}x \geq 0 \quad \text{for all } \boldsymbol{T} \in \bar{\mathcal{S}}. \tag{4.22}$$

due to (4.20a). Let $x \in \Omega$ be arbitrary and define $\theta(x) \in \mathbb{R}$ and $\boldsymbol{M}_0(x) \in \mathbb{S}^2$ with $\boldsymbol{M}_0(x) : \mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x) = 0$, i.e. $\boldsymbol{M}_0(x)$ is in the orthogonal complement of $\operatorname{span}(\mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x))$, by

$$\boldsymbol{M}(x) = -\theta(x)\,\mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x) + \boldsymbol{M}_0(x). \tag{4.23}$$

Since $\bar{\mathcal{S}}$ does not involve any condition on the inactive set $\bar{\mathcal{I}}$ (apart from the regularity condition $\sqrt{\bar{\lambda}}\,\mathcal{D}\boldsymbol{T} \in S$), we are allowed to insert $\chi_{\bar{\mathcal{I}}}\boldsymbol{T}$ with arbitrary $\boldsymbol{T} \in L^\infty(\Omega, \mathbb{S})^2$ as test function, giving in turn $\boldsymbol{M} = \boldsymbol{0}$ a.e. in $\bar{\mathcal{I}}$. In particular, this implies $\theta(x) = 0$ and $\boldsymbol{M}_0(x) = \boldsymbol{0}$ for almost all $x \in \bar{\mathcal{I}}$. Next we investigate the regularity of θ. By multiplying (4.23) with $\mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x)$ and taking $\phi(\bar{\mathbf{\Sigma}}) = 0$ a.e. in $\bar{\mathcal{A}}$ into account, we obtain f.a.a. $x \in \bar{\mathcal{A}}$

$$\begin{aligned} -2\,\tilde{\sigma}_0\,\theta(x) &= -\theta(x)\,\mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x) : \mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x) = \boldsymbol{M}(x) : \mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x) \\ &= (A\mathbf{\Upsilon} + B^\star\boldsymbol{w})(x) : \mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x) + \bar{\lambda}(x)\,\mathcal{D}^\star\mathcal{D}\mathbf{\Upsilon}(x) : \mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x) \\ &= (A\mathbf{\Upsilon} + B^\star\boldsymbol{w})(x) : \mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x) + 2\,\bar{\lambda}(x)\,\mathcal{D}\mathbf{\Upsilon}(x) : \mathcal{D}\bar{\mathbf{\Sigma}}(x). \end{aligned}$$

Because of (4.20c) we have $\mathcal{D}\mathbf{\Upsilon}(x) : \mathcal{D}\bar{\mathbf{\Sigma}}(x) = 0$ a.e. in $\bar{\mathcal{A}}_s = \{x \in \Omega : \bar{\lambda}(x) \neq 0\}$ and thus

$$\theta(x) = \frac{-1}{2\,\tilde{\sigma}_0}(A\mathbf{\Upsilon} + B^\star\boldsymbol{w})(x) : \mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}}(x) \in L^2(\bar{\mathcal{A}})$$

due to $\mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}} \in L^\infty(\Omega, \mathbb{S})$. The sign conditions in $\bar{\mathcal{S}}$ are satisfied by definition of $\boldsymbol{M}_0$. We define $\Lambda_s = \{x \in \bar{\mathcal{A}} : \bar{\lambda} \leq s\}$ for $s > 0$. Due to $\theta\,\mathcal{D}^\star\mathcal{D}\bar{\mathbf{\Sigma}} \in S^2$, (4.23), the definition of $\boldsymbol{M}$ and $\bar{\lambda} \leq s$ on Λ_s, we have $-\chi_{\Lambda_s}\boldsymbol{M}_0 \in \bar{\mathcal{S}}$. Therefore, (4.22) together with (4.23) implies $\boldsymbol{M}_0 = \boldsymbol{0}$ on Λ_s. Since $s > 0$ was arbitrary, $\boldsymbol{M}_0 = \boldsymbol{0}$ a.e.

in Ω. Now we extend θ on $\bar{\mathcal{I}}$ by zero. Therefore, $\theta \in L^2(\Omega)$ holds. Using $\boldsymbol{M} = \mathbf{0}$ a.e. on $\bar{\mathcal{I}}$, $\boldsymbol{M}_0 = \mathbf{0}$ and (4.23) we obtain

$$-\theta\, \mathcal{D}^\star \mathcal{D}\bar{\boldsymbol{\Sigma}} = \boldsymbol{M} = A\boldsymbol{\Upsilon} + B^\star \boldsymbol{w} + \bar{\lambda}\, \mathcal{D}^\star \mathcal{D}\boldsymbol{\Upsilon},$$

which is (4.5a).

Step (3): It remains to prove the complementarity relation (4.7c) and the sign condition on θ in (4.7d).

Since $\theta = 0$ in $\bar{\mathcal{I}} = \{x \in \Omega : \phi(\bar{\boldsymbol{\Sigma}}(x)) \neq 0\}$ by construction, $\theta\, \phi(\bar{\boldsymbol{\Sigma}}) = 0$ a.e. in Ω, i.e. (4.7c) is trivially fulfilled. To verify the sign condition on the biactive set, let $E \subset \bar{\mathcal{B}}$ be some measurable subset. If we insert $-\chi_E\, \bar{\boldsymbol{\Sigma}} \in \bar{\mathcal{S}}$ as test function in (4.22), then, in view of $\boldsymbol{M} = -\theta\, \mathcal{D}^\star \mathcal{D}\bar{\boldsymbol{\Sigma}}$ on $\bar{\mathcal{B}}$, we obtain

$$0 \leq \int_E \theta\, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\bar{\boldsymbol{\Sigma}}\, \mathrm{d}x = \tilde{\sigma}_0 \int_E \theta\, \mathrm{d}x.$$

Since $E \subset \bar{\mathcal{B}}$ was arbitrary, we have $\theta \geq 0$ on $\bar{\mathcal{B}}$. □

4.3. Strong Stationarity implies B-Stationarity.

In this subsection, we state strong stationarity conditions for the original problem (**P**) (without proving their necessity) and show that they imply the B-stationarity conditions from Theorem 3.10. In view of (4.4)–(4.7), the strong stationarity conditions for the original problem are defined as follows.

Definition 4.9. *We say that an optimal control $\bar{\boldsymbol{g}} \in U_{\mathrm{ad}}$ with associated state $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}, \bar{\lambda}) \in S^2 \times V \times L^2(\Omega)$ satisfies the* strong stationarity condition *for* (**P**) *if an adjoint state $(\boldsymbol{\Upsilon}, \boldsymbol{w}) \in S^2 \times V$ and Lagrange multipliers $\mu, \theta \in L^2(\Omega)$ exist such that*

$$A\bar{\boldsymbol{\Sigma}} + \bar{\lambda}\, \mathcal{D}^\star \mathcal{D}\bar{\boldsymbol{\Sigma}} + B^\star \bar{\boldsymbol{u}} = \mathbf{0} \tag{4.24a}$$
$$B\bar{\boldsymbol{\Sigma}} = -\tau_N^\star \bar{\boldsymbol{g}} \tag{4.24b}$$
$$0 \leq \bar{\lambda} \perp \phi(\bar{\boldsymbol{\Sigma}}) \leq 0 \quad \text{a.e. in } \Omega \tag{4.24c}$$
$$A\boldsymbol{\Upsilon} + B^\star \boldsymbol{w} + \bar{\lambda}\, \mathcal{D}^\star \mathcal{D}\boldsymbol{\Upsilon} + \theta\, \mathcal{D}^\star \mathcal{D}\bar{\boldsymbol{\Sigma}} = \mathbf{0} \tag{4.25a}$$
$$B\boldsymbol{\Upsilon} = -(\bar{\boldsymbol{u}} - \boldsymbol{u}_d) \tag{4.25b}$$
$$\int_{\Gamma_N} (\nu\, \bar{\boldsymbol{g}} + \boldsymbol{w}) \cdot (\boldsymbol{g} - \bar{\boldsymbol{g}})\, \mathrm{d}s \geq 0 \quad \text{for all } \boldsymbol{g} \in U_{\mathrm{ad}} \tag{4.26a}$$
$$\mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon} - \mu = 0 \tag{4.27a}$$
$$\mu\, \bar{\lambda} = 0 \quad \text{a.e. in } \Omega \tag{4.27b}$$
$$\theta\, \phi(\bar{\boldsymbol{\Sigma}}) = 0 \quad \text{a.e. in } \Omega \tag{4.27c}$$
$$\theta \geq 0, \quad \mu \geq 0 \quad \text{a.e. in } \bar{\mathcal{B}} \tag{4.27d}$$

holds true.

Remark 4.10. *As already mentioned in Remark 4.4, we cannot prove that (4.24)–(4.27) are necessary for the local optimality of $\bar{\boldsymbol{g}}$. The reason is that a verification of the regular point condition for the auxiliary problems ($\widetilde{\mathbf{P}}_1$) and ($\widetilde{\mathbf{P}}_2$) does not seem to be possible in case of* (**P**). *Note that a regularization technique would yield C-stationarity conditions that coincide with (4.24)–(4.27) except that (4.27d) has to be replaced by $\theta\, \mu \geq 0$ in $\bar{\mathcal{B}}$, cf. [Herzog et al., 2010a, Section 3.3].*

The next proposition shows that strong stationarity implies B-stationarity.

Proposition 4.11. *Assume that $\bar{\boldsymbol{g}} \in U_{\mathrm{ad}}$ with associated state $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}}, \bar{\lambda}) \in S^2 \times V \times L^2(\Omega)$ fulfills the strong stationarity condition (4.24)–(4.27). Then $\bar{\boldsymbol{g}}$ satisfies the B-stationarity condition (3.27), i.e. the variational inequality*

$$\int_\Omega (\bar{\boldsymbol{u}} - \boldsymbol{u}_d) \cdot \boldsymbol{u}' \,\mathrm{d}x + \nu \int_{\Gamma_N} \bar{\boldsymbol{g}} \cdot (\boldsymbol{g} - \bar{\boldsymbol{g}}) \,\mathrm{d}s \geq 0 \quad \text{for all } \boldsymbol{g} \in U_{\mathrm{ad}},$$

where $(\boldsymbol{\Sigma}', \boldsymbol{u}')$ solves the derivative problem (3.2) with $\delta\ell := -\tau_N^\star(\boldsymbol{g} - \bar{\boldsymbol{g}})$ as right hand side.

Proof. According to Proposition 3.12 the variational inequality for $(\boldsymbol{\Sigma}', \boldsymbol{u}')$ can equivalently be expressed in terms of (3.28) (with $(\bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{u}})$ instead of $(\boldsymbol{\Sigma}, \boldsymbol{u})$), which involves a multiplier $\lambda' \in L^2(\Omega)$. If we test (3.28b) with $\boldsymbol{w}$ and set $\delta\ell = -\tau_N^\star(\boldsymbol{g} - \bar{\boldsymbol{g}})$, we arrive at

$$\begin{aligned}
(\boldsymbol{w}, \boldsymbol{g} - \bar{\boldsymbol{g}})_{\Gamma_N} &= -\langle B^\star \boldsymbol{w}, \boldsymbol{\Sigma}' \rangle \\
&= \langle A\boldsymbol{\Upsilon}, \boldsymbol{\Sigma}' \rangle + (\bar{\lambda}, \mathcal{D}\boldsymbol{\Upsilon} : \mathcal{D}\boldsymbol{\Sigma}')_\Omega + (\theta, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}')_\Omega && \text{by (4.25a)} \\
&= -\langle B^\star \boldsymbol{u}', \boldsymbol{\Upsilon} \rangle - (\lambda', \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon})_\Omega + (\theta, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}')_\Omega && \text{by (3.28a)} \\
&= (\bar{\boldsymbol{u}} - \boldsymbol{u}_d, \boldsymbol{u}')_\Omega - (\lambda', \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon})_\Omega + (\theta, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}')_\Omega && \text{by (4.25b)}.
\end{aligned}$$

For the last two addends in the previous equation, we obtain the following sign conditions:

$$\theta\, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}' \begin{cases} = 0 \text{ in } \bar{\mathcal{I}}, & \text{since } \theta = 0 \text{ a.e. in } \bar{\mathcal{I}} \text{ by (4.27c)} \\ \leq 0 \text{ in } \bar{\mathcal{B}}, & \text{since } \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}' \leq 0 \text{ by (3.28d)} \\ & \text{and } \theta \geq 0 \text{ a.e. in } \bar{\mathcal{B}} \text{ by (4.27d)} \\ = 0 \text{ in } \bar{\mathcal{A}}_s, & \text{since } \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Sigma}' = 0 \text{ a.e. in } \bar{\mathcal{A}}_s \text{ by (3.28c)} \end{cases}$$

and

$$\lambda'\, \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon} \begin{cases} = 0 \text{ in } \bar{\mathcal{I}}, & \text{since } \lambda' = 0 \text{ a.e. in } \bar{\mathcal{I}} \text{ by (3.28e)} \\ \geq 0 \text{ in } \bar{\mathcal{B}}, & \text{since } \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon} = \mu \geq 0 \text{ by (4.27a) and (4.27d)} \\ & \text{and } \lambda' \geq 0 \text{ a.e. in } \bar{\mathcal{B}} \text{ by (3.28d)} \\ = 0 \text{ in } \bar{\mathcal{A}}_s, & \text{since } \mathcal{D}\bar{\boldsymbol{\Sigma}} : \mathcal{D}\boldsymbol{\Upsilon} = \mu = 0 \text{ a.e. in } \bar{\mathcal{A}}_s \text{ by (4.27b)}. \end{cases}$$

These sign conditions imply

$$(\boldsymbol{w}, \boldsymbol{g} - \bar{\boldsymbol{g}})_{\Gamma_N} \leq (\bar{\boldsymbol{u}} - \boldsymbol{u}_d, \boldsymbol{u}')_\Omega.$$

Inserting this into (4.26a) yields the desired result. □

A Auxiliary results

In this section, $|B|$ denotes the Lebesgue measure of a set $B \subset \Omega$.

Lemma A.1. *Let $M \subset \Omega$ be some measurable set and let $b \in L^1(M)$ with $b > 0$ a.e. in M be given. Then, for all $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\int_B b \,\mathrm{d}x \geq \delta \quad \text{for all } B \subset M \text{ s.t. } |B| \geq \varepsilon$$

holds.

Proof. Let us define $g : \mathbb{R}^+ \to \mathbb{R}^+$, $g(\gamma) := |\{x \in M : b \le \gamma\}|$. We obtain that g is monotone increasing and $g(0) = 0$. Therefore $\lim_{\gamma\searrow 0} g(\gamma)$ exists and we have

$$\lim_{\gamma\searrow 0} g(\gamma) = \lim_{n\to\infty} g(2^{-n}) = \lim_{n\to\infty} |\{b \le 2^{-n}\}|$$
$$= \Big|\bigcap_{n=1}^{\infty} \{b \le 2^{-n}\}\Big| = |\{b = 0\}| = 0.$$

This shows that g is continuous at 0. Hence there is $\delta_0 > 0$ with $g(\delta_0) \le \varepsilon/2$. For $G := \{b \le \delta_0\}$ we obtain $|G| \le \varepsilon/2$. Let $B \subset M$ with $|B| \ge \varepsilon$ be arbitrary. We have

$$\int_B b \,\mathrm{d}x \ge \int_{B\setminus G} b \,\mathrm{d}x \ge \int_{B\setminus G} \delta_0 \,\mathrm{d}x = \delta_0\, |B \setminus G| \ge \delta_0\, \varepsilon/2.$$

With $\delta := \delta_0\, \varepsilon/2$ the lemma is proved. ☐

Lemma A.2. *Let $M \subset \Omega$ be measurable and $\{f_n\} \subset L^1(M)$ a sequence with $f_n \to f \in L^1(\Omega)$. If $f > 0$ a.e. in M, then $|\{x \in M : f_n = 0\}| \to 0$.*

Proof. We prove this lemma by contradiction. Let us assume that there is $\varepsilon > 0$ and a subsequence n_k such that

$$|\{f_{n_k} = 0\}| \ge \varepsilon \quad \forall\, k \in \mathbb{N}.$$

According to Lemma A.1 there exists $\delta > 0$ with

$$\int_B f \,\mathrm{d}x \ge \delta \quad \text{for all } B \subset M \text{ with } |B| \ge \varepsilon.$$

Now we have

$$\|f - f_{n_k}\|_{L^1(\Omega)} \ge \int_{\{f_{n_k}=0\}} |f - f_{n_k}| \,\mathrm{d}x = \int_{\{f_{n_k}=0\}} f \,\mathrm{d}x \ge \delta \quad \forall\, k \in \mathbb{N}.$$

This is a contradiction to $f_n \to f$ in $L^1(\Omega)$. ☐

Lemma A.3. *Let $\{f_n\} \subset L^1(\Omega)$ with $f_n \ge 0$ a.e. in Ω and $f_n \rightharpoonup f$ in $L^1(\Omega)$ be given. If $|\{x \in \Omega : f_n(x) > 0\}| \to 0$ as $n \to \infty$ holds, then $f \equiv 0$.*

Proof. Let us abbreviate $A_n = \{x \in \Omega : f_n(x) > 0\}$. Since $|A_n| \to 0$, there is a subsequence $\{n_k\} \subset \mathbb{N}$ with $\sum_{k=1}^{\infty} |A_{n_k}| < \infty$. For the sets $B_j := \bigcup_{k=j}^{\infty} A_{n_k}$ we obtain $|B_j| \to 0$ as $j \to \infty$. By construction, $f_{n_i} \equiv 0$ holds on $\Omega \setminus A_{n_i}$ and hence

$$\int_{\Omega\setminus B_j} f_{n_i} \,\mathrm{d}x = 0 \quad \text{if } i > j.$$

The weak convergence $f_n \rightharpoonup f$ implies

$$0 = \int_{\Omega\setminus B_j} f_{n_i} \,\mathrm{d}x \to \int_{\Omega\setminus B_j} f \,\mathrm{d}x \quad \text{as } i \to \infty$$

and thus

$$\int_{\Omega\setminus B_j} f \,\mathrm{d}x = 0 \quad \text{for all } j \in \mathbb{N}.$$

The set of nonnegative functions is weakly closed in $L^1(\Omega)$, hence $f \ge 0$ holds and we obtain $f \equiv 0$ in $\Omega \setminus B_j$ for all $j \in \mathbb{N}$. Using De Morgan's Law, $|B_j| \to 0$ yields $\bigcup_{j=1}^{\infty} \Omega \setminus B_j = \Omega \setminus \bigcap_{j=1}^{\infty} B_j = \Omega$ and thus we conclude $f \equiv 0$ on Ω. ☐

Acknowledgment. This work was supported by a DFG grant within the Priority Program SPP 1253 (*Optimization with Differential Equations*), which is gratefully acknowledged.

References

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

P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston, 1985.

K. Gröger. A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations. *Mathematische Annalen*, 283:679–687, 1989.

R. Haller-Dintelmann, C. Meyer, J. Rehberg, and A. Schiela. Hölder continuity and optimal control for nonsmooth elliptic problems. *Applied Mathematics and Optimization*, 60(3):397–428, 2009.

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

R. Herzog and C. Meyer. Optimal control of static plasticity with linear kinematic hardening. *submitted*, 2009.

R. Herzog, C. Meyer, and G. Wachsmuth. C-stationarity for optimal control of static plasticity with linear kinematic hardening. *submitted*, 2010a.

R. Herzog, C. Meyer, and G. Wachsmuth. Integrability of displacement and stresses in linear and nonlinear elasticity with mixed boundary conditions. *submitted*, 2010b.

R. Herzog, C. Meyer, and G. Wachsmuth. Existence and regularity of the plastic multiplier in static and quasistatic plasticity. *submitted*, 2010c.

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 and I. Kopacka. Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm. Technical Report IFB-Report No. 11, Institute of Mathematics and Scientific Computing, University of Graz, 2008.

Jiří Jarušek and Jiří Outrata. On sharp necessary optimality conditions in control of contact problems with strings. *Nonlinear Analysis*, 67:1117–1128, 2007.

Jiří Jarušek, Jiří Outrata, and Jana Stará. On optimality conditions in control of elliptic variational inequalities. *Set-Valued Analysis*, 2010. To appear.

K. Jittorntrum. Solution point differentiability without strict complementarity in nonlinear programming. *Mathematical Programming Study*, 21:127–138, 1984.

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

H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Mathematics of Operations Research*, 25(1):1–22, 2000.

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.

Chemnitz University of Technology, Faculty of Mathematics, D–09107 Chemnitz, Germany

E-mail address: roland.herzog@mathematik.tu-chemnitz.de

URL: http://www.tu-chemnitz.de/herzog

TU Darmstadt, Graduate School CE, Dolivostr. 15, D–64293 Darmstadt, Germany

E-mail address: cmeyer@gsc.tu-darmstadt.ce

URL: http://www.graduate-school-ce.de/index.php?id=128

Chemnitz University of Technology, Faculty of Mathematics, D–09107 Chemnitz, Germany

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

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